High resolution quadrupole time of flight mass spectrometry in pharmaceutical bioanalysis

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Adrian Sievers-Engler

aus Bad Urach

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	03.03.2020
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Michael Lämmerhofer
2. Berichterstatter:	Jun.-Prof. Dr. Matthias Gehringer
3. Berichterstatter:	Prof. Dr. Maria Kristina Parr

Gewidmet meinen Vorfahren, meiner Familie und meinen Freunden

„Und es ist eine große Lust aufgekommen die Ursachen aller Dinge zu erforschen. Jeden Tag wird etwas gefunden. Selbst die Hundertjährigen lassen sich noch von den Jungen ins Ohr schreien, was Neues entdeckt wurde. Da ist schon viel gefunden, aber da ist noch mehr, was noch gefunden werden kann. Was in alten Büchern steht, das genügt nicht mehr. Eine neue Zeit ist angebrochen, ein großes Zeitalter, in dem zu leben eine Lust ist"

Galilei in Brechts „Leben des Galilei"

Danksagung

Zuvorderst möchte ich Herrn Prof. Dr. Michael Lämmerhofer von Herzen danken, für die recht spontane Einstellung „auf Handschlag" in seinem damals noch taufrischen Arbeitskreis, für die Möglichkeit am Aufbau des Arbeitskreises und insbesondere der Beschaffung des hochauflösenden Massenspektrometers mitzuwirken sowie für die nachfolgende Überlassung und Unterstützung bei der Bearbeitung der breitgefächerten Themen, welche diese Dissertationsschrift ausmachen und der Gelegenheit so ausgiebig zu publizieren und auf Konferenzen zu präsentieren.

Jun.-Prof. Dr. Matthias Gehringer möchte ich für die Zweitbegutachtung dieser Dissertationsschrift und für die Kooperation bei den Forschungsvorhaben danken.

Auch allen anderen Arbeitsgruppen, mit welchen zusammen ich forschen und veröffentlichen konnte sei Dank, stellvertretend für diese ihren jeweiligen Professores: Laufer, Böckler, Koch, Groß, Merfort, Weisz und Lindner

Meinen Kollegen ein Dankeschön für die wunderbare Zeit auf der Morgenstelle, für fachliche Diskussionen aber auch Kurzweil: Dr. Jeannie Horak, Dr. Aleksandra Zimmermann, Corinna Sanwald, Stefanie Bäurer, Eva Döring, Dr. Kirsten Storch, Dr. Heike Gerhart, Dr. Jörg Schlotterbeck, Bernhard Drotleff, Dr. Stefan Neubauer, Christian Geibel, Malgorzata Cebo und Carlos Calderón Castro. Und natürlich alle anderen von der Morgenstelle die hier nicht aufgelistet, aber gewiss nicht vergessen sind.

Gans besonders in Erinnerung bleiben wird mir meine 4A32 Büro-Crew: Mike Kaupert, Markus Höldrich, Dr. Stefan Polnik und Dr. Ulrich Woiwode.

Zuletzt möchte ich mich bei meiner Familie bedanken, welche mir das Studium und die Promotion überhaupt erst ermöglicht hat, insbesondere von Herzen bei meiner Mutter Cordula Sievers, welche mich stets moralisch, logistisch und finanziell unterstützt hat.

Table of contents

Abbreviations

μESI:	micro electrospray ionization
ADC:	analog digital converter
BIRD:	blackbody infrared radiation-induced dissociation
CCS:	collisional cross section
CD:	circular dichroism spectroscopy
CE:	collision energy
CID:	collision induced fragmentation
CLOGP:	logP predicted by group contribution calculation method
CLP:	cyclic lipopeptides
CoMFA:	comparative molecular field analysis
CoMSIA:	comparative molecular similarity analysis
CP/MAS 29Si- NMR:	cross polarization/magic angle spinning nuclear magnetic resonance spectrometry
CS:	chiral selective stationary phase
CSP-NMR:	chemical shift perturbation NMR
Da, kDa:	Dalton
DFT:	density functional theory calculations
DP:	declustering potential
DP:	Declustering potential
ECD:	electron capture dissociation
EDD:	electron detachment dissociation
ELISA:	enzyme-linked immunosorbent assays
EMH-Horn:	electro multiplier horn
ESI:	electrospray ionization
ESI-MS:	electrospray ionization mass spectrometry
ESI-TOF-HRMS:	electrospray ionization time of flight high resolution mass spectrometry
FMM:	faradioldimyristate
FMP/FPM:	mixed esters of faradiol and myristic/pamitic acid
FPP:	faradioldipalmitate
FTICR:	fourier transform ion cyclotron resonance
FTIR:	fourier-transform infrared spectroscopy
FWHM:	full width at half maximum of peak
GC-MS:	gas chromatography coupled to mass spectrometry
GLM:	generalized linear model
HDX:	hydrogen-deuterium exchange
HILIC:	hydrophilic interaction chromatography
IDA:	information dependent acquisition
IEM:	ion evaporation model
IEX:	ion exchange chromatography
IMS-MS:	ion mobility mass spectrometry
ITC:	isothermal titration calorimetry
kDa:	kilodalton
kV:	kilovolts
LOD:	limit of detection
logP:	octanol-water-partition coefficient

LOQ:	limit of quantification
LSER:	linear solvation energy relationships
MALDI-TOF:	matrix assisted laser desorption ionization time of flight mass spectrometry
MC:	monte Carlo simmulation
MCA:	multi channel alignment
MCP:	multi channel plate
MD:	molecular dynamics simulation
NMR:	nuclear magnetic resonance spectroscopy
NMR:	nuclear magnetic resonance spectrometry
NOE-NMR:	nuclear overhauser effect NMR
pKa:	negative decadic logarithm of acid constant
PS/DVB:	polystyrene-divinylbenzene copolymer
QSAR:	quantitative structure activity relationship
QSERR:	quantitative structure enantioseparation relationships
QSRR:	quantitative structural retention relationships
SA:	selectand
SFG:	sum frequency generation
SJT:	superconducting tunnel junction
TDC:	time digital converter
UV/VIS:	ultraviolet–visible spectrophotometry

Summary of this PhD-thesis

This cumulative PhD thesis put forward several aspects of the analysis of biomolecules employing high resolution quadrupole time of flight mass spectrometry (HR-QTOF-MS). Particular analytical challenges in the context of the analysis of various classes of analytes, i.e. amino acids and peptides, oligonucleotide complexes, triterpenoid esters, intact proteins have been addressed and new analytical solutions by either liquid chromatographic separation or HR-QTOF-MS, respectively, their hyphenation have been suggested.

This PhD-Thesis is comprised of four thematically distinct parts:

The first part dealt with the stereoselective analysis of amino acids and peptides. In one study, the complete stereoconfiguration of an antimicrobial active lipopeptide, poaeamide, was determined. Lipopeptides are typically synthesized by a non-ribosomal enzymatic peptide synthesis machinery. As result, they frequently contain several D-amino acids providing hydrolysis resistance towards target organism peptidases. As lipopeptides are of general interest for research on and development of antimicrobial compounds, complete structural elucidation is essential, which encompasses also determination of the absolute configurations of the amino acids constituting the respective peptide, which is presented in **Publication II.** The analysis strategy enveloped the incomplete hydrolysis of the peptide yielding overlapping sequence fragments, micro-scale preparative liquid chromatography and stereoconfiguration analysis of hydrolysis fragments by chiral GC-MS, ultimately providing determination of the stereoconfigurations of its comprising amino acids enantiomers. As the latter cannot be distinguished by mass spectrometry alone, HR-QTOF-MS has been hyphenated with appropriate enantioselective chromatography using cinchonan carbamate based chiral stationary phases. This work involved the optimization of the chromatographic and MS conditions and the demonstration of the feasibility of aforementioned phases in providing complementary chromatographic selectivity when compared to RP and HILIC type phases, in detail emphasized in **Publication VIII**. A major challenge was the determination of configurations of amino acids with more than one stereogenic centers (Thr/allo-Thr, Ile/allo-Ile) and those also present as constitutional isomers (Leu/Ile).

A particular intricacy solved by the analysis strategy of a combined LC and GC approach employing reversed phase (RP), hydrophilic interaction (HILIC) and enantioselective stationary phases was the localization of D-Leu discovered in the peptide, as its position was disguised by presence of several Leu residues in the lipopeptide.

To advance insight into the enantioselective interaction between the employed LC stationary phases and amino acids and derivatives thereof, **Publication I** reported the application of quantitative structure-retention and structure-enantioselectivity relationships to quantitatively study enantioselective molecular recognition mechanisms. By employing a Free-Wilson type generalized linear modelling approach, Publication I validated hypotheses that describe binding energy contribution of individual molecular moieties as being linear independent of each other. Major contribution to retention of the analytes could be attributed to pi-interacting derivatization groups, a finding that stands in congruence to experimental findings reported amongst others by Publication II.

With mass spectrometry today representing the chromatographic detection method of choice and consequently entailing the desideratum of stationary phases compatible with this technology, the first part of this thesis was concluded by **Publication VII** enhancing mass spectrometric compatibility of the employed chiral stationary phases. With a hydrolysis stable crosslinked methylpolysiloxane type surface chemistry, also providing a scaffold for various surface ligand modifications by the employed thiol-ene click chemistry, significant enhancement of mass spectrometric compatibility could be demonstrated. Using the enantioselective cinchonan carbamate based chiral stationary phases as an example ligand, ameliorated phase stability and resultant enhancement in mass spectrometric sensitivity was assessed and confirmed by high resolution quadrupole time of flight mass spectrometry.

In the second part, challenging analysis, both from chromatographic and mass spectrometric perspective, of regioisomers of pentacyclic triterpenoid fatty acid esters instable even under soft ionizing conditions was addressed by **Publication V**. Novel esters of triterpenoids with anti-inflammatory potential, amongst them mixed esters of faradiol, myristic and palmitic acid could be confirmed to be present in extracts of   by employment of orthogonal analysis methods, namely NMR, GC-MS and LC-HR-QTOF-MS. In order to address the

challenging liquid chromatographic separation of mixed regioisomeric diesters, molecular shape selective chromatography was employed using C30-type RP-stationary phases tailored for the task. Chromatographic and mass spectrometric requirements, the latter stemming from the astonishing instability of the analytes during ionization in the presence of water, even under the soft ionization conditions encountered in electrospray (ESI) or atmospheric pressure chemical ionization (APCI), could be harmonized by application of a non-aqueous binary eluent system, cold LC column temperatures facilitating entropic optimization of regioisomer separation and adequate application and tuning of parameters of APCI-QTOF-MS for sensitivity, mass accuracy and resolution.

The third part was dedicated to intact protein mass spectrometric analysis **(Publications IV IX X and XI).** Employing the Sciex 5600+ TripleTOFs capability in mass accuracy, mass resolution and sensitivity even for large molecular species by adequate mass spectrometric and chromatographic method development, analytical questions revolving around analysis of intact proteins could be addressed, including antibody characterization and, in a straightforward approach demonstrating mass accuracy and resolution of QTOF, direct confirmation of attachment and correct target location of covalent kinase inhibitors with sub kDa molecular weight to >40kDa protein targets.

Finally, the **fourth part** of this thesis includes two studies **(Publications III and XII)** that examine interaction between G4-DNA-selective ligands and G4-DNA-quadruplexes, the latter representing an in vivo form of a DNA that is of oncological research interest as it is frequently encountered in promoter regions of oncogenes. Ligands specifically binding to this DNA form is subject of research aimed at cancer imaging or potential anticancer drugs.

Study of such non-covalent complexes in solution is preferably performed by NMR. However, NMR spectra interpretation is both regularly and in case of Publication III and XII severely hampered by extensive peak broadening and overlapping as consequence of fast to intermediate exchange rates relative to the NMR chemical shift timescale of ligands occupying different binding sites Fluorescence titration, employed as orthogonal method in

both studies, also could not unequivocally unveil the stoichiometry of the complexes studied in the two publications. The author's contribution was the development of a native electrospray ionization high resolution quadrupole time of flight mass spectrometry (ESI-HR-QTOF) method to elucidate stoichiometry and binding mode of aforementioned complexes. The intricacy to address for both studies was the provision of mass spectrometric method capable of mapping non-covalent complex stoichiometries and properties from solution to the vacuum of the mass spectrometric ion path without distortion for example of secondary structure or ligand binding by the ionization process, by the atmosphere to vacuum transition or flight through the ion path. The native-ESI-QTOF-MS method developed was capable of providing these requirements for noncovalent DNA-ligand complexes of several kDa molecular mass, yet still allowing to quantitively monitor specific binding of very low molecular (e.g. ammonium NH_4^+) species to the complex.

Overall, the studies summarized in **this Thesis**, demonstrated the great utility and wide area of application of high-resolution quadrupole time of flight mass spectrometry, either in its hyphenated form with liquid chromatography or as direct infusion-MS, to solve challenging analytical questions in the context of (bio)pharmaceutical analysis.

Zusammenfassung der Dissertationsschrift

Diese kumulative Dissertationsschrift beschreibt Beiträge zu verschiedenen Aspekten der Analytik von Biomolekülen unter Anwendung hochauflösender quadrupol Flugzeit-massenspektrometrie. Spezielle Herausforderungen im Kontext der Analytik verschiedener Analytenklassen, z.B. Aminosäuren und Peptide, Oligonukleotidkomplexe, Triterpenoidester und Intaktproteine wurden adressiert und neue analytische Lösungswege mittels Flüssigchromatographie oder HR-QTOF-MS oder der Kopplung beider wurden vorgeschlagen.

Diese Dissertationsschrift gliedert sich in vier thematisch abgegrenzte Teile:

Der erste Teil beschäftigt sich mit der stereoselektiven Analyse von Aminosäuren und Peptiden. In einer Studie wurde die Stereokonfiguration des antimikrobiell wirksamen Lipopeptides Poaeamid zur Gänze aufgeklärt. Lipopeptide werden regulär auf nicht-ribosomalem Wege synthetisiert. Daher enthalten sie häufig mehrere D-Aminosäuren, welche ihnen Hydrolysebeständigkeit gegenüber Peptidasen der Zielorganismen verleihen. Da Lipopeptide von generellem Interesse der Forschung an und Entwicklung von antimikrobiellen Substanzen sind, ist ihre vollständige Strukturaufklärung, welche die Bestimmung der absoluten Konfiguration der Aminosäuren beinhaltet, essentiell und wird in **Publikation II** vorgestellt. Die Analysenstrategie umfasste die partielle Hydrolyse, die überlappende Sequenzfragmente bereitstellte, preparative Flüssigchromatographie im Mikromaßstab und Analyse der Stereokonfiguration der Fragmente via chiraler GC-MS, was letztlich die Bestimmung der Stereokonfigurationen der Aminosäurenenantiomere gestattete. Da letztere nicht massenspektrometrisch unterschieden werden können, erfolgte Kopplung eines HR-QTOF-MS an eine geeignete enantioselektive Flüssigchromatographie unter Verwendung von Cinchonancarbamat basierten chiralen stationären Phasen. Die Arbeiten umfassten die Optimierung der chromatographischen und MS-Parameter und es wurde die Brauchbarkeit der vorgenannten Phasen für die Bereitstellung komplementärer chromatographischer Selektivität in Bezug auf RP und HILIC demonstriert, letzteres detaillierter beschrieben in **Publikation VIII.** Eine große Herausforderung war die

Bestimmung der Konfiguration der Aminosäuren mit mehr als einem stereogenem Zentrum (Thr/allo-Thr, Ile/allo-Ile) und derer, die als Konstitutionsisomere vorlagen (Leu/Ile).

Eine besonderes Problem, gelöst von der Analysenstrategie durch kombinierten Einsatz von LC und GC, unter Verwendung von Umkehrphasen (RP), Phasen mit hydrophiler Interaktion (HILIC) und enantioselektiver stationärer Phasen, war die Lokalisierung des im Peptid festgestellten D-Leu, da seine Position durch die Gegenwart mehrerer Leu im Lipopeptid verschleiert wurde.

Um das Verständnis der enantioselektiven Interaktion zwischen den eingesetzten flüssigchromatographischen stationären Phasen und Aminosäuren und Derivaten letzterer zu vertiefen, berichtet **Publikation I** über die Anwendung quantitativer Struktur-Retentions- und Struktur-Enantioselektivitäts-Beziehungen für die quantitative Untersuchung enantio- selektiver molekularer Erkennungsmechanismen. Durch Anwendung generalisierter linearer Modellierung vom Free-Wilson-Typ, konnte Publikation I die Hypothese validieren, dass der Beitrag individueller Molekülteile zur Bindungsenergie linear unabhängig ist. Der Hauptbeitrag zur Retention der Analyten konnte auf π-Interaktionen mit den Derivatisierungssubstituenten zurückgeführt werden, ein Befund der sich kongruent zu den Ergebnissen experimenteller Studien, unter anderen auch jenen von Publikation II, verhält.

Da Massenspektrometrie heute die chromatographische Detektionsmethode der Wahl darstellt, was folgerichtig einen Bedarf an zu dieser Technologie kompatibler stationärer Phasen nach sich zieht, schließt der erste Teil dieser Dissertationsschrift mit **Publikation VII**, in welcher die Massenspektrometriekompatibilität der eingesetzten chiralen Phasen verbessert wurde. Mit einer Oberflächenchemie auf Basis von hydrolysestabilen quervernetzten Methylpolysiloxan, welche mittels Thiol-ene-click-Reakivität eine Plattform für Modifikation mit verschiedenen Oberflächenliganden bereitstellt, konnte eine signifikante Verbesserung der Massenspektrometriekompatibilität demonstriert werden. Anhand enantioselektiver Cinchonacarbamate als Beispielliganden wurde die erhöhte Phasenstabilität und die daraus resultierende Verbesserung der massenspektrometrischen Sensitivität via hochauflösender quadrupol-Flugzeitmassenspektrometrie untersucht und bestätigt.

Im zweiten Teil dieser Dissertationsschrift wurde die anspruchsvolle Analytik, sowohl aus chromatographischer als auch massenspektrometrischer Perspektive, von Regioisomeren pentazyklischer Triterpenoidfettsäurenester, welche auch unter milden Ionisationsbedingungen instabil sind, von **Publikation V** adressiert. Bisher unbekannte Triterpenoidester mit entzündungshemmenden Potential, unter ihnen gemischte Ester von Faradiol, Myristinsäure und Palmitinsäure, konnten unter Anwendung orthogonaler analytischer Methoden, namentlich NMR, GC-MS und LC-HR-QTOF-MS, in Extrakten von *Calendula officinalis* nachgewiesen werden. Um die herausfordernde chromatographische Trennung von gemischten regioisomeren Diestern zu adressieren, wurden speziell an diese Aufgabe angepasste C30-RP stationäre Phasen für die molekülformselektive Chromatographie eingesetzt. Chromatographische und massenspektrometrische Herausforderungen, letztere resultierend aus der bemerkenswerten Instabilität der Analyten während der Ionisation in Gegenwart von Wasser, sogar unter den milden Bedingungen der Elektrospray (ESI) oder der chemischen Ionisation bei Atmosphärendruck (APCI), konnten miteinander harmonisiert werden. Erreicht wurde dies durch Anwendung eines binären, nicht-wässrigen Eluentensystems, kalter und somit entropisch für die Regioisomerseparation günstiger LC-Säulentemperatur und adäquater Anwendung und Anpassung der Parameter des APCI-QTOF-MS für Sensitivität, Massengenauigkeit und Auflösung.

Der dritte Teil der vorliegenden Dissertationsschrift war der Massenspektrometrie intakter Proteine gewidmet **(Publikationen IV IX X and XI).** Unter Einsatz der Massengenauigkeit, Massenauflösung und auch für große molekulare Spezies ausreichende Sensitivität des Sciex 5600+ TripleTOF konnten, mit angepasster massenspektrometrischer und chromatographischer Methodenentwicklung, analytische Fragestellungen rund um Intaktproteine adressiert werden. Unter diesen, u.a. Antikörpercharakterisierung und eine, die Fähigkeit zur Massengenauigkeit und Auflösung unterstreichende Anwendung des QTOFS, um Bindung und korrekte Lokalisation kovalent bindender Kinaseinhibitoren mit sub-kDa Molekülgewicht an intakte Zielproteine mit >40kDa Molekulargewicht direkt nachzuweisen.

Schließlich enthält der **vierte Teil** dieser Dissertationsschrift zwei Studien **(Publikationen III and XII)**, welche die Interaktion zwischen G4-DNA-selektiven Liganden und G4-DNA-quadruplexen untersuchen. Letztere ist eine in vivo Form der DNA, welche von onkologischem Forschungsinteresse ist, da sie häufig in den Promotorregionen von Onkogenen anzutreffen ist. Liganden, welche spezifisch an diese DNA-Form binden, sind Gegenstand der Forschung welche auf Onko-Imaging und potentielle Krebstherapeutika abzielt.

Die Untersuchung solcher nichtkovalenten Komplexe würde bevorzugt durch NMR vorgenommen. Erschwert wird die Auswertung solcher NMR-Spektren regulär und auch im Falle von Publikation III und XII durch beträchtliche Peakverbreiterung und -überlagerung, beides resultierend aus relativ zur NMR-Zeitachse schnellen bis mittelschnellen Austauschraten von Liganden, welche an verschiedene Bindestellen binden.

Fluoreszenztitration, welche in beiden Studien als orthogonale Technik eingesetzt wurde, konnte ebenfalls nicht die Stöchiometrie der Komplexe aufklären. Der Beitrag des Autors war die Entwicklung einer Methode zur hochauflösenden nativen Elektrospray quadrupol-Flugzeitmassenpektrometrie (ESI-HR-QTOF) um Stöchiometrie und Bindungsmodus vorgenannter Komplexe zu erhellen. Die zu überwindende Hürde war die Bereitstellung einer Methode, welche in der Lage ist, Stöchiometrie und Eigenschaften nichtkovalenter Komplexe unverändert aus der Lösung in das Vakuum des massenspektrometrischen Ionenpfads zu übertragen, ohne hierbei durch Ionisationsprozess, Übergang von Atmosphärendruck zu Vakuum oder den Flug durch den Ionenpfad zum Beispiel Sekundärstrukturen oder Ligandenbindung zu verfälschen.

Die entwickelte native-ESI-QTOF-MS Methode kann diese Anforderungen für nichtkovalente DNA-Ligand-Komplexe von mehreren kDa Molekülgewicht erfüllen, und dennoch die spezifische Bindung niedermolekularer Spezies (z.B. Ammonium NH_4^+) an den Komplex nachweisen.

Insgesamt konnten die in **dieser Dissertationsschrift** zusammengefassten Studien die große Nützlichkeit und Anwendungsbreite der hochauflösenden quadrupol-Flugzeitmassenspektrometrie, entweder gekoppelt an entsprechende Flüssigchromatographie oder als direkt-infusions-MS, zur Lösung verschiedener herausfordernder analytischer Fragestellungen im Kontext der (bio)pharmazeutischen Analytik veranschaulichen.

List of publications

According to §6 Abs. 2 Satz 3 Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Tübingen

Accepted Manuscripts

Publication I	
Title	Ligand-receptor binding increments in enantioselective liquid chromatography
Authors	***Sievers-Engler, Adrian****; Lindner, Wolfgang; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1363), **2014** p. 79-88
DOI	http://dx.doi.org/10.1016/j.chroma.2014.04.077

Publication II	
Title	Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides
Authors	*Gerhardt, Heike; **Sievers-Engler, Adrian**; Jahanshah, Ghazaleh; Pataj, Zoltan; Ianni, Federica; Gross, Harald; Lindner, Wolfgang; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1428) 2016, p. 280-291
DOI	http://dx.doi.org/10.1016/j.chroma.2015.05.065

Publication III	
Title	Molecular Recognition and Visual Detection of G-Quadruplexes by a Dicarbocyanine Dye
Authors	*Karg, Beatrice; Funke, Andrea; Ficht, Anna; **Sievers-Engler, Adrian**; Laemerhofer, Michael; Weisz, Klaus*
Published in	Chemistry - A European Journal (21), 2015, Issue 39, p. 13802-13811
DOI	http://dx.doi.org/10.1002/chem.201502118

Publication IV	
Title	Targeting the Gatekeeper MET146 of C-Jun N-Terminal Kinase 3 Induces a Bivalent Halogen/Chalcogen Bond
Authors	*Lange, Andreas; Guenther, Marcel; Buettner, Felix Michael; Zimmermann, Markus O.; Heidrich, Johannes; Hennig, Susanne; Zahn, Stefan; Schall, Christoph; **Sievers-Engler, Adrian**; Ansideri, Francesco; Koch, Pierre; Laemmerhofer, Michael; Stehle, Thilo; Laufer, Stefan A.; Boeckler, Frank M*
Published in	Journal of The American Chemical Society (137),(2015, Issue 46, p. 14640-14652
DOI	http://dx.doi.org/10.1021/jacs.5b07090

Publication V

Title	Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS
Authors	*Nicolaus, Christoph;* ***Sievers-Engler, Adrian;*** *Murillo, Renato; D'Ambrosio, Michele; Laemmerhofer, Michael; Merfort, Irmgard*
Published in	Journal of Pharmaceutical and Biomedical Analysis (118), 2016, p. 195-205
DOI	http://dx.doi.org/10.1016/j.jpba.2015.10.025

Publication VI

Title	Preparation of fluorescent labeled gentamicin as biological tracer and its characterization by liquid chromatography and high resolution mass spectrometry
Authors	*Woiwode, Ulrich;* ***Sievers-Engler, Adrian;*** *Laemmerhofer, Michael*
Published in	Journal of Pharmaceutical and Biomedical Analysis (121), 2016 p. 307-315
DOI	http://dx.doi.org/10.1016/j.jpba.2015.12.053

Publication VII

Title	Surface-crosslinked poly(3-mercaptopropyl)methylsiloxane-coatings on silica as new platform for low-bleed mass spectrometry-compatible functionalized stationary phases synthesized via thiol-ene click reaction
Authors	*Zimmermann, Aleksandra; Horak, Jeannie;* ***Sievers-Engler, Adrian;*** *Sanwald, Corinna; Lindner, Wolfgang; Kramer, Markus; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1436) 2016, p. 73-83
DOI	http://dx.doi.org/10.1016/j.chroma.2016.01.058

Publication VIII

Title	Peptide Analysis: Zwitterionic Chiral Ion-Exchangers as Complementary Option to HILIC and to Reversed-Phase Chromatography
Authors	*Zhang, Tong; Holder, Emilie; Franco, Pilar; Laemmerhofer, Michael;* ***Sievers-Engler, Adrian;*** *Gerhardt, Heike; Gross, Harald; Lindner, Wolfgang*
Published in	LC GC chromatographyonline.com (29), 2016, issue 3, p. 112-128
Link	http://www.chromatographyonline.com/peptide-analysis-zwitterionic-chiral-ion-exchangers-complementary-option-hilic-and-reversed-phase-ch

Publication IX

Title	Gold nanoparticle-conjugated pepsin for efficient solution-like heterogeneous biocatalysis in analytical sample preparation protocols
Authors	*Hoeldrich, Markus;* ***Sievers-Engler, Adrian;*** *Laemmerhofer, Michael*
Published in	Analytical and Bioanalytical Chemistry (408), 2016, issue 20, p. 5415-

	5427
DOI	http://dx.doi.org/10.1007/s00216-016-9657-y

Publication X

Title	Tri- and Tetra-substituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3
Authors	*Muth, Felix; Ahmed El-Gokha, Ahmed; Ansideri, Francesco; Eitel, Michael; Döring, Eva;* **Sievers-Engler,Adrian**; *Lange, Andreas, Boeckler, Frank M.;Lämmerhofer, Michael; Koch, Pierre, Laufer, Stefan A.*
Published in	Journal of Medicinical Chemistry (60), 2017, issue 2, p. 594-607.
DOI	http://dx.doi.org/10.1021/acs.jmedchem.6b01180

Publication XI

Title	Papain-functionalized gold nanoparticles as heterogeneous biocatalyst for bioanalysis and biopharmaceuticals analysis.
Authors	*Liu,Siyao,; Höldrich, Markus;* **Sievers-Engler, Adrian**; *Horak, Jeannie; Lämmerhofer, Michael*
Published in	Analytica Chimica Acta (963), 2017, p. 33-43
DOI	https://doi.org/10.1016/j.aca.2017.02.009

Publication XII

Title	Selective Targeting of G-Quadruplex Structures by a Benzothiazole-Based Binding Motif
Authors	*Buchholz, Ina; Karg, Beatrice; Dickerhoff, Jonathan;* **Sievers-Engler, Adrian**; *Lämmerhofer, Michael; Weisz, Klaus*
Published in	Chemistry - A European Journal (23), 2017, issue 24, p. 5814–5823
DOI	http://dx.doi.org/10.1002/chem.201700298

Publication XIII

Title	Surface-anchored counterions on weak chiral anion-exchangers accelerate separations and improve their compatibility for mass-spectrometry-hyphenation
Authors	*Woiwode, Ulrich;* **Sievers-Engler, Adrian**; *Zimmermann, Aleksandra; Lindner, Wolfgang; Sanchez-Munoz, Orlando L.; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1503), 2017, Bd. 1503, p. 21-31
DOI	http://dx.doi.org/10.1016/j.chroma.2017.04.054

Publication XIV

Title	Trisubstituted Imidazoles with a Rigidized Hinge Binding Motif Act As Single Digit nM Inhibitors of Clinically Relevant EGFR L858R/T790M and L858R/T790M/C797S Mutants: An Example of Target Hopping
Authors	*Juchum, Michael; Guenther, Marcel; Doering, Eva;* **Sievers-Engler, Adrian;** *Laemmerhofer, Michael; Laufer, Stefan*
Published in	Journal of Medicinal Chemistry (60), 2017, issue 11, p. 4636-4656
DOI	http://dx.doi.org/10.1021/acs.jmedchem.7b00178

Publication XV

Title	Design, Synthesis, and Biological Evaluation of Novel Type I1/2 p38α MAP Kinase Inhibitors with Excellent Selectivity, High Potency, and Prolonged Target Residence Time by Interfering with the R-Spine.
Authors	*Walter,Niklas M.;Wentsch,Heike K.; Bührmann, Mike;Bauer, Silke M.;Döring,Eva;Mayer-Wrangowski,Svenja;* **Sievers-Engler, Adrian**; *Willemsen-Seegers,Nicole; Zaman,Guido; Buijsman, Rogier; Lämmerhofer, Michael; Rauh,Daniel; Laufer,Stefan A.*
Published in	Journal of Medicinal Chemistry (60), 2017, issue 19, p. 8027-8054
DOI	http://dx.doi.org/10.1021/acs.jmedchem.7b00745

Submitted manuscripts

Manuscripts in preparation

Author contributions

Publications with first and shared first authorship

Publication I

Title	Ligand-receptor binding increments in enantioselective liquid chromatography
Author	**Contribution**
Sievers-Engler, Adrian	**Programing of software Framework** **Calculations and interpretation of data** **Manuscript writing**
Lindner, Wolfgang	Proofreading of manuscript Discussion
Lämmerhofer, Michael	General Idea Discussion

	Manuscript writing
	Proofreading of Manuscript
	Coordination an financing of project

Publication II

Title	Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides
Author	**Contribution**
Gerhardt, Heike	Gas chromatography coupled mass spectrometry, interpretation of analytical data, optimisation of chemical digestion
Sievers-Engler, Adrian	**QTOF-Mass spectrometry** **Method development for chiral liquid chromatography with Chiralpak Zwix(+/-) phases,** **Optimization of chemical digestion, preparation of digest products, Interpretation of analytical data** **Manuscript writing**
Jahanshah, Ghazaleh	Preparation of Lipopeptides
Pataj, Zoltan	Method development LC-MS
Ianni, Federica	Method development LC-MS
Gross, Harald	Generation and initiation of the project
Lindner, Wolfgang	Proofreading of manuscript Discussion
Lämmerhofer, Michael	Proofreading of manuscript Discussion Corresponding author

Publication V

Title	Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS
Author	**Contribution**
Nicolaus, Christoph	Extraction
Sievers-Engler, Adrian	**QTOF-Mass spectrometry** **preparation of regioselective C30 silica bound phases** **Method development for non-aqueous regioselective chromatography** **Interpretation of analytical data** **Manuscript writing**
Murillo, Renato	Scientific counsel structural elucidation via NMR

D'Ambrosio, Michele	Providing Fardiol-3-monoesters and HPLC-ELS analysis, LC-MS-Analysis
Laemmerhofer, Michael	Proofreading of manuscript Discussion Corresponding author
Merfort, Irmgard	Generation and initiation of the project Proofreading of manuscript Discussion Corresponding author

Publications with substantial contribution

Publication III

Title	Molecular Recognition and Visual Detection of G-Quadruplexes by a Dicarbocyanine Dye
Author	**Contribution**
Karg, Beatrice	NMR spectroscopy and interpretation, manuscript writing
Funke, Andrea	Steady-state fluorescencemeasurements
Ficht, Anna	UV/Vis absorbance measurement
Sievers-Engler, Adrian	**Native non covalent ESI-QTOF-Mass spectrometry** **Interpretation of analytical data** **Manuscript writing**
Laemmerhofer, Michael	Proofreading of manuscript Discussion
Weisz, Klaus	Generation and initiation of the project Proofreading of manuscript Discussion Corresponding author

Publication VII

Title	Surface-crosslinked poly(3-mercaptopropyl)methylsiloxane-coatings on silica as new platform for low-bleed mass spectrometry-compatible functionalized stationary phases synthesized via thiol-ene click reaction
Author	**Contribution**
Zimmermann, Aleksandra	Synthesis of vinyl silica gel, brush type and polysiloxane bonded tBuCQN stationary phases, Stationary phase characterization Manuscript writing
Horak, Jeannie	Manuscript proofreading, Discussion
Sievers-Engler, Adrian	**MS-compatibility tests, LOD/LOQ determinations** **Interpretation of analytical data** **Manuscript writing/proofreading**
Sanwald, Corinna	Stationary phase characterization
Lindner, Wolfgang	Proofreading of manuscript

Kramer, Markus	Cross-polarization magic angle spinning (CP/MAS)NMR characterization of stationary phases
Laemmerhofer, Michael	General Idea Proofreading of manuscript Discussion Corresponding author

Publication VIII

Title	Peptide Analysis: Zwitterionic Chiral Ion-Exchangers as Complementary Option to HILIC and to Reversed-Phase Chromatography
Author	**Contribution**
Zhang, Tong, Holder, Emilie Franco, Pilar	Manuscript writing
Sievers-Engler, Adrian	**QTOF-Mass spectrometry** **Method development for chiral liquid chromatography with Chiralpak Zwix(+/-) phases,** **Optimization of chemical digestion, preparation of digest products, Interpretation of analytical data** **Manuscript proofreading**
Gerhardt, Heike	Lipopeptide digestion, **Manuscript proofreading**
Gross, Harald	Providing Lipopeptide samples, Manuscript proofreading
Lämmerhofer, Michael	Proofreading of manuscript Discussion Corresponding author
Lindner, Wolfgang	Manuscript writing Proofreading of manuscript Discussion

Publication IX

Title	Gold nanoparticle-conjugated pepsin for efficient solution-like heterogeneous biocatalysis in analytical sample preparation protocols
Author	**Contribution**
Hoeldrich, Markus	Preparation of pepsin bearing gold nanoparticles and characterization thereof by VIS Spectroscopy and zeta potential measurement, chromatographic characterization of pepsin activity, Protease activity study of GNP-immobilized Pepsin Interpretation of analytical data Manuscript writing/proofreading
Sievers-Engler, Adrian	**Intact protein mass spectrometry via electrospray quadrupole time of flight experiments** **QTOF Mass spectrometry of protein digests** **Interpretation of analytical data**

	Manuscript writing/proofreading
Laemmerhofer, Michael	General Idea Proofreading of manuscript Discussion Corresponding author

Publication X

Title	Tri- and Tetra-substituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3
Author	**Contribution**
Muth, Felix	Design and Synthesis of Tri- and Tetra-substituted Pyridinylimidazoles, Manuscript writing
Ahmed El-Gokha	Design and Synthesis of Tri- and Tetra-substituted Pyridinylimidazoles, Manuscript writing
Ansideri, Francesco	
Eitel, Michael	
Döring, Eva	In vitro metabolism studies Screening of Metabolites by LC-MS Analysis
Sievers-Engler, Adrian	**Covalent binding Assay for mass spectrometric analysis** **µLC-µESI-QTOF mass spectrometry of intact JNK-inhibitor complexes,** **HR-UHPLC-MS of ligands** **manuscript proofreading**
Lange, Andreas	Cloning/expression/purification of JNK mutant
Boeckler, Frank M	Proofreading of manuscript Discussion
Laemmerhofer, Michael	Proofreading of manuscript Discussion
Koch, Pierre & *Laufer, Stefan A.*	General Idea Corresponding authors

Publication XI

Title	Papain-functionalized gold nanoparticles as heterogeneous biocatalyst for bioanalysis and biopharmaceuticals analysis.
Author	**Contribution**
Liu, Siyao	Preparation of papain bearing gold nanoparticles and characterization thereof by VIS Spectroscopy and zeta potential measurement, papain activity assay Interpretation of analytical data Manuscript writing/proofreading
Höldrich, Markus	
Sievers-Engler, Adrian	**Antibody papain activity assessment via µLC-electrospray quadrupole time of flight mass spectrometry experiments** **QTOF Mass spectrometry of protein digests** **Interpretation of analytical data** **Partial manuscript writing / proofreading**

Horak, Jeannie	Proofreading of manuscript Discussion
Laemmerhofer, Michael	General Idea Proofreading of manuscript Discussion Corresponding author

Publication XII	
Title	Selective Targeting of G-Quadruplex Structures by a Benzothiazole-Based Binding Motif
Author	**Contribution**
Buchholz, Ina	Synthesis of 2-aryl-benzothiazole Ligands, Preparation of oligonucleotides, NMR spectroscopy , Fluorescence measurements, Isothermal titration calorimetry, Manuscript writing
Karg, Beatrice	NMR spectroscopy and interpretation, manuscript writing Manuscript writing
Dickerhoff, Jonathan	Circular Dichroism measurements Manuscript writing
Sievers-Engler, Adrian	**Native non covalent ESI-QTOF-Mass spectrometry of DNA-Ligand complexes and Interpretation of mass spectrometric data Manuscript proofreading**
Laemmerhofer, Michael	Proofreading of manuscript Discussion
Weisz, Klaus	Generation and initiation of the project Proofreading of manuscript Discussion Corresponding author

Publications with minor scientific contributions

Publication	**Title and DOI**	**Contribution**
IV	Targeting the Gatekeeper MET146 of C-Jun N-Terminal Kinase 3 Induces a Bivalent Halogen/Chalcogen Bond https://doi.org/10.1021/jacs.5b07090	• **UPHLC-HR-QTOF-MS of Compounds Ligands** • **Intact protein QTOF-MS of JNK3 wildtype and mutants**
VI	Preparation of fluorescent labeled gentamicin as biological tracer and its characterization by liquid chromatography and high resolution mass spectrometry https://doi.org/10.1016/j.jpba.2015.12.053	• **Scientific advice for application of LC-HR-QTOF-MS and interpretation of analytical data**

		• **Manuscript writing** • **Manuscript proofreading**
XIII	Surface-anchored counterions on weak chiral anion-exchangers accelerate separations and improve their compatibility for mass-spectrometry-hyphenation https://doi.org/10.1016/j.chroma.2017.04.054	• **Discussion,** • **LC-ESI-IT-MS** • **Manuscript proofreading**
XIV	Trisubstituted Imidazoles with a Rigidized Hinge Binding Motif Act As Single Digit nM Inhibitors of Clinically Relevant EGFR L858R/T790M and L858R/T790M/C797S Mutants: An Example of Target Hopping https://doi.org/10.1021/acs.jmedchem.7b00178	• **UPHLC-HR-QTOF-MS of Compounds Ligands**
XV	Design, Synthesis, and Biological Evaluation of Novel Type I1/2 p38α MAP Kinase Inhibitors with Excellent Selectivity, High Potency, and Prolonged Target Residence Time by Interfering with the R-Spine. https://doi.org/10.1021/acs.jmedchem.7b00745	• **UPHLC-HR-QTOF-MS of Compounds Ligands**

List of conference contributions

Oral Presentations

Conference	**Pharmaceutical and Biomedical Innovations, Minisymposium, 2015, Tübingen**
Title of Presentation	**Potential and applications of high-resolution mass spectrometry in pharmaceutical (bio)analysis**

Poster Presentations

Conference	40th International Symposium on Capillary Chromatography (ISCC), 2016, Riva Del Garda
Title of Poster	Screening of target binding of irreversible JNK3 inhibitors on peptide and intact protein level by MicroFlow-LC-QTOF-MS
Authors	Adrian Sievers-Engler, Felix Muth, Silke Bauer, Stefan Laufer, Michael Lämmerhofer

Conference	Anakon 2015, Graz
Title of Poster	Determination of faradiol fatty acid diesters via UPLC-QTOF-MS
Authors	Adrian Sievers-Engler, Christoph Nicolaus , Irmgard Merfort, Michael Lämmerhofer

Conference	International Symposium on Chromatography (ISC) 2014, Salzburg
Title of Poster	Determination of faradiol fatty acid diesters via UPLC-QTOF-MS
Authors	Adrian Sievers-Engler, Christoph Nicolaus , Irmgard Merfort, Michael Lämmerhofer

Conference	DPhG-Jahresversammlung 2013, Freiburg
Title of Poster	Linear Modeling of Ligand-Receptor Binding in Enantioselective Liquid Chromatography
Authors	Adrian Sievers-Engler, Wolfgang Lindner, Michael Lämmerhofer

Conference	HPLC 2013, Amsterdam
Title of Poster	Linear Modeling of Ligand-Receptor Binding in Enantioselective Liquid Chromatography
Authors	Adrian Sievers-Engler, Wolfgang Lindner, Michael Lämmerhofer

Conference	Anakon 2013, Essen
Title of Poster	Free-Wilson analysis of LC-enantioseparation of amino acid derivatives by Quinine based chiral stationary phases.
Authors	Adrian Sievers-Engler, Wolfgang Lindner, Michael Lämmerhofer

Conference Publications

Conference Publication I

Title	UHPLC-QTOF-MS as a valuable tool for the identification of novel faradiol fatty acid diesters of Calendula officinalis flowers
Authors	*Nicolaus, Christoph; **Sievers-Engler, Adrian**; Murillo, Renato; Laemmerhofer, Michael; Merfort, Irmgard*
Published in	Planta Med 2015; 81 - PW_177
DOI	https://doi.org/10.1186/s13321-016-0119-5

Conference Publication II

Title	11th German Conference on Chemoinformatics (GCC 2015)
Authors	*Fechner et al.*
Published in	Journal of Cheminformatics2016 8(Suppl 1):18
DOI	https://doi.org/10.1186/s13321-016-0119-5

General introduction I

Short primer on quadrupole time of flight mass spectrometry

Time of flight mass spectrometry represents, predecessed only by sector field type instruments[8], the second oldest mass analyzer[10] technology in mass spectrometric research, with its original description by Stephens et al.[20] dating back to 1946 and its first practical implementation in 1948 by Cameron and Eggers[3]. As with quadrupole or ion trap (Paul and Penning types), time of flight mass spectrometry required the advent of soft ionization techniques like electrospray (ESI)[21, 9] and matrix assisted laser desorption ionization (MALDI)[11] in order to fully integrate the technique into chemical, biological and pharmaceutical fields of research[10]. In contrast to triple quadrupole (QqQ) type instruments as most prominent exponent of tandem mass spectrometers, already devised in the late 1970s by Yost and Enke[22], tandem-TOF systems became of interest not before 20 years later with the development of TOF/TOF[7] and quadrupole-TOF (QTOF)[17]. While TOF/TOF remained restricted mainly to offline analysis (MALDI) of biological macromolecules, QTOF was rapidly embraced and its use spread into various application fields of analytical chemistry/mass spectrometry[5], especially as QTOF systems could be readily hyphenated to gas and liquid chromatographic separation, with the latter using electrospray ionization as coupling ion source resembling the standard configuration (LC-ESI-QTOF) encountered in contemporary mass spectrometric laboratories.
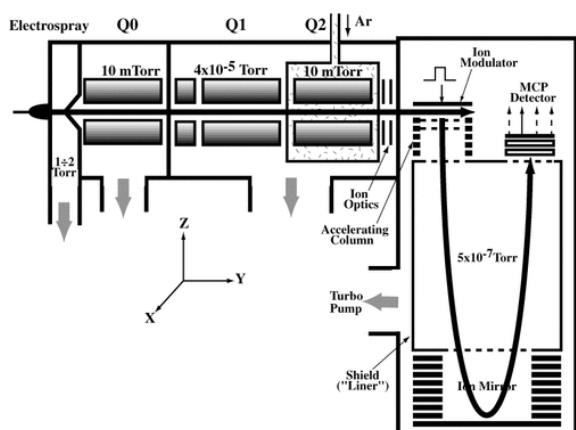

Figure TOF 1 QTOF ion path. Qudrupole (Q1,q2) section to the left, orthogonal extraction drift tube with pusher, reflectron and detector on the right side. The back arrow denotes the path of ions during scan. Reprint from Chernushevich, I. V., Loboda, A. V. and Thomson, B. A. (2001), An introduction to quadrupole–time-of-flight mass spectrometry. J. Mass Spectrom., 36: 849-865An introduction to quadrupole–time-of-flight mass spectrometry, Copyright (2014) with permission from Wiley

As of date of this thesis, the typical bioanalytical or pharmaceutical analytical LC-MS laboratory is equipped with triple quadrupole (QqQ), quadrupole-time of flight (QTOF) or

orbitrap type[15, 14] mass spectrometers. Of those three, the QqQ is still the predominant form, especially when targeted quantitative analysis is performed. This bias towards the less accurate and resolving system is founded, as assessed by Chernushevich et al. [5] 6 years after the introduction of the first QTOF by Morris et al.[17], that QqQ are the more sensitive platform for quantitative LCMS analysis, with generally lower limit of detection (LOD) and quantification (LLOQ) and greater linear range when operated in tandem mode. The latter describes the operational mode of isolation of the analytes of interest m/z by the first quadrupole(Q), fragmentation by collision induced dissociation (CID) by the second quadrupole (q2, collision chamber) and isolation of specific fragments by the third quadrupole (Q3) prior to detection, termed single reaction monitoring (SRM) or multi reaction monitoring (MRM) if performed sequential for multiple analytes. As of date of this thesis, this assertion still holds true in general, yet there are analytical situations where the QTOF excels: when interferences are unresolved by chromatography and subsequent quadrupole ion selection i.e. discrimination of quasi-isobaric species with mass differences of less than 1 Da. Especially with very complex samples of biological origin, the substantially better mass accuracy and mass peak resolution of QTOF relative to QqQ systems, can provide quantitative equivalence or occasionally even an advantage for the time of flight systems.

QTOF systems excel at three performance characteristics, that in addition enable operational modes inaccessible to QqQ sytems**:   mass accuracy, mass resolution**, and in full mass spectra **scan speed** of the time of flight compartment.

Often quoted **high mass range** is, however, not amongst these. As a matter of fact, the large mass range stated by the vendors, e.g. 40000 Da for the Sciex 5600+ used in this thesis, is a theoretical value for a single charged ion, resultant of ion path parameters like pusher acceleration voltage, ion optic geometry and drift path length and detector speed, and largely unused by actual ions entering a QTOF. Ions have to pass multiple quadrupoles Q1, q2 and transmission RF quadrupoles, that restrict efficient passage of ions with m/z ratios larger than 4000, requiring the multiple charging provided by the electrospray ion source as prerequisite to actually reach the detector. The reason QTOF is superior to QqQ for large molecule analysis is not due to the transmissions characteristics, but provided by high mass accuracy and resolution required to discriminate isotopically and or chemically polydisperse

multicharged ion species. See introduction of Part III (intact protein mass spectrometry) of this thesis for reference.

Mass accuracy, as key function of HR-QTOF-MS, needs careful attention during operation and data analysis. For the time of flight compartment of QTOF systems it is specified in the low ppm range, e.g. below 2 ppm for the Sciex 5600+, which accounts for mass differences in the range of one to four electron masses for small molecules (1-1000 Da). Mass accuracy is calculated via:

$$[ppm] = \frac{m_{measured} - m_{theoretical}}{m_{theoretical}} \cdot 10^6$$

Where $m_{measured}$ is the actual signal observed and $m_{theoretical}$ is the calculated monoisotopic mass. However, mass accuracy is not a fixed value, but instable over short periods of time, in the range of minutes to one hour as opposed to quadrupole that retain mass accuracy over longer periods. Temperature deviations subsequent to the preceding mass axis calibration are the main cause by introducing changes in high power supply output voltages and effective TOF drift path length that result in offset between theoretical and observed ion m/z[5].

Technical countermeasures reduce the observed mass accuracy drift to approximately 50 ppm/h, still necessitating frequent mass axis calibration up to one calibration per sample analyzed to maintain maximum mass accuracy. This is performed by intermitted measurement of a calibration mix spanning the m/z range of interest. Various calibrants have been described, amongst them dedicated mixtures, polymers[13, 6] or noncovalent salt clusters[2]. For the studies of this thesis either polypropylene mixtures (PPG) or low concentrated sodium acetate forming non covalent clusters were employed with the latter having the benefit of less memory effect in the ion source[12]. Furthermore, mass axis stability can be achieved by constant introduction of a "lock mass", i.e. a compound of known m/z ratio used for real time, also termed internal or online, mass axis stabilization by single point calibration amenable when drift tube flight time and m/z are in linear relationship[18, 4] or by a posteriori, termed external / offline mass axis recalibration in the acquired raw data prior to data analysis[19, 16]

Mass resolution as the second key function of QTOF-systems provides usually between R 30000-40000 resolution of mass peaks for MS^1-TOF scan mode. In MS/MS mode, this value can differ, e.g. for the Sciex 5600$^+$ in a tradeoff between resolution and sensitivity the operator can choose between R=20000 (high sensitivity) and R=35000 (high resolution). Since MS/MS spectra are less densely populated, lower resolution is often acceptable. Mass resolution is calculated via:

$$R = \frac{m}{\Delta m_{FWHM}}$$

Where R is the mass resolution, m/z value at the mass peak apex and Δm_{FWHM} is the mass peak with at 50% of the maximum intensity.

Mass accuracy and resolution together yield, when compared to QqQ systems, far greater **fidelity of analyte identification or targeting,** respective. It must be noted, that to the quadrupole-part of a QTOF system the same performance characteristic as in a QqQ system applies, i.e. approximately 300ppm mass accuracy and R=1000 resolution, implying, that interferences quasi-isobaric under quadrupole selection during SIM/MRM parent ion selection are fragmented together in the q2. Yet, due to the high resolution and accuracy of the time of flight compartment as opposed to a Q3, even unspecific fragments, e.g. loss of water by both interfering species, can be discriminated. An example of such an analytical challenge resolved by higher analyte targeting fidelity by QTOF was encountered in the experiments for Publication V, where the QTOF exhibited more actual sensitivity than the QqQ usually regarded superior in this respect:
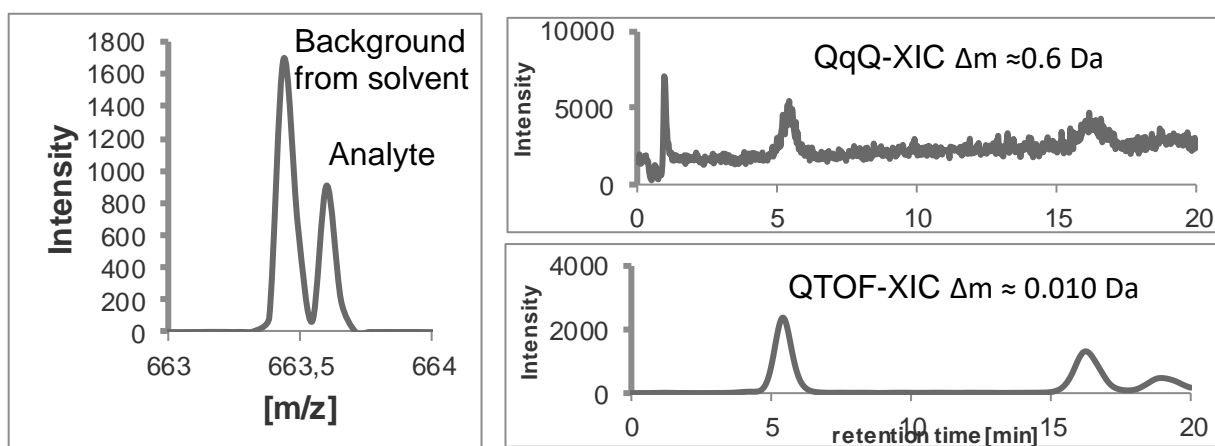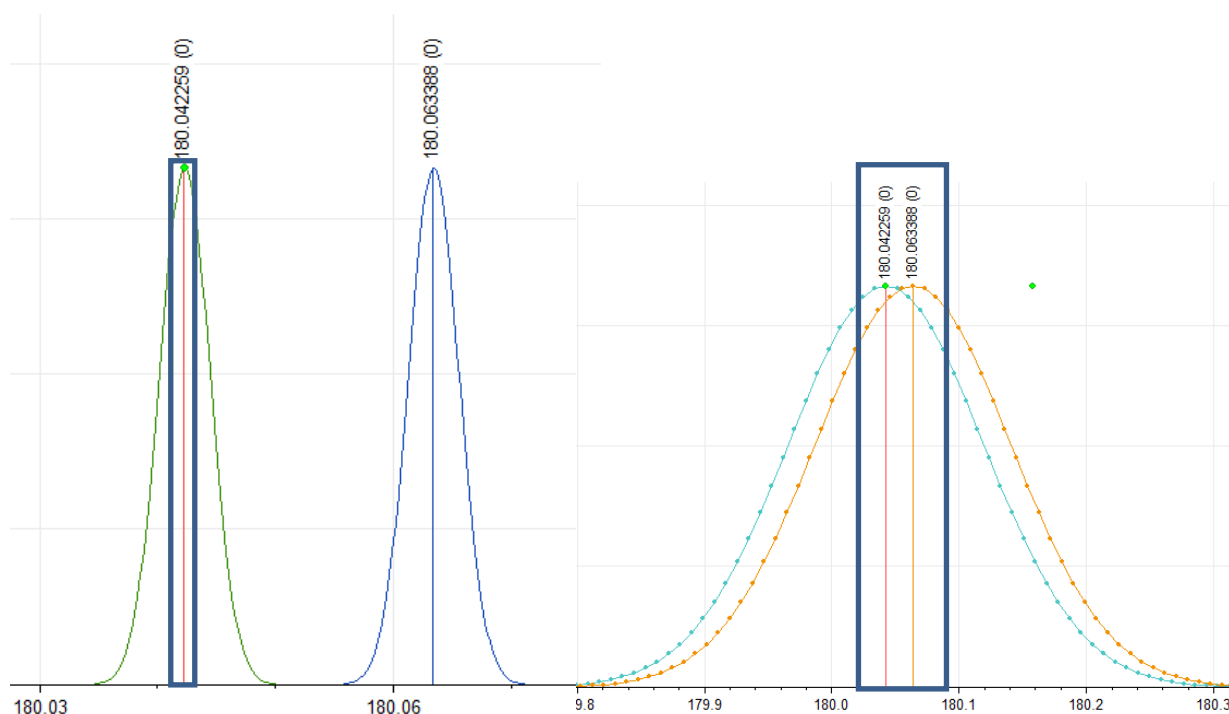
Figure TOF 2 Top: Calculated theoretical spectra of uncharged acetyl salicylic acid (m=180.0423) and Glucose (m= 180.0634). Left: QTOF at R=35000 and 2ppm mass accuracy, left QqQ at R=1000 and 300 ppm. Box indicates mass accuracy corridor, i.e. range within overlapped quasi-isobaric species would not be discriminated. Bottom: Example of increased analyte targeting fidelity and effective sensitivity form Publication V, where quasi isobaric background complicates analysis. Albeit QTOFs are generally inferior regarding sensitivity when compared to QqQ systems, presence of interference could be addressed by mass accuracy and resolution, thereby actually surpassing the latter performance.

Specific operational modes of QTOF systems

The third key function of a QTOF system is the **scan speed** achieved by its time of flight compartment. Flight time of ions for a single raw spectrum (transient) in the drift tube lies in the μs range[5]. To reduce noise and gain viable ion statistics of the mass peaks, multiple transients are summed to yield one spectrum. Effective scan rates therefore achievable are approximately 50 spectra per second, with the Sciex 5600⁺ excelling with up to 100Hz acquisition rate. This high acquisition rate allows operational modes for QTOF systems, that are inaccessible by triple quadrupole instruments and mainly employed for untargeted analysis[23]:

Information dependent acquisition (IDA): The QTOF operates in a cycle of MS1-TOF-Scan and MS/MS scans triggered as a result of the initial MS scan which is examined by an algorithm that identifies mass peaks of interest, e.g. intensity charge state, isotopic pattern and inclusion and exclusion lists. Mass peaks selected are then subjected to MS/MS analysis, i.e. setting of the Q1 transmission window to the respective mass an acquisition of a full product ion spectrum after fragmentation in q2. Collision energy can here be automatically adjusted to m/z ratio of the analyte (rolling collision energy, especially used for peptides). Due to the mode of operation a provision of a few production ion spectra per MS cycle (e.g. typically top 4 to top 20), this mode is employed for untargeted qualitative analysis. MS/MS spectra then allow the identification of the detected compounds (precursors).

Comprehensive MS/MS (SWATH): Ideally, each compound in an untargeted analysis would be accessible quantitively with MS/MS data. Yet, relative to the required performance, slow target mass adjusting time, accuracy and resolution of the Q1 quadrupole prohibit such an approach of completely scanning through each component individually. Instead, the mass spectrum is divided into intervals, by standard 25Da but also individual Q1 precursor isolation widths are possible, that are transmitted by Q1 to fragmentation in q2 and spectrum acquisition in the TOF compartment sequentially.
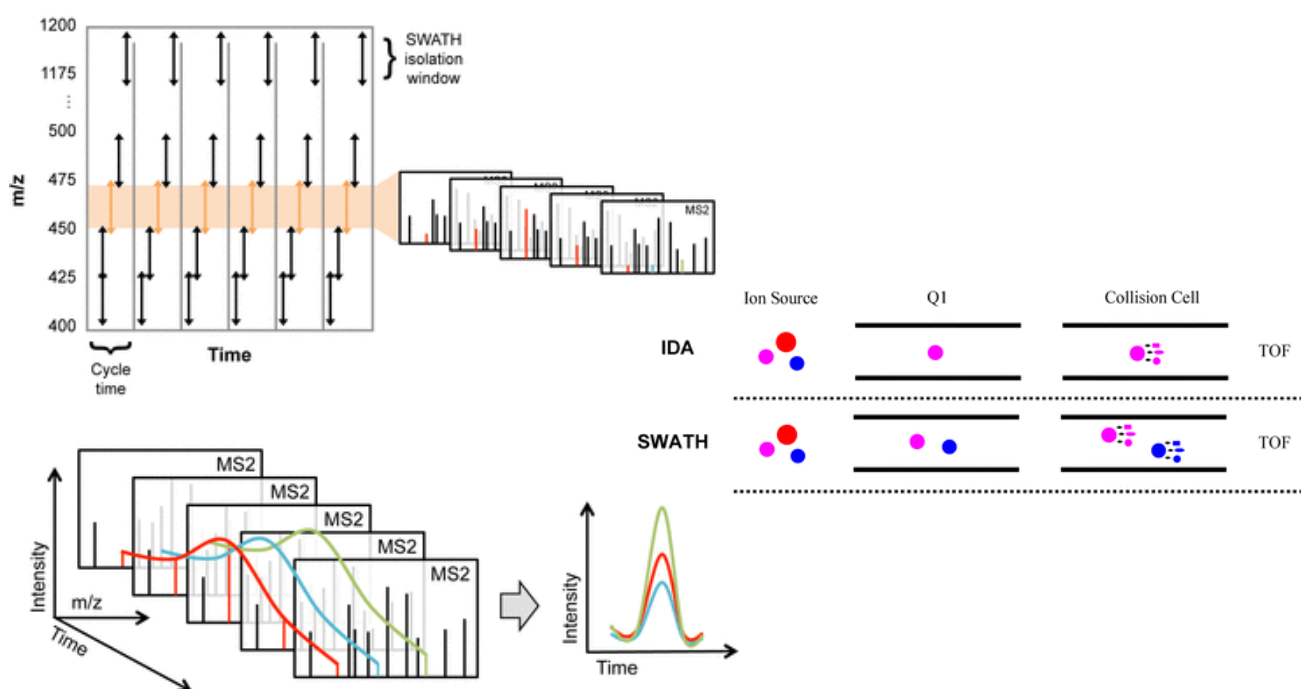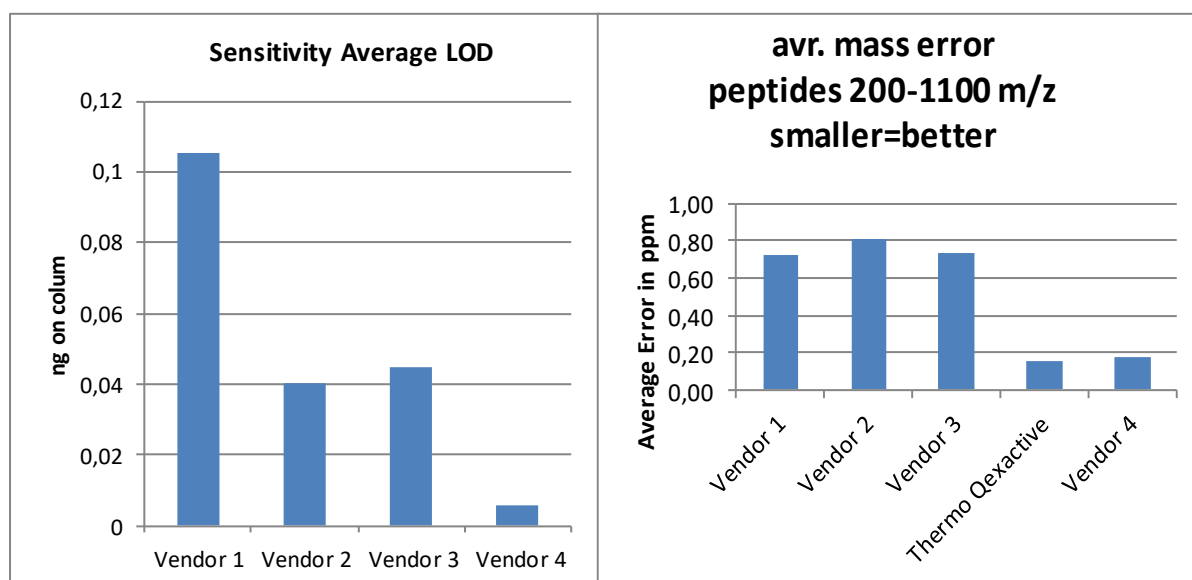

Figure TOF 3Fig TOF: Left SWATH and IDA: Reprint from Comparison of Information-Dependent Acquisition, SWATH, and MSAll Techniques in Metabolite Identification Study Employing Ultrahigh-

Performance assessment of QTOF systems on the market - what to expect

At the beginning of this thesis, the participation in procurement[1] of a research grade quadrupole time of flight mass spectrometer was included. During application procedures for the grant, a system check was developed and submitted to the vendors in order to assess performance indicators like e.g. mass accuracy, mass resolution, scan speed, sensitivity using a variety of analytes ranging from small molecules and metabolites to intact proteins. The overall averages of the results of this assessment are portrayed below. Due to legal considerations, the results are displayed in an anonymized form. It should be noted, that sensitivity, also depicted as overall average over all small molecules, is a strongly analyte and chromatographic method specific result influenced by various factors and should therefore be perceived *cum grano salis*.

Linear Range all systems, bigger=better

Mass error, Lactoglobulin smaller=better

Average Mass error Metabolites, smaller=better
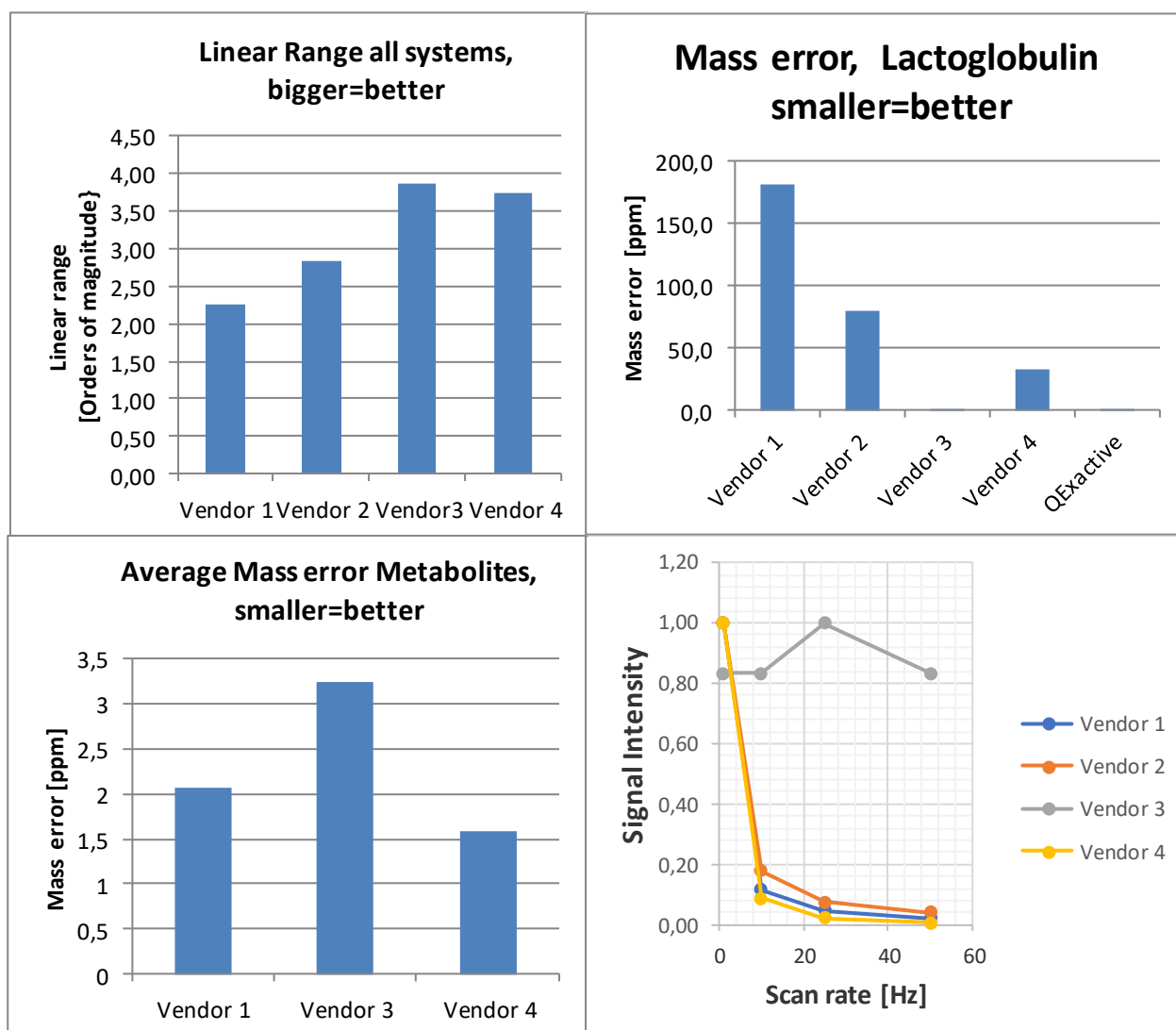
Figure TOF 4Fig. TOF4: Results of performance assessment of research grade quadrupole time of flight mass spectrometers. Note the missing decline of signal vs scan speed only observed with one vendor is due to the employed detection technology that performs independent of scan speed.

References

[1]   Deutsche Forschungsgemeinschaft (DFG) Projektnummer 225950036. (hybrid) tandemmassenspektrometer.

[2]   JF Anacleto, S Pleasance, and RK Boyd. Calibration of ion spray mass spectra using cluster ions. *Organic mass spectrometry*, 27(6):660–666, 1992.

[3]   AE Cameron and DF Eggers Jr. An ion"velocitron". *Review of Scientific Instruments*, 19(9):605–607, 1948.

[4]   Laurence Charles. Flow injection of the lock mass standard for accurate mass measurement in electrospray ionization time-of-flight mass spectrometry coupled with

liquid chromatography. *Rapid communications in mass spectrometry*, 17(13):1383–1388, 2003.

[5]   Igor V Chernushevich, Alexander V Loboda, and Bruce A Thomson. An introduction to quadrupole–time-of-flight mass spectrometry. *Journal of mass spectrometry*, 36(8):849–865, 2001.

[6]   Robert B Cody, Jun Tamura, and Brian D Musselman. Electrospray ionization/magnetic sector mass spectrometry: calibration, resolution, and accurate mass measurements. *Analytical Chemistry*, 64(14):1561–1570, 1992.

[7]   RJ Cotter and TJ Cornish. A tandem time-of-flight (tof/tof) mass spectrometer. *Anal. Chem*, 65:1043–1047, 1993.

[8]   AJ Dempster. A new method of positive ray analysis. *Physical Review*, 11(4):316, 1918.

[9]   John B Fenn, Matthias Mann, Chin Kai Meng, Shek Fu Wong, and Craig M Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926):64–71, 1989.

[10]   Jennifer Griffiths. A brief history of mass spectrometry. *Anal Chem*, 80(15):5678–5683, 2008.

[11]   Franz Hillenkamp and Michael Karas. [12] mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. In *Methods in enzymology*, volume 193, pages 280–295. Elsevier, 1990.

[12]   Simone König and Henry M Fales. Calibration of mass ranges up to m/z 10,000 in electrospray mass spectrometers. *Journal of the American Society for Mass Spectrometry*, 10(3):273–276, 1999.

[13]   Barbara S Larsen and Charles N McEwen. An electrospray ion source for magnetic sector mass spectrometers. *Journal of the American Society for Mass Spectrometry*, 2(3):205–211, 1991.

[14]   Alexander Makarov. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical chemistry*, 72(6):1156–1162, 2000.

[15]   Alexander Makarov, Eduard Denisov, Alexander Kholomeev, Wilko Balschun, Oliver Lange, Kerstin Strupat, and Stevan Horning. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Analytical chemistry*, 78(7):2113–2120, 2006.

[16]   Velitchka V Mihaleva, Oscar Vorst, Chris Maliepaard, Harrie A Verhoeven, Ric CH de Vos, Robert D Hall, and Roeland CHJ van Ham. Accurate mass error correction in liquid

chromatography time-of-flight mass spectrometry based metabolomics. *Metabolomics*, 4(2):171–182, 2008.

[17]   Howard R Morris, Thanai Paxton, Anne Dell, Jean Langhorne, Matthias Berg, Robert S Bordoli, John Hoyes, and Robert H Bateman. High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Communications in Mass Spectrometry*, 10(8):889–896, 1996.

[18]   Jesper V Olsen, Lyris MF de Godoy, Guoqing Li, Boris Macek, Peter Mortensen, Reinhold Pesch, Alexander Makarov, Oliver Lange, Stevan Horning, and Matthias Mann. Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a c-trap. *Molecular & Cellular Proteomics*, 4(12):2010–2021, 2005.

[19]   Craig D Wenger, Graeme C McAlister, Qiangwei Xia, and Joshua J Coon. Sub-part-per-million precursor and product mass accuracy for high-throughput proteomics on an electron transfer dissociation-enabled orbitrap mass spectrometer. *Molecular & Cellular Proteomics*, 9(5):754–763, 2010.

[20]   MM Wolff and WE Stephens. A pulsed mass spectrometer with time dispersion. *Review of Scientific Instruments*, 24(8):616–617, 1953.

[21]   Masamichi Yamashita and John B Fenn. Electrospray ion source. another variation on the free-jet theme. *The Journal of Physical Chemistry*, 88(20):4451–4459, 1984.

[22]   RA Yost and CG Enke. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Analytical chemistry*, 51(12):1251–1264, 1979.

[23]   Xiaochun Zhu, Yuping Chen, and Raju Subramanian. Comparison of information-dependent acquisition, swath, and msall techniques in metabolite identification study employing ultrahigh-performance liquid chromatography–quadrupole time-of-flight mass spectrometry. *Analytical chemistry*, 86(2):1202–1209, 2014.

General Introduction II

Electrospray Ionisation – History, theory and practical implications

Historic development

Observations regarding electrospraying of liquids can be traced back as far as 1750, to French physicist Jean-Antonie Nollet reporting on aerosol formation of water dripping from an electric charged metallic vessel placed near electrical ground. The first physical and mathematical treatment of this phenomenon was given by Rayleigh[60] , devising the first general theory of afore mentioned aerosolization observed in electrospray processes deriving from electrical charge amount, geometry and fluid viscosity a threshold of droplet fission, the latter named after its discoverer Rayleigh-limit.

Further investigations by Zeleny[78, 79, 80] on forms of electrospray emission from capillaries first described the formation of a liquid cone emitting a liquid jet under appropriate electric conditions, providing the experimental basis for Taylors theoretical description of cone jet emission what now is known as the Taylor-cone[68]
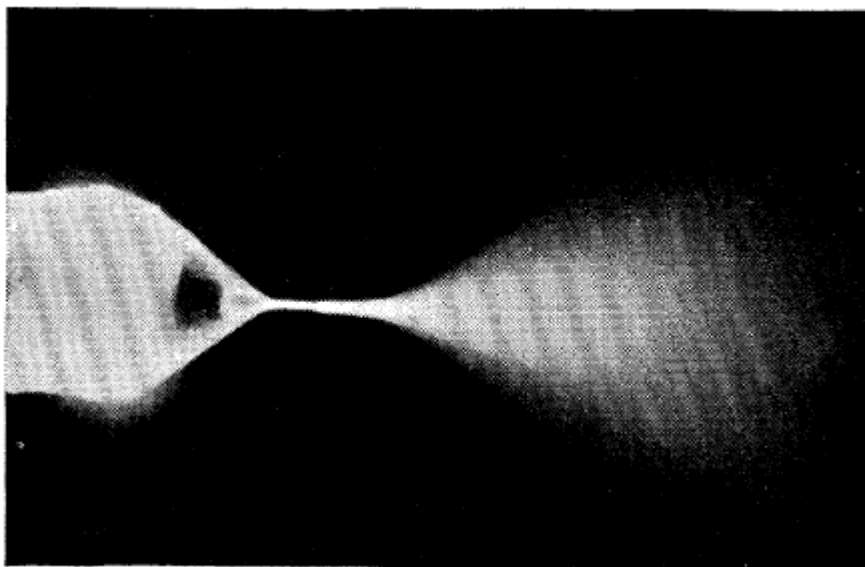
Figure ESI 1 Picture from Zelenys work showing Taylor cone, jet emission and droplet fission as predicted by Rayleigh. Reprint from. "Instability of electrified liquid surfaces." Physical review 10.1 (1917): 1. Copyright (2017) with permission from American Physical Society

Until the mid 1960s electrospray research and development was, from a mass spectrometrists perspective, macroscopically focused such as liquid dynamics or technical applications[7, 8] . The first publication regarding molecular weight determination using

electrospray on volatile solvents can be attributed to Dole[24]. On the basis of earlier theoretical and applied research[34, 70, 60] , however, independent of works in the field of desorption ionization (FDI)[41, 55, 56] also demonstrating ionization on small diameter capillary tips in vacuo, he also proposed macromolecular ion formation by a repetitive solvent evaporation/droplet fission mechanism, later established under the term "Charged residue model" (CRM)[23].
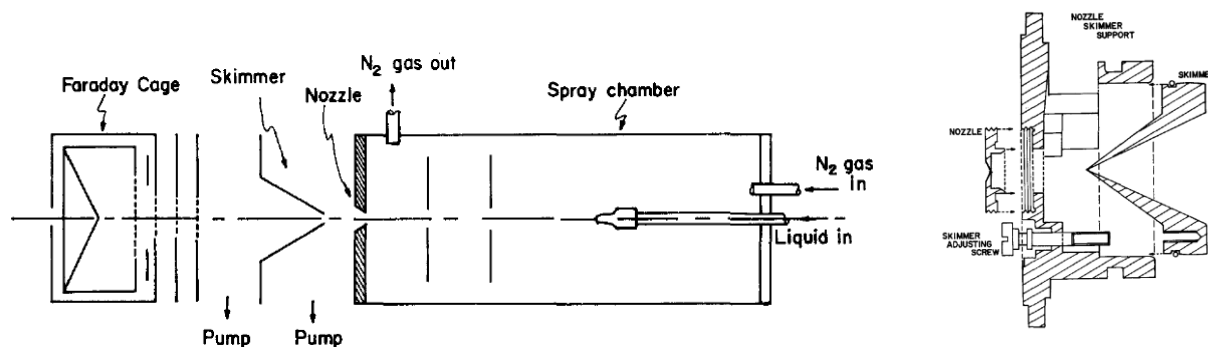

Figure ESI 2Fig ESI3: Doles original electrospray setup and skimmer system. Reprint from: Molecular Beams of Macroions, The Journal of Chemical Physics 49:5, 2240-2249, Copyright (1968), with permission from AIP Publishing

Although featuring already atmospheric pressure, hot nitrogen desolvation assistance, and a 3 stage nozzle/skimmer vacuum system (atmospheric pressure, 0.1mBar, 13µBar), this pioneering setup besides providing a method for generation of molecular beams of both polarities, the findings of Dole and Coworkers[51, 18] are not considered the birthmark of ESI-mass spectrometry, as his detector consisted of a simple Faraday cup and a repeller grid for m/z- filtering to some extent yielding ion current vs. repeller voltage rather than m/z-values. More than the following decade passed till the event of electrospray ionization, as result of the works of Fenn et al.[77].

The strong development of high-performance liquid chromatography HPLC during the 1970s also increased the demand for a suitable coupling technique to mass spectrometry. Various techniques have been proposed. Online eluent drying and ionization, like the moving band interface[52], atmospheric ionization techniques employing radionuclide (^{63}Ni) beta radiation[38, 39, 36, 37] or corona discharge ionization[16] of evaporated analytes, diverting

a small fraction, usually 1/1000, of LC eluent into a standard chemical ionization(CI) interface[9, 4, 53], inlet systems using membrane separators for maintaining CI Vacuum[49, 43] or laser assisted vaporization[13].
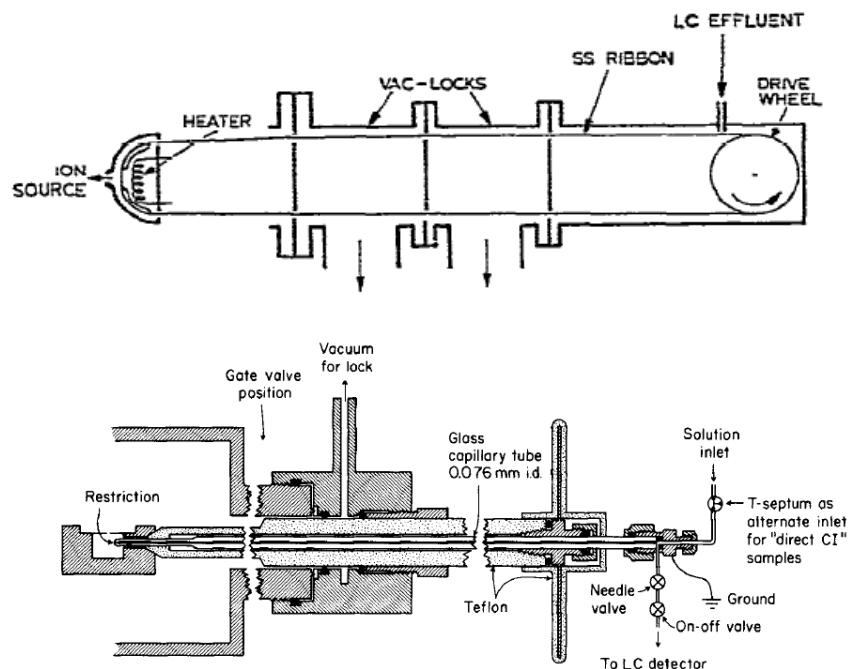

Figure ESI 3Fig ESI3 Two early LC-MS approaches, Moving Band interface and split flow CI interface Reprints from: Direct analysis of liquid chromatographic effluents Journal of Chromatography A, Volume 122, 7 July 1976, Pages 389-396, Copyright (1976) with permission from Elsevier; and from: Liquid chromatography-mass spectrometry. II—continuous monitoring. Biomed. Mass Spectrom., 1: 80-82. 1974, Copyright (1974) with permission from Wiley.

At the end of the decade, a review[5] by Arpino and Guiochon summarizes and evaluate the proposed techniques and, despite the retrospective obvious usability of electrospray for liquid chromatography mass spectrometry coupling (LCMS), consider the split flow interfacing and moving band interfaces the most promising approach.

All of the aforementioned ionization methods differ in the procedure of atomization prior to the common use chemical Ionization to generate molecular ions thereby suffer from CI related limitations like LC-solvent flow and composition constraints, analyte volatility and stability requirements excluding especially biooligomers and –polymers from analysis and unwanted CI side effects i.e. solvent ionization and clustering resulting in crowded lower m/z-spectra region.

Vestal and coworkers established a more straightforward approach with the thermospray ion (TS) source[12, 11, 14] , allowing for LC flow rates up to 2ml/min. In its basic configuration the eluent is simply nebulized into a sonic jet of eluent vapor by electric heating and expanding through a capillary into a vacuum of 3 Torr[14]. Using ionization promoting eluent additives e.g. formic acid, no external ionization source is needed, ion formation results from uneven charge distribution during eluent vaporization and subsequent drying of particles. The original TS source also featured orthogonal spraying, i.e. the path of the vapor jet was perpendicular to the mass spectrometer orifice and ion path[14], enhancing signal to noise ratio. For analytes less prone to ionization by adduct formation, the source could be outfitted with either electron ionization arrangement or an electric discharge electrode pair, the latter resembling an atmospheric pressure chemical ionization (APCI) precursor, albeit operating at reduced pressure. TS had some disadvantages, mainly exerting thermal stress on the analytes and limited solvent compatibility.
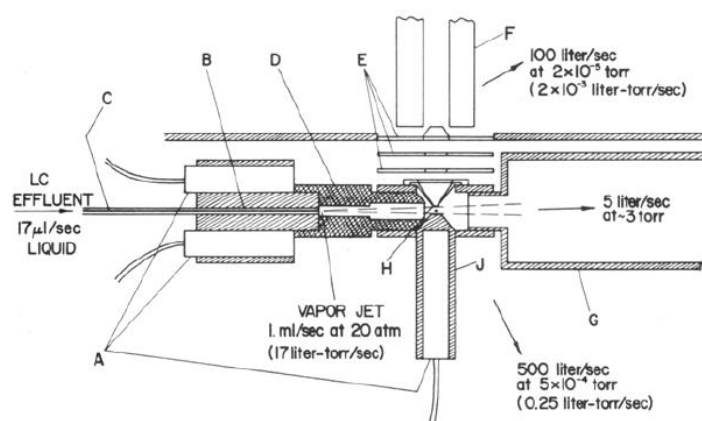
Figure ESI 4 Vestals thermospray (TS) ion source. Reprint from: Thermospray interface for liquid chromatography/mass spectrometry, Analytical Chemistry 1983 55 (4), 750-754, Copyright (1983) with Permission from American Chemical Society

In 1984 Fenn and coworkers[77] demonstrated the first coupling of Doles electrospray ion source to a quadrupole mass spectrometer, showing evidence for the gentle nature of the ionization process by identifying a multitude of noncovalent adduct ion cluster consisting of solvent molecules and a charged central ion. These findings also provided proof for Iribarne-Thomson hypothesis,[42, 69], later known as Ion-evaporation-Model (IEM) of solute ion to gas phase ion transition, as a mechanism concurrent to Doles CRM. In contrast do Doles

macromolecular ions, Fenn also demonstrated the usefulness of electrospray ionization for mass spectrometric analysis of low molecular weight, organic molecules being too instable or nonvolatile to be analyzed by at that time common ionization methods like CI, EI or TS. According to Fenn[73] , and with respect to the Iribarne-Thomson hypothesis, the ionization process of ESI proceeds in a mirror like fashion to Thermo spray ionization. While TS in basic mode produces net charging by atomization of solvent, ESI producing atomization of solvent by electrochemical net charging at the capillary tip.
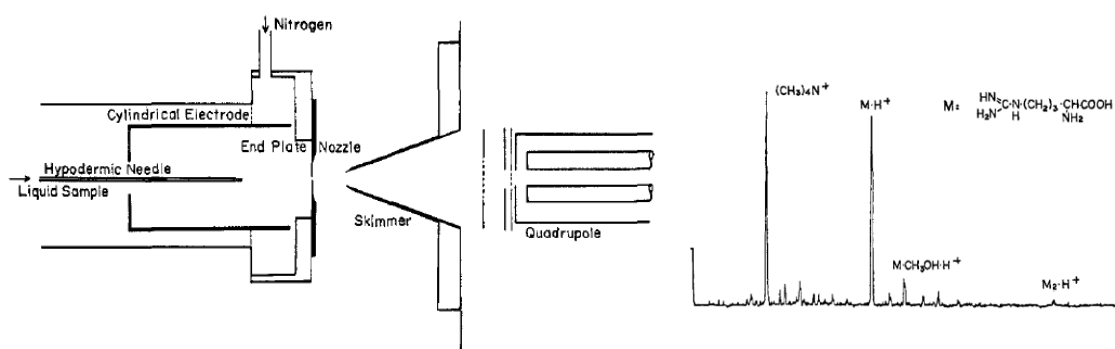

Figure ESI 5 Fenns schematic of the ESI source and example spectrum. Spray jet still parallel to MS orifice and no thermal or pneumatic assistance, yet already counterflow nitrogen. Spectrum showing quasimolecular ions as +H$^+$ adducts and noncovalent adducts of arginine. Reprint from: Electrospray ion source. Another variation on the free-jet theme, The Journal of Physical Chemistry 1984 88 (20), 4451-4459, Copyright (1984) with Permission from American Chemical Society

The full capability of ESI was demonstrated in 1988 on the *36th Annual Conference on Mass Spectrometry* and Allied Topics in the well recognized (7290 cites as of Nov 2016) 1989 publication[73] of Fenn's group, for which he got later awarded the Nobel Prize in 2002. The paper demonstrated the use of ESI, using both polarity, for small molecule of below 100 Da, peptides, oligonucleotides and oligomers in the 3kDa range to various proteins up to 76kDa. The work also again demonstrated the soft nature of the electrospray ionization by demonstrating spectra showing no signs of ionization related fragmentation and the capability of producing noncovalent Coulomb type adducts e.g. [PEG+n*Na$^+$]$^{n+}$ adducts. Fenn and especially his coworker Mann also addressed the challenge of analyte charge multiplicity, at the same time beneficial for low m/z transmitting quadrupoles and maleficial with regard to spectra complexity, by devising a deconvolution algorithm reconstructing the

original molecular weight of a given multiplex signal. With the results of Fenn confirmed by Covey[19] , the "esi revolution" was sparked in 1989, as indicated by Fig ESI7
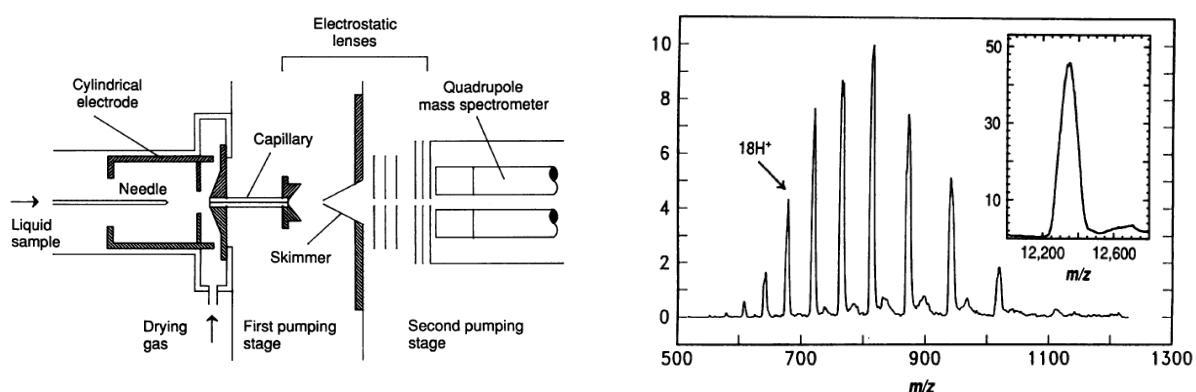

Figure ESI 6 Fenns improved source utilizing a transfer capillary; Mass spectrum of cytochrome c showing ESI typical charge multiplicity and the results of deconvolutions by Fenn&Manns algorithm. Reprint from: Electrospray ion source. Another variation on the free-jet theme, The Journal of Physical Chemistry 1984 88 (20), 4451-4459, Copyright (1984) with Permission from American Chemical Society
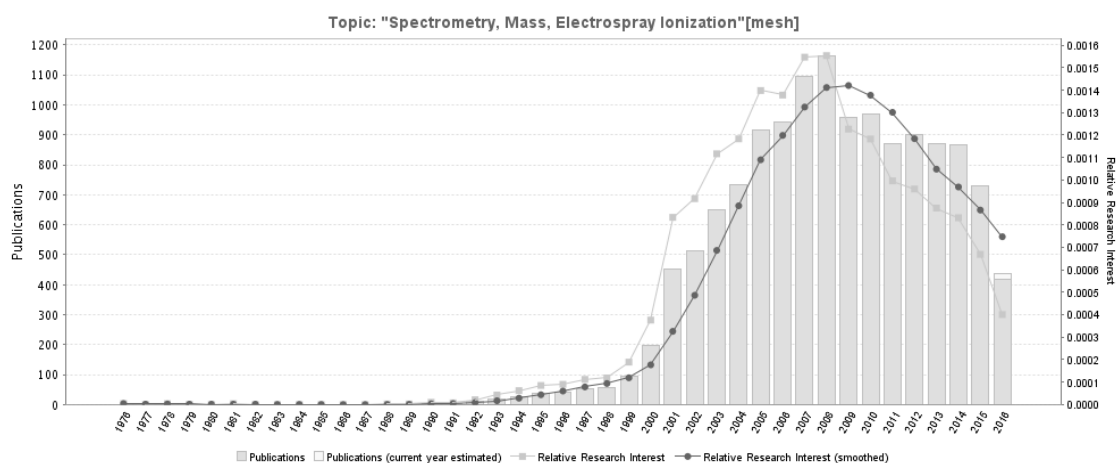
Figure ESI 7 Depicting publication statistics, showing the exponential rise in publications after Fenns paper. Source: gopubmed.org, search words as denoted in title

Fenn's ESI-Source was completed to its common present technical design by enabling pneumatic assistance of electrospray[15], thermal assistance[26] and adoption of the orthogonal design, i.e. the ESI spray jet is introduced orthogonally to the axis of the mass spectrometer orifice entrance. Besides from manufacturer specific minor variations, there have been no essential changes to the electrospray source design.   ESI was enabled as a mass spectrometric standard, as all commercial available general purpose research grade mass spectrometers are equipped with at last an ESI-Source.

Theory of electrospray ionization and practical implications

The process generating ions in an electrospray consists of three[45] sub-processes, (1) dispersion of the analyte containing solvent at a hypodermic capillary into net charge containing fine primary droplets, ~10µm in diameter[15]. (2) evaporation of solvents and Coulomb partition ("explosion") to smaller sub-droplets capable of releasing gas phase ions and (3) production of gas-phase ions.

General setup
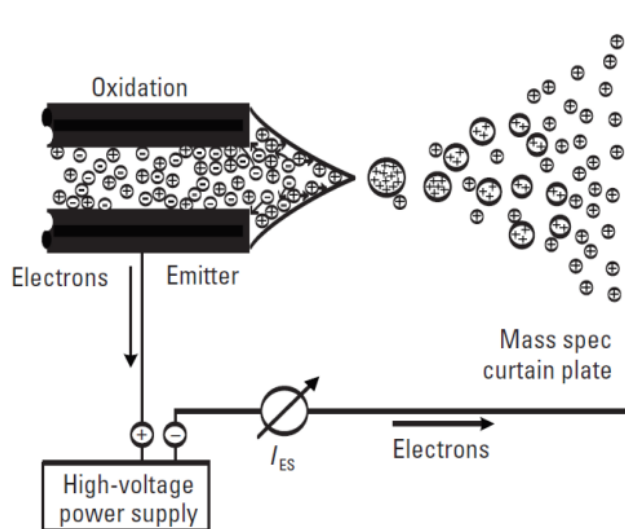
Figure ESITOP 1 Electric flow scheme of an electro spray source. a) general scheme, positive polarity shown, in case of negative ESI mode, electron flow is reversed and reduction is occurring ad emitter tip. Reprinted from Analytical chemistry, 79, 15, Kertesz et al. Using the electrochemistry of the electrospray ion source, 5510--5520, Copyright (1994), with permission from ACS Publications

Dispersion to primary droplets

A potential of usually 2-5kV[17] is applied between the capillary tip and MS interface plate.

The field strength can be approximated by a relation given by Loeb et al.[50]:

$$E_c = \frac{2V_c}{r_c \ln(4d/r_c)}$$

With V_c the applied potential, d the distance between tip and interface plate and r_c the capillary outer radius. As shown by the formula, the most influential geometry parameter is r_c [45]
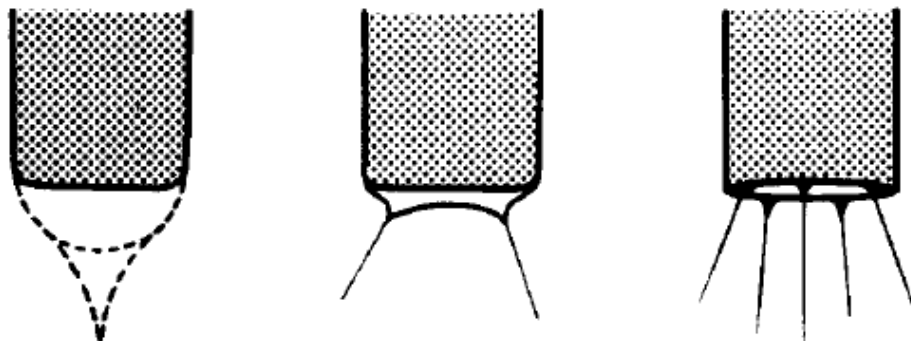
Figure ESITOP 2Fig ESITOP2: Left: Taylor cone formation as result of electrophoretic charge separation. Gray: Electric field direction and isopotential lines. Source: own work. Right: Tylor cone, droplet formation as a result of varicose waves on charged surface leading to pinching of the cone jet. Reprinted from Journal of Mass Spectrometry, 35/7, Kebarle et al. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry, 804-817, Copyright (2000), with permission from John Wiley and Sons.

With common geometries of ESI-ion sources, i.e. d=2-4 cm, r_c~0.5mm and V_c=2-5kV, field strengths in the order of several 10^6 V/m are present. This high field strength results in an electrophoretic charge separation[7, 32, 65, 67], which provides the basis for the deformation of the liquid at the capillary tip into the Taylor cone by electrostatic repulsion[68].

If the potential V_c is sufficiently high, the electrostatic repulsion caused by the charge separation is high enough to overcome the surface tension of the solvent, as predicted by Rayleigh[60]. A jet of charged solvent is subsequently emitted from the Taylor cone. Raising V_c higher changes the cone-jet emission to multiple-mode (s Fig ESITOP 2)

The actual jet breaks down into the aforementioned primary droplets as result of varicose waves induced by interaction of viscosity and surface tension leading to instabilities that "pinch" of droplets[54], as already theoretical postulated by Rayleigh[60]   Further theoretical[63, 58] and experimental investigation[80, 57, 27] revealed initial droplet size in the tens of μm range and that their size can be derived from the jet-cone diameter approximately (ØD/ØJ≈1.95, see Fig ESITOP3) and that initial dispersion of droplet diameter is basically monodispersed[57].

Electrochemistry of electrospray process and consequences

An ESI-ion source is basically and in its most encountered form a two-electrode arrangement (s. Fig ESITOP1) operated in directed current, where the polarity can be chosen by the operator. Fast polarity switching for quasi-synchroneous interlaced acquisition of both polarity species can be performed. Paulie-design based ion traps often offer this option, TOF and QTOF system, however, are not capable of this feature. The reason lies in the high capacitance of the TOF ion optics that would need to be discharged and recharged, a process that usually needs several seconds.
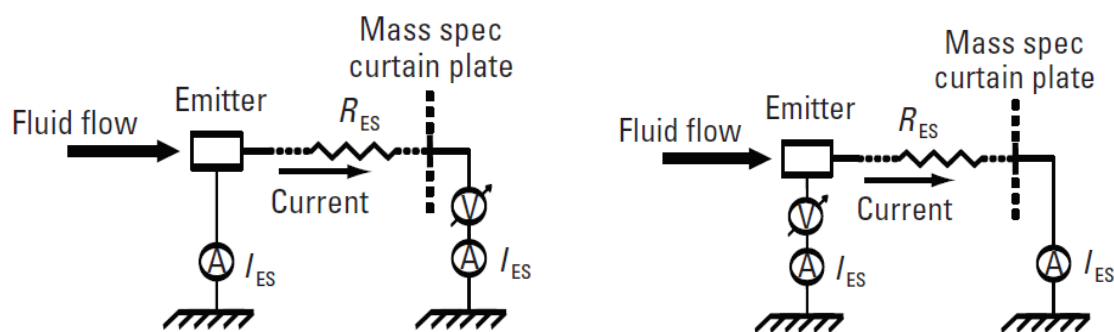
Figure ESITOP 5 Electric flow scheme of an electro spray source. Left: grounded emitter configuration. Right: floated emitter configuration. R_{ES}: resistance of electrospray, I_{ES}: overall current of electrospray. Reprinted from Analytical chemistry, 79, 15, Kertesz et al. Using the electrochemistry of the electrospray ion source, 5510--5520, Copyright (1994), with permission from ACS Publications

Two main configurations are used: grounded and floating emitter. The former keeps the emitter tip at ground potential, whereas in floating configuration electrospray voltage is set at the emitter tip. Both configurations are considered equal regarding ionization, albeit the floating emitter variant is more often encountered, both in literature and in commercial available hardware.

Due to safety considerations, the floating emitter design is commonly used in a modified form, where an upstream ground connection ensures operator and upstream equipment protection from the usual high voltages (kV) used in electrospray ionization (see Fig ESITOP5). The additional ground connection opens a second current loop quantified by the current I_{EXT}, so any calculations based on ion source current, e.g. attempts to provide signal correction for matrix effects, i.e. ion suppression or ion enhancement effects have to take into account that $I_{Total}=I_{EXT} + I_{ES}$.
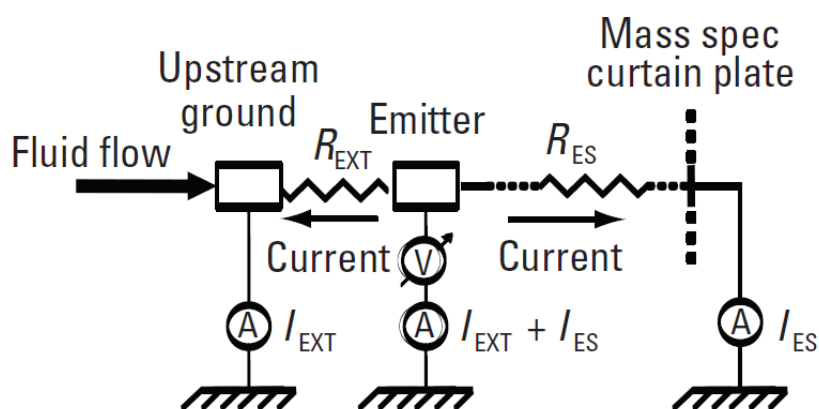
Figure ESITOP 6 Modified floating emitter configuration with upstream ground R_{ES}: resistance of electrospray, I_{ES}:electrospray current I_{EXT}: external ground current . Reprinted from Analytical chemistry, 79, 15, Kertesz et al. Using the electrochemistry of the electrospray ion source, 5510--5520, Copyright (1994), with permission from ACS Publications

As result of high field induced electrophoretic charge separation and ultimately emission of net charged droplets, a net charge transportation i.e. current flow from the emitter capillary tip toward the MS interface place occurs. The opposite charge is enriched in the emitter capillary and charge balance requires a redox-chemical conversion. As Fenn pointed out[54] solely for the generation of atmospheric pressure gas phase ions this conversion process would not be necessary. Albeit it is immanent, due to the nature of the electrospray functioning principle and source construction.

The overall process resembles in fact an electrolysis, as demonstrated by Blades et al[10]. Experimental work by Blades[10] pinpointed this considerations by using a Zinc (Zn) emitter needle, utilizing this metals low reduction potential of E_0= -0.76V. Blades showed qualitatively and quantitatively generation of Zn2+ ions by anodic oxidation in the emitter capillary explained the total current observed in the ESI source. These findings could be reproduced with stainless steel capillaries, finding actual Fe2+ release in congruency with amounts predicted by calculations based on measured total current[46]. Using other emitter materials redox potentials higher than that of H2O reduced the amount of metal cations, and as Blades assumed in favor of H^+ generation from water.

Table ESITOP 1 Major electrochemical reactions occurring in an typical ESI-source is summarized by Kertesz and Vilmos. SHE: Standard hydrogen electrode. Reprinted from Analytical chemistry, 79, 15, Kertesz et al. Using the electrochemistry of the electrospray ion source, 5510--5520, Copyright (1994), with permission from ACS Publications

Oxidation (positive ion mode)	$E^0(V)$ vs. SHE	Reduction (negative-ion mode)	$E^0(V)$ vs. SHE
$4OH^- \rightarrow 2H_2O + O_2 + 4e^-$	0.4	$O_2 + 4H+ \; 4e^- \rightarrow 2H_2O$	1.23
$2OH^- \rightarrow H_2O_2 + 2e^-$	0.88	$O_2 + 2H+ + 2e^- \rightarrow H_2O_2$	0.7
$2H_2O \rightarrow O_2 + 4H + 4e^-$	1.23	$2H_2O + O_2 + 4e^- \rightarrow 4OH^-$	0.4
$2H_2O \rightarrow H_2O_2 + 2H+ + 2e^-$	1.77	$2H_2O + 2e- \rightarrow H_2 + 2OH-$	0.07
$OH^- \rightarrow OH* + e^-$	1.89	$2H+ \; 2e^- \rightarrow H_2$	0
		$O_2+ 2H_2O + 2e^- \rightarrow H_2O_2 + 2OH^-$	-0.13
		$O_2 + e^- \rightarrow O2^-$	-0.33
Emitter electrode reactions			
$Fe + 2OH^- \rightarrow Fe(OH)_2 + 2e^-$	-0.87		

Fe → Fe²⁺ + 2e⁻	-0.44	**Emitter electrode reactions**	
Fe → Fe³⁺ + 3e⁻	-0.03	Pt(OH)₂ + 2e⁻ → Pt + 2OH⁻	0.16
Pt + 2OH⁻ → Pt(OH)₂ + 2e⁻	0.16	Fe(OH)₃ + e⁻ → Fe(OH)₂ + OH⁻	-0.56
Pt + 2H₂O → Pt(OH)₂ + 2H+2e⁻	0.98	Fe(OH)₂ + 2e⁻ → Fe + 2OH⁻	-0.87

Although reactive species like H_2O_2 are produced and direct electrochemical reactivity of analytes under ESI conditions have been reported[71, 76, 72, 61, 29, 28], generally the alteration of analytes and matrix compounds is thought as of minor concern[54]. However, regarding the field of untargeted and comprehensive analysis (metabolomics, lipidomics, etc.) it might be necessary to reevaluate this résumé, as electrochemical processes might introduce unwanted artificial bias or inconsistency of results in conjunction with the use of different vendors systems on equal samples. Furthermore, alteration of ionization efficiency as consequence inevitable emitter erosion caused by the electrolysis process introduces such bias likewise and the sprayer assembly must be inspected regularly.

Ion generation - primary droplet fission

The stability criterion for a charged droplet was first derived by Rayleigh[60] as:

$$q^2 = 64\pi^2\epsilon_0\gamma a^3$$

Where q is the maximum charge up to which the droplet is stable, a the radius of the droplet, γ the surface tension of the liquid and ϵ_0 the electrical permittivity in vacuo. The Rayleigh-model describes primary droplet fission as result of the dispersive force exerted by the charged species overcoming the surface tension of the droplet liquid. With charge density as driving moment, the process is hence also referred to as Coulomb fission or explosion. Starting in the 1960s, studies[33, 25, 1, 6, 64] using Millikan-type condensers or quadrupole-droplet-traps, attempted to confirm the Rayleigh formula. These studies reported varying methodical accuracy (4-15%), yet concurred to a loss of droplet mass in the low percentage (1-5%) and to a loss of roughly 25% of charge. Hence, the fission of droplets does not occur as symmetric process but rather as asymmetric ejection of daughter drops. Furthermore, fission events below the Rayleigh-limit were reported. Gomez et al. reported

images of the fission process clearly depicting asymmetric multiple emission (fig ESITOP6). High precision determining mass (1.0-2.3%) and charge loss (10-18%) was provided by Taflin et. al.[66] , on the basis of optical resonance measurements, confirming also occurrence of pre-Rayleigh-limit fissions.
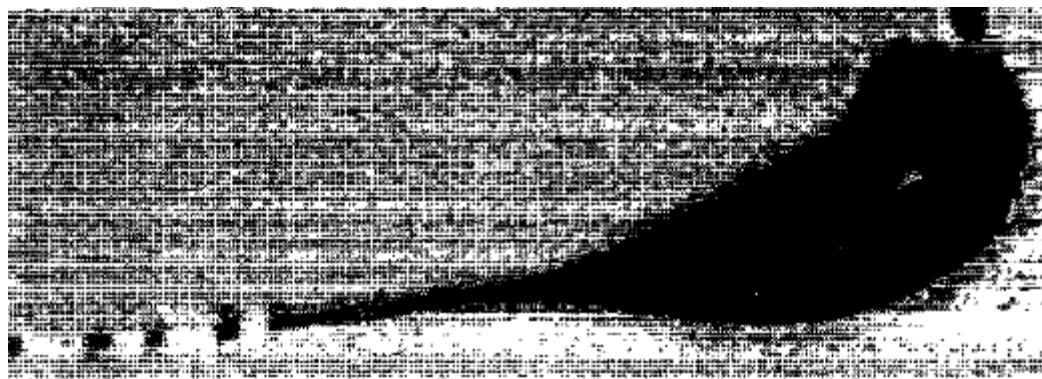
Figure ESITOP 7 Flash shadowgraph of a primary solvent droplet at the moment of coulomb fission. Visible is a parent drop and multiple offspring drops, each considerably smaller. Reprinted from Physics of Fluids, 6 Gomez et al. Charge and fission of droplets in electrostatic sprays, 404--414, Copyright (1994), with permission from AIP Publishing

Ion generation - Charged residue model (CRM)

The most straightforward hypothesis of gas phase ion formation in light of the recursive evaporation-fission cycles is the charged residue model proposed by Dole et al. [24]: Primary droplets emitted from the Taylor cone are ultimately reduced to droplets containing one analyte molecule with an additional monolayer of residual solvent and charge carriers (H_3O^+, NH_4^+ etc.)[21, 24], from which after final evaporation of the solvent the ion is transferred to the gas phase[45, 40]. Due to electrostatic considerations, it is assumed that charge transfer to the analyte occurs during the final evaporation step[45, 40]. Support for the CRM Model was provided especially for macromolecular species, by molecular dynamic calculations (MD) and charge series observation. MD Simulations[3, 59] supported solvation and centered embedding of large molecular species inside the ESI droplets, rendering direct emission (IEM, *vide infra*) of such large species kinetically inviable[47]. Experimentally, support of the CRM accrued from observation of maximum charge of proteins and macromolecules which is described by $[M+z_rH]^{zr+}$, where z_r is the Rayleigh charge of a water droplet resembling the globular diameter of the protein [30, 17].

Ion generation - Ion evaporation model (IEM)

The ion evaporation model (IEM) as introduced by Iribarne and Thomson[42], proposes a solution to gas phase transition prior to complete evaporation of the solvent. Instead of reaching the Rayleigh limit with subsequent Coulomb fission/explosion, charge load of a droplet shrinking by evaporation is reduced beforehand by emission of ions from the surface, giving rise to an independent gas phase transition mechanism. In contrast to the CRM, ion evaporation is energetically only feasible for ions of small size, for example NH_4^+, metabolites or drugs[2].   As MD simulations suggest, ejection of ions occurs not as an unsolvated ion, but by emission of a small nanodrop comprised of a charged small molecular ion solvated by few layers or a mono layer of solvent with subsequent evaporation of the latter.[20, 2]

Figure ESITOP 8 Left: MD simulation of emission of nanodroplet containing one NH4+ ion during ion evaporation process. Right: Distance and Energy at emission event. Reproduced with permission from: Ejection of Solvated Ions from Electrosprayed Methanol/Water Nanodroplets Studied by Molecular Dynamics Simulations, Journal of the American Chemical Society 2011 133 (24), 9354-9363, Copyright (2011), American Chemical Society
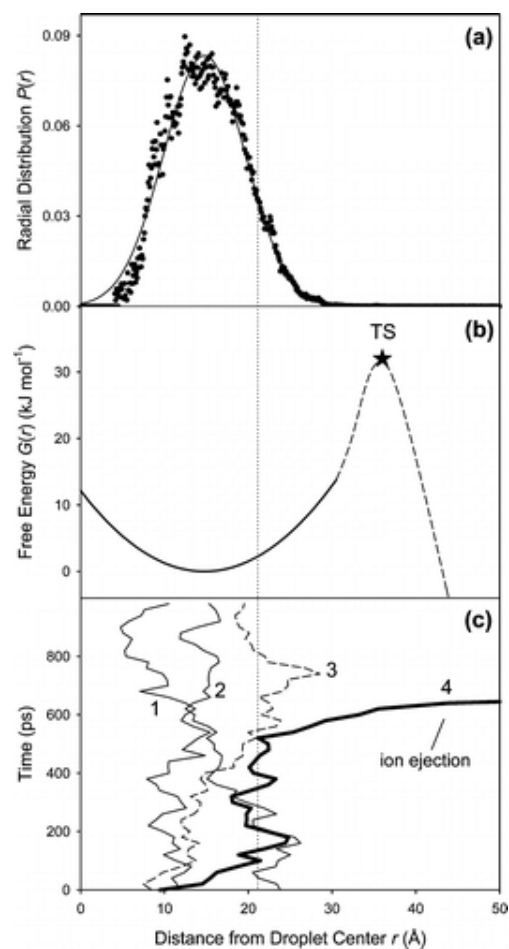
Ion generation - Chain ejection model (CEM)

As mentioned in the previous section, intact protein ions are transferred to the gas phase via the CRM mechanism. Yet, this holds true for proteins in solvents allowing for native conformation, where hydrophobic residues are buried in the core of the tertiary structure[47] . If denaturing conditions are present, MD simulations[3, 48] suggest that due to the exposure of hydrophobic residues, proteins are not embedded in the core of the droplet anymore[3, 59] and are expelled from the droplet surface process somewhat resembling a repetitive ion evaporation mechanism along the protein chain.
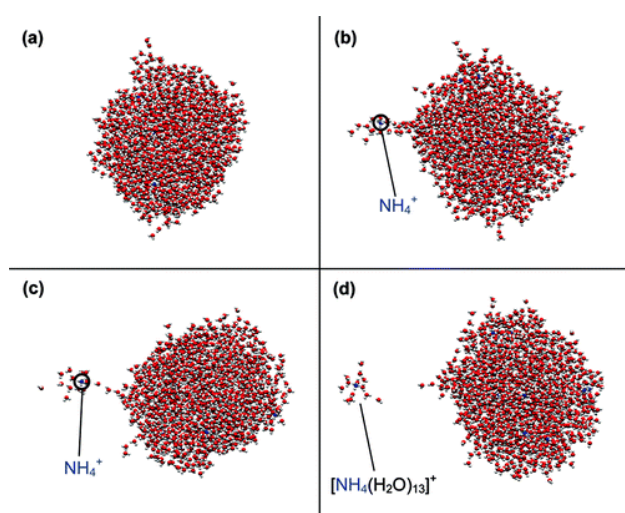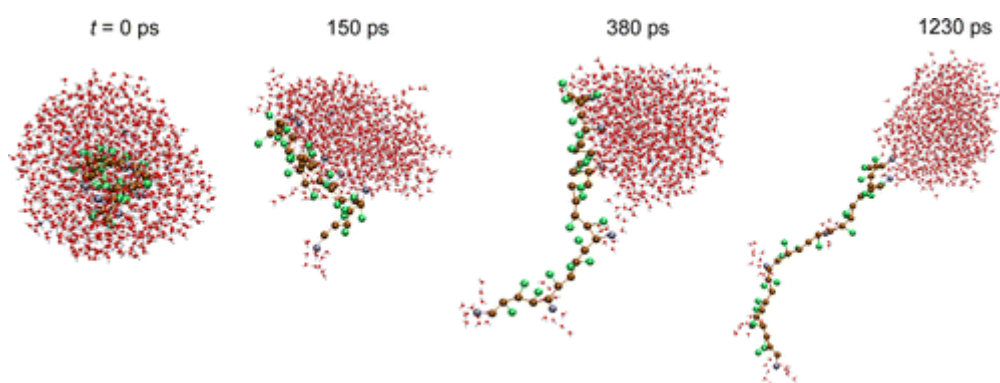

Figure ESITOP 9 MD simulation of emission of nanodroplet containing one NH4+ ion during ion evaporation process Reproduced with permission from: Ejection of Solvated Ions from Electrosprayed Methanol/Water Nanodroplets Studied by Molecular Dynamics Simulations, Journal of the American Chemical Society 2011 133 (24), 9354-9363, Copyright (2011), American Chemical Society

Practical implications of the ionization models

For **The first part** of this thesis (**Publication II and VIII)**, from the mass spectrometric perspective, dealing with LCMS analysis of small molecules and ion suppression, intricacies of the IEM are relevant. IEM efficiency and therefore sensitivity is influenced by eluent composition e.g. organic content, pH, salt concentration, viscosity and source conditions as temperature, gas flows and voltages[17, 2] but also due to its nature as surface emission process by competition of analytes for droplet surface and charge[35]. Droplet surface access is not only governed by analyte hydrophobicity alone, yet is also result of a partitioning process dependent on organic content, as nanodroplets of eluents typical for liquid chromatography, i.e. water/organic mixtures, undergo a demixing process under ESI conditions[2]. The essence for the practitioner is, that ideally conditions are to be

established, where the analyte of interest is the only species entering the ESI process under eluent composition and ion source settings finetuned for maximum sensitivity. Actually, the resultant LCMS method is usually a tradeoff determined by chromatographic requirements and constraints, sample complexity and origin(matrix), eluent quality and analyte of interest properties. Distinct ion source settings can be applied on the chromatographic timescale (e.g. scheduled LCMS), if multiple analytes are to be assessed in a single run, with the unfortunate exception of source temperature, that usually cannot be stabilized on the chromatographic timescale.

For **protein mass spectrometry** as in the **third part** (Publications IV IX X and XI) of this thesis, subtleties of CRM and CEM are relevant due to their characteristics in intact protein ion production, where CRM tends to produce lower and CEM higher charged states[47]. As aforementioned, whether CRM or CEM applies, is dependent on protein characteristics and the conditions in solution. i.e. the eluent composition if liquid chromatography is coupled to the mass spectrometer. Thus, chromatographic conditions at time of elution influence charge state, charge distribution and therefore intensity distribution, complicating harmonization of chromatographic and mass spectrometric requirements by adding protein parameters, e.g. pI, hydrophobic vs hydrophilic residue content, sequence, tertiary structure, post translational modifications and overall solubility amongst others to the equation. Furthermore, observed high charge states of a protein does not exclusively reflect the titration state in solution in a simple manner[47, 17, 44] but are a result of charge equilibration during the CEM process[48]. The distribution of charges determining the charge series encountered in a protein mass spectrum is also a consequence of initial droplet size distribution elicit by ESI sprayer dimension, source gas temperature and, voltage settings, eluent composition and flow[45].
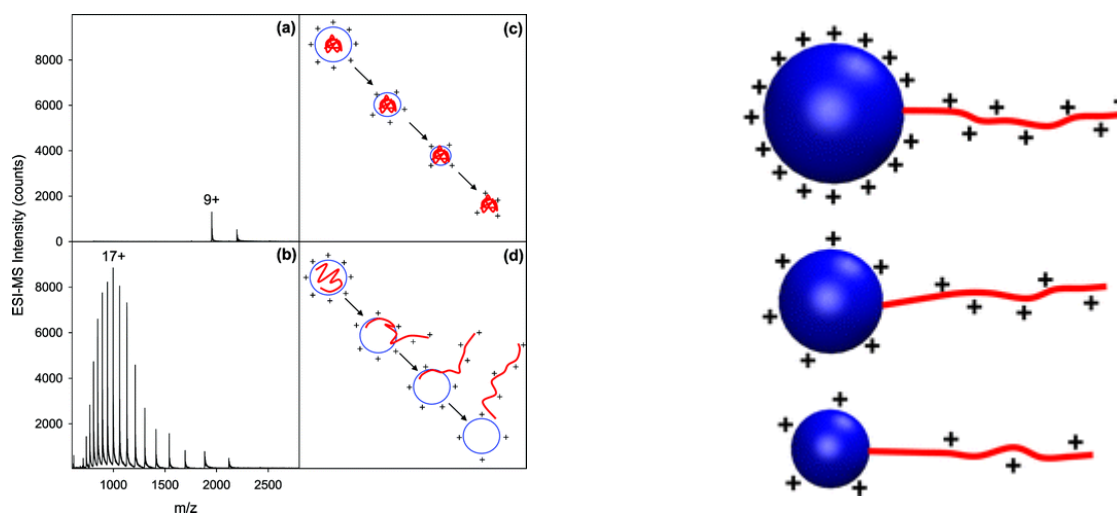
Figure ESITOP 10Fig ESITOP9: Left: Influence of ionization mechanism governed by protein properties and solvent composition on charge distribution and intensity. a and c: CRM, b and d: CEM. Reprinted from "Modeling the behavior of coarse-grained polymer chains in charged water droplets: implications for the mechanism of electrospray ionization." The Journal of Physical Chemistry B 116.1 (2011): 104-112. Copyright (2011), with permission from American Chemical Society. Right: Charge distribution a consequence of distribution of primary droplet size, Reprint from: Unraveling the Mechanism of Electrospray Ionization, Analytical Chemistry 2013 85 (1), 2-9, Copyright (2013), with permission from American Chemical Society

Furthermore, not only primary drop size distribution and thereby charge distribution are of importance, yet also average initial droplet size has a major effect on sensitivity, as smaller primary droplets facilitate more efficient ionization[31, 47, 74]. Smaller droplet size can be achieved by reducing the eluent flow rate (see Fig ESITOP10). [75, 22]. Albeit in order to maintain chromatographic performance, post column dispersion volumes. i.e. dwell volumes, valve inner diameters, tubing and ESI sprayer needle inner diameter must be adapted, i.e. reduced. Unfortunately, sensitivity increase is described by an inverse exponential law[62], requiring very low flow rates on the chromatographic flow scale and hence small inner diameters in order to harvest maximum signal increase (see Fig ESITOP10). In practice, robustness of flow path of especially the ESI sprayer needle behaves inversely, resulting in short maintenance intervals, spontaneous ESI spray breakdown and low operator walk away capacity. Additionally, analyte quantitation or analyte detection at all

when performing qualitatively analysis at limit of detection (LOD) suffers from low reproducibility during e.g. ESI needle exchange[62].

In a tradeoff between sensitivity requirements, instrumentation robustness and results reproducibility the fourth part of this thesis implemented a microflow-ESI(µESI) methodology (50µl/min, 50µm ESI needle ID).
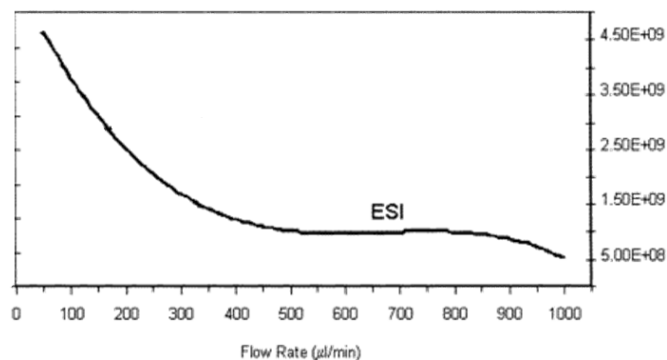
Figure ESITOP 11Fig ESITOP10: Signal gain via reduction primary droplets as consequence of reduced flow rate. Adapted reprint from: Combined Electrospray Ionization−Atmospheric Pressure Chemical Ionization Source for Use in High-Throughput LC−MS Applications Analytical Chemistry 2003 75 (4), 973-977 Copyright (2003), with permission from American Chemical Society.

For the **fourth part** dealing with native electrospray mass spectrometry of noncovalent quadruplex DNA-ligand complexes, basically all ionization models or rather their prevention or finetuning by appropriate methodology are relevant. As Hogan et al. demonstrated[35], IEM is occurring also parallel to CRM ionization essential for native ESI mass spectrometry of macromolecules, implying the risk of ligand depletion during ionization of the complex, whereas occurrence of CEM processes would represent disruption of all but the primary structure of the quadruplex DNA. The challenge with regard to ionization for Publications III and XII was to synergize solute and ion source conditions in a manner that provided both stability of analyte and sensitivity of the method.

Abbreviations

IEM: Ion emission model

CRM: charged residue model

ESI: electrospray ionization

LOD: limit of detection

References

[1]   MA Abbas and J Latham. The instability of evaporating charged drops. *Journal of Fluid Mechanics*, 30(04):663–670, 1967.

[2]   Elias Ahadi and Lars Konermann. Ejection of solvated ions from electrosprayed methanol/water nanodroplets studied by molecular dynamics simulations. *Journal of the American Chemical Society*, 133(24):9354–9363, 2011.

[3]   Elias Ahadi and Lars Konermann. Modeling the behavior of coarse-grained polymer chains in charged water droplets: implications for the mechanism of electrospray ionization. *The Journal of Physical Chemistry B*, 116(1):104–112, 2011.

[4]   Patrick Arpino, MA Baldwin, and FW McLafferty. Liquid chromatography-mass spectrometry. iiâ€"continuous monitoring. *Biological Mass Spectrometry*, 1(1):80–82, 1974.

[5]   Patrick J Arpino and Georges Guiochon. Lc/ms coupling. *Analytical chemistry*, 51(7):682–701, 1979.

[6]   Selcuk Ataman and DN Hanson. Measurement of charged crops. *Industrial & Engineering Chemistry Fundamentals*, 8(4):833–836, 1969.

[7]   Adrian G Bailey. *Electrostatic spraying of liquids*. Wiley New York etc, 1988.

[8]   AG Bailey. The theory and practice of electrostatic spraying. *Atomisation Spray Technology*, 2:95–134, 1986.

[9]   MA Baldwin and FW McLafferty. Liquid chromatography-mass spectrometry interface–i: The direct introduction of liquid solutions into a chemical ionization mass spectrometer. *Organic Mass Spectrometry*, 7(9):1111–1112, 1973.

[10]   Arthur T Blades, Michael G Ikonomou, and Paul Kebarle. Mechanism of electrospray mass spectrometry. electrospray as an electrolysis cell. *Analytical Chemistry*, 63(19):2109–2114, 1991.

[11]   CR Blakley, JJ Carmody, and ML Vestal. Liquid chromatograph-mass spectrometer for analysis of nonvolatile samples. *Analytical Chemistry*, 52(11):1636–1641, 1980.

[12]   CR Blakley, JJ Carmody, and ML Vestal. A new soft ionization technique for mass spectrometry of complex molecules. *Journal of the American Chemical Society*, 102(18):5931–5933, 1980.

[13]   CR Blakley, MJ McAdams, and ML Vestal. Crossed-beam liquid chromatoraph—mass spectrometer combination. *Journal of Chromatography A*, 158:261–276, 1978.

[14]   CR Blakley and ML Vestal. Thermospray interface for liquid chromatography/mass spectrometry. *Analytical Chemistry*, 55(4):750–754, 1983.

[15]   Andries P Bruins, Thomas R Covey, and Jack D Henion. Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry. *Analytical Chemistry*, 59(22):2642–2646, 1987.

[16]   DI Carroll, I Dzidic, RN Stillwell, KD Haegele, and EC Horning. Atmospheric pressure ionization mass spectrometry. corona discharge ion source for use in a liquid chromatograph-mass spectrometer-computer analytical system. *Analytical Chemistry*, 47(14):2369–2373, 1975.

[17]   Nadja B Cech and Christie G Enke. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrometry Reviews*, 20(6):362–387, 2001.

[18]   GA Clegg and M Dole. Molecular beams of macroions. iii. zein and polyvinylpyrrolidone. *Biopolymers*, 10(5):821–826, 1971.

[19]   Thomas R Covey, Ronald F Bonner, Bori I Shushan, Jack Henion, and RK Boyd. The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2(11):249–256, 1988.

[20]   Christopher D Daub and Natalie M Cann. How are completely desolvated ions produced in electrospray ionization: insights from molecular dynamics simulations. *Analytical chemistry*, 83(22):8372–8376, 2011.

[21]   J Fernandez De La Mora. Electrospray ionization of large multiply charged species proceeds via dole€™s charged residue mechanism. *Analytica chimica acta*, 406(1):93–104, 2000.

[22]   J Fernández De La Mora and Ignacio González Loscertales. The current emitted by highly conducting taylor cones. *Journal of Fluid Mechanics*, 260:155–184, 1994.

[23]   Malcolm Dole, RL Hines, LL Mack, RC Mobley, LD Ferguson, and MB Alice. Gas phase macroions. *Macromolecules*, 1(1):96–97, 1968.

[24]   Malcolm Dole, LL Mack, RL Hines, RC Mobley, LD Ferguson, and MB d Alice. Molecular beams of macroions. *The Journal of Chemical Physics*, 49(5):2240–2249, 1968.

[25]   Arnold Doyle, D Read Moffett, and Bernard Vonnegut. Behavior of evaporating electrically charged droplets. *Journal of Colloid Science*, 19(2):136–143, 1964.

[26]   Lee E.D. and Henion J.D., editors. *Proceedings of the 36th ASMS Conference on Mass Spectrometry and Allied Topics*, 1988.

[27]   Alessandro Gomez and Keqi Tang. Charge and fission of droplets in electrostatic sprays. *Physics of Fluids (1994-present)*, 6(1):404–414, 1994.

[28]   Thais Guaratini, Ricardo Vessecchi, Ernani Pinto, Pio Colepicolo, and Norberto P Lopes. Balance of xanthophylls molecular and protonated molecular ions in electrospray ionization. *Journal of mass spectrometry*, 40(7):963–968, 2005.

[29]   Thais Guaratini, Ricardo L Vessecchi, Francisco C Lavarda, Patrcia MBG Maia Campos, Zeki Naal, Paul J Gates, and Norberto P Lopes. New chemical evidence for the ability to generate radical molecular ions of polyenes from esi and hr-maldi mass spectrometry. *Ánalyst*, 129(12):1223–1226, 2004.

[30]   Albert JR Heck and Robert HH van den Heuvel. Investigation of intact protein complexes by mass spectrometry. *Mass spectrometry reviews*, 23(5):368–389, 2004.

[31]   Anthonius AM Heemskerk, Jean-Marc Busnel, Bart Schoenmaker, Rico JE Derks, Oleg Klychnikov, Paul J Hensbergen, Andreì• M Deelder, and Oleg A Mayboroda. Ultra-low flow electrospray ionization-mass spectrometry for improved ionization efficiency in phosphoproteomics. *Analytical chemistry*, 84(10):4552–4559, 2012.

[32]   CD HENDRICKS, JR and RJ Pfeifer. Parametric studies of electrohydrodynamic spraying. *Aiaa Journal*, 6(3):496–502, 1968.

[33]   Charles D Hendricks. Charged droplet experiments. *Journal of Colloid Science*, 17(3):249–259, 1962.

[34]   RL Hines. Electrostatic atomization and spray painting. *Journal of Applied Physics*, 37(7):2730–2736, 1966.

[35]   Christopher J Hogan Jr, James A Carroll, Henry W Rohrs, Pratim Biswas, and Michael L Gross. Combined charged residue-field emission model of macromolecular electrospray ionization. *Analytical chemistry*, 81(1):369–377, 2008.

[36]   EC Horning, DI Carroll, I Dzidic, KD Haegele, MG Horning, and RN Stillwell. Atmospheric pressure ionization (api) mass spectrometry. solvent-mediated ionization of samples introduced in solution and in a liquid chromatograph effluent stream. *Journal of chromatographic science*, 12(11):725–729, 1974.

[37]   EC Horning, DI Carroll, I Dzidic, S-N Lin, RN Stillwell, and J-P Thenot. Atmospheric pressure ionization mass spectrometry: Studies of negative ion formation for detection and quantification purposes. *Journal of Chromatography A*, 142:481–495, 1977.

[38]   EC Horning, MG Horning, DI Carroll, I Dzidic, and RN Stillwell. New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure. *Analytical Chemistry*, 45(6):936–943, 1973.

[39]   EC Horning, MG Horning, DI Carroll, RN Stillwell, and I Dzidic. Nicotine in smokers, non-smokers and room air. *Life sciences*, 13(10):1331–1346, 1973.

[40]   Anthony T Iavarone and Evan R Williams. Mechanism of charging and supercharging molecules in electrospray ionization. *Journal of the American Chemical Society*, 125(8):2319–2327, 2003.

[41]   Mark G Inghram and Robert Gomer. Mass spectrometric analysis of ions from the field microscope. *The Journal of Chemical Physics*, 22(7):1279–1280, 1954.

[42]   JV Iribarne and BA Thomson. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64(6):2287–2294, 1976.

[43]   Patrick R Jones and Shen K Yang. Liquid chromatograph/mass spectrometer interface. *Analytical Chemistry*, 47(7):1000–1003, 1975.

[44]   Igor A Kaltashov and Rinat R Abzalimov. Do ionic charges in esi ms provide useful information on macromolecular structure? *Journal of the American Society for Mass Spectrometry*, 19(9):1239–1246, 2008.

[45]   Paul Kebarle and Udo H Verkerk. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass spectrometry reviews*, 28(6):898–917, 2009.

[46]   Gary J Van Berkel Kertesz and Vilmos. Using the electrochemistry of the electrospray ion source. *Analytical chemistry*, 79(15):5510–5520, 2007.

[47]   Lars Konermann, Elias Ahadi, Antony D Rodriguez, and Siavash Vahidi. Unraveling the mechanism of electrospray ionization. *Analytical chemistry*, 85(1):2–9, 2012.

[48]   Lars Konermann, Antony D Rodriguez, and Jiangjiang Liu. On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. *Analytical chemistry*, 84(15):6798–6804, 2012.

[49]   P.M. LLewllyn. Staged gas inlet for gas analyzers and gas analyzing system for employing same, 1969.

[50]   LB Loeb, AF Kip, GG Hudson, and WH Bennett. Pulses in negative point-to-plane corona. *Physical Review*, 60(10):714, 1941.

[51]   Lawrence Lloyd Mack, P Kralik, A Rheude, and M Dole. Molecular beams of macroions. ii. *The Journal of Chemical Physics*, 52(10):4977–4986, 1970.

[52]   WH McFadden, HL Schwartz, and S Evans. Direct analysis of liquid chromatographic effluents. *Journal of Chromatography A*, 122:389–396, 1976.

[53]   FW McLafferty, Ruedi Knutti, Rengachari Venkataraghavan, PJ Arpino, and BG Dawkins. Continuous mass spectrometric monitoring of a liquid chromatograph with subnanogram sensitivity using an on-line computer. *Analytical Chemistry*, 47(9):1503–1505, 1975.

[54]   Juan Fernandez de la Mora, Gary J Van Berkel, Christie G Enke, Richard B Cole, Manuel Martinez-Sanchez, and John B Fenn. Electrochemical processes in electrospray ionization mass spectrometry. *Journal of Mass Spectrometry*, 35(8):939–952, 2000.

[55]   Erwin W Müller. Das feldionenmikroskop. *Zeitschrift für Physik*, 131(1):136–142, 1951.

[56]   Erwin W Müller. Field ionization and field ion microscopy. *Advances in Electronics and Electron Physics*, 13:83–179, 1960.

[57]   Masayuki Mutoh, Shozo Kaieda, and Kuniaki Kamimura. Convergence and disintegration of liquid jets induced by an electrostatic field. *Journal of Applied Physics*, 50(5):3174–3179, 1979.

[58]   A Neukermans. Stability criteria of an electrified liquid jet. *Journal of Applied Physics*, 44(10):4769–4770, 1973.

[59]   Alexandra Patriksson, Erik Marklund, and David van der Spoel. Protein structures under electrospray conditions. *Biochemistry*, 46(4):933–945, 2007.

[60]   Lord Rayleigh. On the equilibrium of liquid conducting masses charged with electricity', lond. *Edinb. Dublin Philos. Mag. J. Sci*, 14:87, 1882.

[61]   David Rondeau, David Kreher, Michel Cariou, Piétrick Hudhomme, Alain Gorgues, and Pascal Richomme. Electrolytic electrospray ionization mass spectrometry of c60-ttf-c60 derivatives: high-resolution mass measurement and molecular ion gas-phase reactivity. *Rapid Communications in Mass Spectrometry*, 15(18):1708–1712, 2001.

[62]   Andrea Schmidt, Michael Karas, and Thomas Dülcks. Effect of different solution flow rates on analyte ion signals in nano-esi ms, or: when does esi turn into nano-esi? *Journal of the American Society for Mass Spectrometry*, 14(5):492–500, 2003.

[63]   JM Schneider, NR Lindblad, CD Hendricks Jr, and JM Crowley. Stability of an electrified liquid jet. *Journal of Applied Physics*, 38(6):2599–2605, 1967.

[64]   John W Schweizer and DN Hanson. Stability limit of charged drops. *Journal of Colloid and Interface Science*, 35(3):417–423, 1971.

[65]   David PH Smith. The electrohydrodynamic atomization of liquids. *IEEE Transactions on Industry Applications*, (3):527–535, 1986.

[66]   Daniel C Taflin, Timothy L Ward, and E James Davis. Electrified droplet fission and the rayleigh limit. *Langmuir*, 5(2):376–384, 1989.

[67]   Liang Tang and Paul Kebarle. Effect of the conductivity of the electrosprayed solution on the electrospray current. factors determining analyte sensitivity in electrospray mass spectrometry. *Analytical Chemistry*, 63(23):2709–2715, 1991.

[68]   Geoffrey Taylor. Disintegration of water drops in an electric field. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 280, pages 383–397. The Royal Society, 1964.

[69]   BA Thomson and JV Iribarne. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *The Journal of Chemical Physics*, 71(11):4451–4463, 1979.

[70]   R Tilney. Electrostatic coating processes. *British Journal of Applied Physics*, 4(S2):S51, 1953.

[71]   Gary J Van Berkel, Scott A McLuckey, and Gary L Glish. Electrospray ionization of porphyrins using a quadrupole ion trap for mass analysis. *Analytical Chemistry*, 63(11):1098–1109, 1991.

[72]   Gary J Van Berkel, J Martin E Quirke, Rudolfo A Tigani, Anja S Dilley, and Thomas R Covey. Derivatization for electrospray ionization mass spectrometry. 3. electrochemically ionizable derivatives. *Analytical chemistry*, 70(8):1544–1554, 1998.

[73]   CRAIG M Whitehouse, RN Dreyer, M Yamashita, and JB Fenn. Electrospray ionization for mass-spectrometry of large biomolecules. *Science*, 246(4926):64–71, 1989.

[74]   Matthias Wilm and Matthias Mann. Analytical properties of the nanoelectrospray ion source. *Analytical chemistry*, 68(1):1–8, 1996.

[75]   Matthias S Wilm and Matthias Mann. Electrospray and taylor-cone theory, dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes*, 136(2-3):167–180, 1994.

[76]   Xiaoming Xu, Steven P Nolan, and Richard B Cole. Electrochemical oxidation and nucleophilic addition reactions of metallocenes in electrospray mass spectrometry. *Analytical Chemistry*, 66(1):119–125, 1994.

[77]   Masamichi Yamashita and John B Fenn. Electrospray ion source. another variation on the free-jet theme. *The Journal of Physical Chemistry*, 88(20):4451–4459, 1984.

[78]   John Zeleny. The electrical discharge from liquid points, and a hydrostatic method of measuring the electric intensity at their surfaces. *Physical Review*, 3(2):69, 1914.

[79]   John Zeleny. On the conditions of instability of electrified drops, with applications to the electrical discharge from liquid points. In *Proc. Camb. Phil. Soc*, volume 18, pages 71–83, 1915.

[80]   John Zeleny. Instability of electrified liquid surfaces. *Physical review*, 10(1):1, 1917.

Part I - Publications I, II, VII and VIII

Introduction to Publication II and VIII
Structural analysis employing chromatography and mass spectrometry

Background of Publication II and VIII and introduction to the analyte of interest

Both publications are the result of a cooperation between the working groups of Prof. Gross and Prof. Lämmerhofer, both affiliated at the Institute of Pharmaceutical Sciences of University of Tübingen. Publications had its origin in the Gross group's interest in new antimicrobial cyclic lipopeptides (CLPs). Generally, CLPs are promising as new type of antibiotics due to the rising occurrence of multi resistant bacteria [15, 19, 27] and due to an observed low resistance induction [26]. Also, CLPs and their linear analogs are examined as potential biocontrol agents in agriculture, as they not only exert antagonistic activity against bacteria but against fungi as well [8, 21, 24, 38]. A recent review of discovery of antimicrobial peptides, since 2000 has been provided by Xue et al. [37].

Different targets for antimicrobial activity were identified, inhibition of cell wall biosynthesis [10, 22, 31], pore formation in [34] or depolarization of cell membrane [20], or by degeneration of the specificity of members of the intracellular *Clp* protease family , inducing uncontrolled proteolysis [16, 17]. Modes of action of CLPs were reviewed by Schneider et al. [27].

CLPs are encountered as secondary metabolites in bacteria of the genera *Actinomyces*, *Streptomyces*, *Bacillus*, and in *Pseudomonas* and are synthesized by nonribosomal secondary metabolite pathways [27]. Nonribosomal biosynthesis allows for one distinctive feature of CLPs, the incorporation of non-proteinogenic and modified amino acids, mainly D-amino acids and other chemical classes like carboxy acids, especially fatty acids [25]. This property of CLPs, combined with the intramolecular cyclization, facilitates enhanced general and proteolytic stability, circumventing extracellular peptidases secreted by the addressee organisms for metabolite acquisition and protection [5, 9, 15, 27].

As one representative of this class of bacterial secondary metabolites, the lipopeptide Poaeamide originating from *Pseudomonas poae* [2] was first described by Zachow et al [40] in the strain *Pseudomonas poae* RE*1-1-14. By comparison of deletion mutants to wildtype, Poaeamide was identified as antagonist towards the soilborne pathogen *Rhizoctonia solani,*

a pathogenic fungus affecting a variety of agricultural plants [40]. As with other CLPs, the mode of action and structure vs activity relationships are of interest, at which foundations lies the determination of the exact molecular structure of Poaeamide.

Structural analysis of cyclic lipopeptides

To completely characterize a CLP, its amino acid composition, sequence (primary structure), position of ring closure (secondary structure), site of carboxy acid modification and tentative other amino acid modifications have to be elucidated. Furthermore, as lipopeptides usually contain non-proteinogenic amino acids, chiral analysis of the individual amino acids and determination of location of D-amino acids are required. As some amino acids frequently found in CLPs represent constitutional isomers, e.g. allo-Threonine or allo-Isoleucine, a complete structural determination would comprise information about locations of allo-isoforms.

Classically, structural determination of oligopeptides in the size range of CLPs could be performed by crystallography or nuclear magnetic resonance (NMR) spectroscopy or combination of both methods, respectively, with the prerequisite that enough adequately purified compound for both techniques can be accumulated [12, 23]. In absence of crystallographic data, albeit, as NMR is an achiral technique, complete investigation down to the configuration level of amino acids needs to be assisted by complementary methods like chiral gas chromatography coupled to mass spectrometry (GC-MS). Such an approach was followed by Zachow et al. [40]. Yet, enantioselective GC-MS of derivatized amino acids can only confirm absence or presence and quantity of respective enantiomers of amino acids, as sequence information is lost during sample preparation, i.e. total peptide hydrolysis and derivatization. The sequence position of D-amino acids is derived indirectly, by deduction of D-amino acid positions from fragments generated by enzymatic peptide cleavage sites, as peptide bonds with D-amino acids are not cleaved by usual proteomic enzymes [40].

Publication II and VIII describe enhanced procedures for determination of absolute configurations by coaction of enantioselective GC-MS, enantio-selective enzymatic and

chemical digestion, HR-QTOF-MS assisted micropreparative reversed phase chromatography. Also feasibility and orthogonality of ZWIX(+/-) enantionselective phases was demonstrated.

General strategy of analysis employed in Publication II and VIII
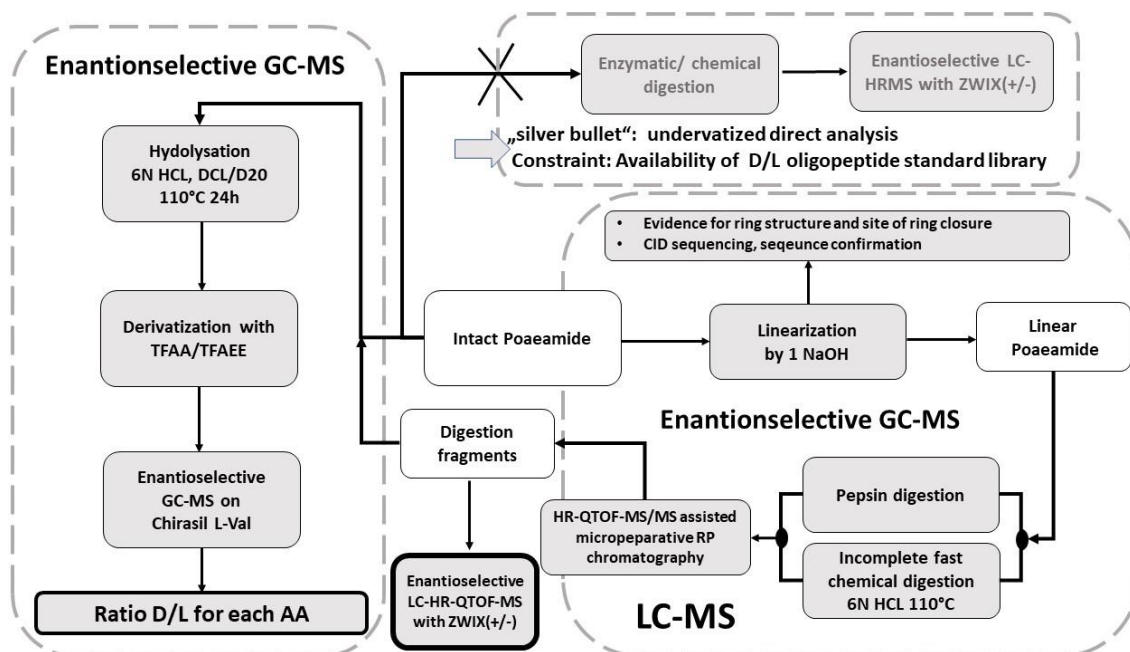
Figure FP2 1 Analytical strategy of Publication II and VIII. Note that CID sequencing as mentioned under the "silver bullet" analysis pathway was not performed, as sequence of Poaeamide was already determined by Zachow et al. [40] Source: own work.

Necessity of digestion for analysis of absolute configuration

A direct method for analyzing oligopeptides ($n_{stereocenters} \geq n_{monomers}$) regarding configuration of each stereocenter is not known besides crystallography. Albeit it can be hypothesized, that chromatographic stationary phase material and method are unlikely to be developed for this task, as expounded for cinchona alkaloid based enantioselective stationary phases, such as the ZWIX(+/-) phases employed in Publication II.

Mechanistically, enantioselective separation can be described by differences in binding affinity of both enantiomers (selectands) towards the stationary phase ligand (selector) that manifest themselves in different overall retention time on the chromatographic column, see

also introduction to Publication I for reference. These discriminative affinities result from subtle differences in Gibbs free energy of binding of the respective selectand-selector complexes, in the range of few kJ/mol or less [7, 29, 30]. Linear independency of individual contributions of intermolecular bonds in the selectand-selector complex to the overall interaction, i.e. independent, noncooperative, linear additive contribution of molecular moieties, has been demonstrated [30]. Therefore, for an oligopeptide, overall affinity could be described in terms of Gibbs free energy by a linear combination of each amino acid residue increment. The binding energy for each amino acid would be modified by stereo-configuration at the α-carbon, on average by 0.1 kJ/mol [30]. The resulting overall Gibbs free energy would be ambiguous regarding amino acid stereo-configuration. Furthermore, this scenario would require every amino acid to interact *pari passu*, which is improbable due to secondary structure formation of oligopeptides and distance between stationary phase ligands: with practically achievable ligand densities ≤50% on silica based phases, average distance from ligand to ligand, when measured as distance of silanol groups is 6.6 Å [3], whereas the distance from one amino acid side chain to the next has 3.5 Å [14]. The observed interaction and hence retention can therefore be regarded as the result of statistical superposition of interactions of varying number of amino acids of the peptide with chiral selectors along the path through the column.

Experimental support for this hypothesis is provided by the decline of selectivity α between all-R and all- S oligopeptides with increasing peptide chain length [6]. Czerwenka et al. could separate oligomers of Alanine up to n=10, yet α decreased to ~1.5 (baseline separation) or lower above a chain length of 3. Similar selectivity has been reported for stereoisomeric D/L permutations for di-, tri and tetrapeptides [11, 13, 36]. Mixed stereo-configurations of the respective peptides would therefore lay in between the two retention times of the all-R and all-S oligopeptides with the possibility of inconclusive results comprising determination due to identical binding energy.

Constraints from a practical standpoint would be, that for an exact determination of an intact oligopeptide of unknown stereo-configuration, chromatographic standards providing all D/L permutations must be available, hence a library of n^2 compounds, where n=length of peptide, and chromatographic accuracy and precision (retention time) would be challenging if not unachievable in order to unequivocally determine absolute configurations. If *de novo*

analysis with unknown sequences would be targeted, the demands would be amplified by the number of expected different amino acids.

The previous paragraphs expound the necessity of a hydrolysis step for the determination of absolute amino acid configuration in oligopeptides, as depicted in Fig FP2.1. Usually enzymatic proteolytic protocols (Pepsine, Trypsine, etc) as common in proteomic workflows are employed, yet it can be beneficial to employ chemical hydrolysis protocols. On the one hand, depending on D-amino acid position and peptide sequence, the fragments generated by enzymatic hydrolysis may not be suited for identification of D-amino acid position, as proteases have specific cleavage sites. Publication II demonstrated this limitation of enzymatic fragmentation, as position of a D-Leu was determined to be either 1 or 4, a curtailed and inconclusive result, that needed further investigation.

On the other hand, chemical digestion can be performed much faster. As a matter of fact, for Publication II incubation time for chemical digestion as employed for full hydrolysis (6N HCl, 110°C, 24h) had to be reduced to the order of minutes (20 min vs 90 for enzymatic hydrolysis) to yield a library of fragments with a length distribution expedient for D-amino acid determination. Consequently, chemical digestion resembles a faster and more generic method, independent of oligopeptide sequence, e.g. with chemical digestion, identification of the D-Leu position left open by the enzymatic approach could be completed for Publication II.

Direct analysis of digestion product mix

As aforementioned, enantioselective cinchona alkaloid based stationary phases like ZWIX(+/-) are capable of readily separating small oligopeptides up to n=3 or 4 in some cases allowing for straightforward analysis of oligopeptide mixes yielded by chemical or enzymatic digestion [6, 11, 13, 36], if suitable standard libraries are available. In case of Publication II, suitable standards were not available due to cost reasons, yet analysis with ZWIX(+/-) phases provided beneficial orthogonal chromatographic selectivity.

ZWIX (+/-) as complimentary selectivity to RP C18

Peptide separation via reversed phase chromatography as part of proteomic workflows is the standard approach [32, 35, 39]. Albeit, as these workflows usually target the generation of longer oligopeptides (n=10-20), that are separated by C18 reversed phase chromatography for identification, i.e. sequencing via mass spectrometric CID-MS/MS fragmentation analysis and subsequent sequence mapping [1]. Small oligopeptides are not in the focus of these workflows, due to their sequence ambiguity regarding large oligopeptides and proteins and as a result employed chromatographic methods are not optimized for separation of these small peptides. Also C18 resembles a poor choice for di- to pentapeptides, as presence of a single hydrophilic residue (Ser, Thr, Glu , Asp) or even more charged residue (Lys, His, Asn, Gln) under the usual acidic RP conditions (0.1% formic acid in eluent, pH ~ 2.9) often result in non-retention and elution with column dwell time t_0. For sole identification via coupled tandem high resolution mass spectrometry, this still may suffice, yet quantitative information is usually distortet by ion suppression (see Publication VII for reference) and baseline separation is mandatory for workflows relying on preparative chromatography steps for pure compounds used in downstream analysis like in Publication II.

Enantioselective zwitterionic stationary phase material like ZWIX(+), based on trans-(1''S,2''S)-*N*-[[[(8S,9R)-6'-methoxycinchonan-9-yl]oxy]carbonyl]-2''-aminocyclohexanesulphonic acid, and ZWIX(-), based on trans-(1''R,2''R)-*N*-[[[(8R,9S)-6'-methoxycinchonan-9-yl]oxy]carbonyl]-2''-aminocyclohexane-sulphonic acid) bear selectivity that is not provided and orthogonal to C18, as depicted in Fig FP2.2

AA and dipeptide: 1 I; 2 L; 3 LE; 4 LL; 5 LS; 6 SI; 7 SL; 8 TL;

tri- to pentapeptides: 9 ETL; 10 SLL; 11 TLL; 12 ETLL; 13 ETLLS; 14 LETL; 15 LETLL; 16 LSLL; 17 LLSLL; 18 SLLS; 19 SLLSI; 20 TLLS; 21 TLLSL;

hepta- to nonapeptides: 22 ETLLSL; 23 ETLLSLL; 24; ETLLSLLSI; 25 LETLLS; 26; LETLLSL; 27 LETLLSLLS; 28 LLSLLSI; 29 LSLLSI; 30 TLLSLL; 31 TLLSLLS; 32 TLLSLLSI;

3-HDA-peptides: 33 3-HDA-LETLL; 34 3-HDA-LETLLS; 35 3-HDA-LETLLSL; 36 3-HDA-LETLLSLL; 37 3-HDA-LETLLSLLSI

Figure FP2 2Fig FP2.2 Left: Orthogonality Plot for C18 and ZWIX. Right: Sequence of Poaeamide hydrolysis fragments. C18 is indiscriminative against peptides of 5 residues and smaller, which elute with dwell time t_0, at 2 min. ZWIX performs with orthogonal selectivity, due to the underlying retention mechanism. Reproduced with permission from Publication II, Copyright (2016), Elsevier

Figure FP2 3 Structure of ZWIX (+/-) and ionic/pi-interaction binding mode Reproduced with permission from Publication II and from Ilisz I., Bajtai A., Péter A., Lindner W. (2019) Cinchona Alkaloid-Based Zwitterionic Chiral Stationary Phases Applied for Liquid Chromatographic Enantiomer Separations: An Overview. In: Scriba G.K.E. (eds) Chiral Separations. Methods in Molecular Biology, vol 1985. Humana, New York, NY

ZWIX (+/-) and anomeric analysis

A full stereoconfiguration analysis would include position of allo-Isoforms of amino acids, as the presence of this isoforms has been shown to exert influence on peptide backbone rigidity and biological activity [4, 18]. In case of Poaeamide, the amino acids in question would be Leucine (Leu), Isoleucine (Ile) and Threonine (Thr). Literature on the separation of allo-isoforms is scarce. For allo-Threonine, complete separation was demonstrated by Zhao et al [41], however, by using an offline two dimensional method, with the first dimension comprising of reversed phase and the second of a chiral stationary phase.

As two dimensional approaches are accompanied by increased technical effort and reduced sensitivity [28, 33], a comprehensive one-dimensional and straightforward chromatographic method for separation all 4 configurations (D/L; allo/normal), ideally without prior derivatization would still be desirable. The capabilities of ZWIX-type stationary phases were examined in this regard. Although no complete separation was achievable, the results of Publication II could be basis for further investigation and improvement. In accordance with the separation mechanism (s. Fig FP2. 3 and introduction Publication I), the separation of Thr is more complete (Fig FP2.4), as an inverted configuration of a hydroxy group is expected to have a greater discriminatory effect on a separation mediated by polar / ionic interaction. Isoleucine with its isomer Leucine was less readily separated, as the different configurations of the short aliphatic sidechain are anticipated to have only minor steric effects. For Threonine, with derivatization by Sangers reagent (1-fluoro-2,4-dinitrobenzene) to DNP-Thr, a separation feasible for qualitative interpretation could be provided. To conclude, for ZWIX-type and related chiral selectors, the employment of other derivatization agents, as well as different stationary phase materials (e.g. core shell) and column dimensions, linear velocity etc. would be a worthwhile examination, as optimization of chromatographic performance could yield baseline separation.
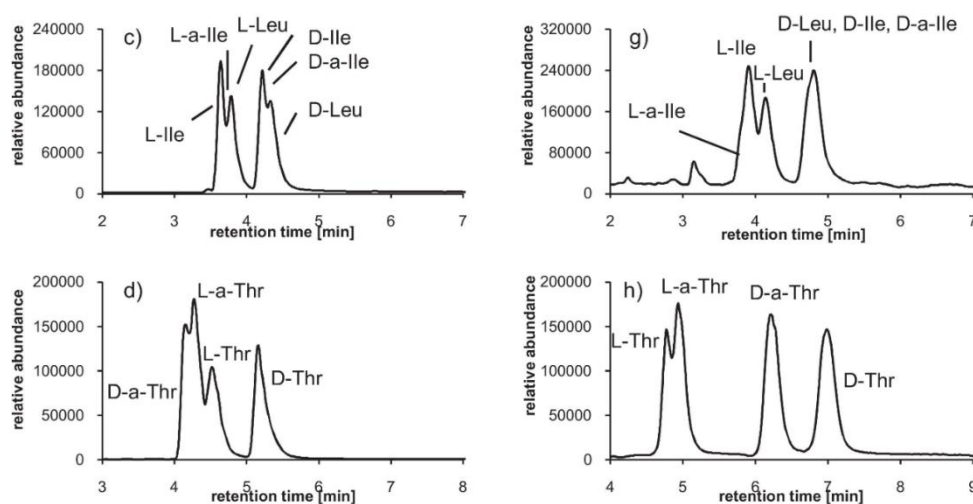

Figure FP2 4 Left column: Separation of underivatized allo-isoforms of Thr and Ile/Leu. Right: column: increase of selectivity by derivatization with Sangers reagent. Reproduced with permission from Publication II, Copyright (2016) Elsevier.

Results of Publication II / VIII

The absolute configuration of Poaeamide with exception of the allo-isomeric configurations could be demonstrated. Although not necessary for stereo-configuration analysis in case of Publication II, as results generated from pepsin and chemical digests via C18 chromatography were sufficient to solve all D and L configurations, the usefulness of Cinchona Alkaloid-Based zwitterionic chiral stationary phases (ZWIX +/-) as valuable tools providing orthogonality to C18 and HILIC type columns could be demonstrated. The latter was more detailed discussed in Publication VIII.

Abbreviations

CLP: cyclic lipopeptides

NMR: nuclear magnetic resonance spectroscopy

GC-MS: gas chromatography coupled to mass spectrometry

HILIC: Hydrophilic interaction chromatography

References

[1]   R. Aebersold and M. Mann. Mass spectrometry-based proteomics. *Nature*, 422 (6928): 198, 2003.

[2]   U. Behrendt, A. Ulrich, and P. Schumann. Fluorescent pseudomonads associated with the phyllosphere of grasses; pseudomonas trivialis sp. nov., pseudomonas poae sp. nov. and pseudomonas congelans sp. nov. *International journal of systematic and evolutionary microbiology*, 53 (5): 1461–1469, 2003. .

[3]   H. Bergna. The colloid chemistry of silica: An overview in the colloid chemistry of silica, 1994.

[4]   D. W. Carney, K. R. Schmitz, J. V. Truong, R. T. Sauer, and J. K. Sello. Restriction of the conformational dynamics of the cyclic acyldepsipeptide antibiotics improves their antibacterial activity. *Journal of the American Chemical Society*, 136 (5): 1922–1929, 2014. .

[5]   M. Chandrasekaran, B. Thangavelu, S. C. Chun, and M. Sathiyabama. Proteases from phytopathogenic fungi and their importance in phytopathogenicity. *Journal of General Plant Pathology*, 82 (5): 233–239, 2016. .

[6]   C. Czerwenka, M. Lämmerhofer, N. M. Maier, K. Rissanen, and W. Lindner. Direct high-performance liquid chromatographic separation of peptide enantiomers: study on chiral recognition by systematic evaluation of the influence of structural features of the chiral selectors on enantioselectivity. *Analytical chemistry*, 74 (21): 5658–5666, 2002.

[7]   C. Czerwenka, M. M. Zhang, H. Kählig, N. M. Maier, K. B. Lipkowitz, and W. Lindner. Chiral recognition of peptide enantiomers by cinchona alkaloid derived chiral selectors: mechanistic investigations by liquid chromatography, nmr spectroscopy, and molecular modeling. *The Journal of organic chemistry*, 68 (22): 8315–8327, 2003. .

[8]   J. D'aes, G. K. H. Hua, K. De Maeyer, J. Pannecoucque, I. Forrez, M. Ongena, L. E. Dietrich, L. S. Thomashow, D. V. Mavrodi, and M. Höfte. Biological control of rhizoctonia root rot on bean by phenazine-and cyclic lipopeptide-producing pseudomonas cmr12a. *Phytopathology*, 101 (8): 996–1004, 2011. .

[9]   D. Ellis, J. Gosai, C. Emrick, R. Heintz, L. Romans, D. Gordon, S.-E. Lu, F. Austin, and L. Smith. Occidiofungin's chemical stability and in vitro potency against candida species. *Antimicrobial agents and chemotherapy*, 56 (2): 765–769, 2012. .

[10]   X. Fang, K. Tiyanont, Y. Zhang, J. Wanner, D. Boger, and S. Walker. The mechanism of action of ramoplanin and enduracidin. *Molecular BioSystems*, 2 (1): 69–76, 2006. .

[11]   T. Z. E. H. P. Franco, M. L. A. S.-E. Heike, and G. H. G. W. Lindner. Peptide analysis: Zwitterionic chiral ion-exchangers as complementary option to hilic and to reversed-phase chromatography. 2016.

[12]   S. Goı̂tze, R. Herbst-Irmer, M. Klapper, H. Goı̂rls, K. R. Schneider, R. Barnett, T. Burks, U. Neu, and P. Stallforth. Structure, biosynthesis, and biological activity of the cyclic lipopeptide anikasin. *ACS chemical biology*, 12 (10): 2498–2502, 2017. .

[13]   F. Ianni, A. Carotti, M. Marinozzi, G. Marcelli, A. Di Michele, R. Sardella, W. Lindner, and B. Natalini. Diastereo-and enantioseparation of a nα-boc amino acid with a zwitterionic quinine-based stationary phase: focus on the stereorecognition mechanism. *Analytica chimica acta*, 885: 174–182, 2015. .

[14]   H.-D. Jakubke and H. Jeschkeit. Aminosäuren, peptide, proteine. verlag chemie, 1982.

[15]   R. Jerala. Synthetic lipopeptides: a novel class of anti-infectives. *Expert opinion on investigational drugs*, 16 (8): 1159–1169, 2007. .

[16]   J. Kirstein, A. Hoffmann, H. Lilie, R. Schmidt, H. Rübsamen-Waigmann, H. Brötz-Oesterhelt, A. Mogk, and K. Turgay. The antibiotic adep reprogrammes clpp, switching it

from a regulated to an uncontrolled protease. *EMBO molecular medicine*, 1 (1): 37–49, 2009.
.

[17]   B.-G. Lee, E. Y. Park, K.-E. Lee, H. Jeon, K. H. Sung, H. Paulsen, H. Rübsamen-Schaeff, H. Brötz-Oesterhelt, and H. K. Song. Structures of clpp in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nature structural & molecular biology*, 17 (4): 471, 2010. .

[18]   I. Malik and H. Brötz-Oesterhelt. Conformational control of the bacterial clp protease by natural product antibiotics. *Natural product reports*, 34 (7): 815–831, 2017. .

[19]   S. M. Mandal, A. E. Barbosa, and O. L. Franco. Lipopeptides in microbial infection control: scope and reality for industry. *Biotechnology Advances*, 31 (2): 338–345, 2013. .

[20]   C. T. Mascio, L. I. Mortin, K. T. Howland, A. D. Van Praagh, S. Zhang, A. Arya, C. L. Chuong, C. Kang, T. Li, and J. A. Silverman. In vitro and in vivo characterization of cb-183,315, a novel lipopeptide antibiotic for treatment of clostridium difficile. *Antimicrobial agents and chemotherapy*, 56 (10): 5023–5030, 2012..

[21]   I. Mora, J. Cabrefiga, and E. Montesinos. Cyclic lipopeptide biosynthetic genes and products, and inhibitory activity of plant-associated bacillus against phytopathogenic bacteria. *PLoS One*, 10 (5): e0127738, 2015. .

[22]   A. Müller, D. Münch, Y. Schmidt, K. Reder-Christ, G. Schiffer, G. Bendas, H. Gross, H.-G. Sahl, T. Schneider, and H. Brötz-Oesterhelt. Lipodepsipeptide empedopeptin inhibits cell wall biosynthesis through ca2+-dependent complex formation with peptidoglycan precursors. *Journal of Biological Chemistry*, 287 (24): 20270–20280, 2012. .

[23]   T. Nielsen, C. Thrane, C. Christophersen, U. Anthoni, and J. Sørensen. Structure, production characteristics and fungal antagonism of tensin–a new antifungal cyclic lipopeptide from pseudomonas fluorescens strain 96.578. *Journal of applied microbiology*, 89 (6): 992–1001, 2000.

[24]   M. Ongena and P. Jacques. Bacillus lipopeptides: versatile weapons for plant disease biocontrol. *Trends in microbiology*, 16 (3): 115–125, 2008. .

[25]   J. M. Raaijmakers, I. De Bruijn, and M. J. de Kock. Cyclic lipopeptide production by plant-associated pseudomonas spp.: diversity, activity, biosynthesis, and regulation. *Molecular Plant-Microbe Interactions*, 19 (7): 699–710, 2006. .

[26]   H. S. Sader, D. Farrell, and R. Jones. Antimicrobial activity of daptomycin tested against gram-positive strains collected in european hospitals: results from 7 years of resistance surveillance (2003-2009). *Journal of Chemotherapy*, 23 (4): 200–206, 2011. .

[27]   T. Schneider, A. Müller, H. Miess, and H. Gross. Cyclic lipopeptides as antibacterial agents–potent antibiotic activity mediated by intriguing mode of actions. *International Journal of Medical Microbiology*, 304 (1): 37–43, 2014. .

[28]   M. R. Schure. Limit of detection, dilution factors, and technique compatibility in multidimensional separations utilizing chromatography, capillary electrophoresis, and field-flow fractionation. *Analytical Chemistry*, 71 (8): 1645–1657, 1999. .

[29]   G. K. Scriba. Recognition mechanisms of chiral selectors: An overview. In *Chiral Separations*, pages 1–33. Springer, 2019. .

[30]   A. Sievers-Engler, W. Lindner, and M. Lämmerhofer. Ligand–receptor binding increments in enantioselective liquid chromatography. *Journal of Chromatography A*, 1363: 79–88, 2014. .

[31]   E. Somner and P. Reynolds. Inhibition of peptidoglycan biosynthesis by ramoplanin. *Antimicrobial agents and chemotherapy*, 34 (3): 413–419, 1990. .

[32]   T. Stasyk and L. A. Huber. Zooming in: fractionation strategies in proteomics. *Proteomics*, 4 (12): 3704–3716, 2004.

[33]   D. R. Stoll and P. W. Carr. Two-dimensional liquid chromatography: a state of the art tutorial, 2016.

[34]   S. K. Straus and R. E. Hancock. Mode of action of the new antibiotic for gram-positive pathogens daptomycin: comparison with cationic antimicrobial peptides and lipopeptides. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1758 (9): 1215–1223, 2006. .

[35]   K. Štulk, V. Pacáková, J. Suchánková, and H. A. Claessens. Stationary phases for peptide analysis by high performance liquid chromatography: a review. *Analytica chimica acta*, 352 (1-3): 1–19, 1997. .

[36]   S. Wernisch and W. Lindner. Versatility of cinchona-based zwitterionic chiral stationary phases: Enantiomer and diastereomer separations of non-protected oligopeptides utilizing a multi-modal chiral recognition mechanism. *Journal of Chromatography A*, 1269: 297–307, 2012.

[37]   Y. Xue, M. Wang, P. Zhao, C. Quan, X. Li, L. Wang, W. Gao, J. Li, X. Zu, D. Fu, et al. Gram-negative bacilli-derived peptide antibiotics developed since 2000. *Biotechnology letters*, 40 (9-10): 1271–1287, 2018. .

[38]   M.-M. Yang, S.-S. Wen, D. V. Mavrodi, O. V. Mavrodi, D. von Wettstein, L. S. Thomashow, J.-H. Guo, and D. M. Weller. Biological control of wheat root diseases by the clp-producing strain pseudomonas fluorescens hc1-07. *Phytopathology*, 104 (3): 248–256, 2014. .

[39]   T. Yoshida. Peptide separation by hydrophilic-interaction chromatography: a review. *Journal of biochemical and biophysical methods*, 60 (3): 265–280, 2004. .

[40]   C. Zachow, G. Jahanshah, I. de Bruijn, C. Song, F. Ianni, Z. Pataj, H. Gerhardt, I. Pianet, M. Lämmerhofer, G. Berg, et al. The novel lipopeptide poaeamide of the endophyte pseudomonas poae re* 1-1-14 is involved in pathogen suppression and root colonization. *Molecular Plant-Microbe Interactions*, 28 (7): 800–810, 2015. .

[41]   H. Zhao, K. Hamase, A. Morikawa, Z. Qiu, and K. Zaitsu. Determination of d-and l-enantiomers of threonine and allo-threonine in mammals using two-step high-performance liquid chromatography. *Journal of Chromatography B*, 810 (2): 245–250, 2004. .

Introduction to Publication I

Linear modeling of binding increments in enantioselective chromatography

As mass spectrometry is a nonchiral analysis technology, enantioselective chromatographic methods are required when examining analytical questions whose analytes can be present with different stereo configurations, as for example retrieving individual D/L amino acid configuration of antimicrobial active oligopeptides examined by Publication II, as described further down in this thesis. At the foundation of enantioselective method development and application lies the theoretical and mechanistic understanding of the process of chromatographic enantioseparation, i.e. of the interaction of enantioselective stationary phases (selector) and the respective chiral analytes (selectands).

Publication I contributes to the state of knowledge by examining enantioseparation of amino acids and related compounds on cinchona alkaloid carbamate based enantioselective stationary phases by quantitatively examining the energetic contribution of individual substituents of the analytes to the overall binding mechanism via quantitative structural retention relationships (QSRR) implemented in the form of Free-Wilson analysis of binding increments.

Liquid chromatography as a tool for studying biological QSAR

Originating from pharmaceutical studies of relationship between molecular structure and biological activity, the term *quantitative structure activity relalionship* (QSAR) refers to mathematical models relating molecular constitution, structure, functional moieties and intrinsic properties like dipole moment, polarity et cetera to biological or chemical activity.

Two different fundamental approaches were published 1964, Hansch-Analysis and Free-Wilson-Analysis, named after their respective Nestors, Corvin Hansch [25], Spencer Free and James Wilson [19]. The former approach uses quantitative descriptors of physicochemical properties of a compound as predictor variables and biological activities as response value, the latter a binary scheme of substituent presence and position in a molecule as predictor variables [36].

Initially, liquid chromatography was utilized as a tool assessing physico-chemcial properties for biological QSAR, rather than itself being studied using the latter, providing for fast probing of hydrophobic parameters of compounds and their substituents respectively [67, 68] At this juncture liquid chromatography yielded retention volumes [27] or indices [7] for direct correlation to biological activity or supplied estimations of QSAR relevant parameters like logP [46, 69] and CLOGP [33], pKa [27]. As of the date of Publication I (2014) there are still studies published correlating retention data to biological activity [35]. An early review on these applications is given by Kaliszan [31].

The term quantitative structure retention relationship (QSRR) was coined by Nasal et al. [51] in a biological QSAR study correlating the liquid chromatographic retention of beta-adrenolytic and antihistamine drugs on a stationary phase of alpha 1-acid glycoprotein to their physiological protein binding data.

QSAR as a tool studying retention processes in liquid chromatography

After a pioneering work of Puech et al [55] in the field of quantitative structure retention relationship (QSRR) in thin layer chromatography, the first application of Hansch-Analysis studying retention processes in liquid chromatography was given by Baker et al. [8], demonstrating on a set of morphine and fentanyl derivatives the feasibility of Hansch-Analysis for predicting the retention index of drugs and for assigning the stereoconfiguration from liquid chromatograms.

Characterization of stationary phases

Zhao et al. [75], focusing more on the properties of stationary phases for liquid chromatography rather than on the analytes, devised the application of linear solvation energy relationships (LSER) to the examination of retention characteristics of aromatic and aliphatic stationary phases. LSERs, introduced first by Altomare et al [3] to liquid chromatography research, generally by means of an Abraham solvation parameter model [1] relate retention behavior of a set of analytes on a given stationary phase to excess molar refraction, the analyte dipolarity/polarizability, the overall or effective hydrogen-bond

acidity and basicity of the analyte, and the McGowan [44] characteristic volume by multiple linear regression, hence the term LSER [64].

The LSER used by Zhao correlated the analytes retention to its molecular volume, dipolarity, hydrogen bond acidity and basicity and molar refraction as well as parameters describing the mobile and stationary phase. While being more complex than a Hansch-Analysis, Zhao's work demonstrated the applicability of molecular structure based approaches to the examination of chromatographic interaction and the characterization of stationary phases.

Predicting retention times – Evaluation of LCMS-data as research driving force

Major research interest regarding QSRR was elicited by the desideratum of predicting retention times. A literature database search refined for results treating only QSRR on reversed phase chromatography (RPLC) on Pubmed-services [72] using combinations of search terms like "retention" "prediction"/"predicting", "chromatography", "modeling" yielded 20 publications studying the prediction of small molecule retention and, driven by the interest in the proteomic analytical key technology liquid chromatography coupled mass spectrometry (LCMS) 45 works handling the accurate prediction of peptide retention times. Amongst the techniques employed for predicting retention times were simple Hansch and Free Wilson-like additive models of molecular residue or hydrophobicity coefficients and peptide chain length [22, 26, 43] and combination of aforementioned additive models with calculated molecular properties like van der Waals volume and cLOGP [32, 6, 5], artificial neuronal networks [53, 61, 18], machine learning approaches like support vector machines [40, 58], regression trees and random forests [57, 66], bayesian regression [66] and uninformative variable elimination [56].

Although Publication I dealing with enantioseparation on quinine-carbamate based chiral ion-exchange columns, being the most extensively researched subfield in this discipline, peptide related QSRR on C18 columns serves as an exemplification of what can be expected from modeling of chromatographic processes and attempts on predicting retention. In the scope of this introduction it adjusts expectance for a rather simple linear regression based Free-Wilson- type of analysis.

In a recent review Moruz and Käll [49] recapitulate predictive accuracy which various retention prediction techniques for peptides. However, as Motuz and Käll demonstrate referencing various works ([34, 54, 59], to cite just a few) even with overall correlation coefficients of $R^2 > 0.9$ an accurate prediction cannot be accomplished. Accurate prediction in this regard would be a retention time forecast with deviations not significantly exceling experimental retention fluctuations. Albeit some peptide sequences can be predicted accurately, alteration in single amino acid sequence position [29], secondary structure [76] and mobile phase can lead to outliers in experimental vs. predicted retention. Especially, as Moruz points out, de novo QSRR prediction, i.e. accurate retention time forecast relying solely on sequence, synonymous to molecular descriptor information in the terminology of QSSR, and description of stationary and mobile phase is yet to be achieved.

The same holds true for QSRR of organic molecules other than peptides, like metabolites, drugs, xenobiotics, toxins, etc, as demonstrated by Wang et al. [71]. Comparing multiple linear regression, partial least squares and random forests methods on a set of over 1700 compounds in C18 chromatography, Wang demonstrated general applicability of aforementioned QSRR techniques, yet lack of exact retention prediction and retention order compared to experimental chromatograms.

Value of QSRR in LCMS / chromatographic fundamental research and application

Rather than accurately predicting retention times, the value of QSRR, especially for mass spectrometry coupled to liquid chromatography, currently lies elsewhere: Improved positive identification of analytes [52, 63], minimizing of false positive identification [34], reduction of interferences [48], optimization of information dependent acquisition (IDA) mass spectrometry in complex mixtures [45], determination of nonlinear chromatographic gradients for optimal usage of gradient peak capacity for whole cell digests [50], bottom-up proteomics [41], and studying of retention mechanisms [2, 74]. As Moruz and Käll [49] also pointed out, data independent (mass spectrometric) analysis (SWATH) [70] in order to reduce complexity and amount of data generated thereby would be a highly desirable application of QSRR, given that the de novo prediction obstacle could be overcome sufficiently.

More recently, with the advent of high resolution mass spectrometry based metabolomics, employing especially non-targeted LCMS-strategies as its major tool, increased research interest in application of QSRR to untargeted metabolite identification [21, 73, 39, 77, 13, 23] emerged.

QSRR and stereoselective chromatography

Whereas the major research interest in QSRR is focused on non-stereoselective, especially reversed-phase chromatography, the very nature of stereoselective chromatography discriminating between structural species as a matter of course leads to examination of relationships between structure and retention. The term QSRR was extended to quantitative structure enantioseparation relationships (QSERR) by Kaliszan et al. [30]. Among the published studies, in QSERR models like Hansch-Analysis [30, 4, 11, 12] linear solvation energy relationships [14, 42, 10, 47], comparative molecular field analysis (CoMFA)/comparative molecular similarity analysis (CoMSIA) [14, 60, 17, 16] and neural networks [65] have been utilized.

Chiral retention - molecular interaction

Before describing the application of Free-Wilson Analysis to rationalize the individual group contributions to the retention and enantioseparation of N-derivatized amino acids on quinine carbamate chiral stationary phases [30], a brief description of basic chiral recognition principles in enantioselective chromatography is given. A substantial review of mechanisms of chiral recognition in liquid chromatography is given by Lämmerhofer [37].

In contrast to the solvophobic theory of Horvath [28] or the lipophilic theory of Carr [15] applicable to the description of reversed phase chromatography, chiral recognition in enantioseparation cannot be explained by macroscopic observables like bulk solubility or partition between aqueous and organic phases, respectively, as enantiomers exhibit identical physicochemical properties. Neither the thermodynamics of exclusion from the bulk mobile phase (solvophobic theory) nor the adsorption to or partition into the bulk stationary phase (analogous to Carrs lipophilic theory for RP) are suitable to adequately describe enantioseparation. Instead, molecular spatial differences between enantiomers

affecting attractive and repulsive forces, like H-bonds, ionic interactions pi-pi-interactions or dipole stacking [37] give rise to a net difference in free energy of transition between the mobile and stationary phase, hence the difference in retention.

The most common yet simplistic model of chiral recognition is the "three point interaction model" describing the formation of a complex between chiral selective stationary phase (CS) and each of the two R/S-enantiomers (selectands SA: R-SA and S-SA). As depicted in Fig FP1.1, the ideal three point interaction is only achieved by one of the selectands, while the other is only capable of less selector-selectand interactions. The complex formation is governed by intermolecular forces resulting in a complex stabilising enthalpic contribution (ΔH_{int}). According to Fig FP1.1 the complex formation/stabillisation is thermodynamically affected by destabilizing entropic contributions (loss of translational, rotational and vibrational degrees of freedom $\Delta S_{r/t}$, ΔS_{rot}, ΔS_{vib}), destabilizing enthalpic contributions (desolvation of CS and SA, ΔH_{solv}) on the one hand, and stabilizing enthalpic and entropic contributions (intermolecular forces ΔH_{int}, entropy increase by desolvation ΔS_{solv}). The respective steric configurations of both selectands allow the intermolecular bonds/forces and subsequent enthalpic and entropic contributions to be established to different magnitudes. This results in an "ideal fit" of one enantiomer, a complex possessing lower Gibbs free energy, resulting in a more stable complex, stronger interaction and thus longer retention on the stationary phase.
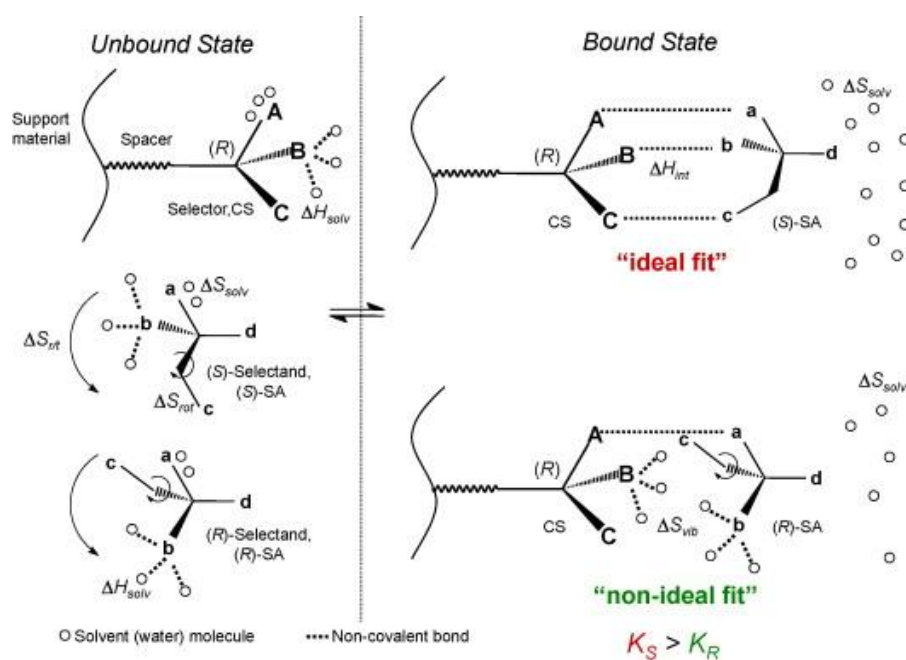
Figure FP1 1 Three-point interaction model illustrating molecular recognition of two chiral selectands (SA) R-SA and S-SA with one chiral center. Reproduced from Journal of Chromatography A, 1217,6, Lämmerhofer, M., Chiral recognition by enantioselective liquid chromatography: Mechanisms and modern chiral stationary phases, 814-856, Copyright (2010), with permission from Elsevier.
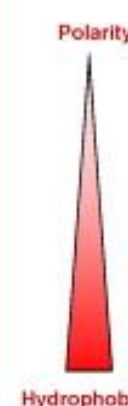

Figure FP1 2 Magnitude of intermolecular forces governing chiral recognition. Values are in vacuo, actual energies are modulated by experimental parameters, e.g. mobile phase composition, temperature. Reproduced from Journal of Chromatography A, 1217,6, Lämmerhofer, M., Chiral recognition by enantioselective liquid chromatography: Mechanisms and modern chiral stationary phases, 814-856, Copyright (2010), with permission from Elsevier.

Thermodynamic description of chiral recognition

The binding of each of the selectands SA to the chiral selector CS reversible forming the [CS-SA] complex is an equilibrium process, thus being quantified by the standard Gibbs free energy. Let Ki be the equilibrium constant of chiral selector and selectand interaction:

$$CS + SA \rightleftharpoons [CS - SA] \quad (1)$$
$$\Delta G_i^0 = -RT \cdot lnK_i \quad (2)$$
$$K_i = \frac{[CS - SA]}{[CS] + [SA]} \quad (3)$$

As described before referring to [37], the Gibbs free energy of the binding process consists of enthalpic (ΔH_{int}, ΔH_{solv}) and entropic contribution ($\Delta S_{r/t}$, ΔS_{rot}, ΔS_{vib}, ΔS_{solv})

$$\Delta H_i^0 = \Delta H_{int}^0 + \Delta H_{solv}^0 \quad (4)$$

$$\Delta S0_i = \Delta S_{r/t}^0 + \Delta S_{rot}^0 + \Delta S_{vib}^0 + \Delta S_{solv}^0 \quad (5)$$

$$\Delta G_i^0 = \Delta H_i^0 - T\Delta S_i^0 \quad (6)$$

Both equations can be combined to the van't Hoff equation

$$lnK_i = -\frac{1}{T}\frac{\Delta H_i^0}{R} + \frac{\Delta S_i^0}{R} \quad (7)$$

Now the equilibrium constant K_i is not directly observable in liquid chromatography. Yet a linear relationship between K_a and the retention factor k_i can be established using the phase ratio $\phi= V_{stationary} /V_{mobile.,}$ where V is the volume of the respective phase in the chromatographic column:

$$k_i = \phi K_i \quad (8)$$

Although other equations describing the relation between k_i and $K_{a,,}$ are known this one is the most widely used in basic chromatographic theory.

Therefore, equation 7 can also be written as

$$\ln k_i = -\frac{1}{TR}(\Delta G^0) + \ln \phi \quad (9)$$

This describes a linear relationship between retention factors and the Gibbs free energy of complex formation. Here, the basic assumption for Publication I emerges with Equation 9: If all Gibbs free energy contributions to the complex formation of individual molecular groups responsible for binding of SA to CS behave additively, i.e. non-cooperatively, the retention factors are a linear combination of the individual group contributions and the magnitude of the individual group contributions can be accessed by a general linear regression model.

$$\ln k_i \propto \Delta G_i^0 = \sum_1^n \Delta G_j^0 \quad (10)$$

As Publication I treats chiral separation, besides k_i also the selectivity α can be expressed accordingly and enantioseparation is the given by

$$\ln \alpha \propto \Delta\Delta G_i^0 = \sum_1^n \Delta\Delta G_j^0 \qquad (11)$$

An early study employing this thermodynamic framework was published by Berthod et al. [9].

Free-Wilson-Analysis – QSAR method originating from pharmaceutical research

In the original work [19], SM Free and JW Wilson use a linear equation l to describe and to some extend predict the biological activity of several classes of compounds. Other than Hansch-analysis [hansch1964p], the Free-Wilson approach uses binary coefficients describing the presence or absence of a molecular group at given position in the compounds structure. Let Y_i be the observable biological activity of the compound i the equation can be written as:

$$Y_i = \sum_{i=1}^n b_i \cdot C_i$$

Where C_i is the functional groups' individual contribution to the observable Y and b_i is a binary coefficient, denoting presence by 1 and absence by 0 and n being the overall number of different substituents. Commonly the group contributions C_i are calculated either relative to the average contribution of all residues tested [19] or relative to an unsubstituted reference compound [20], e.g. having only H-substituents in all examined positions, where μ is the reference compounds observable:

$$Y_i = \sum_{i=1}^n b_i \cdot C_i + \mu$$
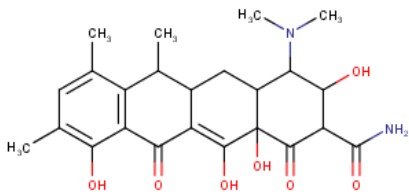
The latter approach resembles somewhat close [36] the Hammet-equation for reactivity of organic compounds [24], also assigning the H-substituted compound as reference.

Also commonly, the response is found to be linear on a logarithmic scale. Consistently, as shown in the previous paragraph, the responses k_R, K_s and $\alpha_{R,S}$ of chiral separations are expected to be found on the logarithmic scale according to the thermodynamic equations.

The group contributions are solved for by linear regression analysis on a training set of compounds, which, provided there is indeed a linear independence of group contributions, eliminate the need for comprehensive testing of all possible combinations of functional groups and positions in the molecule: Table TP1.1 demonstrates the principle of Free Wilson Analysis, by means of a Free-Wilson-Table and shows the results published in the original study of Free and Wilson on the inhibitory potencies of a series of tetracyclines against Staphylococcus aureus [19]. A review on Free-Wilson QSAR principles and studies is given by Kubinyi et. al. [36].

Table TP1 1Free-Wilson QSAR analysis of biological activity of tetracyclines. Reproduced from Journal of Medicinal Chemistry, 7,4, Free and Wilson, A mathematical contribution to structure-activity studies, 395--399, Copyright (1964), with permission from ACS Publications

	R		X			Y			Biological activity
Compound	H	CH3	NO2	Cl	Br	NO2	NH2	CH3CCONH	
III	1		1			1			60
IV	1			1		1			21
V	1				1	1			15
VI	1			1			1		525
VII	1				1		1		320
VII	1		1				1		275
IX		1	1				1		160
X		1	1					1	15
XI		1			1		1		140
XII		1			1			1	75

	R	H	CH3	
		75	-112	
X		NO2	Cl	Br
		-26	84	-16
Y		NO2	NH2	CH3CCONH
		-218	123	123

Utility of Free-Wilson-Analysis

As Free and Wilson state, the main utility of such an approach is to study the underlying additivity principle of group contributions, i.e. independence of individual group contributions and gain a quantitative measure of each group contribution. Molecular moieties violating the additivity principle, e.g. those with cooperative effects are revealed by their poor fit to linear model, i.e. high regression residuals.

Functional groups can be assessed by their influence on chiral separation, especially those specifically introduced by derivatization in order to modify enantioseparation. Also groups actually diminishing enantioseparation can be revealed.

Binding hypotheses of CS-SA complexes derived from structural chemistry methods (X-ray diffraction, NOE-NMR) can be examined quantitatively without the need of a comprehensive library of compounds, as long as the linear independence of group contributions holds true.

Limitations of Free-Wilson-Analysis

Inherently to its principle using only presence or absence of groups instead of physicochemical or spatial parameters, Free-Wilson-Analysis can only provide group contributions for groups being examined in the training set. De novo calculations are impossible and estimations on compounds bearing untrained groups or having an altered molecular structure of the basis compound are not feasible, as this would require invalid presumptions of the untrained groups enthalpic and entropic contributions. The Free-

Wilson-Model is thus no approach used for de novo challenges, but rather in a descriptive and explorative application.

The assessment of the predictive quality Free-Wilson models is dependent furthermore on acceptance criteria for successful prediction. Where accounting for ca. 90% of observed variance by a linear model can be sufficient in biological activity analysis, like in the original work [19], it can be inadequate for calculations demanding better accuracy like retention time prediction in liquid chromatography.

Goal and approach of this work

Goal of publication I was to examine and rationalize the individual group contributions to the retention and enantioseparation of N-derivatized amino acids on quinine carbamate chiral stationary phases [30][38]. The set of compounds consisted of 142 training compounds, mostly amino acids which were N-carbonyl derivatized with the protecting groups shown in fig FP1.3. The chromatography was conducted on a quinine carbamate chiral stationary phase, also shown in Fig. FP1.3:
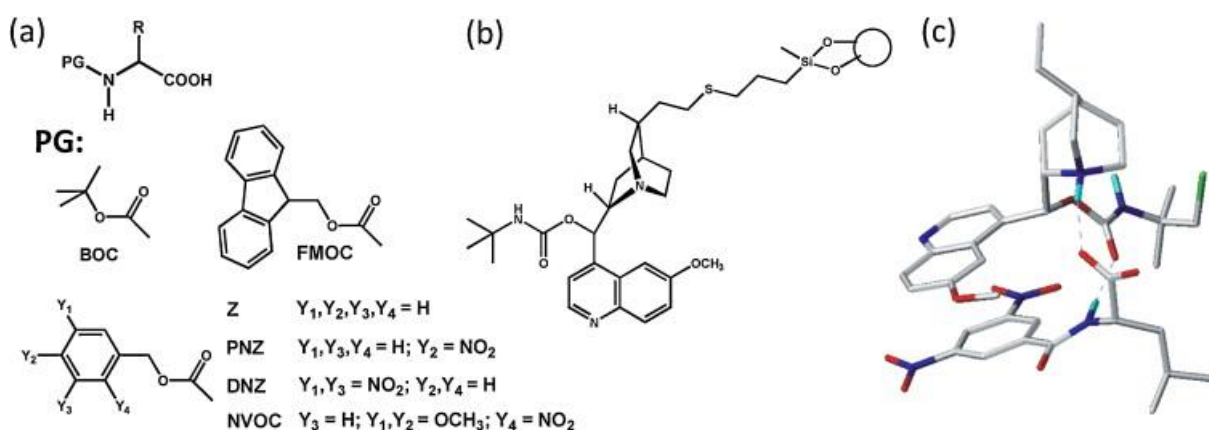
Figure FP1 3 Protecting groups (a), chiral selector (b) and (c) X-ray crystal structure of co-crystallized quinine carbamate and 3,5-dinitrobenzoylamino acid leucine and illustrating the CS-SA complex and the intermolecular bonds stabilizing the complex. Reproduced from Journal of Chromatography A, 1363, Sievers-Engler et. al, Ligand-receptor binding increments in enantioselective liquid chromatography, 79-88, Copyright (2014), with permission from Elsevier.

In order to handle the size of the dataset and facilitate statistical procedures, a software network as shown in Fig FP1.4 was implemented using Java, R, and MySQL:
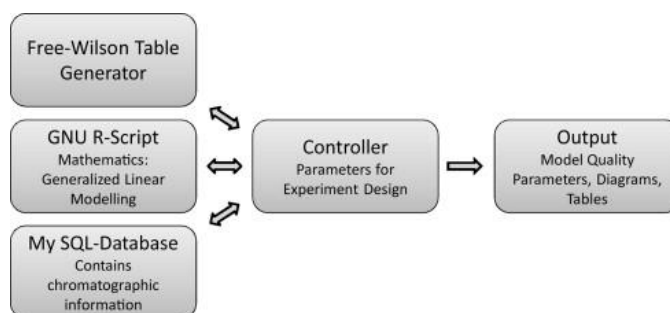
Figure FP1 4 Software network implemented for Publication I. Reproduced from Journal of Chromatography A, 1363, Sievers-Engler et. al, Ligand--receptor binding increments in enantioselective liquid chromatography, 79-88, Copyright (2014), with permission from Elsevier.

Results and discussion

The successful application of Free-Wilson-Analysis to derive the retention increments on the enantiomers could be shown in Publication 1. This success was based on the finding, that the assumed independence of contributions of the respective molecular groups to the binding to the stationary phase holds true. Otherwise, the application of a linear model would result in insufficient performance expressed in low correlation coefficients. These findings were substantiated by cross validation by means of scrambling tests and by leave-n-out. The scrambling test was conducted by randomly reassigning retention coefficients k1 and k2 to tested compounds, resulting as expected in low quality linear models and predictive capability of the latter. The leave-n-out test was derived from the frequently used leave-one-out cross-validation test in order to deal with the extent of the data set. Monitoring the correlation coefficient R^2 and the predictive capability while repeatedly (n=10) removing an increasing number of randomly chosen training compounds, the Free-Wilson approach maintained modeling performance up to over 40% of the initial 142 training compounds removed.

Whilst the models exhibiting good correlation (R>0.95), the standard deviations of the predicted retention indices were too large measured by the demanding requirements of an exact prediction of retention time and enantioseparation, a finding in accordance to the simplistic approach of the Free-Wilson model.

Albeit good overall fit, a deviation from the additivity principle could be revealed: Derivatives of basic, i.e. positive charged compounds e.g. Arginine or Histidine, showed poor fit to the

linear model. These findings were confirmed by establishing a GLM with only neutral charged training compounds und subsequently adding an increasing number of negative, positive or neutral charged compounds to the model. Whilst negative and neutral charged compounds had no effect on model quality, the addition of positive charged training compounds deteriorated model quality.

While the examination of the mechanistic nature of nonconformance of basic compounds to the GLM was not in the scope of Publication I, the disclosure of these compounds behavior attests to the usefulness of the Free-Wilson modeling approach. As linear modeling revealed inapt, this could be an indication for a cooperative binding mechanism. Although other reasons could account for this deviation from linearity as well, the results demonstrate how linear modeling can serve as cross-check for mechanistic assumptions revealing demand for further mechanistic investigation.

With regard to the different stabilization of CS-R-SA / CS-S-SA complexes, hence the difference in retention and enantioseperative selectivity, the Gibbs free energy contribution of the protecting groups was found to be the driving factor, while the influence of the amino acid side chain itself is not significant. While three of the aliphatic amino acid residues, tert-Leucine, Isoleucine and allo-Isoleucine, exerted an increasing effect on selectivity, all other side chains contributed via a reduction of enantionselectivity. Furthermore, a second acidic group in the amino acid side chain revealed to be competitive to the primary one, though only when in close proximity to the stereogenic center.

Finally, one finding directly assent to the binding model derived from the x-ray structure for co-crystallized CS-SA complexes and further showcasing the mechanistic utility of Free-Willson-Analysis: As illustrated in Fig FP1-3 there is a stabilizing in H-bond between the chiral selectors carbamate-carbonyl oxygen and the selectands alpha-N-H-moiety, its formation sterically affected by the configuration of the stereogenic centre. Accordingly, beta-amino acids with reduced interaction possibility due to longer distance and secondary amino acids lacking the possibility of the H-Bond completely, exert almost no difference in contribution of the side chain between the two enantiomers.

Overall the derived models provide useful information on the chiral recognition mechanism and importance of the group contributions to retention and enantiomer separation.

Abbreviations

QSRR: quantitative structural retention relationships

QSAR: quantitative structure activity relationship

logP: octanol-water-partition coefficient

CLOGP: logP predicted by group contribution calculation method

pKa: negative decadic logarithm of acid constant

LSER: linear solvation energy relationships

IDA: information dependent acquisition

QSERR: quantitative structure enantioseparation relationships

CoMSIA: comparative molecular similarity analysis

CoMFA: comparative molecular field analysis

CS: chiral selective stationary phase

SA: selectand

GLM: generalized linear model

References

[1]   Michael H Abraham and Gary S Whiting. Hydrogen bonding: Xxi. solvation parameters for alkylaromatic hydrocarbons from gas-liquid chromatographic data. *Journal of Chromatography A*, 594(1-2):229–241, 1992.

[2]   Tânia MG Almeida, Andrei Leitão, Maria Luiza C Montanari, and Carlos A Montanari. The molecular retention mechanism in reversed-phase liquid chromatography of meso-ionic compounds by quantitative structure–retention relationships (qsrr). *Chemistry & biodiversity*, 2(12):1691–1700, 2005.

[3]   C Altomare, S Cellamare, A Carotti, and M Ferappi. Linear solvation energy relationships in reversed-phase liquid chromatography. examination of rp-8 stationary phases for measuring lipophilicity parameters. *Farmaco (Societa chimica italiana: 1989)*, 49(6):394–401, 1994.

[4]   Cosimo Altomare, Angelo Carotti, Saverio Cellamare, Francesca Fanelli, Francesco Gasparrini, Claudio Villani, Pierre-Alain Carrupt, and Bernard Testa. Enantiomeric resolution of sulfoxides on a dach-dnb chiral stationary phase: A quantitative structure–enantioselective retention relationship (qserr) study. *Chirality*, 5(7):527–537, 1993.

[5]   Tomasz Baì§czek and Roman Kaliszan. Predictive approaches to gradient retention based on analyte structural descriptors from calculation chemistry. *Journal of Chromatography A*, 987(1):29–37, 2003.

[6]   Tomasz Baì§czek, Pawe Wiczling, Micha Marsza, Yvan Vander Heyden, and Roman Kaliszan. Prediction of peptide retention at different hplc conditions from multiple linear regression models. *Journal of proteome research*, 4(2):555–563, 2005.

[7]   John K Baker, David O Rauls, and Ronald F Borne. Correlation of biological activity and high-pressure liquid chromatographic retention index for a series of propranolol, barbiturate, and anthranilic acid analogs. *Journal of medicinal chemistry*, 22(11):1301–1306, 1979.

[8]   John K Baker, R-E_ Skelton, Thomas N Riley, and Jerome R Bagley. Estimation of high pressure liquid chromatographic retention indices of narcotic analgetics and related drugs. *Journal of chromatographic science*, 18(4):153–158, 1980.

[9]   Alain Berthod, San Chun Chang, and Daniel W Armstrong. Empirical procedure that uses molecular structure to predict enantioselectivity of chiral stationary phases. *Analytical chemistry*, 64(4):395–404, 1992.

[10]   Alain Berthod, Clifford R Mitchell, and Daniel W Armstrong. Could linear solvation energy relationships give insights into chiral recognition mechanisms?: 1. π–π and charge interaction in the reversed versus the normal phase mode. *Journal of Chromatography A*, 1166(1-2):61–69, 2007.

[11]   Tristan D Booth and Irving W Wainer. Investigation of the enantioselective separations of α-alkylarylcarboxylic acids on an amylose tris (3, 5-dimethylphenylcarbamate) chiral stationary phase using quantitative structure-enantioselective retention relationships identification of a conformationally driven chiral recognition mechanism. *Journal of Chromatography A*, 737(2):157–169, 1996.

[12]   Katarna Boronová, Jozef Lehotay, Katarna Hrobonová, and Daniel W Armstrong. Study of physicochemical interaction of aryloxyaminopropanol derivatives with teicoplanin

and vancomycin phases in view of quantitative structure–property relationship studies. *Journal of Chromatography A*, 1301:38–47, 2013.

[13]   Tobias Bruderer, Emmanuel Varesio, and Gérard Hopfgartner. The use of lc predicted retention times to extend metabolites identification with swath data acquisition. *Journal of Chromatography B*, 2017.

[14]   Angelo Carotti, Cosimo Altomare, Saverio Cellamare, AnnaMaria Monforte, Giancarlo Bettoni, Fulvio Loiodice, Nicola Tangari, and Vincenzo Tortorella. Lfer and comfa studies on optical resolution of α-alkyl α-aryloxy acetic acid methyl esters on dach-dnb chiral stationary phase. *Journal of Computer-Aided Molecular Design*, 9(2):131–138, 1995.

[15]   Peter W Carr, Jianjun Li, Andrew J Dallas, David I Eikens, and Lay Choo Tan. Revisionist look at solvophobic driving forces in reversed-phase liquid chromatography. *Journal of Chromatography A*, 656(1-2):113–133, 1993.

[16]   Sun Ok Choi, Seok Ho Lee, and Hea-Young Park Choo. Prediction on the chiral behaviors of drugs with amine moiety on the chiral cellobiohydrolase stationary phase using a partial least square method. *Archives of pharmacal research*, 27(10):1009–1015, 2004.

[17]   Walter MF Fabian, Wolfgang Stampfer, Monika Mazur, and Georg Uray. Modeling the chromatographic enantioseparation of aryl-and hetarylcarbinols on ulmo, a brush-type chiral stationary phase, by 3d-qsar techniques. *Chirality*, 15(3):271–275, 2003.

[18]   MH Fatemi, MH Abraham, and CF Poole. Combination of artificial neural network technique and linear free energy relationship parameters in the prediction of gradient retention times in liquid chromatography. *Journal of Chromatography A*, 1190(1):241–252, 2008.

[19]   Spencer M Free and James W Wilson. A mathematical contribution to structure-activity studies. *Journal of Medicinal Chemistry*, 7(4):395–399, 1964.

[20]   Toshio Fujita and Takashi Ban. Structure-activity relation. 3. structure-activity study of phenethylamines as substrates of biosynthetic enzymes of sympathetic transmitters. *Journal of medicinal chemistry*, 14(2):148–152, 1971.

[21]   Krzysztof Gorynski, Barbara Bojko, Alicja Nowaczyk, Adam Bucinski, Janusz Pawliszyn, and Roman Kaliszan. Quantitative structure–retention relationships models for prediction of high performance liquid chromatography retention time of small molecules: Endogenous metabolites and banned compounds. *Ánalytica chimica acta*, 797:13–19, 2013.

[22]   Dacheng Guo, Colin T Mant, Ashok K Taneja, and Robert S Hodges. Prediction of peptide retention times in reversed-phase high-performance liquid chromatography ii. correlation of observed and predicted peptide retention times factors and influencing the retention times of peptides. *Journal of Chromatography A*, 359:519–532, 1986.

[23]   Takashi Hagiwara, Seiji Saito, Yoshifumi Ujiie, Kensaku Imai, Masanori Kakuta, Koji Kadota, Tohru Terada, Kazuya Sumikoshi, Kentaro Shimizu, and Tatsunari Nishi. Hplc retention time prediction for metabolome analysi. *Bioinformation*, 5(6):255, 2010.

[24]   Louis P Hammett. The effect of structure upon the reactions of organic compounds. benzene derivatives. *Journal of the American Chemical Society*, 59(1):96–103, 1937.

[25]   Corwin Hansch and Toshio Fujita. p-σ-π analysis. a method for the correlation of biological activity and chemical structure. *Journal of the American Chemical Society*, 86(8):1616–1626, 1964.

[26]   MTW Hearn, MI Aguilar, CT Mant, and RS Hodges. High-performance liquid chromatography of amino acids, peptides and proteins: Lxxxv. evaluation of the use of hydrophobicity coefficients for the prediction of peptide elution profiles. *Journal of Chromatography A*, 438:197–210, 1988.

[27]   Doug Henry, John H Block, John L Anderson, and Garnet R Carlson. Use of high-pressure liquid chromatography for quantitative structure-activity relationship studies of sulfonamides and barbiturates. *Journal of medicinal chemistry*, 19(5):619–626, 1976.

[28]   Csaba Horvath, Wayne Melander, and Imre Molnar. Solvophobic interactions in liquid chromatography with nonpolar stationary phases. *Journal of Chromatography A*, 125(1):129–156, 1976.

[29]   Richard A Houghten and Stephen T DeGraw. Effect of positional environmental domains on the variation of high-performance liquid chromatographic peptide retention coefficients. *Journal of Chromatography A*, 386:223–228, 1987.

[30]   R Kaliszan, TAG Noctor, and IW Wainer. Quantitative structure-enationselective retention relationships for the chromatography of 1, 4-benzodiazepines on a human serum albumin based hplc chiral stationary phase: An approach to the computational prediction of retention and enantioselectivity. *Chromatographia*, 33(11-12):546–550, 1992.

[31]   Roman Kaliszan. Chromatography in studies of quantitative structure-activity relationships. *Journal of Chromatography A*, 220(1):71–83, 1981.

[32]   Roman Kaliszan, Tomasz Baczek, Anna Cimochowska, Paulina Juszczyk, Kornelia Wisniewska, and Zbigniew Grzonka. Prediction of high-performance liquid chromatography retention of peptides with the use of quantitative structure-retention relationships. *Proteomics*, 5(2):409–415, 2005.

[33]   Roman Kaliszan, Anna Kaliszan, and Irving W Wainer. Deactivated hydrocarbonaceous silica and immobilized artificial membrane stationary phases in high-performance liquid chromatographic determination of hydrophobicities of organic bases:

relationship to log p and clogp. *Journal of pharmaceutical and biomedical analysis*, 11(6):505–511, 1993.

[34]   Aaron A Klammer, Xianhua Yi, Michael J MacCoss, and William Stafford Noble. Improving tandem mass spectrum identification using peptide retention time prediction across diverse chromatography conditions. *Analytical Chemistry*, 79(16):6111–6118, 2007.

[35]   Marcin Koba, Tomasz Baczek, and Micha Piotr Marsza. Importance of retention data from affinity and reverse-phase high-performance liquid chromatography on antitumor activity prediction of imidazoacridinones using qsar strategy. *Journal of pharmaceutical and biomedical analysis*, 64:87–93, 2012.

[36]   Hugo Kubinyi. Free wilson analysis. theory, applications and its relationship to hansch analysis. *Molecular Informatics*, 7(3):121–133, 1988.

[37]   Michael Lämmerhofer. Chiral recognition by enantioselective liquid chromatography: mechanisms and modern chiral stationary phases. *Journal of Chromatography A*, 1217(6):814–856, 2010.

[38]   Michael Lämmerhofer, Pilar Franco, and Wolfgang Lindner. Quinine carbamate chiral stationary phases: Systematic optimization of steric selector-selectand binding increments and enantioselectivity by quantitative structure-enantioselectivity relationship studies. *Journal of separation science*, 29(10):1486–1496, 2006.

[39]   B Law. Analysis of basic compounds on a silica column with an aqueous methanol eluent: The use of quantitative structure-retention relationships in metabolite identification. *Journal of Chromatography A*, 407:1–18, 1987.

[40]   HX Liu, CX Xue, RS Zhang, XJ Yao, MC Liu, ZD Hu, and Bo Tao Fan. Quantitative prediction of log k of peptides in high-performance liquid chromatography based on molecular descriptors by using the heuristic method and support vector machine. *Journal of chemical information and computer sciences*, 44(6):1979–1986, 2004.

[41]   Anna A Lobas, Anatoly N Verenchikov, Anton A Goloborodko, Lev I Levitsky, and Mikhail V Gorshkov. Combination of edman degradation of peptides with liquid chromatography/mass spectrometry workflow for peptide identification in bottom-up proteomics. *Rapid Communications in Mass Spectrometry*, 27(3):391–400, 2013.

[42]   Jana Lokajová, Eva Tesarová, and Daniel W Armstrong. Comparative study of three teicoplanin-based chiral stationary phases using the linear free energy relationship model. *Journal of Chromatography A*, 1088(1-2):57–66, 2005.

[43]   Colin T Mant, TW Lorne Burke, James A Black, and Robert S Hodges. Effect of peptide chain length on peptide retention behaviour in reversed-phase chromatogrphy. *Journal of Chromatography A*, 458:193–205, 1988.

[44]   JC McGowan. Molecular volumes and structural chemistry. *Recueil des Travaux Chimiques des Pays-Bas*, 75(2):193–208, 1956.

[45]   Peter McQueen, Vic Spicer, Thomas Rydzak, Richard Sparling, David Levin, John A Wilkins, and Oleg Krokhin. Information-dependent lc-ms/ms acquisition with exclusion lists potentially generated on-the-fly: Case study using a whole cell digest of clostridium thermocellum. *Proteomics*, 12(8):1160–1169, 2012.

[46]   Margaret S Mirrlees, Stuart J Moulton, Christopher T Murphy, and Peter J Taylor. Direct measurement of octanol-water partition coefficients by high-pressure liquid chromatography. *Journal of medicinal chemistry*, 19(5):615–619, 1976.

[47]   Clifford R Mitchell, Daniel W Armstrong, and Alain Berthod. Could linear solvation energy relationships give insights into chiral recognition mechanisms?: 2. characterization of macrocyclic glycopeptide stationary phases. *Journal of Chromatography A*, 1166(1-2):70–78, 2007.

[48]   Luminita Moruz, Michael R Hoopmann, Magnus Rosenlund, Viktor Granholm, Robert L Moritz, and Lukas KaÌˆll. Mass fingerprinting of complex mixtures: protein inference from high-resolution peptide masses and predicted retention times. *Journal of proteome research*, 12(12):5730–5741, 2013.

[49]   Luminita Moruz and Lukas Käll. Peptide retention time prediction. *Mass spectrometry reviews*, 2016.

[50]   Luminita Moruz, Daniela Tomazela, and Lukas KaÌˆll. Training, selection, and robust calibration of retention time models for targeted proteomics. *Journal of proteome research*, 9(10):5209–5216, 2010.

[51]   Antoni Nasal, Aleksandra Radwanska, Krzysztof Osmiaówski, Adam Bucinski, Roman Kaliszan, Geoffrey E Barker, Peng Sun, and Richard A Hartwick. Quantitative relationships between the structure of β-adrenolytic and antihistamine drugs and their retention on an α1-acid glycoprotein hplc column. *Biomedical Chromatography*, 8(3):125–129, 1994.

[52]   Magnus Palmblad, Margareta Ramström, Christopher G Bailey, Sandra L McCutchen-Maloney, Jonas Bergquist, and Loreen C Zeller. Protein identification by liquid chromatography–mass spectrometry using retention time prediction. *Journal of Chromatography B*, 803(1):131–135, 2004.

[53]   Konstantinos Petritis, Lars J Kangas, Patrick L Ferguson, Gordon A Anderson, Ljiljana Paša-Tolic, Mary S Lipton, Kenneth J Auberry, Eric F Strittmatter, Yufeng Shen, Rui Zhao, et al. Úse of artificial neural networks for the accurate prediction of peptide liquid chromatography elution times in proteome analyses. *Analytical chemistry*, 75(5):1039–1048, 2003.

[54]   Nico Pfeifer, Andreas Leinenbach, Christian G Huber, and Oliver Kohlbacher. Statistical learning of peptide retention behavior in chromatographic separations: a new kernel-based approach for computational proteomics. *BMC bioinformatics*, 8(1):468, 2007.

[55]   A Puech, G Kister, and J Chanal. Structure-mobility relationships in thin-layer chromatography. applications to some progestogens (author's transl). *Journal of chromatography*, 108(2):345–353, 1975.

[56]   R. Put, M. Daszykowski, T. Baczek, and Y. Vander Heyden. Retention prediction of peptides based on uninformative variable elimination by partial least squares. 5:1618–1625, 2006.

[57]   R Put, C Perrin, F Questier, D Coomans, DL Massart, and Y Vander Heyden. Classification and regression tree analysis for molecular descriptor selection and retention prediction in chromatographic quantitative structure–retention relationship studies. *Journal of Chromatography A*, 988(2):261–276, 2003.

[58]   Raf Put and Yvan Vander Heyden. The evaluation of two-step multivariate adaptive regression splines for chromatographic retention prediction of peptides. *Proteomics*, 7(10):1664–1677, 2007.

[59]   Janice Reimer, Vic Spicer, and Oleg V Krokhin. Application of modern reversed-phase peptide retention prediction algorithms to the houghten and degraw dataset: Peptide helicity and its effect on prediction accuracy. *Journal of Chromatography A*, 1256:160–168, 2012.

[60]   Sabine Schefzick, Michael Lämmerhofer, Wolfgang Lindner, Kenny B Lipkowitz, and Mehran Jalaie. Comparative molecular field analysis of quinine derivatives used as chiral selectors in liquid chromatography: 3d qsar for the purposes of molecular design of chiral stationary phases. *Chirality*, 12(10):742–750, 2000.

[61]   Kosaku Shinoda, Masahiro Sugimoto, Nozomu Yachie, Naoyuki Sugiyama, Takeshi Masuda, Martin Robert, Tomoyoshi Soga, and Masaru Tomita. Prediction of liquid chromatographic retention times of peptides generated by protease digestion of theescherichia coliproteome using artificial neural networks. 5:3312–3317, 2006.

[62]   Adrian Sievers-Engler, Wolfgang Lindner, and Michael Lämmerhofer. Ligand–receptor binding increments in enantioselective liquid chromatography. *Journal of Chromatography A*, 1363:79–88, 2014.

[63]   Eric F Strittmatter, Lars J Kangas, Konstantinos Petritis, Heather M Mottaz, Gordon A Anderson, Yufeng Shen, Jon M Jacobs, David G Camp, and Richard D Smith. Application of peptide lc retention time information in a discriminant function for peptide identification by tandem mass spectrometry. *Journal of proteome research*, 3(4):760–769, 2004.

[64]   S Studzinska and B Buszewski. Linear solvation energy relationships in the determination of specificity and selectivity of stationary phases. *Chromatographia*, 75(21-22):1235–1246, 2012.

[65]   Maciej Szaleniec, Agnieszka Dudzik, Marzena Pawul, and Bartomiej Kozik. Quantitative structure enantioselective retention relationship for high-performance liquid chromatography chiral separation of 1-phenylethanol derivatives. *Journal of Chromatography A*, 1216(34):6224–6235, 2009.

[66]   Feifei Tian, Li Yang, Fenglin Lv, and Peng Zhou. Predicting liquid chromatographic retention times of peptides from the drosophila melanogaster proteome by machine learning approaches. *Analytica chimica acta*, 644(1):10–16, 2009.

[67]   Eric Tomlinson. Chromatographic hydrophobic parameters in correlation analysis of structure-activity relationships. *Journal of Chromatography A*, 113(1):1–45, 1975.

[68]   Eric Tomlinson and JC Dearden. Chromatographic substituent constants and their use in quantitative structure activity relationships: The binding of acetanilides to serum albumin. *Journal of Chromatography A*, 106(2):481–484, 1975.

[69]   Klára Valkó. General approach for the estimation of octanol/water partition coefficient by reversed-phase high-performance liquid chromatography. *Journal of liquid chromatography*, 7(7):1405–1424, 1984.

[70]   John D Venable, Meng-Qiu Dong, James Wohlschlegel, Andrew Dillin, and John R Yates. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature methods*, 1(1):39–45, 2004.

[71]   Chunlei Wang, Michael J Skibic, Richard E Higgs, Ian A Watson, Hai Bui, Jibo Wang, and Jose M Cintron. Evaluating the performances of quantitative structure-retention relationship models with different sets of molecular descriptors and databases for high-performance liquid chromatography predictions. *Journal of Chromatography A*, 1216(25):5030–5038, 2009.

[72]   David L Wheeler, Tanya Barrett, Dennis A Benson, Stephen H Bryant, Kathi Canese, Vyacheslav Chetvernin, Deanna M Church, Michael DiCuccio, Ron Edgar, Scott Federhen, et al. Database resources of the national center for biotechnology information. *Nucleic acids research*, 36(suppl_1):D13–D21, 2007.

[73]   Liang Wu, Ping Gong, Yuzheng Wu, Ke Liao, Hanyuan Shen, Qu Qi, Huiying Liu, Guangji Wang, and Haiping Hao. An integral strategy toward the rapid identification of analogous nontarget compounds from complex mixtures. *Journal of Chromatography A*, 1303:39–47, 2013.

[74]   Yanling Zhang, Yiran Guo, Yun Wang, and Yanjiang Qiao. Study on relationship between three-dimensional quantitative structure and retention parameters of chromatography. *Se pu= Chinese journal of chromatography*, 23(3):223–228, 2005.

[75]   Jianhong Zhao and Peter W Carr. Comparison of the retention characteristics of aromatic and aliphatic reversed phases for hplc using linear solvation energy relationships. *Analytical chemistry*, 70(17):3619–3628, 1998.

[76]   NE Zhou, CT Mant, and RS Hodges. Effect of preferred binding domains on peptide retention behavior in reversed-phase chromatography: amphipathic alpha-helices. *Peptide research*, 3(1):8–20, 1990.

[77]   Chrysostomi Zisi, Ioannis Sampsonidis, Stella Fasoula, Konstantinos Papachristos, Michael Witting, Helen G Gika, Panagiotis Nikitas, and Adriani Pappa-Louisi. Qsrr modeling for metabolite standards analyzed by two different chromatographic columns using multiple linear regression. *Metabolites*, 7(1):7, 2017.

Introduction to Publication VII
Low bleed mass spectrometry compatible stationary phase
Minimizing noise and continuous ion suppression

Background of Publication VII

Publication VII put forward a platform technology for preparation of stable functionalized silica based stationary phases for liquid chromatography utilizing a thiol-ene click reaction to graft surface-crosslinked poly(3-mercaptopropyl)methylsiloxane-coatings and ligands onto vinyl-modified silica particles. The main objective was to devise a surface chemistry that, under conventional chromatographic conditions, exhibits minimal ligand bleeding interfering with subsequent coupled detection technologies. The feasibility of the employed bonding chemistry to attain this objective was exemplified on stationary phases with enantioselective tert-butylcarbamoylquinine (tBuCQN) surface ligand chemistry[38]

Low ligand bleed stationary phases are beneficial for the most common detection technologies applied at present in liquid chromatography, namely UV-VIS, fluorescence, charged aerosol and mass spectrometric detection [24, 2, 45]

UV-VIS detection and fluorescence are the least affected by ligand bleeding when referring to alkylsilane surface chemistry, the most common being of C18-type as usually no interference of absorption wavelengths of ligands and analytes are encountered. Yet, with surface chemistries that provide additional binding mechanisms, especially offering π-interaction by plain or substituted phenylalkyl surface chemistries, or more complex ligands that contain molecular moieties of interfering absorbance either by functional requirements or simply by necessity of the synthetic route building the ligand scaffold, raised baseline signal levels can become of concern. If actually disturbing UV-VIS detection, signal interference is additive, i.e. noticeable by increased baseline or background noise. Albeit and still, the relative insensitivity of optical spectroscopy based coupled detection technologies towards ligand bleeding manifested itself in these technologies not expediting research interest in the field of stable stationary phases, especially what isocratic elution is concerned. For gradient elution the situation may be different and the problem also

significant manifested by baseline drifts. Consequently, with the increasing dominance of mass spectrometry as major detection technology, in liquid chromatographic terminology "low bleed stationary phase" is regarded synonymical for "mass spectrometry compatible"

Charged Aerosol detection (CAD) and its predecessor technology evaporative light scattering detection (ELSD) in contrast are highly susceptible towards column bleeding and require stable stationary phases [45]. By nature of both detection processes, i.e. evaporation to a solid aerosol cloud and detection by lateral light scattering (ELSD) or electric aerosol detection (CAD) and due to the lack of an analytical discriminatory dimension of this technologies [33, 11] other than the required nonvolatility of responsive analytes, essentially signals of all compounds regardless of their constitution or chemical nature are affected. Lower limit of detection (LLOD), lower limit of quantitation (LLOQ) are severely affected by instable stationary phases. Conversely CAD is a valuable method for investigation of column bleeding.

Mass spectrometry, representing the most selective of the aforementioned detection technologies by nature of its signal generation (ionization and specific fragmentation) and general signal resolution (mass peak broadness) especially with high resolution mass analyzers time of flight systems (TOF, QTOF, Orbitrap) is affected by column bleeding in a more specific way. With the rare exception of analytes of interest as well as lysed ligands being isobaric/possessing identical MS^n transitions, stationary phase ligand bleeding affects mass spectrometry not by overlaying the desired signals by increased baseline and noise, but rather by quelling the analyte signal in an undesired process termed ion suppression [24].

As the authors contribution to publication VII was the mass spectrometric examination of the reduced ligand bleeding of the new presented stationary phase platform, the following paragraphs highlight some of the aspects of ion suppression and stationary phase stability in liquid chromatography coupled mass spectrometry.

Ion Suppression in liquid chromatography coupled mass spectrometry - general aspects

In general, the term ion suppression when referring to liquid chromatography coupled mass spectrometry describes a deleterious effect of compounds coeluting to the analytes of interest, hence the absolute analyte signal is governed by more factors than analyte concentration itself. The signal of interest is reduced or even subject to complete extinction [2]. The opposite effect of elevated signal levels, named ion enhancement, although less frequently encountered, can also be observed [19, 16]. Even though the latter effect can be exploited beneficially e.g. by usage of a dopant additive in the eluent (e.g. NH_4F in LC-MS of estradiol [15]), generally any kind of distortion of ion signals is undesired and aim of respective countermeasures. The undesirability of any signal alteration is especially aggravated for untargeted comprehensive mass spectrometric workflows (e.g. of SWATH-type [1]) as it is not ascertainable nor ensurable by the very nature of an untargeted approach that signal enhancement or suppression affects all analytes in an uniform manner, thereby introducing problematic signal bias for subsequent statistical evaluation procedures of such workflows [22, 48].

Worse and in opposite to targeted mass spectrometric approaches where signal deviations are usually detected by use of internal standards, or to ion suppression events caused by singular causes revealing themselves during appropriate statistical treatment enabling removal of such outlier results, signal alteration in untargeted approaches may remain unnoticed in spite of method validation and QC efforts, especially for low abundant analytes [39]. In consequence, ion suppression is a main concern regarding the fidelity of untargeted LC-MS^n approaches [17] and measures reducing or even eliminating know sources of ion suppression, like ion suppression caused by constant elution of lysed column ligand as result of instable ligand-scaffold binding chemistry, are of paramount importance [46, 10, 23]

As described in the following paragraphs (*vide infra*), such a signal bias is essentially to be expected when operating chromatography in gradient mode, thereby altering ionization

conditions by means of the changing eluent composition itself, or when complete compound separation is expected to be incomplete, as, for example, in samples containing biological matrix and therefore a plethora of different compounds.

As aforementioned, ion suppression can be caused by coeluting compounds causing peak shaped regions/intervals of certain extent of ion suppression in the chromatogram (momentary ion suppression) or by compounds eluting and affecting ionization permanently (continuous ion suppression) either as part of the mobile phase, contaminants captured by and eluted from various parts of the LC-MS-system (fluidic path, column, ion source) or originating from stationary phases susceptible to degradation, especially by lysis of the stationary phase surface modification, e.g. for a common C18-type column the C18-ligand or alkyl endcapping moieties of lower molecular weight. Combination of both sources of ion suppression are also encountered, depending on the sample matrix, eluent system, stationary phase type and especially in gradient mode [[2, 19, 16].

To eliminate one source of ion suppression, stationary phases employed in liquid chromatography coupled to mass spectrometry should exert minimal lysis and elution of ligand. i.e. a bonding chemistry providing excellent chemical stability under the chromatographic conditions (eluent compositions, column operating temperatures, sample composition). Stationary phases featuring this property are labeled as mass spectrometry compatible. Regarding the influence of the surface chemistry, besides the minimized chronic ion suppression, also chromatographic performance parameters like retention capacity, retention times, peak shape and asymmetry are preserved for longer time by low bleed columns, thereby reducing data distortion by column aging-induced shifts of elution parameters in untargeted approaches.

Mechanisms of ion suppression

The predominant ionization method employed in liquid chromatography coupled mass spectrometry especially in the field of analysis of biological samples is the electrospray ionization (ESI) [25]. Besides this major ionization technique, atmospheric pressure chemical ionization (APCI) and atmospheric pressure photoionization are employed if examined analytes are less compatible with the ESI process (s. general introduction for reference). All described techniques share in common charging and transfer to the gas phase of analytes,

yet with ESI charging occurs in the primary step, opposed to APCI/APPI gas phase transfer is the initial process. Furthermore, as described in the general introduction of this thesis, for ESI two models describe the transition of charged analytes to the gas phase, the ion evaporation model (IEM) for small molecules, and the charged residue model (CRM) for large molecules, e.g. large peptides, proteins or polymers. This crystalizes four vantage points for ion suppression processes [3]: I) competitive inhibition of charging of the analyte by matrix solutes, II) gas phase ion charge neutralization, inhibition of gas phase transition by III) coprecipitation or IV) by deferral of efficient electrospray ionization via interference with the iterative droplet-fission process or with the ion evaporation processes by alteration of viscosity and surface tension of the generated droplets.

Charge competition (I) between different solutes during the primary ESI-droplet formation was described first by Kerbale et.al [25], stating a maximum amount of 10^{-5} charged ions in one primary ESI-droplet.

Ion suppression by gas phase charge neutralization (II), foremost by proton transfer reaction were observed by Chin et al [7] and Buhrman et al [6].

Evidence for precipitation hypothesis (III) was provided by King et al[27] by collecting residues generated under ion suppressing conditions from the mass spectrometers orifice spray shield surface and reanalyzing these under non suppressing conditions, revealing that a fraction several times the detected analyte amount is excluded from the ESI process and deposited on the orifice surface. King et al. also confirmed that ion suppression by coprecipitation is, as expected, facilitated by sulfates or phosphates.

Interference with the ion generation process (IV) has been reported regarding analyte vs matrix competition for surface access necessary for IEM [9, 44]

To conclude from the enumerated vantage points for ion suppression described by the literature cited above, the only copious remedy for ion suppression is the removal of matrix components, adequate chromatographic separation where the latter is not achievable and, as in Publication VII, minimizing of any further introduction of extraneous compounds like surface ligands form chromatographic stationary phases.

Momentary Ion Suppression

Acute ion suppression evokes from compounds coeluting with a chromatographic peak profile simultaneously with the respective analytes of interest. A plethora of ion suppressing compounds have been identified [3, 16], e.g. detergents, ion pairing reagents [4, 13, 20] , inorganic anions or other buffer compounds [27], polymers [35] and compounds facilitating proton exchange [34]. By their very nature being complex mixtures of a large variety of compounds that are often charged as essential property of their metabolic function or for containment inside the cell membrane [49], raw biological samples generally give rise to chromatograms with extended regions of ion suppression and appropriate assessment thereof and countermeasures against during method development must be applied [24].

The source of acute ion suppression usually lies within the sample matrix compounds themselves. Yet other origins are known. One example is accumulation of eluent or additive impurities [26] on the column head during weak elution conditions and separation during the following gradient, essentially mimicking an injection. This is aggravated by non RP-types of chromatography, like HILIC or IEC chromatography, as strong ion suppressing ionic compounds are particularly enriched on the column during initial weak eluting conditions [21]. Furthermore, introduction of ion suppression by sample preparation workflows or inappropriate materials in contact with samples matrix must be avoided.

As stated above, with targeted approaches acute ion suppression can be detected, in some cases evaded by apt measures during method development or to some extent corrected for. Consequently, examination of matrix/ion suppression effects during bioanalytical method development and validation is advised for the research laboratory [2] and required by regulatory bodies [14] for clinical, monitoring, forensic etc. applications. On the contrary, with untargeted comprehensive mass spectrometric approaches, the actual presence of ion suppression remain unnoticed.

Continuous Ion Suppression

In opposite to momentary ion suppression, chronic ion suppression is the result of ion suppressing compounds permanently being introduced to the ESI source. Potential sources are unretained eluent constituents/contaminants, contaminations of the fluidic path of the chromatographic system or column perpetually desorbed by the eluent stream, or, more

inherently, ligands eluting from the column as result of lysis reactions of the stationary phase surface modification, mainly by lysis of the Si-O-Si-ligand bond. Albeit being introduced permanently and uniformly into the ion source, the extent of ion suppression is dependent on respective analyte species and can be modified by eluent composition in case of gradient elution mode. Lower limit of detection (LLOD), lower limit of quantification (LLOQ), method sensitivity are affected always, and additionally and linear range can be reduced by chronic ion suppression [3].

Countermeasures towards ion suppression - Objective of Publication VII

Momentary ion suppression can be reduced by appropriate sample preparation reducing matrix compounds with adverse effect. Usually this can be achieved by solvent or solid extraction protocols, the latter also applicable as online-SPE performed by the chromatographic system. If not possible to remove ion suppression factors, chromatographic separation thereof from the analytes of interest should be achieved by tuning the chromatographic parameters like column selectivity, temperature, eluent composition, gradient shape, or by two dimensional liquid chromatography (2DLC).

Chronic ion suppression can be counteracted by ensuring adequate chromatographic eluent quality, integration of cleaning procedures in chromatographic methods and sample sequences, preemptive ion suppression assessment and preemptive cleaning procedures.

Fundamentally, as focused in Publication VII, stable stationary phases with minimized column bleeding allowing for labeling as LC-MS compatible are a necessity in order to minimize chronic ion suppression.

Chemical instability of silica based columns – column bleeding

Main contributors to degradation of silica particle based stationary phases are temperature and foremost pH. While high pH (>8) eluent compositions are known to rapidly degrade the silica support material hydrolytically, effectively altering the column bed in the sense of critical chromatographic properties of the internal eluent pathways (altering A and C terms of van-Deemter equation when using isocratic conditions for assessment) to an extent of physical collapse of stationary phase particles and column bed as whole [28], acidic conditions decompose porous silica particle based stationary phases more gradually yet

permanently by catalyzing hydrolysis of the Si-O-Si siloxane bonds of the surface ligand [28] while leaving the silica support intact during the usual operational life cycle (few thousand injections) of chromatographic columns [47, 29]. Ligand lysis has been described extensively by Kirkland et al [28, 31, 30, 18].

Ligand siloxane bond lysis results in a continuous deterioration of chromatographic performance, observable by decline of chromatographic retention factors k and peak resolution and symmetry, hence reduced peak capacity of the column. Unfortunately, as observed by Kirkland [29] the loss of chromatographic performance can be described by an exponential decay law, i.e. the adverse interanalysis effects are most strongly observed during the initial phase of chromatographic column lifetime and impact on chromatographic results in between analytical runs and sequences/batches of runs is particularly large during the initial decay phase. This property of instable silica based stationary phases is especially undesirable for untargeted mass spectrometric applications, as bias between samples is introduced by chromatographic performance and simultaneous ion suppression decrease.

Stable surface chemistry for LC-MS applications

Generally, in order to provide for stable stationary phases for liquid chromatography, two major approaches have been pursued in the past: use of alternate scaffold materials, based on inorganic metal oxides e.g. alumina, zirconia or titanium based stationary phases and as second approach, development of stable surface chemistry for silica based stationary phases [8]. The first approach has been studied intensively and yielded stationary phase scaffolds with overall pH stability (0-14) exceeding those of silica materials [36, 8, 12]. Albeit, as result of a less stable M-O-Si bond (M= Zr, Al) of alternate metal-oxide scaffolds [40] , the general superior stability of the scaffold material relative to silica based materials does not apply to surface ligand modification stability, actually the contrary is reported [37, 8]. In order to achieve enhanced ligand stability, polymer coating or cladding techniques have been applied [42]. Yet, alternate metal-oxide based materials suffer from reduced achievable theoretical plates due to reduced specific surface area [5] or reduced chromatographic efficiency as result of restricted mass transfer [42].

A major disadvantage described for the alternate scaffold materials results from strong Lewis base and acid activity effectively providing stationary phase scaffolds with inherent ion exchange (IEX) capabilities, in contrast to an ideal stationary phase material whose chromatographic selectivity would solely be governed by surface modification, as a matter of speaking possessing chromatographic inertness. While the inherent IEX mechanism can be used beneficially for particular analytical questions [38], generally it must be masked by mobile phase additives when RP retention mechanism are required [42]. These additives, necessarily ionic by mode of action, e.g. phosphates, are incompatible with mass spectrometric detection and with surface ligands that rely on specific ionic interactions adjusted by mobile phase ionic activity as the quinine carbamates [38] employed in Publication VII.

To summarize, up to the present day, alternate scaffold materials provide no substitute for mass spectrometric compatible silica based stationary phases, an observation the majority of chromatography aimed research literature in this field [8] and availability of commercially stationary phases and columns for LC-MS applications stands testimony to.

As silica still provides the principal scaffold for stationary phases in liquid chromatography [8] the second general approach aims to enhance ligand stability by application of surface ligand chemistries that circumvent or reduce the susceptibility of the Si-O-Si-ligand bond to hydrolysis under chromatographic conditions. A straightforward method hereby is, to utilize organo-silanes with bulky moieties bonded directly to the silicone atom in the ligand in order to sterically hinder the hydrolysis, while still resembling a brush type surface modification [29].
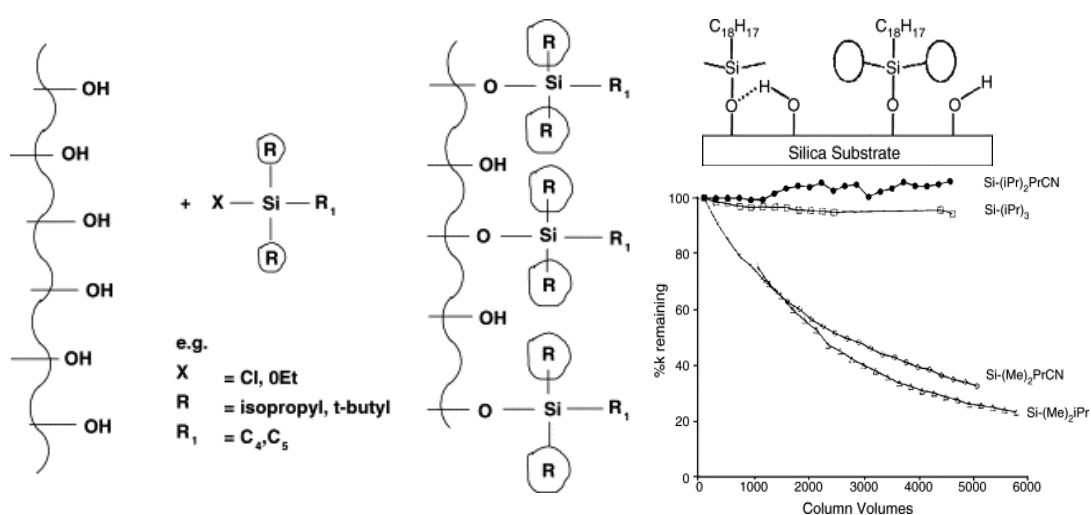
Figure FP7 1 Steric protection as simple and effective approach to protect the labile Si-O-Si bond in silica based stationary phases Left: Schematic of silica surface and employed ligands, Right top: Steric protection from hydrogen bonding of residual unmodified silanols. Right bottom: phase stability expressed via perpetuation of k of 1-phenylheptane vs column volumes of eluent flow applied to column.   Reproduced from Review on the chemical and thermal stability of stationary phases for reversed-phase liquid chromatography Journal of Chromatography A Volume 1060, Issues 1–2, 10 December 2004, Pages 23-41 Copyright (2004), with permission from Elsevier

The stability enhancement observed by chromatographic assessment, e.g. monitoring of evolution of retention factors of appropriate indicator analytes [29] could be confirmed by the orthogonal methodology of CP/MAS ^{29}Si-NMR by Scholten et al. [43]. According to the findings of Scholten et al. steric protection by suitable groups at the ligand Si shield the Si-O-Si hydrogen bonding by vicinal unmodified silanol groups, as depicted in Fig. FP7.1. which supports the assumption of effective steric protection against acidic eluents.

A further variant of the steric protection concept is the use of bidendate stationary phases. Instead of monomeric brush type ligand, two ligands are interconnected via Si-Q-Si bridge, where Q is an ethyl or propionyl chain.
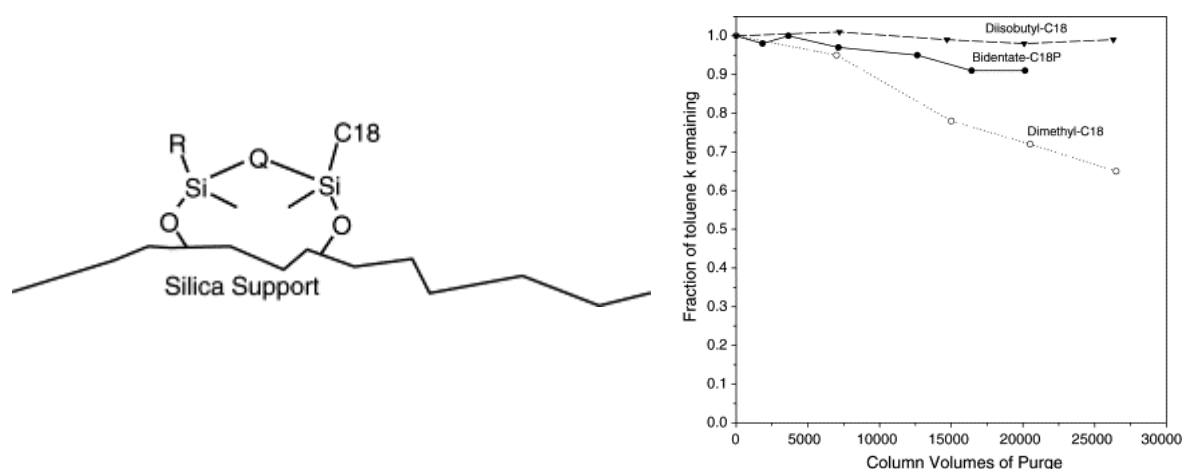
Figure FP7 2 Bidendate stationary phase. R= steric protection group or second ligand   Reproduced from Review on the chemical and thermal stability of stationary phases for reversed-phase liquid chromatography Journal of Chromatography A Volume 1060, Issues 1–2, 10 December 2004, Pages 23-41 Copyright (2004), with permission from Elsevier

An advancement of the bidendate concept for futher improvement of column stability could be achieved by combining mechanical resilience of silica based materials with chemical inertness by immobilization of polymer coatings on the material surface, amongst others mainly polysiloxane coatings [41, 8]. Usually, these coatings represent crosslinking of the brush type ligands by aforementioned polymers, e.g. polymethylsiloxane in case of

Publication VII. The crosslinked immobilized linkers employ the same bonding chemistry as the brush type ligands e.g. Si-substituted alkylchlorosilanes incorporating also the steric protection of the Si-O-Si-Bond as in the stable brush type ligands, see Fig FP7.2 for reference. These types of stable stationary phases were assessed to have excellent stability in neutral conditions with up to 50000 column volumes and enhanced stability in alkaline pH, indicating a protecting effect for the silica scaffold that is usually the cause of column failure under these conditions rather than ligand hydrolysis.
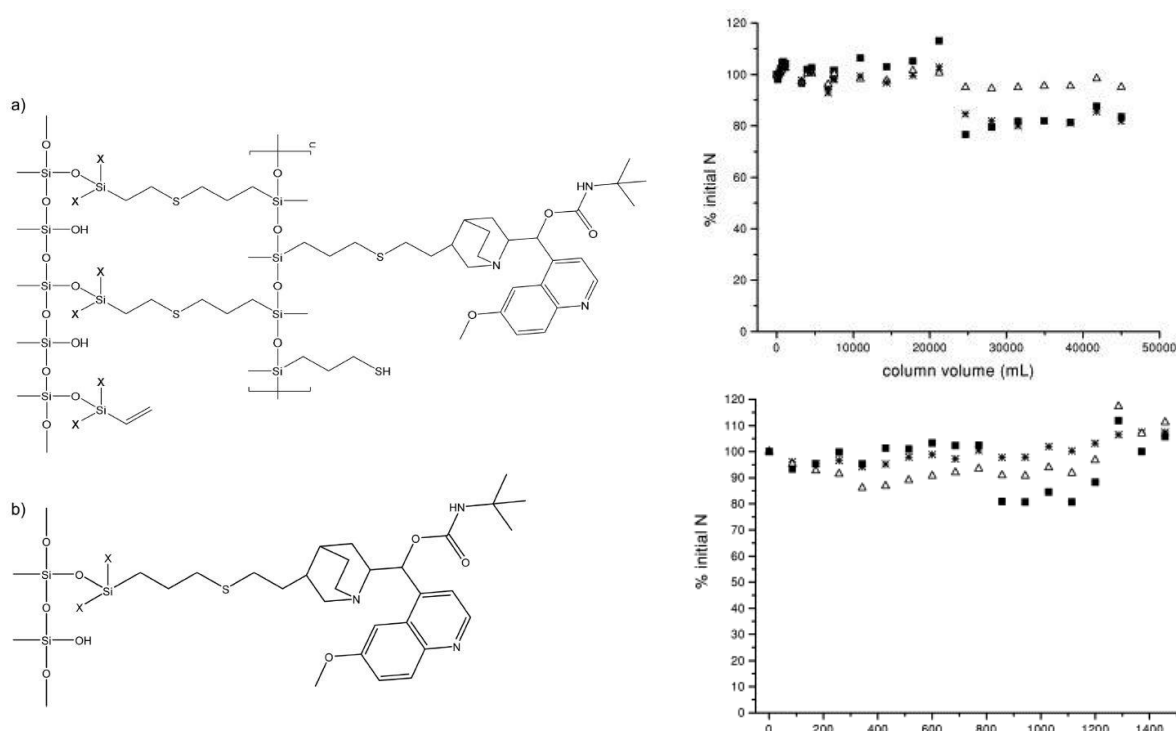
Figure FP7 3 Left a: Cross linking via polymethylsiloxane; b: sterically protected brush type ligand. Reproduced from Publication VII, Copyright (2016), with permission from Elsevier Right: Stability assessment in neutral (top) and alkaline (bottom) eluent for different immbilisation and posttreatments of polymer coated columns. see reference for details. Reproduced from High-performance liquid chromatographic stationary phases based on poly(methyloctylsiloxane) immobilized on silica: III. Stability evaluations, Journal of Chromatography A Volume 987, Issues 1–2, 14 February 2003, Pages 93-101 Copyright(2003) with permission from Elsevier.

Publication VII contributed to stable stationary phase research in advancing the polymethylsiloxane polymer coated type of column by providing an easy modifiable platform bearing as polymer coating poly(3-mercaptopropyl)methylsiloxane that could easily be modified with ligand via a thiol-ene click reaction. For demonstration of concept, enantioselective stationary phases based on quinine carbamates [38] were synthesized and examined regarding column stability and mass spectrometric compatibility.

Methods for evaluating phase stability and MS compatibility

Stationary phase stability can be assessed either indirect by functional assessment via an appropriate chromatographic test case, usually probing retention factors (k), peak resolution (R) and asymmetry, theoretical plates (N) or plate height (H), peak capacity etc. of amenable analytes or direct i.e. by monitoring of the surface ligand lysis, usually by charged aerosol detection or mass spectrometry [45, 8]. The first route of assessment is independent of coupled detection technology and gives a straightforward estimation of maximum achievable useful column lifetime under given eluent composition (pH, salt concentration) and column operation temperature. However, by only evaluating chromatographic readouts, no conclusions can be drawn regarding compatibility with particular detection technologies, i.e. mass spectrometry or charged aerosol detectors, that can readily be affected by ligand concentrations in the eluate generated by ligand bleeding from a new manufactured column, as small amounts of ligand lysis products can exert strong influence on signal generation e.g. ion suppression [45, 2, 24].

The assessment of detector technology compatibility e.g. mass spectrometry compatibility of a stationary phase is therefore best done by the direct approach. However, it should be noted, that chromatographic lifetime of the column as revealed by the functional assessment of chromatographic parameters might not be derivable in a straightforward manner from the direct measurement approach.

Publication VII employed two general test methods, first direct monitoring of ligand bleed intensity by MRM and TIC monitoring and second comparison of LLOQ for selected compounds on brush type vs. immobilized polymethylsiloxane coated stationary phase. The examination of LLOQ improvements gives direct insight to stationary phase LC-MS compatibility.

Results of Publication IV

Publication IV could provide a platform technology for the preparation of functionalized silica modified by immobilized poly(3-mercaptopropyl) methylsiloxane instead of plain polymethylsiloxane. This scaffold can easily be functionalized to various column selectivities by thiol-ene click reaction with respective alkene mojety bearing ligands. Demonstration of the basic principle was performed with the preparation of enantioselective tert-butylcarbamoylquinine carbamates based stationary phases [38]. The ligand stability was assessed by high resolution time of flight mass spectrometry (HRMS QTOF). Relative to brush type ligands, the background intensity measured by total ion current (TIC) could be lowered by 70-80%, indicating significant reduction of column bleeding and increase of mass spectrometric compatibility. This finding was confirmed by enhanced sensitivity of the polymer coated columns relative to the brush type variant with comparable chromatographic performance, as LOD as well as LLOQ were also decreased by 70%

Abbreviations

IEM: Ion evaporation model

IEX: Ion exchange chromatography

CP/MAS ^{29}Si-NMR: cross polarization/magic angle spinning nuclear magnetic resonance spectrometry

References

[1]   Genna L Andrews, Brigitte L Simons, J Bryce Young, Adam M Hawkridge, and David C Muddiman. Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (tripletof 5600). *Analytical chemistry*, 83(13):5442–5446, 2011.

[2]   Thomas M Annesley. Ion suppression in mass spectrometry. *Clinical chemistry*, 49(7):1041–1044, 2003.

[3]   Jean-Philippe Antignac, Katia de Wasch, Fabrice Monteau, Hubert De Brabander, François Andre, and Bruno Le Bizec. The ion suppression phenomenon in liquid chromatography–mass spectrometry and its consequences in the field of residue analysis. *Analytica Chimica Acta*, 529(1-2):129–136, 2005.

[4]   Alex Apffel, Steven Fischer, Gerson Goldberg, Paul C Goodley, and Frank E Kuhlmann. Enhanced sensitivity for peptide mapping with electrospray liquid chromatography-mass spectrometry in the presence of signal suppression due to trifluoroacetic acid-containing mobile phases. *Journal of chromatography A*, 712(1):177–190, 1995.

[5]   U Bien-Vogelsang, A Deege, H Figge, J Köhler, and G Schomburg. Syntheses of stationary phases for reversed-phase lc using silanization and polymer coating. *Chromatographia*, 19(1):170–179, 1984.

[6]   Deborah L Buhrman, Piper I Price, and Patrick J Rudewicz. Quantitation of sr 27417 in human plasma using electrospray liquid chromatography-tandem mass spectrometry: a study of ion suppression. *Journal of the American Society for Mass Spectrometry*, 7(11):1099–1105, 1996.

[7]   Christina Chin, ZP Zhang, and HT Karnes. A study of matrix effects on an lc/ms/ms assay for olanzapine and desmethyl olanzapine. *Journal of pharmaceutical and biomedical analysis*, 35(5):1149–1167, 2004.

[8]   HAI Claessens and MA Van Straten. Review on the chemical and thermal stability of stationary phases for reversed-phase liquid chromatography. *Journal of Chromatography A*, 1060(1-2):23–41, 2004.

[9]   Terri L Constantopoulos, George S Jackson, and Christie G Enke. Effects of salt concentration on analyte response using electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 10(7):625–634, 1999.

[10]   Katja Dettmer, Pavel A Aronov, and Bruce D Hammock. Mass spectrometry-based metabolomics. *Mass spectrometry reviews*, 26(1):51–78, 2007.

[11]   Roy W Dixon and Dominic S Peterson. Development and testing of a detection method for liquid chromatography based on aerosol charging. *Analytical chemistry*, 74(13):2930–2937, 2002.

[12]   Christopher J Dunlap, Peter W Carr, Clayton V McNeff, and Dwight Stoll. Peer reviewed: Zirconia stationary phases for extreme separations, 2001.

[13]   Jamshid Eshraghi and Swapan K Chowdhury. Factors affecting electrospray ionization of effluents containing trifluoroacetic acid for high-performance liquid chromatography/mass spectrometry. *Analytical chemistry*, 65(23):3528–3533, 1993.

[14]   Food FDA et al. Guidance for industry: bioanalytical method validation. *http://www. fda. gov/cder/Guidance/4252fnl. pdf*, 2001.

[15]   Tom Fiers, Bruno Casetta, Brigitte Bernaert, Eric Vandersypt, Martine Debock, and Jean-Marc Kaufman. Development of a highly sensitive method for the quantification of estrone and estradiol in serum by liquid chromatography tandem mass spectrometry without derivatization. *Journal of Chromatography B*, 893:57–62, 2012.

[16]   Ambrose Furey, Merisa Moriarty, Vaishali Bane, Brian Kinsella, and Mary Lehane. Ion suppression; a critical review on causes, evaluation, prevention and applications. *Talanta*, 115:104–122, 2013.

[17]   Helen G Gika, Euan Macpherson, Georgios A Theodoridis, and Ian D Wilson. Evaluation of the repeatability of ultra-performance liquid chromatography–tof-ms for global metabolic profiling of human urine samples. *Journal of Chromatography B*, 871(2):299–305, 2008.

[18]   JL Glajch, JJ Kirkland, and J Köhler. Effect of column degradation on the reversed-phase high-performance liquid chromatographic separation of peptides and proteins. *Journal of Chromatography A*, 384:81–90, 1987.

[19]   Fabio Gosetti, Eleonora Mazzucco, Davide Zampieri, and Maria Carla Gennaro. Signal suppression/enhancement in high-performance liquid chromatography tandem mass spectrometry. *Journal of Chromatography A*, 1217(25):3929–3937, 2010.

[20]   Sven Åke Gustavsson, Jenny Samskog, Karin E Markides, and Bengt Långström. Studies of signal suppression in liquid chromatography–electrospray ionization mass spectrometry using volatile ion-pairing reagents. *Journal of Chromatography A*, 937(1-2):41–47, 2001.

[21]   Yunsheng Hsieh. Potential of hilic-ms in quantitative bioanalysis of drugs and drug metabolites. *Journal of separation science*, 31(9):1481–1491, 2008.

[22]   Qiang Huang, Lu Yang, Ji Luo, Lin Guo, Zhiyuan Wang, Xiangyun Yang, Wenhai Jin, Yanshan Fang, Juanying Ye, Bing Shan, et al. Swath enables precise label-free quantification on proteome scale. *Proteomics*, 15(7):1215–1223, 2015.

[23]   Yusuke Iwasaki, Takahiro Sawada, Kentaro Hatayama, Akihito Ohyagi, Yuri Tsukuda, Kyohei Namekawa, Rie Ito, Koichi Saito, and Hiroyuki Nakazawa. Separation technique for the determination of highly polar metabolites in biological samples. *Metabolites*, 2(3):496–515, 2012.

[24]   Lori Lee Jessome and Dietrich A Volmer. Ion suppression: a major concern in mass spectrometry. *Lc Gc North America*, 24(5):498, 2006.

[25]   Paul Kebarle and Liang Tang. From ions in solution to ions in the gas phase-the mechanism of electrospray mass spectrometry. *Analytical chemistry*, 65(22):972A–986A, 1993.

[26]   Bernd O Keller, Jie Sui, Alex B Young, and Randy M Whittal. Interferences and contaminants encountered in modern mass spectrometry. *Analytica chimica acta*, 627(1):71–81, 2008.

[27]   Richard King, Ryan Bonfiglio, Carmen Fernandez-Metzler, Cynthia Miller-Stein, and Timothy Olah. Mechanistic investigation of ionization suppression in electrospray ionization. *Journal of the American Society for Mass Spectrometry*, 11(11):942–950, 2000.

[28]   James J Kirkland, MA Van Straten, and HA Claessens. High ph mobile phase effects on silica-based reversed-phase high-performance liquid chromatographic columns. *Journal of Chromatography A*, 691(1-2):3–19, 1995.

[29]   JJ Kirkland, JL Glajch, and RD Farlee. Synthesis and characterization of highly stable bonded phases for high-performance liquid chromatography column packings. *Analytical Chemistry*, 61(1):2–11, 1989.

[30]   J Köhler, DB Chase, RD Farlee, AJ Vega, and JJ Kirkland. Comprehensive characterization of some silica-based stationary phase for high-performance liquid chromatography. *Journal of Chromatography A*, 352:275–305, 1986.

[31]   J Köhler and JJ Kirkland. Improved silica-based column packings for high-performance liquid chromatography. *Journal of Chromatography A*, 385:125–150, 1987.

[32]   Michael Lämmerhofer, Pilar Franco, and Wolfgang Lindner. Quinine carbamate chiral stationary phases: Systematic optimization of steric selector-selectand binding increments and enantioselectivity by quantitative structure-enantioselectivity relationship studies. *Journal of separation science*, 29(10):1486–1496, 2006.

[33]   Benjamin YH Liu and David YH Pui. On the performance of the electrical aerosol analyzer. *Journal of Aerosol Science*, 6(3-4):249–264, 1975.

[34]   Claude R Mallet, Ziling Lu, and Jeff R Mazzeo. A study of ion suppression effects in electrospray ionization from mobile phase additives and solid-phase extracts. *Rapid Communications in Mass Spectrometry*, 18(1):49–58, 2004.

[35]   Hong Mei, Yunsheng Hsieh, Cymbylene Nardo, Xiaoying Xu, Shiyong Wang, Kwokei Ng, and Walter A Korfmacher. Investigation of matrix effects in bioanalytical high-performance liquid chromatography/tandem mass spectrometric assays: application to drug discovery. *Rapid Communications in Mass Spectrometry*, 17(1):97–103, 2003.

[36]   J Nawrocki, C Dunlap, A McCormick, and PW Carr. Part i. chromatography using ultra-stable metal oxide-based stationary phases for hplc. *Journal of chromatography A*, 1028(1):1–30, 2004.

[37]   J Nawrocki, M Rigney, A McCormick, and PW Carr. Chemistry of zirconia and its use in chromatography. *Journal of Chromatography A*, 657(2):229–282, 1993.

[38]   Jacek Nawrocki, Christopher J Dunlap, Peter W Carr, and John A Blackwell. New materials for biotechnology: Chromatographic stationary phases based on zirconia. *Biotechnology progress*, 10(6):561–573, 1994.

[39]   Shama Naz, Maria Vallejo, Antonia Garcá, and Coral Barbas. Method validation strategies involved in non-targeted metabolomics. *Journal of Chromatography A*, 1353:99–105, 2014.

[40]   Walter Noll. *Chemistry and technology of silicones*. Elsevier, 2012.

[41]   M Petro and D Berek. Polymers immobilized on silica gels as stationary phases for liquid chromatography. *Chromatographia*, 37(9-10):549–561, 1993.

[42]   MP Rigney, TP Weber, and PW Carr. Preparation and evaluation of a polymer-coated zirconia reversed-phase chromatographic support. *Journal of Chromatography A*, 484:273–291, 1989.

[43]   Alex B Scholten, Jan W De Haan, Henk A Claessens, Leo JM Van de Ven, and Carel A Cramers. 29-silicon nmr evidence for the improved chromatographic siloxane bond stability of bulky alkylsilane ligands on a silica surface. *Journal of Chromatography A*, 688(1-2):25–29, 1994.

[44]   Markus Stüber and Thorsten Reemtsma. Evaluation of three calibration methods to compensate matrix effects in environmental analysis with lc-esi-ms. *Analytical and Bioanalytical Chemistry*, 378(4):910–916, 2004.

[45]   T Teutenberg, J Tuerk, M Holzhauser, and TK Kiffmeyer. Evaluation of column bleed by using an ultraviolet and a charged aerosol detector coupled to a high-temperature liquid chromatographic system. *Journal of Chromatography A*, 1119(1-2):197–201, 2006.

[46]   Georgios A Theodoridis, Helen G Gika, Elizabeth J Want, and Ian D Wilson. Liquid chromatography–mass spectrometry based global metabolite profiling: a review. *Analytica chimica acta*, 711:7–16, 2012.

[47]   Klaus K Unger. *Porous silica*, volume 16. Elsevier, 1979.

[48]   Nawaporn Vinayavekhin and Alan Saghatelian. Untargeted metabolomics. *Current protocols in molecular biology*, 90(1):30–1, 2010.

[49]   Frank H Westheimer. Why nature chose phosphates. *Science,* 235(4793):1173–1178, 1987.

Part 2 - **Publication V**

Introduction to Publication V / Conference Publication I:
Shape selective chromatography coupled to atmospheric pressure chemical ionization mass spectrometry for the analysis of instable triterpenoid ester regioisomers

Background of Publication V / Conference Publication I

Lipophilic extracts and preparations thereof from *Asteraceae Calendula officinalis are* applied for centuries [31] as well as worldwide [32], due to their assumed anti-inflammatory and wound healing properties. Although still under discussion [5, 27] it is considered, that the anti-inflammatory properties arise from 20-taraxastene derivatives, especially monols, diols, and triols and fatty acid esters thereof [19, 58] foremost of the extracts' major triterpenoid compound faradiol [10, 57, 34].

Publication V originated in cooperation with and from works of Nicolaus et al. studying the properties of triterpene alcohols and esters [35, 37, 36] with the group of Prof. I. Merfort, Institute of Pharmaceutical Sciences, Pharmaceutical Biology and Biotechnology, University of Freiburg. During these works, elucidating presence and the intact structure of potential fatty acid diesters was targeted. Albeit their general presence reported earlier [56] exact intact structures of presumed diesters were not established and verified previously.

The 20-taraxastene derivatives focused on in Publication V were faradiol, arnidiol, arnitriol, maniladiol and lupane-3β,16β,20-triol.

Due to the complex composition of lipophilic extracts of Calendula officinalis, structural elucidation has to be performed by employing a concerted multiplicity of analytical methods like NMR, MALDI-MS, GC-EI-MS and APCI- and ESI-LC-MS/MS [35] i.e. complementary techniques that in summa compensate sufficiently for the lack of a single preferable analytical method for each compound constituting the extract.

The author's contribution to Publication V comprised the development of suitable mass spectrometric and liquid chromatography coupled mass spectrometric methods and procedures to complement the NMR based findings of Nicolaus et al.

As described in the following paragraphs, instability of compounds of interest under the respective analytical technique conditions, ambiguity of their decay products and

regioisomerism impairing standard chromatographic methods had to be addressed by adequate methodological countermeasures and preparation of customized chromatographic materials and columns.

In source decay instability and implications on chromatographic requirements

A major difficulty in mass spectrometric detection of regioisomers of the faradiol esters and diesters examined in Publication V was the instability of generated $[M+H]^+$ pseudomolecular ions. Instability of terpenoids in electrospray ionization (ESI) has been reported earlier by [25] and by Ma et al.[30]. Both reported instability predominantly for terpenoids bearing hydroxyl groups, with the extent of instability depending on the molecular structure. Congruent with the findings of Publication V, Ma et al [30] reported instability in extenso for terpenoid diols.

Whereas the general intricacy is, that all connate species give rise to similar fragmentation patterns in collision induced dissociation tandem mass spectrometry (CID-MS²), this ramification is especially aggravated in case in source decay gives rise to isobaric species, as depicted in Fig FP5.XX:
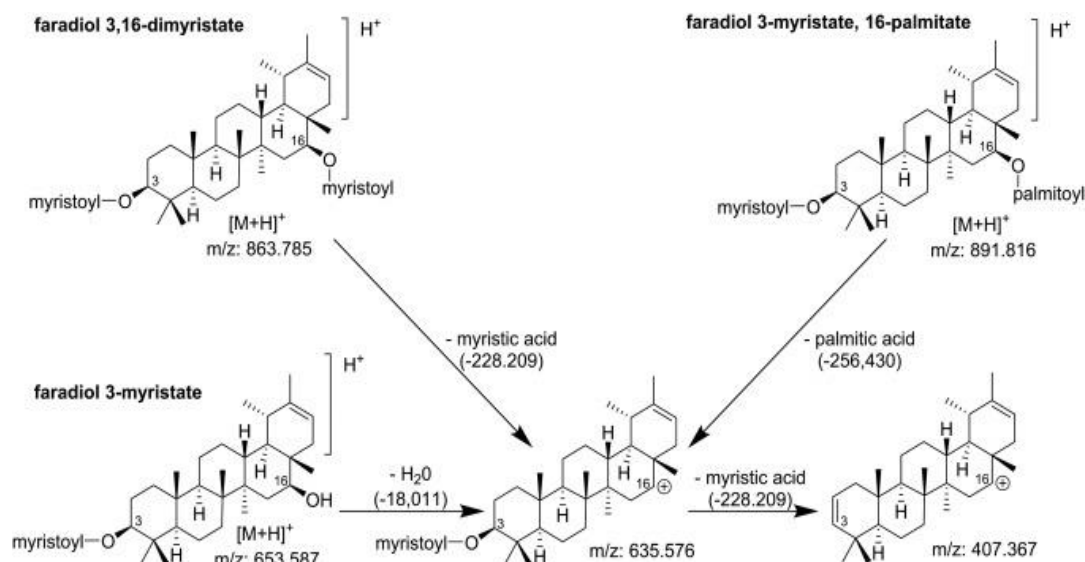
Figure FP5 1 Fragmentation pattern of [M+H]⁺ homogeneous and heterogeneous faradiol esters examined in Publication V. Reproduced from Journal of Pharmaceutical and Biomedical Analysis, 118, 195-205 Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS, Copyright (2016), with permission from Elsevier

For the faradiol esters, it was perceived during experimentation with mono and diester standards, that the 16-hydroxy-group was particularly prone to decay into a $[M+H-H2O]^+$ As depicted in Fig FP5.XX, this leads to a common in source fragment ion with m/z=635.576 for the 3-O-myristoyl derivatives, and for m/z=663.6074 for the 3-O-palmitoyl derivatives. The ambiguity of origin of both respective decay products precluded these ions to be used for confirmation of presence of the diesters by direct infusion experiments without prior chromatographic separation, particularly from a complex extract.

What is referred to as "in source decay" can occur at two locations in common commercial mass spectrometers (ref. Fig FP5.2): In the source itself and in the transition zone between rough vacuum and high vacuum. The decay processes in the transition zone are influenced by rough vacuum pressure and voltages accelerating the ions in the ion beam, and resemble collision induced dissociation (CID). Albeit ion path layout and tuning of voltages is commonly optimized for neutral particles removal and maximum ion transmission in order to provide maximum sensitivity, some manufacturers designs permit the user to intentionally tune specific voltages in order to achieve fragmentation, in particular with single quadrupole mass spectrometers, where no other means of CID are present (e.g. "Fragmentor voltages" between transfer capillaries and skimmer in Agilent mass spectrometers). The processes at this stage are largely unaffected by the source chemistry. Covalent pseudomolecular ions ($[M+nH]^{n+}$ / $[M-nH]^{n-}$) can be fragmented reproducibly and predictably in the same manner as in QqQ-CID, whereas noncovalent, coulomb type clusters like metal cation, acid anion or water adducts ($[M+nX]^{n+/-}$)/ $[M+nX+xH2O]^{n+/-}$) predominantly are dissociated without breaking of covalent bonds, in consequence lowering sensitiivity.

 The Sciex 5600+ TripleTof mass spectrometer used in Publication V employs neither a transfer capillary nor a skimmer at the transition from rough to high vacuum due to sensitivity enhancement reasons (higher transmittance). In source decay occurs besides the ion source in the QJet^{TM}-compartment and in the transition zone right after the orifice assembly. Relevant voltages are ion source floating voltage (ISFV) and declustering potential (DP), as both modulate the ions longitudinal kinetic energy (ref. Fig FP5.2).
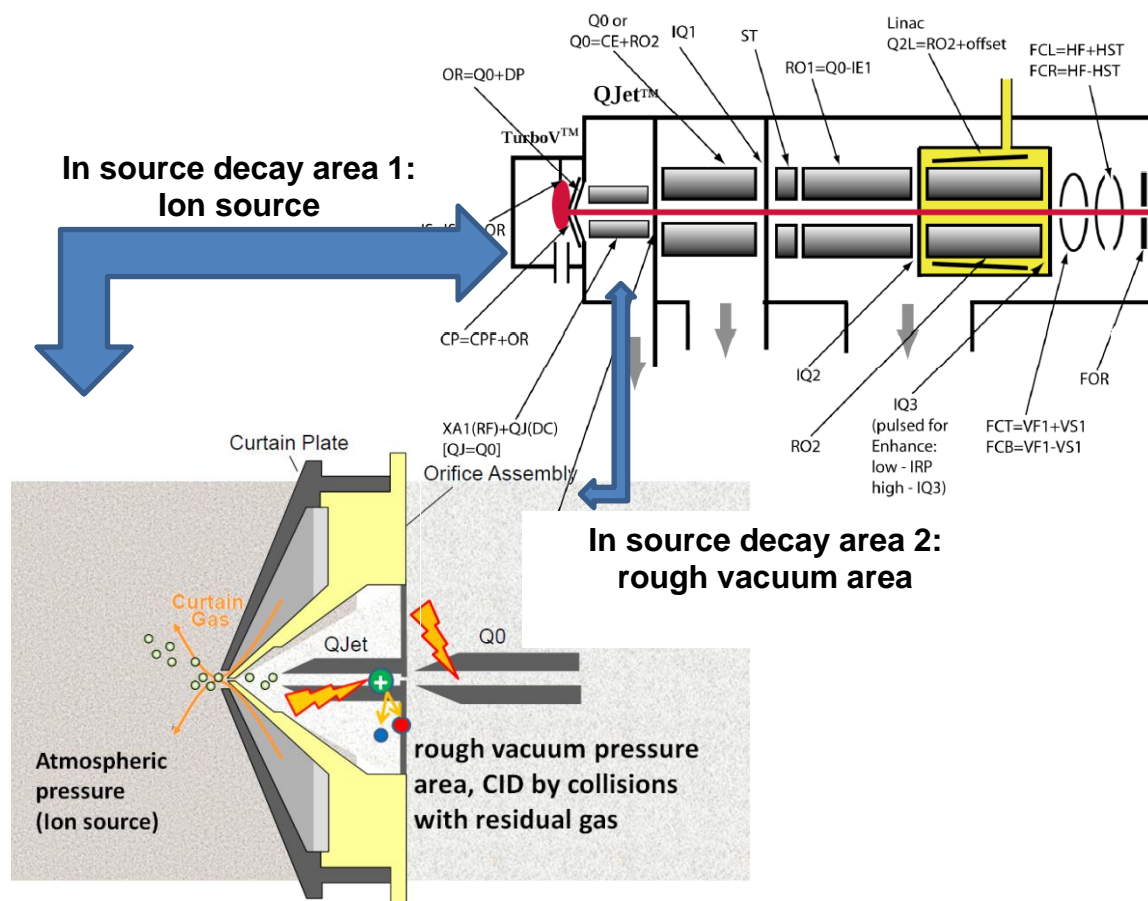
Figure FP5 2 Quadrupole compartment of Sciex 5600+ ion path, applied voltages and locations of in source decay. Note that the only user adjustable voltages are DP and ISVF. Source: own work and reproduced from J. Jasak, AB SCIEX TripleTOF[TM] 5600 System Basic Training, Part 1 Technology, Copyright (2014), with permission from Sciex.

The decay processes in the source on the other hand is dependent on source parameters and chemistry [25]. The faradiol esters of Publication V exhibited neutral loss of water or a fatty acid, that can be tuned for maximum sensitivity or at last relative minimization of decay by solvent composition and ion source parameters like voltages, gas flows and temperatures.

In source fragmentation was found to be extensive and complete using electrospray ionization as no $[M+H]^+$ species of the triterpene diesters were detectable. Albeit still extensive, with atmospheric pressure chemical ionization (APCI) and optimized settings the $[M+H]^+$ species could be identified. For this, ion collision energy was kept at low levels (CE=5V for MS, 17 V for MS/MS), employing high sensitivity MS/MS mode, long acquisition time (1Hz) and multichannel averaging (MCA, 80 time bins), the latter reducing noise at the cost of resolution. Results are summarized in Table TP5.1

Table TP5 1 Results for direct infusion experiments. Reproduced from Journal of Pharmaceutical and Biomedical Analysis, 118, 195-205 Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS. Copyright (2016), with permission from Elsevier.

Compound	Sum formula	Calculated m/z	Found m/z	Δ mDa	Δ ppm
FMM	$C_{58}H_{102}O_4$	863.7851	863.7815	−3.60	−4.17
FMP /FPM	$C_{60}H_{106}O_4$	891.81639	891.8181	1.71	1.92
FPP	$C_{62}H_{110}O_4$	919.84769	919.8358	−11.89	−12.93

While these findings confirmed the presence of dimyristoyl (FMM) and dipalmitoyl (FPP) diesters in the *C. officinalis* extract complementary to the NMR results, the actual presence of the mixed esters FMP/FPM needed to be further elucidated as the 16-O deacetylation was common, yet not exclusive. Comparison (Fig FP5.3) of 3-O-deacylation intensities exhibited by diester standards synthesized from 3-O-monoesters with the extract could not rule out presence of both diesters.
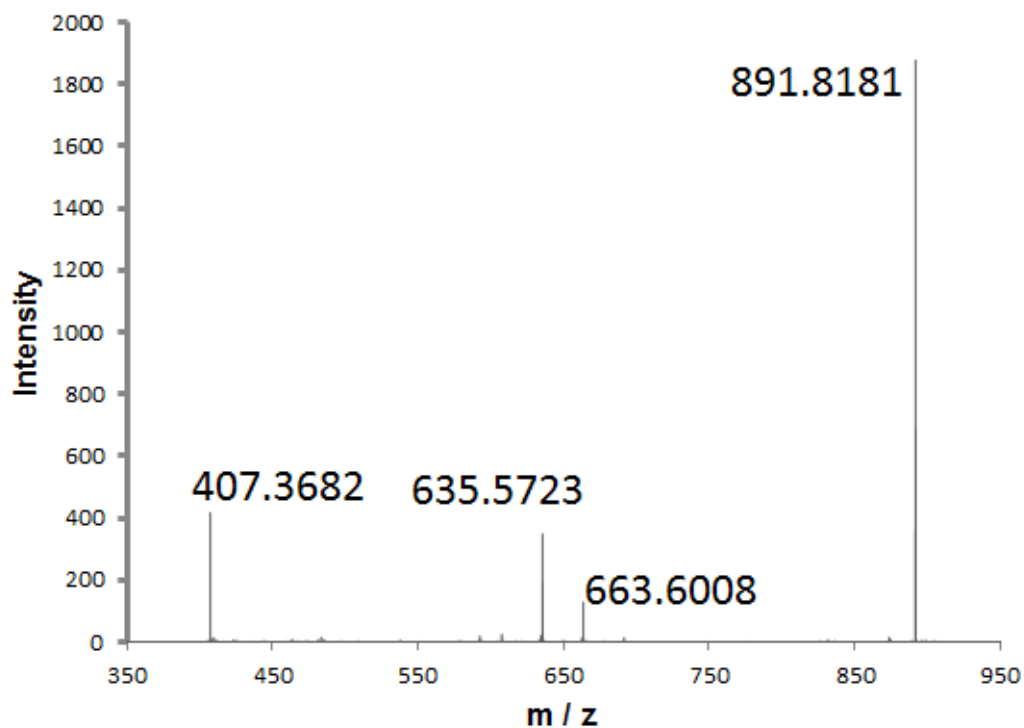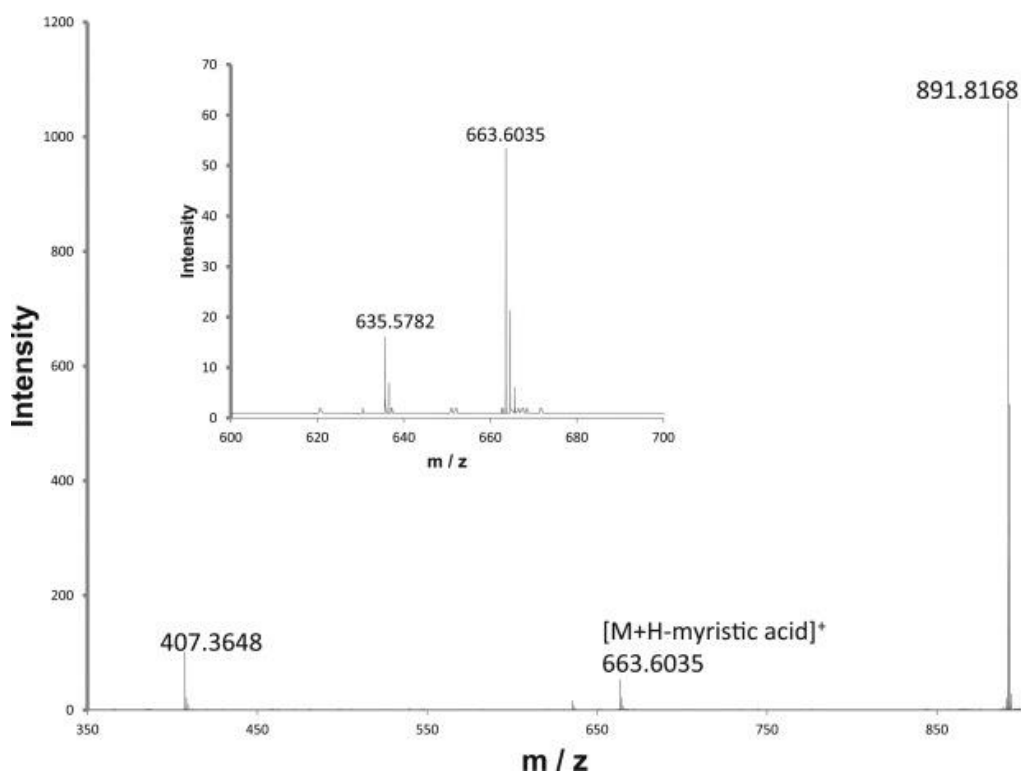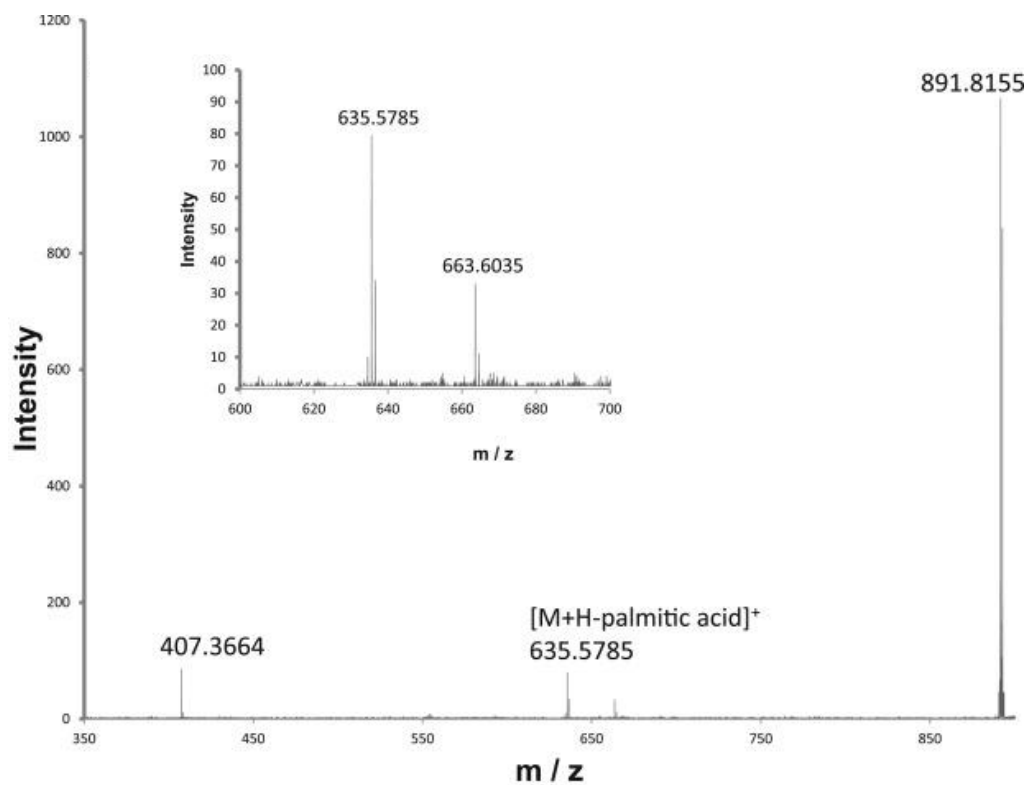
Figure FP5 3 Results for direct infusion experiment for mixed esters. Top: Extract, middle: FMP standard, Down: FMP standard, down: extract. Reproduced from Journal of Pharmaceutical and Biomedical Analysis, 118, 195-205 Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS, Copyright (2016), with

Hence, chromatographic separation prior to mass spectrometry was mandatory. As FMP and FPM are regioisomers with identical methylene (CH2) count, the chromatography could not only rely on the well-established methylene selectivity of C18-type phases [50, 21], yet had to constitute regioselective/shape selective RP-chromatography. Foundations of the C30 chromatography will be described in the following paragraphs.

C30-n-alkyl silica as favourable shape selective material

Early accounts of shape selectivity are given by Sander et al.[47] demonstrating the shape selectivity of monomeric and polymeric phases of various alkyl chain length separating phenanthro[3,4-c]phenanthrene (PhPH),1,2:3,4:5,6:7,8-tetrabenzonapthalene (TBN) and benzo[a]pyrene (BaP). Interest in chromatographic separation of PAHs in field of oil analysis and other fields in which PAHs are of importance, led to the development of according Standard Reference Materials (SRM) at the National Institute of Standards and Technology (NIST) [54]. Of these, SRM 869 was established to judge shape selectivity of chromatographic phases [6].
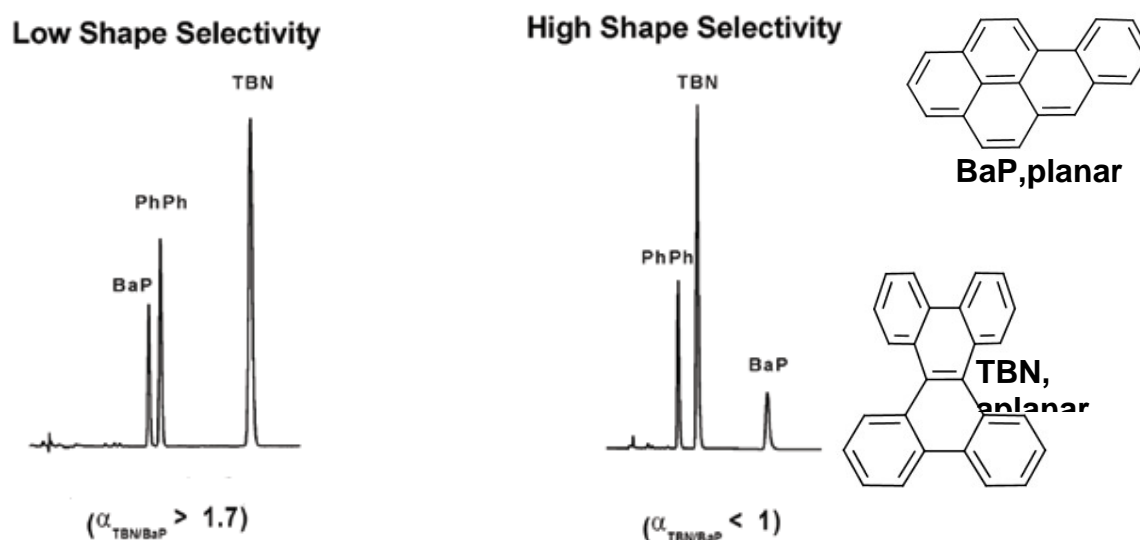
Figure FP5 4 Use of SRM 869 for examination of shape selectivity. Reproduced from LCGC North America, 26, 10, Shape Selectivity in Reversed-Phase Liquid Chromatography, 948-998, Copyright (2008), with permission from Advantsar Communications.

Shape selectivity is judged by the separation factor $\alpha_{TBN / BaP}$, where lower values of alpha indicate higher shape selectivity, see Rimmer et al. [6]. for reference. One of Sanders

findings were, that longer alkyl chain lengths on the immobilized selector increased shape selectivity [47]. An early review on shape selectivity regarding planar and aplanar molecules like PAHs or triterpenoids as examined in Publication V is also given by Sander [48].

Following these early works, after being established by Sander [45], a plethora (selected examples [17, 42, 53, 52]) of papers dealing with application of C30 phases has been published. A notable amount of these works is dedicated to analysis of carotenes and similar compounds, establishing the C30 bond silica phases as "carotenoid phase".

Albert [1] in 1998 reviewing some of the intervening works, summarizing that the C30 phases, albeit offering no benefit over C18 regarding small molecules, indeed possess superior shape selectivity, as long as mobile phase and separation temperature are customized to the respective analytical question.

Basic requirements on the structure of n-alkyl bond silica surfaces regarding shape selectivity

In order to provide for shape selectivity, the stationary phase must provide some kind of scaffold discriminating solutes possessing similar polarity, solubility and lipophilicity but unequal spatial configuration. Considering the most fundamental case of shape selectivity, i.e. shape selective separation of solutes with different shape yet otherwise identical properties regarding RP chromatography, in analogy (yet not identical!) to chiral recognition (ref. to introduction of Publication I) Gibbs free energies of both analytes binding to the shape selective phase must be different from each other.

A model for this mode of action where the difference in Gibbs free energy arises from differences in steric hindrance was given by Wise and Sanders [55], named the "slot model". According to this model, the stationary phase is comprised of "slots" of defined diameter and length. The first discriminant of the analytes is the length-to-breath (L/B) ratio, those with high L/B ratio are capable do permeate further in the alkyl layer, establish more interactions with the alkyl phase and thus are longer retained. With SRM 869, as expected a shape selective phase will retain the planar BaP longer, hence $\alpha_{TBN / BaP}$, <1. See Fig FP5.4 for reference.
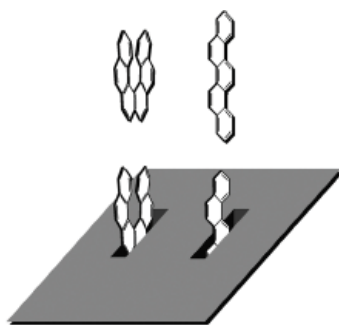
Figure FP5 5 "Slot model" illustrating molecular recognition of analytes with varying L/B-ratio. Reproduced from LCGC North America, 26, 10, Shape Selectivity in Reversed-Phase Liquid Chromatography, 948-998, Copyright (2008), with permission from Advanstar Communications.

These properties can result neither from a phase mode of binding exactly following a solvophobic mechanism / partitioning mechanism [28] nor by the opposing binding enthalpy based view by Carr and coworkers [41] as both models explain retention based on isotropic interaction of eluents, stationary phase and solutes [44]. Concordantly, the n-alkyl layer of shape selective RP stationary phases cannot solely interact with the solutes in an equivalent manner than to corresponding bulk phase of free n-alkanes.

Insights on n-Alkyl bond silica by FTIR

The surface density of alkyl ligands was found to be half that of the corresponding bulk liquids [8] indicating, that the properties of the respective bulk alkane phase are not applicable [9]. An earlier work by Sander et al. [43] employing FT-IR for the study of RP stationary phases revealed, that, judged by the matching IR-spectra, dry C18 bound silica show indeed some similarity to the same silica material with only physically associated alkylsilane or n-alkanes. This behavior is however altered by the mobile phase: under aquatic conditions, the alkyl layer assumes a liquid or glass like phase (indistinguishable by IR) whereas in a high organic environment a brisle / brush like conformation with organic solvent in between is assumed.

An in depth study of the n-alkyl bond silica employing FTIR was provided by Singh et al [49]. Monitoring CH2 stretching modes for conformational order, deuterated CD2 rocking modes for gauche conformation monitoring at selected sites and wagging modes to assess kink (gtg), double gauche (gg) and end gauche (tg) sequences, his work elucidated that, the length of the alkyl chain is proportional to conformational order, the amount of double gauche conformers is reduced with increasing chain length and by lowering temperature.
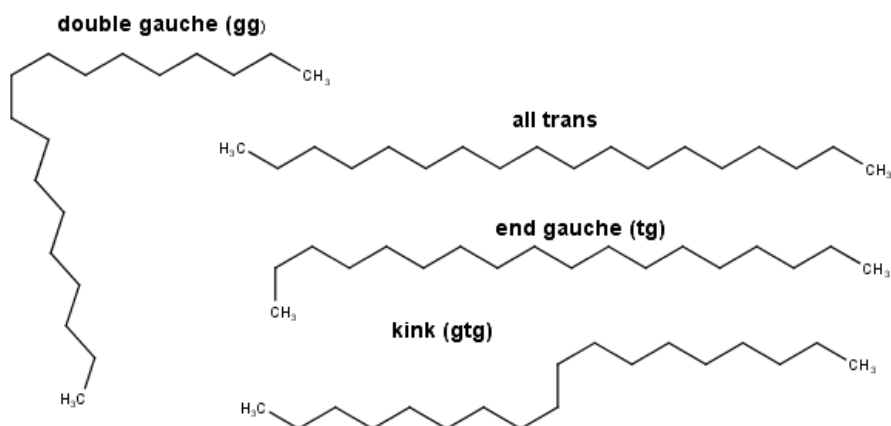
Figure FP5 6 Conformers of n-alkane chains. Source: own work.

Singh also demonstrated, that of all chain lengths (C8-C30) examined, C30 exhibits the lowest number of gtg- and gg-conformers at a given temperature and that the fraction of both conformers is minimized rapidly with decreasing temperature, making it the most ordered phase studied. These findings regarding C30 have been confirmed later by [51].
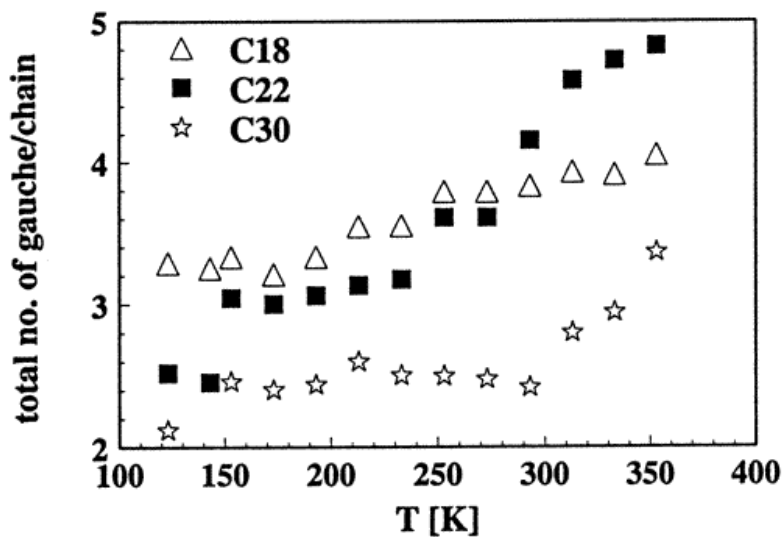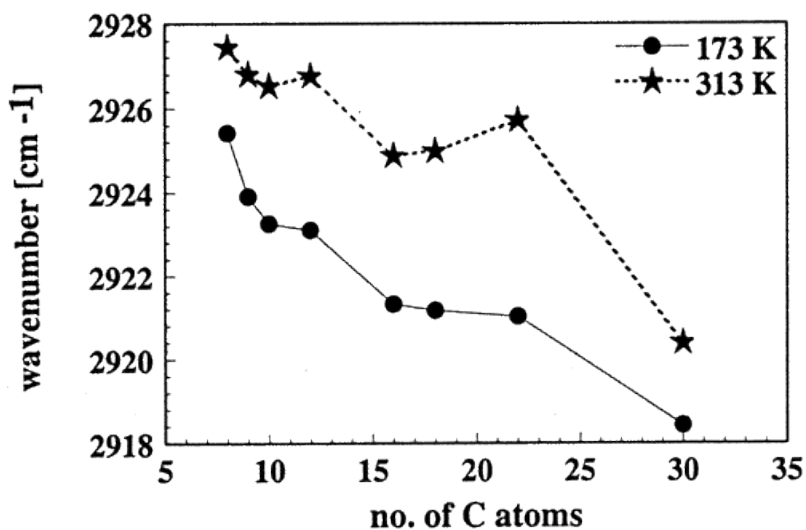
Figure FP5 7 Top: Change of CH2-anti-symmetric stretch band maxima as result of decreasing number gauche conformers. Bottom: C30 exhibiting highest order i.e. lowest gauche defect number over broad temperature range. Reproduced from Singh et al., J. Phys. Chem. B, 2002, 106 (4), pp 878–888, Shape Selectivity in Variable Temperature FT-IR Studies of n-Alkyl Modified Silica Gels, Copyright (2002), with permission from American Chemical Society.
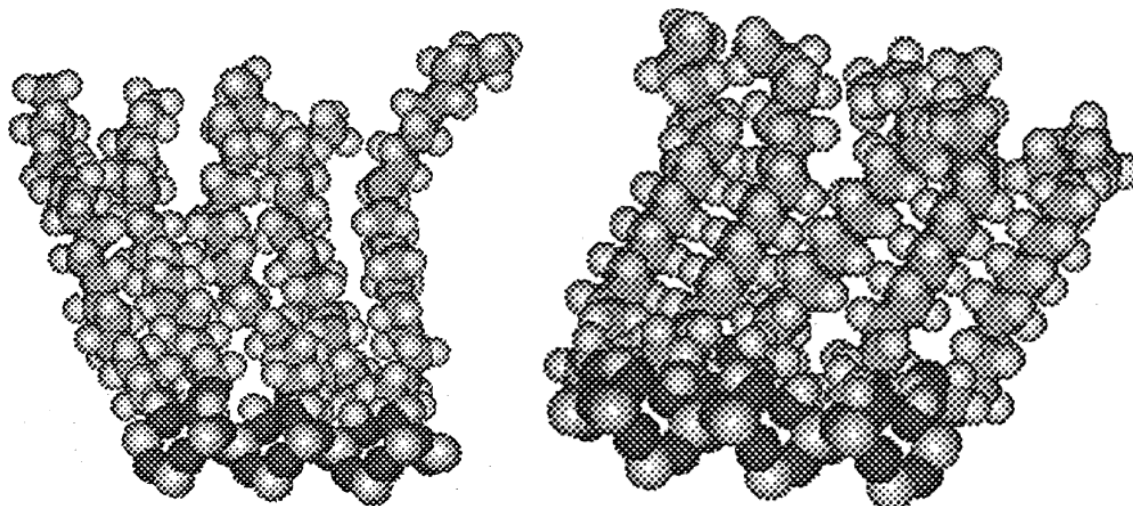
Figure FP5 8 Changes in conformational order in C18 phases in high (left) and low (right) temperature, as portrayed by Singh et al. Reproduced from Singh et al., J. Phys. Chem. B, 2002, 106 (4), pp 878–888, Shape Selectivity in Variable Temperature FT-IR Studies of n-Alkyl Modified Silica Gels,, Copyright (2002), with permission from American Chemical Society.

Raman spectroscopy – a method complementary to FTIR, same results.

Employing mostly the ratio of antisymmetric to symmetric methylene bands $I[v_s(CH_3)]/I[v_s(CH_2)]$ [26] as a measure of conformational order, Raman spectroscopy as method complementary to FTIR was used to confirm the latter findings on the influence of alkyl chain length [11], column temperature [13] and mobile phase composition [12]. Being an emission type spectroscopy, Raman spectroscopy could even be employed to study the properties of stationary phases / RP ligands in actual chromatographic columns [11].

Influence of solvent organic content - SFG spectroscopy

Studying influence of solvent composition on structure and topology of alkyl silica phases by sum frequency generation (SFG) spectroscopy Henry et al. [20] demonstrated C18 phases are unaffected by high aqueous conditions in terms of the "phase collapse myth", i.e. that under high concentrations of water in the mobile phase the alkyl phase, collapses to a "oil

drop" liquid / glass like state conformation. Henry's findings confirm Raman spectroscopic observations by Doyle et al [11]. Yet, presence of water reduces order in the alkyl layer. In contrast, pure organic conditions were found to facilitate additional order of the alkyl chains.

NMR of alkyl silica phases

Nuclear magnetic resonance can be used to study surface coverage and conformational conditions of alkyl silica phases. Due to the insoluble nature, NMR on alkyl silica phases is performed either as solid state or suspended state NMR [2]. The modification of the surface, ligand density can be monitored by 29Si-CP/MAS NMR spectroscopy. Insights to the tethered alkyl chain can be retrieved by 13C- CP/MAS or HR/MAS NMR spectroscopy. C30 phases in particular have been studied by NMR by Albert and coworkers at University of Tübingen [3, 16, 38, 40]. One of these works by Pursch et al. [39] yielded the same results as other spectroscopic methods: Examining the ~2.6 ppm distant signals arising from trans and gauche conformers [15] (see Fig FP5.8) yielded: increased surface bonding density decreases gauche conformers, whereas temperature and gauche ratio are positively correlated.
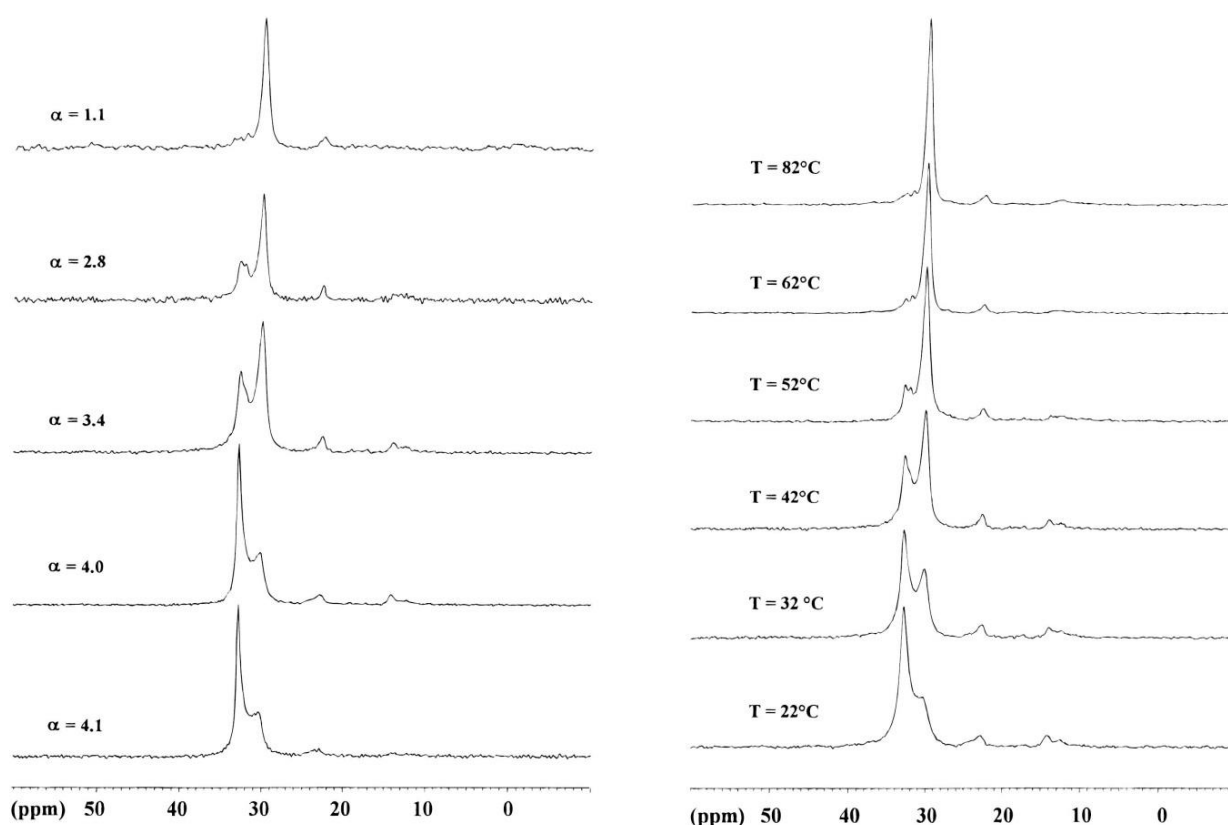
Figure FP5 9 C30 alkyl chain conformation shift for main chain CH2 groups (C4-C27), from trans δ=32.6 ppm to gauche δ=30 ppm with increasing temperature and decreasing surface density α. Reproduced from Anal. Chem., 1996, 68 (2), pp 386–393, Temperature-Dependent Behavior of C30 Interphases. A Solid-State NMR and LC–NMR Study Copyright (1996), with permission American Chemical Society.

In silico chemistry – Simulations of RP-Phases

Most in silico molecular studies on chromatographic systems are performed by molecular dynamics (MD) simulations, as Monte Carlo (MC) approaches are too computationally intensive for the extent of the simulated systems. Two fundamental works regarding alkyl silica surface order employing MD simulations on C8, C18 and C30 type of RP-Phases by Lippa et al. examined influence of chain length and ligand surface density [28] and of mobile phase and column temperature [29]. The results of these studies are congruent with the spectroscopic and chromatographic experimental findings:
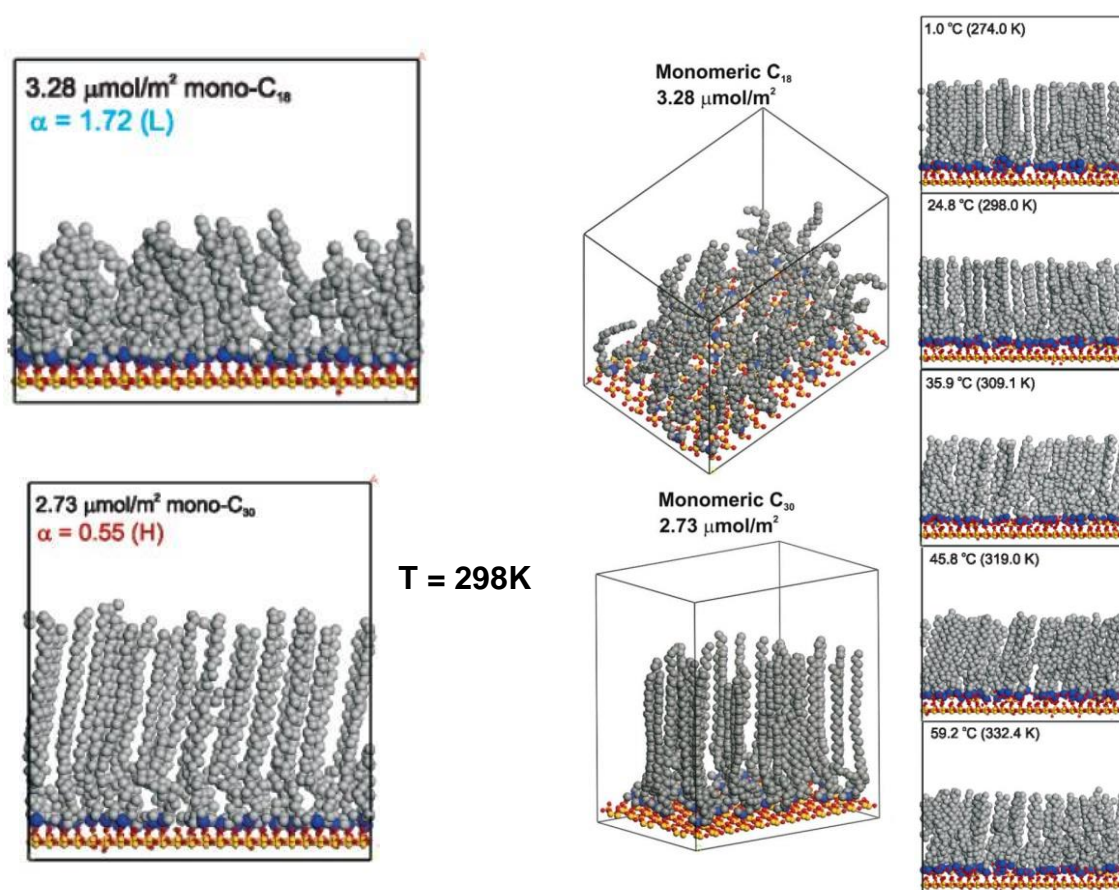
Figure FP5 10 Results from Lippa et al. MD calculations, Anal. Chem., 2005, 77 (24), pp 7862–7871, Copyright (2005), with permission from American Chemical Society Publications.

Order, governed by a low ratio of gauche (tg), double gauche (gg), and kink (gtg) conformers to trans conformers of the chains of the alkyl phase increases with chain length and surface density but decreases with temperature. See Fig FP5.4 for reference).

Furthermore, the work of Lippa [29] revealed, that C30 ligands require less surface density, see fig FP5.9. As depicted, the same order can be achieved at room temperature and with lower surface density in C30 phases as with C18 at higher surface densities and sub ambient temperatures.

Facilitating shape selectivity – C30, high surface density and low temperature

To summarize, the in silico results of Lippa and others [4, 24, 29, 28], as well as the congruent experimental results [7], [14], [13, 43, 49, 23] by FTIR, Raman, NMR and chromatographic experimentation revealed why C30 phases possess advanced capabilities as shape selectors / carotenoid phase:

- **Gauche dihedrals and shape selectivity:** Gauche dihedrals and of the alkyl chain relative to silica surface are reciprocal to each other. Up to 20% gauche dihedrals do not impair high shape selectivity ($\alpha_{TBN / BaP}$ <1.0), 20%-30% result in intermediate ($\alpha_{TBN / BaP}$ 1.0-1.7), above 30% yield low shape selectivity ($\alpha_{TBN / BaP}$ >1.7).

- **Gauche dihedrals and temperature**: Increasing temperature for C18 from 274K to 332K more than doubles gauche dihedrals (12% to 16). Directly Comparing trifunctional C18 to C30, increasing temperature from 24° to 60° affects order of C18 (19% to 26.5% gauche), but does not affect C30 (9.5% to 10.7% gauche).

- **Chain density**: Chain density does not affect gauche concentration in C30 as strong as it affects C18. In order to achieve low gauche concentration (10%-15%) surface density has to be increased to over 5 µmol/m² (5.9 for 10% gauche), an occupancy difficult to achieve experimentally. With C30 this value is already reached with at 2.7 µmol/m² (16% gauche) and high shape selectivity (10%gauche, $\alpha_{TBN / BaP}$ 0.55), can be achieved still below 5 µmol/m².

- **Monomeric vs polymeric alkylsilanes**: With C18, in order to achieve low gauche concentrations, trifunctional C18 (polymeric phase) and therefore a more extensive

synthetic route [46] must be employed. In contrast, with C30 10% gauche and $\alpha_{TBN/BaP}$ 0.55,   is readily achievable using monofunctional chlorotriacontylsilane, hence the preferable easier synthetic route to a monomeric alkyl silica phase can be followed.

- **Phase diameter**: With C18 phase thickness varies with temperature, albeit the effect is stronger with monomeric phases (25% vs 4% reduction of phase thickness between 273 and 298 K). With C30, phase thickness is, as expected from low temperature influence on gauche concentration and tilt angle not significantly affected.

- **Topography**: Smoothness of the alkyl phases surface, described by tile angle of the alkyl chains relative to the silica surface and by the peak to valley distance on the top of the alkyl surface (for graphical ref. see FP5.9). C30 materials have lower tilt angles, and are as aforementioned, less susceptible to temperature influence.

A comprehensive more recent review of order, disorder and shape selectivity in alkyl stationary phases is given by Sander [44].

Results of Publication V

The more straightforward task of confirming the presence of faradiol as well as the myristate and palmitate esters of arnidiol, arnitriol A, lupane-3β,16β,20-triol, and maniladiol, was achieved by non-aqueous reversed phase C18 chromatography. Due to the high lipophilicity of the triterpene monoesters, no elution could be achieved employing a classic RP water/acetonitrile gradient. Hence, a non-aqueous gradient was applied, running from 100 % acetonitrile delivered by Channel A to 100% methanol delivered by Channel B in 25 min. Non-aqueous gradients, while uncommon to RP chromatography, have been applied by previous works [18, 33] .

As described in the sections on shape selective chromatography, C30 was considered as the eligible stationary phase. Concluded from the cited previous works, a monomeric type ligand chemistry would suffice, achieving high shape selectivity ($\alpha_{TBN/BaP}$ =0.55) already with low

surface density.   Different surface coverages have been synthesized, employing surface modification by reflux of Kromasil 100 Å 5 µm silica gel with triacontyldimethylchlorosilane.
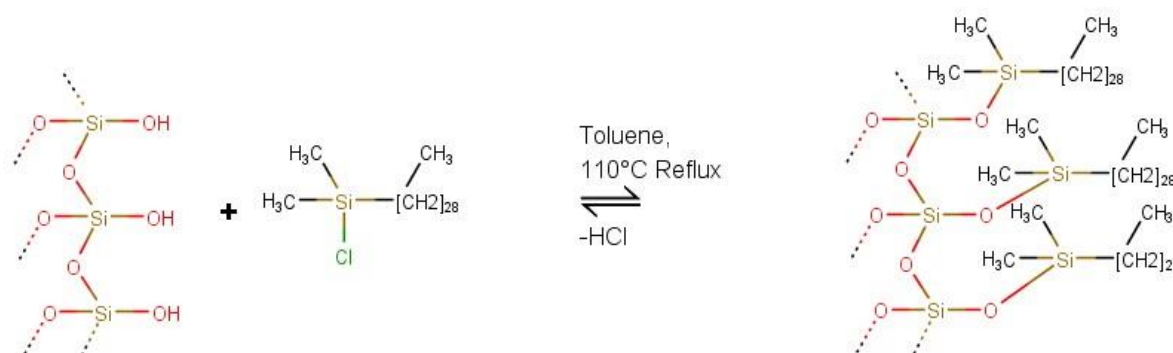
Figure FP5 11 Kromasil 100, 5µm modification as employed in Publication V. Molar ratio of triacontyldimethylchlorosilane to silanol groups was varied to yield several surface densities, of which 1.41 µmol/m² was evaluated as suitable. Source: own work. See Publication V for reference.

Implementing the reciprocal temperature/shape selectivity relationship, chromatography was conducted at low temperature (8°C). Albeit it has to be ascertained, that lowering temperature was limited by backpressure of the chromatographic column and the given flow rate and binary gradient from methanol to isopropanol. Instrumentation hardware (Agilent 1290 Infinity I Series) limits backpressure to a maximum of 120 MPa. Under these conditions, FMM, FMP/FPM, FPP could be resolved. The pronounced 16-O-instability of the mixed esters could be utilized for structural assignement. For cross checking, the trace of the 20-taraxastene skeleton (m/z = 407.3673) was compared to the results, revealing not only the FMP/FPM regioisomers could be partially resolved by C30-RP chromatography, but also the regioisomers of the faradiol monoesters.
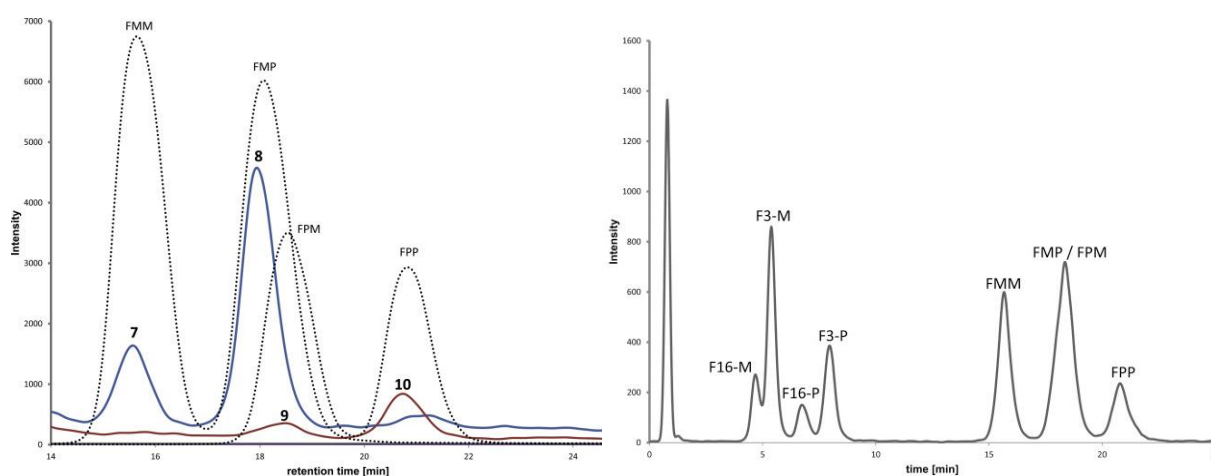
Figure FP5 12 Left: High-Res TOF-MS C30 low temperature, nonaqueous gradient extracted ion chromatograms. Blue: XIC for FMP, m/z 635.5761±2.5 mDa; Red; FPM in source decay, m/z 663.6074

±2.5 mDa Dotted trace: synthesized standards. Right: cross check on 20-taraxastene skeleton (m/z = 407.3673). Reproduced from Journal of Pharmaceutical and Biomedical Analysis, 118, 195-205 Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30

The gradient employed was a non-aqueous type, running from 100 % methanol delivered by Channel A to 100% isopropanol delivered by Channel B in 25 min, as with classic RP water to organic gradients separation could not be achieved. It was assumed, that this results on the one hand from the strong lipophilicity of the diesters hence long retention time and peak broadening (longitudinal diffusion term of van-Deemter equation B/v) under aqueous/organic conditions. On the other hand, it can be hypothesized, that the reported (ref. previous section, Doyle et al. [11]) reduction of order in alkyl silica phases under high water concentration reduces the shape selectivity of the given in house manufactured C30 phase below a threshold required for separation of the diesters, especially FMP\FPM.

Additionally, the confirmation of the presence of monoesters of faradiol, arnidiol, arnitriol, maniladiol and lupane-3β,16β,20-triol could be achieved.

To conclude, advanced chromatographic separation by shape selective C30 phase and particular MS settings with APCI ion source and MCA feature could lead to the adequate analysis of these difficult to characterize analytes and sample mixtures. Only the smart combination of the appropriate LC and finetuned MS could accomplish this success.

Abbreviations

FTIR: Fourier-transform infrared spectroscopy

FMM: Faradioldimyristate

FPP: Faradioldipalmitate

FMP/FPM: Mixed esters of faradiol and myristic/pamitic acid

SFG: sum frequency generation

MC: Monte Carlo simmulation

References

[1]   Klaus Albert. Correlation between chromatographic and physicochemical properties of stationary phases in hplc: C30 bonded reversed-phase silica. *TrAC Trends in Analytical Chemistry*, 17(10):648–658, 1998.

[2]   Klaus Albert. Nmr investigations of stationary phases. *Journal of separation science*, 26(3-4):215–224, 2003.

[3]   Klaus Albert, Tanja Lacker, Martin Raitza, Matthias Pursch, Hans-Joachim Egelhaaf, and Dieter Oelkrug. Investigating the selectivity of triacontyl interphases. *Angewandte Chemie International Edition*, 37(6):777–780, 1998.

[4]   Kazuhiro Ban, Yoshihiro Saito, and Kiyokatsu Jinno. Characterization of the microscopic surface structure of the octadecylsilica stationary phase using a molecular-dynamics simulation. *Analytical sciences*, 20(10):1403–1408, 2004.

[5]   Ethan Basch, Steve Bent, Ivo Foppa, Sadaf Haskmi, David Kroll, Michelle Mele, Philippe Szapary, Catherine Ulbricht, Mamta Vora, and Sophanna Yong. Marigold (calendula officinalis l.) an evidence-based systematic review by the natural standard research collaboration. *Journal of herbal pharmacotherapy*, 6(3-4):135–159, 2006.

[6]   Ronald E. Majors Katrice A. Lippa Catherine A. Rimmer, Lane C. Sander. Shape selectivity in reversed-phase liquid chromatography. *LCGC North America*, 26(10):984–998, 2008.

[7]   Jinlong Cheng, Matilda Fone, and Mark W Ellsworth. Solid state nmr study on the conformation and mobility of n-octadecyl chains in a silane coupling agent attached to the surface of colloidal silica. *Solid state nuclear magnetic resonance*, 7(2):135–140, 1996.

[8]   Henri Colin and Georges Guiochon. Introduction to reversed-phase high-performance liquid chromatography. *Journal of Chromatography A*, 141(3):289–312, 1977.

[9]   Henri Colin and Georges Guiochon. Comparison of some packings for reversed-phase high-performance liquidâ€"solid chromatography: Ii. some theoretical considerations. *Journal of Chromatography A*, 158:183–205, 1978.

[10]   Roberto Della Loggia, Aurelia Tubaro, Silvio Sosa, H Becker, O Isaac, et al. The role of triterpenoids in the topical anti-inflammatory activity of calendula officinalis flowers. *Planta medica*, 60(06):516–520, 1994.

[11]   Charles A Doyle, Thomas J Vickers, Charles K Mann, and John G Dorsey. Characterization of liquid chromatographic stationary phases by raman spectroscopy effect of ligand type. *Journal of Chromatography A*, 779(1-2):91–112, 1997.

[12]   Charles A Doyle, Thomas J Vickers, Charles K Mann, and John G Dorsey. Characterization of c18-bonded liquid chromatographic stationary phases by raman spectroscopy:: the effect of mobile phase composition. *Journal of Chromatography A*, 877(1):25 − 39, 2000.

[13]   Charles A Doyle, Thomas J Vickers, Charles K Mann, and John G Dorsey. Characterization of c18-bonded liquid chromatographic stationary phases by raman spectroscopy:: the effect of temperature. *Journal of Chromatography A*, 877(1):41 − 59, 2000.

[14]   Michael W Ducey, Christopher J Orendorff, Jeanne E Pemberton, and Lane C Sander. Structure- function relationships in high-density octadecylsilane stationary phases by raman spectroscopy. 1. effects of temperature, surface coverage, and preparation procedure. *Analytical chemistry*, 74(21):5576–5584, 2002.

[15]   William L Earl and DL VanderHart. Observations in solid polyethylenes by carbon-13 nuclear magnetic resonance with magic angle sample spinning. *Macromolecules*, 12(4):762–767, 1979.

[16]   H-J Egelhaaf, D Oelkrug, M Pursch, and K Albert. Combined fluorescence and nmr studies of reversed hplc stationary phases with different ligand chain lengths. *Journal of Fluorescence*, 7(4):311–315, 1997.

[17]   Curt Emenhiser, Lane C Sander, and Steven J Schwartz. Capability of a polymeric c30 stationary phase to resolve cis-trans carotenoid isomers in reversed-phase liquid chromatography. *Journal of Chromatography A*, 707(2):205–216, 1995.

[18]   Chiara Fanali, Marco Beccaria, Simona Salivo, Peter Tranchida, Giusy Tripodo, Sara Farnetti, Laura Dugo, Paola Dugo, and Luigi Mondello. Non-polar lipids characterization of quinoa (chenopodium quinoa) seed by comprehensive two-dimensional gas chromatography with flame ionization/mass spectrometry detection and non-aqueous reversed-phase liquid chromatography with atmospheric pressure chemical ionization mass spectrometry detection. *Journal of separation science*, 38(18):3151–3160, 2015.

[19]   M Hamburger, S Adler, D Baumann, A Förg, and B Weinreich. Preparative purification of the major anti-inflammatory triterpenoid esters from marigold (calendula officinalis). *Fitoterapia*, 74(4):328–338, 2003.

[20]   Matthew C Henry, Lauren K Wolf, and Marie C Messmer. In situ examination of the structure of model reversed-phase chromatographic interfaces by sum-frequency generation spectroscopy. *The Journal of Physical Chemistry B*, 107(12):2765–2770, 2003.

[21]   S Heron and A Tchapla. Description of retention mechanism by solvophobic theory: influence of organic modifiers on the retention behaviour of homologous series in reversed-phase liquid chromatography. *Journal of Chromatography A*, 556(1-2):219–234, 1991.

[22]   Csaba Horvath, Wayne Melander, and Imre Molnar. Solvophobic interactions in liquid chromatography with nonpolar stationary phases. *Journal of Chromatography A*, 125(1):129–156, 1976.

[23]   K Jinno, T Ibuki, N Tanaka, M Okamoto, JC Fetzer, WR Biggs, PR Griffiths, and JM Olinger. Retention behaviour of large polycyclic aromatic hydrocarbons in reversed-phase liquid chromatography on a polymeric octadecylsilica stationary phase. *Journal of Chromatography A*, 461:209–227, 1989.

[24]   Steven J Klatte and Thomas L Beck. Molecular dynamics simulations of tethered alkane chromatographic stationary phases. *The Journal of Physical Chemistry*, 99(43):16024–16032, 1995.

[25]   Yoshiharu KOBAYASHI, Kayoko SAIKI, and Fukuko WATANABE. Characteristics of mass fragmentation of steroids by atmospheric pressure chemical ionization-mass spectrometry. *Biological and Pharmaceutical Bulletin*, 16(11):1175–1178, 1993.

[26]   Kåre Larsson and RP Rand. Detection of changes in the environment of hydrocarbon chains by raman spectroscopy and its application to lipid-protein systems. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism*, 326(2):245–255, 1973.

[27]   Matthew J Leach. ï»¿ï»¿ï»¿ calendula officinalis and wound healing: A systematic review. *Wounds: a compendium of clinical research and practice*, 20(8):236–243, 2008.

[28]   Katrice A Lippa, Lane C Sander, and Raymond D Mountain. Molecular dynamics simulations of alkylsilane stationary-phase order and disorder. 1. effects of surface coverage and bonding chemistry. *Analytical chemistry*, 77(24):7852–7861, 2005.

[29]   Katrice A Lippa, Lane C Sander, and Raymond D Mountain. Molecular dynamics simulations of alkylsilane stationary-phase order and disorder. 2. effects of temperature and chain length. *Analytical chemistry*, 77(24):7862–7871, 2005.

[30]   Yee-Chung Ma and Hee-Yong Kim. Determination of steroids by liquid chromatography/mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 8(9):1010–1020, 1997.

[31]   Czygan F.-C. Mayer, J.G. Marigold - calendula officinalis l.: A cultural-historical portrait of a medicinal plant. *Zeitschrift fur Phytotherapie*, 21(3):170–178, 2000.

[32]   BP Muley, SS Khadabadi, and NB Banarase. Phytochemical constituents and pharmacological activities of calendula officinalis linn (asteraceae): a review. *Tropical Journal of Pharmaceutical Research*, 8(5), 2009.

[33]   HJCF Nelis and AP De Leenheer. Isocratic nonaqueous reversed-phase liquid chromatography of carotenoids. *Analytical Chemistry*, 55(2):270–275, 1983.

[34]   Hannes Neukirch, Michele D'Ambrosio, Silvio Sosa, Gianmario Altinier, Roberto Della Loggia, and Antonio Guerriero. Improved anti-inflammatory activity of three new terpenoids derived, by systematic chemical modifications, from the abundant triterpenes of the flowery plant calendula officinalis. *Chemistry & biodiversity*, 2(5):657–671, 2005.

[35]   C Nicolaus, S Junghanns, R Murillo, and I Merfort. Triterpene alcohols from calendula officinalis l. flowers and in vitro studies on their wound healing activity. *Planta Medica*, 80(16):P2B63, 2014.

[36]   Christoph Nicolaus, Susanne Junghanns, Anja Hartmann, Renato Murillo, Markus Ganzera, and Irmgard Merfort. In vitro studies to evaluate the wound healing properties of calendula officinalis extracts. *Journal of ethnopharmacology*, 196:94–103, 2017.

[37]   Christoph Nicolaus, Adrian Sievers-Engler, Renato Murillo, Michele D'ambrosio, Michael Lämmerhofer, and Irmgard Merfort. Mastering analytical challenges for the characterization of pentacyclic triterpene mono-and diesters of calendula officinalis flowers by non-aqueous c30 hplc and hyphenation with apci-qtof-ms. *Journal of pharmaceutical and biomedical analysis*, 118:195–205, 2016.

[38]   Matthias Pursch, Lane C Sander, and Klaus Albert. Peer reviewed: understanding reversed-phase lc with solid-state nmr. *Analytical chemistry*, 71(21):733A–741A, 1999.

[39]   Matthias Pursch, Sabine Strohschein, Heidrun Händel, and Klaus Albert. Temperature-dependent behavior of c30 interphases. a solid-state nmr and lcâˆ'nmr study. *Analytical Chemistry*, 68(2):386–393, 1996.

[40]   Martin Raitza, Martin Herold, Arndt Ellwanger, Günter Gauglitz, and Klaus Albert. Solid-state nmr and ellipsometric investigations of c30 chains bonded to sio2 surfaces. *Macromolecular Chemistry and Physics*, 201(7):825–829, 2000.

[41]   Ravindra PJ Ranatunga and Peter W Carr. A study of the enthalpy and entropy contributions of the stationary phase in reversed-phase liquid chromatography. *Analytical chemistry*, 72(22):5679–5692, 2000.

[42]   Russell Rouseff, Lee Raley, and Hans-Jurgen Hofsommer. Application of diode array detection with a c-30 reversed phase column for the separation and identification of saponified orange juice carotenoids. *Journal of Agricultural and Food Chemistry*, 44(8):2176–2181, 1996.

[43]   Lane C Sander, James B Callis, and Larry R Field. Fourier transform infrared spectrometric determination of alkyl chain conformation on chemically bonded reversed-phase liquid chromatography packings. *Analytical Chemistry*, 55(7):1068–1075, 1983.

[44]   Lane C Sander, Katrice A Lippa, and Stephen A Wise. Order and disorder in alkyl stationary phases. *Analytical and bioanalytical chemistry*, 382(3):646–668, 2005.

[45]   Lane C Sander, Katherine Epler Sharpless, Neal E Craft, and Stephen A Wise. Development of engineered stationary phases for the separation of carotenoid isomers. *Analytical Chemistry*, 66(10):1667–1674, 1994.

[46]   Lane C Sander and Stephen A Wise. Synthesis and characterization of polymeric c18 stationary phases for liquid chromatography. *Analytical Chemistry*, 56(3):504–510, 1984.

[47]   Lane C Sander and Stephen A Wise. Effect of phase length on column selectivity for the separation of polycyclic aromatic hydrocarbons by reversed-phase liquid chromatography. *Analytical chemistry*, 59(18):2309–2313, 1987.

[48]   Lane C Sander and Stephen A Wise. Shape selectivity in reversed-phase liquid chromatography for the separation of planar and non-planar solutes. *Journal of Chromatography A*, 656(1-2):335–351, 1993.

[49]   Shashikala Singh, Jürgen Wegmann, Klaus Albert, and Klaus Mueller. Variable temperature ft-ir studies of n-alkyl modified silica gels. *The Journal of Physical Chemistry B*, 106(4):878–888, 2002.

[50]   Roger M Smith. Functional group contributions to the retention of analytes in reversed-phase high-performance liquid chromatography. *Journal of Chromatography A*, 656(1-2):381–415, 1993.

[51]   Gokulakrishnan Srinivasan, Matthias Pursch, Lane C Sander, and Klaus Mueller. Ftir studies of c30 self-assembled monolayers on silica, titania, and zirconia. *Langmuir*, 20(5):1746–1752, 2004.

[52]   S. Strohschein, G. Schlotterbeck, J. Richter, M. Pursch, L.-H. Tseng, H. Händel, and K. Albert. Comparison of the separation of cis/trans isomers of tretinoin with different stationary phases by liquid chromatography-nuclear magnetic resonance coupling. *Journal of Chromatography A*, 765(2):207 – 214, 1997.

[53]   Sabine Strohschein, Matthias Pursch, Heidrun Händel, and K. Albert. Structure elucidation of β-carotene isomers by hplc-nmr coupling using a c30 bonded phase. *Fresenius' Journal of Analytical Chemistry*, 357(5):498–502, Feb 1997.

[54]   Stephen A Wise, Karen W Phinney, Lane C Sander, and Michele M Schantz. Role of chromatography in the development of standard reference materials for organic analysis. *Journal of Chromatography A*, 1261:3–22, 2012.

[55]   Stephen A Wise and Lane C Sander. Factors affecting the reversed-phase liquid chromatographic separation of polycyclic aromatic hydrocarbon isomers. *Journal of High Resolution Chromatography*, 8(5):248–255, 1985.

[56]   Zdzisaw Wojciechowski, Monika Bochenska-Hryniewicz, Bozena Kucharczak, and Zofia Kasprzyk. Sterol and triterpene alcohol esters from calendula officinalis. *Phytochemistry*, 11(3):1165–1168, 1972.

[57]   K Zitterl-Eglseer, S Sosa, J Jurenitsch, M Schubert-Zsilavecz, R Della Loggia, A Tubaro, M Bertoldi, and Ch Franz. Anti-oedematous activities of the main triterpendiol esters of marigold (calendula officinalis l.). *Journal of ethnopharmacology*, 57(2):139–144, 1997.

[58]   Karin Zitterl-Eglseer, Gottfried Reznicek, Johann Jurenitsch, Johannes Novak, Werner Zitterl, and Chlodwig Franz. Morphogenetic variability of faradiol monoesters in marigold calendula officinalis l. *Phytochemical Analysis*, 12(3):199–201, 2001.

Part 3 – Publication IV, IX, X and XI

Introduction to Publications IV, IX, X, and XI

Intact protein mass spectrometry

Background

This part will be focused on some fundamental properties of commercial QTOF-systems regarding large ion species mass spectrometry, in particular mass spectrometry of intact proteins, as they are usually optimized for small molecules and peptides. As Compton et al[15] describe, expectations regarding mass accuracy, mass resolution and sensitivity must be adjusted, as these performance parameters are subject to effects resulting the chemistry and physics of ion motion in mass spectrometers. For the experimenter, these properties must be understood and applied to instrumental settings in order to achieve maximum analytical performance with intact protein mass spectrometry.

Deconvolution errors, isotopic peak broadening and chemical noise

Deconvolution Error: Mass resolution, and thereby peak broadness, is described by $R = m/FWHM$ meaning the absolute amount of the relative measure R is dependent on mass of the ion. e.g. FWHM at 500 dalton is 14mDa, where at 150kDa it is already 4.2 Da for single charged ions at the same resolution. Fortunately with ESI, proteins are multiply charged, i.e. their mass spectrum consists of an array of m/z signals described by $m/z = [M \pm nH]^{\pm n}/z$, termed a charge envelope (s. Fig FP4.1). This multi charging shifts their m/z signals in the m/z=1000-4000 range, implying apparent peak broadness (FWHM) of ~30-100 mDa during measurement. However, a drawback is introduced by data interpretation. For spectral interpretation, the charge of each peak z_{cal} is derived from the charge envelope and used to calculate the mass M of the protein as shown in Fig FP4.1 for a H-adducted charged 42kDa protein.

$$z_{cal} = \frac{A - M_{Proton}}{B - A}$$

$$M = z_{cal}(A - M_{Proton})$$
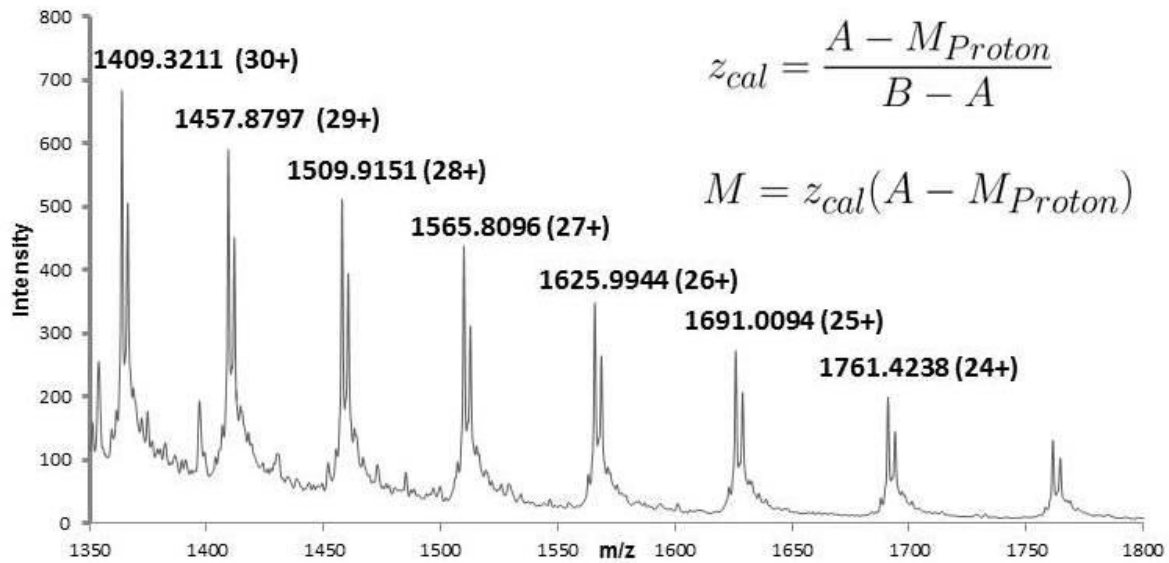
Figure FP4 1 Charge envelope of 42kDa JNK3 kinase with m/z values and corresponding charge. Source: unpublished data, method according to Publication X. A and B denote two subsequent m/z signals (with m/z B > m/z A), z_{cal} the calculated charge in the series, M_{Proton}, mass of proton (1.007276466583 Da, usually 1.0073 or even 1 suffices Note that peak is broad (~10-20Da at the basis) as result of isotope envelope and adduct subspecies

As shown in figure FP4.1 the charge and mass of the protein is reconstructed from the m/z-domain by multiplication, also multiplying deviation measured from theoretical m/z-value arising from actual measurement errors, peak maximum detection errors or calibration errors. Randomly distributed errors can be rectified by averaging over charge envelope as opposed to calibration errors.

Resolution is more severely affected when reconstructing mass peaks during deconvolution. As the m/z peaks are comprised of a finite amount of data points their spacing is increased by the factor identical to charge state, resulting in a reconstructed peak that appears broader than the actual peak would be detected. Also, each signal has a different "broadening factor", as it has a different charge

147
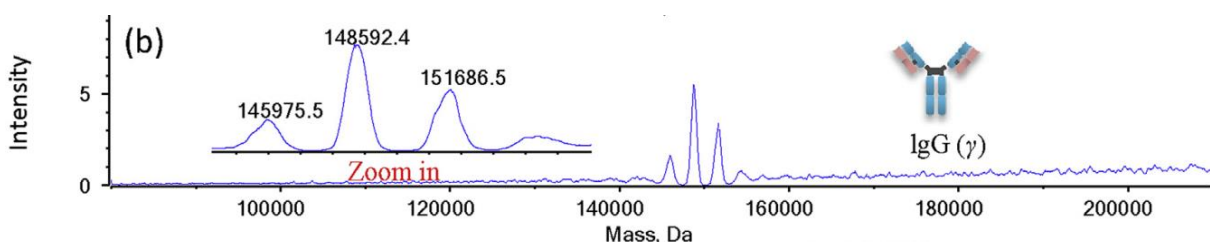
Figure FP4 2 Deconvolution result of an IgG antibody. Note the reconstructed peak is much broader as resolution (35000) of the Sciex 5600+ would imply. Reproduced from Publication XI Papain-functionalized gold nanoparticles as heterogeneous biocatalyst for bioanalysis and biopharmaceuticals analysis, Analytica Chimica Acta, Volume 963, Pages 33-43 Copyright (2017), with permission from Elsevier.

The problem of charge multiplicity induced calculation errors cannot be remedied physically on QTOF-Systems, as the transition window yielding sufficient S/N is determined by Q1 transmission characteristic fixing the useable m/z–corridor (for Sciex 5600[+]: ~900-4000 m/z). This requires higher charge states for larger proteins to pass Q1 and thereby introducing a mass dependency of mass accuracy and resolution, degrading with increasing mass. An exact and frequent calibration of TOF mass axis is mandatory to prevent systematic errors, e.g. a drift of 50ppm, typically in the time scale of one hour, for the 42kDa protein of Fig FP4.1 would give rise to 2 Da deviation, inacceptable for many applications.

As a consequence, data interpretation/deconvolution algorithms usually operate on the m/z data, rather than reconstructing an uncharged molecular peak[32, 40, 35].

Isotopic peak broadening and asymmetry) A major effect on mass spectroscopic resolution is elicited by isotopic distributions[8, 26]. Proteins consist mainly from elements with the following number of sotopes: Carbon(3), Hydrogen(3), Nitrogen(2) and Oxygen(3). With increasing atom count, the monoisotopic peak becomes less predominant, e.g. for an alkane with chain length of 90 the monoisotopic and ($1x^{13}C$)-Species have already same intensity. With large Proteins, what is actually ionized and detected, is a series of Isotopologues each observed separated by $\Delta(m/z) = 1/z$. With typical High-Res-QTOF systems, the threshold where the peaks cannot be resolved any more lies in the low 20kDa region. Above, isotope patterns converge to one peak that resembles the average mass instead of monoisotopic mass. This effect can be observed in Fig FP4.1 where the base of the peaks of the charge envelope is much broader as resolution at this m/z value (R=35000, ~ 40mDa FWHM) would suggest. Fig FP4.3 gives a depiction of theoretical calculations demonstrating the effect.
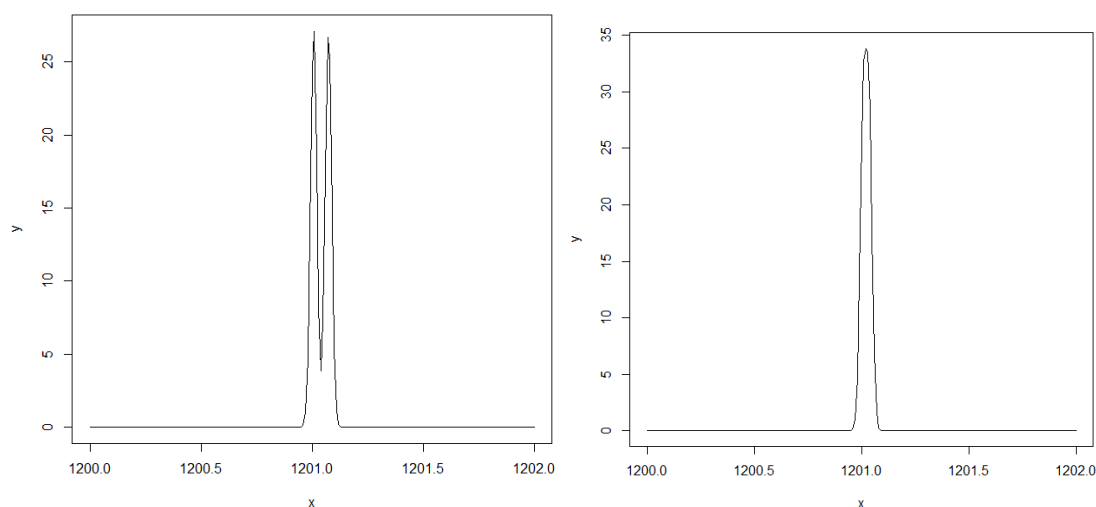
Figure FP4 3 Model of monoisotopic peak and 1x13C isotope pea. Calculated for a QTOF-System with R=35000. Left: Protein with 18kDa, 15fold charged at m/z=1201.0072. Isotope peak can be clearly distinguished. Right: Protein with 42kDa, 35fold charged no isotopic peaks can be resolved at this molecular weight due to high charge of protein moving individual peaks too close together. Source: calculation by R-script.

Isotopic peak broadening contributes to sensitivity decreasing with protein mass additionally to decreasing detector response (*vide infra*) as a respective amount of protein is distributed into a plethora of Isotopologues, of which a certain amount manifests as noise at the peak flanks rather than as signal.

Some deviation in mass accuracy can occur as result of asymmetry of Isotopologue distribution when using data interpretation software that employs symmetric gaussian smoothing. Peaks arising from Isotopologue distributions are described by a Poisson distribution. Standard mass spectrometric data interpretation software albeit usually employs gaussian based signal filter algorithms for mass spectra analysis, e.g. for data representation or smoothing steps during peak detection[21]). Some distortion of peak maxima position is introduced and S/N increase by filtering is not at maximum[41].

With increasing protein mass the isotopic distribution approaches gaussian shape [48] which reduces the effect. For smaller proteins, this effect can be minimized by choosing appropriate calibration standards, i.e. macromolecules, ideally proteins/peptides of similar masses and subjecting calibrant signals and mass spectra to identical post measurement recalibration. Furthermore, adapted smoothing and peak detection algorithms can be

applied. If not required by sample amount/concentration, instrument settings (*vide infra*) yielding isotopic resolution completely eliminate this issue.
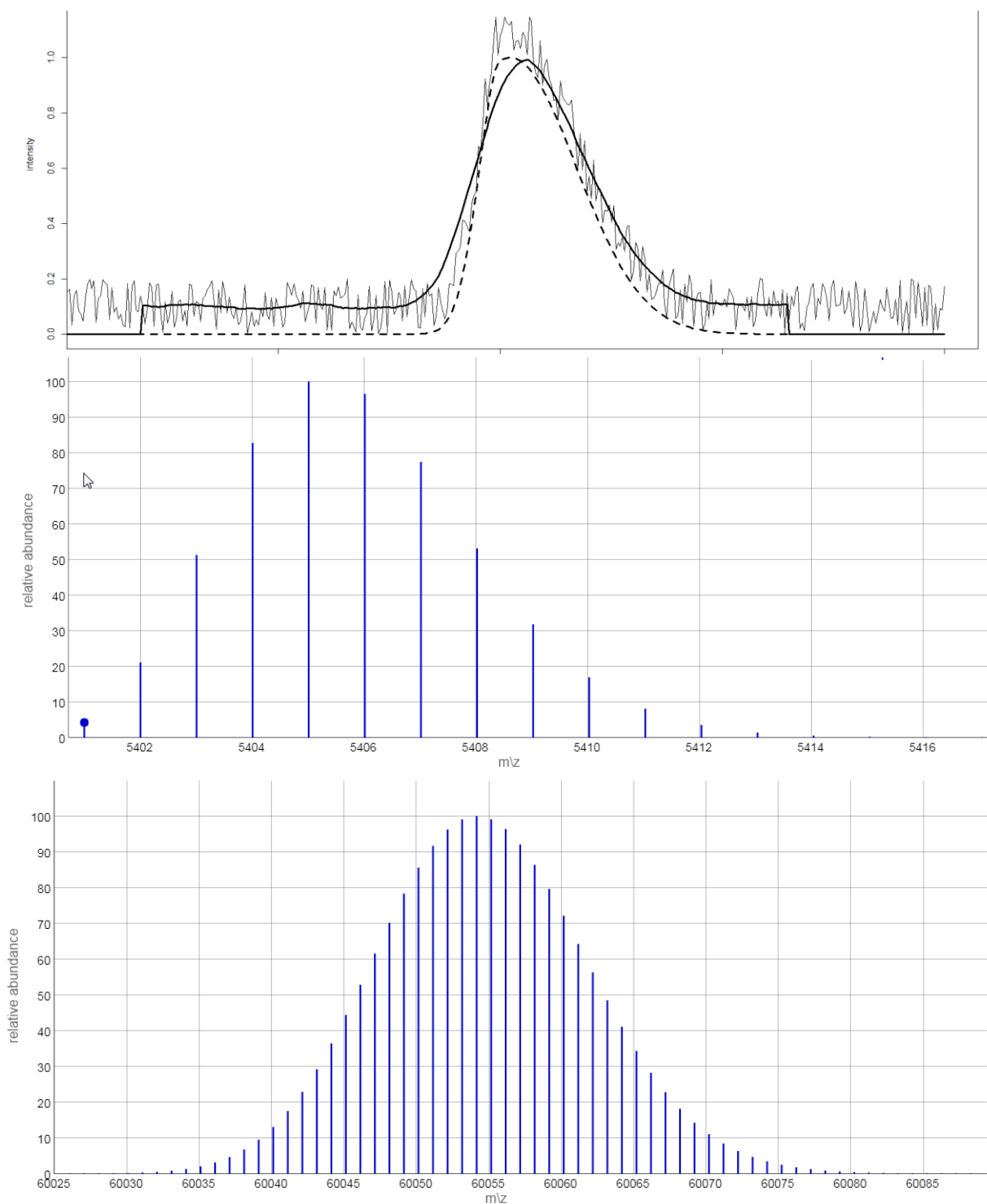
Figure FP4 4 Top: Asymmetric Isotopologue distribution and influence of gaussian filtering on mass accuracy. Dotted line) simulated theoretical peak shape by combining gaussians with σ= 0.5 and 1.5 respectively, x_0=10. Gray line) simulated noise added, solid line) Gaussian smoothed data exhibiting

divergence from actual peak maximum. Middle: Small peptide/protein with asymmetrical isotope envelope, yielding mass determination error when measured as overlapped m/z signal due to instrument settings. Bottom: large protein (~60kDa). Poisson distribution of Isotopologues is symmetrical enough to be approximated by Gaussian filter without mass accuracy distortion

Chemical noise) As charge state increases with size, so does the probability of forming mixed adducts that not only contain H+ but non-covalent charge carriers like Na+ and NH3+ or neutral adducts in some instances. This results in a decreased S/N ratio by generating multiple closely spaced m/z signals for each charge state. For example, when using a High-Resolution QTOF with R=35000 the calculated FWHM peak broadness in antibodies is 50 to 100mDa.   Measuring a 150kDa IgG with a charge state of z=100 gives a signal at m/z 1501.0072, with one proton exchanged for Na^+ at m/z 1501.22702, with NH_4^+ at m/z 1501.17746, with a H_2O molecule as neutral adduct at m/z 1501.17723.

ESI source noise) Mass spectra generated by ESI always exhibit a noticeable amount of noise, due to instabilities in the ionization process and[9], especially with low S/N spectra as encountered often due to analyte amount restrictions. Fig. FP4.5 demonstrates this with a mass spectrum of IgG. Especially high signal spikes on the flanks of the charge envelope m/z signals affect peak detection and achievable mass accuracy.
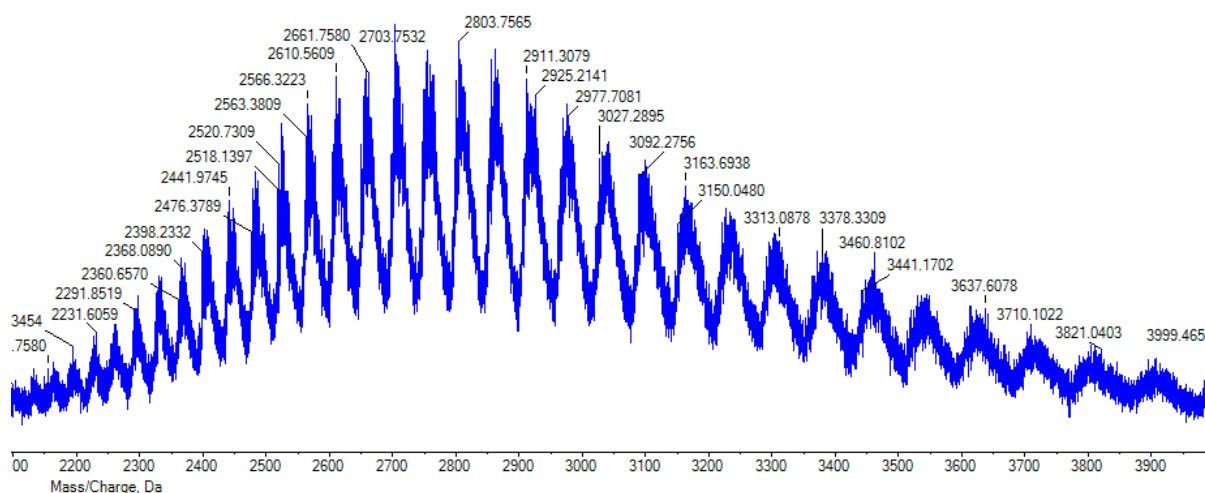
Figure FP4 5 Typical example of intact IgG mass Spectrum 1.8µg of Intact IgG on column. Note peaks are not Gaussian symmetric as opposed to the smoothing and detection filters commonly employed. With low signal intensity, ESI source fluctuations manifest as spikes that can give rise to artifacts by smoothing and peak maximum detection. Source: unpublished data, method according to Publication IX

Ion transmission in QTOF-mass spectrometers and its practical implications

For ion trajectories in the time of flight mass spectrometric compartment itself it has been demonstrated that mass resolution and mass accuracy are independent of m/z, with a slight deviation from this rule on the lower end of the m/z ratio due to increased impact of detector pulse duration variations and voltage jitter[13]. However, these studies have employed in vacuo ion generation. With an atmospheric pressure ion source, this finding gets invalidated by the processes during transition from atmospheric pressure to the vacuum of the ion path and the perpendicular extraction employed in QTOF-systems.

During the ESI-process, ions get accelerated by the ion spray potential (ISFV, ion spray floating voltage with Sciex 5600+) towards and through the orifice (opening of few μm) into of the mass spectrometers ion path. Naturally, driven by the large pressure gradient between ESI-source housing and ion path, not only ions but large amounts of gas from the ion source (usually N2) transit through the aperture forming a defocussed free jet expansion plume[18]. During this pneumatic influx, ions are further accelerated and are unfavorable diverted from the axis of the ion path. Ion velocities at this stage are comprised of a perpendicular component $v_\perp$ as result of free jet expansion and a parallel component $v_\parallel$ comprised of pneumatic acceleration additionally to the electrostatic acceleration by ion source voltage and(ISFV) and declustering potential(DP)[29]. Ion energies at this stage are proportional to m/z and amount to several hundred keV for multiply charged protein ions[12].

As only a minority of ion trajectories follow the intended ion path and the downstream portion of the ion path (Q1 to detector) is sealed by a small aperture refocusing has to be performed in order to transmit as much ions as possible through the latter orifice via a focusing ion optic. The methodology therefore was described first by Krutchinsky et al. [29] employing a quadrupole in RF mode constraining ion movement and collisional dampening by the elevated vacuum pressure in this ion path stage. The Sciex 5600+ TripleTof includes an implementation of Krutchinskys Ion-Optics (Qjet^{TM} and Q_0).

This type of interface refocusing approach has some important implications for intact protein mass spectrometry. The focusing relies on the two factors: collision with residual gas (collisional cooling/dampening) and containment/focusing by RF voltage/frequency both aspects have been studied extensively[4] It was shown in silico[29] and experimentally[12] that increasing the pressure in the interface region facilitates focusing of intact proteins. Yet this option is usually not implemented in commercial type mass spectrometers, like the Sciex 5600+ used in this thesis.
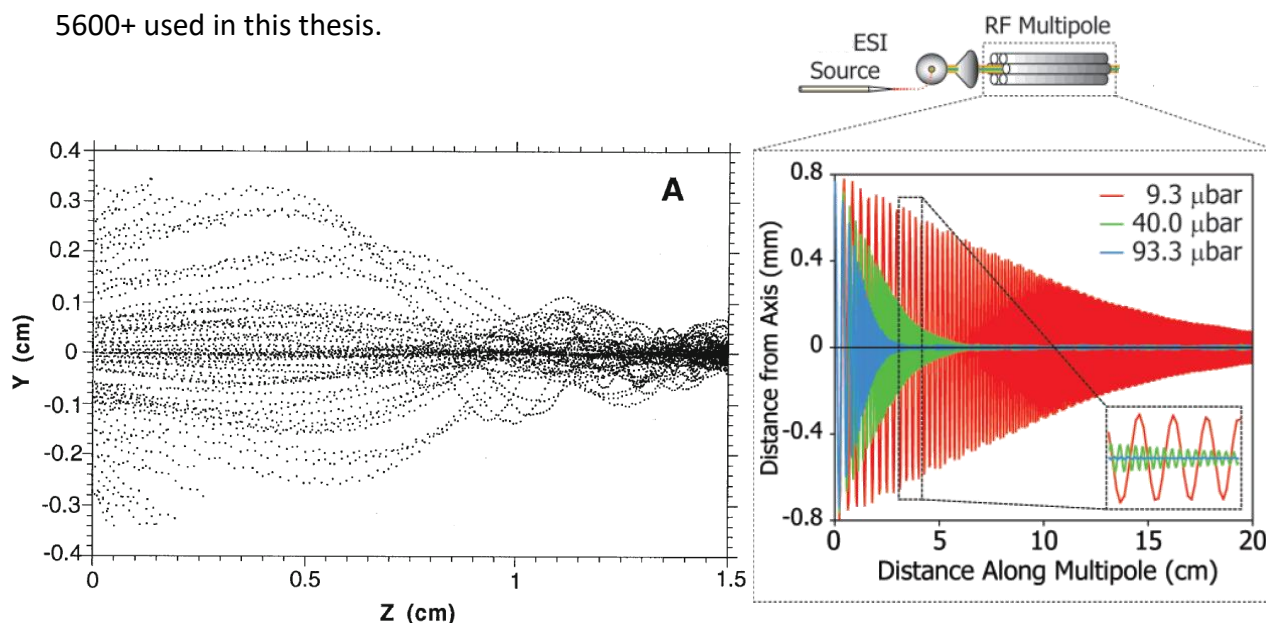
Figure FP4 6 Left calculated ion trajectories in the interface region transitioning form atmospheric pressure to vacuum of myoglobin(17kDa) ions. Reproduced from Collisional Damping Interface for an Electrospray Ionization Time-of-Flight Mass Spectrometer, Journal of the American Society for Mass Spectrometry, Volume 9, Issue 6, Pages 569-579, Copyright (1998), with permission from American Society for Mass Spectrometry. Right: Collisional dampening and residual gas pressure in the interface region. Reproduced from Protein complexes in the gas phase: technology for structural genomics and proteomics, Chemical reviews, Volume 107, Pages 3544-3567, Copyright (2007), with permission from ACS Publications.

The other option is to optimize the RF frequencies and voltages for the transmission of large ions. Hang and Coworkers [27] demonstrated this approach. As they pointed out with (Fig FP4.7) a_z set to zero as in RF-Mode, ion transmission would be governed by the AC-potential determining q_z alone(V=0), theoretically, i.e. all ions with q_z from 0 to 0.908 pass the RF-quadrupole. However as potentials are applied along the RF-Quadrupole (biased RF) in commercial systems to facilitate ion transmission down the ion path, the actual a_z-line in the Matthieu-Diagram would be a parallel to a_z=0, narrowing the operational window along 0< q_z<0.908. Ions with q_z outside of this corridor are on instable trajectories and do not pass the RF-Quadrupole. This applies especially as at the lower end of the q_z scale, as the

equations and stability diagram imply (Fig FP4.7). RF voltage amplitude and frequency must be tuned to match the respective ion m/z. yet at the same time allow for sufficient energy dissipation due to collisional dampening. As Hang et al. [27] demonstrated, due to the high kinetic energy gained by large ions during vacuum transition, high RF voltage amplitudes and low frequencies are required. Albeit on the cost of shifting small mass ions outside of the stable q_z regions, i.e. complete loss of signal for smaller ions.
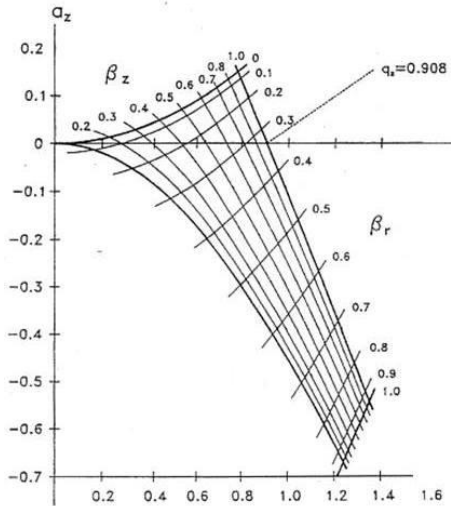
$$q_z = \frac{2zV}{\omega^2 r_0^2 m}$$

$$a_z = \frac{4zU}{\omega^2 r_0^2 m}$$
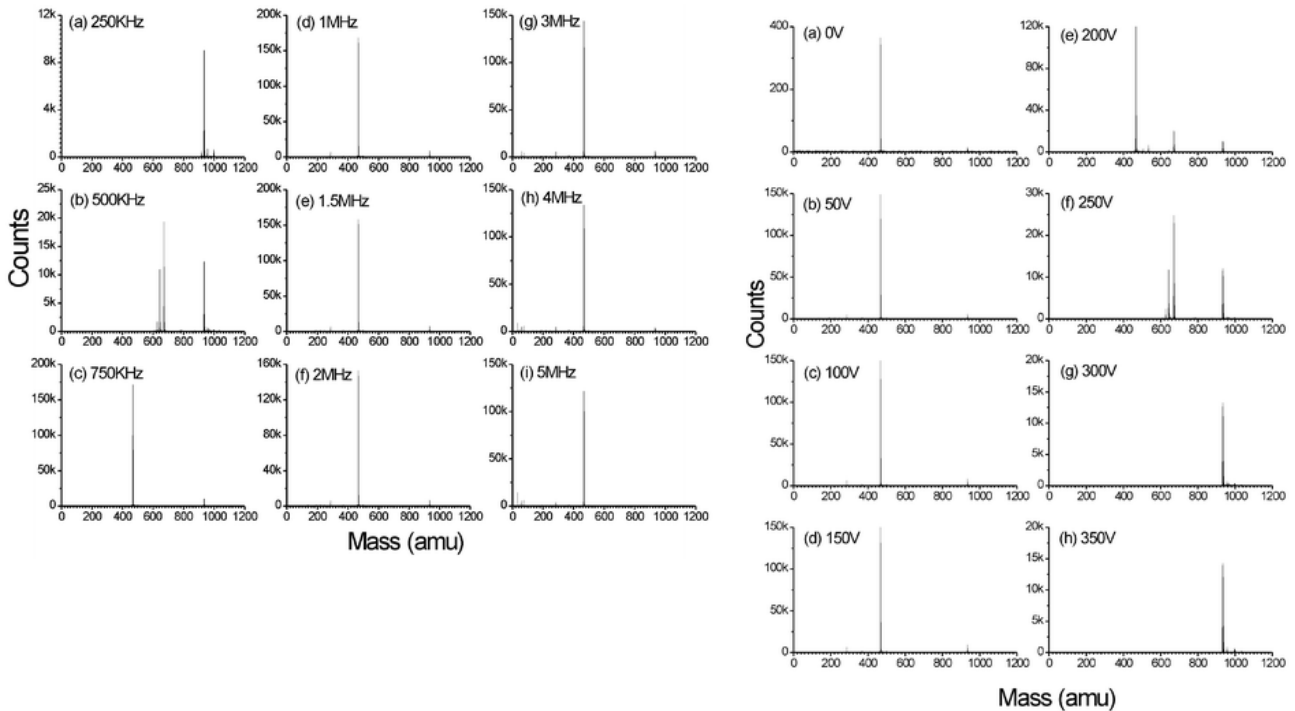
Figure FP4 7 Top-left) Matthieu diagram and solutions for a_z and q_z. For RF-mode, a_z is set to zero. Reproduced from Quadrupole ion trap mass spectrometry: a view at the turn of the century,

A further consequence of the transmission characteristics of RF-quadrupoles affects QTOF mass spectrometry of large intact proteins. As shown in the previous paragraph, transmission conditions are optimized by RF voltage and amplitude specifically for each m/z value. As not only the interface quadrupole(s) but also the Q1-quadrupole is operated in RF-Mode during a MS1-TOF-Scan this affects ion transmission efficiency and therefore signal intensity. To compensate for that, with small molecules the TOF scans are actually composite scans of subsequent transients with different Q1 settings, yielding a m/z independent transmission characteristic. With the Sciex 5600+ this is set by the Q1 transmission settings, usually employing two setting, one low m/z and one high m/z. Yet, the Q1 transmission setting is constraint by a maximum setting, e.g. m/z 1250 for the 5600+ QTOF. Species with larger m/z are affected by the diminishing ion transmission efficiency the farther they are from the optimal m/z. Albeit for the given m/z a_z and q_z might indicate transmission stability,   non-uniform transmission efficiency is also induced by insufficient focusing for a fraction of the energy dispersed ions [5, 24].

As intact proteins give rise to charge envelopes of a multiplicity of charges this leads to a distortion of the otherwise symmetrical charge envelope. The larger an intact protein gets, the further away are the individual charge state m/z ratios from the optimum m/z regarding transmission efficiency, hence, sensitivity worsens with protein size by insufficient RF-quadrupole transmission.

The distortion effect is demonstrated by Fig FP4.1. On the opposite, as slope of the function describing ion transmission efficiency of RF-quadrupoles diminishes with m/z distance to the optimum, Fig. FP4.5 exhibits no indication of envelope shape distortion indicating an almost even distributed, but with an average above m/z 2500 low absolute ion transmission.

As QTOF-systems usually possess perpendicular extraction configuration, i.e. the flight path in the TOF–drift tube is perpendicular to the quadrupole compartment (QJet,Q0,Q1,Q2),

spatial focussing, uniform trajectory vector orientation and especially horizontal kinetic energy distribution perpendicular to TOF flight path) also take effect in the time of flight compartment. Ions with velocities unfit to the TOF transmission time window overshoot the detector plates [12, 4].

Fig FP4.6 depicts this issue and influence of optimal pressure in interface region (collisional dampening). As multi-purpose commercial systems do not offer the possibility of vacuum pressure regulation, RF-quadrupole (QJet, Q0, Q1 and Q2) settings must provide for optimal ion impact distribution across the detector plates. These settings, as in the interface region are exclusive for large ions.
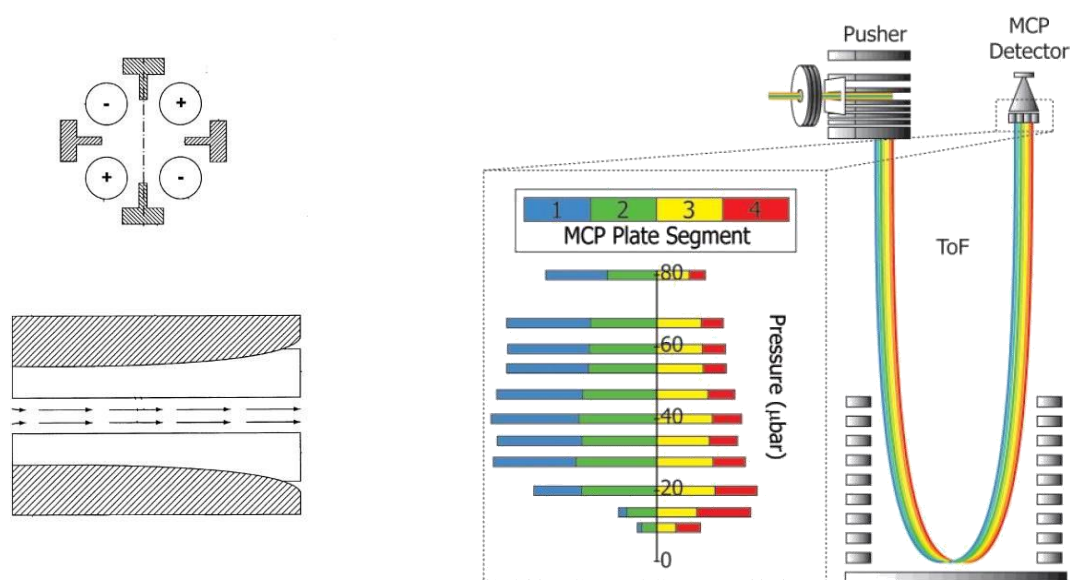
Figure FP4 8 Left: Linac: linear acceleration rods additional to Q2(collision cell). For intact proteins, collision energy applied to this rods must be increased to ~30V, Reproduced from Methods and apparatus for reducing artifacts in mass spectrometers US6909089, USPTO Right: Detector overshoot as result of kinetic energy distribution and possible remedy by collisional dampening/ uniforming. Reproduced from Protein complexes in the gas phase: technology for structural genomics and proteomics, Chemical reviews, Volume 107, Pages 3544-3567, Copyright (2007), with permission from ACS Publications.

Aforementioned collisional dampening reoccurs in the collision cell (Q2) due to the increased pressure in this region. To facilitate ion transmission, a linear acceleration potential is applied (s. Fig FP4.8). Experimentally, it can be found that the adequate linear acceleration potential for transmission of small molecules without fragmentation through the Q2 is insufficient for large molecular species. The user accessible setting for this linear acceleration voltage with the Sciex 5600+ is the collision energy (CE, actually a voltage

setting). For small molecule, CE=5V is sufficient for transmission, settings above result in collision induced fragmentation (CID), depending on ion stability and mass/size CE=10V- 40V are typically employed for fragmentation. With intact proteins, CE=30V is recommended. Settings below result in insufficient ion transmission. This requirement results in further incompatibility of parameters for simultaneous small molecule and intact protein mass spectrometry, as with the large CE, small molecules are fragmented.

The essence of this paragraph for the experimental practice of this thesis is that the ion path voltages and frequencies of quadrupole time of flight mass spectrometers must be tuned to provide for efficient transmission and detection of large species like intact proteins, as these systems are usually optimized for and provided with standard settings for small molecules and peptides. Actually, with standard settings, no signal for intact proteins could be detected at all. For the employed Sciex 5600+ TripleTof, the operator is provided with the "intact protein script" adjusting aforementioned settings, leaving only ion source parameters, declustering potential (DP 150-300V for large proteins), collision energies (CE, 20-30V) to be optimized by the operator for the respective analyte of interest. However, as described previously, it has to be kept in mind that large and small molecule ion path tunings are mutually exclusive. Hence, TOF mass axis calibration cannot be performed with the usual low molecular species (e.g. polypropylenglycol, peptide standards), as they are usually not detected. Yet, if no eligible calibration standards are available, mass axis calibration can be done sufficiently in small molecule mode prior to activation of protein settings. Furthermore, it is not possible to detect proteins and small ligands or other respective matrix components in the same experiment, except when differences in measured protein mass is used to derive the ligands mass, subjected of course to the reduced mass resolution and mass accuracy conditions characteristic for large ions.

Properties of mass spectrometric detectors important for protein mass spectrometry

Mass spectrometric ion detection is primarily based on one of three mechanisms: direct charge generation(Faraday Cup), inductive detection (Orbitrap, FTICR) and secondary/cascade electron detection (MCP, EMH)[22]. Of these, only the latter two are relevant for mass spectrometry of large macroions, due to their sensitivity. As the author's work was based on a Sciex 5600+ TripleTof mass spectrometer, which detector utilizes a

stack of two multichannel plates (MCP), the properties of MCP detectors and implications on the experimental approach in protein analysis shall be focused on in the following.

Multichannel plates are fused arrays of drawn glass tubes each consisting of an etchable core glass cladded in nonetchable lead glass. After cutting the array at an angle (~8°) to provide an angle for later perpendicular ion impact relative to MCP disc surface, the core class is removed by etching and the remaining lead glass surface is reduced in a hydrogen surface resulting in a conductive, yet resistive surface, yielding a continuous potential gradient (overall resistance : 10^9 Ω) rather than a stepwise as in a discrete dynode[3]. The overall 1-10mm thick arrangement which exhibits approximately 10000-450000 channels per square centimeter, depending on channel diameter (10-100µm), each resembling an individual dynode[47] albeit the impact of all resulting output electrons is a single composite signal.
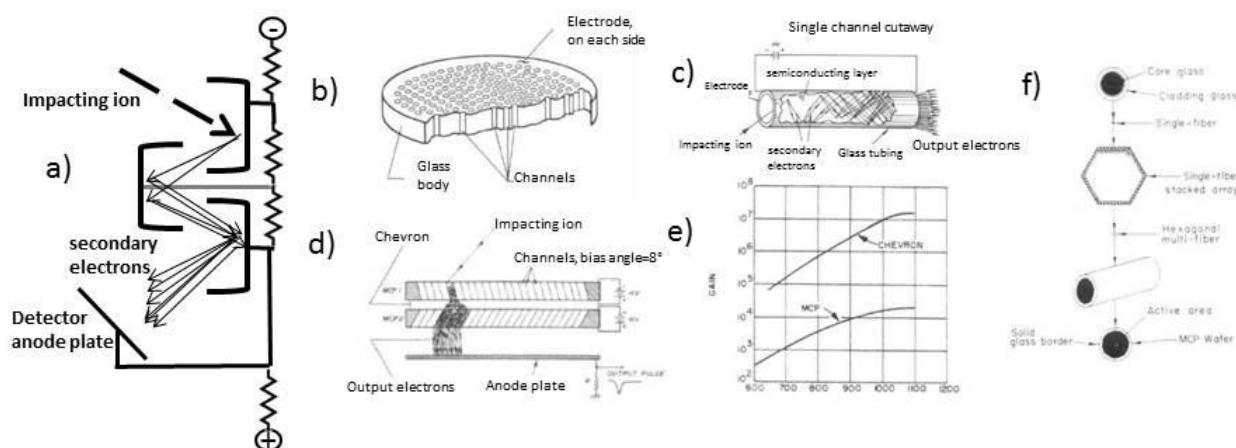
Figure FP4 9: a) Discrete Dynode. Resistors between individual plates generate potential gradient necessary for secondary electron acceleration. b) Multichannel plate. c) Single channel of straight walled MCP; ion impact angle has to be non-perpendicular to surface as usually achieved by angled cut of MCP disc. Potential is provided by semiconductive/resistive lining of glass channel. d) Chevron arrangement. Reproduced from Nucleic Acids Research, Volume 34, Issue 19, 1 November 2006, Pages 5402–5415, Copyright (2006), with permission from Oxford academic.

The gain of a single MCP is limited to 10^3 to 10^5 by positive ion feedback mechanism. At the output region, even under low pressure of mass spectrometers, residual gas molecules are ionized by the high energetic (~30-100eV) output electrons and subsequently accelerated towards the MCP cathode, interfering with and limiting generation of secondary electrons. This restriction in gain can be circumvented by bending[25] or twisting of channels[38] (both

difficult to manufacture) or, as commonly implemented by manufacturers, a chevron arrangement of two MCP as depicted in Fig FP4.9, allowing for gains of up to in the 10^7 order[14].

For detection of large molecular ions, MCPs have specific properties with respect to sensitivity that have to be considered. The gain of the MCP detector array, expressing the number of output electrons per impacting ion, is dependent on various factors such as angle of impingement, acceleration voltage across the MCP plates (often named detector voltage or MCP voltage in commercial mass spectrometers), but also on ion species, molecular weight, charge state and velocity thereof[34]. Influence of molecular weight and impact velocity/energy was established employing MALDI-TOF[23, 33, 10] rather than ESI-QTOF systems, as MALDI yields mostly singly charged ions and MALDI-TOF systems contain no quadrupoles, which would   bias sensitivity by its m/z dependent transmission characteristic[16].

Geno et al[23] derived a function calculating the secondary emission coefficient γ, where ion mass contributes linear, ion velocity contributes exponentially. Aforementioned secondary emission coefficient γ describes the probability of a single ion to produce secondary electrons upon impact, i.e. being detected by a MCP (P=1-e$^{-\gamma}$ derived by Beuhler et al.[5]). Even for the in protein terms lightweight bovine insulin peptide(5.8kDa) an acceleration voltage of 40kV would be necessary[23]   based in calculations, yet are difficult to achieve experimentally[23]   with sufficient supply stability for accurate TOF-mass spectrometry. It should be noted, that the probability functions uses ion mass and not mass to charge ratio.
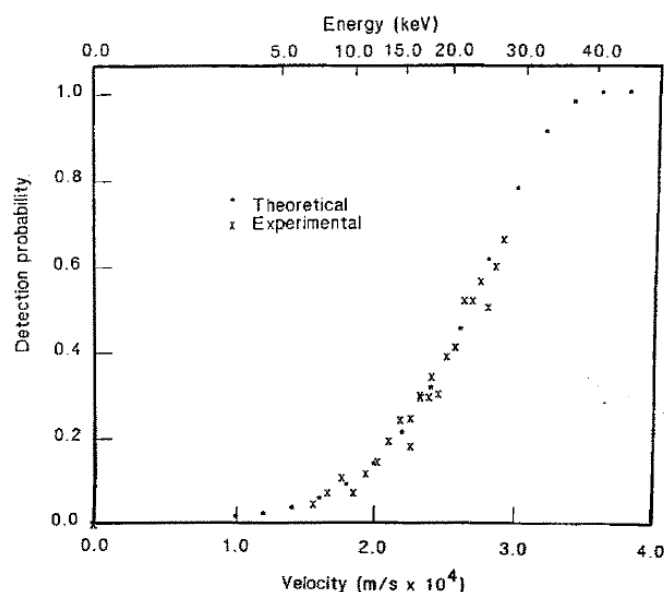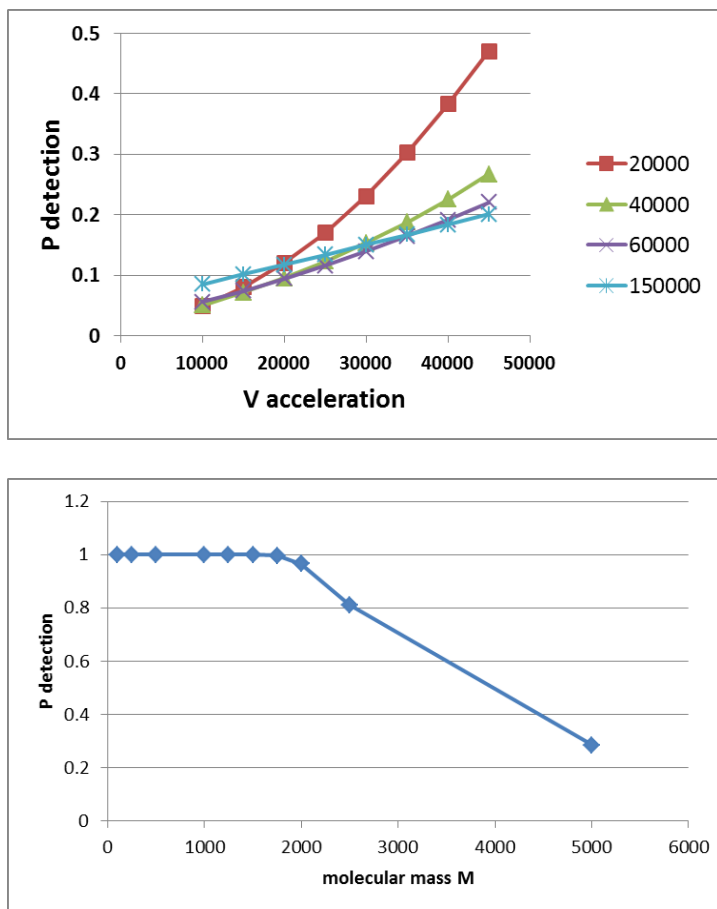
Figure FP4 10 Left) Overlay of measured and calculated detection probability of bovine insulin vs. acceleration voltage. Reproduced with permission from International Journal of Mass Spectrometry and Ion Processes, Issue 92, 1989, Pages 195–210, Copyright (1989), with permission from Oxford academic. Right-top) Detection probabilities for different molecular weights (up to 150kDa for intact antibodies) and acceleration voltages calculated according to formula derived by Geno et al[23]. Right-bottom) Detection probability as function of molecular mass, acceleration voltage 15kV (e.g. Sciex 5600+)

The formula of Geno et al [23] was extended by Westmacott et al [44, 45], deriving function for secondary electron yield $y_e = 2.8 \times 10^{-18} mv^{4.3}$. This has been confirmed by others [37, 31, 10]], with a range of velocity exponents form 3.2 to 4.4 attributed to differences in mathematical fit and MCP material and setup.

These findings indicate that MCP secondary electron multipliers are indeed inferior for macroions like proteins and other detectors have been devised for mass spectrometry like conversion dynode produced secondary ion detection[39] or cryogenic superconducting tunnel junction(SJT) detectors, the latter achieving 100% detection probability for individual

macroions [20], [45], albeit on the cost of more intricate instrumental setup. The drawbacks and complicacies of these alternate detection technologies manifest themselves in the observation that commercially available research grade TOF instruments (Sciex, Bruker, Agilent, Waters) even targeting large molecular analysis still are designed with MCP based ion detection. Above findings are directly relevant to experimental work. As depicted in Fig FP4.10, for larger proteins detection probability and therefore sensitivity of mass spectrometric detection is severely reduced for MCP-detectors. Increase of acceleration voltage, if actually permitted on user level access of commercial mass spectrometric instrumentation is no remedy for this problem as shown in FP4.10. Proteins as small as 20kDa would require >80kV acceleration voltage, let alone intact antibodies, where 500kV would be necessary, both difficult to achieve. It can also be seen that below 2000 Dalton detection probability reaches P=1, indicating that virtually all ions reaching a MCP channel are detected. In contrast to macromolecules, where detection efficiency drops off starting with 20mers concerning peptides/proteins.

What has been demonstrated as beneficial regarding detection of proteins or other macromolecules is a reduction of the MCP voltage, i.e. potential across the multichannel plate by 5-10% (e.g. Sciex 5600+: reduction by 100V). This measure is apparently counterintuitive, as rather an increase of MCP voltage to release more secondary electrons per impact as last resort to achieve sensitivity or mandated by degradation of MCP detector has to be applied. According to Sciex, S/N enhancement by MCP voltage decrease results from a disproportionate decrease of noise generated by singly charged species, while response to multiply charged ions is unaffected.

Time digital converter – Sciex 5600 TripleTOF– impact on experimental settings

Apart from other vendors (Bruker, Agilent, Waters) the Sciex 5600+ employs a four channel time to digital converter (TDC) operating at 40GHz for data acquisition instead of an analog-digital (ADC) converter. While ADC measure the current on the detector anode induced by the secondary electrons output and thereby gain information on the number of ions impacting in a given time window, the TDCs record the impact of individual ions qualitatively. Therefore, TDC have to operate on higher frequencies (40GHz vs 2-4GHz for

ADCs) to provide for quantitative (intensity, peak height, peak shape) information, otherwise multiple ions impacting in a too short time frame are detected as one ion, a process called pileup. To do so, the time after the start signal (simultaneously with rising edge of pusher voltage) is divided into small, picoseconds long time bins. If an ion fires the MCP within such a time bin, the arrival time is recorded and associated with that time bin (multi-hit TDC). To form a spectrum, this process called transition is repeated, ten to a few thousand times and data are combined to yield one spectrum, depending on user set spectral acquisition rate. This requires a certain temporal dilatation of the ion packets of identical m/z. As individual microchannels are thereunto inactive for a dead time in the order of µs during charge restoration, sensitive detection relies both on distribution lateral in longitudinal distribution of the ion packet.

As described in the previous section, with MCPs the detection probability of an impacting ion triggering a pulse on the detector anode is decreasing with ion velocity, which in turn is dependent on the ion mass during TOF acceleration. As shown in Fig FP4.10, the probability at 15kV acceleration voltage as fixed in the Sciex 5600+ TripleTof is around P=0.1. As detection with a TDC is a binary event, the probability to detect all of them in their respective time bins is significantly lower ($P_{all} = (P_{single})^n$) and most of the ions are effectively omitted during detection. This scenario can be improved, if multiple time bins are combined into one, increasing the probability that one of the incident ions triggers detection of this combined bin, albeit on the cost of spectral resolution, as the underlying time grid is set less granular. For small molecules, time bins are usually 1-4, for macromolecules, up to 80 time bins can be combined.
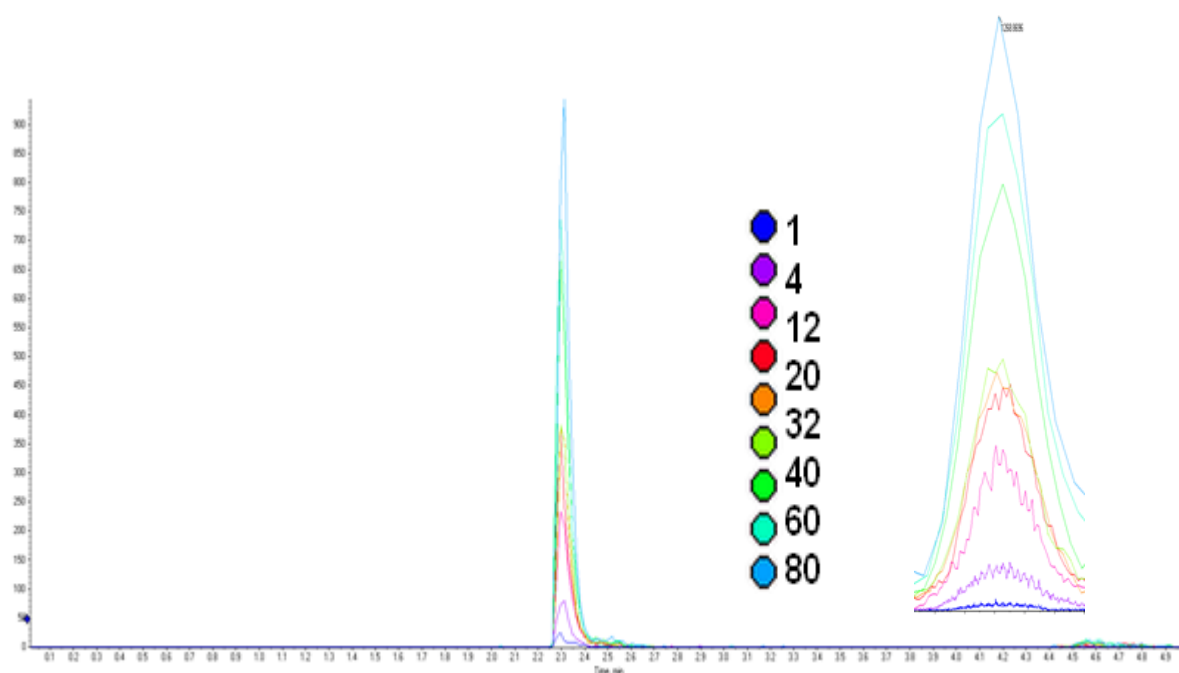
Figure FP4 11 Above: MS-chromatogram of Trypsin demonstrating influence of combining time bins from 1-80 DC bins on sensitivity. All runs were performed as C4-RP trap&elute in backflush mode to remove polar and low MW compounds. By combining multiple time bins peak height could be increased by a factor of 20-30. Right: Individual mass spectral peak, demonstrating the effect of combining time bins on resolution and granularity of mass spectra

Experimental approaches to intact protein mass spectrometry with Sciex 5600+

The following considerations were employed in publications IV, IX, X, and XI to conduct sensitive protein mass spectrometry:

a) For rather pure samples, direct infusion with low flow rates in the µl/min regime and composite spectra generation by MCA can be beneficial. Otherwise, reversed-phase trap and elute in backflush mode should be employed, in order to purify the protein from all nonvolatile low molecular compounds causing either ion suppression or increased level of unwanted adducts (e.g. Na$^+$) and to achieve a sharp and high peak, as with mass chromatogram peak height is of importance for sensitivity and high S/N. In case of protein mixtures, adequate chromatography is mandatory as overlapping charge envelope complicate spectra.

b) Trap and elute was performed employing C4-RP-trap columns, these exhibit good retaining of proteins, yet allow removal of low molecular weight and polar contaminants.

c) Declustering potential (DP): Increased values reduce amount of adduct species of respective proteins. The absolute signal intensity reduction caused by higher DP is overcompensated by depopulation of mass spectrum, overall increasing sensitivity.

d) Collision energy increase: Usually higher voltages are necessary, e.g. for Sciex 5600+: 30V, a voltage where many small molecular species would already be fragmented.

e) Q1 transmission window: 100% on m/z=1250 (instrument limit) to achieve maximum transmission for multiply charged protein ions, usually with m/z from 900-2000.

f) Intact protein mode script: Tuning ion path voltages and frequencies to accommodate for efficient transmission of macro ion beams should be employed

g) Decrease of MCP potential by 100V considerably increases S/N ratio.

h) Time bins: Increasing with target mass (not m/z), from 10 to 80 bins combined.
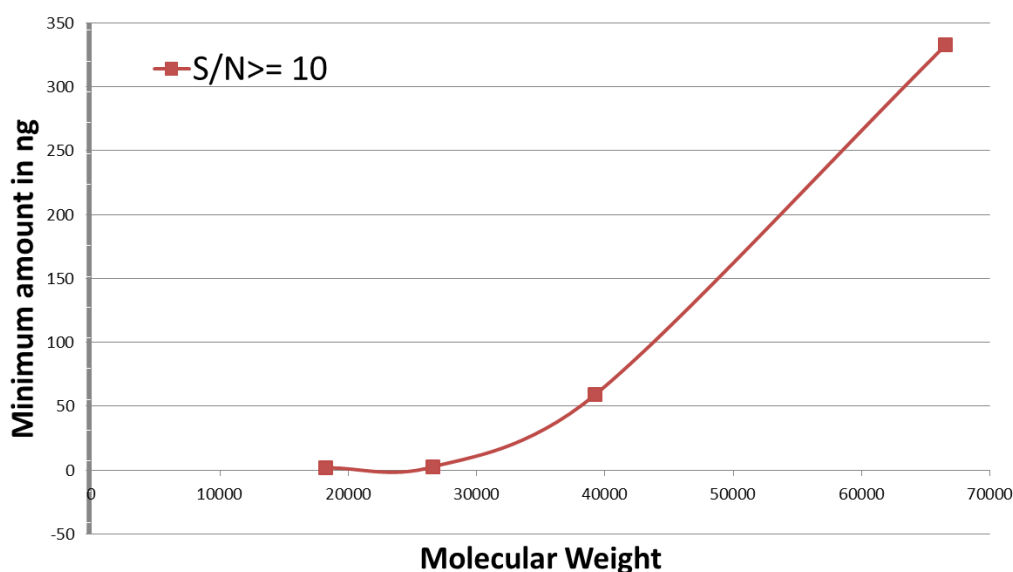
Figure FP4 12 The minimal absolute amount of protein in ng on column to yield a S/N > 10 versus molecular weight of protein. For intact antibodies (~150kDa) 2-3µg on column are required. Results correspond to the literature opinion. Proteins used for this demonstration graph were Lactoglobulin(18kDa), Triosephosphate isomerase(26kDa), Glyceraldehyde-3-Phosphatase, Aldolase(39kDa), Human serum albumin(66kDa)

Objective and results of Publication IX and XI

Both Publications are focused on protease conjugated gold nanoparticles (IX: Pepsin; XI Papain) and their utilization as heterogeneous proteolytic catalysts for proteomic workflows and in middle-down/up proteomics[50]. In general, heterogeneous enzyme catalysts provide for easy and fast removal of catalyst by centrifugation or filtration, or if magnetic carriers are employed, by magnetic separation[42]. Although dependent on the carrier system employed[49, 30], immobilization can enhance enzymatic stability and reaction rate, especially if carrier systems in the nanoscale are chosen [28].

Additionally to exemplifying the classic bottom up approach, Publications IX and XI demonstrated the application of gold nanoparticle proteases in middle-up proteomics. The latter term refers to a proteomic workflow starting with enzymatic disassembly of large proteins inaccessible by Top-Down proteomics into larger peptides (>5kDa) and proteins that are readily accessible by such an approach

[19, 50, 26]. This methodology enables faster characterization, e.g. determination of post translational modifications and sequencing. Regarding intact IgG antibodies, the most common approach is to cleave them into their subunits Fab and Fc prior to intact protein mass spectrometry.
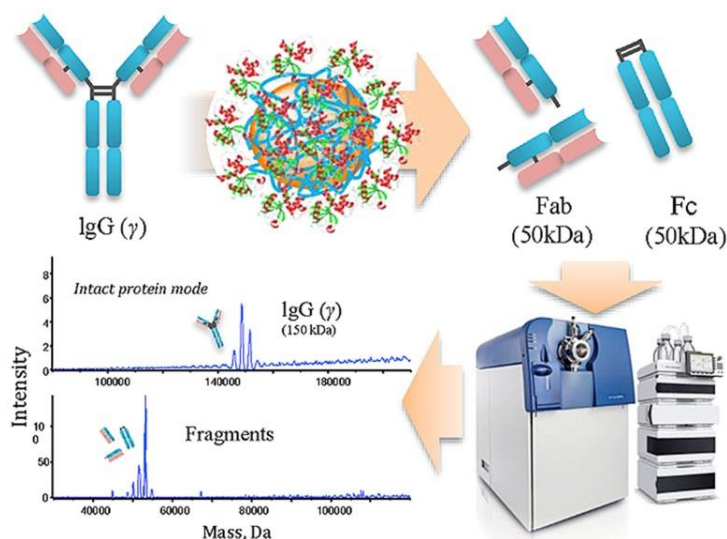
Figure FP4 13 Middle down proteomics employing a nanocarrier with immobilized proteases. Reproduced from Publication XI Papain-functionalized gold nanoparticles as heterogeneous biocatalyst for bioanalysis and biopharmaceuticals analysis, Analytica Chimica Acta, Volume 963, Pages 33-43, Copyright (2017), with permission from Elsevier.

Mass spectrometric top down proteomics have advantage of being less time consuming as protracted protein digestions (usually 12h). Yet, as shown in the previous sections, a major factor restricting the applicability of mass spectrometry on intact proteins is the instrumentation itself with respect to sensitivity, resolution and mass accuracy.

Consequently, a middle down approach yielding lower MW fragments increases sensitivity exponentially (see Fig FP4.10) thereby reducing required concentrations/absolute amounts, especially for ex vivo bioanalytical purposes (less important in analysis of biotechnological products due to large amounts generated). Also, with a middle up approach, digesting IgGs to their smaller subunits, mass accuracy of intact mass spectrometry is significantly increased, in case of the Fc/2(~24kDa) subunits even isotopic resolution can be achieved.

For Publication IX and XI, intact protein mass spectrometry using a high resolution µESI-QTOF-System was successfully employed to demonstrate the catalytic activity of nanoparticle conjugated proteases Both works also confirmed the utility of gold nanoparticle conjugated proteases for a middle-up approach with heterogeneous proteolysis providing fast(high throughput amenable) and convenient sample preparation protocol for analysis of antibodies, circumventing instrumental restrictions of contemporary commercial QTOF-mass spectrometers for increasing result fidelity.

Objective and results of Publication X

Publication X was masterminded by the Group of Prof. Laufer, Institute for Pharmaceutical Sciences, University of Tübingen. This work provided covalent bonding inhibitors of c-Jun N-terminal kinase 3 (JNK3), a member of the mitogen activated protein (MAP) kinase family. The inhibitors are based on a pyridinylimidazole scaffold exhibiting an electrophilic warhead covalently bonding to Cys-154.

 JNK3 is a part of the stress activated SAP/JNK signaling pathway and responsible for activation of transcription factors c-Jun/AP-1 by phosphorylation[7]. As this transcription factor is involved in regulating apoptosis and misregulation is shown to affect cell proliferation it is well established as oncogene, indeed the first oncogene discovered[43]. With JNK3 being located upstream in the signal cascade, it is a target for anticancer strategies. Also, knock-out experiments could demonstrate that it represents a promising therapeutical target for neurodegenerative disorders like Alzheimer's, Parkinson's and Huntington's desease [1], especially as its expression being specific to neuronal and cardiac tissue as well as testis as opposed to its ubiquitous isoforms JNK1 and JNK2[6, 7].
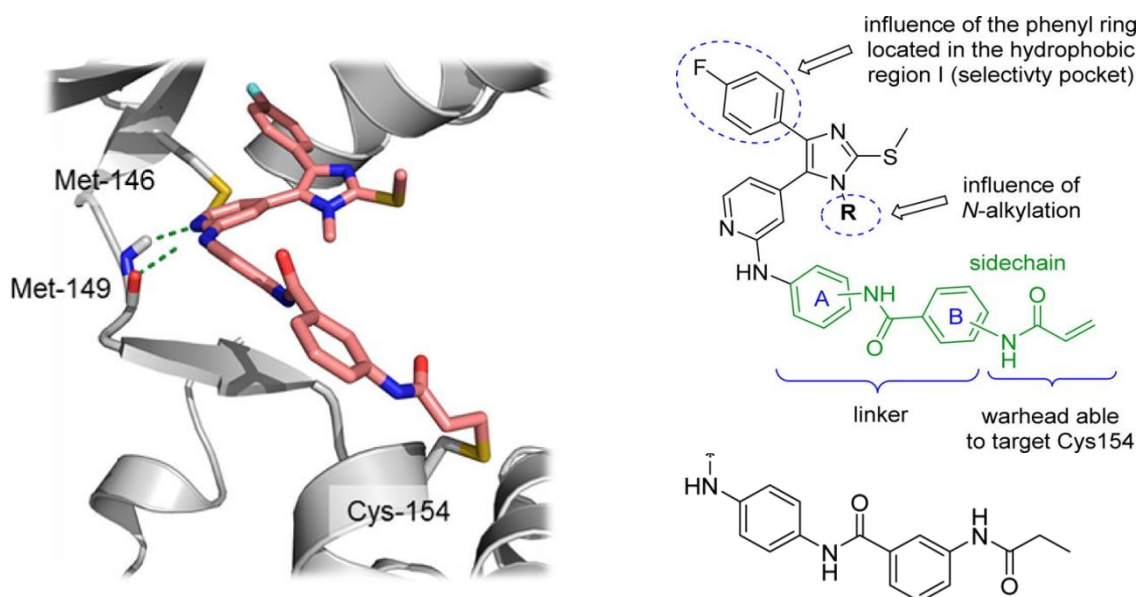
Figure FP4 14 Left) Docking of pyridinylimidazole derivative from Publication X into crystal structure of JNK3(PDB: 1PMN) Right) Scaffold structure and moieties studied. Right bottom: Example of Inactive (noncovalent) warhead used for comparison.   Reproduced from Publication X, Tri- and Tetrasubstituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3, Journal of Medicinal Chemistry, 60 (2), pages 594–607 Copyright (2017), with permission from ACS Publications.

Publication X studied affinity and specificities regarding covalent inhibition of JNK3 of various derivatives of a pyridinylimidazole scaffold(s) (Fig FP4.14). Biological enzymatic inhibition (IC_{50}) was evaluated with enzyme-linked immunosorbent assays (ELISA), in vitro metabolism was accessed via LC/MS. The task of the author of this thesis in Publication X was to confirm the covalent bonding by mass spectrometry.

Examining covalent protein modification can be performed by two approaches, bottom-up or top down. The bottom-up-approach is the classic approach of proteolytically digesting the protein to small peptides, analyzing them by reversed phase LC-MS/MS, identifying the peptides by their fragmentation pattern, and reconstructing the intact protein from these results. Modifications can be identified on peptide fragment level and, with the low mass of the analyzed, lower accuracy mass spectrometric systems, as triple quad (QqQ) systems can be employed. Actually, covalent modification analysis is implicit to proteomic bottom-up workflows, as many proteins are covalently modified in vivo (PTM).

Yet, bottom up has some constraints for covalent analysis of inhibitors. Relatively large sample amounts have to be subjected to lengthy digestion protocols usually 8-12h digestion additional to pre- and postprocessing(purification) steps. Also, proteolytic protocols are not

universally applicable as their feasibility depends presence and distribution of respective protease cleavage sites, which in worst case may result either in peptides too long to be accessible by some mass spectrometric systems when operating in small molecule mode as usual in proteomics, or in short ambiguous fragments of few amino acids attached to the ligand, prohibiting to pinpoint the exact location of the amino acid residue derivatized with the covalent binder. The latter constraint however applies also to top-down approaches while the former (long peptides) could be resolved by intact protein analysis of the long peptide fragments, albeit requiring ample unambiguous peptides for identification. Furthermore, covalent modification may alter or prohibit proteolytic cleavage, or the relevant covalent bond between protein and ligand may even be instable under proteolytic conditions.

In contrast, top down approaches are feasible for fast and high throughput and can in principle be adapted for automated library screening. For in vitro experiments with pure recombinant protein, chromatography(minutes to hours) can be replaced by short trap and elute protocols employing reversed phase trap columns (seconds), where the time constraint is usually of technical nature(autosampler) Sample amount is dependent on size as elaborated previously (Fig FP4.10), but is usually in the ng to μg (absolute amount on column) range.

With top down approaches, noncovalent as well as covalent interactions can be examined. Non covalently bound ligands require careful tuning of ionization and ion path parameters (see also Introduction to Publication III/XII) and bear the risk of false positive evaluation based on nonspecific association. Still, noncovalent protein-ligand interactions are inaccessible by bottom up approaches, as interactions dependent on secondary and tertiary structures are disrupted during proteolysis. In contrast, assessment of covalent ligand bonding requires the opposite instrumental settings, i.e. dissociation of noncovalent interactions must be assured. This is achieved by appropriate temperature and gas flow settings in the ESI ion source ensuring complete desolvation and by sufficiently high declustering potential at the ion path entrance.

For Publication X, a μLC (60μl/min) chromatographic gradient protocol using a monolithic PS/DVB based capillary column (0.5 mm ID) was established. Ionization was performed with

a µLC emitter tip for increased sensitivity. For exact assay conditions, the reader is referred to supplemental material of publication X. Albeit employing larger volumes (200µl of 1pmol/µl JNK3) for convenience, actual injection consumed only 5 pmol(210ng) absolute amount of recombinant JNK3 (42kDa) per assay. This demonstrated the feasibility of such a top-down approach for inhibitor library screening, as 1000 candidates could be screened within less than 5 days, consuming approx. 200µg of recombinant JNK3. This could further be optimized by sample volumes, automated handling, and pre concentration via trap and elute chromatography.

As previously indicated, exact bonding location of positive covalent inhibitor candidates must be evaluated to confirm specific binding. In theory, this could be assessed directly with covalently inhibited JNK3 by EDD /ETD fragmentation similar to PTM analysis[46]. Yet the Sciex 5600+ TripleTof offers only CID fragmentation, inferior for MS/MS analysis of large proteins, suffering from low sequence coverage[36] and limited structural information [17, 2]. For covalent modifications, a special caveat is reported: Long Peptides/Proteins with covalent modifications tend to lose this modification upon CID. Simultaneously, most backbone fragmentation of peptide/protein is suppressed in favor of dissociation of the modification, masking location information and hampering peptide fragment identification. Therefore, direct CID fragmentation of intact protein was deemed unfeasible for proving covalent binding location. When employing a bottom up approach, CID of modified peptides could of course be employed as usual.

Instead, as in publication X, in house expressed and purified recombinant JNK3 was employed, it was the most straightforward approach to use a JNK3-C154A mutant generated by directed mutagenesis. This mutant has the active residue Cys154 reacting with the warhead of the inhibitors by Michael-addition replaced by an inactive alanine. As expected, none of the inhibitors found to bind with the wildtype JNK3 was reactive towards the JNK3-C154A mutant.

Overall, the covalent binding to JNK3 could be successfully proved by intact mass spectrometry. Masses of covalently inhibited JNK3 could be determined with less than 10ppm mass error, molecular masses of calculated inhibitors from the charge envelope could be determined with subdalton mass accuracy. To cross check, noncovalent inhibitors

were incubated with the target, without yielding modified JNK3 (i.e. non-covalent inhibitor-JNK3 associates).

Table TP4 1 Overview of results. Reproduced from Publication X supplemental material, Tri- and Tetrasubstituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3, Journal of Medicinal Chemistry, 60 (2), pages 594–607 Copyright (2017), with permission from ACS Publications.

Compound #	active warhead	Protein Average Mass [Da]			Reconstructed inhibitor	
		Calculated	Measured	error ppm	Expected	Found
7	Yes	42784.5576	42784.2239	-7.8	578.328	578.6617
8	no	42205.8959	42205.5105	9.1	0	0
15	no	42205.8959	42205.9047	0.2	0	0
21	yes	42690.471	42690.103	-8.6	484.2071	484.2071
23	no	42205.8959	42206.342	10.5	0	0

Objective and results of Publication IV

Publication IV is classified within this thesis as a minor scientific contribution (s. List of author contributions). Therefore, it shall be only briefly described. Publication IV examines halogen bonding to the Met146 residue with regard to affinity and selectivity in the human kinome. Ligands substituted with Chlorine, Bromine and Iodine are compared to unsubstituted (H-) compounds. The author's contribution to publication IV is the intact protein analysis of C‑Jun N‑Terminal Kinase 3 (JNK3) and mutants demonstrating the selectivity of methionine, namely M146A, M146L, and M146T.

Mass accuracy (<2ppm) and resolution (35000) of the Sciex 5600+ TripleTOF allowed for subdalton accurate MW determination. This enabled confirmation of mutation success in a straightforward and fast trap and elute LC-MS analysis, with minimal amounts of protein required (50ng per measurement) and without the need for a time-consuming bottom up proteolytic workflow and subsequent LC-MS/MS analysis. The experimental mass spectrometric and chromatographic conditions are described in supplemental material of Publication X.

Table TP4 2 Results from Publication I. Mass error is in the lower mDa and accuracy of 4 ppm and lower as usually would be expected of low MW mass spectrometry could be achieved.

	M146A	M146L	M146T
M_{calc}	42145.7821	42175.8084	42175.8084
M_{meas}	42145.5927	42175.7227	42175.7227

Error mDa	-189	-86	-86
Error ppm	-4.5	-2.0	-2.0

Abbreviations

FWHM: Full width at half maximum of peak

FTICR: Fourier transform ion cyclotron resonance

MCP: Multi channel plate

EMH-Horn: electro multiplier horn

MALDI-TOF: Matrix assisted laser desorption ionization time of flight mass spectrometry

Da, kDa: Dalton, Kilodalton

kV: Kilovolts

LOD: limit of detection

LOQ: limit of quantification

DP: Declustering potential

MCA: Multi channel alignment

SJT: superconducting tunnel junction

ADC: Analog digital converter

TDC: time digital converter

ELISA: enzyme-linked immunosorbent assays

PS/DVB: polystyrene-divinylbenzene copolymer

CID: collision induced fragmentation

References

[1]   Xanthi Antoniou, Mattia Falconi, Daniele Di Marino, and Tiziana Borsello. Jnk3 as a therapeutic target for neurodegenerative diseases. *Journal of Alzheimer's Disease*, 24(4):633–642, 2011.

[2]   Andrea Armirotti, Umberto Benatti, and Gianluca Damonte. Top-down proteomics with a quadrupole time-of-flight mass spectrometer and collision-induced dissociation. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, 23(5):661–666, 2009.

[3]   RM BEASLEY, V DUCHENOIS, R POLAERT, and D WASHINGTON. Technology of channel plate manufacture(glass multichannel perforated plate manufacture technology, discussing drawing and hollow fiber techniques). *ACTA ELECTRONICA*, 14:201–224, 1971.

[4]   Justin LP Benesch, Brandon T Ruotolo, Douglas A Simmons, and Carol V Robinson. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chemical reviews*, 107(8):3544–3567, 2007.

[5]   RJ Beuhler and L Friedman. Threshold studies of secondary electron emission induced by macro-ion impact on solid surfaces. *Nuclear Instruments and Methods*, 170(1-3):309–315, 1980.

[6]   Marie A Bogoyevitch and Bostjan Kobe. Uses for jnk: the many and varied substrates of the c-jun n-terminal kinases. *Microbiology and Molecular Biology Reviews*, 70(4):1061–1095, 2006.

[7]   Marie Cargnello and Philippe P Roux. Activation and function of the mapks and their substrates, the mapk-activated protein kinases. *Microbiology and molecular biology reviews*, 75(1):50–83, 2011.

[8]   Adam D Catherman, Owen S Skinner, and Neil L Kelleher. Top down proteomics: facts and perspectives. *Biochemical and biophysical research communications*, 445(4):683–693, 2014.

[9]   Nadja B Cech and Christie G Enke. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass spectrometry reviews*, 20(6):362–387, 2001.

[10]   Xiaoyu Chen, Michael S Westphall, and Lloyd M Smith. Mass spectrometric analysis of dna mixtures: instrumental effects responsible for decreased sensitivity with increasing mass. *Analytical chemistry*, 75(21):5944–5952, 2003.

[11]   Igor V Chernushevich, Alexander V Loboda, and Bruce A Thomson. An introduction to quadrupole–time-of-flight mass spectrometry. *Journal of mass spectrometry*, 36(8):849–865, 2001.

[12]   Igor V Chernushevich and Bruce A Thomson. Collisional cooling of large ions in electrospray mass spectrometry. *Analytical Chemistry*, 76(6):1754–1760, 2004.

[13]   JN Coles and M Guilhaus. Resolution limitations from detector pulse width and jitter in a linear orthogonal-acceleration time-of-flight mass spectrometer. *Journal of the American Society for Mass Spectrometry*, 5(8):772–778, 1994.

[14]   WB Colson, J McPherson, and FT King. High-gain imaging electron multiplier. *Review of scientific instruments*, 44(12):1694–1696, 1973.

[15]   Philip D Compton, Leonid Zamdborg, Paul M Thomas, and Neil L Kelleher. On the scalability and requirements of whole protein mass spectrometry. *Analytical chemistry*, 83(17):6868–6874, 2011.

[16]   Thomas C Ehlert. Determination of transmission characteristics in mass filters. *Journal of Physics E: Scientific Instruments*, 3(3):237, 1970.

[17]   Rong Feng and Yasuo Konishi. Collisionally-activated dissociation of multiply charged 150-kda antibody ions. *Analytical Chemistry*, 65(5):645–649, 1993.

[18]   John B Fenn. Mass spectrometric implications of high-pressure ion sources. *International Journal of Mass Spectrometry*, 200(1-3):459–478, 2000.

[19]   Luca Fornelli, Daniel Ayoub, Konstantin Aizikov, Alain Beck, and Yury O Tsybin. Middle-down analysis of monoclonal antibodies with electron transfer dissociation orbitrap fourier transform mass spectrometry. *Analytical chemistry*, 86(6):3005–3012, 2014.

[20]   Matthias Frank, Simon E Labov, Garrett Westmacott, and W Henry Benner. Energy-sensitive cryogenic detectors for high-mass biomolecule mass spectrometry. *Mass Spectrometry Reviews*, 18(3-4):155–186, 1999.

[21]   Mattias J Fredriksson, Patrik Petersson, Bengt-Olof Axelsson, and Dan Bylund. An automatic peak finding method for lc-ms data using gaussian second derivative filtering. *Journal of separation science*, 32(22):3906–3918, 2009.

[22]   P.W Geno. *Ion detection in mass spectrometry, mass spectrometry in the biological sciences: a tutorial*. Springer, 1992.

[23]   PW Geno and RD Macfarlane. Secondary electron emission induced by impact of low-velocity molecular ions on a microchannel plate. *International Journal of Mass Spectrometry and Ion Processes*, 92:195–210, 1989.

[24]   Dieter Gerlich. Inhomogeneous rf fields: A versatile tool for the study of processes with slow ions. *Advances in Chemical Physics: State-Selected and State-To-State Ion-Molecule Reaction Dynamics, Part 1. Experiment*, 82:1–176, 1992.

[25]   AJ Guest. A computer model of channel multiplier plate performance. *Acta Electronica*, 14(1):79–97, 1971.

[26]   Xuemei Han, Aaron Aslanian, and John R Yates III. Mass spectrometry for proteomics. *Current opinion in chemical biology*, 12(5):483–490, 2008.

[27]   Wei Hang, Cris Lewis, and Vahid Majidi. Practical considerations when using radio frequency-only quadrupole ion guide for atmospheric pressure ionization sources with time-of-flight mass spectrometry. *Analyst*, 128(3):273–280, 2003.

[28]   Brandy J Johnson, W Russ Algar, Anthony P Malanoski, Mario G Ancona, and Igor L Medintz. Understanding enzymatic acceleration at nanoparticle interfaces: approaches and challenges. *Nano Today*, 9(1):102–131, 2014.

[29]   AN Krutchinsky, IV Chernushevich, VL Spicer, W Ens, and KG Standing. Collisional damping interface for an electrospray ionization time-of-flight mass spectrometer. *Journal of the American Society for Mass Spectrometry*, 9(6):569–579, 1998.

[30]   Jia Liu, Qihua Yang, and Can Li. Towards efficient chemical synthesis via engineering enzyme catalysis in biomimetic nanoreactors. *Chemical Communications*, 51(72):13731–13739, 2015.

[31]   Ranran Liu, Qiyao Li, and Lloyd M Smith. Detection of large ions in time-of-flight mass spectrometry: effects of ion mass and acceleration voltage on microchannel plate detector response. *Journal of the American Society for Mass Spectrometry*, 25(8):1374–1383, 2014.

[32]   Jonathan Lu, Michael J Trnka, Soung-Hun Roh, Philip JJ Robinson, Carrie Shiau, Danica Galonic Fujimori, Wah Chiu, Alma L Burlingame, and Shenheng Guan. Improved peak detection and deconvolution of native electrospray mass spectra from large protein complexes. *Journal of The American Society for Mass Spectrometry*, 26(12):2141–2151, 2015.

[33]   J Martens, W Ens, KG Standing, and A Verentchikov. Secondary-ion and electron production from surfaces bombarded by large polyatomic ions. *Rapid communications in mass spectrometry*, 6(2):147–157, 1992.

[34]   R Meier and P Eberhardt. Velocity and ion species dependence of the gain of microchannel plates. *International journal of mass spectrometry and ion processes*, 123(1):19–27, 1993.

[35]   Nina Morgner and Carol V Robinson. Mass ign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. *Analytical chemistry*, 84(6):2939–2948, 2012.

[36]   Bryan A Parks, Lihua Jiang, Paul M Thomas, Craig D Wenger, Michael J Roth, Michael T Boyne, Patricia V Burke, Kurt E Kwast, and Neil L Kelleher. Top-down proteomics

on a chromatographic time scale using linear ion trap fourier transform hybrid mass spectrometers. *Analytical chemistry*, 79(21):7984–7991, 2007.

[37]   Hui Qiao, Vladimir Collado, Gamini Piyadasa, Alexandre Loboda, Viatcheslav Kozlovski, Victor Spicer, Kenneth G Standing, and Werner Ens. Comparison of electron and ion emission efficiencies in a hybrid detector in an orthogonal tof instrument. In *Proceedings of the 53rd American Society for Mass Spectrometry Annual Conference, San Antonio*, 2005.

[38]   TA Somer and PW Graves. Spiraltron matrices as windowless photon detectors for soft x-ray and extreme uv. *IEEE Transactions on Nuclear Science*, 16(1):376–380, 1969.

[39]   Bernhard Spengler, Dieter Kirsch, Raimund Kaufmann, Michael Karas, Franz Hillenkamp, and Ulrich Giessmann. The detection of large molecules in matrix-assisted uv-laser desorption. *Rapid Communications in Mass Spectrometry*, 4(9):301–305, 1990.

[40]   Yao-Hsin Tseng, Charlotte Uetrecht, Shih-Chieh Yang, Arjan Barendregt, Albert JR Heck, and Wen-Ping Peng. Game-theory-based search engine to automate the mass assignment in complex native electrospray mass spectra. *Analytical chemistry*, 85(23):11275–11283, 2013.

[41]   George Turin. An introduction to matched filters. *IRE transactions on Information theory*, 6(3):311–329, 1960.

[42]   Hamideh Vaghari, Hoda Jafarizadeh-Malmiri, Mojgan Mohammadlou, Aydin Berenjian, Navideh Anarjan, Nahideh Jafari, and Shahin Nasiri. Application of magnetic nanoparticles in smart enzyme immobilization. *Biotechnology letters*, 38(2):223–233, 2016.

[43]   Peter K Vogt. Fortuitous convergences: the beginnings of jun. *Nature reviews Cancer*, 2(6):465, 2002.

[44]   G Westmacott, W Ens, and KG Standing. Secondary ion and electron yield measurements for surfaces bombarded with large molecular ions. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 108(3):282–289, 1996.

[45]   G Westmacott, Matthias Frank, SE Labov, and WH Benner. Using a superconducting tunnel junction detector to measure the secondary electron emission efficiency for a microchannel plate detector bombarded by large molecular ions. *Rapid Communications in Mass Spectrometry*, 14(19):1854–1861, 2000.

[46]   Julia Wiesner, Thomas Premsler, and Albert Sickmann. Application of electron transfer dissociation (etd) for the analysis of posttranslational modifications. *Proteomics*, 8(21):4466–4483, 2008.

[47]   Joseph Ladislas Wiza et al. Microchannel plate detectors. *Nucl. Instrum. Methods*, 162(1-3):587–601, 1979.

[48]   James Yergey, David Heller, Gordon Hansen, Robert J Cotter, and Catherine Fenselau. Isotopic distributions in mass spectra of large molecules. *Analytical Chemistry*, 55(2):353–356, 1983.

[49]   Yifei Zhang, Jun Ge, and Zheng Liu. Enhanced activity of immobilized or chemically modified enzymes. *Acs Catalysis*, 5(8):4503–4513, 2015.

[50]   Zhongqi Zhang, Hai Pan, and Xiaoyu Chen. Mass spectrometry for structural characterization of therapeutic antibodies. *Mass spectrometry reviews*, 28(1):147–176, 2009.

Part 4 – Publication III and XII

Introduction to Publications III and XII

Native ESI-MS of DNA of complexes with DNA binders

Background of Publications III and XII

Both publications examine interactions of DNA-G-quadruplexes with dicarbocyanine and benzothiazole ligands being potential candidates for analytical or pharmaceutical applications. Above-named quadruplexes are one form of less abundant secondary structures of DNA besides the prevalent duplexes, i.e. double stranded DNA or B-DNA. Especially repetitive sequences facilitate the formation of hairpin, triplex, cruciform, left-handed Z-form and quadruplexes, as reviewed by Kervin et al. [26]. These non-B-DNA structures are thought to induce genetic instability and hence thought to contribute to or cause human diseases [26].

In particular, guanine (G)-rich DNA sequences exert the tendency to form into tetrastrand assemblies named G-quadruplexes (G4). These complexes are formed by stacked planar arrangements of four guanine bases, bound by Hoogsten base pairing stabilized by π-π-interactions between the planes and interaction with alkali cations. The latter are located inside a channel formed by the four guanine bases as shown in Fig P3-1:
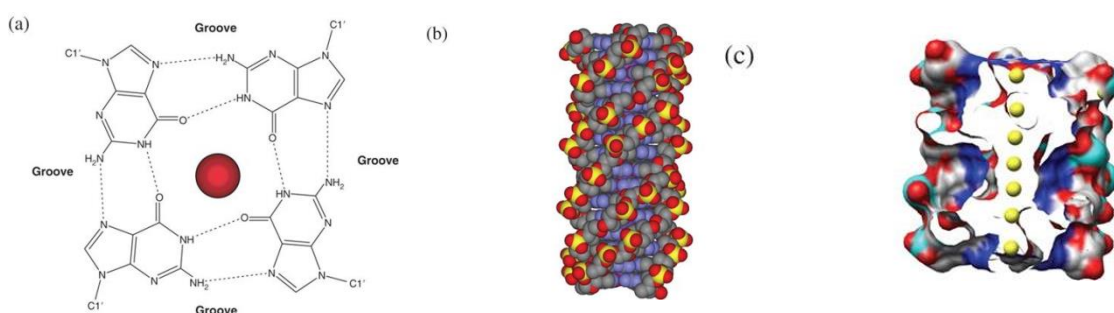
Figure FP3 1 Structure of G-tetrads. A) Hoogsten-base pairing of guanine bases forming one G4-quadriplegic plane. B) section of poly(d)G. Helix formed from multiple planes of G4, stabilisation by pi-pi stack, helix is right handed. C) Interior channel with alkali metal ions further stabilizing structure. Reproduced from Nucleic Acids Research, Volume 34, Issue 19, 1 November 2006, Pages 5402–5415, Copyright (2006), with permission from Oxford academic.

From the pharmaceutical perspective, as the major prerequisite for the formation of G4-quadruplexes, namely G-rich sequences are found in the telomeric regions of chromosomes

and promotor regions of oncogenes, e.g. MYC and BCL2, G4-quadruplexes have become targets for therapeutic approaches, particularly for anticancer strategies [37, 38, 10].

In this perspective the ligand studied in Publication III is the brainchild of the Weisz-Group at the Institute of Biochemistry at the University of Greifswald. It presents and examines 3,3'-diethylthiadicarbocyanine (DiSC$_2$(5)) as selective fluorescence probe for DNA G4-quadruplexes.

Publication XII, also masterminded by the Weisz–Group, identified a benzothiazole derivative as selective ligand candidate for G4-quadruplexes and thoroughly characterized it as a basis for further chemical derivatization in order to yield fluorescence probes for detection of G4-quadruplexes.

Both works examine the binding of their respective ligands to a the MYC-G4-quadruplex (amongst others) by various methods to check topology of targets (CD) initially and after ligand binding, examining interaction sites (NOE-NMR, CSP-NMR), determining binding stoichiometry and affinities (fluorescence titration, UV/VIS, native ESI-TOF-HRMS) and studying thermodynamic aspects (ITC) to provide understanding of ligand-target interaction and as basis for deriving derivatives with optimized structure regarding selectivity and affinity. For structures of ligands and sequences of targets see Fig FP3.2.
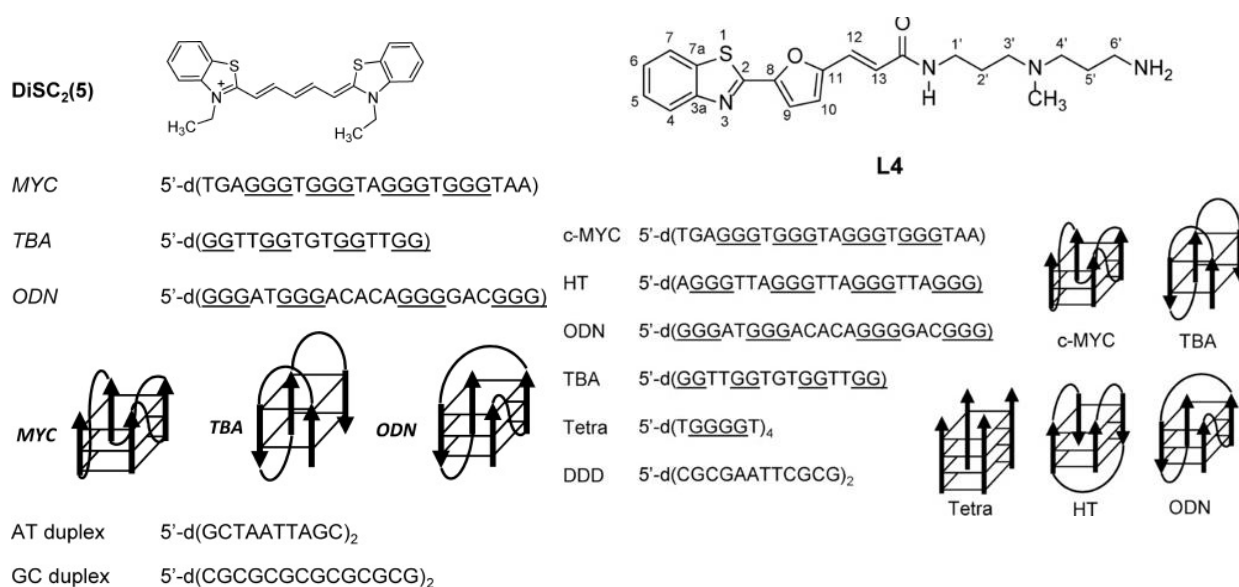
Figure FP3 2 Structure of ligands and targeted oligonucleotides studied by Publication III (DiSC2(5). Publication XII (L4)   Reproduced from Publication III and XII, Copyright (2015, 2017), with permission from Wiley.

Stoichiometries and binding mode - limitations of fluorescence titration

As fluorescence titration yielded binding affinities, yet not unequivocal stoichiometries, this was further investigated by method of continuous variation providing a job plot [25]. Albeit still applied in the field, this method is regarded to be of limited suitability concerning supramolecular complexes and more applicable as a posteriori verification experiment confirming results procured by other techniques [50, 22]. As Hibbert et al. [22] point out, whenever possible from practical reasons, repeated titrations should be performed and the gained data fit to various models, revealing the most suitable model by statistical procedures like shape analysis of residual scatter plot or sum-of-squares F-test[50].
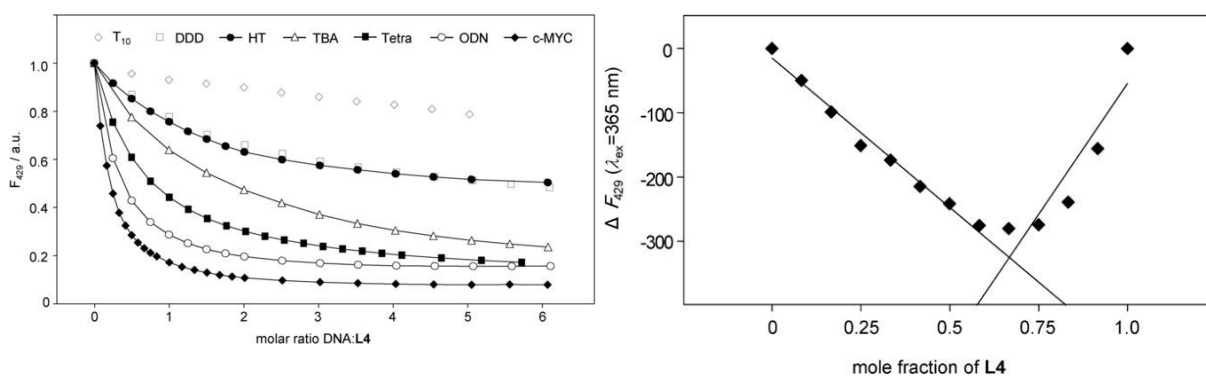
Figure FP3 3 Example of fluorescence titration results (left) and a job plot(right): Job plot obtained by plotting the difference (ΔF_{429}) of L4 fluorescence between c-MYC/L4 mixtures and L4 without DNA, providing information on binding stoichiometry. Publication XII (L4). Reproduced from Publication III and XII, Copyright (2015, 2017), with permission from Wiley.

If due to practical reasons an extensive investigation by titration and model fitting as suggested by Hibbert is impeded, stoichiometry and binding model must be confirmed by orthogonal and preferably direct methods. In case of Publication III and XII, high resolution mass spectrometry of noncovalent bound ligand-target complexes, ionized under native conditions (native ESI) was utilized for this purpose.

Native electrospray ionization mass spectrometry

The capability of the electrospray ionization process (ESI) to produce non-covalently bound species has been reported already in the Nestor publication by Fenn et al. [9]. Shortly after, first publications give account on ESI's aptitude to ionize even supramolecular structures like peptides or protein dimers [46], receptor-ligand or enzyme-substrate complexes [13, 14].

The term "native mass spectrometry" was coined later in 2004 in a review by van den Heuvel and Heck [52], referring to mass spectrometric experimental approaches for ionization and ion transmission that retain all structure types, from primary to quarternary under native (physiological) conditions. Ideally, also binding modes and affinities are preserved during transition from solution to gas phase. The latter of course does not hold true with absolute thermodynamic stringency, as the supramolecular complexes are stripped off their solvent environment and are multiply charged by adduct ions (e.g. H^+, Na^+ NH_4^+) regularly.

Initially and still predominantly, the method is applied in the study of intact protein complexes, both homogenous and heterogeneous and protein-ligand interactions [33, 52, 30, 40, 47]. Yet also other biological supramolecular structures have been subject to study by native ESI-MS, e.g. ribosomal [41] and protein-lipid complexes [29, 56] and relevant to Publications III and XII drug-nucleic acid interactions [42] and G-quadruplexes [35].

Independent of the class of the respective analytes, several prerequisites exist in order to obtain veridical results: 1) To preserve the complexes' native state the chemical environment must be kept as close as possible to in vivo, mandating buffered solutions. 2) The ionization method must be of the "soft type", i.e. unintended in-source fragmentation of the analytes (e.g. DNA oligomers) shall not be induced and the even more dissociation susceptible noncovalent interactions shall not be affected. 3) While transfer into vacuum shall preserve the complex, unspecific binding ionic and neutral adducts shall be reduced or eliminated.

Electrospray ionization (ESI) and variations thereof (µESI, nanoESI, emitter array ESI) is currently the only source type fulfilling first two prerequisites to such extent, that it is implicit referred to by the term "native mass spectrometry" [52]. See also general introduction for description of the ESI ion source. However, as of paramount importance for data interpretation, fundamental research is still focused on fidelity between solute and gas

phase conformation of analytes, structure and stoichiometry of complexes [31, 6], and the effects of the electrospray charging process thereon [27].

A considerable complication results experimentally as well as regarding the fidelity of generated data from necessities of the ESI process, limiting the range of admissible buffers. On the one hand, buffer substances have to be volatile under ESI ion source conditions, predominantly ammonia salts of formic, acetic or carbonic acid. The concentrations of these buffers on the other hand have to be kept in the lower millimolar range to prevent ion suppression and even excessive disintegration and thereby introduction of cations like $Fe^{2+/3+}$, Ni^{2+} or other alloy components of ESI source parts (e.g. ESI sprayer needle) into the solution by the electrochemical process, potentially interfering with the supramolecular complex.

Furthermore, native pH of the supramolecular complex might not be the ideal pH for ionization depending on the chemical moieties of analytes requiring addition of modifiers or dopants, e.g. organic solvents in order to facilitate the ESI process and in order to achieve sufficient ionization. Method development gets even more complex, as an online chromatographic separation may be required rather than a direct infusion to the ion source. Yet, whenever ionization efficiency and supramolecular complex integrity contradict each other, the validity of results must be the essential.

Native mass spectrometry of DNA in general

The concerns regarding preservation of noncovalent complexes during ionization and gas phase transition, described in the previous paragraph and crucial for Publication III and XII, were rebutted by various studies in the past [15, 32]. An early work by Ligh-Wahl et al. [32] reports stability of a 20-mer Watson-Crick-DNA-duplex under ESI-conditions, previously annealed in 10 mM ammonium acetate buffer. Ligh-Wahl and coworkes also give account to influences of buffer composition, as replacing buffer for distilled water yielded no stable duplex in the gas phase. Yet, the gas phase stability of isolated duplexes stripped off all adducts was shown even under the unfavorable conditions of a quadrupole ion trap (collision with damping gas in trap) [7] and it could be demonstrated that base pairing is stable enough under EDD/ECD fragmentation [55].

Thus, native ESI-MS of noncovalent complexes could successfully be applied to various analytes of interest, like noncovalent binding dyes [4], antibiotics [39], metal ions [34, 3] and especially platinum antitumor drugs [24, 5, 53, 28] and DNA-protein ligand complexes, as reviewed by Beck et al. [2]. Further reviews of mass spectrometry of nucleic acids under native conditions are given by Hofstadler et al. [23], Beck et al. [2] and more recently by Fabris et al. [8].

The increased application of IMS-MS to structural biology in the past decade also increased efforts to investigate concordance of solute and gas-phase structure, as the mass spectrometric methodology holds advantages compared to other techniques, regarding required sample amount (usually pmol to nmol range), setup time, speed and thereby throughput. These efforts also brought validation for native nucleic acid mass spectrometry.

These recent approaches regularly employed a variety of complementary methods [1]. Examples are preservation of Watson-Crick pairing in duplex DNA (BIRD[45]), hair-pin structures (HDX,[36]), oligonucleotide-drug complexes (DFT and IMS, [30]). The latter approach correlating calculated and measured collisional cross sections *in vacuo* is deemed promising [1]. DFT can only be applied for small molecules, in this regard bases of nucleic acids. Larger molecules, e.g. oligonucleotides, are simulated with MD simulations. A recent (2014) review by Abi-Ghanem and Gabelica [1] highlights the application of this complementary approach to single stranded, duplex, triplex and quadruplex DNA and DNA-drug complexes forms, all consistent with the overall corollary that the solute structure is preserved upon transition to gas phase or into vacuum, respectively.

Native mass spectrometry of DNA G-Quadruplexes

After the demonstration of feasibility of MD calculations to investigate the structural rigidity of DNA duplexes, this was intensively applied in research of gas phase structure of DNA quadruplexes. Studies investigating G-quadruplex formation and stability in solution by molecular modelling / dynamics [9, 48, 49] and comparison to experimental data [20, 18, 17, 19, 54, 16] established that G-tetrads exhibit large stability and structural rigidity and a significant stabilization by cations in the central channel [44].

Yet, it was also validated by cross section comparison [16], binding and H/D exchange with intra- and intermolecular ligands [12] , MD calculations [44] and ESI-MS experiments [51, 43] that G-quadruplexes retain their solute structure including relevant binding sites during the gas phase transition under electrospray ionization conditions.

It can be concluded from the previously cited literature, that native ESI-MS is an appropriate and valid method for probing complex formation between ligands and G-quadruplexes. As multiply charged 10-mer to 20-mer of DNA (~3000-6000 Da) are in focus, the use of a high resolution mass spectrometer is mandatory. Previous studies often employed FTICR type mass spectrometers, yet technological advance in quadrupole time-of-flight mass spectrometers yielding increased sensitivity, linearity, sub ppm mass accuracy, resolution up to 40000, and scanning speed established this type of mass spectrometer as feasible for high resolution native MS of the G-quadruplexes examined in Publication III and XII.

Experimental approach

As commonly employed in the preceding works cited above, the target DNA was dissolved in 150 mM ammonium acetate as electrospray compatible volatile buffer. To increase ionization efficiency for increased sensitivity without altering buffer composition too much, 10% (v/v) methanol was added. Data were generated under mild ionization conditions, regarding temperature, gas flows, and voltages of ion source and ion path, particularly collision energy (CE) and declustering potential (DP), which are essential to tune accordingly due to their intended mode of action, i.e. dissociation of covalent and noncovalent bonds. To enhance S/N-ratio, the measurements were conducted as direct infusion experiments employing multichannel alignment mode (MCA) generating composite spectra of the entire infusion run. For exact description of experiments, reader is referred to the respective publications.

Results of Publication III and XII

Results for both publications are rather similar and corroborate the findings from the spectroscopic measurements and are in agreement with the NMR results of both works. It was possible to confirm the stoichiometry for the ligands' binding to c-MYC-G-quadruplexes, as model for oncogene promotor regions. The molar stoichiometry was found in both cases to be MYC:ligand 1:1 and 1:2, depending on the concentrations of the respective ligand, conforming two binding sites. Furthermore, it could be derived by titration that the binding on both sites occur non-cooperatively, i.e. independent of each other and that L4 exerts different affinities for each binding site.

In conclusion, the results document the power of native ESI-MS to study DNA-ligand complexes and characterize their binding stoichiometry in support to other spectroscopic methods. Native ESI-MS provides a straightforward and fast method to assess stoichiometries of noncovalent complexes.

Abbreviations

CD: circular dichroism spectroscopy

UV/VIS: ultraviolet–visible spectrophotometry

NMR: Nuclear magnetic resonance spectrometry

NOE-NMR: Nuclear overhauser effect NMR

CSP-NMR: Chemical shift perturbance NMR

ESI: Electrospray ionization

ESI-TOF-HRMS: Electrospray ionization time of flight high resolution mass spectrometry

ITC: Isothermal titration calorimetry

μESI: micro Electrospray ionization

ESI-MS: Electrospray ionization mass spectrometry

DFT density functional theory calculations

IMS-MS: Ion mobility mass spectrometry

BIRD: Blackbody infrared radiation-induced dissociation

HDX: Hydrogen-deuterium exchange

MD: molecular dynamics simulation

CE: collision energy

DP: declustering potential

CCS: Collisional cross section

EDD: electron detachment dissociation

ECD: electron capture dissociation

References

[1]   Josephine Abi-Ghanem and Valérie Gabelica. Nucleic acid ion structures in the gas phase. *Physical Chemistry Chemical Physics*, 16(39):21204–21218, 2014.

[2]   Jennifer L Beck, Michelle L Colgrave, Stephen F Ralph, and Margaret M Sheil. Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass spectrometry reviews*, 20(2):61–87, 2001.

[3]   Jennifer L Beck, Amy Humphries, Margaret M Sheil, and Stephen F Ralph. Electrospray ionisation mass spectrometry of ruthenium and palladium complexes with oligonucleotides. *European Mass Spectrometry*, 5(6):489–500, 1999.

[4]   Ying Cao and Xi-wen He. Studies of interaction between safranine t and double helix dna by spectral methods. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 54(6):883–892, 1998.

[5]   Nathalie Carte, Franck Legendre, Emmanuelle Leize, Noelle Potier, Franziska Reeder, Jean-Claude Chottard, and Alain Van Dorsselaer. Determination by electrospray mass spectrometry of the outersphere association constants of dna/platinum complexes using 20-mer oligonucleotides and ([pt (nh3) 4] 2+, 2cl-) or ([pt (py) 4] 2+, 2cl-). *Analytical biochemistry*, 284(1):77–86, 2000.

[6]   Paul WA Devine, Henry C Fisher, Antonio N Calabrese, Fiona Whelan, Daniel R Higazi, Jennifer R Potts, David C Lowe, Sheena E Radford, and Alison E Ashcroft. Investigating the structural compaction of biomolecules upon transition to the gas-phase using esi-twims-ms. *Journal of The American Society for Mass Spectrometry*, 28(9):1855–1862, 2017.

[7]   Mitchel J Doktycz, Sohrab Habibi-Goudarzi, and Scott A McLuckey. Accumulation and storage of ionized duplex dna molecules in a quadrupole ion trap. *Analytical chemistry*, 66(20):3416–3422, 1994.

[8]   D. Fabris. Ms analysis of nucleic acids in the post-genomic era. *Analytical Chemistry*, 83(15):5810–5816, 2011. PMID: 21651236.

[9]   Eva Fadrná, Nadâ€™a Špacková, Richard Štefl, Jaroslav Koca, Thomas E Cheatham, and Jir Šponer. Molecular dynamics simulations of guanine quadruplex loops: advances and force field limitations. *Biophysical journal*, 87(1):227–242, 2004.

[10]   Kenneth M Felsenstein, Lindsey B Saunders, John K Simmons, Elena Leon, David R Calabrese, Shuling Zhang, Aleksandra Michalowski, Peter Gareiss, Beverly A Mock, and John S Schneekloth Jr. Small molecule microarrays enable the identification of a selective, quadruplex-binding inhibitor of myc expression. *ACS chemical biology*, 11(1):139–148, 2015.

[11]   John B Fenn, Matthias Mann, Chin Kai Meng, Shek Fu Wong, and Craig M Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926):64–71, 1989.

[12]   Valérie Gabelica, Frederic Rosu, Matthias Witt, Gökhan Baykut, and Edwin De Pauw. Fast gas-phase hydrogen/deuterium exchange observed for a dna g-quadruplex. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, 19(2):201–208, 2005.

[13]   Bruce Ganem, Yu Tsyr Li, and Jack D Henion. Detection of noncovalent receptor-ligand complexes by mass spectrometry. *Journal of the American Chemical Society*, 113(16):6294–6296, 1991.

[14]   Bruce Ganem, Yu Tsyr Li, and Jack D Henion. Observation of noncovalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry. *Journal of the American Chemical Society*, 113(20):7818–7819, 1991.

[15]   Bruce Ganem, Yu-Tsyr Li, and Jack D Henion. Detection of oligonucleotide duplex forms by ion-spray mass spectrometry. *Tetrahedron letters*, 34(9):1445–1448, 1993.

[16]   Jennifer Gidden, Erin Shammel Baker, Alessandra Ferzoco, and Michael T Bowers. Structural motifs of dna complexes in the gas phase. *International Journal of Mass Spectrometry*, 240(3):183–193, 2005.

[17]   Charles C Hardin, Matthew J Corregan, David V Lieberman, and Bernard A Brown. Allosteric interactions between dna strands and monovalent cations in dna quadruplex assembly: thermodynamic evidence for three linked association pathways. *Biochemistry*, 36(49):15428–15450, 1997.

[18]   Charles C Hardin, Eric Henderson, Thomas Watson, and Joyce K Prosser. Monovalent cation induced structural transitions in telomeric dnas: G-dna folding intermediates. *Biochemistry*, 30(18):4460–4472, 1991.

[19]   Charles C Hardin, Adam G Perry, and Katie White. Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers: Original Research on Biomolecules*, 56(3):147–194, 2000.

[20]   Pascale Hazel, Julian Huppert, Shankar Balasubramanian, and Stephen Neidle. Loop-length-dependent folding of g-quadruplexes. *Journal of the American Chemical Society*, 126(50):16405–16415, 2004.

[21]   Albert JR Heck and Robert HH van den Heuvel. Investigation of intact protein complexes by mass spectrometry. *Mass spectrometry reviews*, 23(5):368–389, 2004.

[22]   D Brynn Hibbert and Pall Thordarson. The death of the job plot, transparency, open science and online tools, uncertainty estimation methods and other developments in supramolecular chemistry data analysis. *Chemical Communications*, 52(87):12792–12805, 2016.

[23]   Steven A Hofstadler and Richard H Griffey. Analysis of noncovalent complexes of dna and rna by mass spectrometry. *Chemical Reviews*, 101(2):377–390, 2001.

[24]   Paula Iannitti-Tito, Allan Weimann, Geoffrey Wickham, and Margaret M Sheil. Structural analysis of drug–dna adducts by tandem mass spectrometry. *Analyst*, 125(4):627–634, 2000.

[25]   Pann Job. Formation and stability of inorganic complexes in solution. 1928.

[26]   Sean Michael Kerwin. G-quadruplex dna as a target for drug design. *Current pharmaceutical design*, 6(4):441–471, 2000.

[27]   Nina A Khristenko, Jussara Amato, Sandrine Livet, Bruno Pagano, Antonio Randazzo, and Valerie Gabelica. Native ion mobility mass spectrometry: when gas phase ion structures depend on the electrospray charging process. 2018.

[28]   Miriam BG Kloster, James C Hannis, David C Muddiman, and Nicholas Farrell. Consequences of nucleic acid conformation on the binding of a trinuclear platinum drug. *Biochemistry*, 38(45):14731–14737, 1999.

[29]   Arthur Laganowsky, Eamonn Reading, Timothy M Allison, Martin B Ulmschneider, Matteo T Degiacomi, Andrew J Baldwin, and Carol V Robinson. Membrane proteins bind lipids selectively to modulate their structure and function. *Nature*, 510(7503):172, 2014.

[30]   Cris Lapthorn, Frank Pullen, and Babur Z Chowdhry. Ion mobility spectrometry-mass spectrometry (ims-ms) of small molecules: separating and assigning structures to ions. *Mass spectrometry reviews*, 32(1):43–71, 2013.

[31]   Kenneth J Laszlo, Eleanor B Munger, and Matthew F Bush. Effects of solution structure on the folding of lysozyme ions in the gas phase. *The Journal of Physical Chemistry B*, 121(13):2759–2766, 2017.

[32]   KJ Light-Wahl, DL Springer, BE Winger, CG Edmonds, DG Camp, BD Thrall, and RD Smith. Observation of a small oligonucleotide duplex by electrospray ionization mass spectrometry. *Journal of the American Chemical Society*, 115(2):803–804, 1993.

[33]   Joseph A Loo. Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *International Journal of Mass Spectrometry*, 200(1-3):175–186, 2000.

[34]   KP Madhusudanan, SB Katti, R Vijayalakshmi, and BU Nair. Chromium (iii) interactions with nucleosides and nucleotides: a mass spectrometric study. *Journal of mass spectrometry*, 34(8):880–884, 1999.

[35]   Adrien Marchand and Valerie Gabelica. Native electrospray mass spectrometry of dna g-quadruplexes in potassium solution. *Journal of the American Society for Mass Spectrometry*, 25(7):1146–1154, 2014.

[36]   Jingjie Mo, Gabrielle C Todd, and Kristina Håkansson. Characterization of nucleic acid higher order structure by gas-phase h/d exchange in a quadrupole-ft-icr mass spectrometer. *Biopolymers: Original Research on Biomolecules*, 91(4):256–264, 2009.

[37]   Deepanjan Panda, Manish Debnath, Samir Mandal, Irene Bessi, Harald Schwalbe, and Jyotirmayee Dash. A nucleus-imaging probe that selectively stabilizes a minor conformation

of c-myc g-quadruplex and down-regulates c-myc transcription in human cancer cells. *Scientific reports*, 5:13183, 2015.

[38]   Y Pavan Kumar, Puja Saha, Dhurjhoti Saha, Irene Bessi, Harald Schwalbe, Shantanu Chowdhury, and Jyotirmayee Dash. Fluorescent dansyl-guanosine conjugates that bind c-myc promoter g-quadruplex and downregulate c-myc expression. *ChemBioChem*, 17(5):388–393, 2016.

[39]   Gabriella Pócsfalvi, Giuseppe Di Landa, Pasquale Ferranti, Alberto Ritieni, Giacomino Randazzo, and Antonio Malorni. Observation of non-covalent interactions between beauvericin and oligonucleotides using electrospray ionization mass spectrometry. *Rapid communications in mass spectrometry*, 11(3):265–272, 1997.

[40]   Carol V Robinson. Protein complexes take flight. *Nature Structural & Molecular Biology*, 9(7):505, 2002.

[41]   Adam A Rostom, Paola Fucini, Dennis R Benjamin, Ralf Juenemann, Knud H Nierhaus, F Ulrich Hartl, Christopher M Dobson, and Carol V Robinson. Detection and selective dissociation of intact ribosomes in a mass spectrometer. *Proceedings of the National Academy of Sciences*, 97(10):5185–5190, 2000.

[42]   Frédéric Rosu, Edwin De Pauw, and Valérie Gabelica. Electrospray mass spectrometry to study drug-nucleic acids interactions. *Biochimie*, 90(7):1074–1087, 2008.

[43]   Manuel Rueda, F Javier Luque, and Modesto Orozco. Nature of minor-groove binders-dna complexes in the gas phase. *Journal of the American Chemical Society*, 127(33):11690–11698, 2005.

[44]   Manuel Rueda, F Javier Luque, and Modesto Orozco. G-quadruplexes can maintain their structure in the gas phase. *Journal of the American Chemical Society*, 128(11):3608–3619, 2006.

[45]   Paul D Schnier, John S Klassen, Eric F Strittmatter, and Evan R Williams. Activation energies for dissociation of double strand oligonucleotide anions: Evidence for watson- crick base pairing in vacuo. *Journal of the American Chemical Society*, 120(37):9605–9613, 1998.

[46]   Richard D Smith, Karen J Light-Wahl, Brian E Winger, and Joseph A Loo. Preservation of non-covalent associations in electrospray ionization mass spectrometry: Multiply charged polypeptide and protein dimers. *Organic Mass Spectrometry*, 27(7):811–821, 1992.

[47]   Frank Sobott and Carol V Robinson. Protein complexes gain momentum. *Current opinion in structural biology*, 12(6):729–734, 2002.

[48]   Nad'a ŠpacÌŒeková, Imre Berger, and Jir Šponer. Structural dynamics and cation interactions of dna quadruplex molecules containing mixed guanine/cytosine quartets

revealed by large-scale md simulations. *Journal of the American Chemical Society*, 123(14):3295–3307, 2001.

[49]   Richard Štefl, Thomas E Cheatham III, Nadâ€™a Špacková, Eva Fadrná, Imre Berger, Jaroslav Koca, and Jir Šponer. Formation pathways of a guanine-quadruplex dna revealed by molecular dynamics and thermodynamic analysis of the substates. *Biophysical journal*, 85(3):1787–1804, 2003.

[50]   Filip Ulatowski, Kajetan Dabrowa, Tomasz Baakier, and Janusz Jurczak. Recognizing the limited applicability of job plots in studying host–guest interactions in supramolecular chemistry. *The Journal of organic chemistry*, 81(5):1746–1756, 2016.

[51]   M Vairamani and Michael L Gross. G-quadruplex formation of thrombin-binding aptamer detected by electrospray ionization mass spectrometry. *Journal of the American Chemical Society*, 125(1):42–43, 2003.

[52]   Robert HH van den Heuvel and Albert JR Heck. Native protein mass spectrometry: from intact oligomers to functional machineries. *Current opinion in chemical biology*, 8(5):519–526, 2004.

[53]   G Wickham, P Iannitti, J Boschenok, and MM Sheil. Electrospray ionization mass spectrometry of covalent ligand-oligonucleotide adducts: evidence for specific duplex ion formation. *Rapid communications in mass spectrometry: RCM*, pages S197–203, 1995.

[54]   James R Williamson, MK Raghuraman, and Thomas R Cech. Monovalent cation-induced structure of telomeric dna: the g-quartet model. *Cell*, 59(5):871–880, 1989.

[55]   Jiong Yang, Jingjie Mo, Julie T Adamson, and Kristina Håkansson. Characterization of oligodeoxynucleotides by electron detachment dissociation fourier transform ion cyclotron resonance mass spectrometry. *Analytical chemistry*, 77(6):1876–1882, 2005.

[56]   Yixuan Zhang, Lan Liu, Rambod Daneshfar, Elena N Kitova, Caishun Li, Feng Jia, Christopher W Cairo, and John S Klassen. Protein–glycosphingolipid interactions revealed using catch-and-release mass spectrometry. *Analytical chemistry*, 84(18):7618–7621, 2012.

List of Figures

Part 1

Part 2

Part 3

List of Tables

Appendix: Publication manuscripts and curriculum vitae

Appendix I

Accepted Manuscripts

Publication I and supplemental material	
Title	Ligand-receptor binding increments in enantioselective liquid chromatography
Authors	*Sievers-Engler, Adrian; Lindner, Wolfgang; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1363), **2014** p. 79-88
DOI	http://dx.doi.org/10.1016/j.chroma.2014.04.077

Reprint with permission from Elsevier

Ligand–receptor binding increments in enantioselective liquid chromatography

Adrian Sievers-Engler [a], Wolfgang Lindner [b], Michael Lämmerhofer [a],*

[a] *Institute of Pharmaceutical Sciences, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany*
[b] *University of Vienna, Department of Analytical Chemistry, Währinger Strasse 38, A-1090 Vienna, Austria*

A B S T R A C T

A set of *N*-derivatized amino acids were separated into enantiomers on a *tert*-butylcarbamoylated quinine-based chiral stationary phase (CSP). Quantitative structure–property relationship (QSPR) studies were then employed to investigate the retention behavior and factors responsible for enantioselectivity. Computations were performed using a general linear model and a Free-Wilson matrix with indicator variables as structural descriptors. The approach allowed calculations of retention increments for first and second eluted enantiomers as well as group contributions to enantioselectivity. The results demonstrated that the additivity principle of group contributions was obeyed for the majority of solutes in the data set. Only a few basic amino acids (Arg, His) needed to be removed as they did not fit to such a linear model leading to outliers. The model was carefully validated and then utilized to investigate retention and enantioselectivity contributions of different protection groups and individual amino acid residues. It turned out that primarily protection groups were driving retention and enantioselectivity. In contrast, the contribution of amino acid residues to enantioselectivity was only significant for secondary amino acids, α-methylated amino acids, aspartic acid and a few sterically bulky aliphatic amino acid residues (Tle, Ile, allo-Ile). Amongst them only the latter group contributed positively to enantioselectivity while the other residues mentioned reduced enantioselectivity significantly. This type of QSPR model may be valuable to analyze retention/selectivity data of closely related congeneric compound series, is illustrative and straightforward to implement. It is thus valuable for interpretation of retention mechanisms, while its utility for prediction of retention and enantioselectivity data is limited to compounds made up of groups included in the solute set used for deriving the increments.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Liquid chromatographic enantiomer separation has become an indispensable tool in pharmaceutical sciences and drug discovery, respectively [1,2]. Nowadays it is a well-developed technology with numerous chiral stationary phases commercially available for separation of virtually any chiral compound [3,4]. For practical reasons, pharmaceutical industries have established column screening platforms to find the most suitable chiral stationary phase (CSP) for a given enantiomer separation problem by automated overnight screenings of a predefined set of chiral stationary phases. While this approach is most likely successful in terms of finding a CSP with satisfactory enantiomer separation factor, information on which parameters and interactions have led to enantiomer

recognition by the chiral selector immobilized on the supporting particles is usually not derived. Due to the systematic nature of such screenings the derived data matrix may be information-rich regarding structural effects on enantiorecognition. Chemometric and chemoinformatic methods [5,6], respectively, might be helpful to extract valuable information on binding increments and molecular recognition. Amongst those, quantitative structure–property relationships (QSPR) are one powerful option. They have become of utmost interest for retention prediction as an additional constraint in comprehensive analysis of complex mixtures by hyphenated methods such as GC–MS/MS and HPLC–MS/MS to enhance the confidence for correct compound identification [7–12]. Various QSPR variants have been utilized for data processing in enantioselective chromatography comprising linear free energy relationship (LFER) studies in form of the Hansch analysis [13–24], linear solvation energy relationships (LSER) [25–29], 3D-QSPR employing comparative molecular field analysis (CoMFA)/comparative molecular similarity analysis (CoMSIA) [5,17,30–35], and neural network
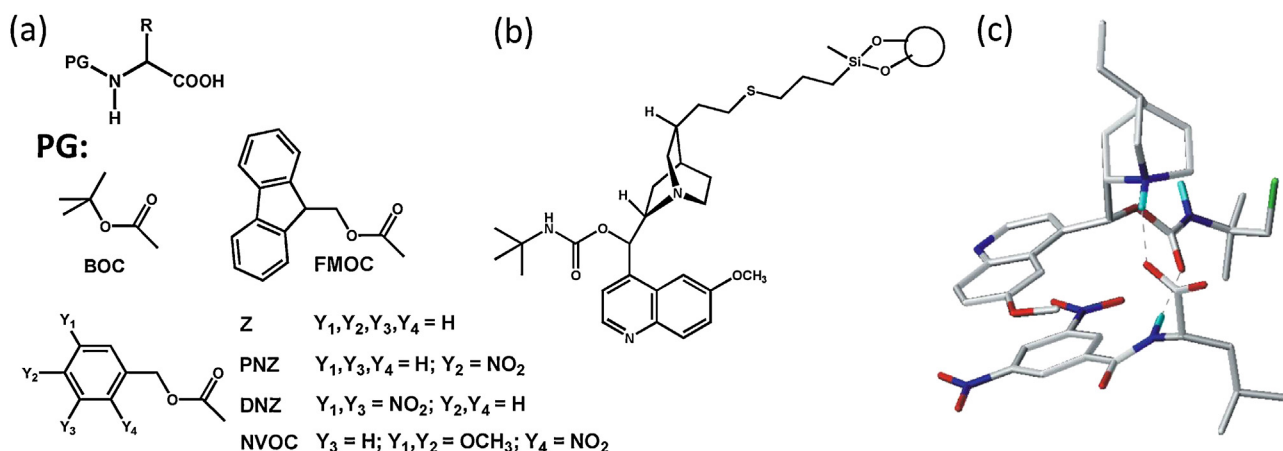
Fig. 1. Structure of (a) data set comprised of amino acid derivatives, (b) chiral stationary phase, (c) X-ray crystal structure of co-crystallized quinine carbamate and 3,5-dinitrobenzoylamino acid leucine [41]. The present aromatic oxycarbonyl-type amino acid derivatives (e.g. DNZ-amino acids) are supposed to reveal a very similar predominant binding mode with an additional —CH$_2$—O— moiety inserted between aromatic and carbonyl group. The carbamate type amino acid derivatives tentatively adopt a similar binding mode maintaining the primary ionic, hydrogen-bond and π–π-interactions, but arranging itself in a slightly wider turn at the selector due to the additional —CH$_2$—O— element.

[23,32,34], or other models [36]. Another straightforward readily amenable, yet informative method is the additivity model (Free Wilson analysis), in particular for certain type of solute sets in which the members of the congeneric series differ in at least two positions by certain structural features from a basic lead structure. It is based on the idea that, in absence of cooperative effects, the free energy for the process of transferring the solute from the mobile to the stationary phase is additively composed of the free energy of individual group contribution [3,37,38]. In chromatography, the group contributions are often termed as binding or retention increments.

Free Wilson analysis [39] is a proper and efficient means to calculate such retention increments. Unlike above mentioned QSPR approaches, it is a simple numerical method directly relating structural features to a dependent variable [39], in chromatography retention or separation factors. Free Wilson analysis is most often adopted in the so-called Fujita-Ban modification. It states that the dependent variable (response y) is made up of the sum of individual group contributions $a_{i,n}$ and a basic contribution provided by the unsubstituted reference compound a_0 (Eq. (1)).

$$\ln y = a_0 + \sum_i a_{i,n} \cdot X_{i,n} \qquad (1)$$

In Eq. (1), the independent variables $X_{i,n}$ are indicator variables which indicate the presence (indicated by the numerical value 1) or absence (value 0) of a specific substituent. The coefficients $a_{i,n}$ represent individual group contributions and can be conveniently obtained by multiple linear regression analysis.

With the list of derived coefficients in hand, the effect of each substituent on the dependent variable (response) becomes readily evident and compounds with substituent combinations not present in the test set can be straightforwardly predicted. Such an approach was not yet investigated for calculating retention increments for chiral separations. Another methodology to calculate group contributions in the context of enantioselective chromatography was presented by Armstrong [40].

In this study, Free Wilson analysis was utilized to rationalize individual group contributions to retention and enantiomer separation of various N-derivatized amino acids (Z, benzyloxycarbonyl; BOC, *tert*-butoxycarbonyl; PNZ, 4-nitrobenzyloxycarbonyl; DNZ, 3,5-dinitrobenzyloxycarbonyl; FMOC, 9-fluorenylmethoxycarbonyl; NVOC, o-nitroveratryloxycarbonyl, i.e. 4,5-dimethoxy-2-nitrobenzyloxycarbonyl) on quinine carbamate chiral stationary phases (Fig. 1). The group of amino acid derivatives shows structural variability of the amino acid residue and the protection group. The results of the Free Wilson analysis can be employed to predict solutes not present in the test set and are used herein to compare the effect of distinct residues on the chromatographic response. The general prediction capability of the model has been tested by leave-out cross-validation. Feasibilities and shortcomings of the methodology are discussed.

2. Experimental

2.1. Materials

The chiral stationary phase based on *tert*-butylcarbamoylquinine immobilized onto 3-mercaptopropyl-modified porous silica Kromasil 100 Å, 5 μm (from EKA-Nobel, Bohus, Sweden), a prototype of Chiralpak QN-AX (Chiral Technologies Europe, Illkirch, France), was synthesized as described elsewhere [42]. The selector coverage was 0.27 mmol/g CSP corresponding to about 0.8 μmol/m^2. The CSP was packed into a stainless-steel column (150 mm × 4.6 mm ID) by a conventional slurry packing procedure.

Amino acids were provided by Bachem (Bubendorf, Switzerland) and Sigma–Aldrich (Vienna), respectively, or were gifts of various colleagues. Amino acids were derivatized following standard protocols as described for instance in Ref. [43]. Elution orders were determined with single enantiomers and well-defined non-racemic mixtures, respectively.

2.2. Chromatography

The utilized chromatographic system was a Hitachi-Merck Liquid Chromatograph and consisted of L-6200 Intelligent Pump, D-6000 Interface, AS-2000A Autosampler, D-6000 Chromatography Data Station Software, HPLC Manager Vers. 2.09 from Merck (Darmstadt, Germany) and a column thermostat of W.O. Electronics (Langenzersdorf, Austria).

The pH of the mobile phase was measured with an Orion pH-meter model 520A (from Orion, Vienna, Austria) and represents the apparent pH (pH$_a$) value measured in the aqueous–organic mobile phase mixture.

Mobile phases for chromatography were prepared from ammonium acetate p.a. (Merck, Darmstadt, Germany), HPLC grade water
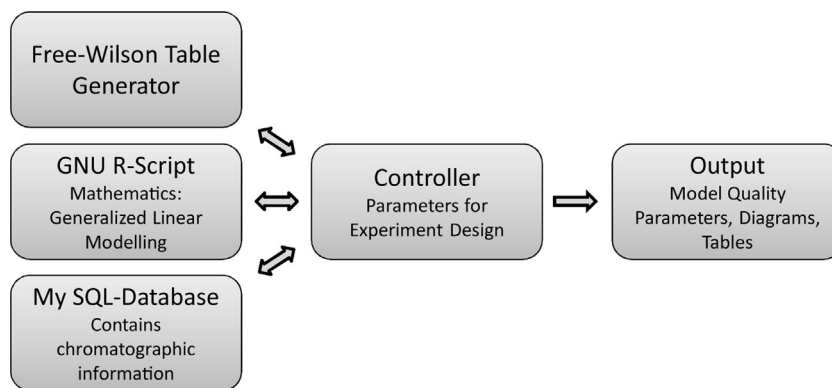
Fig. 2. Building blocks of the software. Arrows depict data/command flow.

(purified by a Milli-Q-Plus filtration unit from Millipore, Bedford, MA, USA) and methanol of HPLC-grade (from Baker).

A mixture of methanol and 0.1 M ammonium acetate buffer (80:20; v/v) was used as mobile phase. The pH_a of this mixture was adjusted to 6.0 by adding glacial acetic acid of analytical grade. Mobile phases were filtered through a Nalgene nylon membrane filter (0.2 μm) from Nalge Co. (New York, USA) and degassed before use by sonication. The flow rate was 1 ml min^{-1} and the column was thermostated at 25 °C. The UV signal was monitored at 254 nm. Some representative chromatograms are given in Supplementary Material, Fig. S1.

2.3. Statistical procedures

To derive functional group contributions for retention and enantioselectivity, generalized linear models (GLM) have been derived by a software network. Thus, statistical computations have been performed by GNU-R using the glm()-function.

The chromatographic data used for computations were retention factors of first (k_1) and second (k_2) eluted enantiomer as well as separation factors α (enantioselectivities) (see supplementary material, Table S1). These data were organized in a MySQL database and served as input values for the dependent variable in the GLMs. Structural descriptors were organized in a Free-Wilson table, in which presence of a certain structural feature was indicated by a value of 1 and its absence by a 0. These independent variables were correlated with the dependent variable by the glm()-function using the GNU-R script to derive statistical meaningful correlations and weights of each structural feature, i.e. individual group contributions. The Controller is a graphical user interface which allows the user to select the data set via SQL-query, to define parameters for the analysis and to control the data flow between the other programs (Fig. 2).

3. Results and discussion

3.1. Test set and characteristics of general linear model

For the sensitive and stereoselective analysis of amino acids, they are often derivatized. Such molecular systems are ideally suited for investigation of molecular recognition contributions of individual structural features. Structural variation in two positions ensures that each substituent can be present more than once in the data set which is deemed to be a prerequisite for statistical meaningful data. Yet, the GLM approach generally allows single point determinations in which a particular residue is present only once in the data set, if there are not too many of such single point

determinations. In this case, of course the entire experimental error is contained in this fragment contribution.

In total 142 distinct analytes were contained in the data set, of which 7 had acidic amino acid residues, 19 basic, 45 aromatic, 23 polar and the remaining 42 apolar aliphatic amino acid side chains. Protection groups employed were all of the oxycarbonyl type structure with aromatic (Z, PNZ, DNZ, NVOC, FMOC) or aliphatic (BOC) residue. Amino acid derivatives with Lys and Orn were derivatized in the side chain leading to neutral derivatives. In total, 62 unique amino acid residues were present in the data set. It might appear that the data matrix is not sufficiently covered for some of the amino acid derivatives (ca. 45% coverage of all possible cases). However, the Free Wilson model can principally deal with such situations and is actually dedicated for such a scenario. If the model is intended to be used for predictions, this is only possible with experimentally partly covered data matrix. Extrapolation to compounds of which the substituents have not been in the original data matrix for model derivation is not possible.

For the computations, Z-Glycine was selected herein as reference compound with basic retention increment a_0, for both first (k_1) and second eluted enantiomer (k_2) sets, because it is the structurally simplest compound from viewpoint of both protecting group and amino acid substituent. Retention increments were then calculated for first and second eluted enantiomer from $k_{1,2}$ as dependent variable and group contributions to enantioselectivity from separation factors α after transformation to the logarithmic scale.

GLMs were derived for the whole set as well as subsets for different parameters like polarity, charge, or π-interaction capability. Resultant estimated coefficients represent the weight (retention increment) by which each structural feature contributes to the overall dependent variable in addition to the intercept a_0. In other words, the corresponding $\ln k$ of a certain compound can be easily calculated from the sum of the coefficients of amino acid residue, protection group and intercept a_0. The statistical significance of the distinct group contributions provides information whether this substituent adds a statistically significant increment to the basic contribution constituted by the intercept a_0. Non-significant terms are therefore not eliminated because they provide also useful information which would be lost otherwise. The model also tolerates few single point determinations. If they would be deleted from the data set, the information about those substituents would be completely lost. If they are included, information is available and can be used for prediction of substituent combinations. In fact, if the intercept is known or fixed and the additivity principle obeyed, a single point determination is in principle suitable for calculation of the substituent increment (like single point calibrations in quantitative photometric analysis where the calibration line is forced
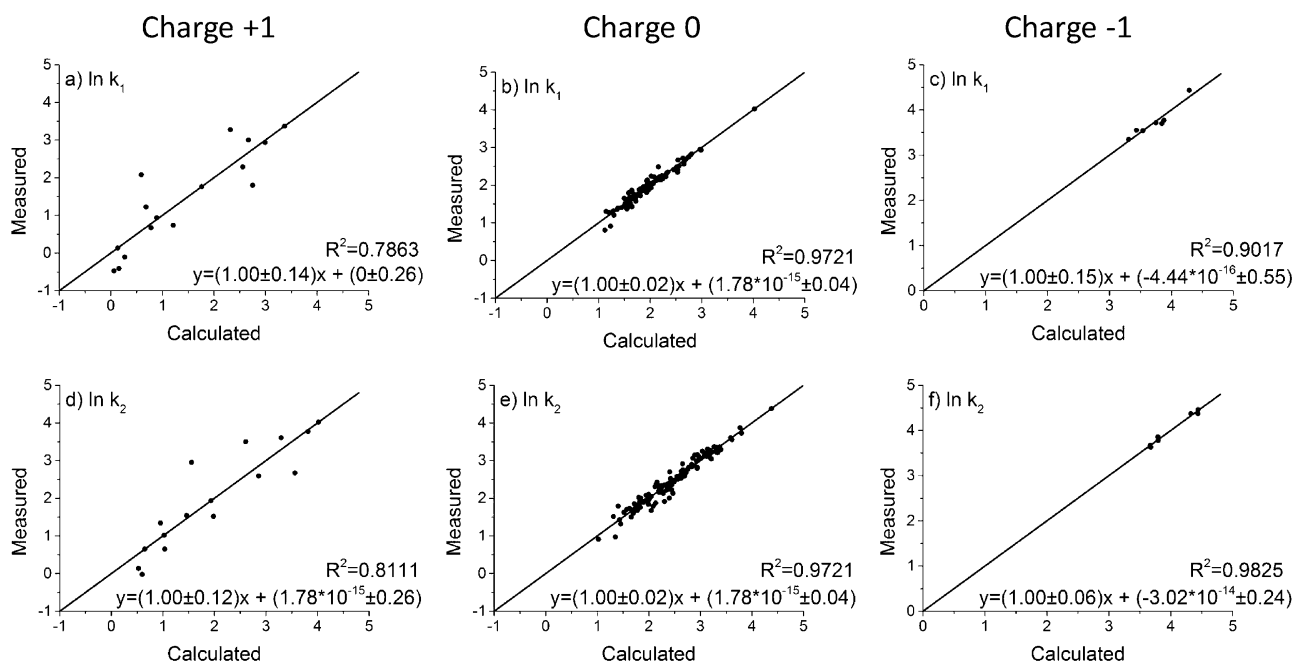
Fig. 3. Influence of side chain character on model quality. (a and d) Basic amino acid residues; b) and e) neutral aliphatic, aromatic and polar residues; and (c and f) acidic amino acids side chains.

through 0). They give a first idea of its effect on the response, but without any statistical significance. One has to be aware that all the experimental error is contained in this increment and not averaged by multiple determinations like for residues which are present several times.

The big advantage of such a QSPR approach is that the table for regression analysis can easily be generated without complicated measurement of physicochemical properties or lengthy computations of structural parameters. The derived group contributions are illustrative, much more than models derived by multivariate statistical methods such as PCA, PCR, and PLS. The derived fragment contributions can be correlated with physicochemical parameters or substituent constants to provide further insight into binding modes. Addition and elimination of compounds is simple and does not change the values of the other compounds. Any compound can serve as reference; its exchange just leads to a linear transformation of coefficients by a certain increment, i.e. it constitutes a linear shift of the scale. The entire method is easy to apply. The main disadvantage is that the utility for prediction is limited. Only substituents present in the data set can be employed for predictions of new combinations not present in the congeneric series. Thus, the number of new analogs that can be predicted is usually low. However, it is very useful for finding factors, i.e. substituent effects which contribute significantly to the response (dependent variable) of interest.

3.2. Analysis with full set and data set refinement

Initially, all 142 compounds for which experimental chromatographic data were available (see supplementary information, Table S1) were subjected to statistical analysis using the glm()-function of GNU-R. For analysis, measured response values were plotted over calculated ones yielding determination coefficients R^2 of 0.86381 for $\ln k_1$, 0.89049 for $\ln k_2$, and 0.89382 for $\ln \alpha$. The variances of the response values cover roughly two orders of magnitudes for retention factors and ca. one order of magnitude for separation factors, while experimental errors for retention and separation factors are usually less than 2% RSD. The parity plots between experimental and calculated responses for the full set clearly documented that the linearity (additivity) principle of group contributions is essentially obeyed but there are a few outliers.

Thus, a refinement of the data set was carried out to remove outliers. To identify those amino acid residues which give poor fit, the full set was divided into subsets according to side chain character (Fig. 3). The fit for neutral and acidic side chains was fairly good while for basic amino acid residues agreement between calculated and experimental values was significantly worse.

In order to examine which cases and structural features were responsible for the poor fit and larger scatter of the data in the plot of the basic amino acid residue subset, a threshold value for outliers of 0.389 as the maximum absolute deviation tolerated between measured and calculated $\ln k$-values was defined (note, this was the maximal deviation in the set with neutral side chain). Table 1 shows the residuals and type of compounds which were found to be outliers in the basic amino acid set according to above defined threshold level. It is striking that all of them had either an Arg or a His residue. On contrary, Orn and Lys with a primary amine in the side chain seemed to obey the additivity principle. This can be readily explained by the fact that these two amino acids are deriva-

Table 1
Outliers within subset of compounds with positive side chain charge. Deviation based on $\ln k_2$.

Deviation (residuals)	Compound	Deviation (residuals)	Compound	Deviation (residuals)	Compound
1.412	NVOC-Arg	0.585	PNZ-Arg	0.389	Z-His
0.902	bis-(NVOC)-His	0.446	NVOC-His	0.389	Z-Arg
0.902	bis-(PNZ)-His	0.438	DNZ-Arg		

Z, benzyloxycarbonyl; PNZ, 4-nitrobenzyloxycarbonyl; DNZ, 3,5-dinitrobenzyloxycarbonyl; FMOC, 9-fluorenylmethoxycarbonyl; NVOC, o-nitroveratryloxycarbonyl (4,5-dimethoxy-2-nitrobenzyloxycarbonyl).
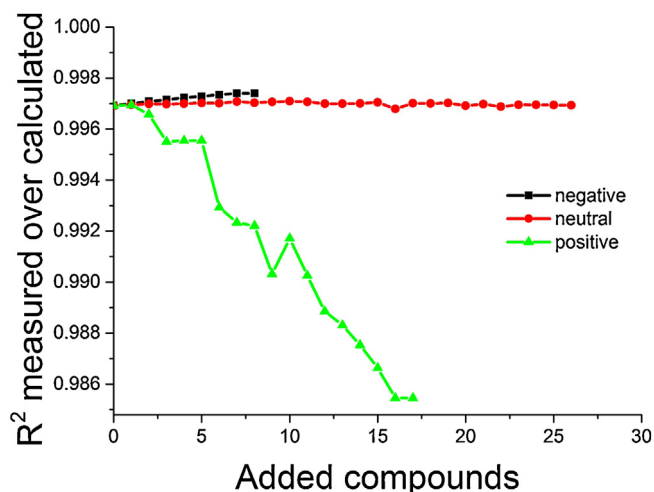
Fig. 4. Effect of polar and charged residues in data set on GLM quality.

tized at the α-amino group as well as the side chain amino group leading to a doubly protected derivative with a neutral side chain. In the refined data set, all cases with Arg and His as amino acid residue were removed.

To further elucidate the effect of polar and charged amino acid residues on the quality of the model, group contributions were calculated for a subset with apolar neutral residues and subsequently an increasing number of particular compounds with polar, basic (positively) or acidic (negatively charged) residues have been added stepwise. The quality of the fit was estimated after each addition by the determination coefficient R^2 of the correlation between measured and predicted response $\ln k_2$. For each number of added compounds, 10 randomized experiments were done and averaged R^2 values of the model are recorded in Fig. 4.

R^2 dropped significantly with each step of addition of a compound with basic amino acid residue from 0.997 to 0.985 when 17 compounds with basic amino acid residues were added to the data set. On contrary, R^2 did not change significantly with addition of polar or acidic amino acids. This experiment clearly demonstrated the adverse effect of basic (positively charged) side chains on model quality and therefore demonstrated the inaptitude of the

GLM to accurately compute retention and enantioselectivity based on the additivity principle of group contributions for some positively charged amino acids (Arg, His). The reason may be disturbing interactions with the silica support or additional cooperative contributions from the basic side chain in the solute-selector complex. It makes thus sense to remove these residues from the data set, while polar and acidic amino acids can be kept in the data set without any negative effect on the model quality. By removing them, the quality of the model gets better but the applicability is restricted to amino acids contained in the test set.

3.3. Refined data set

Fig. 5 illustrates the parity plots between measured and calculated $\ln k_{1,2}$ and $\ln \alpha$ values, respectively, as obtained for the refined set (full set minus all compounds with the basic side chains Arg and His). It can be clearly seen that all the observations nicely scatter around the parity line and no obvious outliers can be detected after removal of derivatives with Arg and His. Residuals (deviations between experimental and predicted response values) are randomly distributed along the 0-deviation line confirming that no systematic deviations exist due to inadequate model (see supplementary material, Fig. S2). The statistical parameters for the computed models are summarized in Table 2. It can be seen that the determination coefficient R^2 is larger than 0.9 for $\ln k_1$ and $\ln k_2$ and close to 0.9 for $\ln \alpha$ demonstrating a sufficient model quality. The prediction performance of the model has been tested by a leave-one-out (LOO) cross-validation. Q^2 ($R^2_{\text{cross-validation}}$) of >0.8 for $\ln k_1$ and $\ln k_2$ as well as >0.6 for $\ln \alpha$ document a reasonable performance for predictions of substituent combinations not in the training set. The refined data set was then used for further model validation.

3.4. Validation

3.4.1. Cross-validation by scrambling test

The performance and robustness of the model was then tested by cross-validation. Various approaches have been considered for this purpose. Amongst others, the prediction capability of the model after scrambling the data set has been examined as cross-validation strategy. For this purpose, $\ln k_1$ and $\ln k_2$ values have been mixed and reassigned en bloc to randomly selected
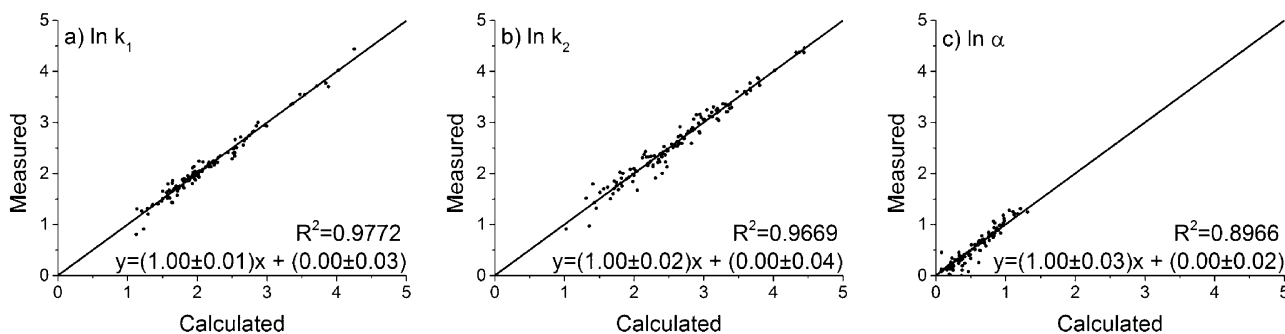
Fig. 5. Parity plots of measured over calculated for the refined set. $N = 130$.

Table 2
Statistical parameters for refined set.

$N = 130$	R^2	Adj. R^2	Q^2	Std. err.	PRESS	F-statistic	p-Value
k_1	0.973	0.978	0.804	0.099	9.303	44.799	4.036E−39
k_2	0.964	0.970	0.817	0.138	12.385	32.403	2.123E−34
α	0.896	0.915	0.654	0.112	5.448	10.547	4.667E−19

Adj. R^2, adjusted R^2; Q^2, R^2 cross-validation; Std. err., standard error; PRESS, prediction residual error sum of squares; F-statistic, Fisher's F-value; p-Value, statistical probability measure.

compounds. Thus, the *y*-responses of the whole data set were scrambled. As quality measure we have chosen the determination coefficient of the linear regression of measured over predicted response values. As expected, the prediction did not work well with the scrambled set. R^2 of measured over calculated dropped to <0.15, depending on the number of compounds removed for prediction (see Figs. 6 and 7).

3.4.2. Leave-out cross-validation

The predictive quality of QSPR models is frequently tested by leave-one-out (LOO) cross-validation and results have been provided above (see also Table 2). For the present large data set LOO cross-validation may lead to too optimistic results (removal of a single compound in a large data set has little effect on the underlying model and prediction capability). Hence, another way of cross-validation was tested as well. We examined the stability of the correlation coefficient for predictions, while iteratively increasing the number of excluded compounds and performing predictions with the derived group contributions. Determination coefficients of measured vs predicted response values as well as corresponding standard deviations are plotted in dependence of the number of excluded and predicted compounds in Fig. 8a and b demonstrate
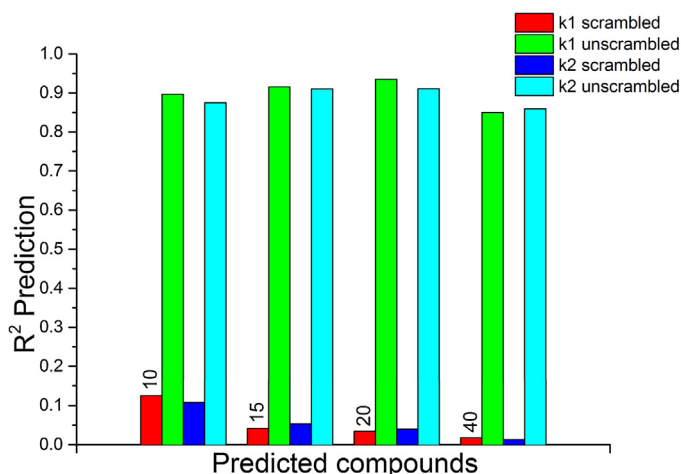
Fig. 6. Average R^2 of predicted over calculated fit for increasing amount of predicted compounds. For averaging, each calculation cycle was repeated 10 times with randomly chosen predicted compounds.
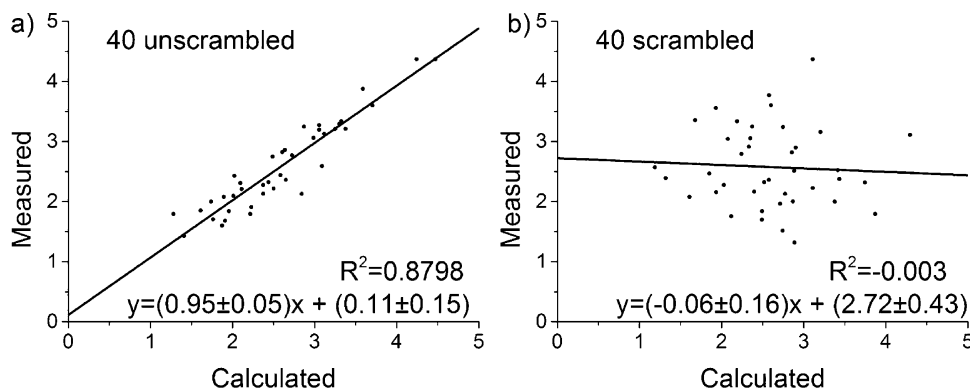
Fig. 7. Typical scatter plot of (a) unscrambled and (b) scrambled data set. $N = 40$.
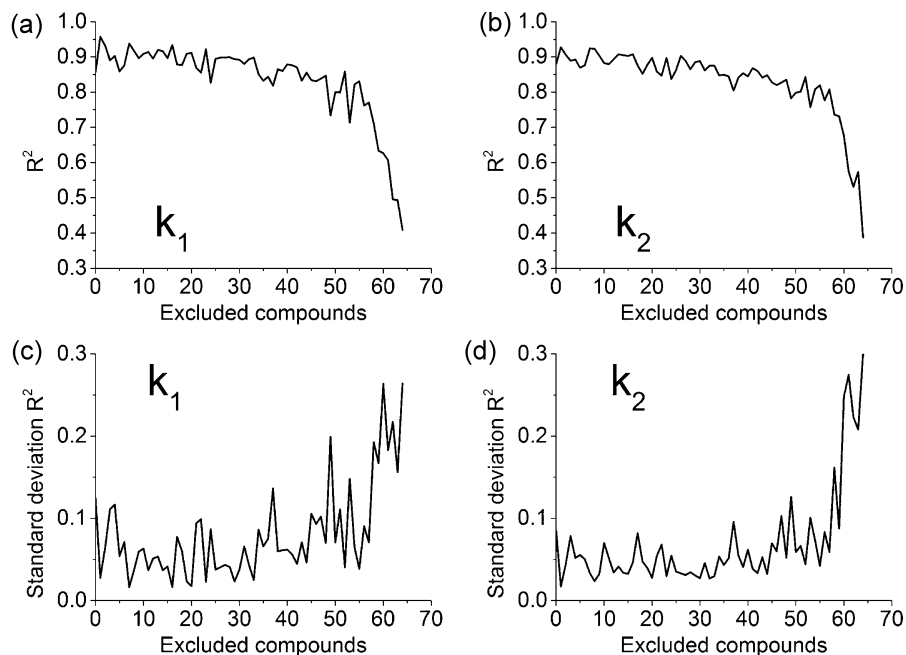
Fig. 8. Left: Average R^2 of 10 prediction experiments for each number of excluded. Right: Standard deviation of the 10 random experiments; at least one side chain per protection group was kept in the model.

Table 3

Group contributions (retention and enantioselectivity increments) (statistically significant values with $p < 0.05$ in bold; for abbreviations of protection groups and amino acids see Table S1 of supplementary material; for full statistics of model see Table 2 and for detailed statistics of increments see supplementary material Tables S2–S4).

Substituent	Retention increment first eluted enantiomer	Retention increment second eluted enantiomer	Enantioselectivity increment	Free energy increments		
	$\ln k_{i,1} \pm SE$	$\ln k_{i,2} \pm SE$	$\ln \alpha_i \pm SE$	$\Delta G_{i,1} \pm SE$	$\Delta G_{i,2} \pm SE$	$\Delta\Delta G_{i,\alpha} \pm SE$
a_0	1.428	1.428	0	−4.21	−4.21	0
PG						
BOC	**0.36 ± 0.15**	**0.57 ± 0.21**	0.2 ± 0.17	−3.13 ± 0.37	−3.63 ± 0.51	−0.5 ± 0.42
Z	**0.71 ± 0.13**	**0.95 ± 0.19**	0.24 ± 0.15	−4 ± 0.33	−4.59 ± 0.46	−0.59 ± 0.37
PNZ	**0.72 ± 0.14**	**1.11 ± 0.2**	**0.38 ± 0.16**	−4.02 ± 0.35	−4.97 ± 0.49	−0.95 ± 0.4
NVOC	**1.07 ± 0.14**	**1.95 ± 0.2**	**0.88 ± 0.16**	−4.88 ± 0.36	−7.06 ± 0.49	−2.17 ± 0.4
DNZ	**1.09 ± 0.14**	**1.76 ± 0.2**	**0.67 ± 0.16**	−4.94 ± 0.35	−6.59 ± 0.49	−1.65 ± 0.4
FMOC	**1.31 ± 0.17**	**1.77 ± 0.24**	**0.46 ± 0.19**	−5.46 ± 0.43	−6.61 ± 0.6	−1.14 ± 0.48
R						
β-Aminobutyric acid	**−1.02 ± 0.16**	**−1.02 ± 0.22**	−0.01 ± 0.18	0.3 ± 0.39	0.31 ± 0.54	0.02 ± 0.44
Citrulline	**−0.8 ± 0.16**	**−0.71 ± 0.22**	0.09 ± 0.18	−0.24 ± 0.4	−0.47 ± 0.55	−0.23 ± 0.45
Homoserine	**−0.8 ± 0.19**	**−0.71 ± 0.27**	0.09 ± 0.22	−0.25 ± 0.48	−0.46 ± 0.67	−0.21 ± 0.55
Gln	**−0.79 ± 0.17**	**−0.82 ± 0.24**	−0.03 ± 0.19	−0.27 ± 0.42	−0.19 ± 0.59	0.08 ± 0.48
β-Aminoisobutyric acid	**−0.71 ± 0.17**	**−0.76 ± 0.24**	−0.05 ± 0.19	−0.46 ± 0.42	−0.35 ± 0.59	0.11 ± 0.48
Met-sulfoxide	**−0.71 ± 0.17**	**−0.74 ± 0.24**	−0.03 ± 0.19	−0.46 ± 0.42	−0.39 ± 0.59	0.07 ± 0.48
Thr	**−0.66 ± 0.16**	**−0.72 ± 0.22**	−0.05 ± 0.18	−0.58 ± 0.4	−0.45 ± 0.55	0.13 ± 0.45
Asn	**−0.65 ± 0.17**	**−0.79 ± 0.24**	−0.14 ± 0.19	−0.63 ± 0.42	−0.28 ± 0.59	0.35 ± 0.48
N-Me-Val	**−0.63 ± 0.19**	**−1.15 ± 0.27**	**−0.51 ± 0.22**	−0.66 ± 0.48	0.61 ± 0.67	1.27 ± 0.55
Val	**−0.59 ± 0.15**	−0.29 ± 0.21	0.3 ± 0.17	−0.76 ± 0.37	−1.5 ± 0.52	−0.74 ± 0.42
α-Aminobutyric acid	**−0.58 ± 0.19**	−0.35 ± 0.27	0.23 ± 0.22	−0.78 ± 0.48	−1.36 ± 0.67	−0.57 ± 0.55
Azetidine carboxylic acid	**−0.57 ± 0.16**	**−1.07 ± 0.22**	**−0.5 ± 0.18**	−0.82 ± 0.39	0.42 ± 0.54	1.24 ± 0.44
Pro	**−0.56 ± 0.15**	**−0.97 ± 0.21**	**−0.41 ± 0.17**	−0.84 ± 0.37	0.18 ± 0.52	1.02 ± 0.42
Met-sulfone	**−0.56 ± 0.17**	**−0.6 ± 0.24**	−0.04 ± 0.19	−0.84 ± 0.42	−0.74 ± 0.59	0.1 ± 0.48
tBu-Gly (Tle)	**−0.55 ± 0.16**	−0.24 ± 0.22	0.31 ± 0.18	−0.86 ± 0.39	−1.63 ± 0.54	−0.77 ± 0.44
Ser	**−0.55 ± 0.15**	**−0.55 ± 0.21**	−0.01 ± 0.17	−0.87 ± 0.38	−0.86 ± 0.53	0.01 ± 0.43
Nval (Nva)	**−0.52 ± 0.16**	−0.26 ± 0.22	0.25 ± 0.18	−0.95 ± 0.4	−1.58 ± 0.55	−0.63 ± 0.45
Leu	**−0.51 ± 0.15**	−0.33 ± 0.21	0.18 ± 0.17	−0.97 ± 0.37	−1.41 ± 0.52	−0.44 ± 0.42
Gly	**−0.5 ± 0.16**			−1 ± 0.41		
Ala	**−0.5 ± 0.15**	**−0.54 ± 0.21**	−0.04 ± 0.17	−1 ± 0.37	−0.89 ± 0.52	0.1 ± 0.42
Dehydro-Pro	**−0.5 ± 0.19**	**−0.99 ± 0.27**	**−0.5 ± 0.22**	−1 ± 0.48	0.24 ± 0.67	1.23 ± 0.55
Ile	**−0.49 ± 0.16**	−0.05 ± 0.22	**0.44 ± 0.18**	−1.03 ± 0.4	−2.11 ± 0.55	−1.09 ± 0.45
allo-Ile	**−0.46 ± 0.17**	0.09 ± 0.24	**0.55 ± 0.19**	−1.08 ± 0.42	−2.44 ± 0.59	−1.36 ± 0.48
N-Me-Leu	**−0.45 ± 0.17**	**−0.91 ± 0.24**	**−0.46 ± 0.19**	−1.11 ± 0.42	0.04 ± 0.59	1.15 ± 0.48
Nleu (Nle)	**−0.44 ± 0.16**	−0.2 ± 0.22	0.25 ± 0.18	−1.13 ± 0.4	−1.74 ± 0.55	−0.61 ± 0.45
β-Ala	**−0.44 ± 0.17**			−1.14 ± 0.42		
α-Me-Leu	**−0.4 ± 0.17**	**−0.82 ± 0.24**	**−0.42 ± 0.19**	−1.24 ± 0.42	−0.19 ± 0.59	1.04 ± 0.48
Pipecolinic acid	**−0.36 ± 0.17**	**−0.79 ± 0.24**	**−0.43 ± 0.19**	−1.34 ± 0.42	−0.27 ± 0.59	1.06 ± 0.48
β-Phe	−0.28 ± 0.16	−0.43 ± 0.22	−0.16 ± 0.18	−1.54 ± 0.39	−1.15 ± 0.54	0.39 ± 0.44
Met	−0.27 ± 0.15	−0.18 ± 0.21	0.09 ± 0.17	−1.56 ± 0.38	−1.78 ± 0.53	−0.22 ± 0.43
p-CF₃-Phe	−0.22 ± 0.2	−0.1 ± 0.28	0.12 ± 0.23	−1.69 ± 0.49	−1.98 ± 0.69	−0.29 ± 0.56
p-Me-Phe	−0.05 ± 0.2	−0.18 ± 0.28	−0.13 ± 0.23	−2.1 ± 0.49	−1.78 ± 0.69	0.32 ± 0.56
Cys	0.01 ± 0.22	0.03 ± 0.3	0.02 ± 0.25	−2.25 ± 0.54	−2.3 ± 0.75	−0.05 ± 0.61
Tyr	0.02 ± 0.15	0.03 ± 0.21	0.01 ± 0.17	−2.27 ± 0.37	−2.29 ± 0.52	−0.02 ± 0.42
Phe	0.02 ± 0.15	0 ± 0.21	−0.01 ± 0.17	−2.27 ± 0.37	−2.24 ± 0.52	0.03 ± 0.42
Phegly (Phg)	0.04 ± 0.16	−0.06 ± 0.22	−0.11 ± 0.18	−2.34 ± 0.4	−2.07 ± 0.55	0.27 ± 0.45
p-iPr-Phe	0.06 ± 0.2	−0.07 ± 0.28	−0.13 ± 0.23	−2.38 ± 0.49	−2.06 ± 0.69	0.32 ± 0.56
o-F-Phe	0.07 ± 0.19	0.24 ± 0.26	0.17 ± 0.21	−2.4 ± 0.47	−2.82 ± 0.65	−0.42 ± 0.53
m-F-Phe	0.1 ± 0.19	0.03 ± 0.26	−0.07 ± 0.21	−2.48 ± 0.47	−2.31 ± 0.65	0.16 ± 0.53
p-tBu-Phe	0.1 ± 0.2	−0.01 ± 0.28	−0.11 ± 0.23	−2.48 ± 0.49	−2.2 ± 0.69	0.28 ± 0.56
p-Cl-Phe	0.14 ± 0.16	0.15 ± 0.22	0.01 ± 0.18	−2.58 ± 0.4	−2.6 ± 0.55	−0.03 ± 0.45

A. Sievers-Engler et al. / J. Chromatogr. A 1363 (2014) 79–88

85

Table 3 (*Continued*)

Substituent	Retention increment first eluted enantiomer	Retention increment second eluted enantiomer	Enantioselectivity increment	Free energy increments		
	$\ln k_{i,1} \pm SE$	$\ln k_{i,2} \pm SE$	$\ln \alpha_i \pm SE$	$\Delta G_{t,1} \pm SE$	$\Delta G_{t,2} \pm SE$	$\Delta\Delta G_{t,\alpha} \pm SE$
p-Nitro-Phe	0.15 ± 0.16	0.18 ± 0.22	0.03 ± 0.18	−2.61 ± 0.4	−2.69 ± 0.55	−0.08 ± 0.45
p-Br-Phe	0.18 ± 0.17	0.16 ± 0.24	−0.02 ± 0.19	−2.68 ± 0.42	−2.63 ± 0.59	0.05 ± 0.48
Lys	0.29 ± 0.16	0.4 ± 0.22	0.11 ± 0.18	−2.95 ± 0.4	−3.22 ± 0.55	−0.27 ± 0.45
Orn	**0.35 ± 0.16**	0.29 ± 0.22	−0.06 ± 0.18	−3.1 ± 0.4	−2.94 ± 0.55	0.15 ± 0.45
Trp	**0.47 ± 0.16**	0.43 ± 0.22	−0.04 ± 0.18	−3.4 ± 0.4	−3.29 ± 0.55	0.11 ± 0.45
p-Amino-Phe	**0.85 ± 0.19**	**0.83 ± 0.27**	−0.02 ± 0.22	−4.34 ± 0.48	−4.29 ± 0.67	0.05 ± 0.55
α-Aminoadipic acid	**1.19 ± 0.17**	**1.14 ± 0.24**	−0.06 ± 0.19	−5.19 ± 0.42	−5.05 ± 0.59	0.14 ± 0.48
Cys-S-trityl	**1.29 ± 0.22**	**1.18 ± 0.3**	−0.11 ± 0.25	−5.43 ± 0.54	−5.16 ± 0.75	0.27 ± 0.61
Glu	**1.32 ± 0.17**	**1.25 ± 0.24**	−0.07 ± 0.19	−5.51 ± 0.42	−5.33 ± 0.59	0.17 ± 0.48
Cysteic acid	**1.39 ± 0.19**	**1.13 ± 0.27**	−0.25 ± 0.22	−5.67 ± 0.48	−5.04 ± 0.67	0.63 ± 0.55
Asp	**1.73 ± 0.17**	**1.26 ± 0.24**	**−0.47 ± 0.19**	−6.52 ± 0.42	−5.34 ± 0.59	1.17 ± 0.48

SE, standard error.

good prediction capability within the range of substituents covered by the test set and remarkable robustness of the model. It becomes evident that a large number of residues can be removed from the set without impairing prediction quality significantly. R^2 for predictions remained nearly constant above 0.8 until 60 compounds had been excluded (Fig. 8a). At the same time, standard deviations for predictions remained smaller than 0.2 and only increased significantly once about 60 compounds or more have been removed and predicted (Fig. 8b). All these results confirm the general validity of the additivity principle of group contributions to retention and enantioselectivity, and of the general linear model employed herein, for the given data set.

3.5. Retention increments and group contributions to enantioselectivity

The refined data set was utilized to calculate group contributions for protection groups and amino acid residues (Table 3). Besides retention increments for first and second eluted enantiomers, group contributions to enantioselectivity and corresponding energy increments for retention and enantioselectivity are reported as well. From these group contributions the corresponding response (*y*-values) can be readily calculated from the sum of increments for protection group and amino acid residue and the respective intercept a_0. For instance, the retention factor of the first eluted enantiomer ($\ln k_1$) of Z-Met can be calculated from $a_0 + \ln k_{i,1}(\text{Z}) + \ln k_{i,1}(\text{Met}) = 1.428 + 0.71 + (-0.27) = 1.868$ (experimental value is 1.914). A negative value of the group contribution means that presence of the substituent reduces the *y*-value relative to the offset and a positive value means that it increases the response. Full statistics of each increment is presented in the supplementary material (Tables S2–S4). *p*-Values in Table S2 and S3 provide information on the statistical significance of the corresponding group contribution. *p*-Values <0.05 signify statistical significance at the 95% level. Table S4 reports calculated responses and corresponding residuals with regards to experimental values. Here it should be mentioned that the experimental data incorporate all kind of variabilities as they have been generated on different days with different batches of mobile phase, with different lots of the chiral stationary phase, by different operators, on different HPLC instruments. If all experiments were performed by the same operator, on the same column with the same mobile phase on the same instrument within one batch of analysis, residuals for predictions were most likely significantly smaller. Overall, run-to-run repeatabilities of chromatographic data as well as batch-to-batch reproducibilities of the chiral stationary phase are good. RSD values <1% for *k*-values and <0.5% for α-values are typically obtained for run-to-run repeatabilities (e.g. RSD values of 0.55%, 0.68%, and 0.15% have been measured for k_1, k_2, and α of Z-Phe; $n = 5$). Further, batch-to-batch reproducibilities ($n = 3$) of 1.24, 1.01, and <0.1% RSD for k_1, k_2, and α of Z-Phe have been determined for 3 consecutive batches of CSP. In Table 3, functional group contributions are divided into two subgroups, namely those of protection groups and amino acid residues. They are listed in order of increasing retention increments for first eluted enantiomers. It can be seen that in any case protection groups significantly contribute to retention with increase of $\ln k_{i,1}$ in the order BOC < Z < PNZ < NVOC < DNZ < FMOC. For second eluted enantiomers this order is essentially obeyed as well, with exception for NVOC protection group which displayed the highest retention contribution of all protection groups investigated. In terms of enantioselectivity increments, this order is perturbed. The FMOC group shows lower enantioselectivity than DNZ and NVOC groups and in particular the latter is remarkably favorable for enantiomer separation of amino acid derivatives on the quinine carbamate CSP. The increase in enantioselectivity increment from

Z < PNZ < DNZ clearly indicates the existence and importance of a stereoselective π–π-interaction between these aromatic groups and the quinolone of the selector (Fig. 1c).

With regards to retention increments for amino acid residues of first eluted enantiomers the following findings can be summarized. Acidic residues (α-aminoadipic acid, Glu, cysteic acid, and Asp) provide the largest retention increments, followed by aromatic amino acid residues. Basic amino acids like Lys, Orn are within the group of aromatic amino acids since the side chain carries a protection group and the same is valid for S-protected Cys (Cys-S-trityl). In general, β-amino acids feature lower retention increments than corresponding α-amino acids (cf. α/β-aminobutyric acid and α/β-Phe; for α/β-Ala the difference is not significant). For the remaining amino acids it holds that polar ones contribute a lower retention increment compared to hydrophobic ones and sterically bulky side chains display smaller group contributions in comparison to the corresponding n-alkyl analogs. In many instances, there is no statistically significant difference between retention increments for first and second eluted enantiomer which indicates that the amino acid side chains do not contribute significantly to enantioselectivity. This can also be explained by the binding model in Fig. 1c which shows that the amino acid residue is oriented toward the tert-butyl group of the selector. While bulky alkyl moieties may be capable of interacting via van der Waals type interactions, hydrophilic residues will not undergo any interaction under employed hydroorganic conditions. It may be striking for readers that there are many cases in which the retention increment for the second enantiomer is even lower than for the first eluted enantiomer. A closer look, however, reveals that in all these cases the difference in retention increments of amino acid residues between first and second eluted enantiomer are not significant and the difference in retention is essentially due to stereoselective interaction of the protection group. Corresponding enantioselectivity increments $\log \alpha_i$ confirm the above findings and interpretations. Besides secondary amino acids (N-Me-Val, azetidine carboxylic acid, Pro, dehydro-Pro, N-Me-Leu, and pipecolinic acid), α-methylated amino acids (α-Me-Leu), Asp and a few aliphatic amino acids (Tle, Ile, and allo-Ile) all other group contributions to enantioselectivity were not really significant from the statistical point of view. Amongst all these group contributions only the bulky aliphatic Tle, Ile, and allo-Ile provided an increase in enantioselectivity (positive sign) while the others significantly reduced enantioselectivities (negative sign of group contributions). Secondary amino acids show greatly reduced separation factors because the supportive enantioselective H-bond interaction does not exist due to lack of the N—H hydrogen donor (Fig. 1c). A competitive acidic group in the amino acid side chain turns out to be negative for enantioselectivity solely if it is located close to the stereogenic center (cysteic acid and Asp, both with considerable negative enantioselectivity increment of -0.25 ± 0.22 and -0.47 ± 0.19, respectively). Once the distance between α-carbon and carboxylic acid moiety increases, the acidic group of the side chain appears to be no longer competitive to the α-carboxylic group. No significant negative enantioselectivity increment is therefore found for α-amino adipic acid and Glu (-0.06 ± 0.19 and -0.07 ± 0.19, respectively). For the remaining amino acids the enantiomer separation is dominated by the protection group with negligible influence of the amino acid residues (see Table 3).

4. Conclusion

Enantioselective analysis of amino acids is an important topic in bioanalysis and quality control of educts in peptide synthesis. A tert-butylcarbamoylated quinine-based chiral stationary phase was utilized to separate the enantiomers of a large set of proteinogenic

and unusual amino acid enantiomers after derivatization of the amino functionality with chloroformate type reagents. Obtained chromatographic retention and separation factors were subjected to quantitative structure–property relationship studies. A general linear model based on a Free-Wilson matrix as structural descriptors was utilized to derive retention increments for first and second eluted enantiomers as well as group contributions to enantioselectivity. The results demonstrated that the additivity principle of group contributions was largely obeyed for the majority of solutes in the data set. Only a few basic amino acids (Arg, His) were eliminated due to departure from a linear model leading to outliers. If significant cooperative effects were present, the employed linear model would not be capable of predicting retention and separation factors with sufficient accuracy.

Derived binding and enantioselectivity increments gave insights into contributions dominating retention and enantioselectivity. Protection groups were largely driving retention and enantioselectivity, particularly in case of aromatic ones. In contrast, the contribution of amino acid residues to enantioselectivity was only significant for secondary amino acids, α-methylated amino acids, aspartic acid and a few sterically bulky aliphatic amino acid residues (Tle, Ile, allo-Ile). Amongst them only the latter group contributed positively to enantioselectivity while the other residues mentioned reduced enantioselectivity significantly. With help of derived group contributions retention and enantioselectivity values of amino acid derivatives not present in the experimental data set can be readily estimated by simple calculations. However, this kind of QSPR has only limited utility for predictions, yet is valuable for interpretation of retention mechanisms. The presented model is illustrative and simple to implement. It may be used for optimization of derivatizing agent and guide researchers in selection of suitable reagents based on a limited set of experimental data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chroma.2014.04.077.

References

[1] H.Y. Aboul-Enein, I.W. Wainer (Eds.), The Impact of Stereochemistry on Drug Development and Use, John Wiley & Sons, New York, 1997.
[2] N.M. Maier, P. Franco, W. Lindner, J. Chromatogr. A 906 (2001) 3.
[3] M. Lämmerhofer, J. Chromatogr. A 1217 (2010) 814.
[4] N.M. Maier, LindnerF W., in: E. Francotte, W. Lindner (Eds.), Chirality in Drug Research, Wiley-VCH, Weinheim, 2006, p. 189.
[5] A. Del Rio, P. Piras, C. Roussel, Chirality 18 (2006) 498.
[6] A. Chmielewska, T. Baczek, J. AOAC Int. 95 (2012) 624.
[7] A. Knorr, A. Monge, M. Stueber, A. Stratmann, D. Arndt, E. Martin, P. Pospisil, Anal. Chem. 85 (2013) 11216.
[8] M. Gilar, H. Xie, A. Jaworski, Anal. Chem. 82 (2009) 265.
[9] J.E. Madden, M.J. Shaw, G.W. Dicinoski, N. Avdalovic, P.R. Haddad, Anal. Chem. 74 (2002) 6023.
[10] J.L. Adcock, M. Adams, B.S. Mitrevski, P.J. Marriott, Anal. Chem. 81 (2009) 6797.
[11] T.H. Miller, A. Musenga, D.A. Cowan, L.P. Barron, Anal. Chem. 85 (2013) 10330.
[12] V. Spicer, A. Yamchuk, J. Cortens, S. Sousa, W. Ens, K.G. Standing, J.A. Wilkins, O.V. Krokhin, Anal. Chem. 79 (2007) 8762.
[13] T.D. Booth, I.W. Wainer, J. Chromatogr. A 741 (1996) 205.
[14] R. Kaliszan, T.A.G. Noctor, I.W. Wainer, Chromatographia 33 (1992) 546.
[15] T.D. Booth, I.W. Wainer, J. Chromatogr. A 737 (1996) 157.
[16] C. Altomare, A. Carotti, S. Cellamare, F. Fanelli, F. Gasparrini, C. Villani, P.A. Carrupt, B. Testa, Chirality 5 (1993) 527.
[17] C. Altomare, S. Cellamare, A. Carotti, M.L. Barreca, A. Chimirri, A.-M. Monforte, F. Gasparrini, C. Villani, M. Cirilli, F. Mazza, Chirality 8 (1996) 556.

[18] M. Lämmerhofer, P. Franco, W. Lindner, J. Sep. Sci. 29 (2006) 1486.
[19] B. Natalini, R. Sardella, A. Macchiarulo, R. Pellicciari, Chirality 18 (2006) 509.
[20] E. Calleri, E. Lorenzi, D. Siluk, M. Markuszewski, R. Kaliszan, G. Massolini, Chromatographia 55 (2002) 651.
[21] C.A. Montanari, Q.B. Cass, M.E. Tiritan, A.L.S.d. Souza, Anal. Chim. Acta 419 (2000) 93.
[22] R. Sardella, F. Ianni, N. Giacche, A. Lisanti, P. Conti, A. Pinto, L. Tamborini, B. Natalini, Trends Chromatogr. 7 (2012).
[23] K. Boronová, J. Lehotay, K. Hroboňová, D.W. Armstrong, J. Chromatogr. A 1301 (2013) 38.
[24] B. Rasulev, M. Turabekova, M. Gorska, K. Kulig, A. Bielejewska, J. Lipkowski, J. Leszczynski, Chirality 4 (2012) 72.
[25] A. Berthod, C.R. Mitchell, D.W. Armstrong, J. Chromatogr. A 1166 (2007) 61.
[26] C.R. Mitchell, D.W. Armstrong, A. Berthod, J. Chromatogr. A 1166 (2007) 70.
[27] J. Lokajova, E. Tesarova, D.W. Armstrong, J. Chromatogr. A 1088 (2005) 57.
[28] J. Vozka, K. Kalíková, C. Roussel, D.W. Armstrong, E. Tesařová, J. Sep. Sci. 36 (2013) 1711.
[29] C. West, Y. Zhang, L. Morin-Allory, J. Chromatogr. A 1218 (2011) 2019.
[30] S. Schefzick, M. Lämmerhofer, W. Lindner, K.B. Lipkowitz, M. Jalaie, Chirality 12 (2000) 742.
[31] W.M.F. Fabian, W. Stampfer, M. Mazur, G. Uray, Chirality 15 (2003) 271.
[32] T. Suzuki, S. Timofei, B.E. Iuoras, G. Uray, P. Verdino, W.M.F. Fabian, J. Chromatogr. A 922 (2001) 13.
[33] A. Carotti, C. Altomare, S. Cellamare, A. Monforte, G. Bettoni, F. Loiodice, N. Tangari, V. Tortorella, J. Comput. Aided Mol. Des. 9 (1995) 131.
[34] M. Szaleniec, A. Dudzik, M. Pawul, B. Kozik, J. Chromatogr. A 1216 (2009) 6224.
[35] S. Choi, S. Lee, H.-Y. Choo, Arch. Pharm. Res. 27 (2004) 1009.
[36] B. Natalini, A. Macchiarulo, R. Sardella, A. Massarotti, R. Pellicciari, J. Sep. Sci. 31 (2008) 2395.
[37] A. Vailaya, C. Horvath, J. Phys. Chem. 100 (1996) 2447.
[38] W. Bicker, I. Chiorescu, V.B. Arion, M. Lämmerhofer, W. Lindner, Tetrahedron: Asymmetry 19 (2008) 97.
[39] H. Kubinyi, QSAR: Hansch Analysis and Related Approaches, VCH, Weinheim, 1993.
[40] A. Berthod, S.C. Chang, D.W. Armstrong, Anal. Chem. 64 (1992) 395.
[41] C. Czerwenka, M. Lämmerhofer, N.M. Maier, K. Rissanen, W. Lindner, Anal. Chem. 74 (2002) 5658.
[42] A. Mandl, L. Nicoletti, M. Lämmerhofer, W. Lindner, J. Chromatogr. A 858 (1999) 1.
[43] C. Czerwenka, M. Lämmerhofer, W. Lindner, J. Sep. Sci. 26 (2003) 1499.

Supplementary Material

Ligand-Receptor Binding Increments in

Enantioselective Liquid Chromatography

Adrian Sievers-Engler [1], Wolfgang Lindner [2], Michael Lämmerhofer [1] ∗

[1] Institute of Pharmaceutical Sciences, University of Tübingen, Auf der Morgenstelle 8, 72076

Tübingen, Germany

[2] University of Vienna, Department of Analytical Chemistry,

Währinger Strasse 38, A-1090 Vienna, Austria

AUTHOR INFORMATION

Corresponding Author

*Author for correspondence: Tel +49 7071 29 78793, Fax: +49 7071 29 4565

E-mail address: Michael.Laemmerhofer@uni-tuebingen.de

Table S1: Experimental retention and separation factors utilized in this study for calculations of retention and selectivity increments

Compound	k_1	k_2	α
BOC-Ala	3.328	3.734	1.122
BOC-Arg	5.836	6.898	1.182
BOC-p-CF$_3$-Phe [a]	4.833	6.633	1.372
BOC-p-Cl-Phe	8.261	10.739	1.300
BOC-p-iPr-Phe [a]	6.393	6.867	1.074
BOC-p-Me-Phe [a]	5.700	6.120	1.074
BOC-p-Nitro-Phe [a]	7.193	10.153	1.411
BOC-p-tBu-Phe [a]	6.653	7.260	1.091
BOC-Leu [a]	3.719	4.496	1.209
BOC-Phe	6.682	8.145	1.219
BOC-Pro	2.486	2.486	1.000
BOC-Tyr	6.270	7.819	1.247
BOC-Val	3.551	4.957	1.396
DNZ-Ala	6.885	13.439	1.952
DNZ-allo-Ile	7.784	28.872	3.709
DNZ-α-Aminoadipic acid	41.000	78.980	1.926
DNZ-α-Aminobutyric acid	6.946	17.054	2.455
DNZ-α-Me-Leu	8.345	9.264	1.110
DNZ-Arg	0.899	1.912	2.128
DNZ-Asn	6.514	10.926	1.677
DNZ-Asp	84.318	86.439	1.025
DNZ-Azetidine carboxylic acid [a]	8.450	10.056	1.190

Compound	k_1	k_2	α
DNZ-β-Ala	9.236		
DNZ-β-Aminobutyric acid	5.209	11.351	2.179
DNZ-β-Aminoisobutyric acid	6.318	12.459	1.972
DNZ-β-Phe [b]	8.730	15.976	1.830
DNZ-Citrulline	5.236	11.811	2.255
DNZ-dehydro-Pro	7.574	8.973	1.185
DNZ-Gln	5.507	10.757	1.953
DNZ-Glu	43.345	79.054	1.824
DNZ-Gly	9.405		
DNZ-His	2.547	4.662	1.830
DNZ-Homoserine	5.595	11.885	2.124
DNZ-Ile	7.784	27.182	3.492
DNZ-p-Amino-Phe	29.128	55.703	1.912
DNZ-p-Br-Phe	14.723	28.115	1.910
DNZ-p-Cl-Phe	12.892	24.784	1.922
DNZ-p-Nitro-Phe	14.027	26.405	1.882
DNZ-Leu	7.649	21.412	2.799
DNZ-Lys	16.892	36.764	2.176
DNZ-Met	9.155	21.291	2.325
DNZ-Met-Sulfone	6.595	12.669	1.921
DNZ-Met-Sulfoxide	5.568	10.622	1.908
DNZ-*N*-Me-Leu	7.703	8.709	1.131
DNZ-*N*-Me-Val	6.601	7.709	1.168
DNZ-NLeu (DNZ-Nle)	7.973	21.216	2.661
DNZ-NVal (DNZ-Nva)	7.378	19.831	2.688
DNZ-Orn	20.115	36.709	1.825

Compound	k_1	k_2	α
DNZ-Phe	11.101	21.595	1.945
DNZ-PheGly (DNZ-Phg)	12.243	22.426	1.832
DNZ-Pipecolinic acid	8.892	10.196	1.147
DNZ-Pro	7.520	9.088	1.208
DNZ-Ser	6.514	13.074	2.007
DNZ-tBu-Gly (DNZ-Tle)	7.700	23.485	3.050
DNZ-Thr	6.520	8.412	1.290
DNZ-Trp	18.736	35.034	1.870
DNZ-Tyr	10.378	20.966	2.020
DNZ-Val	7.142	23.270	3.258
FMOC-Arg [a]	1.143	1.915	1.675
FMOC-Cys [b]	15.514	25.164	1.622
FMOC-Cys-S-Trityl [b]	55.893	79.648	1.425
FMOC-Leu	9.260	17.594	1.900
FMOC-Phe [a]	15.650	24.414	1.560
NVOC-Ala	7.074	18.115	2.561
NVOC-Arg	7.973	19.162	2.403
NVOC-Azetidine carboxylic acid	6.061	6.761	1.116
NVOC-β-Phe	9.128	16.318	1.788
Bis-(NVOC)-His [a]	6.047	14.439	2.388
NVOC-Citrulline	5.757	15.615	2.712
NVOC-His [a]	2.088	4.568	2.188
NVOC-Ile [a]	7.142	24.709	3.460
NVOC-Leu [a]	7.162	25.723	3.592
NVOC-Lys [a]	16.568	48.169	2.907
NVOC-Met	9.466	26.331	2.782

Compound	k_1	k_2	α
NVOC-Met-Sulfone	7.520	16.858	2.242
NVOC-Met-Sulfoxide	6.547	15.128	2.311
NVOC-NLeu (NVOC-Nle) [a]	7.777	25.574	3.288
NVOC-NVal (NVOC-Nva) [a]	7.270	24.426	3.360
NVOC-Orn	18.784	43.412	2.311
NVOC-Phe	11.919	28.784	2.415
NVOC-PheGly (NVOC-Phg) [a]	14.370	28.596	1.990
NVOC-Pro	6.860	7.409	1.080
NVOC-Ser	6.858	17.453	2.545
NVOC-Thr	6.331	18.385	2.904
NVOC-Trp [a]	19.034	41.541	2.182
NVOC-Tyr	11.034	26.872	2.435
NVOC-Val [a]	6.838	22.838	3.340
PNZ-Ala	5.331	7.128	1.337
PNZ-allo-Ile	5.459	12.568	2.302
PNZ-α- Aminoadipic acid	28.392	37.669	1.327
PNZ-α-Me-Leu	5.750	6.378	1.109
PNZ-Arg	0.662	0.980	1.480
PNZ-Asn	4.514	5.818	1.289
PNZ-Asp	40.378	43.703	1.082
PNZ-α-Aminobutyric acid	4.818	7.568	1.571
PNZ-β-Aminobutyric acid	3.682	5.088	1.382
PNZ-β-Aminoisobutyric acid	4.061	5.372	1.323
Bis-(PNZ)-His	26.336	33.025	1.254
PNZ-Citrulline	3.885	5.784	1.489
PNZ-Cysteic acid	34.453	39.189	1.137

Compound	k_1	k_2	α
PNZ-Gln	4.014	5.507	1.372
PNZ-Glu	34.730	47.345	1.363
PNZ-His	1.953	2.764	1.415
PNZ-Ile	5.459	11.541	2.114
PNZ-p-Br-Phe	10.439	15.054	1.442
PNZ-p-Cl-Phe	9.189	13.291	1.446
PNZ-p-Nitro-Phe	10.061	14.574	1.449
PNZ-Leu	6.115	10.182	1.665
PNZ-Lys	11.169	16.764	1.501
PNZ-Met	6.439	9.736	1.512
PNZ-*N*-Me-Leu	5.642	5.642	1.000
PNZ-NLeu (PNZ-Nle)	5.574	9.169	1.645
PNZ-NVal (PNZ-Nva)	5.169	8.426	1.630
PNZ-Orn	9.845	13.351	1.356
PNZ-Phe	9.216	13.128	1.424
PNZ-PheGly (PNZ-Phg)	8.439	11.486	1.361
PNZ-Pipecolinic acid	5.858	6.196	1.058
PNZ-Pro	5.257	5.466	1.040
PNZ-Ser	4.574	6.304	1.378
PNZ-tBu-Gly (PNZ-Tle)	5.324	10.459	1.964
PNZ-Thr	4.291	6.736	1.570
PNZ-Trp	15.074	22.223	1.474
PNZ-Tyr	8.439	12.331	1.461
PNZ-Val	5.095	10.345	2.031
Z-Ala [a]	6.429	7.413	1.153
Z-Arg	0.622	1.149	1.848

Compound	k_1	k_2	α
Z-Azetidine caboxylic acid	4.560	4.560	1.000
Z-β-Aminobutyric acid	2.240	2.643	1.180
Z-β-Phe	7.070	7.989	1.130
Z-Gly	4.169		
Z-His	3.392	3.824	1.127
Z-o-F-Phe [a]	9.133	13.687	1.499
Z-m-F-Phe [a]	9.420	11.167	1.185
Z-p-F-Phe [a]	8.520	10.793	1.267
Z-Leu	4.196	5.331	1.271
Z-Met [a]	6.780	8.633	1.273
Z-Phe [a]	8.950	10.919	1.220
Z-Pro	6.004	6.004	1.000
Z-Ser	6.099	7.374	1.209
Z-tBu-Gly (Z-Tle)	4.220	6.499	1.540
Z-Tyr	11.996	14.875	1.240
Z-Val	3.919	6.007	1.533

[a] elution order: R < S; [b] S < R

Abbreviations: *Protection groups:* Z, benzyloxycarbonyl; BOC, *tert*-butoxycarbonyl; PNZ, 4-nitrobenzyloxycarbonyl; DNZ, 3,5-dinitrobenzyloxycarbonyl; FMOC, 9-fluorenylmethoxycarbonyl; NVOC, o-nitroveratryloxycarbonyl (4,5-dimethoxy-2-nitrobenzyloxycarbonyl). *Amino acids:* 3-letter code of proteinogenic amino acids see biochemistry text book; special amino acids: p-CF$_3$-Phe, *p*-trifluoromethylphenylalanine; p-Cl-Phe, *p*-chlorophenylalanine; p-iPr-Phe, *p*-isopropylphenylalanine; p-Me-Phe, *p*-methylphenylalanine; p-Nitro-Phe, *p*-nitrophenylalanine; p-tBu-Phe, *p*-(*tert*-butyl)phenylalanine; α-Me-Leu, alpha-methyl-leucine; *N*-Me-Leu, *N*-methyl-leucine; *N*-Me-Val, *N*-methyl-valine; NLeu/Nle, norleucine; NVal/Nva, norvaline; PheGly/Phg, α-phenylglycine; tBu-Gly/Tle, α-(*tert*-butyl)glycine/*tert*-leucine
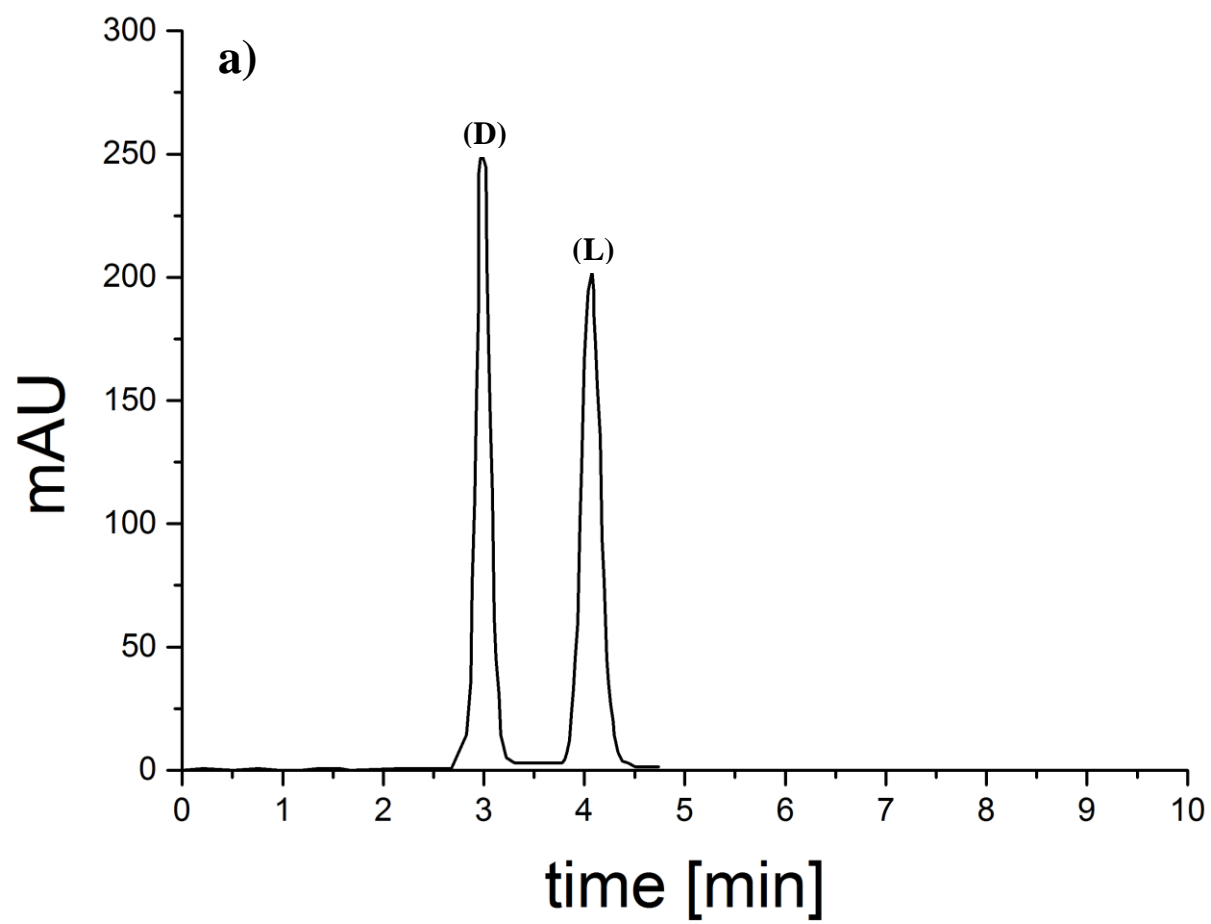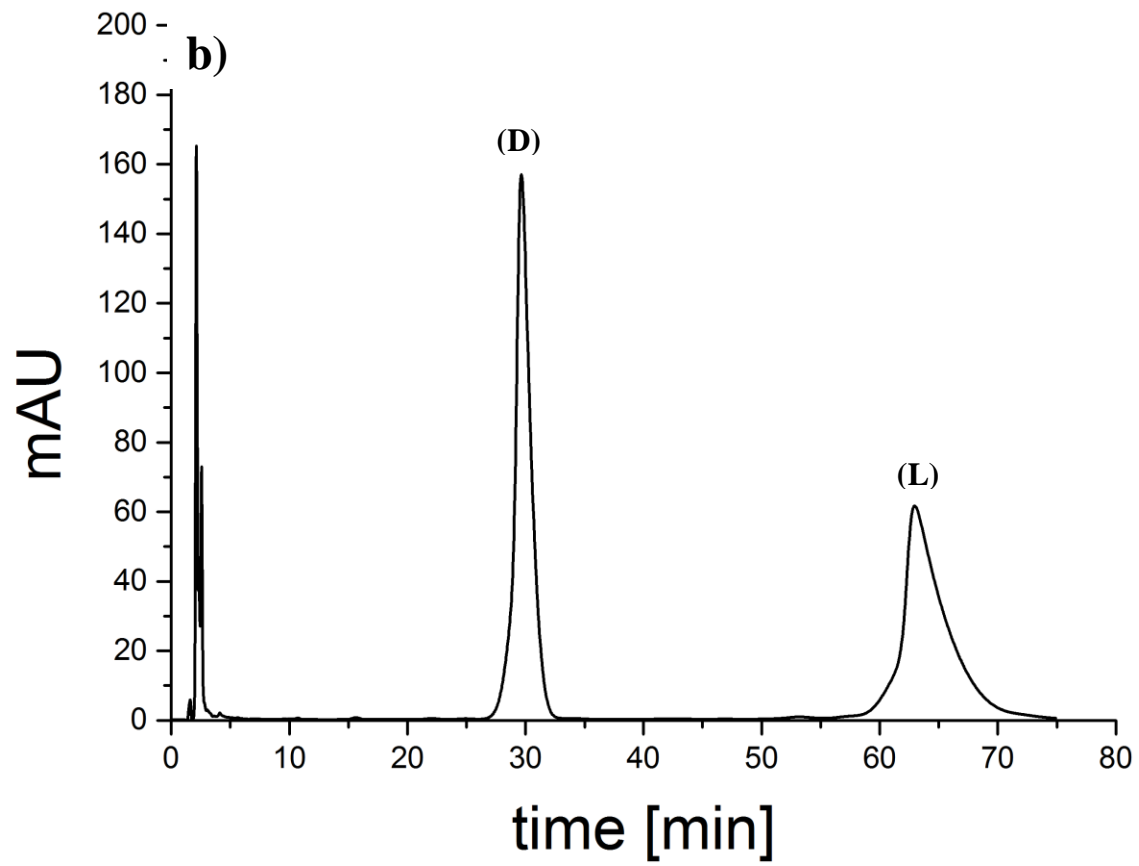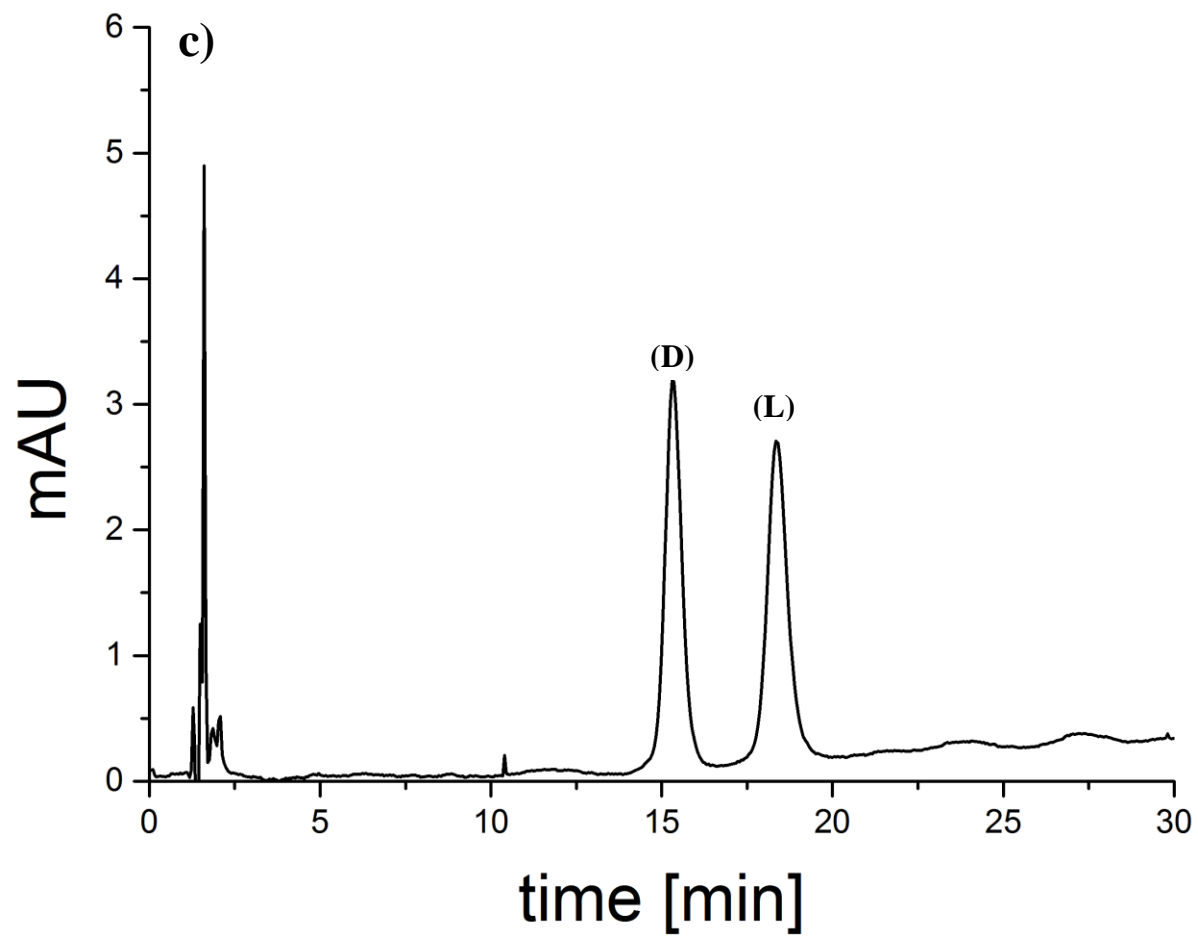
Figure S1: Representative chromatograms: a) FMOC-Arg, b) NVOC-Trp, c) Z-Phe; conditions as described in experimental section of main document.

Table S2: Retention increments and statistical parameters for both enantiomers

Substituent	retention increment first eluting enantiomer				retention increment second eluting enantiomer			
	Incre-ment	95% CI	t-value	p-value	Increment	95% CI	t-value	p-value
a0	1.428	-	-	-	1.428	-	-	-
PG								
BOC	0.36	0.29	2.48	1.55E-02	0.57	0.40	2.75	7.55E-03
Z	0.71	0.26	5.39	8.40E-07	0.95	0.36	5.14	2.29E-06
PNZ	0.72	0.28	5.12	2.41E-06	1.11	0.39	5.61	3.50E-07
NVOC	1.07	0.28	7.56	1.03E-10	1.95	0.39	9.83	6.10E-15
DNZ	1.09	0.28	7.78	3.90E-11	1.76	0.38	8.97	2.38E-13
FMOC	1.31	0.34	7.63	7.58E-11	1.77	0.47	7.39	2.11E-10
R								
β-Aminobutyric acid	-1.02	0.31	-6.53	7.98E-09	-1.02	0.43	-4.70	1.20E-05
Citrulline	-0.80	0.31	-5.05	3.26E-06	-0.71	0.44	-3.19	2.12E-03
Homoserine	-0.80	0.38	-4.14	9.31E-05	-0.71	0.53	-2.64	1.01E-02
Gln	-0.79	0.33	-4.68	1.30E-05	-0.82	0.46	-3.49	8.31E-04
β-Aminoisobutyric acid	-0.71	0.33	-4.24	6.55E-05	-0.76	0.46	-3.23	1.87E-03
Met-Sulfoxide	-0.71	0.33	-4.23	6.72E-05	-0.74	0.46	-3.16	2.33E-03
Thr	-0.66	0.31	-4.18	8.09E-05	-0.72	0.44	-3.22	1.90E-03
Asn	-0.65	0.33	-3.84	2.67E-04	-0.79	0.46	-3.34	1.33E-03
N-Me-Val	-0.63	0.38	-3.28	1.59E-03	-1.15	0.53	-4.25	6.36E-05
Val	-0.59	0.29	-3.96	1.73E-04	-0.29	0.41	-1.40	1.64E-01
α-Aminobutyric acid	-0.58	0.38	-3.02	3.50E-03	-0.35	0.53	-1.31	1.96E-01
Azetidine carboxylic acid	-0.57	0.31	-3.66	4.82E-04	-1.07	0.43	-4.91	5.54E-06
Pro	-0.56	0.29	-3.75	3.51E-04	-0.97	0.41	-4.66	1.41E-05
Met-Sulfone	-0.56	0.33	-3.32	1.42E-03	-0.60	0.46	-2.55	1.29E-02
tBu-Gly (Tle)	-0.55	0.31	-3.55	6.78E-04	-0.24	0.43	-1.11	2.70E-01
Ser	-0.55	0.30	-3.60	5.75E-04	-0.55	0.42	-2.60	1.12E-02

Substituent	Incre-ment	95% CI	t-value	p-value	Increment	95% CI	t-value	p-value
Nval (Nva)	-0.52	0.31	-3.24	1.81E-03	-0.26	0.44	-1.18	2.43E-01
Leu	-0.51	0.29	-3.43	9.97E-04	-0.33	0.41	-1.60	1.14E-01
Gly	-0.50	0.32	-3.03	3.36E-03				
Ala	-0.50	0.29	-3.33	1.38E-03	-0.54	0.41	-2.58	1.19E-02
Dehydro-Pro	-0.50	0.38	-2.57	1.22E-02	-0.99	0.53	-3.69	4.39E-04
Ile	-0.49	0.31	-3.05	3.19E-03	-0.05	0.44	-0.21	8.31E-01
allo-Ile	-0.46	0.33	-2.74	7.73E-03	0.09	0.46	0.36	7.17E-01
N-Me-Leu	-0.45	0.33	-2.67	9.27E-03	-0.91	0.46	-3.89	2.24E-04
Nleu (Nle)	-0.44	0.31	-2.78	6.95E-03	-0.20	0.44	-0.88	3.82E-01
β-Ala	-0.44	0.33	-2.60	1.12E-02				
α-Me-Leu	-0.40	0.33	-2.38	2.00E-02	-0.82	0.46	-3.49	8.17E-04
Pipecolinic acid	-0.36	0.33	-2.14	3.61E-02	-0.79	0.46	-3.35	1.28E-03
β-Phe	-0.28	0.31	-1.77	8.04E-02	-0.43	0.43	-1.99	5.01E-02
Met	-0.27	0.30	-1.77	8.05E-02	-0.18	0.42	-0.85	4.01E-01
p-CF3-Phe	-0.22	0.39	-1.10	2.76E-01	-0.10	0.54	-0.37	7.16E-01
p-Me-Phe	-0.05	0.39	-0.26	7.92E-01	-0.18	0.54	-0.66	5.14E-01
Cys	0.01	0.42	0.04	9.70E-01	0.03	0.59	0.10	9.20E-01
Tyr	0.02	0.29	0.12	9.07E-01	0.03	0.41	0.13	9.00E-01
Phe	0.02	0.29	0.12	9.04E-01	0.00	0.41	0.02	9.84E-01
Phegly (Phg)	0.04	0.31	0.28	7.84E-01	-0.06	0.44	-0.29	7.72E-01
p-iPr-Phe	0.06	0.39	0.32	7.53E-01	-0.07	0.54	-0.24	8.11E-01
o-F-Phe	0.07	0.37	0.37	7.12E-01	0.24	0.51	0.91	3.67E-01
m-F-Phe	0.10	0.37	0.54	5.94E-01	0.03	0.51	0.13	8.97E-01
p-tBu-Phe	0.10	0.39	0.52	6.07E-01	-0.01	0.54	-0.04	9.69E-01
p-Cl-Phe	0.14	0.31	0.88	3.81E-01	0.15	0.44	0.68	4.98E-01
p-Nitro-Phe	0.15	0.31	0.96	3.42E-01	0.18	0.44	0.83	4.09E-01
p-Br-Phe	0.18	0.33	1.08	2.85E-01	0.16	0.46	0.69	4.92E-01
Lys	0.29	0.31	1.84	7.06E-02	0.40	0.44	1.80	7.61E-02
Orn	0.35	0.31	2.20	3.10E-02	0.29	0.44	1.30	1.98E-01
Trp	0.47	0.31	2.97	4.03E-03	0.43	0.44	1.93	5.78E-02
p-Amino-Phe	0.85	0.38	4.41	3.61E-05	0.83	0.53	3.08	2.91E-03
α-Aminoadipic acid	1.19	0.33	7.09	7.46E-10	1.14	0.46	4.84	7.27E-06

Substituent	Incre-ment	95% CI	t-value	p-value	Increment	95% CI	t-value	p-value
Cys-S-Trityl	1.29	0.42	5.96	8.55E-08	1.18	0.59	3.91	2.07E-04
Glu	1.32	0.33	7.86	2.84E-11	1.25	0.46	5.32	1.10E-06
Cysteic acid	1.39	0.38	7.18	5.23E-10	1.13	0.53	4.19	7.70E-05
Asp	1.73	0.33	10.28	9.06E-16	1.26	0.46	5.34	1.02E-06

Table S3: Selectivity increments and statistical parameters.

	Enantioselectivity increment			
Substituent	**Increment**	**95% CI**	**t-value**	**p-value**
a0	0	-	-	-
PG				
BOC	0.20	0.33	1.20	2.36E-01
Z	0.24	0.30	1.57	1.21E-01
PNZ	0.38	0.31	2.39	1.96E-02
FMOC	0.46	0.32	2.37	2.06E-02
DNZ	0.67	0.31	4.17	8.32E-05
NVOC	0.88	0.38	5.42	7.41E-07
R				
N-Me-Val	-0.51	0.35	-2.33	2.26E-02
Azetidine carboxylic acid	-0.50	0.36	-2.81	6.36E-03
Dehydro-Pro	-0.50	0.43	-2.26	2.66E-02
Asp	-0.47	0.38	-2.47	1.59E-02
N-Me-Leu	-0.46	0.38	-2.42	1.80E-02
Pipecolinic acid	-0.43	0.36	-2.24	2.83E-02
α-Me-Leu	-0.42	0.38	-2.20	3.12E-02
Pro	-0.41	0.43	-2.42	1.80E-02
Cysteic acid	-0.25	0.43	-1.16	2.52E-01

Substituent	Increment	95% CI	t-value	p-value
β-Phe	-0.16	0.35	-0.89	3.78E-01
Asn	-0.14	0.33	-0.73	4.68E-01
p-Me-Phe	-0.13	0.38	-0.57	5.68E-01
p-iPr-Phe	-0.13	0.35	-0.57	5.69E-01
p-tBu-Phe	-0.11	0.34	-0.50	6.17E-01
Phegly (Phg)	-0.11	0.36	-0.60	5.52E-01
Cys-S-Trityl	-0.11	0.33	-0.44	6.65E-01
Glu	-0.07	0.37	-0.37	7.16E-01
m-F-Phe	-0.07	0.33	-0.31	7.56E-01
Orn	-0.06	0.43	-0.34	7.38E-01
α-Aminoadipic acid	-0.06	0.36	-0.29	7.70E-01
Thr	-0.05	0.38	-0.28	7.77E-01
β-Aminoisobutyric acid	-0.05	0.38	-0.24	8.11E-01
Trp	-0.04	0.36	-0.24	8.08E-01
Met-Sulfone	-0.04	0.38	-0.22	8.29E-01
Ala	-0.04	0.38	-0.24	8.09E-01
Gln	-0.03	0.38	-0.17	8.66E-01
Met-Sulfoxide	-0.03	0.35	-0.16	8.76E-01
p-Amino-Phe	-0.02	0.34	-0.09	9.31E-01
p-Br-Phe	-0.02	0.44	-0.10	9.22E-01
Phe	-0.01	0.44	-0.08	9.35E-01
β-Aminobutyric acid	-0.01	0.48	-0.04	9.71E-01
Ser	-0.01	0.33	-0.03	9.75E-01
Tyr	0.01	0.33	0.05	9.58E-01
p-Cl-Phe	0.01	0.36	0.06	9.51E-01
Cys	0.02	0.44	0.09	9.29E-01
p-Nitro-Phe	0.03	0.42	0.18	8.60E-01
Homoserine	0.09	0.42	0.39	6.96E-01
Met	0.09	0.44	0.52	6.05E-01
Citrulline	0.09	0.36	0.52	6.05E-01
Lys	0.11	0.36	0.60	5.53E-01
p-CF$_3$-Phe	0.12	0.38	0.52	6.08E-01

Substituent	Increment	95% CI	t-value	p-value
o-F-Phe	0.17	0.36	0.79	4.34E-01
Leu	0.18	0.36	1.05	2.97E-01
α-Aminobutyric acid	0.23	0.36	1.05	2.97E-01
Nleu (Nle)	0.25	0.43	1.36	1.78E-01
Nval (Nva)	0.25	0.38	1.40	1.65E-01
Val	0.30	0.48	1.76	8.33E-02
tBu-Gly (Tle)	0.31	0.38	1.76	8.34E-02
Ile	0.44	0.43	2.42	1.82E-02
allo-Ile	0.55	0.38	2.85	5.64E-03

Table S4: Calculated (predicted) retention factors (ln k_1, ln k_2) and separation factors (ln α) as well as corresponding residuals (calculations based on increments of refined data set; Table 1 of main document)

Compound	ln k_1 ± SE	Residual ln k1	ln k_2 ± SE	Residual ln k2	ln α ± SE	Residual ln α
BOC-Ala	1.3 ± 0.3	0.09	1.45 ± 0.42	0.14	0.16 ± 0.34	0.04
BOC-p-CF$_3$-Phe	1.58 ± 0.35	0	1.89 ± 0.49	0	0.32 ± 0.39	0
BOC-p-Cl-Phe	1.93 ± 0.31	-0.18	2.14 ± 0.43	-0.23	0.21 ± 0.35	-0.05
BOC-p-iPr-Phe	1.86 ± 0.35	0	1.93 ± 0.49	0	0.07 ± 0.39	0
BOC-p-Me-Phe	1.74 ± 0.35	0	1.81 ± 0.49	0	0.07 ± 0.39	0
BOC-p-Nitro-Phe	1.95 ± 0.31	-0.03	2.18 ± 0.43	-0.14	0.23 ± 0.35	-0.11
BOC-p-tBu-Phe	1.9 ± 0.35	0	1.98 ± 0.49	0	0.09 ± 0.39	0
BOC-Leu	1.28 ± 0.3	-0.03	1.66 ± 0.42	0.16	0.38 ± 0.34	0.19
BOC-Phe	1.81 ± 0.3	-0.09	2 ± 0.42	-0.1	0.19 ± 0.34	-0.01
BOC-Pro	1.23 ± 0.3	0.32	1.02 ± 0.42	0.11	-0.21 ± 0.34	-0.21
BOC-Tyr	1.81 ± 0.3	-0.03	2.02 ± 0.42	-0.04	0.21 ± 0.34	-0.01
BOC-Val	1.2 ± 0.3	-0.07	1.7 ± 0.42	0.1	0.5 ± 0.34	0.17
DNZ-Ala	2.02 ± 0.29	0.09	2.65 ± 0.41	0.05	0.63 ± 0.33	-0.04
DNZ-allo-Ile	2.06 ± 0.31	0.01	3.27 ± 0.44	-0.09	1.21 ± 0.35	-0.1

Compound	ln k_1 ± SE	Residual ln k1	ln k_2 ± SE	Residual ln k2	ln α ± SE	Residual ln α
DNZ-α-Aminobutyric acid	1.94 ± 0.33	0	2.84 ± 0.47	0	0.9 ± 0.38	0
DNZ-α-Me-Leu	2.12 ± 0.31	0	2.37 ± 0.44	0.14	0.25 ± 0.35	0.14
DNZ-Asn	1.88 ± 0.31	0	2.4 ± 0.44	0.01	0.53 ± 0.35	0.01
DNZ-Asp	4.25 ± 0.31	-0.18	4.45 ± 0.44	-0.01	0.19 ± 0.35	0.17
DNZ-Azetidine carboxylic acid	1.95 ± 0.3	-0.18	2.12 ± 0.42	-0.19	0.17 ± 0.34	-0.01
DNZ-β-Ala	2.08 ± 0.31	-0.14				
DNZ-β-Aminobutyric acid	1.5 ± 0.3	-0.15	2.16 ± 0.42	-0.27	0.66 ± 0.34	-0.12
DNZ-β-Aminoisobutyric acid	1.81 ± 0.31	-0.04	2.43 ± 0.44	-0.09	0.62 ± 0.35	-0.06
DNZ-β-Phe	2.24 ± 0.3	0.08	2.75 ± 0.42	-0.02	0.51 ± 0.34	-0.09
DNZ-Citrulline	1.72 ± 0.3	0.06	2.48 ± 0.42	0.01	0.76 ± 0.34	-0.05
DNZ-dehydro-Pro	2.02 ± 0.14	0	2.19 ± 0.2	0	0.17 ± 0.16	0
DNZ-Gln	1.73 ± 0.31	0.03	2.37 ± 0.44	-0.01	0.63 ± 0.35	-0.03
DNZ-Glu	3.84 ± 0.31	0.07	4.44 ± 0.44	0.07	0.6 ± 0.35	0
DNZ-Gly	2.02 ± 0.3	-0.22				
DNZ-Homoserine	1.72 ± 0.33	0	2.48 ± 0.47	0	0.75 ± 0.38	0
DNZ-Ile	2.04 ± 0.3	-0.02	3.14 ± 0.42	-0.16	1.11 ± 0.34	-0.15
DNZ-p-Amino-Phe	3.37 ± 0.33	0	4.02 ± 0.47	0	0.65 ± 0.38	0
DNZ-p-Br-Phe	2.7 ± 0.31	0.01	3.35 ± 0.44	0.01	0.65 ± 0.35	0
DNZ-p-Cl-Phe	2.66 ± 0.3	0.1	3.34 ± 0.42	0.13	0.68 ± 0.34	0.03
DNZ-p-Nitro-Phe	2.67 ± 0.3	0.03	3.37 ± 0.42	0.1	0.7 ± 0.34	0.07
DNZ-Leu	2.01 ± 0.29	-0.02	2.86 ± 0.41	-0.21	0.84 ± 0.33	-0.18
DNZ-Lys	2.81 ± 0.3	-0.01	3.59 ± 0.42	-0.02	0.78 ± 0.34	0
DNZ-Met	2.25 ± 0.29	0.04	3.01 ± 0.41	-0.05	0.76 ± 0.33	-0.09
DNZ-Met-Sulfone	1.96 ± 0.31	0.08	2.59 ± 0.44	0.05	0.63 ± 0.35	-0.03
DNZ-Met-Sulfoxide	1.81 ± 0.31	0.09	2.45 ± 0.44	0.08	0.64 ± 0.35	-0.01
DNZ-N-Me-Leu	2.07 ± 0.31	0.03	2.27 ± 0.44	0.11	0.2 ± 0.35	0.08
DNZ-N-Me-Val	1.89 ± 0.33	0	2.04 ± 0.47	0	0.16 ± 0.38	0
DNZ-NLeu (DNZ-Nle)	2.08 ± 0.3	0	2.99 ± 0.42	-0.06	0.91 ± 0.34	-0.06
DNZ-NVal (DNZ-Nva)	2.01 ± 0.3	0.01	2.93 ± 0.42	-0.06	0.92 ± 0.34	-0.07

Compound	ln k_1 ± SE	Residual ln k1	ln k_2 ± SE	Residual ln k2	ln α ± SE	Residual ln α
DNZ-Orn	2.87 ± 0.3	-0.13	3.48 ± 0.42	-0.13	0.61 ± 0.34	0
DNZ-Phe	2.54 ± 0.29	0.13	3.19 ± 0.41	0.12	0.65 ± 0.33	-0.01
DNZ-PheGly (DNZ-Phg)	2.56 ± 0.3	0.06	3.12 ± 0.42	0.01	0.56 ± 0.34	-0.05
DNZ-Pipecolinic acid	2.16 ± 0.31	-0.02	2.4 ± 0.44	0.08	0.24 ± 0.35	0.1
DNZ-Pro	1.96 ± 0.29	-0.06	2.22 ± 0.41	0.01	0.26 ± 0.33	0.07
DNZ-Ser	1.97 ± 0.29	0.1	2.64 ± 0.41	0.07	0.66 ± 0.33	-0.03
DNZ-tBu-Gly (DNZ-Tle)	1.97 ± 0.3	-0.07	2.95 ± 0.42	-0.21	0.98 ± 0.34	-0.14
DNZ-Thr	1.86 ± 0.3	-0.02	2.47 ± 0.42	0.34	0.62 ± 0.34	0.36
DNZ-Trp	2.99 ± 0.3	0.06	3.62 ± 0.42	0.06	0.62 ± 0.34	0
DNZ-Tyr	2.54 ± 0.29	0.2	3.21 ± 0.41	0.17	0.68 ± 0.33	-0.03
DNZ-Val	1.93 ± 0.29	-0.04	2.9 ± 0.41	-0.25	0.97 ± 0.33	-0.22
FMOC-Cys	2.74 ± 0.39	0	3.23 ± 0.54	0	0.48 ± 0.44	0
FMOC-Cys-S-Trityl	4.02 ± 0.39	0	4.38 ± 0.54	0	0.35 ± 0.44	0
FMOC-Leu	2.22 ± 0.32	0	2.86 ± 0.45	0	0.64 ± 0.36	0
FMOC-Phe	2.75 ± 0.32	0	3.2 ± 0.45	0	0.45 ± 0.36	0
NVOC-Ala	2 ± 0.29	0.05	2.84 ± 0.41	-0.06	0.83 ± 0.33	-0.11
NVOC-Azetidine carboxylic acid	1.93 ± 0.3	0.13	2.31 ± 0.42	0.4	0.38 ± 0.34	0.27
NVOC-β-Phe	2.22 ± 0.3	0.01	2.94 ± 0.42	0.15	0.72 ± 0.34	0.14
NVOC-Citrulline	1.7 ± 0.3	-0.05	2.67 ± 0.42	-0.08	0.97 ± 0.34	-0.03
NVOC-Ile	2.01 ± 0.3	0.05	3.33 ± 0.42	0.12	1.31 ± 0.34	0.07
NVOC-Leu	1.99 ± 0.29	0.02	3.04 ± 0.41	-0.2	1.05 ± 0.33	-0.23
NVOC-Lys	2.79 ± 0.3	-0.02	3.78 ± 0.42	-0.1	0.98 ± 0.34	-0.08
NVOC-Met	2.23 ± 0.29	-0.02	3.2 ± 0.41	-0.07	0.97 ± 0.33	-0.06
NVOC-Met-Sulfone	1.94 ± 0.31	-0.08	2.78 ± 0.44	-0.05	0.83 ± 0.35	0.03
NVOC-Met-Sulfoxide	1.79 ± 0.31	-0.09	2.63 ± 0.44	-0.08	0.85 ± 0.35	0.01
NVOC-NLeu (NVOC-Nle)	2.06 ± 0.3	0.01	3.18 ± 0.42	-0.06	1.12 ± 0.34	-0.07
NVOC-NVal (NVOC-Nva)	1.98 ± 0.3	0	3.11 ± 0.42	-0.08	1.13 ± 0.34	-0.08
NVOC-Orn	2.85 ± 0.3	-0.08	3.66 ± 0.42	-0.11	0.82 ± 0.34	-0.02
NVOC-Phe	2.52 ± 0.29	0.04	3.38 ± 0.41	0.02	0.86 ± 0.33	-0.02

Compound	ln k_1 ± SE	Residual ln k1	ln k_2 ± SE	Residual ln k2	ln α ± SE	Residual ln α
NVOC-PheGly (NVOC-Phg)	2.54 ± 0.3	-0.12	3.31 ± 0.42	-0.04	0.77 ± 0.34	0.08
NVOC-Pro	1.94 ± 0.29	0.01	2.4 ± 0.41	0.4	0.46 ± 0.33	0.39
NVOC-Ser	1.95 ± 0.29	0.03	2.82 ± 0.41	-0.04	0.87 ± 0.33	-0.06
NVOC-Thr	1.83 ± 0.3	-0.01	2.66 ± 0.42	-0.25	0.82 ± 0.34	-0.24
NVOC-Trp	2.97 ± 0.3	0.03	3.8 ± 0.42	0.08	0.83 ± 0.34	0.05
NVOC-Tyr	2.52 ± 0.29	0.12	3.4 ± 0.41	0.11	0.88 ± 0.33	-0.01
NVOC-Val	1.91 ± 0.29	-0.01	3.08 ± 0.41	-0.05	1.17 ± 0.33	-0.03
PNZ-Ala	1.65 ± 0.29	-0.02	2 ± 0.41	0.03	0.34 ± 0.33	0.05
PNZ-allo-Ile	1.69 ± 0.31	-0.01	2.62 ± 0.44	0.09	0.93 ± 0.35	0.1
PNZ-α-Aminoadipic acid	3.34 ± 0.31	0	3.67 ± 0.44	0.04	0.33 ± 0.35	0.04
PNZ-α-Me-Leu	1.75 ± 0.31	0	1.71 ± 0.44	-0.14	-0.04 ± 0.35	-0.14
PNZ-Asn	1.51 ± 0.31	0	1.75 ± 0.44	-0.01	0.24 ± 0.35	-0.01
PNZ-Asp	3.88 ± 0.31	0.18	3.79 ± 0.44	0.01	-0.09 ± 0.35	-0.17
PNZ-α-Aminobutyric acid	1.71 ± 0.33	0.14	1.8 ± 0.47	-0.23	0.08 ± 0.38	-0.37
PNZ-β-Aminobutyric acid	1.13 ± 0.3	-0.17	1.51 ± 0.42	-0.12	0.38 ± 0.34	0.05
PNZ-β-Aminoisobutyric acid	1.44 ± 0.31	0.04	1.77 ± 0.44	0.09	0.34 ± 0.35	0.06
PNZ-Citrulline	1.35 ± 0.3	-0.01	1.83 ± 0.42	0.07	0.48 ± 0.34	0.08
PNZ-Cysteic acid	3.54 ± 0.33	0	3.67 ± 0.47	0	0.13 ± 0.38	0
PNZ-Gln	1.36 ± 0.31	-0.03	1.71 ± 0.44	0.01	0.35 ± 0.35	0.03
PNZ-Glu	3.47 ± 0.31	-0.07	3.79 ± 0.44	-0.07	0.31 ± 0.35	0
PNZ-Ile	1.67 ± 0.3	-0.03	2.49 ± 0.42	0.04	0.82 ± 0.34	0.07
PNZ-p-Br-Phe	2.33 ± 0.31	-0.01	2.7 ± 0.44	-0.01	0.36 ± 0.35	0
PNZ-p-Cl-Phe	2.29 ± 0.3	0.07	2.69 ± 0.42	0.1	0.39 ± 0.34	0.03
PNZ-p-Nitro-Phe	2.3 ± 0.3	-0.01	2.72 ± 0.42	0.04	0.42 ± 0.34	0.05
PNZ-Leu	1.64 ± 0.29	-0.17	2.2 ± 0.41	-0.12	0.56 ± 0.33	0.05
PNZ-Lys	2.44 ± 0.3	0.03	2.93 ± 0.42	0.12	0.49 ± 0.34	0.09
PNZ-Met	1.88 ± 0.29	0.02	2.36 ± 0.41	0.08	0.47 ± 0.33	0.06
PNZ-N-Me-Leu	1.7 ± 0.31	-0.03	1.62 ± 0.44	-0.11	-0.08 ± 0.35	-0.08
PNZ-NLeu (PNZ-Nle)	1.71 ± 0.3	-0.01	2.34 ± 0.42	0.12	0.63 ± 0.34	0.13

Compound	ln k_1 ± SE	Residual ln k1	ln k_2 ± SE	Residual ln k2	ln α ± SE	Residual ln α
PNZ-NVal (PNZ-Nva)	1.64 ± 0.3	-0.01	2.27 ± 0.42	0.14	0.64 ± 0.34	0.15
PNZ-Orn	2.5 ± 0.3	0.21	2.82 ± 0.42	0.23	0.32 ± 0.34	0.02
PNZ-Phe	2.17 ± 0.29	-0.05	2.54 ± 0.41	-0.04	0.37 ± 0.33	0.02
PNZ-PheGly (PNZ-Phg)	2.19 ± 0.3	0.06	2.47 ± 0.42	0.03	0.28 ± 0.34	-0.03
PNZ-Pipecolinic acid	1.79 ± 0.31	0.02	1.75 ± 0.44	-0.08	-0.05 ± 0.35	-0.1
PNZ-Pro	1.59 ± 0.29	-0.07	1.56 ± 0.41	-0.14	-0.03 ± 0.33	-0.07
PNZ-Ser	1.6 ± 0.29	0.08	1.98 ± 0.41	0.14	0.38 ± 0.33	0.06
PNZ-tBu-Gly (PNZ-Tle)	1.6 ± 0.3	-0.08	2.29 ± 0.42	-0.06	0.7 ± 0.34	0.02
PNZ-Thr	1.49 ± 0.3	0.03	1.82 ± 0.42	-0.09	0.33 ± 0.34	-0.12
PNZ-Trp	2.62 ± 0.3	-0.09	2.96 ± 0.42	-0.14	0.34 ± 0.34	-0.05
PNZ-Tyr	2.17 ± 0.29	0.04	2.56 ± 0.41	0.05	0.39 ± 0.33	0.01
PNZ-Val	1.56 ± 0.29	-0.07	2.24 ± 0.41	-0.1	0.68 ± 0.33	-0.03
Z-Ala	1.65 ± 0.28	-0.22	1.84 ± 0.4	-0.16	0.2 ± 0.32	0.05
Z-Azetidine carboxylic acid	1.57 ± 0.29	0.06	1.31 ± 0.41	-0.21	-0.26 ± 0.33	-0.26
Z-β-Aminobutyric acid	1.12 ± 0.29	0.32	1.35 ± 0.41	0.38	0.23 ± 0.33	0.06
Z-β-Phe	1.87 ± 0.29	-0.09	1.95 ± 0.41	-0.13	0.08 ± 0.33	-0.04
Z-Gly	1.65 ± 0.29	0.22				
Z-o-F-Phe	2.21 ± 0.32	0	2.62 ± 0.45	0	0.4 ± 0.36	0
Z-m-F-Phe	2.24 ± 0.32	0	2.41 ± 0.45	0	0.17 ± 0.36	0
Z-p-F-Phe	2.14 ± 0.13	0	2.38 ± 0.19	0	0.24 ± 0.15	0
Z-Leu	1.63 ± 0.28	0.2	2.05 ± 0.4	0.37	0.41 ± 0.32	0.17
Z-Met	1.87 ± 0.28	-0.04	2.2 ± 0.4	0.04	0.33 ± 0.32	0.08
Z-Phe	2.16 ± 0.28	-0.03	2.38 ± 0.4	-0.01	0.22 ± 0.32	0.02
Z-Pro	1.58 ± 0.28	-0.21	1.41 ± 0.4	-0.39	-0.18 ± 0.32	-0.18
Z-Ser	1.6 ± 0.28	-0.21	1.83 ± 0.4	-0.17	0.23 ± 0.32	0.04
Z-tBu-Gly (Z-Tle)	1.59 ± 0.29	0.15	2.14 ± 0.41	0.27	0.55 ± 0.33	0.12
Z-Tyr	2.16 ± 0.28	-0.32	2.41 ± 0.4	-0.29	0.25 ± 0.32	0.03
Z-Val	1.55 ± 0.28	0.19	2.09 ± 0.4	0.29	0.54 ± 0.32	0.11
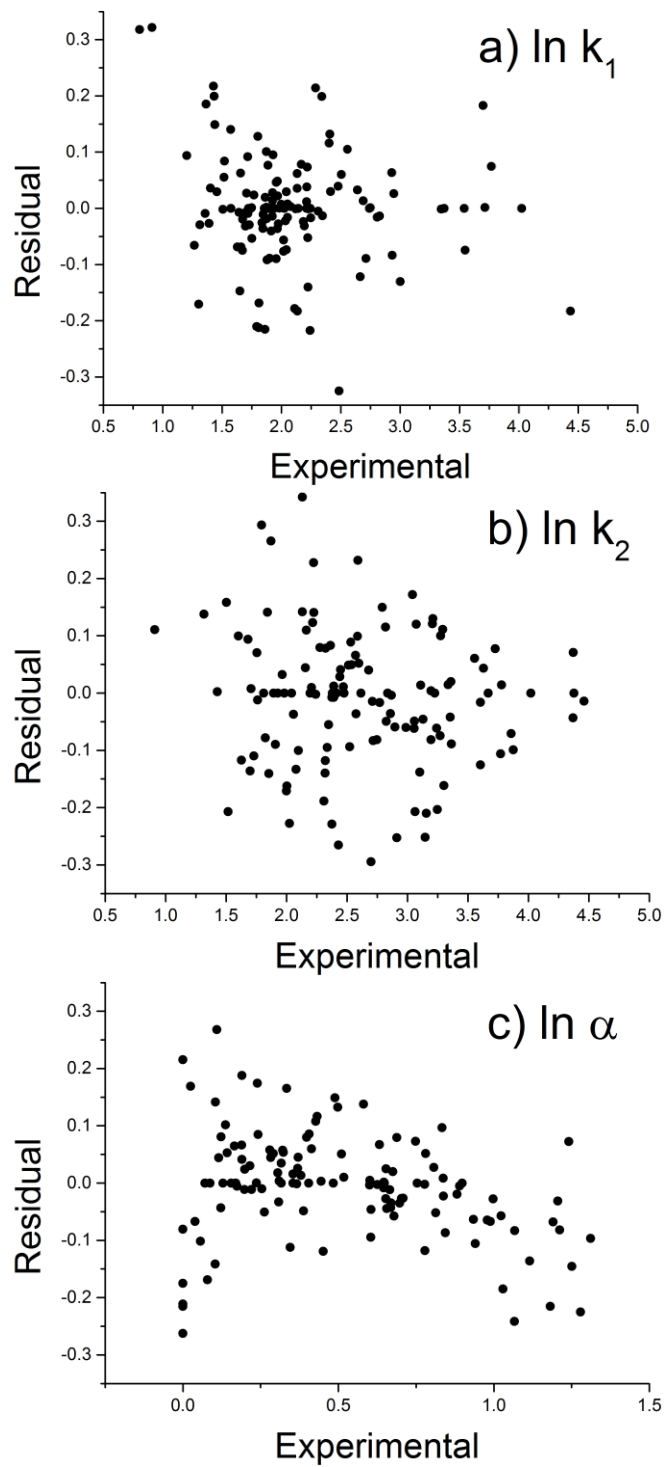
Figure S2: Residual plots indicate good explanation of the data by the linear model and random distribution of the residuals

Appendix II

Accepted Manuscripts

Publication II and supplemental material	
Title	Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides
Authors	*Gerhardt, Heike; **Sievers-Engler, Adrian**; Jahanshah, Ghazaleh; Pataj, Zoltan; Ianni, Federica; Gross, Harald; Lindner, Wolfgang; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1428) 2016, p. 280-291
DOI	http://dx.doi.org/10.1016/j.chroma.2015.05.065

Publication VIII	
Title	Peptide Analysis: Zwitterionic Chiral Ion-Exchangers as Complementary Option to HILIC and to Reversed-Phase Chromatography
Authors	*Zhang, Tong; Holder, Emilie; Franco, Pilar; Laemmerhofer, Michael; **Sievers-Engler, Adrian**; Gerhardt, Heike; Gross, Harald; Lindner, Wolfgang*
Published in	LC GC chromatographyonline.com (29), 2016, issue 3, p. 112-128
Link	http://www.chromatographyonline.com/peptide-analysis-zwitterionic-chiral-ion-exchangers-complementary-option-hilic-and-reversed-phase-ch

Reprint with permission from Elsevier and Advanstra Communications

Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides

Heike Gerhardt [a,1], Adrian Sievers-Engler [a,1], Ghazaleh Jahanshah [a], Zoltán Pataj [a,b], Federica Ianni [a,c], Harald Gross [a], Wolfgang Lindner [d], Michael Lämmerhofer [a,*]

[a] *Institute of Pharmaceutical Sciences, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany*
[b] *Department of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Germany*
[c] *Department of Pharmaceutical Sciences, University of Perugia, Via del Liceo 1, 06123 Perugia, Italy*
[d] *Department of Analytical Chemistry, University of Vienna, Waehringer Strasse 38, 1090 Vienna, Austria*

ARTICLE INFO

ABSTRACT

A panel of methods of general suitability for complete structural elucidation of the stereochemistry of cyclopeptides, depsipeptides and lipopeptides is presented and described in detail. The suitability of the proposed methods was exemplified on the lipopeptide poaeamide from *Pseudomonas poae*. Amino acid configurations have been assigned by direct LC enantiomer separation with Chiralpak ZWIX(+) and were confirmed by GC enantiomer separation on Chirasil L-Val. 3-Hydroxydecanoic acid absolute configuration was analyzed on Chiralpak ZWIX(+) and confirmed by injection on ZWIX(−) which showed opposite elution order. Plenty of D-amino acids have been found in this lipopeptide. It contained in total 5 Leu residues of which one had D-configuration. The position of the D-Leu in the peptide sequence was determined by pepsin and chemical digestions in combination with isolation of diagnostic peptide-fragments and subsequent identification of absolute configurations of the Leu residues. This allowed pinpointing the position of the D-amino acid. The complementarity of the peptide retention profiles on Chiralpak ZWIX column as compared to both RPLC and HILIC suggests its great utility as an alternative peptide separation tool.

1. Introduction

Lipopeptides are amphiphilic natural compounds constituted by an oligopeptide backbone and a fatty acid tail [1–3]. They often exist in macrocyclic form due to lactone-type ring closure between hydroxyl side chains and carboxylic functionalities, thus also termed depsipeptides. Lipopeptides are biosurfactants [4] and there is great interest in these biomolecules due to their potential bioactivities in particular antimicrobial and antitumor activities [1,5,6], usually by exerting their activity via specific interactions with the plasma membrane [7]. Interestingly, D-amino acids are commonly found in these biomolecules which make them more resistant against endopeptidase enzymes and proteolytic degradation, respectively [8].

Complete structure elucidation consequently requires the determination of the absolute configurations of the chiral building blocks. For this purpose, the peptide is hydrolyzed and the configurations of the residues are determined by enantioselective analysis [9]. For amino acid enantiomer separation and absolute configuration assignment, gas chromatography with pre-column derivatization has become a state-of-art methodology [10,11]. On the other hand, in liquid chromatography indirect methods [12,13] using either ortho-phthaldialdehyde (OPA)/chiral thiol as chiral derivatizing agent [14,15] or Marfey's reagent (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide) [16–19] and subsequent achiral RPLC are commonly used. There are only a few methods which allow direct enantiomer separation of amino acids on chiral stationary phases in underivatized form, namely chiral ligand exchange chromatography [20], LC with macrocyclic antibiotics such as teicoplanin or teicoplanin aglycon based CSPs [21–24], LC with chiral crown ether CSPs [25–27], and LC on zwitterionic chiral ion exchangers [28–33] (for recent reviews on this topic cf. ref. [34]). Two-dimensional HPLC assays are sometimes implemented

* Corresponding author. Tel.: +49 7071 29 78793; fax: +49 7071 29 4565.
 E-mail address: michael.laemmerhofer@uni-tuebingen.de (M. Lämmerhofer).
[1] These authors contributed equally.

for the purpose of comprehensive enantioselective amino acid enantiomer analysis, also to overcome limited chemoselectivity of chiral stationary phases [35,36]. Since CLEC is incompatible with MS detection and chiral crown ethers typically require strongly acidic conditions (often perchloric acid in the mobile phase), the two MS compatible CSP classes, namely macrocyclic glycopeptide CSPs and zwitterionic cinchona alkaloid derived CSPs, were considered as first choice. Since the zwitterionic ion-exchangers based on cinchona alkaloids appear in two versions, Chiralpak ZWIX(+) and ZWIX(−) (see Fig. 1), with opposite elution orders [28,31], allowing validation of results by reversal of elution order, they were selected for our study and hyphenated with QTOF MS. Only few reports exist in the literature in which LC–MS in combination with chiral stationary phases was used for direct enantiomer analysis without pre-column derivatization [22] [37]. However, matrix effects, insufficient sensitivities, peak distortions and inadequate chemical selectivity for isobaric amino acids remained some of the unsolved problems complicating this endeavor [20–22,25,28–31,35–37].

Elucidation of the stereochemistry in a lipopeptide is straightforward if individual building blocks are present only once in the biomolecule or only in one configuration. The situation gets more complicated if chiral synthons of one sort are present more than once and both configurations are detected by enantioselective analysis of those constituents. In such cases, not only the sequence of the building blocks but also the position of the particular stereochemistry of the synthons has to be pinpointed for complete structural characterization. Analytically, this is more challenging because upon hydrolysis of the lipopeptide for enantioselective analysis of the building blocks the sequence information is lost. Different strategies have then to be adopted to solve the problem.

In this article, we suggest analytical strategies, methods and complementary stationary phases for the comprehensive structural elucidation of the constitution and stereochemistry of lipopeptides, depsipeptides and therapeutic peptides. Particular focus is given on the elucidation of the potential of a new chiral stationary phase, Chiralpak ZWIX (see Fig. 1), which is a chiral zwitterionic ion-exchanger, for its capability to resolve the enantiomers of the building blocks (amino acids, hydroxy acid) and serve as complementary stationary phase to RPLC and HILIC in separations of linearized and cyclic lipopeptide as well as of peptide fragments generated by enzymatic and chemical digestion. New selectivity profiles in such peptide separations and the stereoselectivity of this separation material for peptide enantiomers, epimers and diastereomers [30,38,39] might facilitate and support the structural elucidation.

It is documented for the comprehensive structural elucidation of a lipopeptide isolated from the endophytic *Pseudomonas poae* strain RE*1-1-14 (Fig. 2). Briefly, the target lipopeptide (LP) with a 10-amino acid peptide moiety was originally isolated from internal root tissue of sugar beet plants and shown to suppress growth of the fungal pathogen *Rhizoctonia solani*. The amino acid sequence and fatty acid side chains by which this LP was constituted could be readily confirmed by MS-based sequencing and NMR [40]. Uncertainty existed for the stereochemistry which has been clarified in this study by the use of Chiralpak ZWIX and a reference GC–MS method using Chirasil L-Val. Furthermore, a combination of enzymatic, chemical and chromatographic methods were utilized for unequivocally pinpointing which one of 5 Leu residues had D-configuration.

2. Materials and methods

2.1. Materials

Amino acid standards, 1-fluoro-2,4-dinitrobenzene (Sanger's reagent) as well as reagents and solvents for GC–MS analysis comprising trifluoroacetic anhydride (TFAA), ethyl trifluoroacetate (TFAEE), acetyl chloride, deuterium oxide, dichloromethane, deuterated ethanol (EtOD) were all purchased from Sigma-Aldrich (Munich, Germany).

Organic solvents, such as methanol and acetonitrile for LC–MS were of Ultra-MS grade and supplied by Carl Roth (Karlsruhe, Germany). Solvent additives (formic acid, ammonium formate) were purchased in trace metal analysis or UPLC-MS grade from Sigma-Aldrich. Ultra-MS-grade water was obtained from Water Purelab Analytics Purification System (from ELGA, Celle, Germany). The lipopeptide poae RE*1-1-14 was prepared and isolated as described previously [40].

2.2. Instrumentation

2.2.1. GC–MS instrumentation

GC–MS analysis was performed as described previously [11] using an Agilent 7890 A GC-system coupled to a single quadrupole mass spectrometer 5975C inert MSD with Triple-Axis (Agilent Technologies, Waldbronn, Germany). A Chirasil L-Val column (20 m × 0.3 mm i.d.; 0.28 μm film thickness; C.A.T., Tübingen, Germany) was used for separation of the amino acid enantiomers after double derivatization as described below. The temperature of the split/splitless injector was set to 220 °C and the injector was operated in the splitless mode. Hydrogen was used as a carrier gas at constant pressure of 0.45 bar. The GC oven temperature was programmed as follows: 50 °C for 0.5 min, 50 °C/min to 77 °C, then 4 °C/min to 110 °C, then 10 °C/min to 140 °C, then 6 °C/min to 195 °C followed by a final hold for 7 min.

EI spectra were recorded in SIM mode determining the ions with a mass to charge ratio of 126, 138, 140, 152, 166, 168, 176, 180, 199, 182, 214, 228, 290, 315 and 421. All measurements were performed with an MS source temperature of 230 °C and a quadrupole temperature of 150 °C at 70 eV.
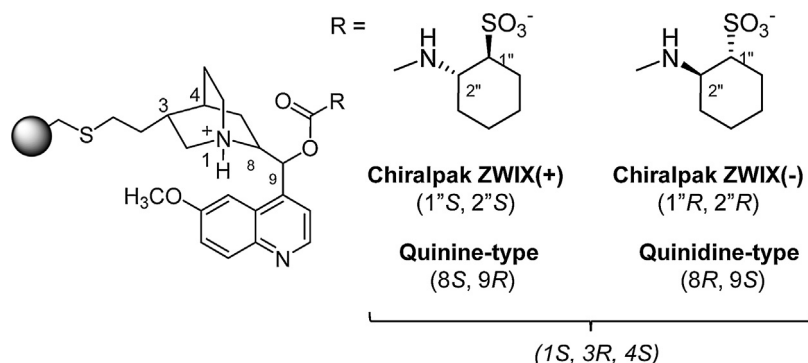
Fig. 1. Structure of utilized chiral stationary phases: (a) zwitterionic chiral ion-exchanger Chiralpak ZWIX(+) and ZWIX(−).
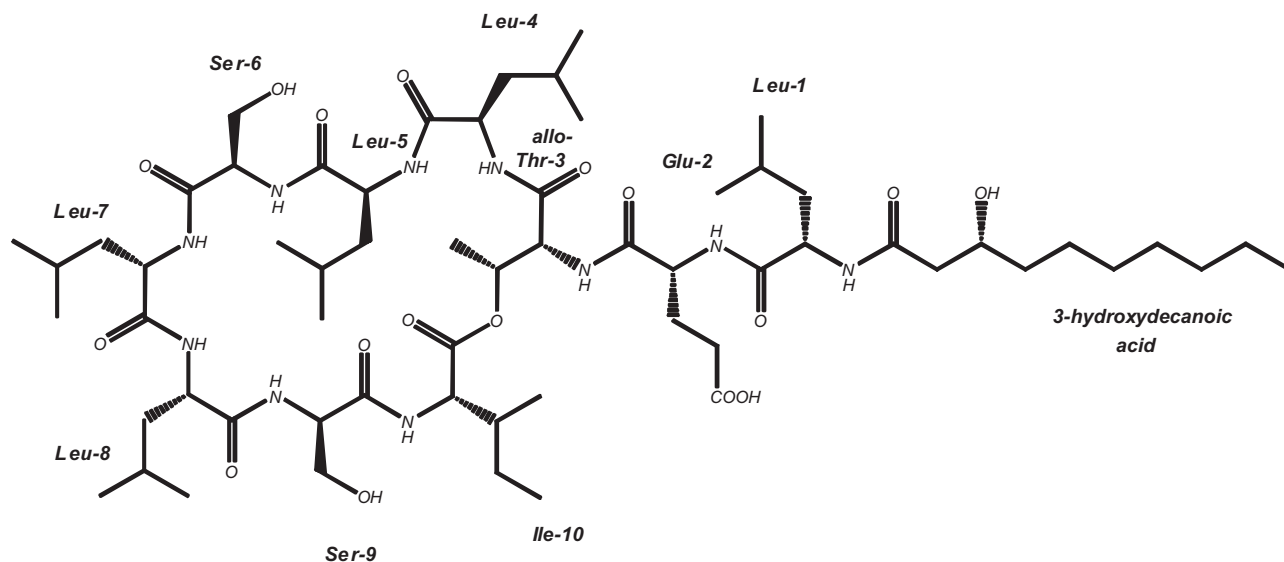
Fig. 2. Structure of lipopeptide poaeamide. Numbering system refers to amino acid residue counted from the *N*-terminus.

2.2.2. HPLC-MS instrumentation

For absolute configuration determination of the 3-hydroxy-decanoic acid an Agilent 1100 series LC system (Agilent Technologies, Waldbronn, Germany) equipped with a binary gradient pump, autosampler, vacuum degasser, temperature-controlled column compartment, UV-detector coupled to a mass-spectrometer Agilent 1100 series LC MSD-SL ion-trap was used. The data were processed with Bruker Daltonics Esquire data analysis software version 2.2.

Other LC–MS experiments were conducted on a TripleTOF 5600+ (Sciex, Ontario, Canada) QTOF MS instrument coupled via Duospray Ion Source operated in ESI mode to an Agilent 1290 series UHPLC pump and column thermostat equipped with a CTC-PAL HTS autosampler (CTC Analytics, Zwingen, Switzerland).

2.3. Methods

2.3.1. Enantioselective analysis of amino acids by GC–MS with Chirasil L-Val

The general procedure for the assignment of absolute configurations of amino acids in peptides by GC–MS is described elsewhere [11]. Briefly, the peptide sample was first hydrolyzed by 6N deuterated hydrochloric acid (DCl/D$_2$O) for 24 h at 110 °C. The hydrochloric acid was evaporated in a Thermo Savant ISS110 SpeedVac (Thermo Scientific, Holbrook, USA) at 43 °C for 1–1.5 h. The residue was dissolved in 350 µL of 3 N EtOD/DCl (prepared from deuterated ethanol and acetyl chloride) and heated for 20 min at 110 °C. The reagent was removed by a gentle stream of nitrogen at 110 °C. For the derivatization of the amino groups 350 µL of a mixture of TFAA/TFAEE (1:2; v/v) were added and the reaction allowed to proceed for 10 min at 130 °C. Afterwards, the reagent was removed at room temperature by a gentle stream of nitrogen. For injection, the residue was dissolved in 100 µL dichloromethane.

2.3.2. Enantioselective analysis of amino acids by LC–MS with Chiralpak ZWIX(+)

The assignment of absolute configurations of amino acid constituents in peptides by LC–MS was performed on a Chiralpak ZWIX(+) (3 µm; 150 mm × 4 mm i.d.) column (Chiral Technologies Europe, Illkirch, France). A mixture of MeOH/H$_2$O (98:2; v/v) containing 9.4 mM ammonium formate and 9.4 mM formic acid was used as mobile phase and the flow rate was 0.7 mL/min.

MS parameters were as follows: the Duospray ion source was operated in ESI(−) mode and the source voltage was −4500 V. The source temperature was 400 °C, declustering potential −100 V, curtain gas at 30 psi, nebulizer and drying gas at 60 psi. For targeted acquisition and enhanced sensitivity, product ion high sensitivity mode was used employing −20 V as collision energy. Analyses of amino acids were performed after hydrolysis of lipopeptide as described above in underivatized and derivatized form. Thus, amino acid standards, single enantiomers and racemates, were injected. For derivatization, Sanger's reagent (1-fluoro-2,4-dinitrobenzene in acetonitrile; 2.5%, v/v) was used. Thus, 200 µL of a solution of amino acids in 0.1 M sodium bicarbonate buffer (pH 9.4) (0.2 mg/mL) were mixed with 100 µL Sanger's reagent and incubated at RT for 2 h. Then, the reaction was quenched by addition of 10 µL concentrated ammonia, pH adjusted for extraction by addition of 100 µL concentrated hydrochloric acid. DNP-derivatives of amino acids were then extracted with 100 µL chloroform. The organic extract was subsequently diluted with MeOH (1:100) for LC–MS analysis.

2.3.3. Enantioselective analysis of 3-hydroxy carboxylic acid by LC–MS with Chiralpak ZWIX(+) and ZWIX(−)

The lipopeptide was hydrolyzed as described for GC–MS analysis. For decreasing ion suppression, the 3-hydroxy fatty acid was extracted using 200 µL of a mixture containing water and chloroform in a ratio of 1:1 (v/v). The chloroform layer was used for the determination of the configuration of the 3-hydroxy decanoic acid. Thus, the chloroform was removed in the SpeedVac and the residue was reconstituted in methanol.

In order to establish the enantiomeric elution order on Chiralpak ZWIX(+) and ZWIX(−), a standard mixture of enantiomerically enriched (R)-3-hydroxydecanoic acid was prepared from a commercial rhamnolipid (R-95 Rhamnolipid, Sigma-Aldrich) by acid hydrolysis according to ref. [41], due to the commercial unavailability of the enantiopure compound of 3-hydroxydecanoic acid. Thus, 10 mg of rhamnolipid were suspended in 0.5 mL 15% (v/v) H$_2$SO$_4$ in a screw-capped glass tube. Then a volume of 0.5 mL of chloroform was added and the obtained biphasic system was heated at 110 °C for 140 min. The chloroform layer containing the fatty acid was collected and evaporated to dryness and, successively, solubilized in MeOH for subsequent chromatographic runs.

Enantioselective analysis was performed using a Chiralpak ZWIX(+) and ZWIX(−) (3 μm; 150 × 4 mm i.d.) column from Chiral Technologies Europe. The column thermostat was set to 10 °C and an isocratic run of 30 min was performed using a mixture of acetonitrile/methanol/acetic acid (95/5/0.025; v/v/v) at a mobile phase flow rate of 0.3 mL/min. The injection volume was 10 μL.

The ESI source of the ion-trap mass spectrometer was operated in the negative ionization mode at 350 °C, 40 psi with a dry gas flow of 7.5 L/min. Analysis was performed in a scan range of $m/z = 180–200$.

2.3.4. Hydrolysis of cyclic ester and analysis of linearized lipopeptide

For enzymatic digestion (see below) the cyclic ester (macrocyclic lactone ring) needed to be hydrolyzed to obtain the linear form of the lipopeptide. Thus, the lipopeptide (2 mg) was hydrolyzed with 1N NaOH (400 μL) at 45 °C for 90 min, the reaction mixture neutralized with 2N HCl and purified by RP-HPLC as described previously [40].

The reaction product was monitored by (U)HPLC-QTOF-MS. For all separations the MS parameters were as follows: curtain gas 30 psi, nebulizer and drying gas 60 psi, ion source floating voltage 5500 V, source temperature 400 °C, and declustering potential 100 V. Acquisition was performed as scheduled targeted product ion scans with collision energy of −25 V, for generating confirmation fragment spectra.

Separation of cyclic and linear form was performed using (i) RPLC, (ii) HILIC and (iii) zwitterionic ion-exchanger Chiralpak ZWIX(+) in polar organic mode under isocratic conditions. Two different RP columns were tested for RPLC: ODS-Hypersil (5 μm; 100 × 2.1 mm i.d.) from Agilent (Palo Alto, USA) and Kinetex C18 (2.6 μm; 100 × 3 mm i.d.) from Phenomenex (Aschaffenburg, Germany). The eluent consisted of water/acetonitrile (40/60; v/v) with 0.1% (v/v) formic acid as modifier. The flow rate was 300 μL/min.

For HILIC separations, a SeQuant ZIC-HILIC column (3.5 μm; 150 × 2 mm i.d.) from MerckMillipore (Darmstadt, Germany) was utilized. The eluent consisted of water/acetonitrile (5/95; v/v) with 0.1% (v/v) formic acid and the flow rate was 300 μL/min.

For polar organic mode separations, the Chiralpak ZWIX(+) column was run with an eluent composed of water/acetonitrile (65/35; v/v) with 0.1% (v/v) formic acid and a flow rate of 700 μL/min. Column temperature was maintained at 25 °C in all experiments.

2.3.5. Location of D-leucine in peptide chain

For the determination of the position of the D-leucine residue in the peptide chain the lipopeptide was digested to smaller peptide fragments which were analyzed by (i) RPLC, (ii) HILIC and (iii) zwitterionic ion-exchange LC. Digests of the lipopeptide were prepared both by enzymatic and chemical digestion.

Enzymatic digestion was performed as follows: 0.5 mg of purified peptide were dissolved in 200 μL of 0.4% (w/v) solution of pepsin in 10 mM HCl and digested for 90 min at 37 °C. Hydrochloric acid was then removed by SpeedVac and the residue reconstituted in 200 μL water/acetonitrile (50/50; v/v).

Chemical digestion was first optimized using 6N HCl at 110 °C monitoring selected peptide intensities every 20 min. Finally, chemical digestion was performed with 6N HCl at 110 °C for 20 min, followed by evaporation of the hydrochloric acid in a SpeedVac and reconstitution of the residue in acetonitrile/water (50/50, v/v).

Both enzymatic and chemical digests were analyzed by RPLC-MS on a Kinetex C18 column with gradient elution using water with 0.1% (v/v) formic acid as eluent A and acetonitrile with 0.1% (v/v) formic acid as eluent B. The following gradient profile was adopted: 20% B from 0 to 2.5 min, 20–64% B from 2.5 to 12.5 min, 64% B from 12.5 to 15 min, 64–80% B from 12.5 to 17 min, 80% B from 17 to 20 min, 20% B from 20.1 to 23 min. Retention times of generated peptides were elucidated mass spectrometrically in information dependent acquisition (IDA) mode yielding retention time, peak width and spectral confirmation. Fractions of target peptides were collected by using a Waters fraction collector II. After hydrolysis of the isolated peptides with DCl/D₂O (as described above), their amino acid enantiomer composition was analyzed by GC–MS as described above.

Chemical digests were additionally analyzed by HILIC with a MerckSeQuant ZIC-HILIC column. Channel A (water with 0.1% (v/v) formic acid) and B (acetonitrile with 0.1% (v/v) formic acid) solvents were the same as specified above for RPLC on Kinetex C18 but the gradient profile was inverted: 80% B from 0 to 2.5 min, 20–64% B from 2.5 to 12.5 min, 64% B from 12.5 to 15 min, 64–80% B from 12.5 to 17 min, 80% B from 17 to 20 min, 20% B from 20.1 to 23 min. Analysis in polar organic mode with Chiralpak ZWIX(+) was conducted in isocratic mode with methanol/acetonitrile/water (49/49/2; v/v/v) containing 25 mM ammonium formate and 25 mM formic acid. MS parameters were as follows: ESI(−) mode, curtain gas 30 psi, nebulizer and drying gas 60 psi, ion source floating voltage −4500 V, source temperature 400 °C, and declustering potential -100 V. Acquisition was performed as product ion-high sensitivity mode with −20 V collision energy, for enhanced sensitivity and generating confirmation fragment spectra.

3. Results and discussion

3.1. Amino acid configurations in peptides

Racemates and single enantiomers of amino acids being constituents of the lipopeptide along with its hydrolysate were injected into a Chiralpak ZWIX(+) column employing conditions as specified in Section 2 [16,17,21,24,28–33]. All amino acids except for Ser could be resolved under adopted conditions without derivatization (Suppl. Material, Figure S1, left panel as well as Table S1). Except for Glu and allo-threonine (a-Thr), the L-enantiomers eluted first on Chiralpak ZWIX(+). The racemates of the amino acids present in the lipopeptide along with allo-forms of Thr and Ile were analyzed as a mixture as well and the results are shown for the individual amino acid traces in Fig. 3. In spite of good enantiomeric separation for Leu, Ile, and allo-isoleucine (a-Ile), it was not possible to distinguish between Ile and a-Ile when they were present in the same mixture due to coelution of these isomeric compounds (Fig. 3c). The problem could be solved by a two-dimensional LC-method, but was beyond the scope of the present study [35,36,42,43]. Instead, achiral derivatization with Sanger's reagent was examined. Sanger's derivatives of all amino acids (N-2,4-dinitrophenyl-amino acids) could be resolved (Suppl. Material, Figure S1, right panel as well as Table S1) with consistent elution order of L before D on Chiralpak ZWIX(+). The current zwitterionic cinchona alkaloid derived chiral ion-exchangers have the advantage over corresponding anion-exchanger based CSPs (Chiralpak QN-AX and QD-AX) [44] that lower buffer concentrations, thus milder elution conditions which is favorable for ESI-MS hyphenation, can be used for analysis of acidic analytes such as N-derivatized amino acids, and that run times are much shorter (here always less than 10 min, see Fig. 3). While Ser was fully baseline resolved and also all 4 isomers of the Thr/a-Thr family could be separated, a-Ile was only partially resolved from Ile and no resolution was observed between the 3 D-leucine isomers (D-Leu, D-Ile, and D-a-Ile) (Fig. 3g). While Chiralpak ZWIX exhibits good enantiomer separation capability for both underivatized and derivatized amino acids, chemoselectivity limitations for the group of leucine isomers requires other strategies such as column coupling.

Free amino acids

DNP-derivatized amino acids
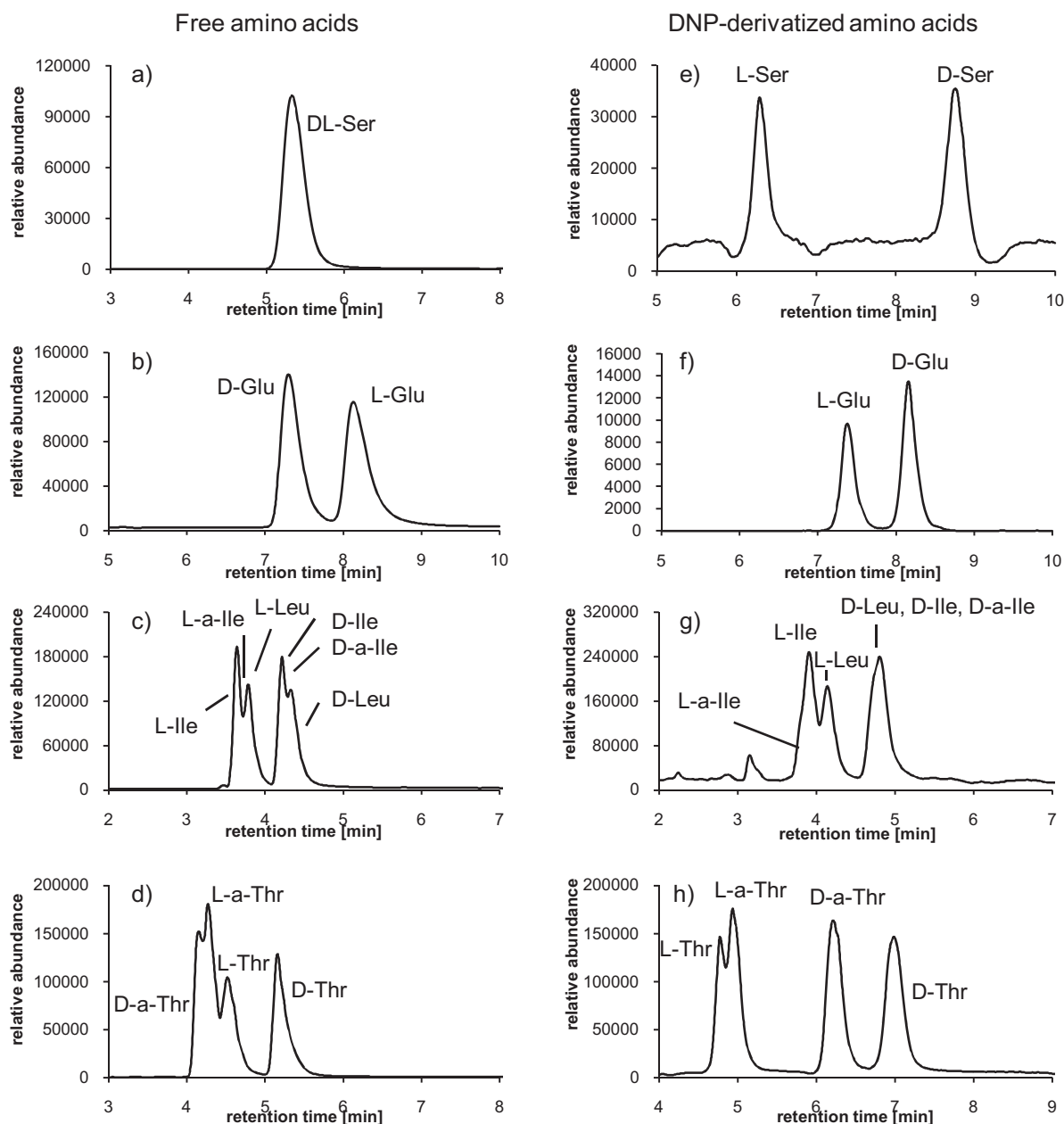
Fig. 3. Amino acid enantiomer separations by HPLC-ESI-QTOF-MS on Chiralpak ZWIX(+): All amino acids analyzed simultaneously in a mixture (XICs: [M-H]⁻ m/z: Ser:104.0354, Thr:118.0510, Glu:146.0460, Leu:130.0874, DNP-Ser:270.0368, DNP-Thr:284.0525, DNP-Glu:312.0474, DNP-Leu:296.0889; ± 2.5mDa).

Thus, as confirmatory method GC–MS using Chirasil L-Val as chiral capillary column was utilized for determination of the absolute configurations in the lipopeptide hydrolysate [10,11]. Amino acids are polar and require double derivatization at both amino and carboxylic acid terminus for GC. Thus, after hydrolysis of the lipopeptide the carboxylic acid group of amino acids was derivatized to the ethyl ester in a first step and the amino group to the *N*-trifluoroacetyl-derivative in a second derivatization step. Although the analysis time is much longer in the GC–MS method than in the HPLC-MS runs, GC–MS with Chirasil L-Val enabled the simultaneous enantiomer separation of all amino acids including isomeric leucines (Leu, Ile, a-Ile) and isomeric threonines (Thr, a-Thr) (Fig. 4a). Fig. 4a shows the extracted ion chromatograms of characteristic diagnostic ions of racemic mixtures of the amino acids which were part of the lipopeptide. It can be seen that all compounds of interest were well resolved and that without any exemption D-amino acids eluted before the corresponding L-amino

acid enantiomers. Analysis of the hydrolyzed lipopeptide by GC–MS finally revealed that the lipopeptide poaeamide is constituted of D-Ser, D-a-Thr, D-Glu, L-Ile. For Leu on the other hand an enantiomer ratio of 4:1 was determined, i.e. 1 of the 5 leucine residues had D-configuration. It becomes evident that a significant portion (50%) of the amino acids in the lipopeptide has D-configuration which is not uncommon for this class of natural compounds.

3.2. Determination of absolute configurations of 3-hydroxy fatty acids

A characteristic feature of lipopeptides is a fatty acid side chain which can be linked via ester or amide bonds to a linear or cyclic peptide structure. Prior MS and NMR structure elucidation showed that a 3-hydroxydecanoic acid side chain is attached to the exocyclic portion of the peptide moiety of poaeamide. Absolute configurations of 3-hydroxy acids in lipopeptides have
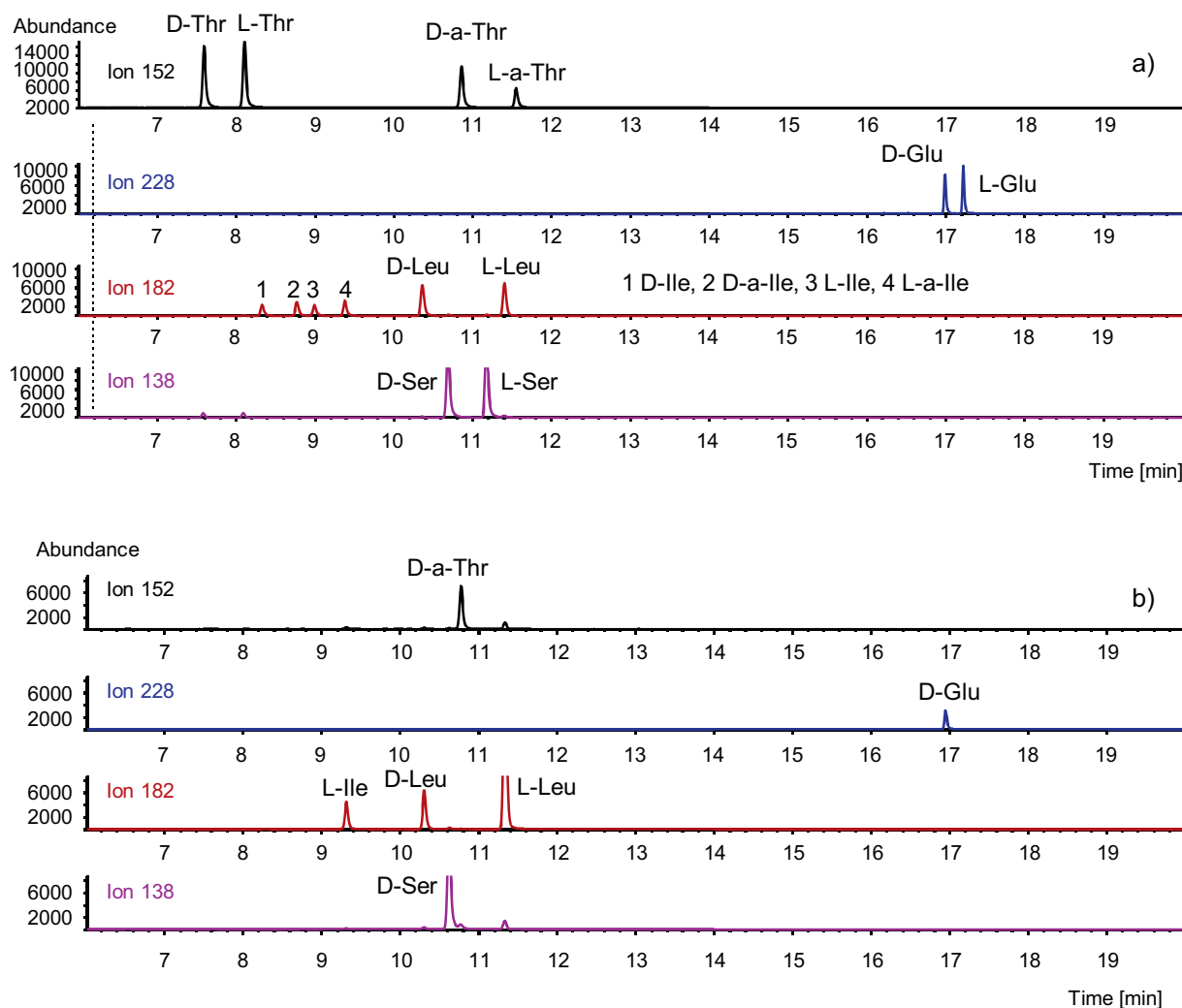
Fig. 4. Amino acid enantiomer separations by GC-EI-MS on Chirasil L-Val: (a) Corresponding EICs of diagnostic ions characteristic for the amino acids present in the peptide. (b) Characteristic EICs of hydrolyzed poaeamide sample.

been frequently assumed to be *R*-configured in the literature or have not been specified at all. 3-Hydroxy fatty acids are formed enzymatically in the course of β–oxidation of fatty acids. After dehydrogenation in α/β-position of the carboxylic acid, enoyl-CoA hydratases (ECHs) catalyze the *syn* hydration of trans-enoyl-CoA to produce 3-hydroxyacyl-CoA in the second step of β-oxidation [41]. Interestingly, two classes of ECHs are known, of which one (ECH-1) is *S*-specific while the other (ECH-2) is *R*-specific. In order to provide a method for experimental confirmation of the absolute configuration in 3-hydroxy fatty acids and in particular 3-hydroxydecanoic acid, the fatty acid building block of poaeamide, an enantioselective HPLC-ESI-MS assay for experimental determination and verification of absolute configurations of 3-hydroxy acids was established. It is a challenging task and only few methods were reported. An HPLC method incorporating pre-column derivatization and subsequent HPLC analysis employing a polysaccharide CSP has been described recently [41]. Ianni et al. reported that 3-hydroxy alkanoic acids can be resolved into enantiomers on Chiralpak ZWIX without derivatization [45]. This method was adopted herein and chromatographic conditions adapted to be compatible with MS detection. In this part of the study, an ion-trap mass spectrometer was used and coupled to the HPLC instrument with the Chiralpak ZWIX columns by an ESI interface.

Since no authentic enantiomerically pure standards were available, enantiomerically enriched (*R*)-3-hydroxydecanoic acid was isolated from a commercially available rhamnolipid known to have largely *R*-configured 3-hydroxy decanoic acid as part of the fatty acid ester moiety attached to the sugar residue, in order to establish enantiomer elution orders on Chiralpak ZWIX(+) and ZWIX(−). This allowed to establish elution orders to be *R* before *S* on Chiralpak ZWIX(+) and *S* before *R* on Chiralpak ZWIX(−).

At this point it must be emphasized that it was of utmost importance to extract the 3-hydroxy decanoic acid into an organic phase after hydrolysis. This is valid for both the acid hydrolysate of the rhamnolipid standard as well as the acid hydrolysate of the lipopeptide sample. Direct injection of hydrolyzed rhamnolipid standard and hydrolyzed lipopeptide sample gave severe retention time shifts on the Chiralpak ZWIX columns which are disturbing when the goal is to pinpoint the stereochemical identity. (Zwitterionic) chiral ion-exchangers such as Chiralpak ZWIX do not well tolerate high salt loads in the sample matrix as such conditions represent strong elution conditions leading to retention time shifts. To avoid wrong assignments, the sample was injected additionally after spiking to the racemate. As can be seen from Fig. 5a, the 3-hydroxy carboxylic acid was determined on Chiralpak ZWIX(+) to have *R*-configuration. In order to validate the assignment, sample and spiked racemate were also injected
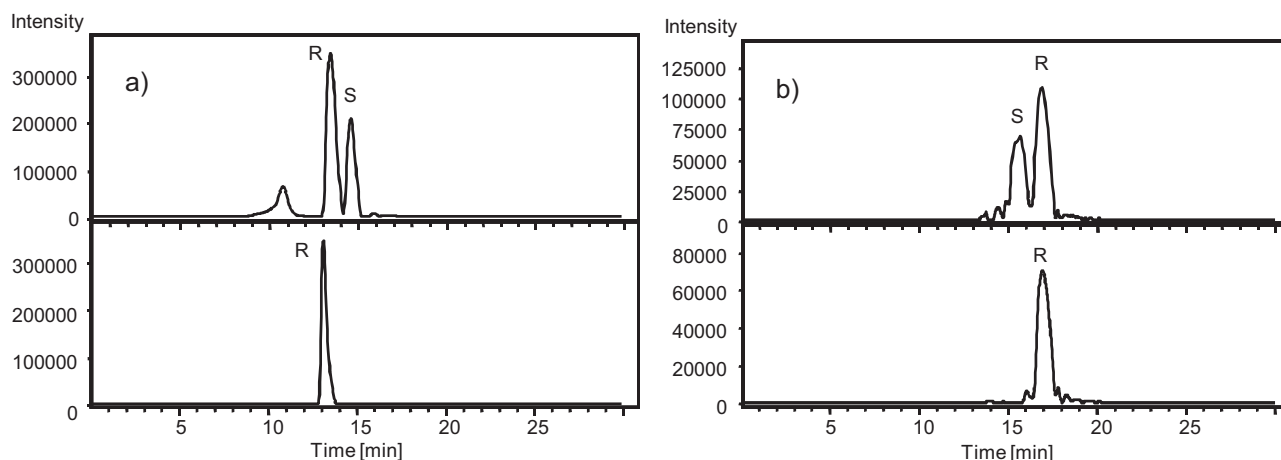
Fig. 5. HPLC-ESI-IT-MS analysis of 3-hydroxydecanoic acid (lipid side chain of poaeamide) on Chiralpak ZWIX(+) (a) and Chiralpak ZWIX(−) (b). Top trace, racemate spiked with sample; bottom trace, sample, i.e. 3-hydroxydecanoic acid isolated from poaeamide.

onto the Chiralpak ZWIX(−) column which showed reversed elution order confirming the *R*-configuration of 3-hydroxydecanoic acid side chain of poaeamide (data summarized in Table S2 of Suppl. Material). Such reversal of elution orders are commonly not possible with CSPs derived from natural chiral pool but is often possible with quinine and quinidine derived ion-exchange type CSPs, such as Chiralpak ZWIX(+) and ZWIX(−) due to their pseudo-enantiomeric character, i.e. opposite configurations in positions C8/C9 of the cinchonan moiety and the cyclohexane moiety as well. Since these are the stereoconfigurations within the active chiral recognition site, enantiomeric behavior is often observed in spite of their actually diastereomeric nature (both chiral selectors have 1*S*,3*R*,4*S*-configurations).

3.3. Hydrolysis of cyclic ester (lactone)

A strategy implementing enzymatic and chemical digestion was utilized for the identification of the position of the D-Leu residue in the peptide sequence (*vide infra*). Enzymatic digestion, however, is only possible on a linear peptide chain and therefore a hydrolysis step of the cyclic ester and lactone group, respectively, was developed. Its success was monitored chromatographically.

Treatment of the cyclic lipopeptide with aqueous 1N NaOH for 90 min at 45 °C gave the hydrolyzed linear peptide with opened lactone ring in good yields. Amide bonds were not cleaved under such conditions, yet a minor side product with *m/z* 1253.7967 was formed. MS spectra confirmed that this impurity originates from water cleavage in the a-Thr residue of the linear peptide (see Suppl. Mat., Figure S2). It may be regarded as an indication for the position of ring closure. While MS spectra in RPLC-MS runs of hydrolyzed poaeamide (Fig. 6a, middle trace) indicated the product formation, RPLC provided insufficient selectivity between hydrolyzed and cyclic lipopeptide which actually coeluted on the ODS Hypersil 5 μm column (Fig. 6a, bottom trace). UHPLC with a Kinetex C18 2.6 μm core–shell particle column provided enough efficiency to furnish baseline separation, at modest selectivity yet (Fig. 6b). Attempts to improve the separation by use of a zwitterionic HILIC column completely failed and only a slight shoulder of the linearized peptide was visible in front of the educt (poaeamide) (Fig. 6c, bottom trace). In contrast, Chiralpak ZWIX exhibited greatly enhanced chromatographic selectivities between the cyclic educt and the linear hydrolysis product (see Fig. 6d, bottom trace).

It seems that retention is dominated by the lipophilic fatty acid moiety both in RPLC and HILIC leading to poor retentivity in HILIC

and (very) similar retention thus poor or lack of resolution in RPLC. Due to specific molecular adsorption and binding to the quinine carbamate selector of the Chiralpak ZWIX stationary phase driven by specific binding domains, this selector appears to be highly sensitive to minor structural changes in the polar domains of the peptide when it is run in polar organic mode elution conditions. The additional C-terminal carboxylic acid in the linear peptide strongly drives retention and leads to recognition of minute charge differences in cyclic and linear peptide forms. The fact that the zwitterionic HILIC phase is not sensitive for these charge differences indicates that the Chiralpak ZWIX column may involve other binding increments for selective recognition as well. In any case, it becomes evident that the Chiralpak ZWIX column may be a valuable complementary tool for studies on synthetic and natural peptide analysis, not least due to its stereoisomer (enantiomers, epimers, diastereomers) recognition capabilities demonstrated for peptides recently [29,30,38,39,46].

3.4. Position of D-Leu residue in the peptide chain

Amino acid configuration analysis showed that 1 out of 5 Leu residues has D-configuration. Since sequence information was lost due to hydrolysis of the sample, the position of the D-Leu amino acid residue in the peptide chain was unclear and had to be determined. Chiralpak ZWIX has proven capability for peptide stereoisomer separation but authentic standards of each possible stereoisomer of this peptide comprised of non-natural amino acids with lipid modification were too expensive. To locate the position of the D-Leu residue, enzymatic and chemical digestion to produce smaller fragments followed by subsequent hydrolysis of isolated fractions and enantioselective analysis at the amino acid level was envisioned as a strategy to solve the stereochemistry of poaeamide.

3.4.1. Pepsin digestion

Pepsin is an endopeptidase which cleaves amide bonds preferentially at the carboxyl side of Phe, Leu and partly Glu, thus showing to some extent sequence selectivity. D-configurations at such residues preclude cleavage and hence it was expected that D-Leu should be present in the interior of a peptide fragment while C-terminal Leu would indicate an L-configuration in pepsin digested peptides.

Along this line, the linearized peptide was subjected to pepsin digestion at 37 °C for 90 min and the RPLC chromatogram of the digest showed several peptide fragments. Identified peptides and their retention times as well as accurate masses are summarized
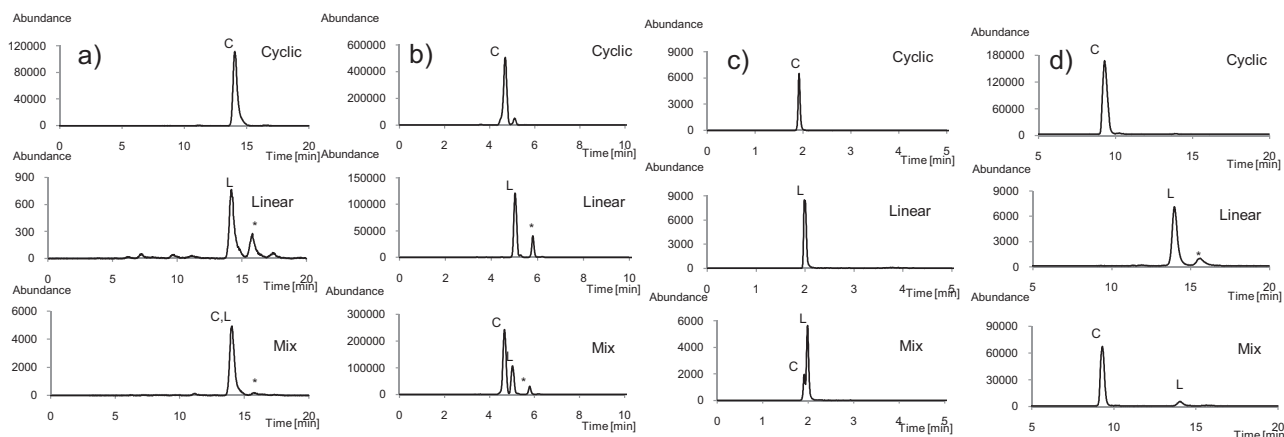
Fig. 6. (U)HPLC-ESI-QTOF-MS Analysis of linearized poaeamide (cyclopeptide hydrolyzed with 1N NaOH). (a) RP-HPLC on ODS-Hypersil 5 μm, (b) RP-UHPLC on Kinetex C18 2.6 μm, (c) HILIC run on ZIC-HILIC 3.5 μm, and (d) HPLC in polar organic mode on Chiralpak ZWIX(+). C, Cyclic poaeamide; L, linear poaeamide; * indicates side product of hydrolysis (allo-threonine with water cleavage, i.e. linear peptide minus H_2O; see supplementary information for further support; it may be regarded as an indication for the position of ring closure) (Note, new standard mix used in c), therefore different intensity ratios of linear and cyclic peptide compared to other chromatograms).

in Table 1 and XICs of relevant peptides used for clarification of the stereochemistry are shown in Fig. 7. Since the digestion was incomplete, linear peptide with m/z 1271.8072 in positive ion mode was found in the digest eluting at 12.283 min on the Kinetex C18 core–shell column with employed gradient elution conditions (Fig. 7a). The peak with m/z 1253.7967 eluting shortly after the linear peptide corresponds to the impurity formed during alkaline ester cleavage due to water loss at the a-Thr residue (see above and Suppl. Mat., Figure S2). Several small peptides, which were hardly retained, amongst them a peak corresponding to LSL, SLL and/or LSI tripeptides with m/z 331.2107, were detected at 0.408 min (the XIC for m/z 331.2107 in Fig. 7b provided also a peak at the retention time of the undigested linear peptide which can be assigned to any of those tripeptides formed by in-source fragmentation). Fig. 7c shows the XIC generated for the peptide fragment 3-hydroxydecanoyl-(3-HDA)-LETLL. Several peaks were detected, many of them originating from in-source fragmentation of larger peptide fragments as indicated. All major peaks were isolated with a fraction collector in a subsequent semi-preparative run. The isolated peptides were then subjected to acid hydrolysis with DCl/D_2O followed by double derivatization and enantioselective GC–MS analysis on Chirasil L-Val. No D-Leu was found in the tripeptides eluting at 4.1 min and thus Leu-residues 7 and 8 were L-configurated. In contrast, D-Leu was detected in the peptide fragment 3-HDA-LETLL which was eluted and isolated at ca. 8.6 min. Assuming that the C-terminal Leu residue bears L-configuration

because otherwise cleavage by pepsin would have been blocked, it can be concluded that either Leu residue 1 or 4 carry the D-amino acid.

3.4.2. Chemical digestion

Pepsin digestion did not yield smaller peptide fragments of the N-terminus of linear poeamide. In order to get smaller fragments of the N-terminus in which only one Leu residue was present, a chemical digestion approach was devised. Thus, first the extent of digestion in dependence of reaction time using acidic hydrolysis with 6N HCl at 110 °C was explored to produce sufficiently small peptide fragments, but avoid complete digestion to amino acid level. Thus, the digestion kinetics was studied. It is evident that a large diversity of peptides is generated by chemical digestion in contrast to enzymatic digestion by pepsin. The relevant peptides being of help for locating the D-Leu residue in the peptide chain are those containing either the Leu-1 or Leu-4 residue. Fig. S3 (Suppl. Material) reveals the decline of the peak area of 3 diagnostic peptide ions containing either Leu-1 or Leu 4 in dependence of the digestion time. The kinetic data (rate constant for degradation k and half-time $t_{1/2}$) are summarized in Table 2. It turns out that maximal peak area is observed soon after start of the chemical reaction and fragments completely disappear after ca. 150 min. Hence, a digestion time of 20 min was finally selected. The peptide fragments formed under such conditions are summarized in Table 3. Fractions were collected for specific peptides as indicated in Table 3. Fig. 8 depicts XICs of RPLC runs on Kinetex C18 of collected fractions. Their amino acid enantiomer analysis by GC–MS on Chirasil L-Val revealed that the Leu residue of 3-HDA-L and 3-HDA-LE had L-configuration. In contrast, D-Leu residue was found in the isolated peptides LETL and TLL. Consequently, the D-Leu residue was confirmed to be present in position 4 of the peptide providing the complete stereochemistry of the lipopeptide poeamide as follows: D-Leu in residue 4 and D-amino acids in position 2, 3, 6 and 9.

Table 1

Pepsin digest of linearized poaeamide analyzed by RP-UHPLC on Kinetex C18 2.6 μm core–shell column indicating also D-Leu presence/absence in different isolated fractions.

Identified fragments	Extraction mass (Da) ±0.5 mDa	RT (min)	Comment	D-Leu
LSL, LSI, SLL	331.2107	0.41	Could be each of these	No
3HDA-LETLL	757.4837	9.06		Yes
3HDA-LETLLSL	957.5998	9.37		Yes
3HDA-LETLLSLL	1070.6838	10.54		Yes
ETLLSLLSI	987.5852	12.28	In source decay intact peptide	-
LSLLSI	644.4109	12.28		-
SL	218.1266	12.29		-
SI	218.1266	12.29		-
3HDA-LETL	644.3996	12.29		-
SLLSI	531.3268	12.29		-
LSLL	444.2948	12.29		-

Table 2

Kinetic data of chemical digestion for 3 different peptide fragments.

	Linear equation[a]	R^2	k [min^{-1}]	$t_{1/2}$ [min]
3HDA-L	$y = -0.0264x + 9.0933$	0.9892	0.026	26.26
TLL	$y = -0.0180x + 8.9011$	0.9480	0.018	38.51
LETL/ETLL	$y = -0.0393x + 9.1352$	0.9520	0.039	17.64

[a] $\ln y = \ln y_0 - kt$.
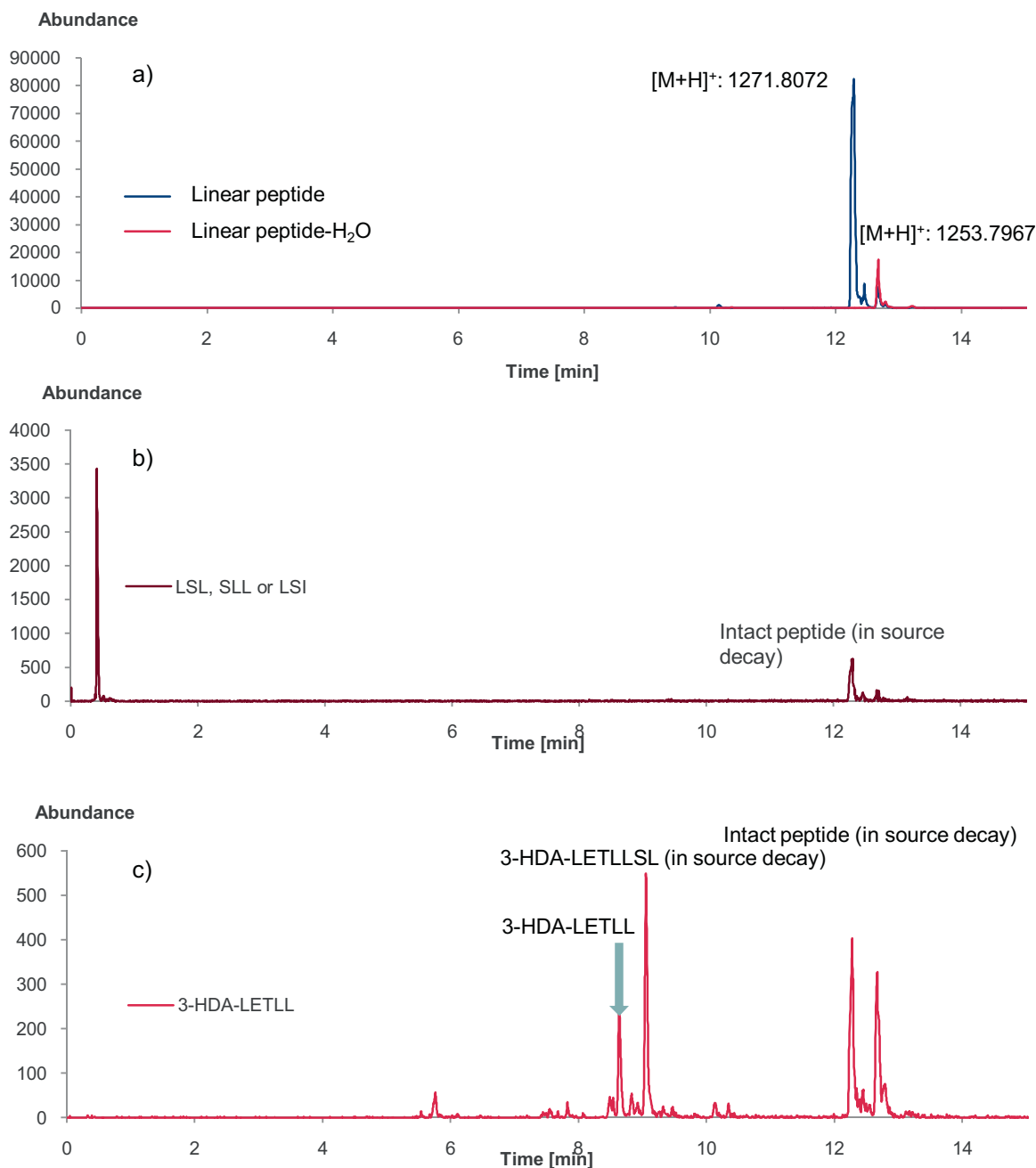
Fig. 7. UHPLC-ESI-QTOF-MS analysis of pepsin digested linear poaeamide. The chromatograms show the collected fractions for the verification of presence of D-Leu in peptide fragments. (a) XIC of m/z 1271.8072 (trace in blue) and of m/z 1253.7967 (trace in red). (b) XIC of m/z 331.2107. (c) XIC of m/z 757.4837. Column: Kinetex C18, 2.6 μm. Extraction window, 5 mDa (±2.5 mDa) (For interpretation of the color information in this figure legend, the reader is referred to the web version of the article.).

3.5. Potential of Chiralpak ZWIX for complementary peptide separations

RPLC and HILIC showed some deficiencies in above peptide separations in particular for small peptides. The last goal of this study was therefore to elucidate whether Chiralpak ZWIX can generate complementary retention profiles to RPLC and HILIC in peptide separations. To demonstrate this, the chemical digest of poaeamide was analyzed by RPLC (on Kinetex C18), HILIC (on ZIC-HILIC) and polar organic mode (on Chiralpak ZWIX). Orthogonality plots can be utilized to display the complementarity of the respective chromatographic modes (Fig. 9; the corresponding orthogonality plot between RPLC and HILIC is given in Figure S4 of Suppl. Material).

Fig. 9a shows the corresponding orthogonality plot between RPLC on Kinetex C18 core-shell column and zwitterionic chiral ion-exchanger Chiralpak ZWIX in polar organic elution mode. It is striking that the data points are relatively evenly distributed in the parity plot indicating good orthogonality. It can be seen that the peaks eluting < 2 min in RPLC are well retained and evenly spread in the Chiralpak ZWIX separation dimension. This chiral zwitterionic ion-exchanger may therefore be regarded as strongly complementary to RPLC. Therefore, it is supposed to be valuable for peptide separations in which selectivity problems of RPLC are observed and highly suitable for 2D-LC peptide separations.

The data points are even more uniformly distributed in the orthogonality plot spanned by the retention times on the ZIC-HILIC
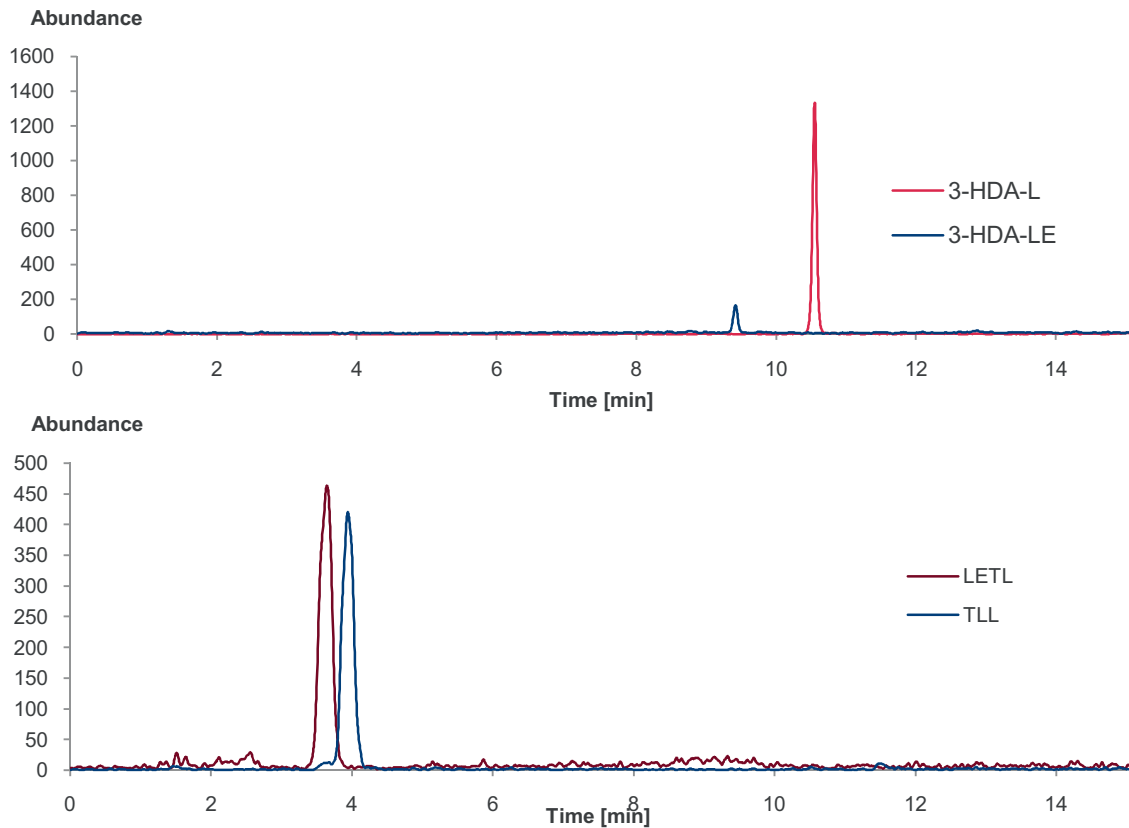
Fig. 8. UHPLC-ESI-QTOF-MS analysis of chemically digested poaeamide. The chromatograms show the collected fractions for the verification of presence of D-Leu in peptide fragments. (a) XIC of m/z 430.2679 (trace in blue) and of m/z 301.2253 (trace in red). b) EIC of m/z 345.2264 (trace in blue) and of m/z 474.2689 (trace in red). Column: Kinetex C18, 2.6 μm. Extraction window, 5 mDa (±2.5 mDa) (For interpretation of the color information in this figure legend, the reader is referred to the web version of the article.).

column and the polar organic mode separation on the Chiralpak ZWIX column (Fig. 9b). It seems that these two chromatographic methods are a perfect combination for two-dimensional peptide separations. Elution conditions of these two complementary modes

(HILIC and polar organic mode) are better compatible and their combination in a 2D-method appears more straightforward than of RPLC and HILIC. It should also be emphasized again that Chiralpak ZWIX has capability to separate stereoisomers of (small) peptides
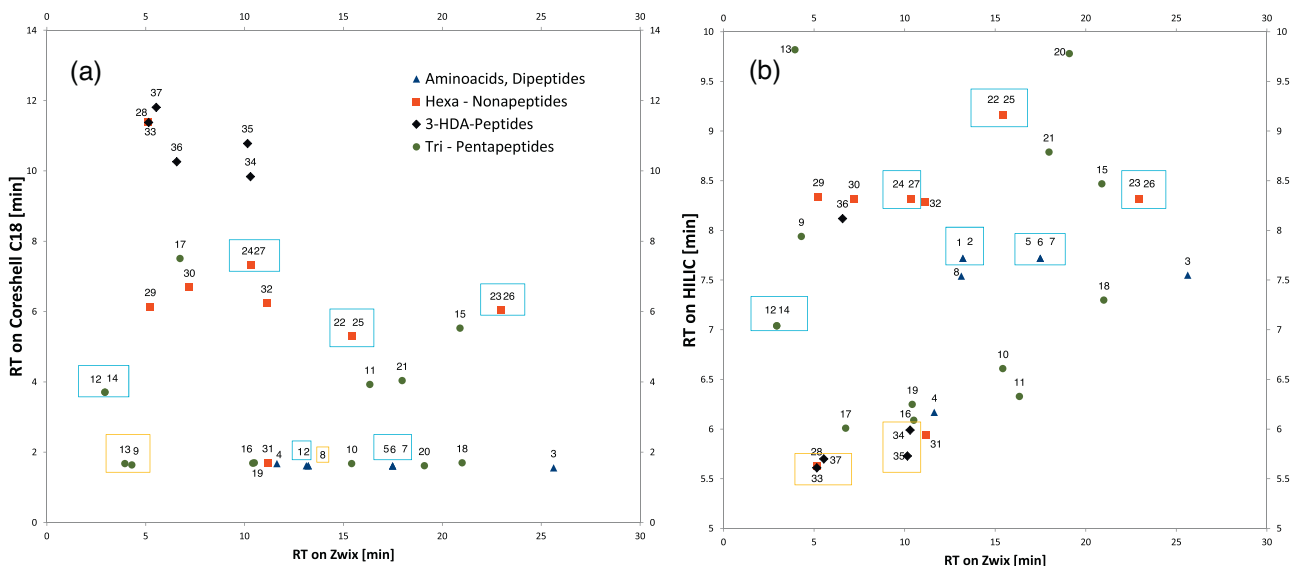
Fig. 9. Orthogonality plots measuring complementarity of retention profiles of distinct chromatographic modes. (a) RPLC (Kinetex C18, 2.6 μm) vs HPLC in polar organic mode on chiral zwitterionic ion-exchanger (Chiralpak ZWIX(+), 3 μm), (b) HPLC in polar organic mode on chiral zwitterionic ion-exchanger (Chiralpak ZWIX(+), 3 μm) vs HILIC (ZIC-HILIC, 3.5 μm). Amino acids or peptides which cannot be distinguished by fragmentation spectra are surrounded by a blue dotted line and could be each of these. Coeluting amino acids or peptides are surrounded by an orange solid line. Amino acids and dipeptides: 1 I; 2 L; 3 LE; 4 LL; 5 LS; 6 SI; 7 SL; 8 TL; tri- to pentapeptides: 9 ETL; 10 SLL; 11 TLL; 12 ETLL; 13 ETLLS; 14 LETL; 15 LETLL; 16 LSLL; 17 LLSLL; 18 SLLS; 19 SLLSI; 20 TLLS; 21 TLLSL; hepta- to nonapeptides: 22 ETLLSL; 23 ETLLSLL; 24; ETLLSLLSI; 25 LETLLS; 26; LETLLSL; 27 LETLLSLLS; 28 LLSLLSLL; 29 LSLLSI; 30 TLLSLL; 31 TLLSLLS; 32 TLLSLLSI; 3-HDA-peptides: 33 3-HDA-LETLL; 34 3-HDA-LETLLS; 35 3-HDA-LETLLSL; 36 3-HDA-LETLLSLL; 37 3-HDA-LETLLSLLSI (For interpretation of the color information in this figure legend, the reader is referred to the web version of the article.).

Table 3
Chemical digest of poaeamide analyzed by RP-UHPLC on Kinetex C18 2.6 μm core–shell column indicating also D-Leu presence/absence in different isolated fractions.

Identified fragments	Extraction mass ±0.5 mDa	RT (min)	Comment	D-Leu
SI	218.1266	1.55	Co-elution	Not collected
SL	218.1266	1.55		
LS	218.1266	1.55		
L	131.0946	1.55		
I	131.0946	1.55		
TLLS	432.2584	1.55		
TL	232.1423	1.57		
LET	361.1849	1.58		
ETL	361.1849	1.58		
ETLLS	561.301	1.63		
LLS/LSL/SLL/LSI	331.2107	2.47	Could be each of these	Not collected
LL	244.1787	2.51		
ETLL	474.2689	3.64	Could be each of these	Yes
LETL	474.2689	3.64		
TLL	345.2264	3.95		Yes
LETLLS	674.385	5.17	Could be each of these	Not collected
ETLLSL	674.385	5.17		
3-HDA-LE	430.2679	9.42		No
3-HDA-L	301.2253	10.55		No

(enantiomers, epimers, and diastereomers) as shown recently on some examples [29,30,38,39]. It is without doubt that the Chiralpak ZWIX column could be a valuable tool for complementary peptide separations, worth to be tested as secondary retention principle.

4. Conclusions

A set of complementary and confirmative methods has been suggested to support complete structural elucidation of the stereochemistry of lipopeptides. Chiralpak ZWIX allows LC enantiomer and stereoisomer separations of amino acids, hydroxy alkanoic acids and peptides and was of central importance. Its ESI compatible elution conditions enabled straightforward coupling to MS a sine qua non in structure elucidations. Chirasil L-Val turned out to be invaluable for confirmation of amino acid configurations. The determination of the position of absolute configurations in the peptide sequence in case of multiple amino acids of one type and detection of both configurations in the peptide is more challenging. Fragmentation of the peptide into smaller fragments by enzymatic and chemical digestion provided diagnostic peptides which could be used to solve this issue. For the entire study, highly pure starting peptides are needed for this purpose otherwise the results may get confusing, e.g. if peptide or protein impurities are present. The proposed methods can be analogously applied to solve similar stereochemical problems in other peptides, cyclopeptides, lipopeptides, depsipeptides and impurity profiling of therapeutic peptides as well.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chroma.2015.05.065

References

[1] T. Schneider, A. Müller, H. Miess, H. Gross, Cyclic lipopeptides as antibacterial agents—potent antibiotic activity mediated by intriguing mode of actions, Int. J. Med. Microbiol. 304 (2014) 37–43.
[2] R.H. Baltz, Combinatorial biosynthesis of cyclic lipopeptide antibiotics: a model for synthetic biology to accelerate the evolution of secondary metabolite biosynthetic pathways, ACS Synth. Biol. 3 (2012) 748–758.
[3] G. Dey, R. Bharti, R. Sen, M. Mandal, Microbial amphiphiles: a class of promising new-generation anticancer agents, Drug Discov. Today 20 (2015) 136–146.
[4] R. Geys, W. Soetaert, I. Van Bogaert, Biotechnological opportunities in biosurfactant production, Curr. Opin. Biotechnol. 30 (2014) 66–72.
[5] J.M. Raaijmakers, I. de Bruijn, M.J.D. de Kock, Cyclic lipopeptide production by plant-associated pseudomonas spp.: diversity, activity, biosynthesis, and regulation, Mol. Plant-Microbe Interact. 19 (2006) 699–710.
[6] E.J. Gudiña, V. Rangarajan, R. Sen, L.R. Rodrigues, Potential therapeutic applications of biosurfactants, Trends Pharmacol. Sci. 34 (2013) 667–675.
[7] M. Deleu, J.-M. Crowet, M.N. Nasir, L. Lins, Complementary biophysical tools to investigate lipid specificity in the interaction between bioactive molecules and the plasma membrane: a review, Biochim. Biophys. Acta 1838 (2014) 3171–3190.
[8] H. Brückner, N. Fujii, D-Amino Acids in Chemistry, Life Sciences, and Biotechnology, Wiley-VCH, Zurich, 2010.
[9] A. Rocco, Z. Aturki, S. Fanali, Chiral separations in food analysis, TrAC, Trends Anal. Chem. 52 (2013) 206–225.
[10] R. Pätzold, H. Brückner, Chiral separation of amino acids by gas chromatography, in: M.-P. Ibolya (Ed.), Journal of Chromatography Library, Elsevier, 2005, pp. 98–118.
[11] J. Horak, H. Gerhardt, J. Theiner, W. Lindner, Correlation between amino acid racemization and processing conditions for various wheat products, oil seed press cakes and lignin samples, Food Bioprod. Process. 92 (2014) 355–368.
[12] I. Ilisz, R. Berkecz, A. Péter, Application of chiral derivatizing agents in the high-performance liquid chromatographic separation of amino acid enantiomers: a review, J. Pharm. Biomed. Anal. 47 (2008) 1–15.
[13] I. Ilisz, A. Aranyi, A. Péter, Chiral derivatizations applied for the separation of unusual amino acid enantiomers by liquid chromatography and related techniques, J. Chromatogr. A 1296 (2013) 119–139.
[14] H. Brückner, S. Haasmann, M. Langer, T. Westhauser, R. Wittner, H. Godel, Liquid chromatographic determination of d- and l-amino acids by derivatization with o-phthaldialdehyde and chiral thiols: applications with reference to biosciences, J. Chromatogr. A 666 (1994) 259–273.
[15] C. Müller, J.R. Fonseca, T.M. Rock, S. Krauss-Etschmann, P. Schmitt-Kopplin, Enantioseparation and selective detection of D-amino acids by ultra-high-performance liquid chromatography/mass spectrometry in analysis of complex biological samples, J. Chromatogr. A 1324 (2014) 109–114.

[16] R. Bhushan, H. Brueckner, Use of Marfey's reagent and analogs for chiral amino acid analysis: assessment and applications to natural products and biological systems, J. Chromatogr. B 879 (2011) 3148–3161.

[17] S. Hess, A universal HPLC-MS method to determine the stereochemistry of common and unusual amino acids, Methods Mol. Biol. (N.Y., NY, U.S.A.) 828 (2012) 63–75.

[18] R. Bhushan, H. Brückner, Marfey's reagent for chiral amino acid analysis: a review, Amino Acids 27 (2004) 231–247.

[19] R. Bhushan, H. Brückner, Use of Marfey's reagent and analogs for chiral amino acid analysis: assessment and applications to natural products and biological systems, J. Chromatogr. B 879 (2011) 3148–3161.

[20] B. Natalini, R. Sardella, F. Ianni, Enantioseparations by high-performance liquid chromatography based on chiral ligand-exchange, in: G.K.E. Scriba (Ed.), Chiral Separations, Humana Press, New York, 2013, pp. 191–208.

[21] A. Berthod, X. Chen, J.P. Kullman, D.W. Armstrong, F. Gasparrini, I. D'Acquaric, C. Villani, A. Carotti, Role of the carbohydrate moieties in chiral recognition on teicoplanin-based LC stationary phases, Anal. Chem. 72 (2000) 1767–1780.

[22] M.J. Desai, D.W. Armstrong, Analysis of native amino acid and peptide enantiomers by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry, J. Mass Spectrom. 39 (2004) 177–187.

[23] I. Ilisz, R. Berkecz, A. Péter, HPLC separation of amino acid enantiomers and small peptides on macrocyclic antibiotic-based chiral stationary phases: a review, J. Sep. Sci. 29 (2006) 1305–1321.

[24] Z. Pataj, I. Ilisz, R. Berkecz, A. Misicka, D. Tymecka, F. Fülöp, D.W. Armstrong, A. Péter, Comparison of performance of Chirobiotic T, T2 and TAG columns in the separation of β2- and β3-homoamino acids, J. Sep. Sci. 31 (2008) 3688–3697.

[25] M.H. Hyun, Enantioseparations of primary amino compounds by high-performance liquid chromatography using chiral crown ether-based chiral stationary phase, Methods Mol. Biol. 970 (2013) 165–176.

[26] E. Bang, J.Y. Jin, J.H. Hong, J.S. Kang, W. Lee, W. Lee, Comparative studies on enantiomer resolution of α-amino acids and their esters using (18-crown-6)-tetracarboxylic acid as a chiral crown ether selector by NMR spectroscopy and high-performance liquid chromatography, Bull. Korean Chem. Soc. 33 (2012) 3481–3484.

[27] W.W. Barnhart, X. Xia, R. Jensen, K.H. Gahm, Comparison of chiral separations of aminophosphonic acids and their aminocarboxylic acid analogs using a crown ether column, Chirality 25 (2013) 369–378.

[28] C.V. Hoffmann, R. Pell, M. Lämmerhofer, W. Lindner, Synergistic effects on enantioselectivity of zwitterionic chiral stationary phases for separations of chiral acids, bases, and amino acids by HPLC, Anal. Chem. 80 (2008) 8780–8789.

[29] T. Zhang, E. Holder, P. Franco, W. Lindner, Method development and optimization on cinchona and chiral sulfonic acid-based zwitterionic stationary phases for enantiomer separations of free amino acids by high-performance liquid chromatography, J. Chromatogr. A 1363 (2014) 191–199.

[30] T. Zhang, E. Holder, P. Franco, W. Lindner, Zwitterionic chiral stationary phases based on cinchona and chiral sulfonic acids for the direct stereoselective separation of amino acids and other amphoteric compounds, J. Sep. Sci. 37 (2014) 1237–1247.

[31] M. Lämmerhofer, Liquid chromatographic enantiomer separation with special focus on zwitterionic chiral ion-exchangers, Anal. Bioanal. Chem. 406 (2014) 6095–6103.

[32] I. Ilisz, Z. Gecse, Z. Pataj, F. Fulop, G. Toth, W. Lindner, A. Peter, Direct high-performance liquid chromatographic enantioseparation of secondary amino acids on Cinchona alkaloid-based chiral zwitterionic stationary phases. Unusual temperature behavior, J. Chromatogr. A 1363 (2014) 169–177.

[33] I. Ilisz, N. Grecso, A. Aranyi, P. Suchotin, D. Tymecka, B. Wilenska, A. Misicka, F. Fulop, W. Lindner, A. Peter, Enantioseparation of β2-amino acids on cinchona alkaloid-based zwitterionic chiral stationary phases. Structural and temperature effects, J. Chromatogr. A 1334 (2014) 44–54.

[34] I. Ilisz, A. Aranyi, Z. Pataj, A. Peter, Recent advances in the direct and indirect liquid chromatographic enantioseparation of amino acids and related compounds: a review, J. Pharm. Biomed. Anal. 69 (2012) 28–41.

[35] Y. Miyoshi, T. Oyama, Y. Itoh, K. Hamase, Enantioselective two-dimensional high-performance liquid chromatographic determination of amino acids; analysis and physiological significance of D-amino acids in mammals, Chromatography 35 (2014) 49–57.

[36] F. Ianni, R. Sardella, A. Lisanti, A. Gioiello, B.T. Cenci Goga, W. Lindner, B. Natalini, Achiral–chiral two-dimensional chromatography of free amino acids in milk: a promising tool for detecting different levels of mastitis in cows, J. Pharm. Biomed. Anal. (2015), http://dx.doi.org/10.1016/j.jpba.2014.12.041, ahead of Print.

[37] E. Barbaro, R. Zangrando, M. Vecchiato, C. Turetta, C. Barbante, A. Gambaro, D-and L-amino acids in Antarctic lakes: assessment of a very sensitive HPLC-MS method, Anal. Bioanal. Chem. 406 (2014) 5259–5270.

[38] S. Wernisch, W. Lindner, Versatility of cinchona-based zwitterionic chiral stationary phases: enantiomer and diastereomer separations of non-protected oligopeptides utilizing a multi-modal chiral recognition mechanism, J. Chromatogr. A 1269 (2012) 297–307.

[39] S. Wernisch, O. Trapp, W. Lindner, Application of cinchona-sulfonate-based chiral zwitterionic ion exchangers for the separation of proline-containing dipeptide rotamers and determination of on-column isomerization parameters from dynamic elution profiles, Anal. Chim. Acta 795 (2013) 88–98.

[40] C. Zachow, G. Jahanshah, I.d. Bruijn, C. Song, F. Ianni, Z. Pataj, H. Gerhardt, I. Pianet, M. Lämmerhofer, G. Berg, H. Gross, J.M. Raaijmakers, The novel lipopeptide poaeamide of the endophyte *Pseudomonas poae* RE*1-1-14 is involved in pathogen suppression and root colonization, Mol. Plant-Microbe Interact. (2015).

[41] A.M. Abdel-Mawgoud, F. Lépine, E. Déziel, A chiral high-performance liquid chromatography–tandem mass spectrometry method for the stereospecific analysis of enoyl-coenzyme A hydratases/isomerases, J. Chromatogr. A 1306 (2013) 37–43.

[42] K. Hamase, Y. Nakauchi, Y. Miyoshi, R. Koga, N. Kusano, H. Onigahara, H. Naraoka, H. Mita, Y. Kadota, Y. Nishio, M. Mita, W. Lindner, Enantioselective determination of extraterrestrial amino acids using a two-dimensional chiral high-performance liquid chromatographic system, Chromatography 35 (2014) 103–110.

[43] Y. Miyoshi, R. Koga, T. Oyama, H. Han, K. Ueno, K. Masuyama, Y. Itoh, K. Hamase, HPLC analysis of naturally occurring free D-amino acids in mammals, J. Pharm. Biomed. Anal. 69 (2012) 42–49.

[44] R.J. Reischl, W. Lindner, Methoxyquinoline labeling—a new strategy for the enantioseparation of all chiral proteinogenic amino acids in 1-dimensional liquid chromatography using fluorescence and tandem mass spectrometric detection, J. Chromatogr. A 1269 (2012) 262–269.

[45] F. Ianni, Z. Pataj, H. Gross, R. Sardella, B. Natalini, W. Lindner, M. Lämmerhofer, Direct enantioseparation of underivatized aliphatic 3-hydroxyalkanoic acids with a quinine-based zwitterionic chiral stationary phase, J. Chromatogr. A 1363 (2014) 101–108.

[46] R.J. Reischl, W. Lindner, The stereoselective separation of serine containing peptides by zwitterionic ion exchanger type chiral stationary phases and the study of serine racemization mechanisms by isotope exchange and tandem mass spectrometry, J. Pharm. Biomed. Anal. (2015), http://dx.doi.org/10.1016/j.jpba.2015.02.014.

Supplementary Material

Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides

Heike Gerhardt[1#], Adrian Sievers-Engler[1#], Ghazaleh Jahanshah[1], Zoltán Pataj[1,2], Federica Ianni [1,3], Harald Gross[1], Wolfgang Lindner[4], Michael Lämmerhofer[1*]

[1] *Institute of Pharmaceutical Sciences, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany*

[2] *Department of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Germany*

[3] *Department of Pharmaceutical Sciences, University of Perugia, Via del Liceo 1, 06123 Perugia, Italy*

[4] *Department of Analytical Chemistry, University of Vienna, Waehringer Strasse 38, 1090 Vienna, Austria*

# These authors contributed equally

* Author for correspondence:
michael.laemmerhofer@uni-tuebingen.de
Institute of Pharmaceutical Sciences
Pharmaceutical (Bio)Analysis
University of Tübingen
Auf der Morgenstelle 8
72076 Tübingen, Germany
T +49 7071 29 78793
F +49 7071 29 4565
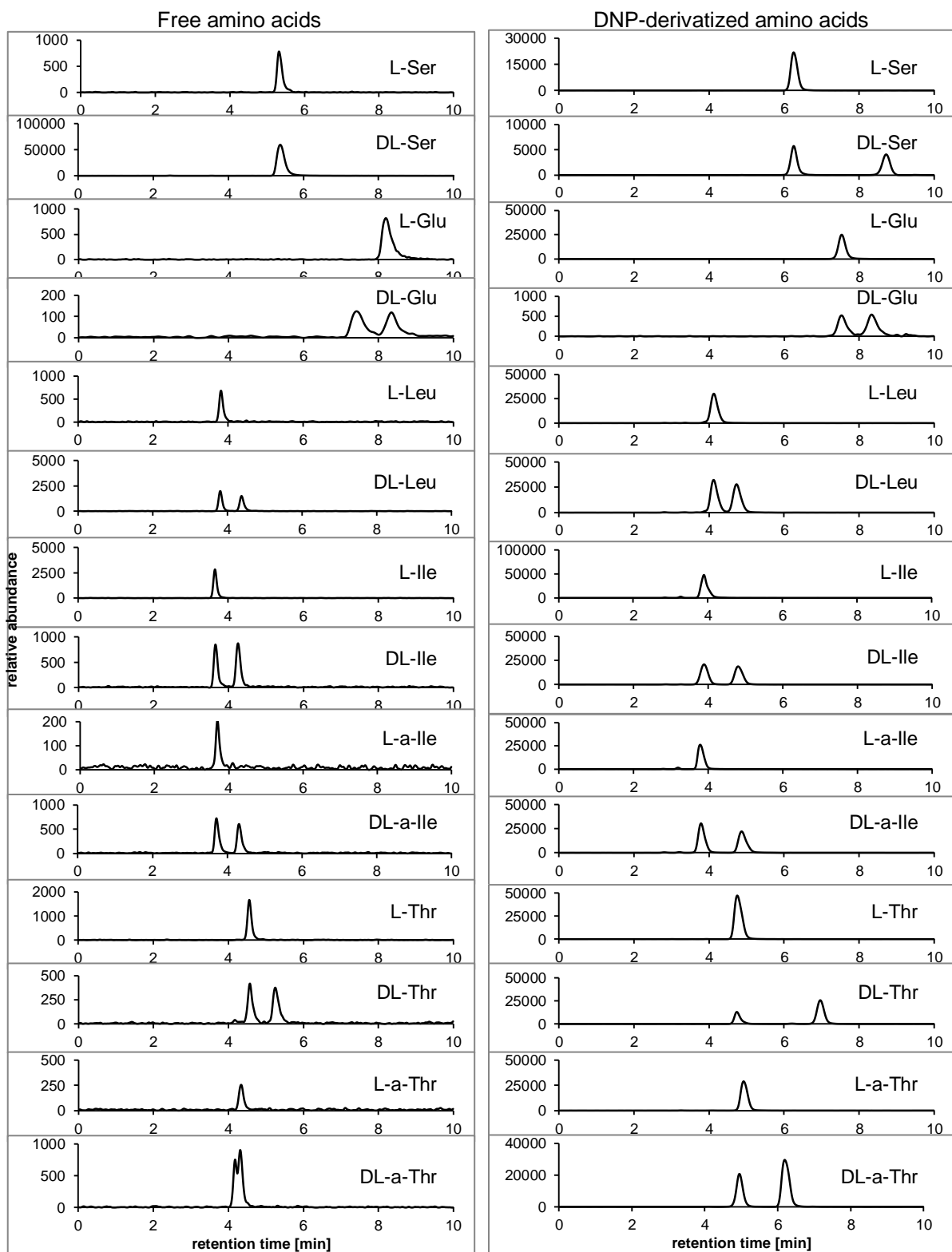
Figure S1. Amino acid enantiomer separations on Chiralpak ZWIX(+): a) underivatized, b) as *N*-2,4-dinitrophenyl (Sanger's) derivatives. Individual racemates and single enantiomers have been injected. Conditions see main document.

Table S1. Liquid chromatographic enantiomer separation data for underivatized and *N*-2,4-dinitrophenyl-derivatized amino acids obtained on the zwitterionic chiral ion-exchanger Chiralpak ZWIX (t_0=2.1 min, 5µl Acetone as t_0 marker).

Amino acid	Free amino acids [a]					DNP-derivatized amino acid [b]				
	D amino acid		L amino acid			D amino acid		L amino acid		
	RT [min]	*k*	RT [min]	*k*	α	RT [min]	*k*	RT [min]	*k*	α
Serine (Ser)	5.33	1.54	5.33	1.54	1.0	8.72	3.15	6.26	1.98	1.6
Glutamic acid (Glu)	7.42	2.53	8.20	2.90	1.1	8.34	2.97	7.54	2.59	1.1
Leucine (Leu)	4.38	1.09	3.81	0.81	1.3	4.74	1.26	4.13	0.97	1.3
Isoleucine (Ile)	4.26	1.03	3.65	0.74	1.4	4.81	1.29	3.90	0.86	1.5
Allo-isoleucine (a-Ile)	4.32	1.06	3.70	0.76	1.4	4.88	1.32	3.80	0.81	1.6
Threonine (Thr)	5.25	1.50	4.56	1.17	1.3	6.97	2.32	4.77	1.27	1.8
Allo-threonine (a-Thr)	4.19	1.00	4.34	1.07	1.1	6.21	1.96	4.93	1.35	1.5

[a] elution order for underivatized amino acids: L before D, except for Glu and a-Thr (D < L)
[b] elution order for DNP-derivatives of amino acids: always L before D

Table S2. Chromatographic results for the enantiomer separation of the *N*-terminally attached fatty acid (FA) moiety of the lipopeptide poaeamide on the zwitterionic chiral ion-exchanger Chiralpak ZWIX.

	Chiralpak ZWIX (+)			Chiralpak ZWIX (-)		
Parameter	Retention time [min]		Elution order	Retention time [min]		Elution order
FA of poaeamide	13.118	-	R	-	16.944	R
Standard + R	13.477	14.669	R < S	15.458	16.764	S < R
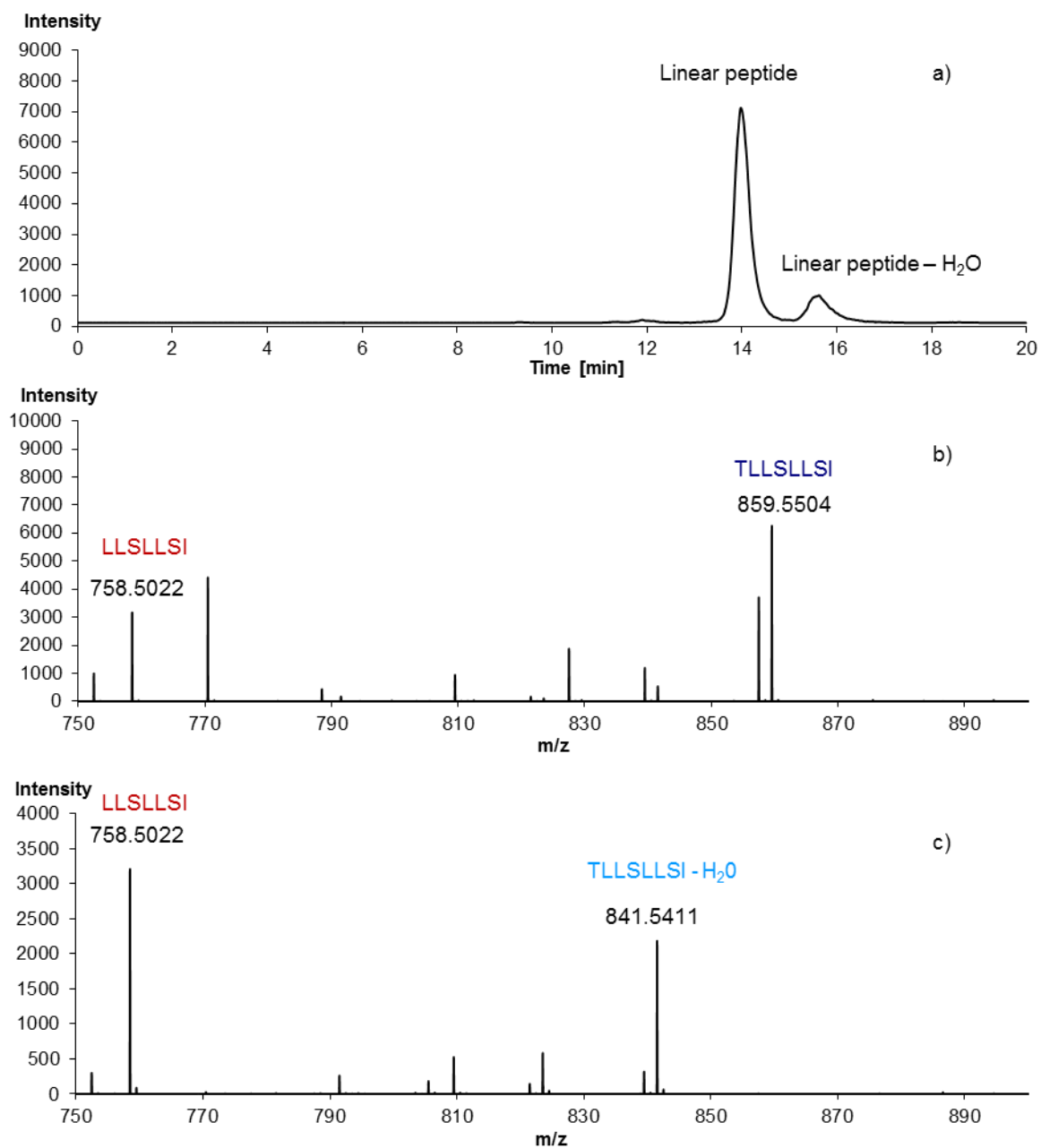
Figure S2: Verification of the structure of the impurity in the sample of the linearized peptide (impurity not detected in the cyclic peptide). a) Chromatogram of the linearized peptide measured on ZWIX (+). b) Fragment mass spectra of the linear peptide (*m/z* 1270.8072). c) Mass spectra of linear peptide with one water cleaved. The spectra show that the water cleavage is on the allo-threonine side chain
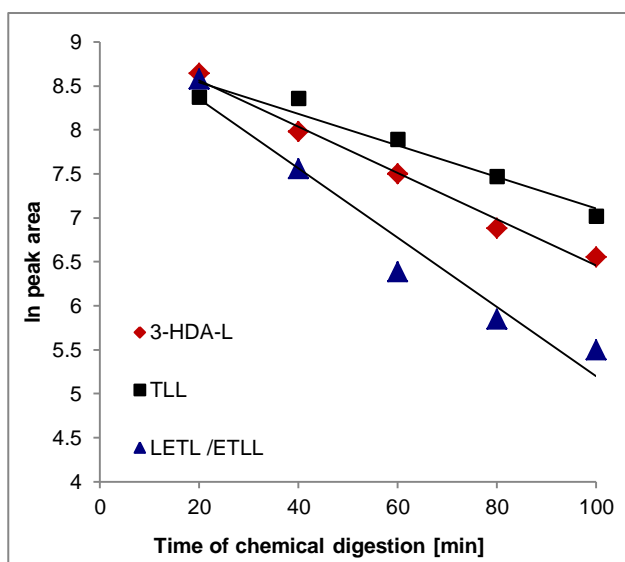
Figure S3. Kinetics of chemical digestion as measured on 3 different peptide fragments.

Figure S4 depicts the orthogonality plot of retention times between HILIC on ZIC-HILIC column and RPLC on Kinetex C18 core-shell column. From this plot a few general patterns become readily evident. The group of amino acids and dipeptides, which elute in a single peak in RPLC (compounds lying on a vertical line and eluting little earlier than 2 min) are reasonably well retained and are better spread in HILIC mode, although still many peptides coelute. The peptides still carrying the 3-hydroxy fatty acid moiety at the *N*-terminus (3-HDA-peptides) are strongest retained in RPLC, yet still sufficiently retained in HILIC mode. Taking the cluster of compounds early eluted in RPLC (< 2 min; except for 18 and 28) aside, the lipopeptide fragments are lying on a negative trend line being indicative for negative correlation of RPLC and HILIC with little complementarity in the two modes. In fact, overall the data points are not widely spread over the two dimensions but essentially clustered into 4 groups.
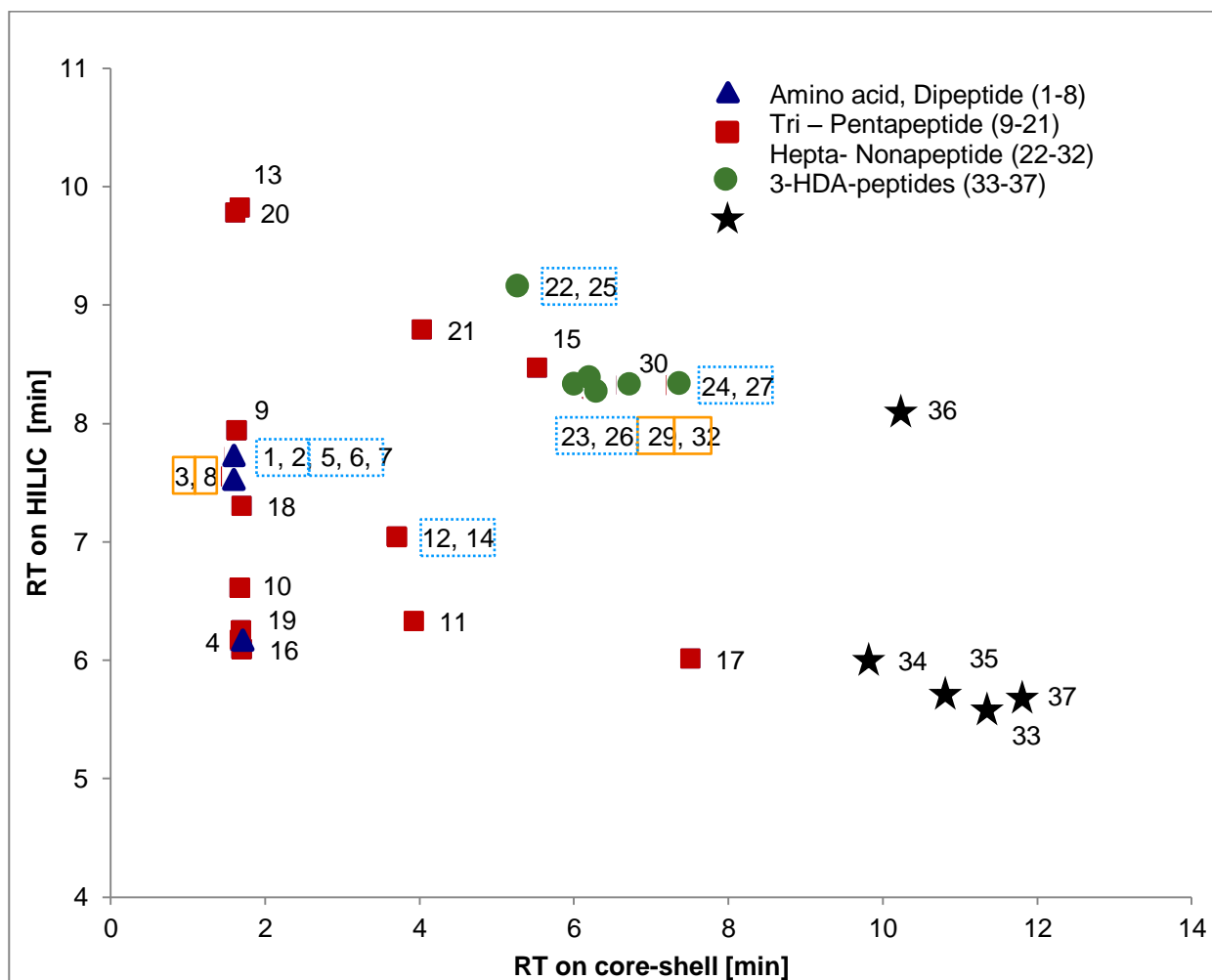
Figure S4. Orthogonality plots measuring complementarity of retention between RPLC (Kinetex C18, 2.6 µm) and HILIC (ZIC-HILIC, 3.5 µm). Amino acids or peptides which cannot be distinguished by fragmentation spectra are surrounded by a blue dotted line and could be each of these. Coeluting amino acids or peptides are surrounded by an orange solid line.

Amino acids and dipeptides: 1 I; 2 L; 3 LE; 4 LL; 5 LS; 6 SI; 7 SL; 8 TL; **tri- to pentapeptides:** 9 ETL; 10 SLL; 11 TLL; 12 ETLL; 13 ETLLS; 14 LETL; 15 LETLL; 16 LSLL; 17 LLSLL; 18 SLLS; 19 SLLSI; 20 TLLS; 21 TLLSL; **hepta- to nonapeptides:** 22 ETLLSL; 23 ETLLSLL; 24; ETLLSLLSI; 25 LETLLS; 26; LETLLSL; 27 LETLLSLLS;; 28 LLSLLSI; 29 LSLLSI; 30 TLLSLL; 31 TLLSLLS; 32 TLLSLLSI; **3-HDA-peptides:** 33 3-HDA-LETLL; 34 3-HDA-LETLLS; 35 3-HDA-LETLLSL; 36 3-HDA-LETLLSLL; 37 3-HDA-LETLLSLLSI.

Peptide Analysis: Zwitterionic Chiral Ion-Exchangers as Complementary Option to HILIC and to Reversed-Phase Chromatography

Photo Credit: Daniel Diaz Santana / EyeEm/Getty Images

Tong Zhang[1], Emilie Holder[1], Pilar Franco[1], Michael Lämmerhofer[2], Adrian Sievers-Engler[2], Heike Gerhardt[2], Harald Gross[2], and Wolfgang Lindner[3], [1]Chiral Technologies Europe, Illkirch, France,[2]Institute of Pharmaceutical Sciences, University of Tübingen, Tübingen, Germany, [3]Lindner Consulting GmbH, Klosterneuburg, Austria.

Therapeutic peptides represent one of the fastest growing segments in the pharmaceutical market. To bring these products to the market in a consistent manner, high quality is a major concern and requires stringent quality control (QC) methods. The quality of peptide therapeutics cannot be comprehensively tested by a single method. Hence, a number of tools are necessary to meet the goals of QC in therapeutic peptides. This article discusses the potential of zwitterionic chiral ion-exchangers to support peptide analysis and quality control as a flexible complementary tool to monitor the stereochemical integrity and chemical modifications.

Peptides occupy important positions in therapeutic research and in healthcare, food, and cosmetic industries, as well as in many other fields. Appropriate analytical methods are required to check the purities in peptide synthesis and meet the regulatory challenges in R&D, manufacturing, and QC of therapeutic peptides.

In practice, multiple analytical aspects have to be considered, such as the chemical and enantiomeric/stereoisomeric purities of starting amino acid and small peptide building blocks, the monitoring and control of stereochemistry and impurities of intermediate molecules from synthetic process, as well as the assessment of the structural integrity (amino acid composition, sequence, chirality) of final peptide products (1–3).

To determine the enantiomeric or stereoisomeric purities of raw materials or for the assignment of the absolute configuration of constituent entities of peptides or other related biopharmaceutical preparations, pre-column derivatization of the hydrolysates, followed by gas chromatography–mass spectrometry (GC–MS) or liquid chromatography (LC) coupled to MS analysis using capillaries or particulate packed columns has been the standard practice for decades.

Nevertheless, direct chiral analysis of amino acid or peptide fragments with no pre-column derivatization represents a more convenient and straightforward approach in terms of sample treatment, method simplicity, and speed of analysis cycle. High performance liquid chromatography (HPLC), often combined to MS detection, is undoubtedly the most attractive technique for this purpose.

A series of chiral selectors or stationary phases have been proven to be valuable for direct enantiomer or stereoisomer analysis of amino acids and small peptides by chromatography. The most important chromatographic supports to be cited in this regard include the chiral ligand exchangers (4–6), chiral crown ethers (7–11), macrocyclic antibiotics (12–15), and zwitterionic chiral ion-exchangers (16–17). These

chiral supports run analysis by chromatography on the basis of various enantio-selective or stereo-selective retention mechanisms.

Cinchona alkaloid-derived zwitterionic chiral ion-exchangers with cyclohexane sulphonic acid moiety (Figure 1[a]) have proved to be effective for the direct analysis of enantiomers and stereoisomers of a wide range of ampholytes such as amino acids (18–21) and peptides (22–24,37). The chiral-recognition ability of these zwitterionic chiral ion-exchanger columns to a great variety of amino acids (18–21, 24–32) and other racemic ionic compounds (33–35) has been investigated and demonstrated.

Figure 1: (a) Structures of zwitterionic chiral ion-exchangers (ZWIX phases) and (b) tentative binding model for zwitterionic analytes exemplified by Gly-Asp (cyan) (zwitterionic chiral ion-exchanger backbone in yellow). Adapted and reproduced with permission from *Chirality* **28**, F. Ianni, R. Sardella, A. Carotti, B. Natalini, W. Lindner, and M. Lämmerhofer, *Quinine-Based Zwitterionic Chiral Stationary Phase as a Complementary Tool for Peptide Analysis: Mobile Phase Effects on Enantio- and Stereoselectivity of Underivatized Oligopeptides*, 5–16 (2016) © John Wiley and Sons.
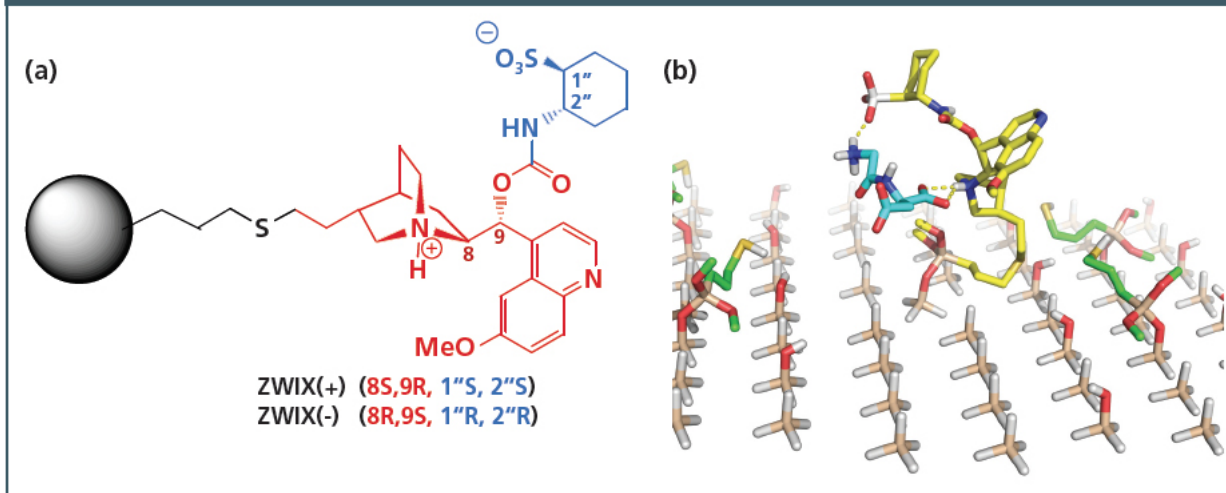
[9]

Mechanistically, three distinct application modes can be distinguished: (i) For chiral acidic analytes the zwitterionic chiral ion-exchanger CSPs (chiral stationary phases) act like an enantioselective anion-exchanger. Thereby, the zwitterionic chiral ion-exchanger columns are typically operated under weakly acidic (polar organic or hydro-organic) mobile phase conditions. Under such conditions the quinuclidine moiety is positively charged and attracts electrostatically acidic analytes (anion-exchange mode). H-bond mediated ionic interaction combined with additional H-bonds (at carbamate moiety), π-π-interactions (with quinoline), dipole–dipole, and van der Waals forces or steric interactions support enantioselective complexation of a wide variety of chiral acids covering essentially the same application spectrum as the corresponding chiral anion-exchangers (with tert-butyl moiety instead of the cyclohexane sulphonic acid residue at the carbamate group). Moreover, the sulphonic acid moiety of the chiral selector plays the role of an intramolecular counterion, which is present in equimolar concentrations with respect to the fixed ion-exchanger moiety. It leads to faster elution and allows the counterion concentration in the mobile phase to be reduced significantly, which is favourable for electrospray ionization (ESI) in LC–MS(–MS) applications; (ii) Vice versa, chiral basic compounds primarily interact at the chiral sulphonic acid moiety following the principles of cation-exchange (cation-exchange mode). In this case, the positively charged quinuclidinium represents the intramolecular counterion, which leads to accelerated separations compared to corresponding chiral cation-exchangers; (iii) For zwitterionic solutes, double ion-pairing at both cation- and anion-exchange sites may occur simultaneously (zwitterionic ion-exchange mode), giving zwitterionic-type CSPs their unique character (see Figure 1[b]). Depending on which of the two domains provides higher affinity the analyte will primarily bind to this site with high strength while the other site may contribute by long-range electrostatic interactions. In any case, increased ionic strength (that is, higher acid/base additive concentrations) of the eluent leads to ionic shielding of charged groups and stronger analyte displacement from the ion-exchanger site, respectively, consequently decreasing retention.

In the current study, their performance in direct stereo-selective separation of small peptides is further explored using LC–MS-compatible mobile phases. The second part of the report focuses on a case study where the zwitterionic chiral ion-exchanger columns investigated are deeply involved in the elucidation of the structural constitution and the stereochemistry of a lipopeptide, as well as to separate the peptide fragments generated from the linearized and the digested lipopeptide. In this case, the application of the zwitterionic chiral ion-exchanger columns is extended from analysis of enantiomers and stereoisomers of the hydrolysate amino acids, over determination of absolute configuration of the hydrolyzed fatty acid side chain, to the orthogonal properties of the zwitterionic chiral selectors with regard to other chromatographic modes such as hydrophilic interaction liquid chromatography (HILIC) and reversed-phase chromatography.

Experimental

Chemicals: Mobile phases for chromatography were prepared from HPLC-grade solvents. Methanol, acetonitrile, tetrahydrofuran (THF), and water were purchased from Carlo Erba Reagents. Formic acid (FA), diethylamine (DEA), ammonium hydroxide (NH4OH, 27–28%), and di- and tri-peptide samples were supplied by Sigma-Aldrich Chimie S.a.r.l. The compressed nitrogen (5.0, Messer France SAS) was used as the nebulizing gas for the evaporative light scattering detector (ELSD).

For Enantio- and Stereo-Selective Separation of the Peptides: The HPLC system used was an Agilent 1100 series apparatus optimized in terms of system void by using micro-flow cell (1.7 μL) and flow capillaries of 0.12-mm i.d. An evaporative light scattering detector (ELSD 2000ES, Alltech) was hyphenated to a diode-array detector (Agilent) via an interface 35900E. The generic parameters of ELSD were: gas flow, 1.7 L/min; drift tube temperature, 70 °C; gain, 1; impactor, OFF.

The zwitterionic chiral ion-exchanger columns used for the study were 150 × 3 mm, 3-μm Chiralpak ZWIX(+) (based on trans-(1''S,2''S)-N-[[[(8S,9R)-6'-methoxycinchonan-9-yl]oxy]carbonyl]-2''-aminocyclohexanesulphonic acid; quinine-derived) and 150 × 3 mm, 3-μm Chiralpak ZWIX(-) (based on trans-(1''R,2''R)-N-[[[(8R,9S)-6'-methoxycinchonan-9-yl]oxy]carbonyl]-2''-aminocyclohexanesulphonic acid; quinidine-derived) (Chiral Technologies Europe) (Figure 1[a]). The chemistries of these zwitterionic chiral ion-exchange columns will be referrred to genericaly as ZWIX(+) and ZWIX(–)] in this article. (The typical mobile phases used in the study were 50 mM FA + 25 mM DEA as additives in methanol/acetonitrile/H2O 49:49:2 v/v/v (MP-I), methanol/THF/H2O 49:49:2 v/v/v (MP-II), methanol/H2O 98:2 v/v (MP-III), and methanol/H2O 90:10 v/v (MP-IV). The flow rate was set at 0.5 mL/min.

For Lipopeptide Analysis: Absolute configurations of amino acid constituents of the lipopeptide (after hydrolysis by 6N DCl in D2O for 24 h at 110 °C and evaporation) were assigned by enantioselective HPLC with a 150 × 4 mm, 3-μm ZWIX(+) column, methanol/H2O (98:2 v/v) containing 9.4 mM ammonium formate and 9.4 mM formic acid as mobile phase, flow rate of 0.7 mL/min and ESI with quadrupole time-of-flight (QTOF)-MS detection. For absolute configuration determination of the 3-hydroxy fatty acid side chain, this lipid was extracted by liquid–liquid extraction (LLE) with chloroform–water (1:1 v/v). The chloroform layer was evaporated and analyzed by 150 × 4 mm, 3-μm ZWIX(+) and 150 × 4 mm, 3-μm ZWIX(-) columns at 10 °C column temperature, acetonitrile/methanol/acetic acid (95/5/0.025 v/v/v) as mobile phase, flow rate of 0.3 mL/min, and MS detection.

Separations of cyclic and linear lipopeptide forms were performed using i) reversed-phase LC using octadecylsilica with a 100 × 2.1 mm, 5-μm ODS-Hypersil (Agilent) column using water/acetonitrile (40:60 v/v) with 0.1% (v/v) formic acid, ii) HILIC on a sulphobetaine stationary phase with a 150 × 2 mm, 3.5-μm SeQuant ZIC-HILIC (Merck Millipore) column using water/acetonitrile (5:95 v/v) with 0.1% (v/v) formic acid, and iii) a 150 × 4 mm, 3-μm ZWIX(+) column with water/acetonitrile (65:35 v/v) containing 0.1% (v/v) formic acid. ESI-QTOF-MS was used as detector.

Complementarity plots were generated on chemical digests of the lipopeptide using reversed-phase LC–MS on a 100 × 3 mm, 2.6-μm core–shell octadecylsilica Kinetex C18 column (Phenomenex) with gradient elution using water (A) and acetonitrile (B) with 0.1% (v/v) formic acid and the following gradient profile: 20% B from 0–2.5 min, 20–64% B from 2.5–12.5 min, 64% B from 12.5–15 min, 64–80% B from 15–17 min, 80% B from 17–20 min, and 20% B from 20.1–23 min. Furthermore, the HILIC method made use of the sulphoalkylbetaine ZIC-HILIC column and the same eluents but inverted gradient: 80% B from 0–2.5 min, 80–36% B from 2.5–12.5 min, 36% B from 12.5–15 min, 36–20% B from 15–17 min, 20% B from 17–20 min, and 80% B from 20.1–23 min. Analysis for complementarity assessment in polar organic mode with the ZWIX(+) column was conducted in isocratic mode with methanol/acetonitrile/water (49:49:2 v/v/v) containing 25 mM ammonium formate and 25 mM formic acid.

Detection in LC–MS experiments was performed on a TripleTOF 5600+ (Sciex) QTOF MS instrument coupled via Duospray Ion Source (Sciex) and operated in ESI mode (curtain gas 30 psi, nebulizer and drying gas 60 psi, source temperature 400 °C) to a 1290 series UHPLC pump (Agilent) and column thermostat equipped with a CTC-PAL HTS autosampler (CTC Analytics). For amino acid analysis (free and derivatized with Sanger's reagent) as well as peptide separations to elucidate complementarity profiles on a ZWIX(+) column in ESI(-) mode was used with an ion-source floating voltage of -4500 V, and declustering potential -100 V. Acquisition was performed in product ion-high sensitivity mode with -20 V collision energy for enhanced sensitivity and to generate confirmation fragment spectra. Peptide separations by reversed-phase LC and HILIC were performed in positive ESI(+) mode, ion-source floating voltage 5500 V, and declustering potential 100 V. Acquisition was performed as scheduled targeted product ion scans with collision energy of 25 V for generating confirmation fragment spectra.

Further details on experimental conditions can be found in reference (36).

Results and Discussion

Two zwitterionic chiral ion-exchange CSPs were investigated for this study. Chemically, they consist of cinchona alkaloid-derived quinine carbamate with sulphonic acid moiety (S,S)-trans-2-aminocyclohexanesulphonic acid, ZWIX(+), and cinchona alkaloid-derived quinidine carbamate with (R,R)-trans-2-aminocyclohexanesulphonic acid, ZWIX(-), (18) (see Figure 1[a]). Stereochemically, these two chiral selectors are diastereoisomers but behave most frequently as pseudo-enantiomers. Such an idiosyncrasy affords the possibility and convenience of reversing elution order of enantiomers by switching the ZWIX column from one to the other (24). The chiral recognition and stereo-selective separation of ampholytic analytes such as free amino acids and peptides is primarily based on the synergistic double ion-pairing between the zwitterionic chiral ion-exchange selector and the zwitterionic analytes (zwitterionic ion-exchange mode) (Figure 1[b]) and assisted by other weaker interactions such as hydrogen bonding, π-π stacking, and van der Waals forces, as outlined above. On account of the zwitterionic ion-exchange mechanism, the co-presence of acidic and basic additives in an appropriate ratio is necessary to regulate the ionic interactions via displacement effects (16–17). The methanol-based mobile phase is recommended as the first choice for separations on the zwitterionic chiral ion-exchangers. The most useful mobile phase modifiers include water, acetonitrile, or tetrahydrofuran (THF). Among these solvents, methanol and water offer the most suitable solvation ability to all the ionized species involved in the ion-exchange equilibria. Exhibiting the strongest elution power, water is normally used at a low percentage (≤20% in volume). In contrast, the aprotic solvents acetonitrile and THF are weak eluting mobile phase components and are usually used up to 50%. Their presence in the mobile phase can efficiently contribute to the retention adjustment of fast-eluting analytes (21).

Stereo-Selective Separation of Di- and Tri-peptides: The experimental scheme or approach is essentially the same for amino acids and peptides on the zwitterionic chiral ion-exchanger columns investigated. For both a zwitterionic ion pairing/ion-exchange mode drives the retention and separation. Both the length of peptide, that is, the number of amino acid residues in the peptide chain, and the side chains modulate enantioselectivity. The first trials of separating enantiomers and stereoisomers of small

peptides were attempted with easily accessible commercial small peptide standard samples. In total, 13 common di-peptides and six typical tri-peptides are involved in the first part of this study.

Table 1: Separation of enantiomers of di- or tri-peptides containing one stereogenic centre.

Compound		t_1	t_2		k_1		α		Rs		MP
βAla-DL-Leu		8.2	11.2		3.6		1.47		7.1		I
Gly-DL-Met		13.8	15.0		6.7		1.11		1.7		I
Gly-DL-Phe		12.2	13.3		5.8		1.10		1.8		II
Gly-DL-Ser		5.1	7.4		1.9		1.66		8.2		III
Gly-DL-Trp		5.0	8.0		1.8		1.91		9.1		IV
Gly-DL-Val		6.5	7.7		2.7		1.26		2.4		I
DL-Leu-Gly		7.6	10.5		3.3		1.49		7.1		II
Gly-Gly-DL-Phe		4.7	5.5		1.6		1.25		2.9		III
Gly-Gly-DL-Val		6.3	7.4		2.5		1.25		3.0		I
DL-Leu-Gly-Gly		4.3	4.8		1.4		1.21		2.4		III

Mobile phase: (50 mM formic acid + 25 mM diethylamine) in acetonitrile/methanol/H_2O 49:49:2 (MP-I); THF/methanol/H_2O 49:49:2 (II); methanol/H_2O 98:2 (III); methanol/H_2O 90:10 (IV).
Column: 150 × 3 mm ZWIX(+), Flow rate: 0.5 mL/min, Temperature: 25 °C, Detection: ELSD.

[10]

The experimental results are summarized in Table 1 for enantiomers of di- or tri-peptides bearing a single stereogenic centre. As indicated by the selectivity data (□), each pair of enantiomers could be recognized on the ZWIX(+) column, with the highest resolution degree for Gly-dl-Trp and Gly-dl-Ser. It was observed that, even with extensive optimization of the chromatographic conditions, the achievement of full enantiomer resolution of dl-Ser proved to be challenging on these CSPs. With a glycine component attached to the N-terminus of serine (Gly-dl-Ser), however, a large and effortless separation of the enantiomers could be obtained. It should be noted that this does not represent a generic scenario and the enantio- or stereo-selectivity is highly structure-dependent.

When it comes to the peptides containing more than one stereogenic centre, the peak configuration in a single analysis becomes more complex. In this case, it would be of high importance to monitor the peak elution order while optimizing the enantio- and stereo-selectivities as well as the resolution degrees between the peaks. Unfortunately no information on elution order could be collected in our study because of the unavailability of enantiomer or diastereoisomer standards.

Again, the challenging configuration of multi-peaks in a single and short chromatographic analysis would require enhanced performance of the column. For the given chiral selector bonded on to the spherical silica matrix with well-defined particle size, pore, and surface properties, the most viable options would be the variation in composition of bulk mobile phase, the reduction of column temperature, and the use of a longer column. While the first two approaches could be adopted in attempts to enlarge the selectivities (therefore the resolution degree), the third alternative would normally be more forthright by offering higher column efficiency without changing the capacity factors.

Table 2: Separation of stereoisomers of di- or tri-peptides containing two stereogenic centres.

Compound	t_1	t_2	t_3	t_4	k_1	$\alpha_{(1-2)}$	$\alpha_{(2-3)}$	$\alpha_{(3-4)}$	$Rs_{(1-2)}$	$Rs_{(2-3)}$	$Rs_{(3-4)}$	MP	T(°C)
DL-Ala-DL-Ala	8.4	9.4	9.8	10.5	3.7	1.15	1.05	1.09	2.1	0.7	1.3	II	10
DL-Ala-DL-Leu	6.8	7.8	8.1	8.9	2.8	1.21	1.06	1.12	3.0	0.8	1.8	II	10
DL-Ala-DL-Ser	6.9	7.5	8.5	14.5	2.9	1.13	1.16	1.90	1.8	1.9	8.4	I	25
DL-Leu-DL-Leu	5.0	5.0	5.7	5.7	1.8	1.00	1.22	1.00	0.0	3.2	0.0	II	25
DL-Leu-DL-Tyr	4.8	5.5	6.6	13.4	1.6	1.24	1.30	2.43	3.1	3.9	14.3	I	10
DL-Leu-DL-Val	4.9	5.6	6.3	8.7	1.7	1.25	1.17	1.53	3.3	2.3	6.8	II	25
DL-Ala-DL-Leu-Gly	11.8	12.9	15.1	23.1	5.6	1.10	1.20	1.60	1.6	3.1	7.5	II	10
Gly-DL-Leu-DL-Ala	3.7	4.1	4.4	4.9	1.0	1.20	1.11	1.20	2.0	1.3	2.3	III	10
DL-Leu-Gly-DL-Phe	3.8	3.8	4.3	4.7	1.1	1.00	1.25	1.13	0.0	2.5	1.5	III	10

Mobile phase: (50 mM formic acid + 25 mM diethylamine) in acetonitrile/methanol/H$_2$O 49:49:2 (MP-I); THF/methanol/H$_2$O 49:49:2 (II); methanol/H$_2$O 98:2 (III). Column: 150 × 3 mm ZWIX(+), Flow rate: 0.5 mL/min, Temperature: 25 °C, Detection: ELSD.

[11]

For separation of diastereoisomers of the di- or tri-peptides containing two stereogenic centres (see Table 2), most of the data were acquired at 10 °C and represented an effective improvement in resolution degree of the adjacent peaks in regard to the conventional temperature at 25 °C. The effect of mobile phase on the resolution of dl-Leu-dl-Val is shown in Figure 2(e). Peak co-elution occurred while methanol/acetonitrile/H2O 49:49:2 (MP-I in [e-1]) or methanol/H2O 98:2 (MP-III in [e-3]) was in use as the bulk mobile phase. A full separation of the four stereoisomers could be achieved by simple replacement of acetonitrile in MP-I with THF (MP-II in [e-2]). The performance of the zwitterionic chiral ion-exchanger columns investigated can be further demonstrated with successful stereoisomeric separations under optimized conditions. For instance, the combination of low temperature and longer columns led to satisfactory separation of four stereoisomers for (a) dl-Leu-dl-Tyr, (b) dl-Ala-dl-Leu, and (d) Gly-dl-Leu-dl-Ala (Figure 2). As far as dl-Ala-dl-Leu-Gly is concerned, the complete resolution of the four stereoisomers could hardly be achieved on the ZWIX(+) column even with extensive optimization of the chromatographic conditions.

Figure 2: Separation of stereoisomers of di- and tri-peptides. Flow rate: 0.5 mL/min; Detection: ELSD. Temperature: 25 °C for (b) and (e); 10°C for (a) and (c)-(d). Mobile phase: MP-I for (a) and (e-1); MP-II for (b), (d), and (e-3); MP-III for (c) and (e-2).

(a) DL-Leu-DL-Tyr ZWIX(+)150x3mm

(b) DL-Ala-DL-Leu ZWIX(+)250x4mm

(c) DL-Ala-DL-Leu-Gly ZWIX(–)150x3mm

(d) Gly-DL-Leu-DL-Ala ZWIX(+)250x34mm

(e) DL-Leu-DL-Val ZWIX(+)150x3mm

(e-1)

(e-2)

(e-3)
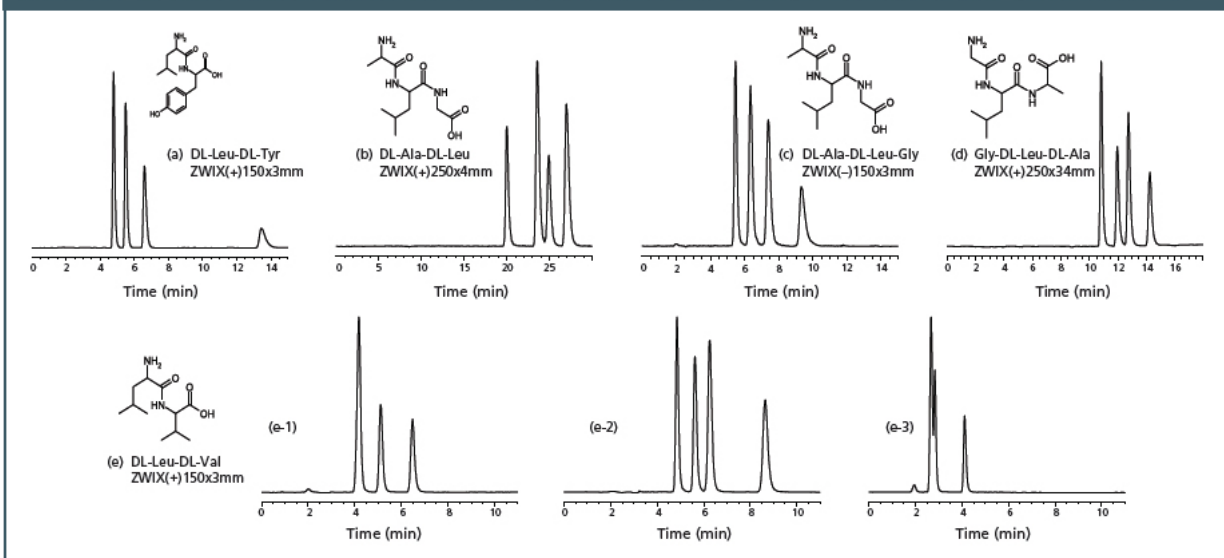
[12]

However, complete resolutions could be obtained with a ZWIX(-) column of the same column size (Figure 2[c]). This specifies the complementarity properties between ZWIX(+) and ZWIX(-) columns in terms of stereoselective recognition performance and resolution power.

Lipopeptide Analysis: The great utility of the cinchonan-based zwitterionic chiral ion-exchanger columns investigated as complementary tools for peptide separation and characterization in quality control and drug discovery, taking benefit from its distinct selectivity profiles ranging from free and derivatized amino acid enantiomers, peptide enantiomer, epimers, diastereomers, and (minor) chemical modifications in peptides and peptide therapeutics, is illustrated in the following practical example from pharmaceutical biology (36). In this study ZWIX(+) was used as a basic tool for comprehensive structural elucidation of a lipopeptide (Figure 3[a]) isolated from the endophytic Pseudomonas poae strain RE*1-1-14, being supported by reversed-phase LC and HILIC. The cyclic lipopeptide poaeamide A constituted by 10 amino acids showed some bioactivity in terms of growth inhibition of the fungal pathogen Rhizoctonia solani. For comprehensive structural elucidation the lipopeptide was analyzed on three structural levels: in intact form, in fully hydrolyzed form, and at intermediate level in digested form (obtained by enzymatic and/or chemical digestion).

Figure 3: (a) Chemical structure of poaeamide A with annotated stereochemistry as identified by complementary analysis including ZWIX(+). (b) HPLC–ESI-QTOF-MS analysis of linearized peptide (poaeamide hydrolyzed with 1N NaOH) by reversed phase HPLC on ODS-Hypersil 5 μm, (c) by HILIC on ZIC-HILIC 3.5 μm, and (d) HPLC in polar organic mode on ZWIX(+). C, cyclic poaeamide; L, linear poaeamide; * indicates side product of hydrolysis. Adapted and reproduced with permission from *Journal of Chromatography A* **1428**, H. Gerhardt, A. Sievers-Engler, G. Jahanshah, Z. Pataj, F. Ianni, H. Gross, W. Lindner, and M. Lämmerhofer, *Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides*, 280–291 (2016) © Elsevier.
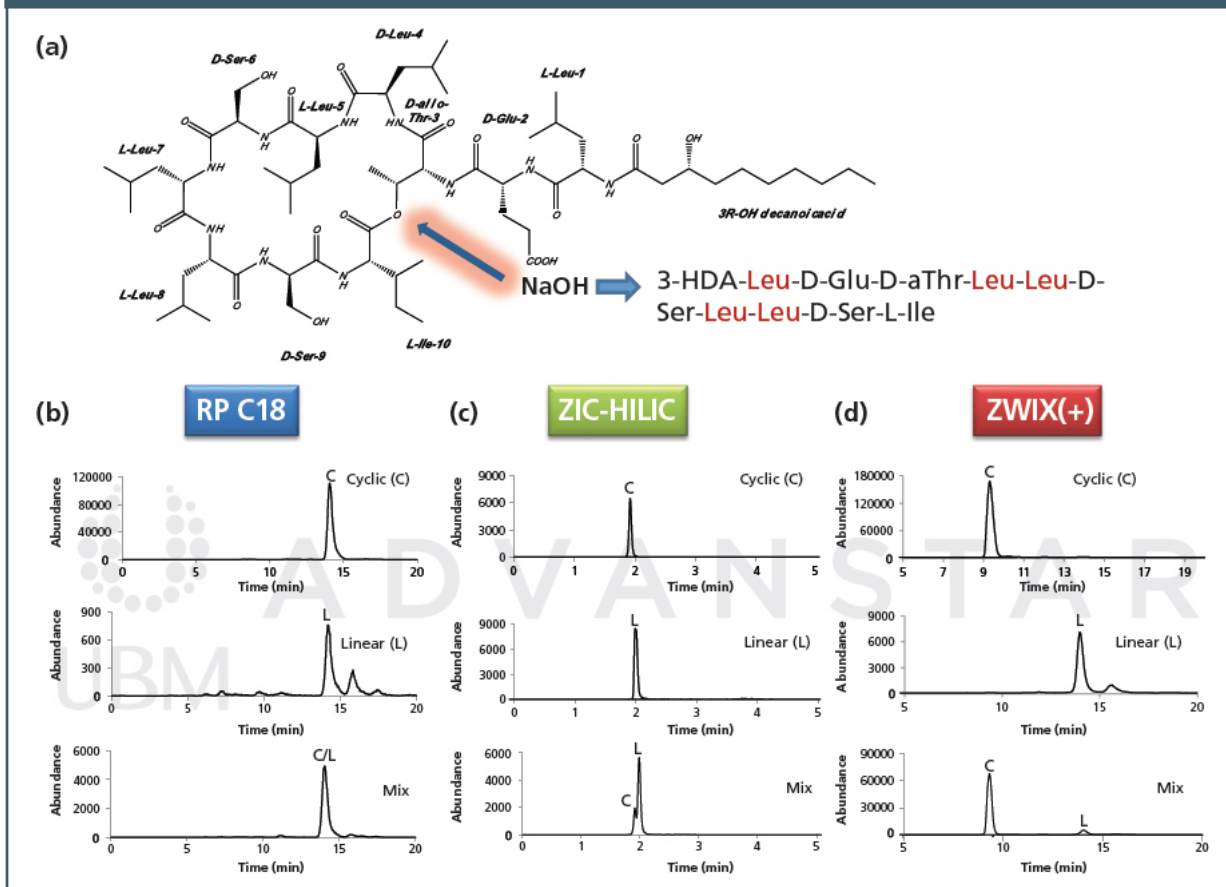
[13]

Amino acid sequence and constitution of fatty acid side chain were readily determined on the intact peptide by nuclear magnetic resonance (NMR) and MS after purification by reversed-phase LC. However, several structural features related to stereochemistry required careful elucidation by stereoselective methods. In a first step, the peptide was fully hydrolyzed by DCl/D2O to obtain free amino acid and 3-hydroxydecanoic acid constituents.

Absolute configurations of amino acids were determined by enantioselective chromatography on ZWIX(+) with mobile phases as specified in the experimental section (36). A hydrolyzed lipopeptide sample was analyzed without derivatization (free amino acids) in zwitterionic ion-exchange mode and after derivatization with Sanger's reagent, which yielded 2,4-dinitrophenyl (DNP) derivatives of amino acids in anion-exchange mode. This was necessary because the current lipopeptide contained all the most challenging amino acids from a stereoisomer separation point of view (as a result of side chain chirality) (Ile/Leu/a-Ile as well as Thr/a-Thr). Chromatographic results for amino acids present in the lipopeptide sample are summarized in Table 3. Comparison of retention factors of amino acid standards and amino acids from the peptide sample revealed that the lipopeptide was constituted by d-Ser (determined as DNP-derivative), d-Glu (free, DNP), d-a-Thr (DNP), l-Leu (DNP), d-Leu (free), and l-Ile (free). Several options for verification and validation of the results exist: Analysis on ZWIX(-) is first choice because elution orders of enantiomers are reversed. Furthermore, other N-derivatives may be used and give distinct elution orders, retention factors, and separation factors allowing the unequivocal identification of the stereochemistry (34), or analysis of N-derivatized forms on tert-butylcarbamoylquinine or quinidine-based chiral anion exchangers might be considered as an appropriate choice for confirmation of results. In our example, we have also used GC–MS with a polysiloxane modified with N-(2-methylpropanoyl)-valyl-tert-butylamide column as the complementary analysis method, revealing that the stereochemistry was correctly identified (36).

Table 3: Liquid chromatographic enantiomer separation data for amino acids present in lipopeptide on ZWIX(+) in free underivatized form and as N-2,4-dinitrophenyl-derivatives.

Amino Acid	Free Amino Acids[a]					DNP-Derivatized Amino Acid[b]				
	D Amino Acid		L Amino Acid			D Amino Acid		L Amino Acid		
Compound	RT (min)	k	RT (min)	k	α	RT (min)	k	RT (min)	k	α
Serine (Ser)	5.33	1.54	5.33	1.54	1.0	8.72	3.15	6.26	1.98	1.6
Glutamic acid (Glu)	7.42	2.53	8.20	2.90	1.1	8.34	2.97	7.54	2.59	1.1
Leucine (Leu)	4.38	1.09	3.81	0.81	1.3	4.74	1.26	4.13	0.97	1.3
Isoleucine (Ile)	4.26	1.03	3.65	0.74	1.4	4.81	1.29	3.90	0.86	1.5
Allo-isoleucine (a-Ile)	4.32	1.06	3.70	0.76	1.4	4.88	1.32	3.80	0.81	1.6
Threonine (Thr)	5.25	1.50	4.56	1.17	1.3	6.97	2.32	4.77	1.27	1.8
Allo-threonine (a-Thr)	4.19	1.00	4.34	1.07	1.1	6.21	1.96	4.93	1.35	1.5

[a] elution order for underivatized amino acids: L before D, except for Glu and a-Thr (D < L)

[b] elution order for DNP-derivatives of amino acids: always L before D

[14]

A significant advantage of the zwitterionic chiral ion-exchanger columns investigated in this study was their broad applicability profile for various classes of compounds, ranging from free amino acids and derivatized amino acids over peptide stereoisomers, to chiral carboxylic acids. It is a peculiar feature of Pseudomonas lipopeptides that they contain a fatty acid side chain with at least one stereogenic centre. In the investigated lipopeptide it was identified as 3-hydroxydecanoic acid (36). Prior studies with racemic mixture showed that the enantiomers of 3-hydroxydecanoic acids can be resolved on ZWIX(+) in the anion-exchange mode when a mobile phase consisting of acetonitrile/methanol/acetic acid (95/5/0.025; v/v/v) was used and column temperature was 10 °C (36). Retention times on ZWIX(+) were 13.48 min for the first eluted enantiomer and 14.67 min for second eluted enantiomer, as well as 15.46 and 16.76 min on ZWIX(-). An R-enantiomer standard of 3-hydroxydecanoic acid isolated from rhamnolipid allowed the elution order to be pinpointed (R before S on ZWIX[+] and S before R on ZWIX[-]). Consequently, this enantioselective HPLC method was a suitable means to identify the stereochemistry in the fatty acid side chain that normally remains undetermined in research on lipopeptides. The chloroform extract was injected with and without a spike of racemic mixture, and it turned out that the absolute configuration of 3-hydroxydecanoic acid in poaeamide has R-configuration (36).

Absolute configurations of all constituents were determined at this point, however, the full structure was not yet identified. Positions of ring closure and the location of d-Leu in the peptide sequence were still unclear. A carboxylic acid derivatization procedure (using EDC and an amine label) gave rise to the conclusion that the C-terminus must be involved in the ring closure because the carboxylic acid moiety of the d-Glu residue was found by MS to be chemically modified by this derivatization step. On the other hand, a side product in the alkaline hydrolysis of the cyclic ester indicated that the a-Thr residue represented the alcoholic component for ester formation and ring closure, respectively (36).

A bit more challenging was to pinpoint the position of the d-Leu residue in the peptide sequence. One out of five Leu residues had d-configuration. Unfortunately, sequence information is lost as a result of hydrolysis of the lipopeptide sample prior to amino acid configuration analysis. Therefore, the determination of the position of the d-Leu amino acid residue in the peptide chain needs other strategies to fix this problem. If authentic standards of each possible stereoisomer were available, their analysis on the investigated zwitterionic chiral ion-exchanger columns had a great opportunity to reveal distinct retention times as a result of the good diastereomer selectivity of this column, as proven with several peptides. Unfortunately, authentic standards of each possible stereoisomer of this lipopeptide comprised of non-natural amino acids with lipid modification were too expensive and therefore a strategy to digest the lipopeptide (enzymatic and chemical digestion) to produce smaller fragments followed by subsequent hydrolysis of isolated fractions and enantioselective analysis at the amino acid level was envisioned to solve the stereochemistry of poaeamide.

For enzymatic digestion with pepsin the cyclic lipopeptide was first hydrolyzed with 1N NaOH to ensure better digestion efficiency of the enzyme (36). Control of the reaction product of hydrolysis revealed another interesting feature of the zwitterionic chiral ion-exchanger columns studied. It appears to be sensitive for minor molecular changes (small structural modifications) in peptides, which are accompanied with changes in ionization state and effective charge, respectively. While the intact cyclic lipopeptide carries a single negatively charged functionality (in the form of a Glu side chain close to the N-terminal end) featuring an anion-exchange mode on zwitterionic chiral ion-exchanger column, it becomes two-fold negatively charged after hydrolysis in the linear form. Since the Gibbs free energy of binding is proportionally decreasing and the log k-values are directly proportionally increasing with the effective charge number of the ionic analyte, retention of the hydrolyzed linear form is significantly different in the prevailing anion-exchange mode on the zwitterionic chiral ion-exchanger column. It becomes evident from the chromatograms in Figure 3 that the zwitterionic chiral ion-exchanger column investigated exhibits better selectivity for distinction between the cyclic and linearized forms of the lipopeptide. While cyclic and linear forms essentially

co-eluted both in reversed-phase LC (Figure 3[b]) and HILIC (Figure 3[c]), the zwitterionic quinine carbamate selector selectively recognized the hydrolyzed form. It was more strongly retained by additional interaction of the carboxylic acid C-terminus, which was available in free form only in the linear lipopeptide (Figure 3[d]). Similar situations may often exist in therapeutic peptides and their impurities in which such chemical modifications may occur because of degradation. Zwitterionic chiral ion-exchanger columns might therefore be valuable complementary tools for impurity profiling studies of therapeutic peptides.

For locating d-Leu within the peptide chain, the lipopeptide was digested under controlled conditions to small peptide fragments that contain only one or a few Leu residue(s). The peptide fragments were then separated by reversed-phase LC and individual peptide fragments containing a Leu residue were isolated micro-preparatively. Hydrolysis and subsequent amino acid enantiomer analysis of chromatographically isolated single peptide fragments allowed the stereochemistry of the peptide to be put together like a puzzle (36).

Pepsin is an endopeptidase that cleaves amide bonds preferentially at the carboxylic side of Phe, Leu, and partly Glu. D-configuration would block enzyme action. The linearized peptide was digested with pepsin at 37 °C for 90 min and subsequently Leu-containing peptides isolated by reversed-phase LC. An early eluted peak corresponding to LSL, SLL, or LSI tripeptides contained only l-Leu-, but no d-Leu-configurated peptides. Only 3 peaks, all with the peptide fragment 3-hydroxydecanoyl-(3-HDA)-LETLL (HDA-LETLL, HDA-LETLLSL, HDA-LETLLSLL), contained d-Leu. This allowed the possibilities for d-Leu positions to be narrowed down to the two Leu residues from the amino terminal side. Controlled chemical digestion (hydrolysis with 6N HCl at 110 °C for 20 min) produced several smaller peptides. Some potentially diagnostic peptides in terms of Leu stereochemistry were again collected by reversed-phase LC. Their amino acid enantiomer analysis revealed that the Leu residue of 3-HDA-L and 3-HDA-LE had L-configuration. In contrast, d-Leu was detected in the isolated peptides LETL and TLL. Consequently, the d-Leu residue was confirmed to be located in residue 4 from the N-terminus of the peptide.

Figure 4: Orthogonality of zwitterionic chiral ion-exchanger columns. (a) reversed phase LC (Kinetex C18, 2.6 µm) vs. ZWIX(+), 3 µm, in polar organic mode, and (b) HILIC (ZIC-HILIC, 3.5 µm) vs. ZWIX(+), 3 µm, polar organic mode. **Amino acids and dipeptides:** 1 I; 2 L; 3 LE; 4 LL; 5 LS; 6 SI; 7 SL; 8 TL; **tri- to pentapeptides:** 9 ETL; 10 SLL; 11 TLL; 12 ETLL; 13 ETLLS; 14 LETL; 15 LETLL; 16 LSLL; 17 LLSLL; 18 SLLS; 19 SLLSI; 20 TLLS; 21 TLLSL; **hexa- to nonapeptides:** 22 ETLLSL; 23 ETLLSLL; 24; ETLLSLLSI; 25 LETLLS; 26; LETLLSL; 27 LETLLSLLS;; 28 LLSLLSI; 29 LSLLSI; 30 TLLSLL; 31 TLLSLLS; 32 TLLSLLSI; **3-HDA-peptides:** 33 3-HDA-LETLL; 34 3-HDA-LETLLS; 35 3-HAD LETLLSL; 36 3- HDA-LETLLSLL; 37 3-HDA-LETLLSLLSI. Adapted and reproduced with permission from *Journal of Chromatography A* **1428**, H. Gerhardt, A. Sievers-Engler, G. Jahanshah, Z. Pataj, F. Ianni, H. Gross, W. Lindner, and M. Lämmerhofer, *Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides*, 280–291 (2016) © Elsevier.
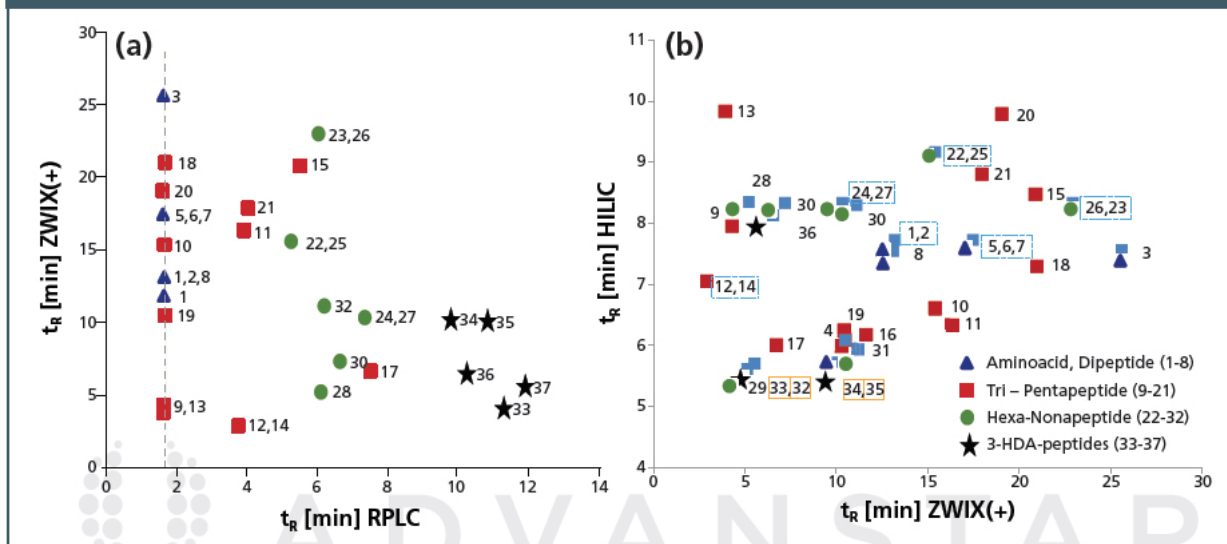
[15]

Reversed-phase LC and HILIC showed some shortcomings in the above peptide separations, in particular for small peptides. Figure 3 has already demonstrated that ZWIX columns can generate complementary retention profiles to reversed-phase LC and HILIC in peptide separations. To demonstrate this more systematically on a wider scope, chemically digested poaeamide was analyzed by reversed-phase LC (on octadecylsilica core–shell column), HILIC (on silica-based sulphobetaine column), and chiral HPLC in polar organic mode (on ZWIX[+]) (36). Orthogonality plots in Figure 4 convincingly document evenly distributed data points in the two-dimensional design space (tR of peptide fragments on ZWIX[+] vs. reversed phase in Figure 4[a] and silica-based sulphobetaine column vs. the ZWIX[+] column in Figure 4[b]) and thus prove excellent complementarity of these chromatographic modes. A number of small peptides are well retained and resolved on the zwitterionic chiral ion-exchanger columns but co-eluted close to t0 in reversed-phase LC (Figure 4[a], indicated by dotted line). On top of this complementary retention profile of the zwitterionic chiral ion-exchanger, it provides additional selectivity for peptide stereoisomers (enantiomers, epimers, and diastereomers), which either do not exist for achiral columns, may be greatly enhanced, or be complementary and confirmatory on the zwitterionic chiral ion-exchanger phase. Furthermore, degradation and side product formation in therapeutic peptides may come along with alterations in the charge state of the peptide; for example, as a result of side chain deamidation (Asn☐Asp, Gln☐Glu), C-terminal deamidation, C-term amidation (COOH☐CONH2), isomerization (Asp☐IsoAsp), and succinimide formation (Asn/Asp☐Suc), which may be readily resolved by a zwitterionic chiral ion-exchanger column while it may be more challenging to resolve such structural modifications with common reversed-phase LC. Zwitterionic chiral ion-exchangers may therefore be regarded as a valuable tool for peptide separations and complement state-of-the-art reversed-phase LC and HILIC. Furthermore, its integration as one orthogonal dimension in 2D LC peptide separations could be another application of great interest and future potential.

Conclusions

Therapeutic peptides are a growing segment in pharmaceutical markets. In order to ensure the quality of such products, a large number of critical quality attributes have to be tested and controlled during production and release. Besides chemical integrity, stereochemical integrity is of central importance and its analytical quality assurance testing is challenging and not a trivial task. A common strategy for stereochemical impurity testing of peptides is to fully hydrolyze the sample and perform enantioselective amino acid analysis. GC–MS on a polysiloxane modified with N-(2-methylpropanoyl)-valyl-tert-butylamide column is one of the established and well accepted protocols for this purpose. Yet, this strategy also has some shortcomings. There are some critical proteinogenic-type amino acids like Cys, His, Asn, and Gln. Asn/Gln are hydrolyzed in the side chain and cannot be distinguished from corresponding acids Asp/Glu if they co-exist in the peptide. Cys and His (partly also Arg) are problematic in GC–MS amino acid analysis. Unfortunately, cysteine, histidine, and aspartic acid were reported to be susceptible to racemization in peptide synthesis and important amino acids to be controlled.

Furthermore, if individual amino acids are present more than once and are found to be present in d- and l-configuration, the correct stereochemistry of the peptide, be it the therapeutic peptide or an impurity, cannot be determined on the amino acid level. Hence, alternative support methods are required. We suggest performing peptide analysis systematically on three levels of structural integrity, namely the intact molecule, its fully hydrolyzed form (amino acid level), and on the intermediate level (partially digested form). On all levels, three complementary LC methods — reversed-phase LC, HILIC, and enantioselective or stereoselective HPLC with a zwitterionic chiral ion-exchanger column — should be used for a comprehensive structural elucidation, ideally in combination with complementary detection such as UV and MS. This would provide the maximum amount of information, allowing the chemical and stereochemical structure of the analyzed peptides to be pieced together. Although the zwitterionic chiral ion-exchanger columns studied were initially designed for enantiomer separation of very diversely structured natural and unnatural □-amino acid but also for □-and □-amino acids, it represents a valuable tool in many other applications as well, in particular peptide separations. Zwitterionic chiral ion-exchanger columns can provide enantioselectivity, epimer selectivity, diastereomer selectivity for peptides, enhanced selectivity for chemical modifications that are accompanied by changes in charge state, and last but not least it can be utilized for raw materials quality control (usually FMOC amino acids but also free □-, □-, □-amino acids, and other amino acid derivatives).

Acknowledgements

References

1. I. Eggen, B. Gregg, H. Rode, A. Swietlow M. Verlander, and A. Szajek, *Pharmaceutical Technology* **38**(3), 27–4 (2014).
2. I. Eggen, B. Gregg, H. Rode, A. Swietlow M. Verlander, A. Szajek, *BioPharm International* **38**(4), 27–4 (2014).
3. B. Gregg, A. Swietlow, A.Y. Szajek, H. Rode, M. Verlander, and I. Eggen, *BioPharm International* **38**(5), (2014).
4. V.A. Davankov and S.V.Rogozhin, *J. Chromatogr.* **60**, 280–283 (1971).
5. M.H. Hyun, S.C. Han, C.W.Lee, and Y.K. Lee, *J. Chromatogr. A* **950**, 55–63 (2002).
6. B. Natalini, R. Sardella, A. Macchiarulo, and R. Pellicciari, *J. Chromatogr. B* **405**, 145–153 (2008).
7. T. Shinbo, T. Yamaguchi, K. Nishimura, and M. Sugiura, *J. Chromatogr.* **405**, 145–153 (1987).
8. A. Péter, G. Török, P. Csomós, M. Péter, G. Bernáth, and F. Fülöp, *J. Chromatogr. A* **761**, 103–113 (1997).
9. M.H. Hyun, J.S. Jin, and W. Lee, *J. Chromatogr. A* **822**, 155–161 (1998).
10. Y. Machida, H. Nishi, K. Nakamura, H. Nakai, and T. Sato, *J. Chromatogr. A* **805**, 85–92 (1998).
11. H.J. Choi and M.H. Hyun, *J. Liq. Chrom. Rel. Tech.* **30**, 853–875 (2007).
12. D.W. Armstrong, Y. Lui, and K.H. Ekborgott, *Chirality* **7**, 474–797 (1995).
13. A. Berthod, Y. Lui, C. Bagwill, and D.W. Armstrong, *J. Chromatogr. A* **731**, 123–137 (1996).
14. A. Péter, G. Török, and D.W. Armstrong, *J. Chromatogr. A* **793**, 283–296 (1998).
15. A. Péter, G. Török, D.W. Armstrong, G. Tóth, and D. Tourwé, *J. Chromatogr. A* **904**, 1–15 (2000).
16. C.V. Hoffmann, R. Pell, M. Lämmerhofer, and W. Lindner, *Anal. Chem.* **80**, 8780–8789 (2008).
17. C.V. Hoffmann, R. Reischl, N.M. Maier, M. Lämmerhofer, and W. Lindner, *J. Chromatogr. A* **1216**, 1157–1166 (2009).
18. T. Zhang, E. Holder, P. Franco, and W. Lindner, *J. Chromatogr. A* **1355**, 191–199 (2014).
19. M. Lämmerhofer, *Anal. Bioanal. Chem.* **406**, 6095–6103 (2014).
20. T. Fukushima, A. Sugiura. I. Furuta, S. Iwasa, H. Lizuka, H. Ichiba, M. Onozato, H. Hikawa, and Y. Yokoyam, *Int. J. Trp. Res.* **8**, 1–5 (2015).
21. I. Ilisz, Z. Gecse, Z. Pataj, F. Fülöp, C. Toth, W. Lindner, and A. Péter, *J. Chromatogr. A* **1363**, 169–177 (2014).
22. S. Wernisch and W. Lindner, *J. Chromatogr. A* **1269**, 279–307 (2012).
23. R.J. Reischl and W. Lindner, *J. Pharm. Biomed. Anal.* **116**, 123–130 (2015).
24. T. Zhang, E. Holder, P. Franco, and W. Lindner, *J. Sep. Sci.* **37**, 1237–1247 (2014).
25. I. Ilisz, N. Grecso, M. Palko, F. Fülöp, W. Lindner, and A. Péter, *J. Pharm. Biochem. Anal.* **98**, 130–139 (2014).
26. Z. Pataj,I. Ilisz,Z. Gecse, Z. Szakonyi, F. Fülöp, W. Lindner, and A. Péter, *J. Sep. Sci.* **37**, 1075–1082 (2014).
27. I. Ilisz, Z. Pataj, Z. Gecse, Z. Szakonyi, F. Fülöp, W. Lindner, and A. Péter, *Chirality* **26**, 385–393 (2014).
28. I. Ilisz, N. Grecso, A. Aranyi, P. Suchotin, D. Tymecka, B. Wilenska, A. Misicka, F. Fülöp, W. Lindner, and A. Péter, *J. Chromatogr. A* **1334**, 44–54 (2014).
29. I. Ilisz, N. Grecso, A. Misicka, D. Tymecka, L. Lazar, W. Lindner, and A. Péter, *Molecules* **20**, 70–87 (2015).
30. I. Ilisz, Z. Gecse, G. Lajko, M. Nonn, F. Fülöp, W. Lindner, and A. Péter, *J. Chromatogr. A* **1384**, 67–75 (2015).
31. I. Ilisz, Z. Gecse, G. Lajko, E. Forro, F. Fülöp, W. Lindner, and A. Péter, *Chirality* **27**, 563–570 (2015).
32. I. Ilisz, N. Grecso, R. Papousek, Z. Pataj, P. Bartak, L. Lazar, F. Fülöp, W. Lindner, and A. Péter, *Amino Acids* **47**, 2279–2291 (2015).
33. M.C.K. Geditz, W. Lindner, M. Lämmerhofer, G. Heinkele, R. Kerb, M. Ramharter, M. Schwab, and U. Hofmann, *J. Chromatogr. B* **968**, 32–39 (2014).
34. F. Ianni, Z. Pataj, H. Gross, R. Sardella, B. Natalini, W. Lindner, and M. Lämmerhofer, *J. Chromatogr. A* **1363**, 101–108 (2014).

35. F. Ianni, A. Carotti, M. Marinozzi, G. Marcelli, A. Di Michele, R. Sardella, W. Lindner, and B. Natalini, *Anal. Chim. Acta* **885**, 174–182 (2015).
36. H. Gerhardt, A. Sievers-Engler, G. Jahanshah, Z. Pataj, F. Ianni, H. Gross, W. Lindner, and M. Lämmerhofer, *J. Chromatogr. A* **1428**, 280–291 (2016).
37. F. Ianni, R. Sardella, A. Carotti, B. Natalini, W. Lindner, and M. Lämmerhofer, *Chirality* **28**, 5–16 (2016).

Tong Zhang received her Ph.D. degree in organic chemistry from University of Bordeaux I (France) in 1992. After three years postdoctoral work at Novartis (Switzerland) under the supervision of Eric Francotte, she joined Chiral Technologies Europe (Strasbourg, France). She is now the R&D Manager in the same company.

Emilie Holder graduated from the Institute of Technology of Strasbourg (France) in 2007. As an organic chemist she worked in the R&D team of Sunovion Pharmaceuticals Inc. (Marlborough, MA, USA) and then joined an organic chemistry laboratory at University of Strasbourg. She is now an R&D assistant at Chiral Technologies Europe in Strasbourg, France.

Pilar Franco received her Ph.D. degree in pharmacy from the University of Barcelona, Spain. After two years postdoctoral work in the University of Vienna, Austria, under the supervision of Wolfgang Lindner, she joined Chiral Technologies Europe (Strasbourg, France). She is now Manager of Technical Operations in the same company.

Michael Laemmerhofer is professor for pharmaceutical (bio-)analysis at the Institute of Pharmaceutical Sciences, University of Tübingen in Germany. His research interests include surface functionalization of micro- and nanoparticulate materials for separation and sample preparation, chiral separation, impurity profiling, bioanalysis (metabolomics and lipidomics), and biopharmaceuticals analysis.

Adrian Sievers-Engler is a PhD student in the pharmaceutical (bio-)analysis group at the University of Tübingen, Germany. He graduated in biochemistry in 2009 from the University of Greifswald, Germany. His research interests include high resolution mass spectrometry and separation techniques in bioanalysis and pharmaceutical bioanalysis.

Heike Gerhardt studied chemistry at the universities of Tübingen and Vienna. She finished her Ph.D. in May 2015 at the University of Tübingen under the supervision of Michael Lämmerhofer working on analysis of insect secretion by GC–MS. Nowadays, she works as COO at the company C.A.T. GmbH & Co KG in Tübingen, Germany.

Harald Gross is professor for pharmaceutical biology at the Pharmaceutical Institute, University of Tübingen, Germany. His research interests include genome mining in microbial genomes for secondary metabolites, biosynthesis research, isolation, 2D-/3D-structure elucidation (NMR, MS), and mode of action investigations of natural products.

Wolfgang Lindner studied chemistry at the KFU-Graz, Austria, where he worked for over 20 years at the Institute of Pharmaceutical Chemistry. In 1996 he was appointed a Chair of Analytical Chemistry at the University of Vienna, Austria, and became Emeritus in 2012. Throughout his career he has been strongly influenced by life sciences themes spanning from pharmaceutical analysis, metabolomics, and proteomics, and separation science methods such as HPLC, SFC, GC, CE/CEC, and LC–MS.

Source URL: http://www.chromatographyonline.com/peptide-analysis-zwitterionic-chiral-ion-exchangers-complementary-option-hilic-and-reversed-phase-ch

Links:
[1] http://www.pharmtech.com/tony-zhang
[2] http://www.pharmtech.com/emilie-holder
[3] http://www.chromatographyonline.com/pilar-franco
[4] http://www.pharmtech.com/michael-l-mmerhofer
[5] http://www.pharmtech.com/adrian-sievers-engler
[6] http://www.pharmtech.com/heike-gerhardt
[7] http://www.pharmtech.com/harald-gross
[8] http://www.chromatographyonline.com/wolfgang-lindner
[9] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/figure%201%20L.jpg
[10] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/table%20L.jpg
[11] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/table%202%20L.jpg
[12] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/figure%202%20L.jpg
[13] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/figure%203%20L.jpg
[14] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/table%203%20L.jpg
[15] http://files.alfresco.mjh.group/alfresco_images/pharma//2016/03/09/04b1313d-02a5-4345-aa0c-7fa9fbc3e21a/figure%204%20L.jpg

Appendix III

Accepted Manuscripts

▌**Energy Transfer**

Molecular Recognition and Visual Detection of G-Quadruplexes by a Dicarbocyanine Dye

Beatrice Karg,[a] Andrea Funke,[a] Anna Ficht,[a] Adrian Sievers-Engler,[b] Michael Lämmerhofer,[b] and Klaus Weisz*[a]

Abstract: The interactions of a dicarbocyanine dye 3,3′-di-ethylthiadicarbocyanine, **DiSC$_2$(5)**, with DNA G-quadruplexes were studied by means of a combination of various spectroscopic techniques. Aggregation of excess dye as a result of its positive charge is promoted by the presence of the polyanionic quadruplex structure. Specific high-affinity binding to the parallel quadruplex of the *MYC* promoter sequence involves stacking of **DiSC$_2$(5)** on the external G-tetrads; the 5′-terminal tetrad is the favored binding site. Significant energy transfer between DNA and the dye in the UV spectral region is observed upon **DiSC$_2$(5)** binding. The transfer efficiency strongly depends on the DNA secondary structure as well as on the G-quadruplex topology. These photophysical features enable the selective detection of DNA quadruplexes through sensitized **DiSC$_2$(5)** fluorescence in the visible region.

Introduction

Motivated by their demonstrated existence in vivo[1] and their potential biological role, including transcriptional regulation,[2] the molecular recognition and detection of G-quadruplex structures has been the subject of many studies in recent years. The G-quadruplexes are formed from guanine-rich sequences through Hoogsteen hydrogen bonding among a cyclic array of four guanine bases.[3] In general, the quadruplex core consists of two to four of such planar G-tetrads that exhibit a helical stacked arrangement and are additionally stabilized through monovalent cations, such as sodium or potassium, positioned within a central channel of negative electrostatic potential. The stacked tetrads present four grooves of the same or different geometry, depending on relative G-tract orientations and guanine glycosidic torsion angles. The physiologically most relevant monomolecular quadruplexes are further characterized through intervening sequences that form different types of loops upon folding into a quadruplex structure. Because G-rich sequences with quadruplex-forming ability are not only found in the human telomeric sequence at the chromosomal ends, but also in various promoter regions of oncogenes such as *MYC* or *BCL2*, they have become attractive targets in anticancer strategies.[4] Alternatively, they have also been explored for use as aptamers, DNAzymes, or sensors of metal ions and small molecules.[5]

The specific interactions of many cyanine dyes with nucleic acids combined with their excellent photophysical properties have resulted in various applications, such as cancer imaging[6] or nucleic acid staining.[7] More recently, the favorable characteristics of cyanines have increasingly triggered efforts aimed at their use as specific G-quadruplex-targeting agents.[8] Studies on the binding of cyanine dyes to quadruplexes that employed UV/Vis, circular dichroism (CD), and fluorescence spectroscopy have suggested various modes of binding, depending on the dye structure and quadruplex topology.[9] Unfortunately, information on binding, as extracted from these low-resolution methods, remains vague in most cases. On the other hand, a detailed characterization of geometries and specific binding modes by solution NMR spectroscopy for small ligands associated with G-quadruplexes has proven notoriously difficult. This is especially true for cyanine dyes, with their strong propensity to self-aggregate, depending on the particular solution conditions, and high-resolution structures of corresponding associates have not been reported to date.

Clearly, the development of quadruplex-specific fluorescent dyes that discriminate against double-helical DNA and also other quadruplex conformers would significantly extend their use as powerful detection and recognition tools not only for in vitro but also for in vivo applications. Thus, cyanine dyes based on the benzothiazole heterocyclic moiety have been shown to inhibit telomerase through their putative G-quadruplex-stabilizing properties.[10] A cyanine dye has also been employed for the visual detection of potassium ions through its differential binding to two telomeric quadruplex motifs.[11] Previously, the interaction of the symmetrical cationic cyanine dye 3,3′-diethylthiadicarbocyanine (**DiSC$_2$(5)**) with duplex DNA was studied by spectroscopic methods.[12] An increase of its blue-

[a] B. Karg, A. Funke, A. Ficht, Prof. Dr. K. Weisz
Institute of Biochemistry, Ernst-Moritz-Arndt University Greifswald
Felix-Hausdorff-Str. 4, 17487 Greifswald (Germany)
E-mail: weisz@uni-greifswald.de

[b] A. Sievers-Engler, Prof. Dr. M. Lämmerhofer
Institute of Pharmacy, Eberhard Karls University Tübingen
Auf der Morgenstelle 8, 72076 Tübingen (Germany)

🖥 Supporting information for this article is available on the WWW under
http://dx.doi.org/10.1002/chem.201502118.

shifted H-aggregate absorption band in the UV/Vis spectra has been attributed to its dimerization within the minor groove of alternating adenine/thymine (A/T) or inosine/cytosine (I/C) sequences, but not guanine/cytosine (G/C) sequences. Duplex-templated dye aggregation was reported to extend in an end-to-end fashion on longer sequences in the duplex minor groove, giving rise to strong biphasic Cotton effects from exciton coupling between neighboring chromophores in CD spectra and to result in a significant quenching of the dye fluorescence. Later, as part of a test series of carbocyanine dyes, its ability to bind a telomeric G-quadruplex structure was also examined by using a Taq polymerase stop assay.[10] However, due to Taq polymerase inhibition through nonspecific effects, no binding data have been reported and G-quadruplex binding by $DiSC_2(5)$ has not been further explored.

Herein, we report on a detailed characterization of the $DiSC_2(5)$ cyanine dye binding to the *MYC* quadruplex by a combination of various spectroscopic techniques. Additional binding studies with other quadruplex and duplex structures reveal significant differences in the photophysical properties of complexes, with energy-transfer processes from DNA to the dye enabling a straightforward and selective visual detection of particular quadruplex topologies.

Results and Discussion

Dicarbocyanine $DiSC_2(5)$

The structure of $DiSC_2(5)$ is shown in Figure 1 together with the DNA oligonucleotide sequences used in the present studies. The UV/Vis spectrum of a 2 μM solution of $DiSC_2(5)$ in aqueous buffer shows a major absorption at $\lambda = 646$ nm followed by a blueshifted absorption at $\lambda \approx 582$ nm and an additional shoulder of low intensity at shorter wavelengths (Figure S1 in the Supporting Information). Depending on the solution conditions, the latter exhibits some absorbance changes with time and we tentatively assign this band to higher-order H-aggregates that are slowly formed in the aqueous buffer, as recently reported for other cyanine dyes.[13] Absorbances at $\lambda = 646$ and 582 nm have previously been shown to result from monomeric dye and H-type dimers, respectively.[12] In line with these assignments, the major absorbance at $\lambda = 646$ nm increases with increasing temperature and decreasing solvent polarity through the stepwise addition of methanol (Figure S1 in the Supporting Information). This is associated with an increasing ratio of absorbances A_{646}/A_{582}, as expected for an equilibrium between monomers and dye aggregates.

Steady-state fluorescence measurements

To follow changes in its fluorescence, the cyanine dye was titrated with the *MYC* quadruplex and corresponding emission spectra were recorded at 20 °C (Figure 2 A). Apparently, dye fluorescence exhibits discontinuous behavior. Initial quenching with the dye in excess (process 1) is followed by a gradual fluorescence enhancement and a significant bathochromic shift at a dye/*MYC* molar ratio of ≤ 2 (process 2). The presence of at
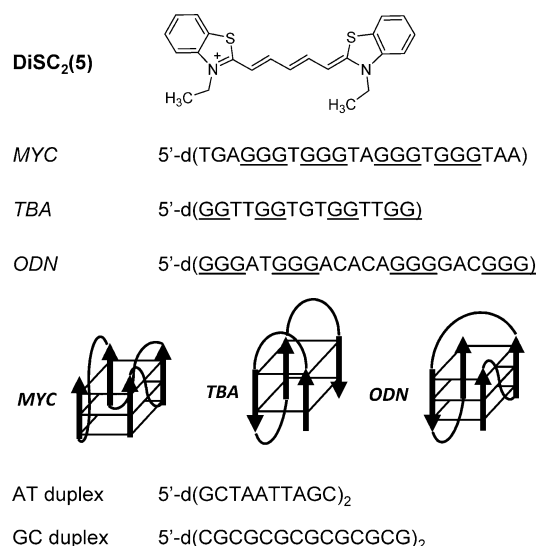
Figure 1. Structure of $DiSC_2(5)$ and oligonucleotide sequences of quadruplexes and double-stranded DNA species; for quadruplexes, guanine bases of the G-tetrads are underlined. Also shown are the topologies of quadruplexes *MYC*, *TBA*, and *ODN*.

least two different binding equilibria with competing interactions have opposing effects on the fluorescence quantum yield. When increasing the temperature in the fluorescence titration experiments, the course of dye fluorescence changes noticeably. With increased DNA-free fluorescence of $DiSC_2(5)$ at higher temperature, as expected for dye deaggregation, the extent of initial quenching for the dye in excess is gradually reduced and fully eliminated at 60 °C (Figure 2 B). At this temperature, only a continuous fluorescence increase together with a bathochromic shift is observed during stepwise titration.

To test any unspecific binding of the cationic cyanine dye to polyanionic DNA, we titrated a single-stranded (ss) oligonucleotide dT_{10} to the dye solution at 20 °C and followed the $DiSC_2(5)$ fluorescence (Figure S2 in the Supporting Information). Interactions of the dye with ssDNA is indicated by a noticeable decrease in the dye fluorescence. Although fluorescence is quenched by 50 % at a 2:1 molar ratio of dT_{10}/$DiSC_2(5)$, saturation is not achieved, even with the ssDNA in large excess, as expected for a nonspecific and rather weak electrostatically driven binding process.

Because $DiSC_2(5)$ has been reported to interact with double-helical [poly(dA−dT)]₂, but not with GC-rich duplexes, through the formation of extended aggregates along the duplex minor groove,[12a] we also titrated short AT- and GC-rich duplexes against the cyanine and examined changes in its fluorescence intensity (Figure S2 in the Supporting Information). As expected for exclusively nonspecific binding, titrating the GC duplex against the dye solution again resulted in a progressive quenching of the $DiSC_2(5)$ fluorescence with a rather steep decrease at the beginning of titration. In contrast, when titrated with the AT duplex, a discontinuous development of fluorescence is again observed with quenching effects at the beginning of titration followed by fluorescence enhancements with DNA in excess. However, the latter are clearly smaller than that
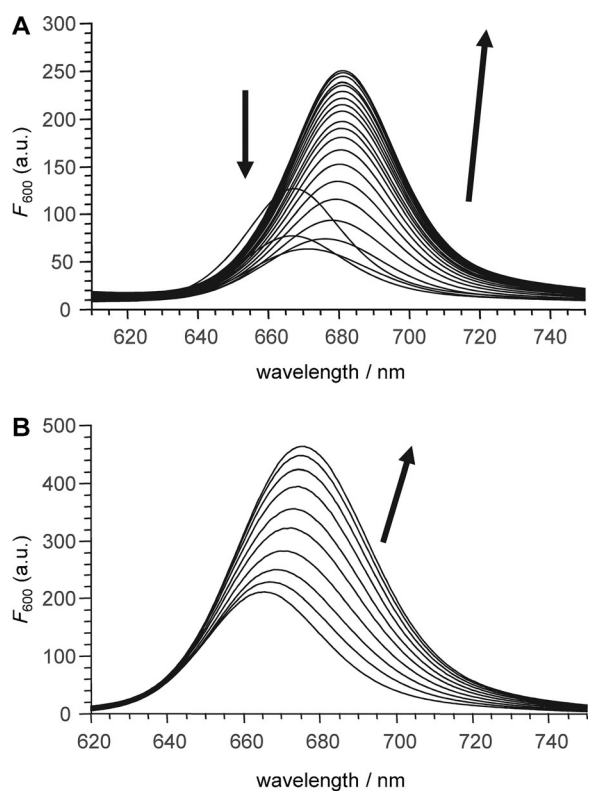
Figure 2. Fluorescence titrations ($\lambda_{exc} = 600$ nm) of 3 μM **DiSC$_2$(5)** in potassium buffer with the *MYC* quadruplex added in 0.75 μM aliquots at: A) 20, and B) 60 °C; the arrows indicate changes to the emission band during titration. Emission spectra were corrected for absorbance changes at $\lambda = 600$ nm.

observed for the *MYC* quadruplex and result in fluorescence intensities that barely surpass the free **DiSC$_2$(5)** fluorescence, even at a 4.5-fold excess of duplex.

UV/Vis titrations with the *MYC* quadruplex at 20 °C closely parallel the fluorescence titrations. Thus, titrating *MYC* against the dye results in an initial decrease in the monomeric **DiSC$_2$(5)** absorbance followed by an increase and bathochromic shift by up to 15 nm at dye/quadruplex molar ratios of <2 (Figure S3 in the Supporting Information). Also, the monomer concentration decreases in favor of H-aggregates at low quadruplex concentrations, but increases for the quadruplex in excess, as determined from the relative monomer and aggregate absorbances. Such more complex spectral changes are not without precedent for quadruplex–cyanine dye interactions.[14,15] Very similar behavior has recently been reported with the cyanine dye 2,2′-diethyl-9-methylselenacarbocyanine upon its interaction with G-quadruplexes and discussed in terms of assembly–disassembly processes of the dye at the DNA template.[14]

In the following, we describe characterizing initial process 1 at the start of the titration as being nonspecific and of low affinity because 1) fluorescence changes parallel quenching effects, as observed upon titration with ssDNA, and 2) these quenching effects are largely eliminated at higher temperatures. Cyanine dyes are known to form either H- or J-aggregates, depending on their structural features and on the particular solvent conditions. Compared with the monomeric spe-

cies, J-aggregates exhibit bathochromic shifts in their absorption band and significant fluorescence emission, whereas H-aggregates show a blueshifted absorption band and decreased fluorescence intensity. Consequently, the results presented above for the fluorescence and UV/Vis titrations suggest binding events that initially involve enhanced aggregation of the dye at the polyanionic nucleic acid to form H-aggregates associated with fluorescence quenching. At higher quadruplex/dye molar ratios, a gradual deaggregation and enhanced specific high-affinity binding of monomeric dye is expected to occur. The latter, which is characterized by an increase and bathochromic shift of the dye emission, starts to dominate the overall impact on the dye fluorescence at a dye/quadruplex molar ratio of <2. Because enthalpy-driven exothermic dye aggregation at the negatively charged quadruplex surface should be considerably weakened with temperature, only specific high-affinity binding may be effective at higher temperature; this is in agreement with experimental data (Figure 2).

Because titration of the cyanine dye with quadruplex DNA yields a complex titration curve due to various competing binding modes with contributions that depend on the dye/quadruplex molar ratio, changes in **DiSC$_2$(5)** fluorescence upon quadruplex binding barely allow for a reliable determination of association constants for the dye-specific binding. In contrast, when titrating dye against DNA, there is an excess of quadruplex at the beginning of such a reverse titration and this situation should favor specific high-affinity binding interactions. In addition, any quadruplex signal used for a direct readout is expected to be less sensitive to unspecific, compared with specific, ligand binding.

Unfortunately, the intrinsic fluorescence of the natural pyrimidine and purine bases in nucleic acids is too low to be effectively used for analytical methods. However, guanine residues in G-quadruplexes were shown to have enhanced photoexcited lifetimes and fluorescence quantum yields that often constituted a unique feature of these DNA structures.[16] Indeed, *MYC* in K$^+$ buffer exhibits noticeable guanine fluorescence, with a wavelength of $\lambda = 387$ nm at maximum emission, that is significantly quenched when interacting with the cyanine dye (Figure 3A). Plotting the G fluorescence in a reverse titration experiment as a function of added **DiSC$_2$(5)** gives a simple hyperbolic binding isotherm. With binding sites of different affinity and likely also of different impact on the quadruplex fluorescence (see below), fitting of this isotherm would require a model with inequivalent binding sites. However, with too many free-floating parameters such curve fittings become increasingly error-prone and random. We therefore resorted to a simplified model based on independent and equivalent binding sites, yielding 2.6 binding sites per quadruplex and an association constant per site of $K_a = 4.8 \times 10^5$ M^{-1} at 20 °C. As seen from Figure 3B and expected from the above considerations, some deviations between calculated and experimental data are clearly noticeable. However, given such a simple analysis, fitting results should nevertheless give a reasonably good idea of affinity and number of specific binding sites for the ligand.

A

B

$K_a = 4.8 \cdot 10^5 \text{ M}^{-1}$

$n = 2,6$
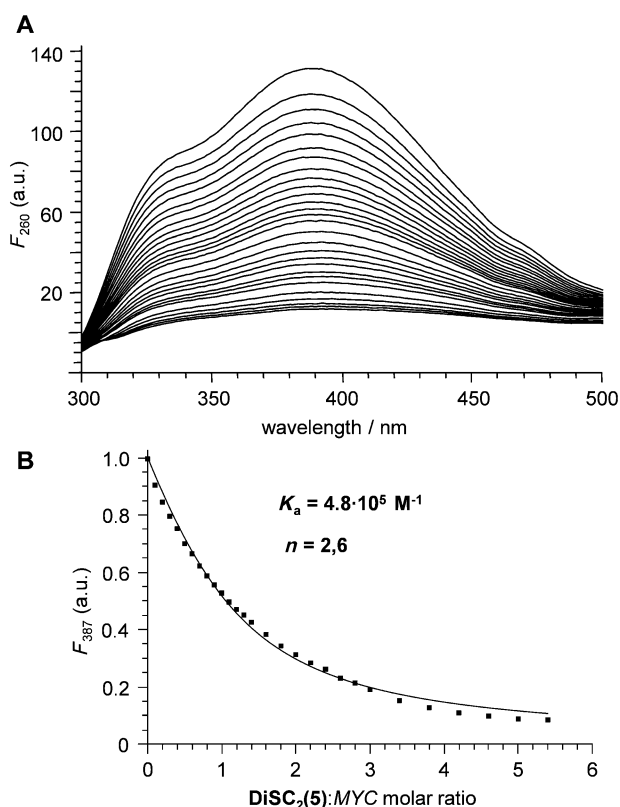
Figure 3. A) Fluorescence emission ($\lambda_{exc} = 260$ nm) of 6 µM *MYC* quadruplex in potassium buffer upon titrating **DiSC₂(5)** up to a final 5.4:1 dye/quadruplex molar ratio at 20 °C. B) Binding isotherm based on the fluorescence intensity at $\lambda = 387$ nm together with the fitted curve.

NMR spectroscopy

Adding **DiSC₂(5)** in small increments to a solution of *MYC* results in considerable signal broadening of quadruplex resonances. As shown in Figure 4 A, guanine imino signals broaden to a variable extent at the initial steps of titration, but sharpen up again at a dye/quadruplex molar ratio of ≥2. Conspicuously, imino protons of G13, G4, and G17 of the 5′-terminal G-tetrad and G6 within the 3′-terminal tetrad broaden out almost completely after the addition of 0.5 dye equivalents, which indicates their significant perturbation upon ligand binding. Unfortunately, as a result of ligands occupying different binding sites with fast to intermediate exchange rates on the NMR chemical shift timescale, resonances are considerably broadened and overlapped, even with the dye

in excess. With only a partial assignment of resonances, a more detailed NMR structural characterization of the complex was rendered impossible.

Nevertheless, to gain information on possible interaction sites, we inspected the differential line-broadening effects not only for the guanine imino protons, but also for 2D NOE contacts observed in the base to sugar proton spectral regions (see the Experimental Section and Figure S4 in the Supporting Information). Signal broadening depends on the population of exchanging species and on chemical shift differences of individual protons in the free and complexed *MYC*. The average extent of broadening for individual nucleotides was therefore used as an indicator for perturbations caused by **DiSC₂(5)** binding; this is mapped on the schematic quadruplex structure in Figure 4 B. Clearly, such footprints, involving varying numbers of different protons, only provide rather poor resolution, but nevertheless give some valuable hints on **DiSC₂(5)** binding. Thus, loops seem to be barely affected by the dye, whereas 5′- and 3′-overhang sequences, as well as external G-tetrads, exhibit more significant perturbations. It is the 5′-terminal tetrad that experiences the largest effects. Consequently, the cyanine may preferentially bind through outer stacking on top of the 5′-terminal tetrad, with additional stacking on the 3′-terminal tetrad, albeit with lower affinity (Figure 5). Such a situation is not without precedent and has also been reported for the interaction of a quindoline drug with the *MYC* quadruplex.[17] Here, stacking of the ligand on the 5′- and 3′-terminal G-tetrads showed different features, but was promoted in both cases by the flanking bases that formed a binding pocket, as also indicated for cyanine binding. Whereas interactions of the cyanine dye with propeller loops can largely be excluded, low-affinity binding to the available groove cannot be completely dismissed based on these data.

A

B
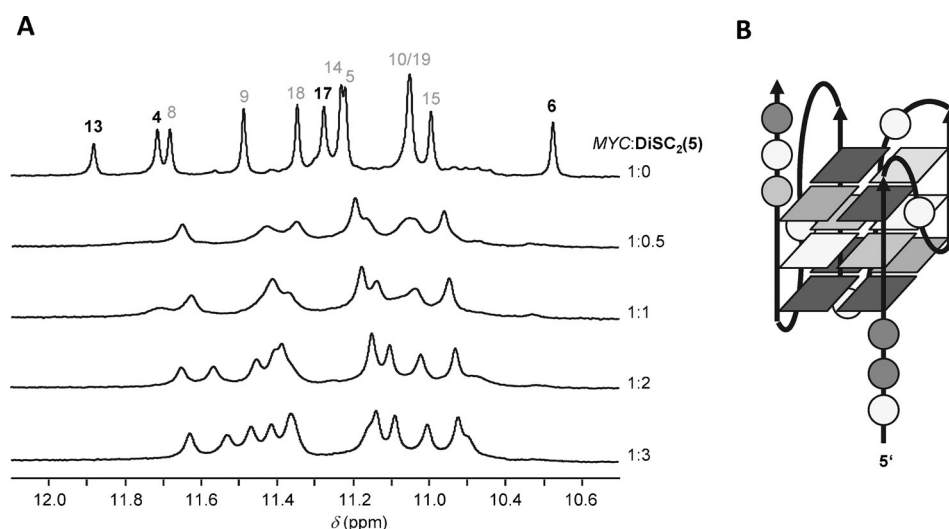
Figure 4. A) Imino proton spectral region of 0.66 mM *MYC* quadruplex with increasing amounts of **DiSC₂(5)** at 40 °C; imino assignments are indicated for the free quadruplex with guanine resonances that exhibit considerable line broadening highlighted. B) Schematic structure of the *MYC* quadruplex with darker residues that experience more significant line-broadening effects upon dye addition; overhang and loop nucleosides are represented by circles.
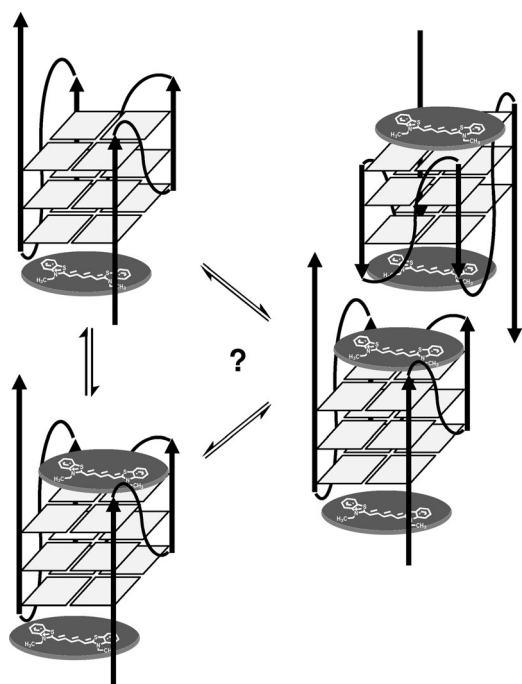
Figure 5. Proposed **DiSC₂(5)**–*MYC* complex structures with populations that depend on concentrations.

crease in intensity with increasing dye/quadruplex molar ratios. Neglecting minute amounts of a possible 3:1 complex only observed at closer inspection at a 4:1 molar ratio, stable complexes with a 1:1 and 2:1 stoichiometry seem to be the major species transferred to the gas phase. Signal intensities are in agreement with uncooperative binding of the dye molecules to two binding sites with differential affinity and corroborate the NMR spectroscopy results that suggest stacking interactions of different strengths on the 5′- and 3′-terminal tetrads.

Circular dichroism

The achiral cyanine dye exhibits no CD signal in the absence of quadruplex DNA. However, upon titrating **DiSC₂(5)** into a solution of *MYC* in potassium buffer, induced Cotton effects in the absorption region of the dye emerge in addition to the typical CD signature of the parallel *MYC* quadruplex below $\lambda = 300$ nm (Figure 7). The latter involves positive and negative amplitudes at $\lambda = 265$ and 245 nm and is barely changed upon addition of the dye; this confirms the conservation of the quadruplex structure in the presence of the ligand. Cotton effects in the long-wavelength absorption region of the bound dye at a 1:1 quadruplex/dye molar ratio include a negative

ESI mass spectrometry

To reduce counterion condensation on the nucleic acid upon evaporation in the presence of sodium or potassium ions, DNA and cyanine dye were dissolved for mass spectral analysis in 150 mm ammonium acetate buffer.[18] Addition of 10 % methanol had a favorable effect on the electrospray and signal stability without altering the buffer conditions to a more significant extent. Initial CD measurements on *MYC*–**DiSC₂(5)** mixtures in the ammonium acetate/MeOH buffer confirmed the formation of a parallel quadruplex (not shown). Also, binding of the dye is apparent by induced Cotton effects that closely match those measured in a potassium phosphate buffer (see below). The ESI mass spectrum of the free oligonucleotide indicate three charge states of $z = -4, -5,$ and -6 of which $z = -5$ is the dominant species. In agreement with cation coordination between G-tetrads, *m/z* 1403.4 represents the free quadruplex with two tightly bound ammonium ions (Figure 6). In the presence of the cyanine dye, new signals at higher *m/z* ratios appear, in addition to free quadruplex signals, and can be assigned to -5 charged complexes with one and two dye molecules bound to the quadruplex·2 NH₄⁺ (*m/z* 1481.4 and 1559.5, respectively). The 1:1 and 2:1 complexes in-

Figure 6. Mass spectra of solutions of *MYC* quadruplex with different *MYC*/**DiSC₂(5)** molar ratios.

high-amplitude signal at $\lambda = 672$ nm and two other positive bands of decreasing intensity centered at $\lambda = 645$ and 592 nm. Clearly, the various Cotton effects indicate a more complex pattern of binding for the cyanine dye. Only in rare cases do induced CD effects allow for a straightforward interpretation in terms of drug-binding modes. However, a closer inspection of the CD effects of the dye induced upon quadruplex binding gives some clues on dye interactions.

Conspicuously, the large negative amplitude of the long-wavelength band seems to be barely compatible with simple groove binding and, in particular, with external stacking. The
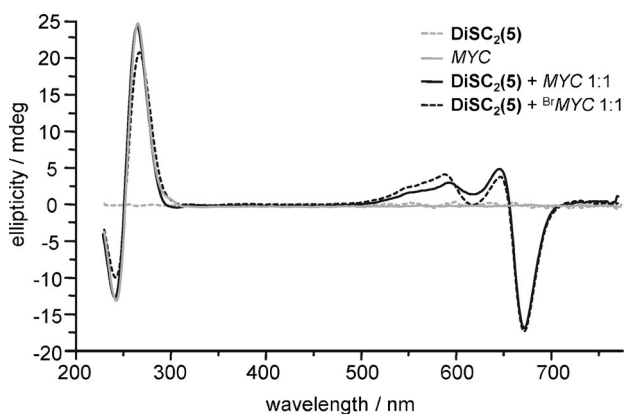
Figure 7. CD spectra of **DiSC₂(5)** (gray, broken line), the *MYC* quadruplex (gray, solid line), and equimolar mixtures of **DiSC₂(5)** with *MYC* (black, solid line) and 8-bromo-G monosubstituted ᴮʳ*MYC* (black, broken line) in potassium buffer; $T = 20\,°C$, c(quadruplex) $= 3\,\mu M$.

latter is expected to only induce very small CD effects on the stacked ligand. Rather, the long-wavelength Cotton effects may constitute a nonconservative CD couplet centered at the absorption of monomeric dye, with a smaller positive band at $\lambda = 645$ nm and a larger negative band at $\lambda = 672$ nm (see also Figure S5 in the Supporting Information). In general, exciton coupling between electronic transitions on neighboring dyes gives rise to a conservative CD signal with two bands of equal intensity and opposite sign located at the wavelength of maximum UV absorbance. In this case, the reduced intensity of the positive band may partially be attributed to other Cotton effects overlapping with the exciton couplet at shorter wavelengths. Clearly, such an interpretation requires at least two exciton-coupled dye molecules bound to the quadruplex. Support comes from CD titration experiments with *MYC* titrated in excess against **DiSC₂(5)** (Figure S5 in the Supporting Information). Starting with a 1:1 molar ratio, the addition of excess quadruplex results in a gradual decrease of all induced circular dichroism (ICD) amplitudes, even with a fixed dye concentration. Because excess DNA should increase the population of bound dye, these results are only compatible with Cotton effects that arise from interacting dyes simultaneously bound to *MYC*. The question arises whether the suggested exciton-split CD is due to short- or long-range exciton couplings. End-stacking of an H-type **DiSC₂(5)** dimer may be discarded because of its more blueshifted absorption. Likewise, a possible side-by-side stacking of two dyes on the outer G-tetrads places the two transition dipoles within the same plane, which prevents exciton coupling.[19] However, it is conceivable that two externally stacked dye molecules are sandwiched between two quadruplexes to form a dimer by the interaction of their outer faces, in analogy to the structural arrangement found for a telomeric G-quadruplex complexed with acridine derivative BRACO-19 (see Figure 5).[20] Indeed, such dimerization can be reconciled with results obtained from nondenaturing gel electrophoresis of *MYC*–dye mixtures (Figure S6 in the Supporting Information). Alternatively, with its large molar absorptivity of $\varepsilon_{660} = 266\,000\,M^{-1}\,cm^{-1}$ for the monomer absorption measured

at a dye/quadruplex molar ratio of 1:2, intense bisignate CD curves with exciton split bands can even be expected for favorably aligned 5′- and 3′-end-stacked dyes with interchromophoric distances of about 14 Å.[21]

We also performed a CD competition experiment by titrating a high-affinity N5-methylated indoloquinoline analogue (11-phenyl-indolo[3,2-*b*]quinoline (PIQ)) to a 2:1 mixture of **DiSC₂(5)** and *MYC*. Indoloquinoline derivatives have been shown to preferably interact with quadruplexes through stacking on their outer G-tetrads and a corresponding solution NMR spectroscopy structure has been reported for a 2:1 complex of an indoloquinoline derivative end-stacked on the *MYC* quadruplex.[17,22] Upon titrating the indoloquinoline, Cotton effects of **DiSC₂(5)** are immediately attenuated and only small amplitudes are left after the addition of one dye equivalent of PIQ (Figure S7 in the Supporting Information). The bound cyanine dye is effectively displaced by the competing ligand. A closer look reveals, however, that the negative band at $\lambda = 672$ nm shows a faster decline when compared with induced CD effects centered at $\lambda \approx 570$ nm; this again indicates the former as being exclusively derived from **DiSC₂(5)** outer stacking.

To independently check for possible groove binding, we also examined the CD spectra of the dye when bound to an 8-bromo-monosubstituted *MYC* quadruplex. With the dG analogue incorporated at position 5 of the central G-tetrad, the 8-bromo substituent projects into the available quadruplex groove and partially blocks groove access. Resulting in only a small decrease of both negative and positive CD amplitudes for the modified *MYC* quadruplex, the bromo substituent does not alter the overall quadruplex conformation (Figure 7). Likewise, Cotton effects in the dye absorption region for a 1:1 dye/quadruplex mixture show only subtle changes relative to the unmodified complex. Whereas the amplitude of the negative long-wavelength band remains completely unaffected in agreement with end-stacked dye molecules, small but noticeable amplitude changes are observed for the positive bands at shorter wavelengths. We tentatively assign these positive Cotton effects in the absorption range of H-aggregates to bound dimers. Their binding is expected to be of lower affinity, if their attenuated CD amplitude in mixtures with *MYC* in excess (Figure S5 in the Supporting Information) is attributed to their deaggregation followed by their transfer to sites of higher affinity, which are likely to be on top of the terminal G-tetrads. Whereas tight groove binding can largely be excluded based on the experimental data, the mode of binding for these dimers at or close to the groove and possibly also through partial stacking on the outer tetrads remains unresolved. Taken together, the overall signature of ICD effects is consistent with the partial superposition of two CD couplets of the same sign: a less intense bisignate signal of a dimer at shorter wavelengths and a stronger couplet at lower energy for end-stacked monomers.

Dye binding to other G-quadruplex structures

To study binding interactions of the cyanine dye with other quadruplexes of different topology, additional fluorescence ti-

trations of **DiSC₂(5)** with the *ODN* quadruplex and thrombin-binding aptamer *TBA* were performed. The *ODN* quadruplex forms a hybrid-type structure with one propeller, one diagonal, and one lateral loop, whereas *TBA* with its core of only two G-tetrads forms an antiparallel quadruplex that features three lateral loops (see also Figure 1). Titrating **DiSC₂(5)** with *ODN* at 20 °C results in a biphasic change of the dye fluorescence that is reminiscent of the corresponding *MYC* titrations (Figure 8). Following the initial decrease in fluorescence, moderate enhancements and bathochromic shifts are observed at later stages of the titration. Similar to *MYC* addition (Figure 2), the transition between processes 1 and 2 occurs at a quadruplex/dye molar ratio of about 0.5:1. Likewise, after correcting for absorbance changes at the excitation wavelength, as obtained by corresponding UV/Vis titration experiments, a fluorescence enhancement of about 1.5 can be determined after adding a sevenfold excess of *ODN*. In contrast, titrating *TBA* against **DiSC₂(5)** only shows a gradual quenching of the dye fluorescence and no fluorescence recovery up to a final threefold excess of quadruplex.

Based on the interpretation that nonspecific and specific interactions are responsible for the initial quenching and final enhancement of the dye fluorescence (see above), the course of the titration curves may be correlated with relative binding affinities. Thus, weak binding requires more quadruplex to overcome the quenching effect of electrostatically driven dye aggregation, and thus, identifies the *TBA* quadruplex as a rather poor dye receptor with low affinity. On the other hand, fluorescence changes upon addition of the parallel *MYC* and hybrid-type *ODN* quadruplex suggest much stronger cyanine binding. Strictly speaking, such a direct relationship requires a similar impact of binding on the **DiSC₂(5)** fluorescence, irrespective of the quadruplex type. However, the latter seems reasonable, especially in the case of similar modes of binding, and is also supported by results obtained from energy-transfer experiments (see below).

Energy transfer from DNA to DiSC₂(5)

Clearly, the intrinsic photophysical characteristics make **DiSC₂(5)** a very poor probe for the optical detection of G-quadruplex structures, even in case of the preferred *MYC* target. However, in addition to its strong absorbance in the visible spectral region, dye **DiSC₂(5)** shows low-intensity absorption bands at short wavelengths below $\lambda = 400$ nm due to higher electronic transitions with corresponding small transition dipole moments (Figure S1 in the Supporting Information). To probe possible energy-transfer processes between the DNA bases and bound dye, we recorded fluorescence excitation spectra from $\lambda = 230$ to 450 nm with observation of the dye fluorescence at $\lambda = 680$ nm in the presence of DNA (Figure 9). Whereas the excitation spectrum of the free dye corresponds to its absorption spectrum and shows barely any intensity in the UV region below $\lambda = 300$ nm, an intense and broad excitation band covering a spectral range from $\lambda = 240$ to 300 nm appears after the addition of *MYC* in a twofold excess. This new band with maxima at $\lambda = 280$ and 250 nm matches the
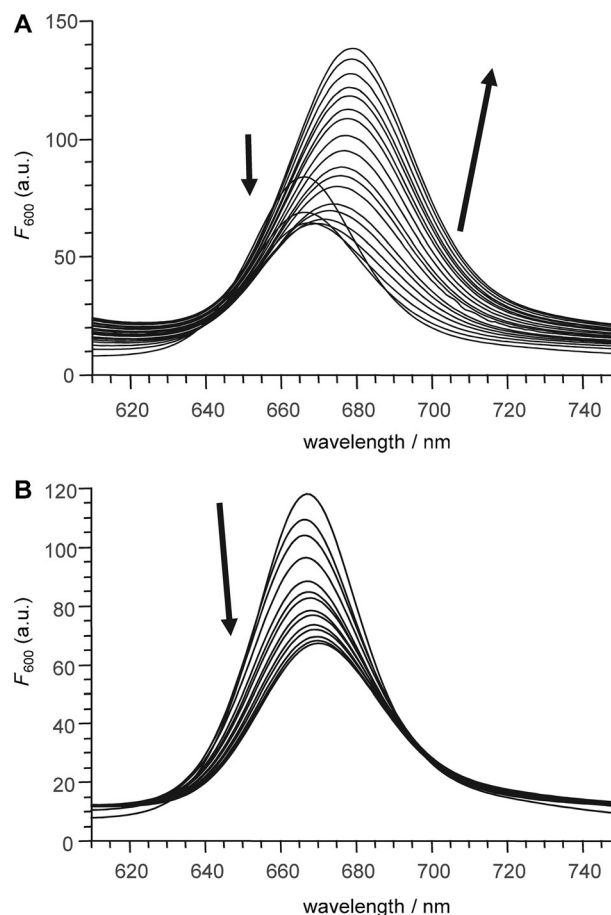
Figure 8. Fluorescence titrations ($\lambda_{exc} = 600$ nm) of: A) 2 µM **DiSC₂(5)** in K⁺ buffer with the *ODN* quadruplex added in 0.4 µM aliquots, and B) 3 µM **DiSC₂(5)** in K⁺ buffer with the *TBA* quadruplex added in 0.75 µM aliquots; $T = 20$ °C. The arrows indicate changes to the emission band during titration. Emission spectra were corrected for absorbance changes at $\lambda = 600$ nm.

DNA absorbance and must be attributed to energy transfer from the G-quadruplex to the dye. A corresponding transfer has also been reported in the past for a 3,3′-diethyloxadicarbocyanine interacting with a dimeric hairpin quadruplex.[23] However, compared with these previous studies, a much higher fluorescence enhancement through energy transfer is observed here for **DiSC₂(5)** when excited in the range $\lambda = 250$–280 nm. As a consequence, there is a much higher enhancement of dye emission upon *MYC* binding when excited at $\lambda = 254$ nm, compared with a direct excitation at $\lambda = 600$ nm (Figure S8 in the Supporting Information).

Energy transfer to **DiSC₂(5)** should depend on the binding affinity of the dye for the DNA target. Also, given a Förster mechanism of transfer, transfer is expected to be strongly promoted by the enhanced guanine fluorescence of the quadruplex above $\lambda = 300$ nm (Figure 3) and on the orientation between donor and acceptor associated with the binding mode and favored for a stacked dye with parallel transition moments. These geometric criteria have been used in the past to infer an intercalative mode of binding to double-helical DNA from the observation of a fluorescence resonance energy transfer from DNA bases to a fluorescent ligand.[24] In agreement with the
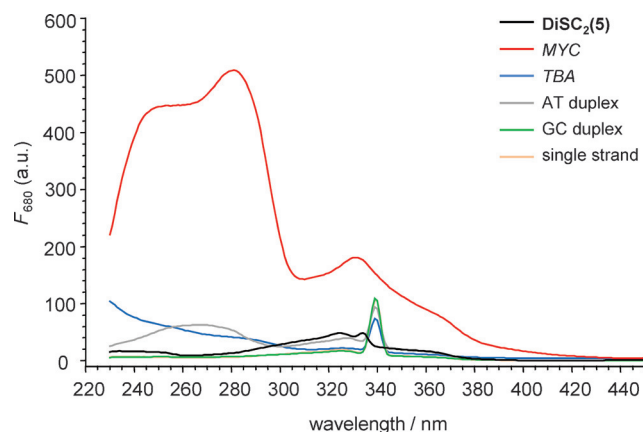
Figure 9. Fluorescence excitation spectra ($\lambda_{em} = 680$ nm) of 3 μM **DiSC$_2$(5)** in the absence and presence of 6 μM *MYC* quadruplex, *TBA* quadruplex, AT duplex, GC duplex, or single-stranded dT$_{10}$ at 20 °C.

Figure 10. Photograph of 3 μM solutions of **DiSC$_2$(5)** mixed with different DNA structures in a twofold excess (in quadruplex, duplex, or single-stranded) and irradiated with a UV lamp at $\lambda = 254$ nm.

absence of specific binding, no energy transfer was observed upon the addition of ssDNA or GC-rich duplex DNA to the cyanine dye. In fact, the very small excitation of free dye in the DNA absorbance region completely vanishes in the presence of these DNA species due to an inner filter effect caused by the nucleic acid.

We also tested the visible-light emission of the dye in the presence of the AT-rich duplex known to bind **DiSC$_2$(5)** in its minor groove, as well as of the *TBA* quadruplex expected to bind **DiSC$_2$(5)** with low affinity (see above). As shown in Figure 9, the excitation spectra in the presence of both DNA structures exhibit enhanced intensity in the UV region below $\lambda = 300$ nm relative to the free dye. These observations also suggest some energy transfer to the dye, albeit with significantly lower efficiency than in case of the *MYC* quadruplex.

To evaluate the potential of **DiSC$_2$(5)** to serve as a simple probe for the visual detection of specific DNA structures based on these energy-transfer processes, we prepared buffer solutions of 3 μM **DiSC$_2$(5)** with various DNA species in a twofold excess and irradiated the solutions with a UV lamp at $\lambda = 254$ nm (Figure 10). Whereas the *MYC* quadruplex is easily identified through its bright-red emission, as a result of sensitized **DiSC$_2$(5)** fluorescence, the *TBA* quadruplex is barely visible to the naked eye, and tubes with the other double- and single-stranded DNA species escape direct visual detection. Consequently, sensitized **DiSC$_2$(5)** fluorescence may provide an efficient tool to detect specific quadruplex structures down to sub-micromolar concentrations and to discriminate not only between quadruplexes and other DNA secondary structures, but also between quadruplex topologies.

Various cyanine derivatives have more recently been developed and tested for their ability to serve as light-up probes for G-quadruplexes.[25] Despite their sometimes significant fluorescence enhancements upon binding, their use as quadruplex-specific diagnostic tools is often hampered by their poor quadruplex selectivity, as exemplified by thiazole orange. On the other hand, energy transfer from DNA bases of the quadruplex has not been reported for these dyes. In a recent study, several [2.2.2]heptamethinecyanine derivatives were shown to bind
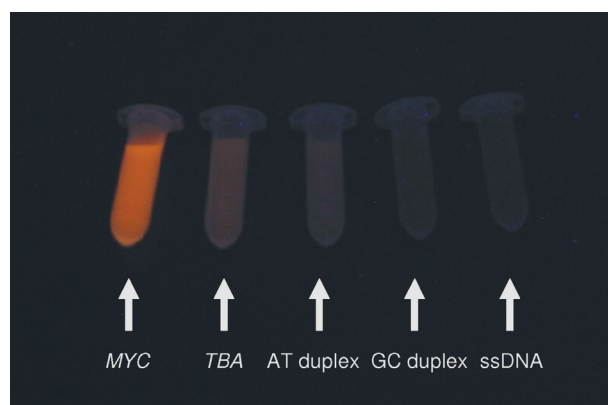
a telomeric quadruplex and a quadruplex composed of parallel and antiparallel topologies with rather high binding constants in the range of $0.2–0.9\times10^6$ and $0.2–17\times10^6$ M^{-1}, respectively.[15] Notably, with a 1:1 binding stoichiometry in a proposed terminal π-stacking mode, these cyanine dyes did not show induced CD signals in CD spectra of the complexes, despite their high quadruplex affinities. This again corroborates direct dye–dye interactions in the *MYC*–**DiSC$_2$(5)** complexes, giving rise to strong Cotton effects, as seen in the absorption region of the bound cyanine dye (Figure 7).

Conclusion

Results from optical spectroscopy on the association of the cyanine dye **DiSC$_2$(5)** with the *MYC* quadruplex indicated multiple binding modes with linked binding equilibria. Complexity arose in part from the pronounced propensity of the dye to form aggregates and aggregation was further promoted by the polyanionic G-quadruplex template. The present data indicated a scenario in which nonspecific outer binding associated with enhanced H-aggregate formation dominated with the dye in excess. Upon the addition of more quadruplex, individual dye molecules increasingly occupied specific high-affinity binding sites. Although the fast dynamics of association and dissociation processes prevented high-resolution structure determination of complexes, NMR spectroscopy results suggested that two dye molecules were tightly bound to the *MYC* quadruplex through end-stacking, albeit with stronger interactions at the 5′-terminal G-tetrad. Such a binding mode was corroborated by independent spectral data, as obtained from CD and ESI-MS.

A typical fluorescence light-up probe for optical G-quadruplex detection relies on its strongly increased fluorescence quantum yield when binding DNA.[25] In contrast, significant resonance energy transfer between the *MYC* quadruplex and the cyanine dye allows for the visual detection of the **DiSC$_2$(5)**–*MYC* complexes after UV irradiation through sensitized dye fluorescence in the long-wavelength region above $\lambda = 640$ nm. Given no or only small transfer efficiencies in the

case of single- and double-stranded DNA, as well as of low-affinity quadruplex topologies, the dye exhibits considerable structural selectivity. Conceivably, the quadruplex versus duplex discrimination could be further improved by established structural modifications of the dye incompatible with duplex minor groove binding. It should also be noted that DNA species lacking efficient energy transfer will even suppress any remaining background signal as a result of the inner filter effect from the DNA UV absorbance. These distinctive optical properties may be exploited in various potential applications. Thus, the use of **DiSC$_2$(5)** as a staining reagent for electrophoresis gels can reveal a quadruplex structure by direct visual inspection of the bands under UV irradiation (see also Figure S6 in the Supporting Information). Also, competition experiments of ligands might benefit from a visual readout of their relative ability to displace **DiSC$_2$(5)** from the quadruplex receptor. Additional studies on this sensing system are presently being carried out to further explore the sensitivity and, in particular, the discriminatory potential for various quadruplex sequences and topologies.

Experimental Section

Materials

All DNA oligonucleotides were purchased from TIB MOLBIOL (Berlin, Germany). Before use, oligonucleotides were precipitated in ethanol and the concentrations were determined spectrophotometrically by measuring the absorbance at $\lambda = 260$ nm. The cyanine dye **DiSC$_2$(5)** was purchased from Sigma–Aldrich with a purity of 98% and used without further purification.

Samples for optical measurements were obtained by dissolving the corresponding oligonucleotides in a buffer with 20 mm potassium phosphate, 100 mm KCl, at pH 7.0. Prior to measurements, the quadruplex samples were annealed by heating to 90 °C followed by slow cooling to room temperature. For the dye, stock solutions in DMSO were employed.

UV/Vis absorbance measurements

UV/Vis absorbances were measured with a Jasco V-650 spectrophotometer (Jasco, Tokyo, Japan). The temperature was controlled by means of a Peltier temperature control unit. For titration experiments, aliquots of the oligonucleotide were stepwise added to a 2–3 μm solution of the dye in 1 cm quartz cuvettes.

Circular dichroism

CD spectra were acquired with a Jasco J-810 spectropolarimeter equipped with a thermoelectrically controlled cell holder (Jasco, Tokyo, Japan). In general, by using 1 cm quartz cuvettes, CD data were collected with a bandwidth of 1 nm, a scanning speed of 100 nm min^{-1}, and 10 accumulations. All spectra were blank-corrected.

Fluorescence measurements

Fluorescence measurements were recorded with a Jasco FP-6500 spectrofluorometer thermostated with a Peltier unit (Jasco, Tokyo, Japan). By using excitation wavelengths of $\lambda = 600$ and 260 nm for **DiSC$_2$(5)** and the oligonucleotide, respectively, dye and G-quadru-

plex emission spectra were acquired from $\lambda = 610$ to 750 nm and from $\lambda = 300$ to 500 nm with a scanning speed of 50–100 nm min^{-1}, emission and excitation bandwidths of 10 and 5 nm, and a response time of 1 s. Fluorescence excitation spectra of the dye were recorded with emission wavelengths of $\lambda = 670$ or 680 nm without or with DNA, respectively. For fluorescence titrations, aliquots of the quadruplex were added stepwise to a 2–3 μm solution of the dye. All emission spectra of the dye were corrected for absorbance changes at $\lambda = 600$ nm during titrations. In reverse titrations, a 6 μm quadruplex solution was titrated with 0.6–1.5 μm aliquots of **DiSC$_2$(5)**.

ESI mass spectrometry

ESI-MS experiments were performed on a quadrupole time-of-flight instrument (TripleTOF5600+, Sciex, Ontario, Canada) by direct infusion. The G-quadruplex samples were measured in negative-ion mode with source and collision voltages set to -3800 and -10 V, respectively. The source temperature was screened from RT to 180 °C; the latter was used without significant deterioration of spectrum quality. The radiofrequency (rf) transmission window was set according to the m/z of the analyte. The infusion was performed by means of a syringe at 20 μL min^{-1}. Other optimized settings included a curtain gas of 16 psi, nebulizer and drying gas of both 17 psi, and a declustering potential of -20 V. Acquisition was performed with an accumulation time of 2000 ms in MCA (multiple channel acquisition) mode. Samples were prepared at a concentration of 2 μm of oligonucleotide in 150 mm ammonium acetate buffer (pH 7.0) containing 10% methanol and varying amounts of the cyanine dye (up to a 4:1 molar ratio of dye/quadruplex).

NMR spectroscopy experiments

For all NMR spectroscopy measurements, a low salt buffer with 10 mm potassium phosphate, pH 7.0, was used. Final concentrations of the *MYC* quadruplex were between 0.5 and 0.7 mm. NMR spectra were acquired on a Bruker Avance 600 MHz spectrometer equipped with an inverse ^{1}H/^{13}C/^{15}N/^{31}P quadruple resonance cryoprobehead and z-field gradients. Data were processed by using Topspin 3.1 and analyzed with CcpNmr Analysis software.[26] Proton chemical shifts were referenced relative to trimethylsilyl propionate (TSP) by setting the H$_2$O signal in 90% H$_2$O/10% D$_2$O to $\delta_H = 4.62$ ppm at 40 °C. A WATERGATE with w5 element was employed for solvent suppression for measurements in 90% H$_2$O/10% D$_2$O. 2D NOE experiments at 40 °C were acquired with mixing times of 150 and 200 ms for free *MYC* and *MYC* in the presence of **DiSC$_2$(5)**.

Differential line-broadening effects were evaluated for all unambiguously assigned intranucleotide base–sugar proton contacts in 2D NOE spectra. Cross-peak amplitudes for the complex were determined as a qualitative measure of the line widths in the ω_1 and/or ω_2 dimensions and related to corresponding cross-peak amplitudes of the unaffected free *MYC* quadruplex acquired and processed by using the same parameters. Cross-peaks observed for the free quadruplex, but completely missing in the complex spectra, should either indicate excessive line broadening or significant conformational readjustments, and corresponding residues were thus categorized as being strongly perturbed nucleotides. The extent of signal broadening was averaged for the individual residues and used as an indicator for perturbations caused by **DiSC$_2$(5)** binding.

[1] a) G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182–186; b) E. Y. N. Lam, D. Beraldi, D. Tannahill, S. Balasubramanian, *Nat. Commun.* **2013**, *4*, 1796.

[2] a) T. A. Brooks, L. H. Hurley, *Nat. Rev. Cancer* **2009**, *9*, 849–861; b) T. A. Brooks, L. H. Hurley, *Genes Cancer* **2010**, *1*, 641–649.

[3] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402–5415.

[4] a) S. M. Kerwin, *Curr. Pharm. Des.* **2000**, *6*, 441–471; b) D. Monchaud, M.-P. Teulade-Fichou, *Org. Biomol. Chem.* **2008**, *6*, 627–636.

[5] J. L. Neo, K. Kamaladasan, M. Uttamchandani, *Curr. Pharm. Des.* **2012**, *18*, 2048–2057.

[6] X. Yang, C. Shi, R. Tong, W. Qian, H. E. Zhau, R. Wang, G. Zhu, J. Cheng, V. W. Yang, T. Cheng, M. Henary, L. Strekowski, L. W. Chung, *Clin. Cancer Res.* **2010**, *16*, 2833–2844.

[7] M. Eriksson, H. J. Karlsson, G. Westman, B. Åkerman, *Nucleic Acids Res.* **2003**, *31*, 6235–6242.

[8] a) I. Lubitz, D. Zikich, A. Kotlyar, *Biochemistry* **2010**, *49*, 3567–3574; b) Q. Yang, J. Xiang, S. Yang, Q. Zhou, Q. Li, A. Guan, X. Zhang, H. Zhang, Y. Tang, G. Xu, *Chin. J. Chem.* **2010**, *28*, 1126–1132; c) R. Nanjunda, E. A. Owens, L. Mickelson, S. Alyabyev, N. Kilpatrick, S. Wang, M. Henary, W. D. Wilson, *Bioorg. Med. Chem.* **2012**, *20*, 7002–7011; d) R. Nanjunda, E. A. Owens, L. Mickelson, T. L. Dost, E. M. Stroeva, H. T. Huynh, M. W. Germann, M. M. Henary, W. D. Wilson, *Molecules* **2013**, *18*, 13588–13607.

[9] a) B. Jin, X. Zhang, W. Zheng, X. Liu, J. Zhou, N. Zhang, F. Wang, D. Shangguan, *Anal. Chem.* **2014**, *86*, 7063–7070; b) Q. Yang, J. Xiang, S. Yang, Q. Zhou, Q. Li, Y. Tang, G. Xu, *Chem. Commun.* **2009**, 1103–1105; c) Q. Yang, J. Xiang, S. Yang, Q. Li, Q. Zhou, A. Guan, X. Zhang, H. Zhang, Y. Tang, G. Xu, *Nucleic Acids Res.* **2010**, *38*, 1022–1033.

[10] S. M. Kerwin, D. Sun, J. T. Kern, A. Rangan, P. W. Thomas, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2411–2411.

[11] H. Sun, J. Xiang, W. Gai, Q. Shang, Q. Li, A. Guan, Q. Yang, Y. Liu, Y. Tang, G. Xu, *Analyst* **2012**, *137*, 5713–5715.

[12] a) J. L. Seifert, R. E. Connor, S. A. Kushon, M. Wang, B. A. Armitage, *J. Am. Chem. Soc.* **1999**, *121*, 2987–2995; b) M. Wang, G. L. Silva, B. A. Armitage, *J. Am. Chem. Soc.* **2000**, *122*, 9977–9986.

[13] T. Mahmood, Y. Wu, D. Loriot, M. Kuimova, S. Ladame, *Chem. Eur. J.* **2012**, *18*, 12349–12356.

[14] a) W. Gai, Q. Yang, J. Xiang, W. Jiang, Q. Li, H. Sun, A. Guan, Q. Shang, H. Zhang, Y. Tang, *Nucleic Acids Res.* **2013**, *41*, 2709–2722; b) W. Gai, Q. Yang, J. Xiang, L. Yu, A. Guan, Q. Li, H. Sun, Q. Shang, W. Jiang, H. Zhang, Y. Liu, L. Wang, Y. Tang, *Phys. Chem. Chem. Phys.* **2013**, *15*, 5758–5761.

[15] D. Dzubiel, H. Ihmels, M. M. A. Mahmoud, L. Thomas, *Beilstein J. Org. Chem.* **2014**, *10*, 2963–2974.

[16] N. T. Dao, R. Haselsberger, M.-E. Michel-Beyerle, A. T. Phan, *FEBS Lett.* **2011**, *585*, 3969–3977.

[17] J. Dai, M. Carver, L. H. Hurley, D. Yang, *J. Am. Chem. Soc.* **2011**, *133*, 17673–17680.

[18] a) F. Rosu, V. Gabelica, H. Poncelet, E. De Pauw, *Nucleic Acids Res.* **2010**, *38*, 5217–5225; b) F. Balthasart, J. Plavec, V. Gabelica, *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1–8.

[19] R. V. Person, K. Monde, H.-U. Humpf, N. Berova, K. Nakanishi, *Chirality* **1995**, *7*, 128–135.

[20] N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J. Am. Chem. Soc.* **2008**, *130*, 6722–6724.

[21] a) S. Matile, N. Berova, K. Nakanishi, J. Fleischhauer, R. W. Woody, *J. Am. Chem. Soc.* **1996**, *118*, 5198–5206; b) N. Berova, D. Gargiulo, F. Derguini, K. Nakanishi, N. Harada, *J. Am. Chem. Soc.* **1993**, *115*, 4769–4775.

[22] F. Riechert-Krause, K. Weisz, *Heterocycl. Commun.* **2013**, *19*, 145–166.

[23] Q. Chen, I. D. Kuntz, R. H. Shafer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2635–2639.

[24] D. Suh, J. B. Chaires, *Bioorg. Med. Chem.* **1995**, *3*, 723–728.

[25] a) E. Largy, A. Granzhan, F. Haman, D. Verga, M.-P. Teulade-Fichou, *Top. Curr. Chem.* **2012**, *330*, 111–178; b) S.-B. Chen, W.-B. Wu, M.-H. Hu, T.-M. Ou, L.-Q. Gu, J.-H. Tan, Z.-S. Huang, *Chem. Commun.* **2014**, *50*, 12173–12176; c) B. Jin, X. Zhang, W. Zheng, X. Liu, C. Qi, F. Wang, D. Shangguan, *Anal. Chem.* **2014**, *86*, 943–952; d) L. Yu, Q. Yang, J. Xiang, H. Sun, L. Wang, Q. Li, A. Guan, Y. Tang, *Analyst* **2015**, *140*, 1637–1646.

[26] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins* **2005**, *59*, 687–696.

CHEMISTRY
A European Journal

Supporting Information

Molecular Recognition and Visual Detection of G-Quadruplexes by a Dicarbocyanine Dye

Beatrice Karg,[a] Andrea Funke,[a] Anna Ficht,[a] Adrian Sievers-Engler,[b] Michael Lämmerhofer,[b] and Klaus Weisz*[a]

chem_201502118_sm_miscellaneous_information.pdf

Table of Contents

Figure S1. A) UV-vis absorption spectra of 2 μM **DiSC$_2$(5)** in potassium buffer at 20 $^{\circ}$C. B) Temperature dependent UV-vis absorption spectra of 3 μM **DiSC$_2$(5)** in potassium buffer at 20, 25, 30, 35, 40 and 45 $^{\circ}$C.
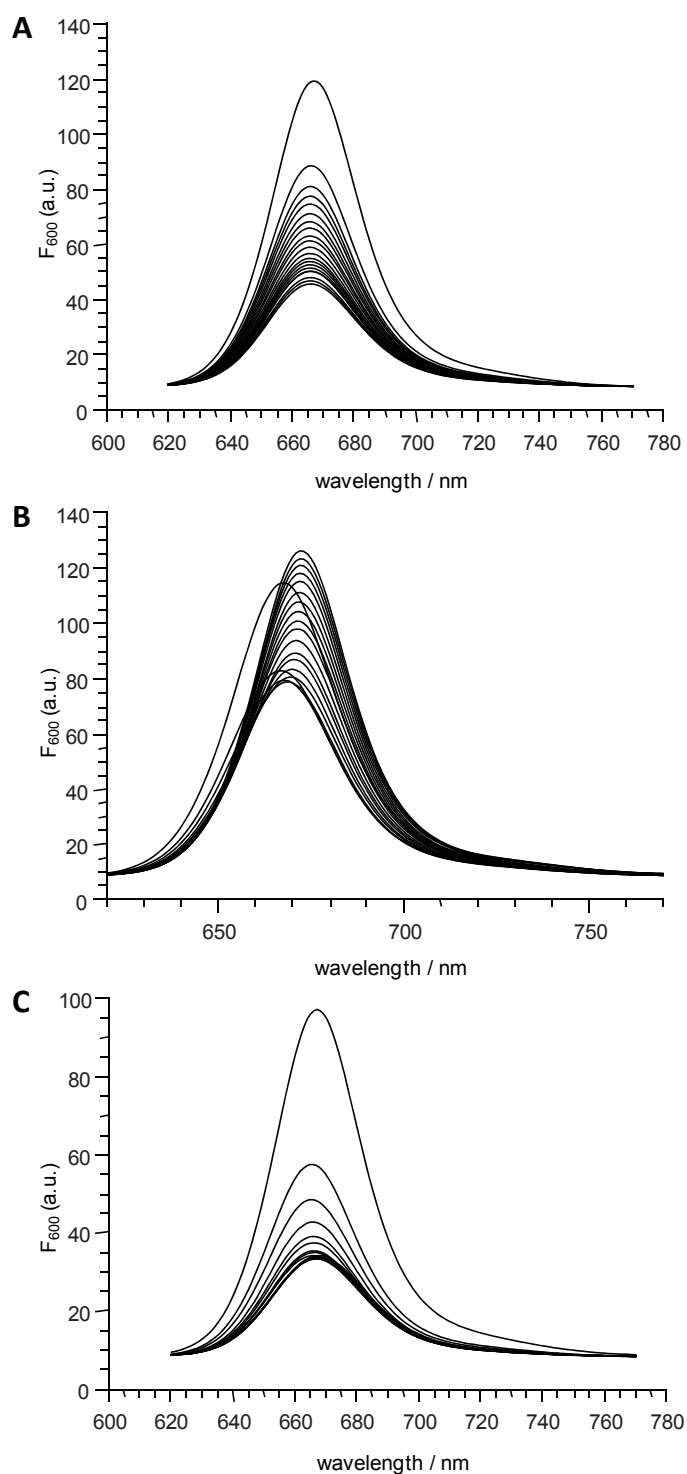
Figure S2. Fluorescence titration (λ_{exc}=600 nm) of 3 μM **DiSC$_2$(5)** in potassium buffer at 20 ^{o}C with A) the single-stranded dT$_{10}$ added in aliquots of 0.25 eq up to a final 5:1 ssDNA: **DiSC$_2$(5)** molar ratio, B) the AT duplex added in aliquots of 0.25 eq up to a final 4.5:1 duplex: **DiSC$_2$(5)** molar ratio, and C) the GC duplex added in aliquots of 0.25 eq up to a final 3.5:1 duplex: **DiSC$_2$(5)** molar ratio.
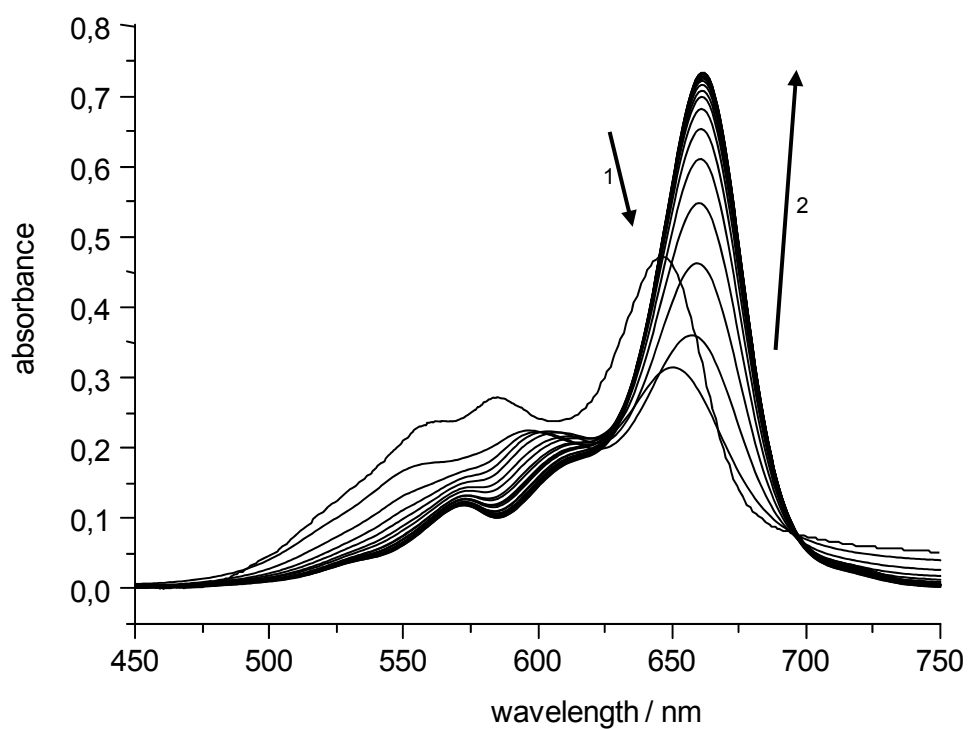
Figure S3. UV-vis titration of 3 μM **DiSC$_2$(5)** in potassium buffer with the *MYC* quadruplex added in 0.75 μM aliquots at 20 $^{\circ}$C; the arrows indicate changes of the monomer absorption band during titration.
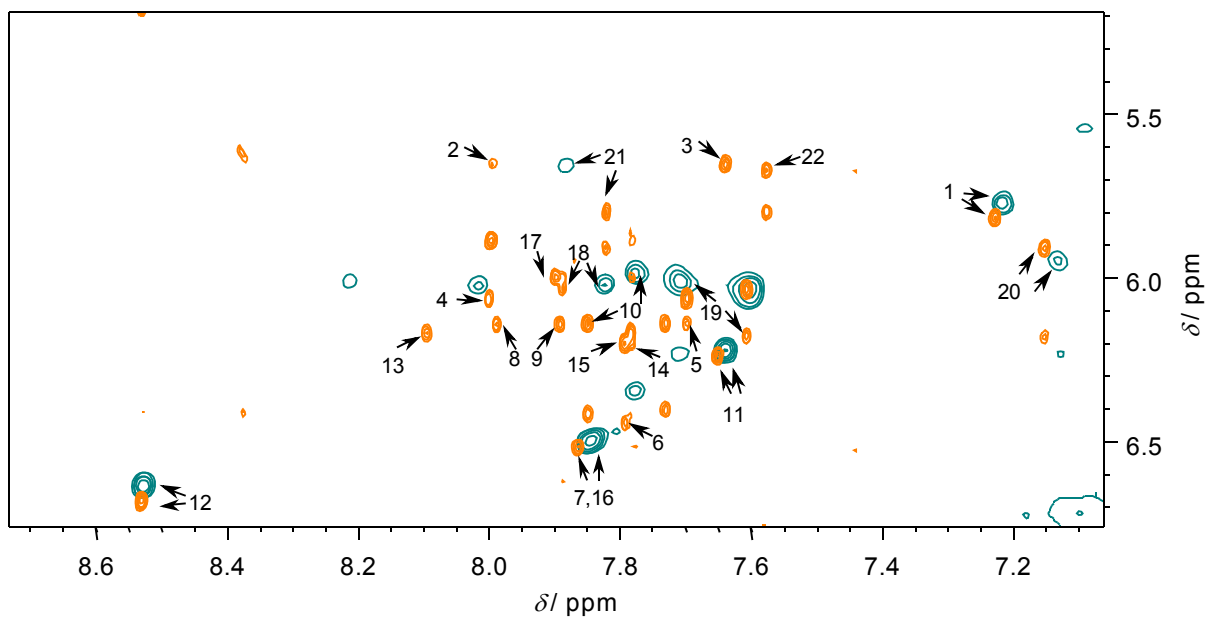
Figure S4. Superimposed 2D NOE H6/H8-H1' spectral region of the *MYC* quadruplex in the absence (orange) and in the presence (green) of a threefold excess of **DiSC$_2$(5)** at 40 °C (200 ms mixing time); assignments to the residue number are indicated for the intranucleotide base-sugar proton crosspeaks.
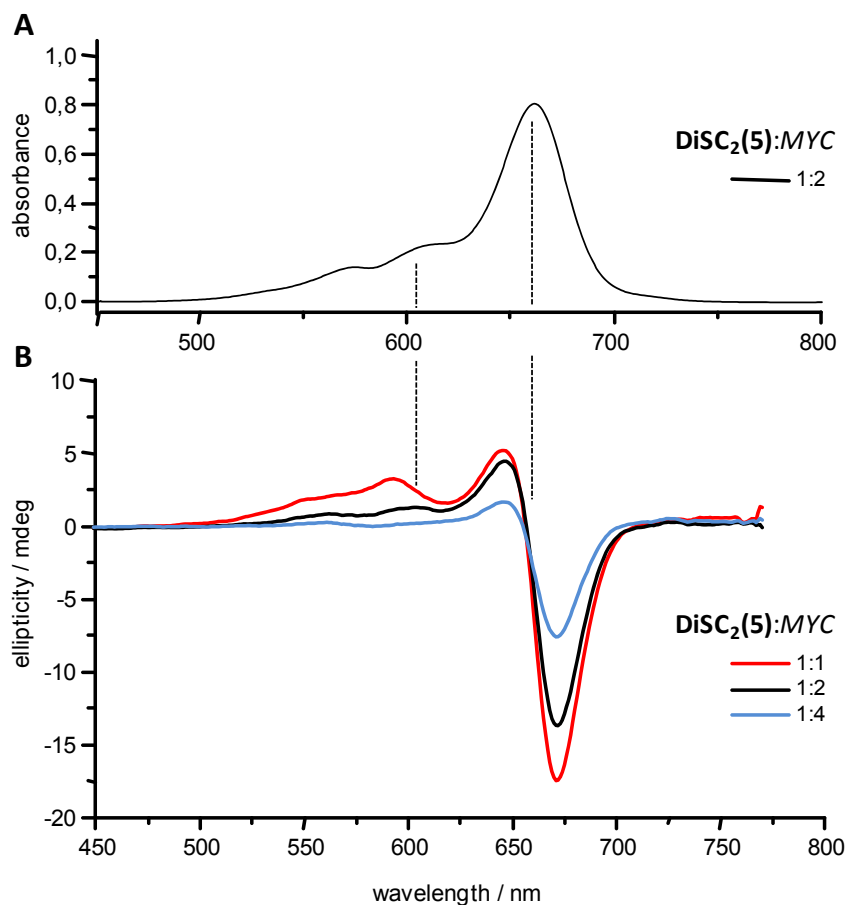
Figure S5. A) Long-wavelength region of UV-vis absorption spectrum for a 1:2 mixture of **DiSC$_2$(5)** with the *MYC* quadruplex and B) CD spectra for mixtures of **DiSC$_2$(5)** with the *MYC* quadruplex in molar ratios of 1:1 (red), 1:2 (black) and 1:4 (blue); T = 20 °C, c(**DiSC$_2$(5)**) = 3 μM.
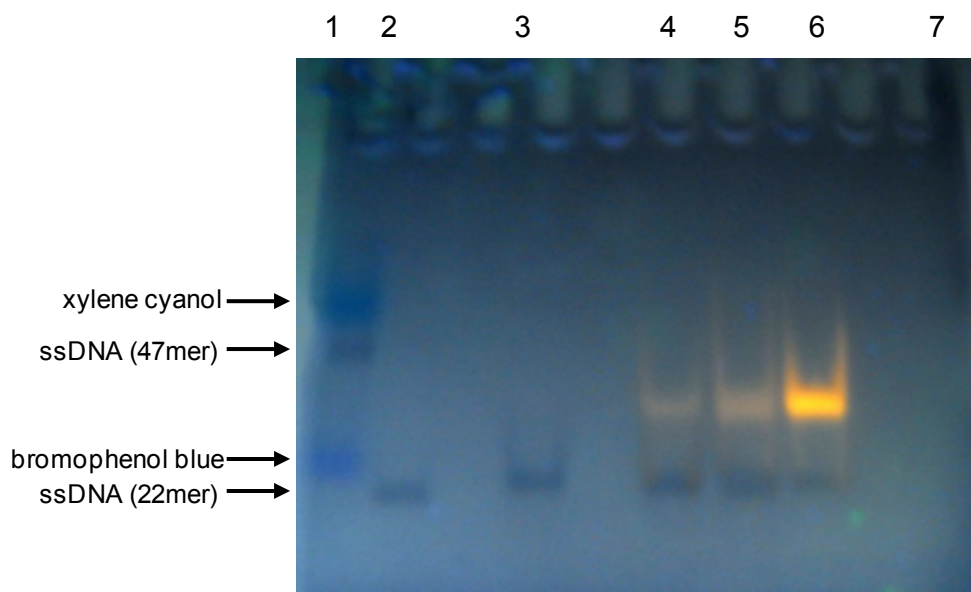
Figure S6. Non-denaturing gel electrophoresis of the *MYC* sequence (10 μM) in the absence and the presence of **DiSC$_2$(5)**. After annealing the DNA in K$^+$ containing buffer (pH 7.0) by heating to 90 °C for 5 min followed by slow cooling to room temperature, the dye was added and the samples were loaded on a 15% polyacrylamide gel. The gel was run with 100 V at 4 °C and bands were photographed under UV light at 254 nm. Lane 1 and 2: marker lanes; lane 3: *MYC*:**DiSC$_2$(5)** 1:0; lane 4: *MYC*:**DiSC$_2$(5)** 1:0.25; lane 5: *MYC*:**DiSC$_2$(5)** 1:1; lane 6: *MYC*:**DiSC$_2$(5)** 1:4; lane 7: *MYC*:**DiSC$_2$(5)** 0:1. Whereas the dye without DNA (lane 7) remains undetectable under these conditions, the *MYC* quadruplex shows a single dark band without added dye (lane 3). For the mixtures with added **DiSC$_2$(5)** an increasingly strong second band appears in addition to the free quadruplex band and is clearly identified as being a quadruplex with bound dye through its red emission. An end-to-end or end-to-tail stacking of this complex to form a dimeric species is compatible with its noticeably slower migration.
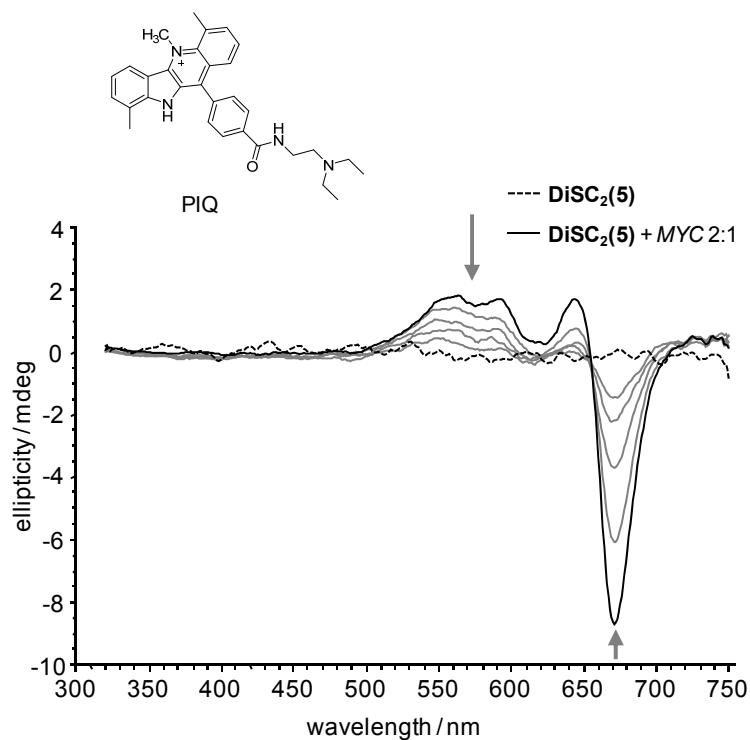
Figure S7. Long-wavelength CD spectral region of a 2:1 **DiSC₂(5)**:*MYC* quadruplex mixture at 20 °C and c(**DiSC₂(5)**) = 3 μM. Ligand displacement experiments were performed by adding the competing PIQ ligand (shown on top) in 0.75 μM aliquots to the dye-quadruplex solution up to a 1:1 PIQ-to-dye molar ratio; arrows indicate changes of CD amplitudes during titration.
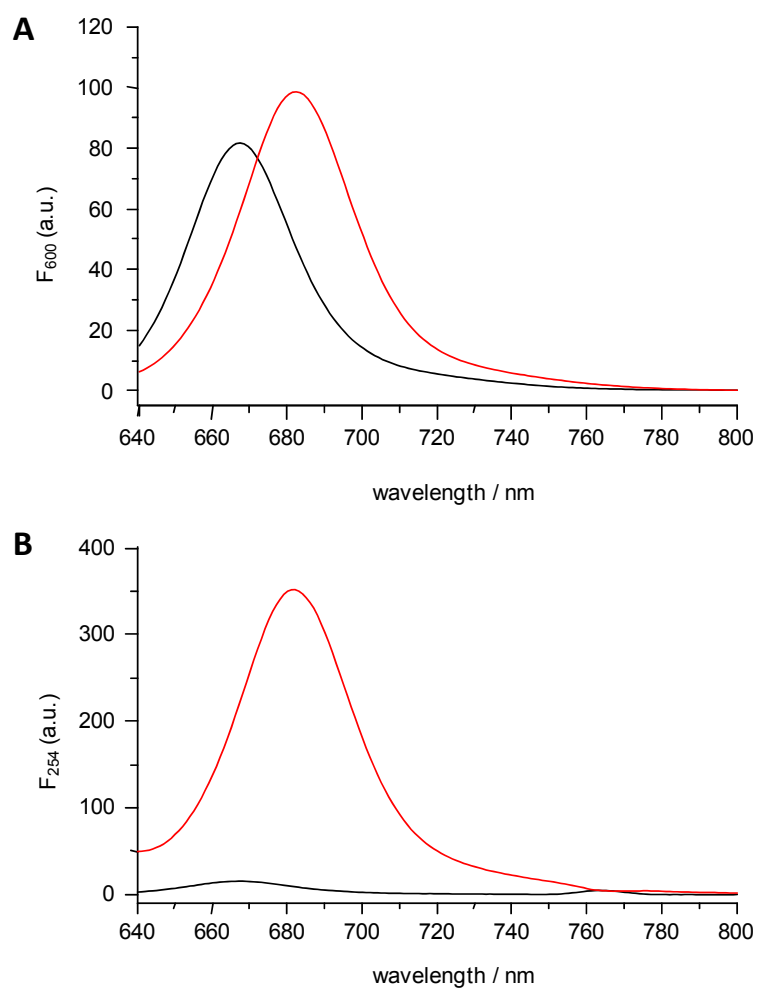
Figure S8. Fluorescence emission of free **DiSC$_2$(5)** (black) and a 1:2 **DiSC$_2$(5)**:*MYC* quadruplex mixture (red) with an excitation wavelength of 600 nm (A) and 254 nm (B); T = 20 °C, c(**DiSC$_2$(5)**) = 3 μM.

█ **Molecular Recognition**

Selective Targeting of G-Quadruplex Structures by a Benzothiazole-Based Binding Motif

Ina Buchholz,[a] Beatrice Karg,[a] Jonathan Dickerhoff,[a] Adrian Sievers-Engler,[b] Michael Lämmerhofer,[b] and Klaus Weisz*[a]

Abstract: A benzothiazole derivative was identified as potent ligand for DNA G-quadruplex structures. Fluorescence titrations revealed selective binding to quadruplexes of different topologies including parallel, antiparallel, and (3+1) hybrid structures. The parallel c-MYC sequence was found to constitute the preferred target with dissociation constants in the micromolar range. Binding of the benzothiazole-based ligand to c-MYC was structurally and thermodynamically characterized in detail by employing a comprehensive set of spectroscopic and calorimetric techniques. Job plot analyses and mass spectral data indicate noncooperative ligand binding to form complexes with 1:1 and 2:1 stoichiometries. Whereas stacking interactions are suggested by optical methods, NMR chemical shift perturbations also indicate significant rearrangements of both 5′- and 3′-flanking sequences upon ligand binding. Additional isothermal calorimetry studies yield a thermodynamic profile of the ligand–quadruplex association and reveal enthalpic contributions to be the major driving force for binding. Structural and thermodynamic information obtained in the present work provides the basis for the rational development of benzothiazole derivatives as promising quadruplex binding agents.

Introduction

Although prevalent in the living cell, DNA duplexes are only one of several secondary structures that may be adopted by nucleic acids.[1] In particular, G-rich sequences also fold into tetrastranded assemblies called G-quadruplexes (G4s) under appropriate conditions. In general, G4 structures comprise a stack of G-tetrads that are composed of a planar arrangement of four hydrogen-bonded guanine bases and additionally stabilized by cations, usually potassium or sodium, located within a central channel of negative potential.[2,3] G4s have attracted rapidly growing interest during the past two decades because of their existence and potential regulatory role in vivo; this makes them promising drug targets, especially in anticancer strategies.[4] Thus, G-rich sequences able to fold into a quadruplex structure are found at the telomere single-stranded ends, but are also over-represented in promoter regions of various oncogenes. G4-selective ligands, with their ability to induce quadruplex formation, may inhibit telomerase action in cancer cells or block expression of specific oncogenes, such as c-MYC.[5–8] However, the quadruplex also constitutes a powerful platform in technological methodologies and has found various applications in aptamer and sensing systems for an increasing number of proteins, metabolites, and metal ions.[9,10]

Apart from favorable chemical and physical properties, therapeutically useful ligands should not only discriminate between G4s and the prevailing genomic double-helical DNA, but also between G4s of different topology and conformation. Such discrimination seems to be feasible because the 3D structure of G4s is far from uniform and depends on the particular sequence, coordinating cations, and environmental conditions. Unfortunately, the type and orientation of loops or the dimension of G4 grooves that vary among different G4s have not yet been fully exploited in G4 recognition. The majority of G4 ligands developed and tested so far have been shown to preferably stack on the G4 outer tetrads with limited discriminatory potential. Thus, G4 binders typically feature extended aromatic surface areas for optimal stacking on a G-tetrad and are often linked with flexible side chains for additional less specific electrostatic and van der Waals interactions.[11,12] Developing a ligand that targets a particular G4 with high affinity and high selectivity is still a challenging task that is also hampered by limited thermodynamic and structural data.[13–15]

Cyanine dyes, such as thiazole orange or oxazole yellow, based on benzothiazole and benzoxazole heterocycles, have been widely used as fluorescent probes for nucleic acids.[16–18] Concomitant with increasing interest in G4 structures, novel benzothiazole or benzoxazole cyanine dyes have been developed for selective G4 detection. Ideally, these should exhibit

[a] I. Buchholz, B. Karg, J. Dickerhoff, Prof. Dr. K. Weisz
Institute of Biochemistry, Ernst-Moritz-Arndt University Greifswald
Felix-Hausdorff-Strasse 4, 17487 Greifswald (Germany)
E-mail: weisz@uni-greifswald.de

[b] A. Sievers-Engler, Prof. Dr. M. Lämmerhofer
Institute of Pharmaceutical Sciences
Eberhard Karls University Tübingen
Auf der Morgenstelle 8, 72076 Tübingen (Germany)

█ Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
http://dx.doi.org/10.1002/chem.201700298.

a combination of high binding affinity, quadruplex selectivity, and significant changes in their fluorescence quantum yield upon complex formation. However, simple cyanine dyes often show indiscriminate affinity for nucleic acid secondary structures and their binding to quadruplex DNA often suffers from complexity and only poorly characterized binding modes. Thus, given the strong propensity of cyanines to aggregate, nucleic acid targets frequently induce additional assembly–disassembly of the bound dye.[19, 20]

Remarkably, G4-selective ligands structurally unrelated to the cyanine dye family, but based on benzo-fused thiazole or oxazole heterocyclic moieties, have rarely been reported.[21] The promising binding properties of the two heterocycles prompted us to explore their use as a structural platform in G4-selective binding motifs. A promising quadruplex-binding 2-aryl-benzothiazole ligand was identified and its binding to various G4 structures was subjected to more detailed structural and thermodynamic characterization. A comprehensive analysis of affinity, selectivity, binding mode, and the thermodynamic profile of association makes use of a diverse array of powerful biophysical methods, including circular dichroism (CD), MS, fluorescence and NMR spectroscopy, and isothermal titration calorimetry (ITC). These provide complementary information on interactions with different G4 structures, and serve as a solid basis for future structural optimizations.

Results and Discussion

Synthesis and spectroscopic characterization of benzothiazole derivative L4

In initial binding assays of 2-aryl-benzothiazoles and 2-aryl-benzoxazoles that have been selected for their availability, drug-like properties, and the presence of functionalities amenable to simple and straightforward derivatization, acryloylfuranyl-substituted benzothiazole L4 not only exhibits significant binding to G4 structures, but also noticeable discrimination against quadruplexes with different topologies (Figure S1 in the Supporting Information). The unfused 2-aryl substituent expands the π-conjugated system and allows for increased stacking interactions with outer G-tetrads, but should be flexible enough to adopt nonplanar conformations for alternative interactions with loops or within the G4 grooves (Figure 1 A).

The L4 derivative was prepared from 3-[5-(1,3-benzothiazol-2-yl)furan-2-yl]prop-2-enoic acid. Amide formation of the acid with the polyamine N,N-bis(3-aminopropyl)methylamine was accomplished through activation as the N-hydroxysuccinimide (NHS) ester. The latter was isolated before reaction with a five-fold excess of polyamine in DMF. The amine side chain attached to the conjugated scaffold for this first-generation ligand is closely related to spermidine and has frequently been employed for ligands interacting with DNA. In addition to possible favorable contributions to the DNA binding properties through amine functionalities with their high proton affinities, it also increases water solubility of the ligand. As an added benefit, the primary amino group may be used to easily tether additional binding motifs in future modifications.
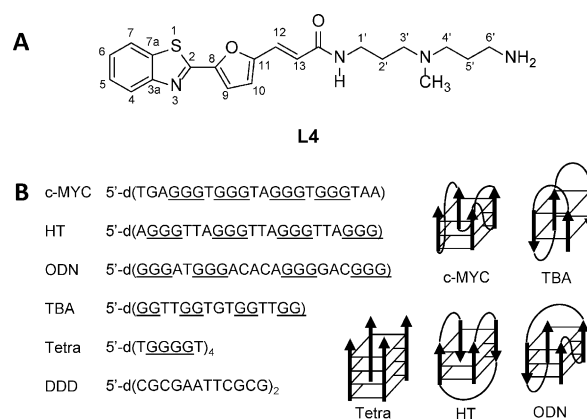
Figure 1. A) Structure of **L4**. B) Sequence and topology of the quadruplexes c-MYC, TBA, Tetra, HT, and ODN, and that of the duplex DDD; guanine bases of G-tetrads are underlined.

The optical properties of the **L4** ligand are summarized in Table 1 (see also Figure S3 in the Supporting Information). Molar absorptivities at the absorption maximum at $\lambda = 368$ nm have been determined from measurements at eight ligand

Table 1. Optical properties of **L4** at 20 °C.

Conditions	Absorption		Fluorescence[b]	
	λ_{max} [nm]	ε [M^{-1} cm^{-1}][a]	λ_{max} [nm]	F [a.u.]
buffer, pH 7.0	368	$18\,364 \pm 294$	429	1.0
DMSO	368	$19\,154 \pm 67$	422	4.8
[a] From the slope of concentration-dependent absorption data; average values with standard deviations from three independent determinations. [b] $c = 3$ μM; $\lambda_{ex} = 365$ nm.				

concentrations and show only small changes upon going from aqueous buffer to a solution in DMSO. Also, the strict linearity of concentration-dependent absorbances over a concentration range of $5\,\mu M < c < 40\,\mu M$ suggest the absence of noticeable aggregation processes in either solvent. In contrast to absorption, the fluorescence intensity of the ligand is heavily influenced by the solvent, with a fluorescence quantum yield increasing by a factor of nearly five upon going from aqueous buffer to DMSO.

Quadruplex-selective L4 binding affinities

To evaluate the binding affinities of **L4** towards G4s with different folds, various well-characterized G4 targets, namely, Tetra, c-MYC, TBA, HT, and ODN, were selected for the binding studies.[22–26] Quadruplex topologies and sequences of all oligonucleotides are shown in Figure 1 B. Initially, CD spectra were acquired that confirmed the expected topologies of the five quadruplexes under the solution conditions employed (see Figure S4 in the Supporting Information). It should be mentioned that spectra were acquired by using potassium buffer, except for HT, which was dissolved in a Na+-containing buffer to ensure folding into an antiparallel topology. Thus, the two

parallel c-MYC and Tetra quadruplexes reveal negative and positive amplitudes centered at $\lambda = 243$ and 264 nm, respectively. In contrast, two maxima at $\lambda \approx 245$ and 294 nm and a minimum at $\lambda \approx 265$ nm identify the antiparallel topology adopted by telomeric HT and TBA quadruplexes, whereas the ODN sequence folds into a $(3+1)$ hybrid structure, in line with a negative amplitude at $\lambda = 239$ nm and two positive amplitudes at $\lambda = 264$ and 287 nm.[27,28] Next, the impact of ligand binding on the G4 structures was examined by ligand titrations. Adding L4 in increasing amounts has only small effects on the G4 CD spectral signatures and only changes amplitudes to a minor extent, even with the ligand in a tenfold excess (Figure S4 in the Supporting Information). Consequently, ligand L4 does not significantly perturb the quadruplex conformation and all G4s preserve their native topology. However, at high ligand concentrations, very weak induced CD (ICD) effects of the ligand are noticeable for TBA and ODN in the L4 absorption range between $\lambda = 325$ and 410 nm; this demonstrates quadruplex binding of the ligand. Typically, the ICD signal is expected to be small for intercalation or outer stacking and contrasts with groove binding, which often induces a positive CD of larger magnitude with the same shape as that of the absorption band.[29] Consequently, the low-amplitude positive ICDs may indicate an outer stacking binding mode. It should be emphasized that, due to their dependence on the relative orientation of transition dipole moments, and thus, on the L4 orientation, the absence of corresponding ICD effects for the other G4 structures does not exclude high-affinity binding of the ligand.

To determine binding affinities, the ligand was titrated with the quadruplexes and its fluorescence was followed during the course of the titration. As shown in Figure 2, the addition of nucleic acid results in quenching of the L4 fluorescence to various degrees, depending on the particular quadruplex. Upon the addition of ODN, and especially c-MYC, a faster decline towards a limiting L4 fluorescence associated with binding-site saturation is clearly apparent relative to the other G4 structures. This already indicates widely differing binding affinities, of which c-MYC is the preferred target. To evaluate binding to single- and double-stranded DNA, titrations were also performed with the Dickerson–Drew dodecamer duplex DDD and
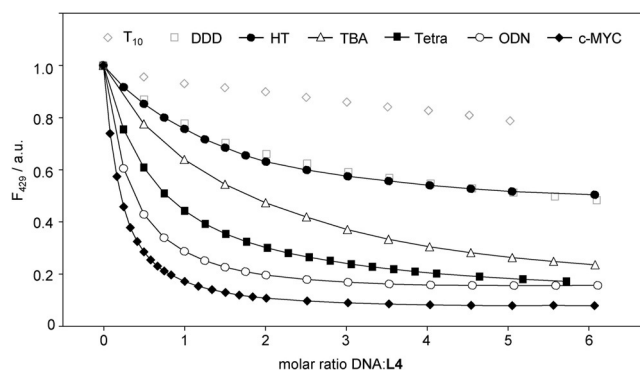
a single-stranded T_{10} decamer. Whereas the duplex titration data suggest weak to moderate L4 binding, only a gradual linear decrease of fluorescence was observed in the case of T_{10}; this indicates some nonspecific electrostatic interactions with cationic L4.

Hyperbolic binding isotherms were fitted based on a model with equivalent and independent binding sites and with limiting fluorescence intensity, binding stoichiometry (n), and association constant (K_a) as free floating parameters. However, with $c_{L4}K_a < 10$, and incomplete saturation at the end of titration, the stoichiometry n and K_a are increasingly interdependent and cannot be separately obtained with sufficient accuracy in the nonlinear least-squares fit of most binding isotherms.[30] On the other hand, the product $nK_a = K_a^{app}$ is reasonably well defined by the isotherms under the present conditions. K_a^{app} allows for a convenient comparison of quadruplex affinities, but with the true number of individual binding sites undetermined. As plotted in Figure 3, K_a^{app} values of L4 range from $1.1 \times 10^5 \, \text{M}^{-1}$ for TBA to $1.1 \times 10^6 \, \text{M}^{-1}$ for the c-MYC target; the
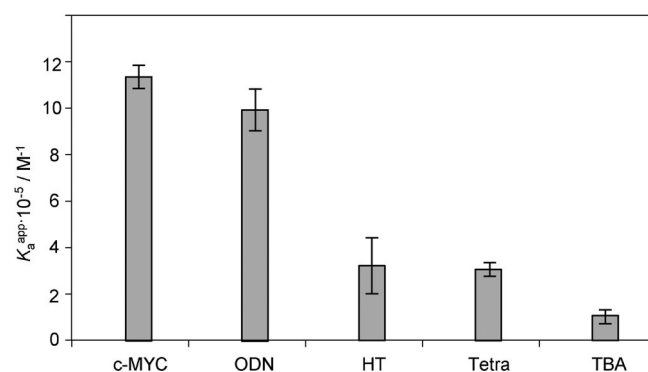
Figure 3. Quadruplex binding constants $K_a^{app} = nK_a$, as determined by fluorescence titrations of ligand L4 with quadruplex in potassium phosphate or sodium phosphate buffer (HT), pH 7, at 20°C; average values from ≥ 3 independent experiments are plotted and standard deviations are indicated by error bars.

latter is in a range reported for the G4 binding of many typical quadruplex ligands.[31–33] Binding to c-MYC is also accompanied by the most pronounced change in L4 fluorescence. Whereas a high L4 affinity is also found for hybrid-type ODN, both HT and Tetra quadruplexes bind L4 with only modest affinity. Due to a weak L4 association, no reliable fits were obtained for single- or double-stranded DNA. However, with an estimated K_a^{app} of 10^4–$10^5 \, \text{M}^{-1}$ for DDD, duplex affinities are one to two orders of magnitude lower than those of the preferred c-MYC target.

Stoichiometry of L4 binding to c-MYC

Given the favored L4 binding to the c-MYC quadruplex and its biological significance as a potential target, we focused on the c-MYC sequence for additional studies on the number and affinity of potential binding sites. However, the artificially designed ODN quadruplex, with its $(3+1)$-hybrid structure and

Figure 2. Fluorescence titrations of L4 (6 μM) with different quadruplexes, duplex DDD, and single-stranded T_{10} in potassium phosphate or sodium phosphate buffer (HT), pH 7; $T = 20°C$, $\lambda_{ex} = 365$ nm.

long diagonal loop, constitutes a G4 target for **L4** that is only slightly less preferred than that of c-MYC.

Because unambiguous binding stoichiometries were not accessible by the fluorescence titration experiments, the method of continuous variation was employed. By mixing the ligand and quadruplex in different molar ratios, while keeping the total concentration constant, a Job plot was constructed (Figure 4). Here, differences in the fluorescence of **L4** between different mixtures and corresponding quadruplex-free solutions are plotted against the molar ratio of **L4** to c-MYC. Two independent measurements yielded an intersection of the two linear regions at ratios of 0.63 and 0.67, in agreement with a 2:1 stoichiometry of ligand–quadruplex.
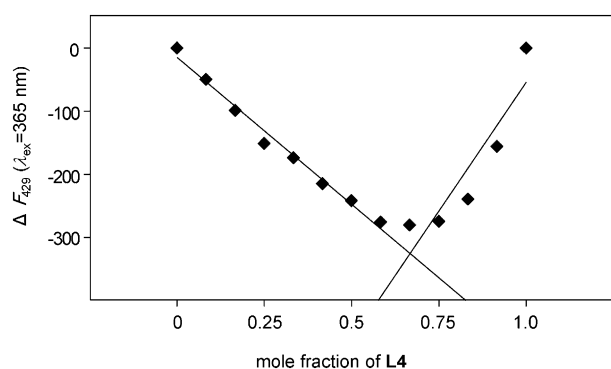
Figure 4. Job plot showing differences in **L4** fluorescence at $\lambda = 429$ nm between mixtures of **L4** with c-MYC at different molar fractions and **L4** without DNA in potassium phosphate buffer at 20 °C. The sum of ligand and quadruplex concentration was fixed to 6 μM. The intersection of the two linear regression lines suggests 2:1 stoichiometry for the ligand–quadruplex.

Because multiple binding sites of the asymmetric c-MYC quadruplex cannot be equal, the question arises if the sites exhibit noticeable differences in their binding affinity. To elaborate on this issue, we also performed a reverse titration, by titrating the quadruplex with aliquots of **L4** (Figure S5 in the Supporting Information). Such a reverse titration is made possible through the intrinsic G4 fluorescence of c-MYC in the range $\lambda = 325$–360 nm when excited at $\lambda = 260$ nm.[34] In contrast to single- and double-stranded DNA, the guanine arrangement in quadruplexes results in enhanced fluorescence, which gradually decreases upon adding the ligand. The combination of forward and reverse titrations gives a more comprehensive picture of binding due to the different initial conditions. Disregarding cooperative binding (see below), changes in fluorescence at the start of the reverse titrations are dominated by the formation of 1:1 complexes and by ligand binding to high-affinity sites on the free quadruplex being in excess. On the other hand, with the ligand in excess, the initial fluorescence upon titrating the quadruplex to the ligand will mostly respond to the formation of higher order complexes that also involve low-affinity sites. Consequently, results obtained for the forward and reverse titrations will only match for stoichiometries of >1 if available binding sites are of similar affinity and trigger similar optical responses. Fitting the data of the reverse titration with a model of equivalent and independent binding sites enabled the separate determination

of $K_a = 6 \times 10^5$ M^{-1} and the number of binding sites $n = 2.4$. These values are in good agreement with previously determined parameters. Given some uncertainties due to a small but non-negligible absorbance of **L4** at the wavelength of G4 excitation, comparable binding constants determined from forward and reverse titrations thus suggest binding sites of broadly similar affinity.

We complemented optical studies by also following changes in ligand absorption upon c-MYC addition (Figure S6 in the Supporting Information). Inspection of the UV/Vis spectra reveals hypso- and bathochromic effects, in agreement with a stacked ligand and an isosbestic point at $\lambda = 395$ nm that is increasingly blurred for later titration steps. These fully reproducible results demonstrate the presence of more than two spectrally distinct species as a result of inequivalent binding sites, but may also indicate additional weak and nonspecific association, mostly driven by electrostatic forces between negatively charged nucleic acid and cationic ligand (see below).

In contrast to optical methods, mass spectrometry allows the separate detection of individual complexes if they are stable enough during the ionization process (here electrospray ionization (ESI)) and can be transferred into the gas phase.[35, 36] To reduce counterion condensation on the nucleic acid upon evaporation in the presence of sodium or potassium ions, DNA and ligand were dissolved for ESI-TOF MS analysis in 150 mM ammonium acetate buffer with the addition of 10 % methanol. The amount of methanol additive ensured a stable signal without changing buffer conditions to a more significant extent. Initial CD measurements on c-MYC in the ammonium acetate/ MeOH buffer confirmed the formation of a parallel quadruplex. ESI-TOF mass spectra of solutions of c-MYC–**L4** are shown in Figure 5 and indicate two predominant charge states of $z = 4-$ and $5-$.

In accordance with cation coordination between G-tetrads, signals at m/z 1403.4 and 1754.6 identify ions [c-MYC]$^{5-}$ and [c-MYC]$^{4-}$, respectively, with two tightly bound ammonium ions, as expected for a conserved quadruplex fold. With one equivalent of the ligand in the spraying solution, [1:1] complex ions appear at m/z 1483.1 ($z = 5-$) and 1854.1 ($z = 4-$), but signals of free quadruplexes still dominate. The addition of more ligand to yield a molar ratio of quadruplex to **L4** of 1:2 results in a further increase of the [1:1] complex ion and the appearance of additional ions of [1:2] complexes with m/z 1562.7 ($z = 5-$) and 1953.6 ($z = 4-$). The 1:1 and 1:2 complexes further grow in intensity at the expense of free quadruplex with an increasing molar ratio of ligand to quadruplex. The intensity distribution of ions at different molar ratios clearly suggests noncooperative **L4** binding to two binding sites. Although different buffer conditions were employed, the mass spectral data corroborated results from the optical measurements in identifying two independent **L4** binding sites on the c-MYC quadruplex.

Interacting sites of L4 and c-MYC

To obtain additional information on the **L4** binding mode, NMR spectra of the c-MYC quadruplex were acquired in the
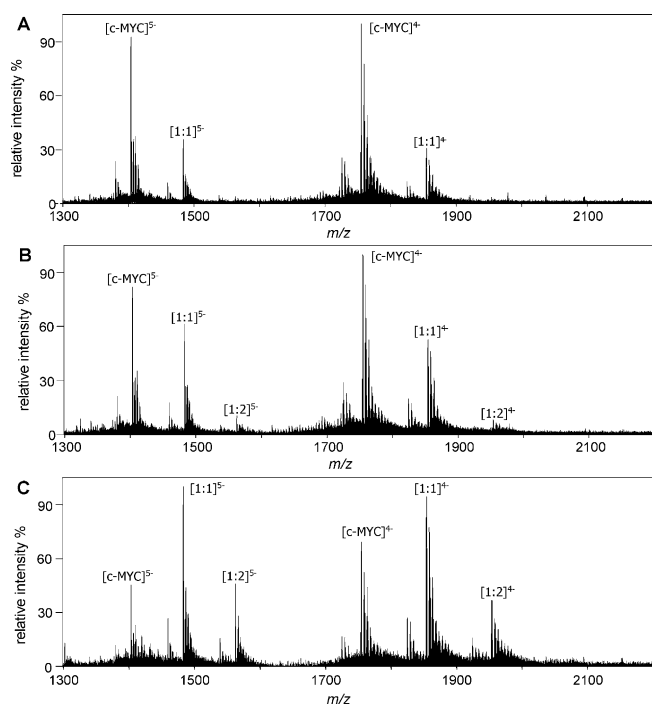
Figure 5. ESI-TOF mass spectra of c-MYC in the presence of **L4** with molar ratios of 1:1 (A), 1:2 (B), and 1:3 (C) in 150 mM ammonium acetate/10% methanol.

absence and presence of the ligand. The imino proton spectral region of 1D ^{1}H NMR spectra are shown in Figure 6 at 40 °C. The free quadruplex exhibits twelve well-resolved guanine imino signals engaged in hydrogen bonding between $\delta = 10.5$ and 12.0 ppm; this demonstrates the formation of a quadruplex structure composed of three G-tetrads under the solution conditions employed.

Adding increasing amounts of **L4** leads to a moderate broadening of the signals, but no additional resonances are observed. This indicates intermediate to fast exchange process-
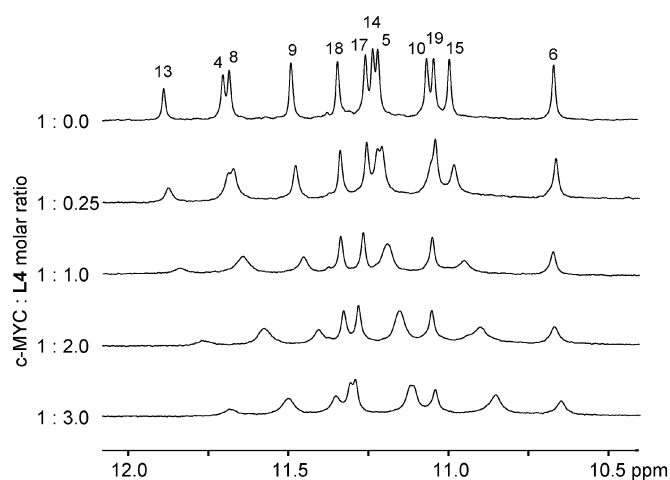
Figure 6. Imino proton spectral region of 0.45 mM c-MYC quadruplex with increasing amounts of **L4** in 10 mM potassium phosphate, pH 7, at 40 °C. Assignments to G residues are given for the free c-MYC.

es between free and bound c-MYC, in agreement with a sharpening of signals due to faster exchange at elevated temperatures. Most imino resonances are gradually shifted upfield during titration to a varying extent. Upfield shifts of G4 imino protons upon binding have frequently been observed and can be attributed to ring-current effects of the aromatic ligand stacked on the outer G-tetrad.[37,38] Notably, saturation is not yet reached, even after the addition of ligand in a threefold excess; this possibly also indicates some ligand-mediated aggregation processes.

For the resonance assignments, standard NMR spectroscopy methodologies were employed on a mixture of c-MYC–**L4** in a 1:3 molar ratio.[39] In particular, 2D NOE spectra enabled the complete assignment of H6/H8 base and H1′/H2′/H2″ sugar protons in the complex, but not those of residue A22 at the 3′-terminus (see Tables S1 and S2 in the Supporting Information). Thus, intranucleotide and sequential H6/H8–H1′ and H6/H8–H2′/H2″ NOE contacts allow for a continuous walk along the sequence starting from T1 of the 5′-terminus and only interrupted by the three propeller loops (see Figure S8 in the Supporting Information). Unambiguous discrimination of the second and third G-tracts was possible through connectivities observed between H1′ of residue G10 and A12 H2 within the second propeller loop. Imino protons were identified by H1–H8 contacts detected within G-tetrads and H1–H1 contacts along the G-tracts. Ambiguities in a few assignments were resolved by direct comparison with assignments made for the free quadruplex, with its largely similar NOE patterns. For the latter, proton chemical shifts were previously published and only some small deviations, in particular, those apparent at the 5′- and 3′-overhang, resulted from different temperature and solution conditions.[23]

Unfortunately, hampered by dynamic processes, no intermolecular NOE contacts between **L4** and c-MYC protons were detected. Nevertheless to localize c-MYC–**L4** interactions within the complex, chemical shift differences of base and sugar protons between the complex and free quadruplex were determined. Plotted as a function of residue, these chemical shift footprints identify sites considerably perturbed by the bound ligand and are likely to be located near the ligand-binding site. Chemical shift changes along the G4 sequence, involving nonexchangeable base and sugar protons, as well as imino protons within the G-tetrads, are shown in Figure 7 and Figure S9 in the Supporting Information, respectively. Conspicuously, 5′- and 3′-overhang residues experience the largest perturbations, whereas loop residues are only affected to a small extent (Figure 7). Significant deshielding effects are observed for the H8 and H1′ protons of A3 flanking the 5′-tetrad, and for the methyl resonance of T20 flanking the 3′-tetrad. The last of these initially broadens and considerably shifts downfield by 0.17 ppm (Figure S7 in the Supporting Information). Chemical shifts of nonexchangeable protons within the four G-tracts are moderately affected by **L4** binding with changes of $\Delta\delta \leq$ 0.1 ppm. Given a largely unaffected TA propeller loop that shields the groove between the second and third G-tracts, more noticeable perturbations along G8-G9-G10 and G13-G14-G15 must result from conformational readjustments, rather
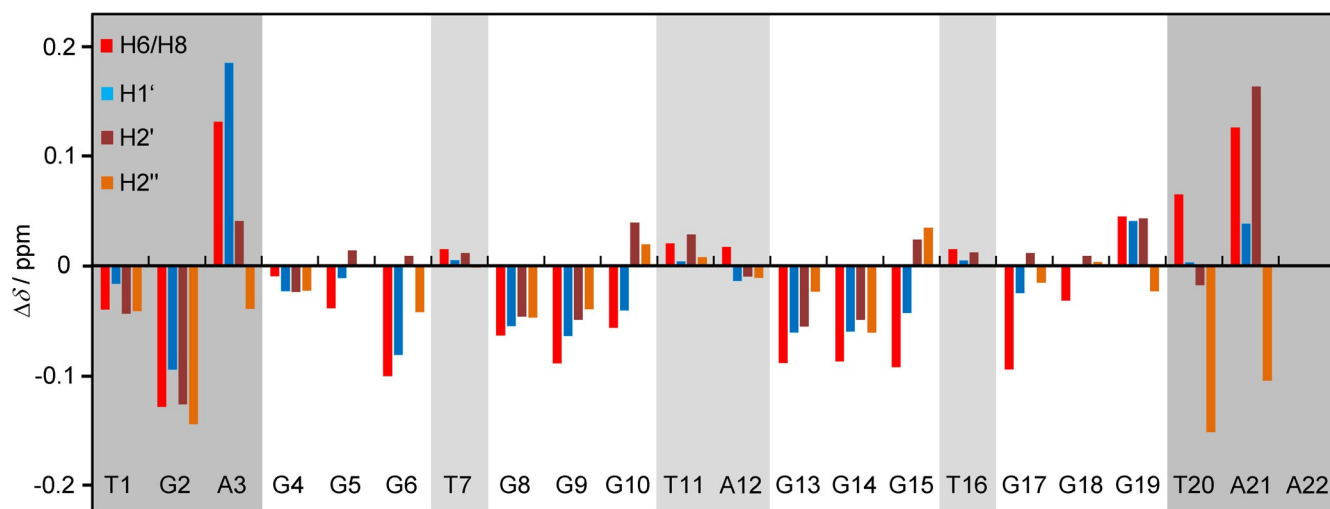
Figure 7. Chemical shift perturbations $\Delta\delta = \delta(\text{complex}) - \delta(\text{c-MYC})$ for H6/H8, H1′, H2′, and H2″ protons along the c-MYC sequence upon the addition of **L4** (3 equiv). Overhang and loop residues are highlighted in dark and light gray, respectively.

than direct ligand contacts. Except for G17, major chemical shift changes of $\Delta\delta \approx 0.2$ ppm are also observed for the imino protons of the 5′-tetrad and indicate their contact to bound **L4** (Figure S9 in the Supporting Information). Whereas imino groups of the central tetrad are moderately affected through mediated long-range effects, only G15 exhibits a significant upfield shift within the 3′-tetrad.

In the following, the impact of **L4** binding to the quadruplex proton chemical shifts was also mapped on a surface model of the c-MYC structure (PDB 1XAV; Figure 8). Changes in chemical shifts for imino, H6/H8, and H1′/H2′/H2″ protons were separately assigned for each type of proton to five classes with different gradations of red color. To ensure unobstructed views of the outer tetrads, both 5′-TGA and TAA-3′ overhangs were removed. As seen in Figure 8, the deeply red-colored 5′-tetrad is strongly affected by the ligand. On the other hand, smaller disconnected areas of chemical shift perturbations are observed at the 3′-terminus and only minor effects of **L4** are seen for the groove and loop regions. These results suggest that the furylbenzothiazole moiety of the ligand preferentially stacks on the G4 5′-tetrad. Fewer and/or weaker contacts with the outer quartet at the 3′-end, mostly involving G15, are indicated by the G imino chemical shift footprints (Figure S9 in the Support-

ing Information). The positively charged aminoalkyl side chain of **L4** may form additional contacts within the freely accessible groove or with loop and overhang nucleotides to support binding on the outer tetrads. However, its impact on quadruplex chemical shifts seems to be small due to its high flexibility, and no further information on corresponding interactions can be obtained based on chemical shift changes.

As shown by the surface model in Figure 8, furanose sugars at the 3′-tetrad protrude beyond the plane and may hamper putative stacking interactions with the ligand. The higher affinity of the better accessible 5′-tetrad is not without precedence and has also been observed for other ligands binding to the c-MYC quadruplex.[40,41] In further analogy to these studies, the strongly impacted overhangs may also indicate large conformational readjustments with the possible formation of a ligand-induced binding cavity. It should be cautioned, however, that large changes in proton chemical shifts upon ligand binding do not necessarily correlate with thermodynamic association constants and only indicate significant ligand-induced perturbations. Thus, whereas chemical shift footprints are very valuable for an initial localization of bound ligand, they do not provide for the unambiguous identification of high- and low-affinity binding sites, especially when differences in microscopic association constants are only moderate.

Thermodynamic signature of binding

A more detailed understanding of the thermodynamic contributions to binding is a prerequisite for the rational structural optimization of ligands. Consequently, additional isothermal calorimetry studies were performed to measure the enthalpy of **L4** binding to the c-MYC quadruplex directly. By using a model-free approach, standard molar enthalpies for the high-affinity binding sites were determined by an excess-site method, yielding $\Delta H^o = (-5.79 \pm 0.25)$ kcal mol^{-1} from three independent measurements (Figure 9A). With a standard free energy of -7.77 kcal mol^{-1} per binding site, as independently

Figure 8. Proton chemical shift perturbations mapped with red indicating variable intensity on a surface model of c-MYC (PDB 1XAV). View of the 5′- (left) and 3′-tetrad (right). Arrows indicate the freely accessible groove.
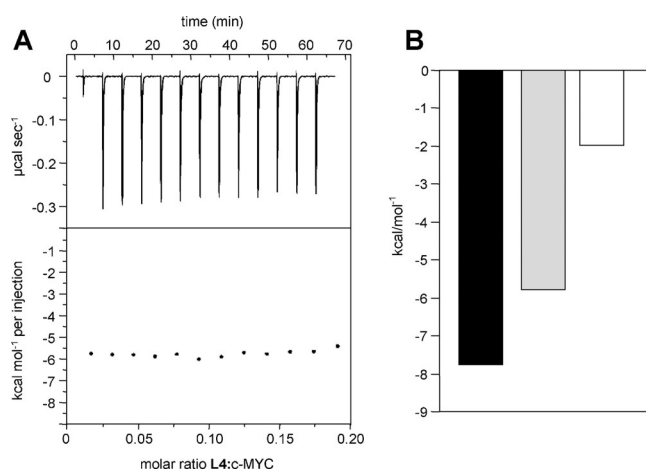
Figure 9. A) Model-independent excess-site ITC titration of **L4** to the c-MYC quadruplex. The upper and lower panels show the heat burst for every injection step and the corresponding normalized heat signal versus molar ratio, respectively. B) Thermodynamic binding profile of **L4** with $\Delta G°$ (black bar), $\Delta H°$ (gray bar), and $-T\Delta S°$ (white bar). Parameters were calculated from standard thermodynamic relationships: $\Delta G° = -RT\ln K_a$ and $\Delta G° = \Delta H° - T\Delta S°$.

determined from fluorescence titration experiments (see above), an entropic contribution of $-T\Delta S° = -1.98$ kcal mol^{-1} can be calculated. Whereas both enthalpic and entropic effects favor binding, it is the enthalpic contribution that mostly drives **L4** association with the c-MYC quadruplex (Figure 9B). This is in contrast with previous ITC studies on the binding of indolenine cyanine dyes to quadruplexes.[42] These have shown that the interaction of the cyanines is predominantly entropy driven, which indicates a significant impact from hydrophobic effects on complex formation. On the other hand, stacking, dipole–dipole, hydrogen-bond, and electrostatic interactions are suggested to considerably contribute to binding of the benzothiazole derivative, with its relatively small hydrophobic surface area.

We also performed a complete ITC titration for a comparison with isotherms acquired by optical methods. Interestingly, ITC titration curves revealed additional low-affinity binding, probably through nonspecific electrostatic interactions at the quadruplex outer surface (Figure S10 in the Supporting Information). Without any partial fixing of parameters, attempts to fit the experimental data based on a model with inequivalent binding sites resulted in only poorly defined fit values. However, fitting the isotherms with fixed stoichiometry and standard molar enthalpy, as determined previously for the specific high-affinity binding, leaves a reduced set of adjustable parameters and enables their more accurate determination. Thus, in addition to parameters for the low-affinity binding with K_a values in the order of 10^4 M^{-1}, a corresponding fit yields $K_a = 6.5 \times 10^5$ M^{-1} for each of the two high-affinity binding sites. This is in excellent agreement with the values obtained from the fluorescence titrations, and demonstrates the reliability of the values extracted and consistency among the different methods for the present binding event.

Conclusion

Cyanine dyes with benzothiazole moieties have found widespread use as optical probes due to their often favorable photophysical properties. Binding G4s with high affinity, but little discriminatory power, thiazole orange has been employed in the past for ligand displacement studies on various G4s.[43, 44] However, benzothiazole compounds not belonging to the cyanine dye family have rarely been examined in more detail as G4 binding agents. We have therefore derivatized a planar furyl-substituted benzothiazole with an extended π-conjugated system and examined its binding to various quadruplexes of different topologies. With a difference in K_a^{app} of up to one order of magnitude, the ligand exhibits significant discriminatory potential for different quadruplex folds, binding the favored c-MYC target with micromolar affinity. At the same time, moderate to good selectivities with respect to double-helical DNA and K_a^{app}(c-MYC)/K_a^{app}(duplex) ratios of 10–100 are estimated from titration experiments. Clearly, despite these promising binding properties, the selectivities of this first-generation ligand have to be improved for advanced applications.

Further developments of G4 ligands on a rational basis will strongly benefit from a better understanding of the modes of interaction and thermodynamic driving forces for binding. Although dynamic exchange processes of the ligand between different binding sites prevent the determination of a high-resolution structure, the present studies not only disclose enthalpic and entropic contributions to binding, but also indicate binding of the benzothiazole-based ligand through partial stacking on the outer G-tetrads. The cationic amine side chain is expected to promote binding through electrostatic interactions with phosphates of the loop or within the groove. Interactions other than stacking forces are also suggested by the thermodynamic signature that indicates a predominantly favorable enthalpy in driving ligand association with parallel c-MYC. Enforcing entropic contributions through an extension of nonpolar surface areas associated with enhanced hydrophobic effects may thus be a viable strategy in the future to further increase affinities. However, selectivity should benefit from more specific interactions, such as those of hydrogen bonding, π stacking, or van der Waals. Corresponding modifications are easily introduced through altered side chains and through coupling with other G4-binding motifs for possible synergistic effects. The easy availability of the present benzothiazole structure combined with its straightforward derivatization should represent key advantages for its further development to efficient G4-binding ligands.

Experimental Section

General

3-[5-(1,3-Benzothiazol-2-yl)furan-2-yl]prop-2-enoic acid and *N,N*-bis(3-aminopropyl)methylamine were purchased from Otava (Kyiv, Ukraine) and TCI (Eschborn, Germany), respectively. All other reagents and solvents were purchased from Sigma–Aldrich (Steinheim, Germany) and Euriso-Top (Saarbrücken, Germany). Solvents

were distilled under reduced pressure prior to use. The progress of reactions was monitored by TLC on plates precoated with silica gel $60F_{254}$ (Merck).

Synthesis of L4

3-[5-(1,3-Benzothiazol-2-yl)furan-2-yl]prop-2-enoic acid (80 mg, 0.295 mmol) was dissolved together with NHS (91.7 mg, 0.797 mmol) and N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (DCI; 152.8 mg, 0.797 mmol) in dry DMF (7 mL). The mixture was stirred in the dark at room temperature under an argon atmosphere and the progress of the reaction was followed by TLC (CH_2Cl_2-MeOH 9:1). After 22 h, the solvent was removed under reduced pressure. The NHS ester was dissolved in CH_2Cl_2 and washed three times with distilled water. The organic phase was dried over Na_2SO_4 and the solvent removed (yield 95%).

The NHS ester intermediate (110.0 mg, 0.28 mmol) was dissolved in DMF and added dropwise to a mixture of N,N-bis(3-aminopropyl)-methylamine (239.4 μL, 1.495 mmol) and DMF (total volume 3 mL). The mixture was stirred in the dark at room temperature, while being monitored by TLC (CH_2Cl_2/MeOH 7:3, 10% NH_3). After 3 h, the solvent was removed under reduced pressure and the product dissolved in CH_2Cl_2. It was washed three times with distilled water, twice with a saturated solution of NaCl, and dried over Na_2SO_4. Amide **L4** was finally purified by column chromatography on silica gel (CH_2Cl_2/MeOH 10:0 to 7:3 followed by 0 to 10% NH_3). Total yield 47%; MW 400.54 g mol^{-1}; $R_f = 0.50$ (CH_2Cl_2/MeOH 7:3, 10% NH_3); ^{1}H NMR (600 MHz, 298 K, [D_6]DMSO): $\delta = 8.33$ (t, 5.5 Hz, 1H; amide NH), 8.18 (d, 7.8 Hz, 1H; H-4), 8.05 (d, 7.8 Hz; 1H, H-7), 7.57 (t, 7.8 Hz, 1H; H-6), 7.49 (t, 7.8 Hz, 1H; H-5), 7.46 (d, 3.6 Hz, 1H; H-9), 7.32 (d, 15.6 Hz, 1H; H-12), 7.05 (d, 3.6 Hz, 1H; H-10), 6.68 (d, 15.6 Hz, 1H; H-13), 3.20 (quart, 6.5 Hz, 2H; H-1′), 2.58 (br, 2H; H-6′), 2.30 (t, 7.0 Hz, 4H; H-3′, H-4′), 2.12 (s, 3H; N-CH$_3$), 1.59 (quin, 7.0 Hz, 2H; H-2′), 1.50 ppm (quin, 7.0 Hz, 2H; H-5′); ^{13}C NMR (150 MHz, 298 K, [D_6]DMSO): $\delta = 164.3$ (C=O), 153.4 (C-3a), 153.3 (C-11), 148.4 (C-8), 133.9 (C-7a), 126.9 (C-6), 125.7 (C-5), 124.9 (C-12), 122.7 (C-7), 122.4 (C-4), 122.2 (C-13), 116.1 (C-10), 114.4 (C-9), 54.9 (C-3′, C-4′), 41.8 (N-CH$_3$), 39.7 (C-6′), 37.2 (C-1′), 30.1 (C-5′), 26.9 ppm (C-2′).

Sample preparation

DNA oligonucleotides were purchased from TIB MOLBIOL (Berlin, Germany). Before use, oligonucleotides were precipitated in ethanol and the concentrations were determined spectrophotometrically by measuring the absorbance at $\lambda = 260$ nm. Optical and calorimetric measurements were performed in a buffer with 20 mM potassium phosphate, 100 mM KCl, pH 7.0. For the telomeric quadruplex HT, a buffer with 20 mM sodium phosphate, 100 mM NaCl, pH 7.0, was used. Prior to measurements, quadruplex samples were annealed by heating to 90 °C for 15 min followed by slow cooling (10 h) to 5 °C. For the **L4** ligand, a concentrated stock solution in DMSO was employed. Generally, the concentration of samples in DMSO for optical measurements was < 1%.

Circular dichroism (CD)

CD spectra were acquired with a Jasco J-810 spectropolarimeter equipped with a thermoelectrically controlled cell holder (Jasco, Tokyo, Japan). Measurements were performed at 20 °C in 1 cm quartz cuvettes on a 5 μM solution of oligonucleotide without or with ligand in a molar ratio of 1:1, 1:5, or 1:10. Spectra were recorded from $\lambda = 230$ to 450 nm with a bandwidth of 1 nm, a re-

sponse time of 1 s, a scanning speed of 50 nm min^{-1}, and 5 accumulations. All spectra were blank-corrected.

CD melting curves were acquired over a temperature range from 20 to 90 °C with a bandwidth of 1 nm, a response time of 4 s, a heating rate of 0.5 °C min^{-1}, and data points every 0.5 °C. By using 1 cm quartz cuvettes, the concentrations of quadruplex and ligand were 5 and 25 μM, respectively. The wavelength employed for recording ellipticities was $\lambda = 264$ nm for c-MYC and $\lambda = 289$ nm for ODN. The intersection of the melting curve with the median of the fitted baselines was used to determine the melting point, T_m, in three independent measurements.

Fluorescence measurements

Fluorescence measurements were performed with a Jasco FP-6500 spectrofluorometer thermostated with a Peltier unit (Jasco, Tokyo, Japan). By using 1 cm quartz cuvettes, ligand emission spectra were acquired from $\lambda = 375$ to 550 nm with an excitation wavelength of $\lambda = 365$ nm, a scanning speed of 100 nm min^{-1}, emission and excitation bandwidths of 5 nm, and a response time of 1 s. For fluorescence titrations, aliquots of the quadruplex were added stepwise to a 6 μM solution of the ligand. Emission spectra of the G4 were acquired from $\lambda = 300$ to 400 nm with a scanning speed of 50 nm min^{-1}, an excitation wavelength of $\lambda = 260$ nm, emission and excitation bandwidths of 10 nm, and the response time set to 2 s. In the reverse titrations, a solution of c-MYC (6 μM) was titrated with aliquots of **L4**. Emission spectra were recorded after a 5 min equilibration period in each case. Volume-corrected fluorescence intensities at $\lambda = 429$ nm for **L4** and $\lambda = 330$ nm for c-MYC in reverse titrations were plotted over the molar ratio and fitted with an appropriate equation based on equivalent binding sites.

Job plots were performed at 20 °C by varying both the ligand and quadruplex concentration, but keeping their sum fixed at 6 μM. Different volumes of equimolar stock solutions of each binding partner were mixed to give different molar fractions of ligand ranging from zero to one. A blank measurement was performed by replacing the quadruplex solution with buffer. The difference in fluorescence between sample and blank at $\lambda = 429$ nm for **L4** was plotted against the molar fraction of ligand. The stoichiometry of the ligand–quadruplex interaction was obtained by the intersection of the two linear regression lines.

UV/Vis spectroscopy

UV/Vis absorbances were measured with a Jasco V-650 spectrophotometer (Jasco, Tokyo, Japan). The temperature was controlled by a Peltier temperature control unit. For titration experiments, aliquots of G4 were added stepwise to a 6 μM solution of the ligand in buffer in 1 cm quartz cuvettes at 20 °C. Absorption spectra were recorded at a scanning speed of 100 nm min^{-1} after a 5 min equilibration period in the range $\lambda = 300$ to 450 nm with a bandwidth of 1 nm for each titration step.

ESI-MS

ESI-TOF MS experiments were performed on a quadrupole time-of-flight instrument (TripleTOF5600+, Sciex, Ontario, Canada) by direct infusion. G4 samples were measured in negative-ion mode with source and collision voltages set to −3800 and −10 V, respectively. The source temperature was screened from RT to 180 °C; the latter was used without significant deterioration of spectrum quality. The radiofrequency (rf) transmission window was set according to the m/z of the analyte. The infusion was performed by using a syringe at 20 μL min^{-1}. Other optimized settings include a curtain

gas of 16 psi, nebulizer and drying gas of both 17 psi, and a declustering potential of -20 V. Acquisition was performed with an accumulation time of 2000 ms in multiple channel analyzer (MCA) mode. Samples were prepared at a concentration of 2 μM HPLC-purified oligonucleotide in 150 mM ammonium acetate buffer (pH 7.0) containing 10% methanol and varying amounts of the **L4** ligand.

NMR spectroscopy

NMR spectra were acquired on a Bruker Avance 600 MHz spectrometer equipped with an inverse $^1H/^{13}C/^{15}N/^{31}P$ quadruple resonance cryoprobe head and z-field gradients. For all NMR spectroscopy measurements on the quadruplex and quadruplex–ligand mixtures, a low-salt buffer with 10 mM potassium phosphate, pH 7.0, was used. Titrating the ligand in a concentrated solution in DMSO to the quadruplex resulted in a final DMSO content of about 7% for **L4**–quadruplex mixtures with a molar ratio of 3:1. Data were processed by using Topspin 3.1 software and analyzed with the CcpNmr Analysis program.[45] Proton chemical shifts were referenced relative to the temperature-dependent chemical shift of H_2O ($\delta_H = 4.62$ ppm at 40 °C). For water suppression in 90% H_2O/10% D_2O, a WATERGATE w5 pulse sequence with gradients was used. NOESY experiments for free c-MYC and c-MYC in the presence of **L4** were performed at 40 °C with a mixing time of 300 ms and a spectral width of 10 kHz. 4 K×1 K data points with 24 transients per t_1 increment and a recycle delay of 2 s were collected in t_2 and t_1. Prior to Fourier transformation, data were zero-filled to give a 4 K×2 K matrix and both dimensions were apodized by squared sine bell window functions.

Isothermal titration calorimetry (ITC)

ITC experiments were performed with a MicroCal PEAQ-ITC instrument (Malvern Instruments, UK). Samples were dissolved in buffer (pH 7.0) containing 5% DMSO. Titrations were performed at 20 °C with a reference power of 5 μcal s^{-1} and a delay between injections of 240 s. The first injection volume was rejected before data analysis. Subsequent titration steps involved injections with 1.5 μL each of a 999 μM solution of ligand to 20 μM quadruplex. For the direct determination of binding enthalpies by the excess-site method, buffer solutions of 200 μM quadruplex and 200 μM ligand were employed with a delay between injections of 300 s and injection volumes of 3 μL.

Blank titrations for the correction of the titration curves were conducted by injecting the ligand into buffer under otherwise identical experimental conditions. In the case of the model-independent protocol, the molar binding enthalpy, $\Delta H°$, was directly determined by signal integration of the power output following each injection, normalization by the number of moles of added ligand, and averaging from three independent experiments with 12 titration steps each. The final data were analyzed with Origin software.

Conflict of interest

The authors declare no conflict of interest.

Keywords: G-quadruplexes · ligand effects · molecular recognition · NMR spectroscopy · noncovalent interactions

[1] J. Choi, T. Majima, *Chem. Soc. Rev.* **2011**, *40*, 5893–5909.
[2] S. M. Kerwin, *Curr. Pharm. Des.* **2000**, *6*, 441–471.
[3] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402–5415.
[4] T. A. Brooks, L. H. Hurley, *Nat. Rev. Cancer* **2009**, *9*, 849–861.
[5] D. Panda, M. Debnath, S. Mandal, I. Bessi, H. Schwalbe, J. Dash, *Sci. Rep.* **2015**, *5*, 13183.
[6] Y. Pavan Kumar, P. Saha, D. Saha, I. Bessi, H. Schwalbe, S. Chowdhury, J. Dash, *ChemBioChem* **2016**, *17*, 388–393.
[7] K. M. Felsenstein, L. B. Saunders, J. K. Simmons, E. Leon, D. R. Calabrese, S. Zhang, A. Michalowski, P. Gareiss, B. A. Mock, J. S. Schneekloth, Jr., *ACS Chem. Biol.* **2016**, *11*, 139–148.
[8] A. Chauhan, R. Paul, M. Debnath, I. Bessi, S. Mandal, H. Schwalbe, J. Dash, *J. Med. Chem.* **2016**, *59*, 7275–7281.
[9] H. Yaku, T. Murashima, D. Miyoshi, N. Sugimoto, *Molecules* **2012**, *17*, 10586–10613.
[10] J. L. Neo, K. Kamaladasan, M. Uttamchandani, *Curr. Pharm. Des.* **2012**, *18*, 2048–2057.
[11] D. Monchaud, M.-P. Teulade-Fichou, *Org. Biomol. Chem.* **2008**, *6*, 627–636.
[12] S. Zhang, Y. Wu, W. Zhang, *ChemMedChem* **2014**, *9*, 899–911.
[13] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.* **2005**, *1*, 167–173.
[14] N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J. Am. Chem. Soc.* **2008**, *130*, 6722–6724.
[15] A. Kotar, B. Wang, A. Shivalingam, J. Gonzalez-Garcia, R. Vilar, J. Plavec, *Angew. Chem. Int. Ed.* **2016**, *55*, 12508–12511.
[16] V. B. Kovalska, K. D. Volkova, M. Y. Losytskyy, O. I. Tolmachev, A. O. Balanda, S. M. Yarmoluk, *Spectrochim. Acta Part A* **2006**, *65*, 271–277.
[17] Q. Chen, I. D. Kuntz, R. H. Shafer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2635–2639.
[18] I. Lubitz, D. Zikich, A. Kotlyar, *Biochemistry* **2010**, *49*, 3567–3574.
[19] Q. Yang, J. Xiang, S. Yang, Q. Li, Q. Zhou, A. Guan, X. Zhang, H. Zhang, Y. Tang, G. Xu, *Nucleic Acids Res.* **2010**, *38*, 1022–1033.
[20] B. Jin, X. Zhang, W. Zheng, X. Liu, J. Zhou, N. Zhang, F. Wang, D. Shangguan, *Anal. Chem.* **2014**, *86*, 7063–7070.
[21] S. Cosconati, A. Rizzo, R. Trotta, B. Pagano, S. Iachettini, S. De Tito, I. Lauri, I. Fotticchia, M. Giustiniano, L. Marinelli, C. Giancola, E. Novellino, A. Biroccio, A. Randazzo, *J. Med. Chem.* **2012**, *55*, 9785–9792.
[22] F. Aboul-ela, A. I. H. Murchie, D. M. J. Lilley, *Nature* **1992**, *360*, 280–282.
[23] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048–2058.
[24] P. Schultze, R. F. Macaya, J. Feigon, *J. Mol. Biol.* **1994**, *235*, 1532–1547.
[25] Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263–282.
[26] M. Marušič, P. Šket, L. Bauer, V. Viglasky, J. Plavec, *Nucleic Acids Res.* **2012**, *40*, 6946–6956.
[27] J. Kypr, I. Kejnovská, D. Renčiuk, M. Vorlíčková, *Nucleic Acids Res.* **2009**, *37*, 1713–1725.
[28] S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, G. P. Spada, *Org. Biomol. Chem.* **2010**, *8*, 2683–2692.
[29] E. Tuite, B. Nordén, *J. Am. Chem. Soc.* **1994**, *116*, 7548–7556.
[30] F. Riechert-Krause, A. Eick, R. Grünert, P. J. Bednarski, K. Weisz, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2380–2383.
[31] B. Pagano, A. Virno, C. A. Mattia, L. Mayol, A. Randazzo, C. Giancola, *Biochimie* **2008**, *90*, 1224–1232.
[32] S. Ghosh, A. Kar, S. Chowdhury, D. Dasgupta, *Biochemistry* **2013**, *52*, 4127–4137.
[33] A. Funke, J. Dickerhoff, K. Weisz, *Chem. Eur. J.* **2016**, *22*, 3170–3181.
[34] N. T. Dao, R. Haselsberger, M.-E. Michel-Beyerle, A. T. Phan, *FEBS Lett.* **2011**, *585*, 3969–3977.
[35] F. Rosu, E. De Pauw, V. Gabelica, *Biochimie* **2008**, *90*, 1074–1087.
[36] A. Marchand, V. Gabelica, *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1146–1154.
[37] K. Jäger, J. W. Bats, H. Ihmels, A. Granzhan, S. Uebach, B. O. Patrick, *Chem. Eur. J.* **2012**, *18*, 10903–10915.
[38] M. Trajkovski, E. Morel, F. Hamon, S. Bombard, M.-P. Teulade-Fichou, J. Plavec, *Chem. Eur. J.* **2015**, *21*, 7798–7807.
[39] J. Feigon, K. M. Koshlap, F. W. Smith, *Methods Enzymol.* **1995**, *261*, 225–256.
[40] J. Dai, M. Carver, L. H. Hurley, D. Yang, *J. Am. Chem. Soc.* **2011**, *133*, 17673–17680.
[41] B. Karg, A. Funke, A. Ficht, A. Sievers-Engler, M. Lämmerhofer, K. Weisz, *Chem. Eur. J.* **2015**, *21*, 13802–13811.

[42] R. Nanjunda, E. A. Owens, L. Mickelson, S. Alyabyev, N. Kilpatrick, S. Wang, M. Henary, W. D. Wilson, *Bioorg. Med. Chem.* **2012**, *20*, 7002 – 7011.

[43] D. Monchaud, C. Allain, H. Bertrand, N. Smargiasso, F. Rosu, V. Gabelica, A. De Cian, J.-L. Mergny, M.-P. Teulade-Fichou, *Biochimie* **2008**, *90*, 1207 – 1223.

[44] E. Largy, F. Hamon, M.-P. Teulade-Fichou, *Anal. Bioanal. Chem.* **2011**, *400*, 3419 – 3427.

[45] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins Struct. Funct. Bioinf.* **2005**, *59*, 687 – 696.

CHEMISTRY

A European Journal

Supporting Information

Selective Targeting of G-Quadruplex Structures by a Benzothiazole-Based Binding Motif

Ina Buchholz,[a] Beatrice Karg,[a] Jonathan Dickerhoff,[a] Adrian Sievers-Engler,[b] Michael Lämmerhofer,[b] and Klaus Weisz*[a]

chem_201700298_sm_miscellaneous_information.pdf

Table of Contents

Figure S1. A) Structure of benzothiazole and benzoxazole derivatives **L4**, **L5**, and **L6**. B-F) Normalized ligand fluorescence upon titrating the ligands (6 μM) with quadruplexes TBA (B), HT (C), Tetra (D), ODN (E), and c-MYC (F) in potassium phosphate or sodium phosphate buffer (HT) at pH 7 and T = 20 °C. Note, that with ligand **L4** saturation is reached more rapidly indicating its higher affinity for all G4 structures. G) Changes in CD melting temperature of c-MYC and ODN quadruplexes upon addition of ligand in a five-fold excess; standard deviations based on three independent measurements are indicated. Due to its high thermal stability, the c-MYC quadruplex was dissolved in a sodium phosphate buffer previously found to conserve its parallel fold. Again, binding of ligand **L4** results in the highest G4 thermal stabilization.
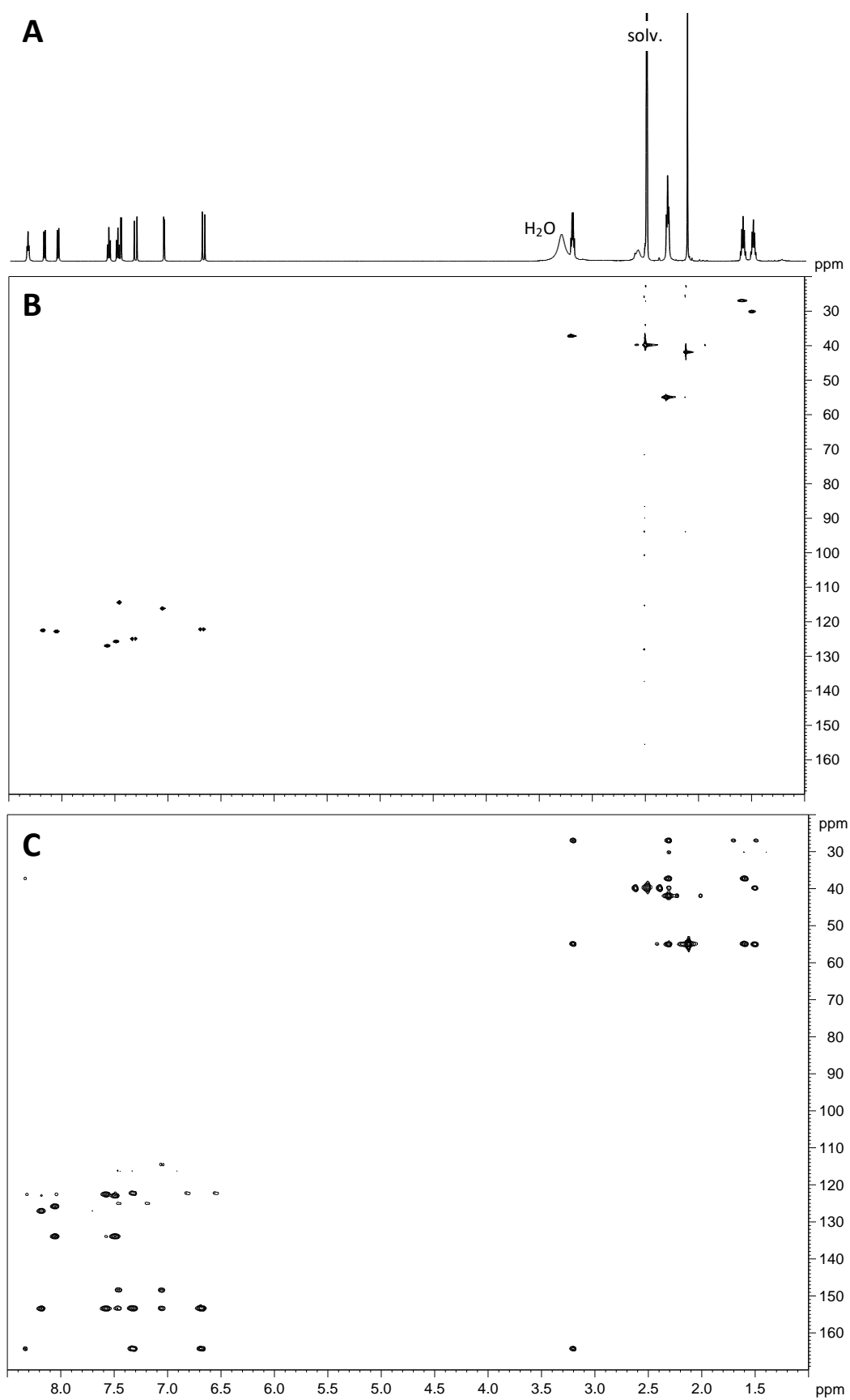
Figure S2. One-dimensional ^{1}H NMR spectrum (A) and two-dimensional ^{1}H-^{13}C HSQC (B) and ^{1}H-^{13}C HMBC spectrum (C) of **L4** in DMSO-d_6 at 298 K.
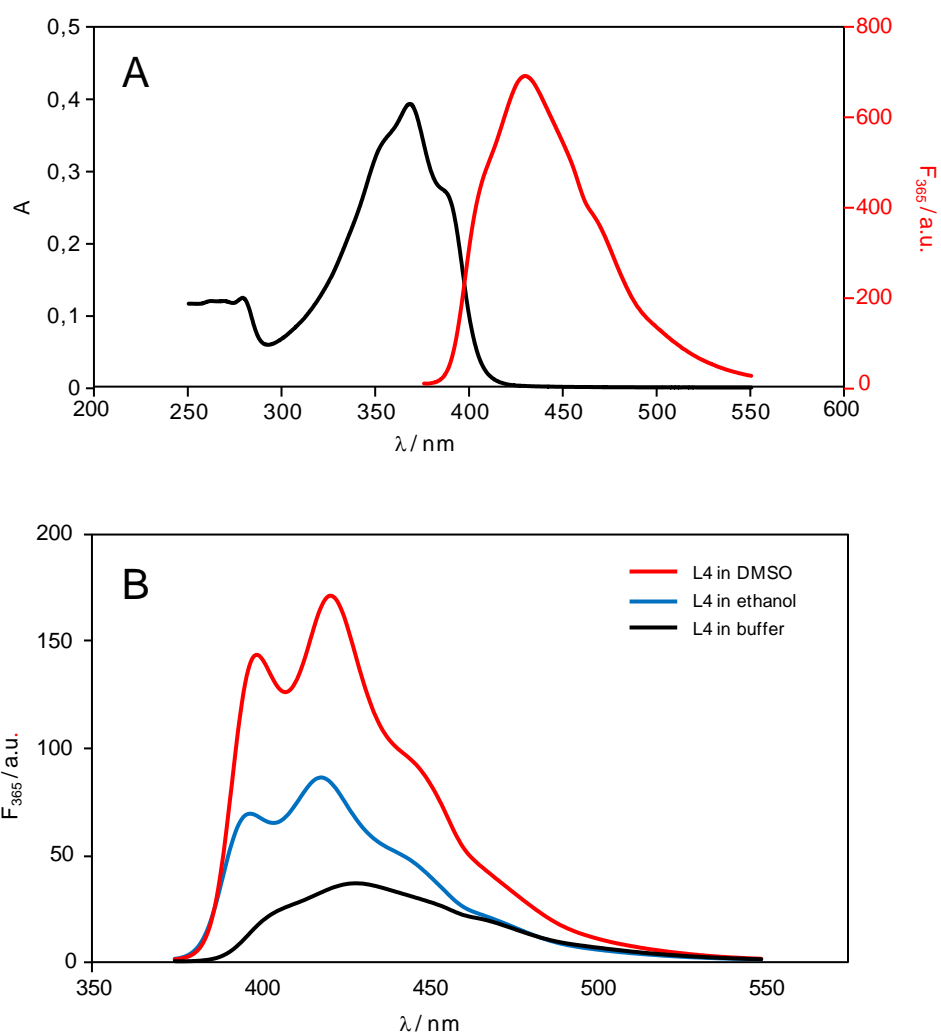
Figure S3. A) Absorption (black) and fluorescence emission spectrum (red) of 20 μM and 6 μM **L4** in potassium phosphate buffer, pH 7.0. B) Fluorescence emission spectra of **L4** (3 μM) in buffer (black), ethanol (blue) and DMSO (red); T = 20 °C.
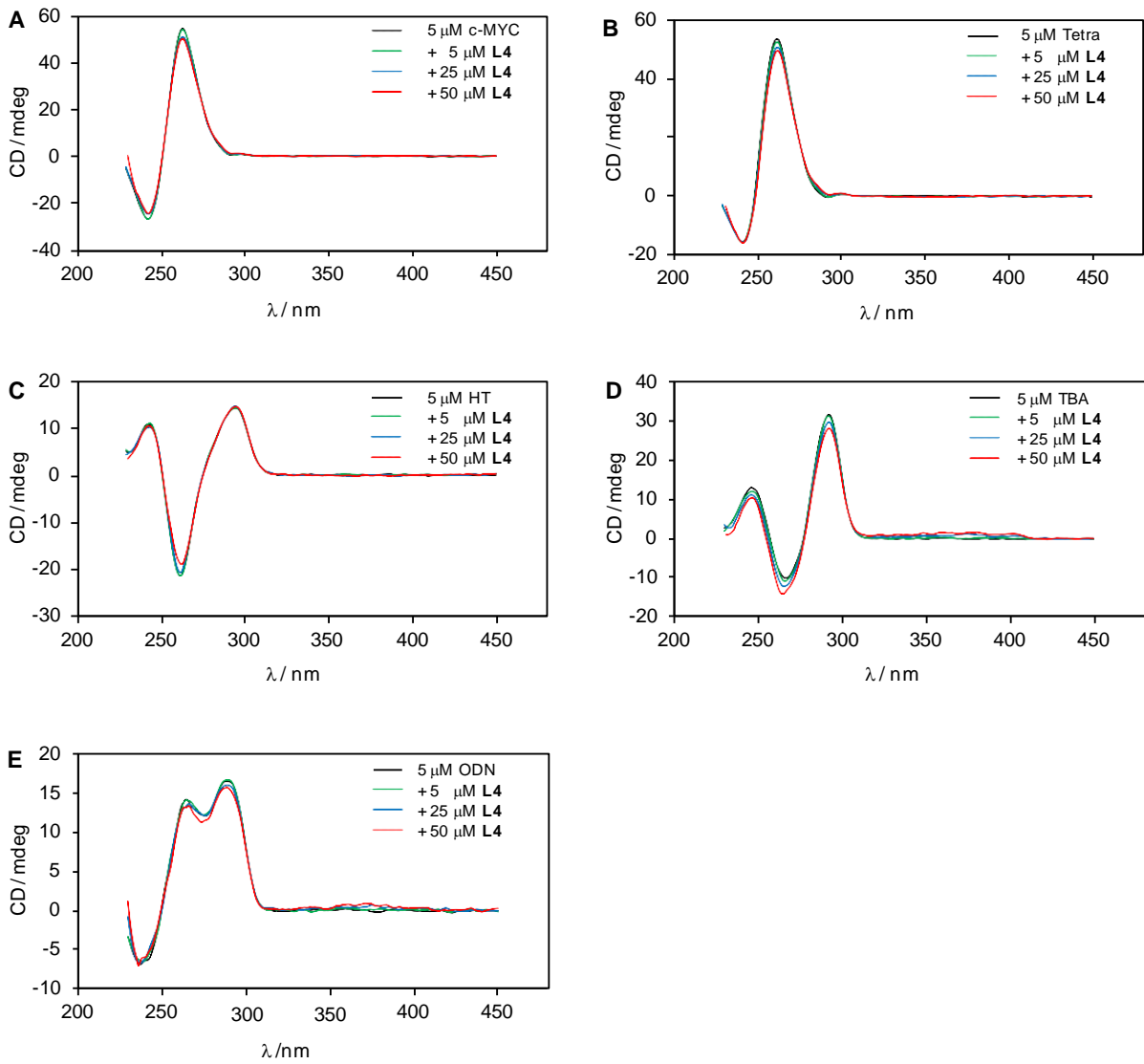
Figure S4. CD spectra of c-MYC (A), Tetra (B), HT (C), TBA (D), and ODN (E) with increasing amounts of ligand **L4** in potassium phosphate (A,B,D,E) or sodium phosphate buffer (C), pH 7, at 20 °C.
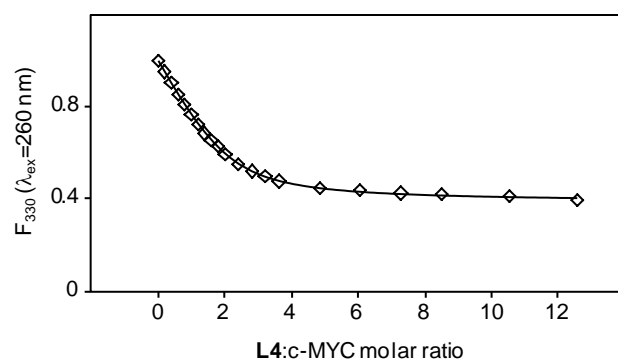
Figure S5. Fluorescence titration of c-MYC (6 μM) with **L4** in potassium phosphate buffer, pH 7; T = 20 °C, λ_{ex} = 260 nm.
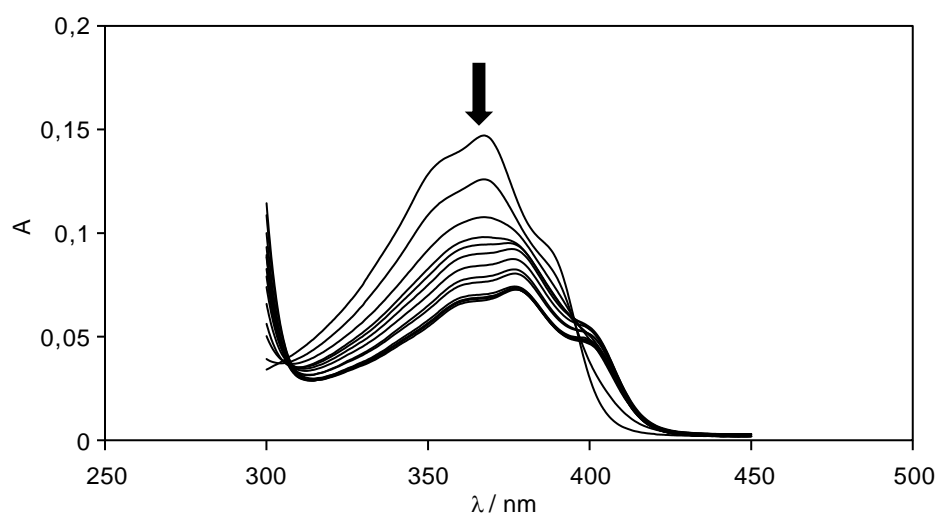
Figure S6. UV-vis titration of **L4** (6 μM) with c-MYC in aliquots of 0.5 eq. in potassium phosphate buffer at 20 °C.
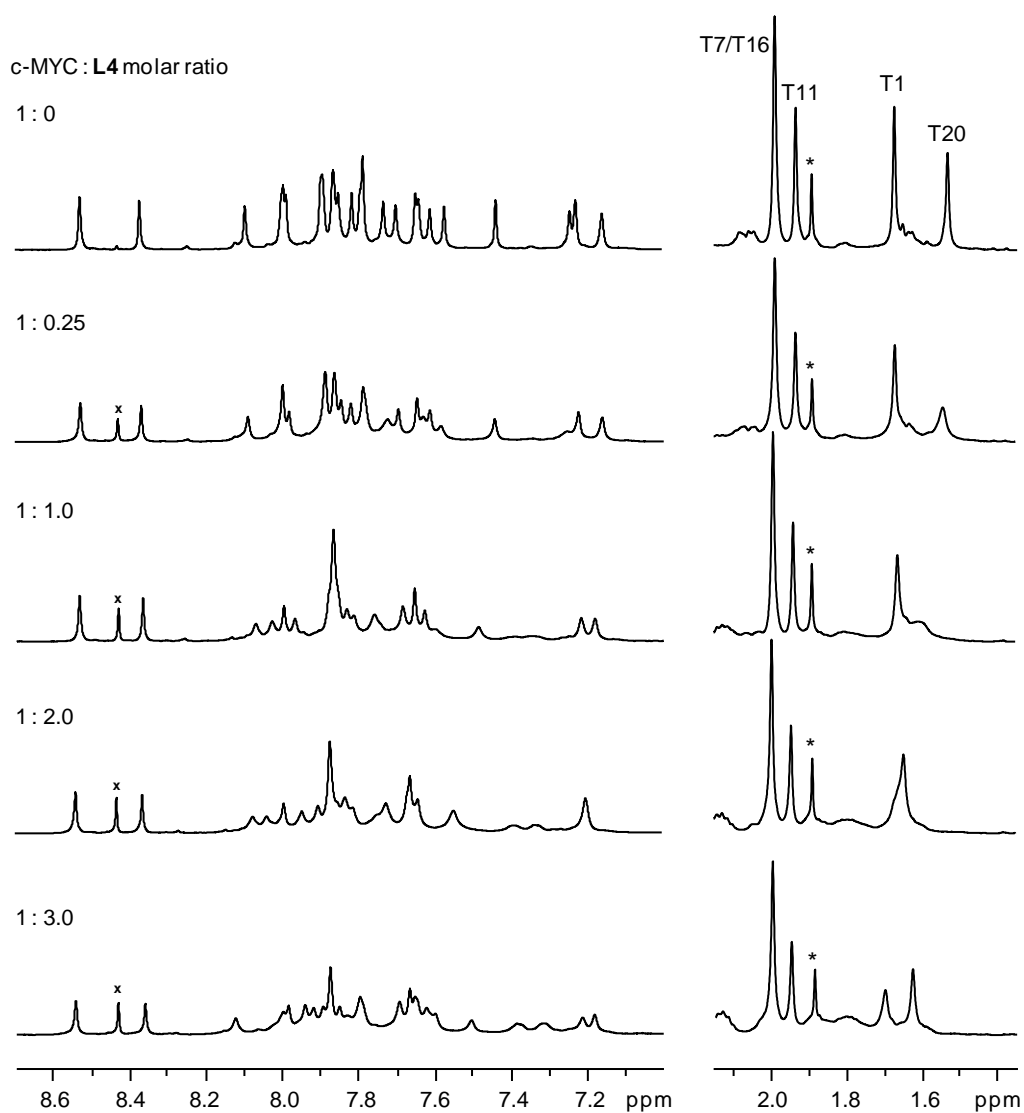
Figure S7. Aromatic and methyl proton spectral regions of 0.45 mM c-MYC quadruplex with increasing amounts of **L4** in 10 mM potassium phosphate, pH 7, at 40 °C. Assignments for T methyl protons are given for the free c-MYC. The resonance marked with an asterisk is due to residual acetate. A resonance of unknown origin not correlated to any other proton of the c-MYC quadruplex through exchange, dipolar, or scalar coupling is marked by 'x'.
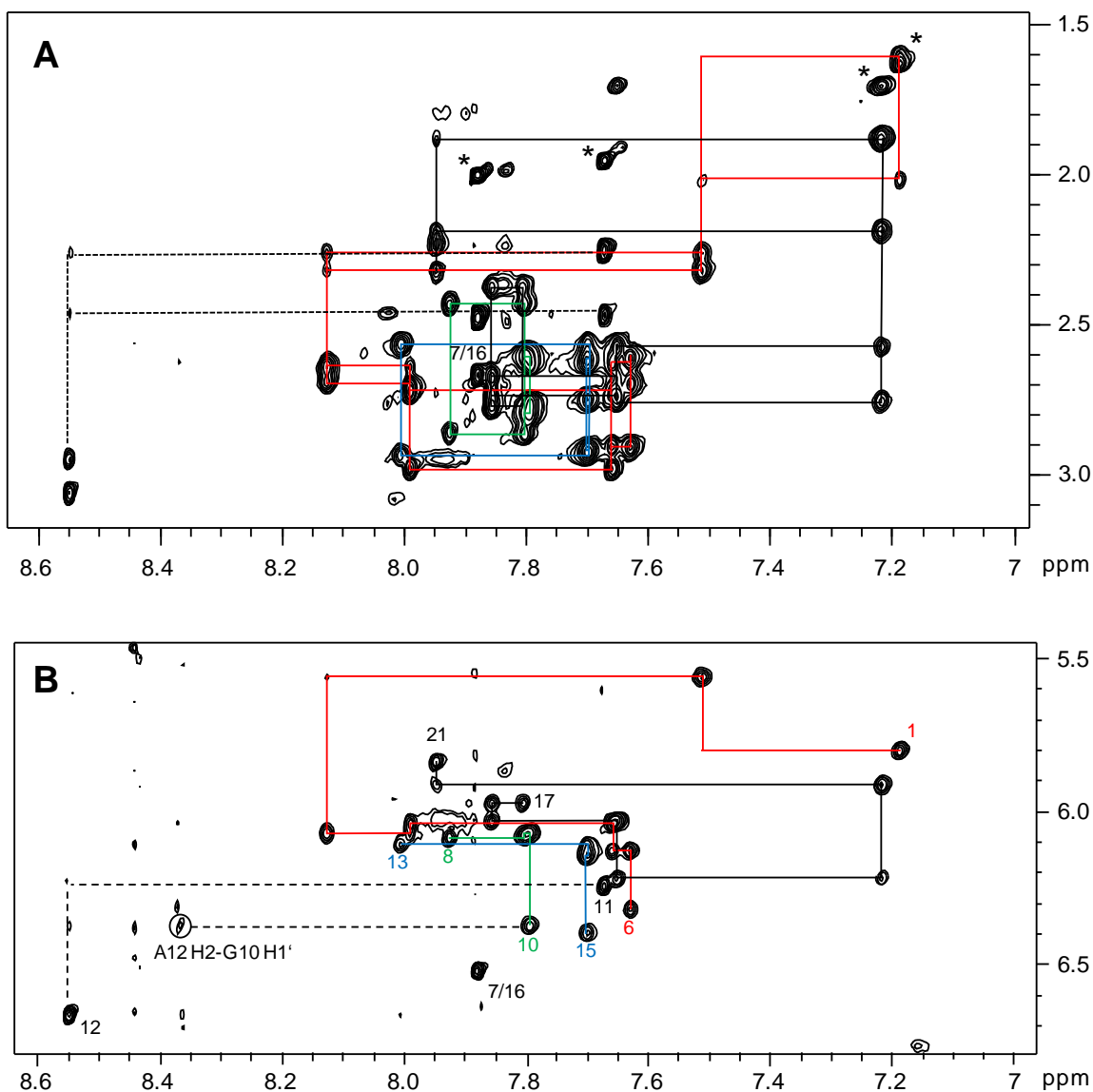
Figure S8. H6/H8-H2'/H2" (A) and H6/H8-H1' (B) spectral region of a 2D NOE spectrum (mixing time 300 ms) of c-MYC (0.45 mM) in the presence of **L4** (1:3 molar ratio) in potassium phosphate buffer at 40 °C. Intranucleotide and sequential NOE connectivities are followed by the red (T1-G6), green (G8-G10), blue (G13-G15), and black (G17-G21) horizontal and vertical lines; broken lines trace NOE contacts in the TA loop while H6-Me contacts in (A) are marked by asterisks.

Table S1. ^{1}H chemical shifts δ (in ppm) of protons in c-MYC[a]

	imino	H6/H8	H1'	H2'1	H2'2
T1	n.d.	7.23	5.82	1.65	2.06
G2	n.d.	7.64	5.65	2.39	2.46
A3	---	8.00	5.89	2.60	2.73
G4	11.70	8.00	6.06	2.74	3.00
G5	11.22	7.70	6.14	2.61	2.91
G6	10.67	7.73	6.40	2.60	2.74
T7	n.d.	7.87	6.52	2.47	2.67
G8	11.68	7.99	6.14	2.48	2.91
G9	11.49	7.89	6.14	2.66	2.85
G10	11.08	7.85	6.41	2.60	2.74
T11	n.d.	7.65	6.24	2.23	2.46
A12	---	8.53	6.68	2.96	3.07
G13	11.89	8.09	6.17	2.62	2.96
G14	11.24	7.78	6.20	2.67	2.98
G15	11.00	7.79	6.44	2.61	2.71
T16	n.d.	7.86	6.52	2.47	2.67
G17	11.25	7.90	6.00	2.37	2.79
G18	11.35	7.89	6.03	2.67	2.73
G19	11.05	7.61	6.18	2.53	2.78
T20	n.d.	7.15	5.91	1.90	2.34
A21	---	7.82	5.80	2.07	2.43
A22	---	n.d.	n.d.	n.d.	n.d.

[a]At 40 °C in 90% H_2O/10% D_2O, 10 mM potassium phosphate buffer, pH 7.0; n.d. = not determined.

Table S2. [1]H chemical shifts δ (in ppm) of protons in c-MYC with **L4** in a 1:3 molar ratio[a]

	imino	H6/H8	H1'	H2'1	H2'2
T1	n.d.	7.19	5.80	1.61	2.02
G2	n.d.	7.51	5.56	2.26	2.32
A3	---	8.13	6.07	2.64	2.69
G4	11.52	7.99	6.04	2.72	2.98
G5	11.12	7.66	6.13	2.63	2.91
G6	10.67	7.63	6.32	2.61	2.70
T7	n.d.	7.88	6.52	2.48	2.67
G8	11.51	7.92	6.09	2.43	2.86
G9	11.36	7.80	6.07	2.61	2.81
G10	11.06	7.79	6.37	2.64	2.76
T11	n.d.	7.67	6.24	2.26	2.47
A12	---	8.55	6.67	2.95	3.06
G13	11.69	8.01	6.11	2.57	2.94
G14	11.14	7.70	6.14	2.62	2.92
G15	10.87	7.70	6.40	2.64	2.75
T16	n.d.	7.88	6.52	2.48	2.67
G17	11.30	7.80	5.97	2.38	2.78
G18	11.32	7.86	6.03	2.68	2.73
G19	11.06	7.65	6.22	2.58	2.76
T20	n.d.	7.22	5.91	1.88	2.19
A21	---	7.95	5.84	2.23	2.33
A22	---	n.d.	n.d.	n.d.	n.d.

[a]At 40 °C in 90% H_2O/10% D_2O, 10 mM potassium phosphate buffer, pH 7.0; n.d. = not determined.
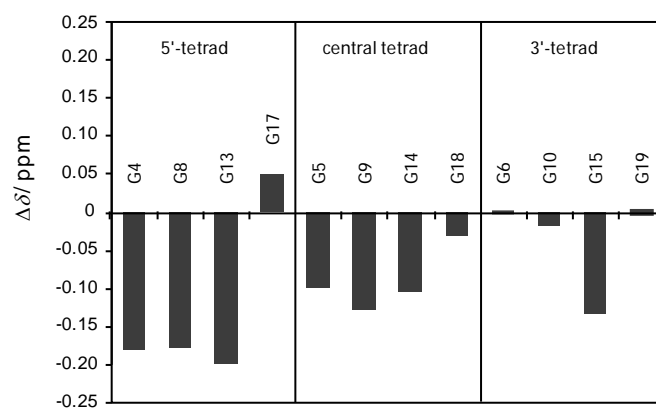
Figure S9. Chemical shift changes for imino protons within the three tetrads of the c-MYC sequence upon binding **L4**.
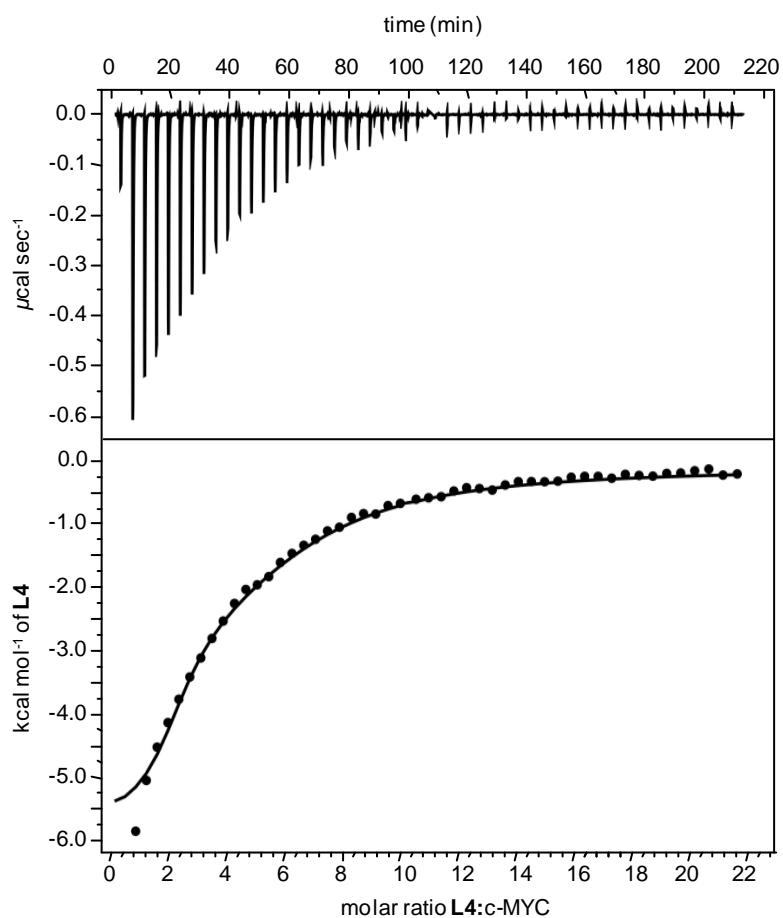
Figure S10. ITC binding isotherm of **L4** titrated to the c-MYC quadruplex. The upper and lower panel shows the heat burst for every injection step and the corresponding normalized heat signal versus molar ratio, respectively. The curve represents a best fit of the experimental data based on two nonequivalent binding sites and with molar enthalpy $\Delta H^{o}{}_1$ and stoichiometry n_1 for the high-affinity binding sites fixed to -5.79 kcal/mol and 2, respectively (see manuscript text). Variable parameters determined from curve fitting yield $K_{a1} = 6.5 \cdot 10^5$ M^{-1}, $K_{a2} = 2.7 \cdot 10^4$ M^{-1}, $\Delta H^{o}{}_2 = -2.4$ kcal/mol, and $n_2 = 6.4$.

Appendix IV

Accepted Manuscripts

<table>
<tr><td colspan="2">Publication V</td></tr>
<tr><td>Title</td><td>Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of Calendula officinalis flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS</td></tr>
<tr><td>Authors</td><td>Nicolaus, Christoph; Sievers-Engler, Adrian; Murillo, Renato; D'Ambrosio, Michele; Laemmerhofer, Michael; Merfort, Irmgard</td></tr>
<tr><td>Published in</td><td>Journal of Pharmaceutical and Biomedical Analysis (118), 2016, p. 195-205</td></tr>
<tr><td>DOI</td><td>http://dx.doi.org/10.1016/j.jpba.2015.10.025</td></tr>
</table>

Reprint with permission from Elsevier

Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of *Calendula officinalis* flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS

Christoph Nicolaus [a,1], Adrian Sievers-Engler [b,1], Renato Murillo [c], Michele D'Ambrosio [d], Michael Lämmerhofer [b,**], Irmgard Merfort [a,*]

[a] *Institute of Pharmaceutical Sciences, Pharmaceutical Biology and Biotechnology, University of Freiburg, Stefan-Meier-Strasse 19, D-79104 Freiburg, Germany*
[b] *Institute of Pharmaceutical Sciences, University of Tübingen, Auf der Morgenstelle 8, D-72076 Tübingen, Germany*
[c] *Escuela de Quimica and CIPRONA, Universidad de Costa Rica, 2060 San Jose, Costa Rica*
[d] *Laboratorio di Chimica Bioorganica, Università degli Studi di Trento, Via Sommarive, 14, Italy*

ARTICLE INFO

ABSTRACT

Pentacyclic triterpene mono- and diesters have been isolated from *Calendula officinalis* flowers. GC–MS, APCI-Exactive Orbitrap HR-MS and NMR allowed to identify the triterpene skeleton in various samples (different triterpene mixtures from *Calendula n*-hexane extract). NMR provided evidence that triterpene diesters are present in the samples as well. However, the corresponding quasi-molecular ions could not be detected by APCI-Exactive Orbitrap HR-MS. Instability of triterpene diesters and loss of a fatty acid residue, respectively, in the ion-source made their MS detection challenging. Thus, a set of new APCI-QTOF-MS methods (using the TripleTOF 5600+ mass spectrometer) were developed which made it eventually possible to solve this problem and confirm the diester structures by MS via quasi-molecular ion [M + H]+ detection. Direct infusion APCI-QTOF MS experiments in MS/MS high sensitivity scan mode with low collision energy and multi-channel averaging acquisition (MCA) allowed the detection of quasi-molecular ions of triterpene diesters for the first time and unequivocally confirmed the presence of faradiol 3,16-dimyristate and -dipalmitate, as well as the corresponding mixed diesters faradiol 3-myristate,16-palmitate and faradiol 3-palmitate,16-myristate. Preferential loss of the fatty acid in 16-position made it possible to distinguish the mixed diesters by MS/MS spectra. Their chromatographic separations turned out to be challenging due to their bulkiness and extended molecular dimensions. However, separation could be achieved by an uncommon non-aqueous RPLC mode with an in-house synthesized C30 phase. Finally, two (U)HPLC-APCI-QTOF-MS methods with C18- and C30-based non-aqueous RPLC provided suitable, sensitive assays to monitor the presence of monoesters and diesters of various triterpenes (faradiol, maniladiol, arnidiol, arnitriol A and lupane-3β,16β,20-triol esters) in the *n*-hexane extract of *C. officinalis* with high mass resolution and good mass accuracy.

1. Introduction

Lipophilic preparations from flowers of the Asteraceae *Calendula officinalis* L. (marigold) are recommended for the treatment of minor inflammation of the skin and wounds [1]. Triterpenes from the 20-taraxastene type, especially its major triterpenoid constituent faradiol seem to be crucial for the anti-inflammatory potential [2–4]. These triterpenes are present as monols (e.g. ψ-taraxasterol) as well as diols (e.g. faradiol and arnidiol) and triols (e.g. heliantriol A-C) and here especially in form of their 3-O-lauroyl, -myristoyl and -palmitoyl esters [5–7]. Although the occurrence of triterpene diol diesters has also been reported [8], their exact structures have never been clarified.

Structure elucidation of the non-esterified pentacyclic triterpenes can be done by NMR combined with MS analysis. Here, spectra obtained by EI-MS or GC–EI-MS technique give characteristic fragments of the respective triterpene skeleton

[9–14]. Additionally, LC–MS using either ESI (electrospray ionization, especially for polar substituted triterpenes) [15–18], APCI (atmospheric-pressure chemical ionization) or APPI (atmospheric-pressure photoionization) has been reported [19–22]. These LC–MS techniques are often suitable for the analysis of complex triterpene mixtures, but they may also cause sensitivity problems depending on the triterpene derivative. Moreover, Vahur et al. [23] analysed a dammar resin and concluded that the combinatorial use of MALDI-FT-ICR-MS and APCI-FT-ICR-MS seems the best option. However, the successful analysis of complex triterpene mono- and diester mixtures is still missing. This is a challenging task as often triterpenes occur in a mixture of isomeric skeletons or with the same triterpene core and differences in the fatty acid side chains. A prominent example is the lipophilic extract from C. officinalis L. flowers.

A major analytical problem encountered with these triterpene esters in developing feasible LC–MS methods is the poor ionization efficiency and/or low stability of the [M+H]+ molecular ion initially formed during commonly used LC compatible ionization methods like ESI, APCI or APPI, resulting in low sensitivity. Thus, adequate ionization parameter optimization is required to allow detection of the intact molecular ion with sufficient sensitivity [24]. Additionally, isobaric and/or isomeric interference from mono- or diester derivatives does not allow unequivocally identifying the target compound requiring suitable selective chromatographic methods which allow separation of the isomeric species.

We here report on (U)HPLC-APCI-QTOF-MS in the positive mode as a MS setup exhibiting acceptable sensitivity, high mass resolution and good mass accuracy. Moreover, the developed (U)HPLC methods allowed the separation of triterpene mono- and diesters in a mixture. The combination of both methods, LC and HR-MS, enabled the identification of various triterpene mono- and diesters, which have not been previously reported for C. officinalis flowers. Moreover, structure elucidation of novel faradiol diesters was possible based on NMR data and this LC–MS method.

2. Experimental

2.1. Materials

Faradiol for preparation of monoester and diester standards was purchased from PhytoLab (Vestenbergsgreuth). Faradiol 3-myristate and faradiol 3-palmitate were previously isolated from C. officinalis L. flowers [25]. Chemicals and solvents for synthesis of standards were obtained from Sigma–Aldrich. Solvents for HPLC–MS were of MS-grade and purchased from Carl Roth (Karlsruhe, Germany). MS-grade water was prepared by a Purelab Ultra Purification system from Elga Labwater (Celle, Germany).

2.2. Experimental methods

2.2.1. NMR

Spectra were recorded in CDCl₃ on a Bruker DRX 400 MHz instrument (Bruker, Bremen, Germany) at 400 MHz (^{1}H) and 100 MHz (^{13}C).

2.2.2. GC–MS

Analysis was performed on an Agilent 6890 Series; detector: Agilent 5973 Network Mass Selective Detector, Agilent Technologies; ionization energy: 70 eV; column: fused-silica capillary column (30 m × 0.25 mm) coated with 0.25 μm dimethyl polysiloxane (RTX®-1 MS, Restek) and helium as carrier gas with a flow rate of 1.0 ml per minute. The temperature profile was isothermal at 270 °C for 50 min. The injector and detector temperatures were 300 °C, the injection volume was 1.0 μl and the split ratio 50:1.

2.2.3. HR-APCI-MS

Spectra in the positive mode were recorded after direct infusion (in MeOH with a flow rate of 100 μl/min) on an Exactive Mass Spectrometer (Thermo Scientific).

2.2.4. LC-APCI-QTOF-MS

Analyses were run on a Sciex TripleTOF 5600+ (Sciex, Ontario, CA) with a Duospray ion source coupled to an Agilent 1290 Series UHPLC instrument (Agilent, Waldbronn, Germany) equipped with a Pal HTC-XS autosampler from CTC (Zwingen, Switzerland).

Source parameters in APCI+-mode were as follows: curtain gas 30 psi, nebulizer gas 60 psi, drying gas 50 psi, source temperature 500 °C, source voltage 5500 V. Cycles consisted of a TOF-MS scan from 100 to 2000 m/z, followed by individual product-ion scans. Parameters for TOF-scan: collision energy (CE) 5 V, declustering potential (DP) 100 V, RF-transmission 400 m/z: 50%, 550 m/z: 50%. Due to strong in-source decay of the analytes (in particular the diesters) owing to neutral loss of fatty acid and/or water, time bins were set to 60 in order to enhance sensitivity compromising, however, mass resolution. For product-ion scan in high sensitivity mode, CE was raised to 17 V. Precursor and enhancement m/z (EMZ) were set according to [M+H]+ for confirmation runs, whilst using otherwise [M+H-H₂O]+ and [M+H-RCOOH]+ monitoring the in-source decay products for enhanced sensitivity. Hold up time t_m was determined by injection of 5 μl of acetone.

2.2.5. Direct infusion APCI-QTOF-MS

Infusion spectra were obtained employing the Sciex TripleTOF 5600+ in APCI+ mode. Spectra were acquired using multichannel averaging (MCA) for enhanced signal to noise ratio. Source temperature was set to 500 °C, curtain gas to 30 psi, nebulizer gas to 50 psi, drying gas to 40 psi and source voltage to 5500 V. CE and DP were set accordingly (see Results and discussion section). Transmission windows were set for MS scans m/z 400 50% and m/z 850 50% for Q1 and for MS2 scans Q2 transmission was set to m/z 400 50% and m/z 550 50%. Time bins were varied accordingly between 20 and 60. For optimal sensitivity, product ion scans were acquired in high sensitivity mode and Q1 resolution was set to low. Flow rate was adjusted accordingly between 15 and 30 μl per minute.

2.2.6. C18-Chromatography for monoesters

Liquid chromatography of monoesters was performed on a core–shell Kinetex C18, 2.6 μm (100 × 3 mm ID) column from Phenomenex (Torrance, USA) using a non-aqueous gradient elution mode. Eluent A consisted of acetonitrile and eluent B of MeOH. The gradient profile was as follows: 0–1 min, 0%B; 1–10 min, 0–100%B; 10–15 min, 100%B. The flow rate was set to 400 μl/min and the column temperature kept constant at 45 °C.

2.2.7. C30-Chromatography for diesters

Liquid chromatography of diesters was performed on an in-house synthesized C30-modified silica column (50 × 4 mm ID). The column packing material was obtained by refluxing Kromasil 100 Å 5 μm silica gel (EKA Chemicals, Bohus, Sweden) with triacontyldimethylchlorosilane (ABCR, Karlsruhe, Germany) leading to a surface coverage of 1.4 μmol/m². A non-aqueous gradient elution mode was employed with the following gradient profile: 0–1 min, 30% B; 1–25 min, 30–100% B; 25–30 min, 100% B. Eluent A contained MeOH and eluent B 2-propanol. The flow rate was set to 1200 μl/min and the column temperature to 8 °C.

2.2.8. Further column chromatography

Open CC was carried out by using Silica gel 60 (0.06–0.20 mm) or Florisil® (60–100 mesh), both from Roth, and flash chromatography by using the Sepacore® chromatography system (Büchi Labortechnik, Flawil, Switzerland) instrumented with two C 605

pump modules, a C 620 control unit, a C 635 UV detector, a C 660 fraction collector, the control software Sepacore® Control 1.2 and pre-packed Silica gel 60 (0.015–0.04 mm, Machery Nagel), polypropylene cartridges.

2.2.9. TLC

Fractions were monitored by analytical TLC (Silica gel 60 F$_{254}$ aluminium sheets, Merck) developed with CH$_2$Cl$_2$-EtOAc (9:1, v/v) and detected at 366 nm and vis after derivatization with anisaldehyde/sulphuric acid and heating at 120 °C for 10 min. Sample application was carried out with the Automatic TLC sampler AST-4 (CAMAG, Muttenz, Switzerland).

2.3. Plant material

C. officinalis L. flowers were obtained from the Dr. Junghanns GmbH, Aschersleben, Germany, and personally identified by Dr. Junghanns, lot no.: 1200812. Flower heads were collected during September 2012 and dried at 45 °C.

2.4. Extraction and isolation

Air-dried and powdered flower heads (502 g) were exhaustively extracted with *n*-hexane (5.8 l) using a Soxhlet apparatus. After removal of the solvent under vacuum at 40 °C and lyophilization, a crude *n*-hexane extract (55.7 g) was obtained. Two portions of 8 g of this extract were separated into 15 fractions from I, and 19 fractions from II by flash chromatography (15 × 4 cm) on silica gel with mixtures of petrol-CH$_2$Cl$_2$ (50:49, v/v) and MeOH (0–60%, v/v) and a flow rate of 10 ml/min. Subsequently, fraction I.13 (1.89 g) and fraction II.11 (2.12 g) were fractionated by open CC (30 × 2 cm and 50 × 3 cm) on Florisil® with mixtures of petrol-acetone (0–90%, v/v) yielding 13 and 9 fractions, respectively. A portion of sub-fraction I.13.II.1 (200 mg) was separated on silica gel by flash chromatography (12 × 1.5 cm) using mixtures of cyclohexane-CH$_2$Cl$_2$ (7:4, v/v) and MeOH (0–8%, v/v) yielding 5 sub-fractions. Fractionation of sub-fraction I.13.II.1.5 (45 mg) over silica by open CC (30 × 1 cm) using *n*-hexane-EtOAc mixtures (20–100%, v/v) afforded a fraction with **4a + b** and **5a + b** (9 mg). All fractions were monitored by TLC (CH$_2$Cl$_2$-EtOAc; 9:1, v/v) for triterpenediol esters. Fractions with this triterpene type were combined to TFa (2.32 g) which was chromatographed by flash chromatography (15.0 × 4.0 cm) on silica gel with mixtures of CH$_2$Cl$_2$-EtOAc (0–30%, v/v) affording fractions with a mixture of **7–10** (11.8 mg), **6a + b** (38.9 mg) and 14 sub-fractions. After monitoring the fractions for triterpenetriol esters using analytical TLC (CH$_2$Cl$_2$-EtOAc; 9:1, v/v) fractions I.13.I.2 (46.9 mg), I.13.II.3 (61.4 mg), I.13.III.2 (39.5 mg), II.12.5 (83.3 mg) and TFa.15 (0.11 g) were combined to TFb (0.34 g). TFb was chromatographed by open CC (38 × 1.5 cm) with mixtures of *n*-hexane/EtOAc (20–100%, v/v) resulting in 14 sub-fractions. Further separation of sub-fraction TFb.10 (38.8 mg) by open CC (30 × 1 cm) with mixtures of CH$_2$Cl$_2$-EtOAc (0–30%, v/v) yielded **2a + b** (10.7 mg) and **1** (3.8 mg). Isolation of **3a + b** (3.7 mg) was afforded by fractionation of sub-fraction TFb.11 (18.5 mg) by open CC (39 × 1 cm) using mixtures of CH$_2$Cl$_2$-EtOAc (5–20%, v/v).

2.5. Synthesis of faradiol mono- and diester standards

Faradiol mono- and diesters were synthesized by a Steglich-type esterification protocol. In brief, myristic or palmitic acid, or a mixture of both, were dissolved in CH$_2$Cl$_2$ (dried and distilled over phosphorous pentoxide). Subsequently, 4-dimethylaminopyridine (DMAP) (18 mmol) was added to 0.24 mmol fatty acid followed by addition of dicyclohexylcarbodiimide (DCC) (0.26 mmol). After refluxing this reaction mixture for 2 h, faradiol (0.7 mmol) was added. The progress of the reaction was monitored and samples

taken after 6 h and after 20 h. The reaction was finally quenched with MeOH and the solvent evaporated. This procedure afforded faradiol 3-myristate (F3-M), faradiol 16-myristate (F16-M) and faradiol 3,16-dimyristate (FMM) from the reaction with myristic acid; faradiol 3-palmitate (F3-P), faradiol 16-palmitate (F16-P) and faradiol 3,16-dipalmitate (FPP) from palmitic acid, as well as a mixture of all of them along with mixed diesters, faradiol 3-myristate, 16-palmitate (FMP) and faradiol 3-palmitate,16-myristate (FPM) from the reaction with a mixture of both myristic and palmitic acid. Besides, a faradiol 3-myristate,16-palmitate standard from faradiol 3-myristate and a faradiol 3-palmitate,16-myristate standard from faradiol 3-palmitate were synthesized by esterification with palmitic and myristic acid, respectively, using the above reported protocol.

3. Results and discussion

3.1. Identification of triterpene monoesters

Re-investigation of the triterpenoid fraction from *C. officinalis* flowers afforded several known triterpenes (Fig. 1). Based on 1D and 2D NMR analyses (^{1}H, ^{13}C, HSQC, HMBC, ^{1}H/^{1}H-COSY) as well as HR-APCI–MS (positive mode after direct infusion) and UHPLC–APCI–QTOF-MS, partly GC–MS analysis after alkaline hydrolysis and comparison to spectral data from the literature the compounds were identified as 3-myristic and 3-palmitic acid esters of arnitriol A (**2a + b**) [26], lupane-3β,16β,20-triol (**3a + b**) [26] and maniladiol (**6a + b**) [10,26,27] as well as faradiol (**1**) [2]. Triterpenes **4a + b** and **5a + b** were isolated in a mixture. Identification of the triterpene skeleton as faradiol and arnidiol was accomplished after alkaline hydrolysis by GC–MS [28], and of the esters by UHPLC-APCI-QTOF MS. Compounds **2a + b** and **3a + b** have not been reported for *C. officinalis* yet.

3.2. UHPLC-APCI-QTOF MS characterization of isolated monoester samples

Identification of the triterpene monoesters was successful in a mixture of the myristic and palmitic acid esters without prior chromatographic separation. However, the conditions used for HR-APCI-MS analyses with Exactive Orbitrap mass spectrometer did not yield [M + H]$^+$ quasi-molecular ions, respectively, but gave an intense peak [M + H-H$_2$0]$^+$ after loss of H$_2$O. This prompted us to test other instruments and optimize ionization parameters on a QTOF instrument (TripleTOF 5600+). Samples of compounds **1** to **6a + b** were analyzed by UHPLC-APCI-QTOF MS using a Kinetex core–shell C18 column. Due to the high lipophilicity of the triterpene fatty acid monoesters a non-aqueous gradient elution mode was applied. Such non-aqueous RP method may be regarded uncommon, but is actually often employed for very lipophilic natural compounds such as carotenoids [29], xanthophylls [30,31], triglycerides [32] and D vitamins [33]. In fact, no elution of triterpene esters was observed when a common RP-type gradient from water-acetonitrile (95:5; v/v) to 100% acetonitrile was used for elution. Next, a gradient was run from 100% MeOH to 100% acetonitrile. Unfortunately, the triterpene monoesters eluted close to t$_0$ unresolved from impurities. Thus, the gradient solvents were exchanged in A and B, and a gradient from 100% acetonitrile to 100% MeOH gave finally an acceptable retention and separation (Fig. 2). Surprisingly, MeOH showed higher elution strength than acetonitrile which may indicate that to some extent H-bond qualities of the elution solvent are of importance to disrupt analyte-sorbent interactions. As shown in Fig. 2 all monoesters eluted as narrow peaks within 10 min after the non-esterified basic faradiol (**1**).
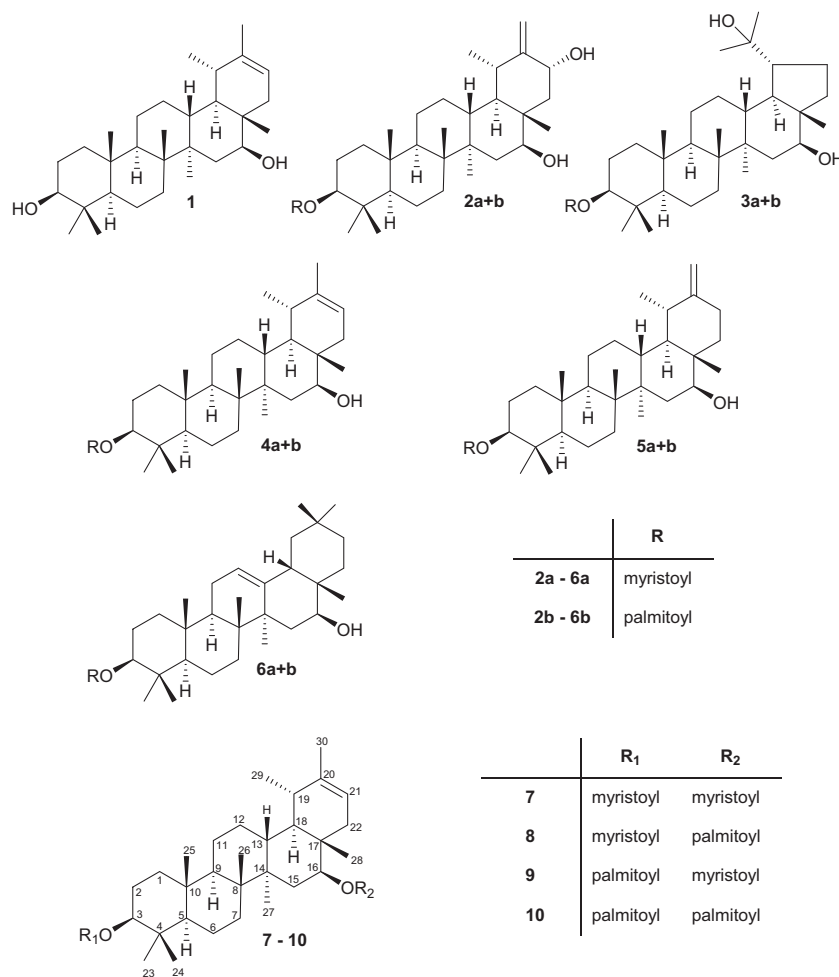
Fig. 1. Structures of pentacyclic triterpene fatty acid mono- and diesters isolated from *C. officinalis* L. flowers.

Table 1
Summary of MS data and retention times (Kinetex C18) of triterpene monoesters. For confirmation see MS/MS spectra in supplemental information.

Compound	Name	Formula	Calculated m/z	Found at m/z	Δ mDa	Δ ppm	t_R (min)
1	faradiol	$C_{30}H_{50}O_2$	443.38836	443.3863	2.06	4.6	2.36
2a	arnitriol A 3-myristate	$C_{44}H_{76}O_4$	669.58164	669.5918	−10.16	−15.2	5.80
2b	arnitriol A 3-palmitate	$C_{46}H_{80}O_4$	697.61294	697.6268	−13.86	−19.9	6.64
3a	lupane-3β,16β,20-triol-3-myristate	$C_{44}H_{78}O_4$	671.59729	671.5955	1.79	2.7	4.97
3b	lupane-3β,16β,20-triol-3-palmitate	$C_{46}H_{82}O_4$	699.62859	699.6293	−0.71	−1.0	5.83
4a	faradiol 3-myristate	$C_{44}H_{76}O_3$	653.58672	653.5912	−4.48	−6.9	6.74
4b	faradiol 3-palmitate	$C_{46}H_{80}O_3$	681.61802	681.6252	−7.18	−10.5	8.02
5a	arnidiol 3-myristate	$C_{44}H_{76}O_3$	653.58672	653.5922	−5.48	−8.4	6.52
5b	arnidiol 3-palmitate	$C_{46}H_{80}O_3$	681.61802	681.6153	2.72	4.0	7.77
6a	maniladiol 3-myristate	$C_{44}H_{76}O_3$	653.58672	653.5948	−8.08	−12.4	6.52
6b	maniladiol 3-palmitate	$C_{46}H_{80}O_3$	681.61802	681.6218	−3.78	−5.5	7.80

The triterpene monoesters are susceptible to in-source decay giving preferentially [M+H-H$_2$O]$^+$ signal in APCI-QTOF MS detection. In order to be able to detect the quasi-molecular ion [M+H]$^+$ which confirms the presence of the respective triterpene monoester, careful optimization of the instrumental settings of the TripleTOF 5600+ mass spectrometer was necessary. For monitoring of the quasi-molecular ion [M+H]$^+$ and the confirmation of the respective monoesters in the samples, a MS/MS scan in the high sensitivity mode with enhancement mass set on the precursor m/z and a CE of 5 V, which provided the best sensitivity for the target mass, was selected. While we were finally able to detect the [M+H]$^+$ quasi-molecular ions of the monoesters with such settings and unequivocally confirm their presence (Table 1), for sensitive detection of the substances in the samples, the [M+H-

H$_2$O]$^+$ fragment ion can be conveniently selected. All monoesters could be identified based on accurate masses (Table 1). Mass errors were typically less than 10 mDa (except for arnitriol A 3-palmitate, **2b**, 13.9 mDa). Product-ion scans (CE at 17 V) provided characteristic fragmentations of the substances and confirmed the correct structural annotations. Detailed information for each sample with chromatogram, fragmentation spectrum and interpretation of characteristic fragments is provided as Supplementary Material. Interestingly, product ion spectra of the monoesters revealed that they are fairly stable and preferentially lose water rather than the fatty acid residue. However, the corresponding fragment ions with loss of R-COOH are also always observed in the product ion spectra.
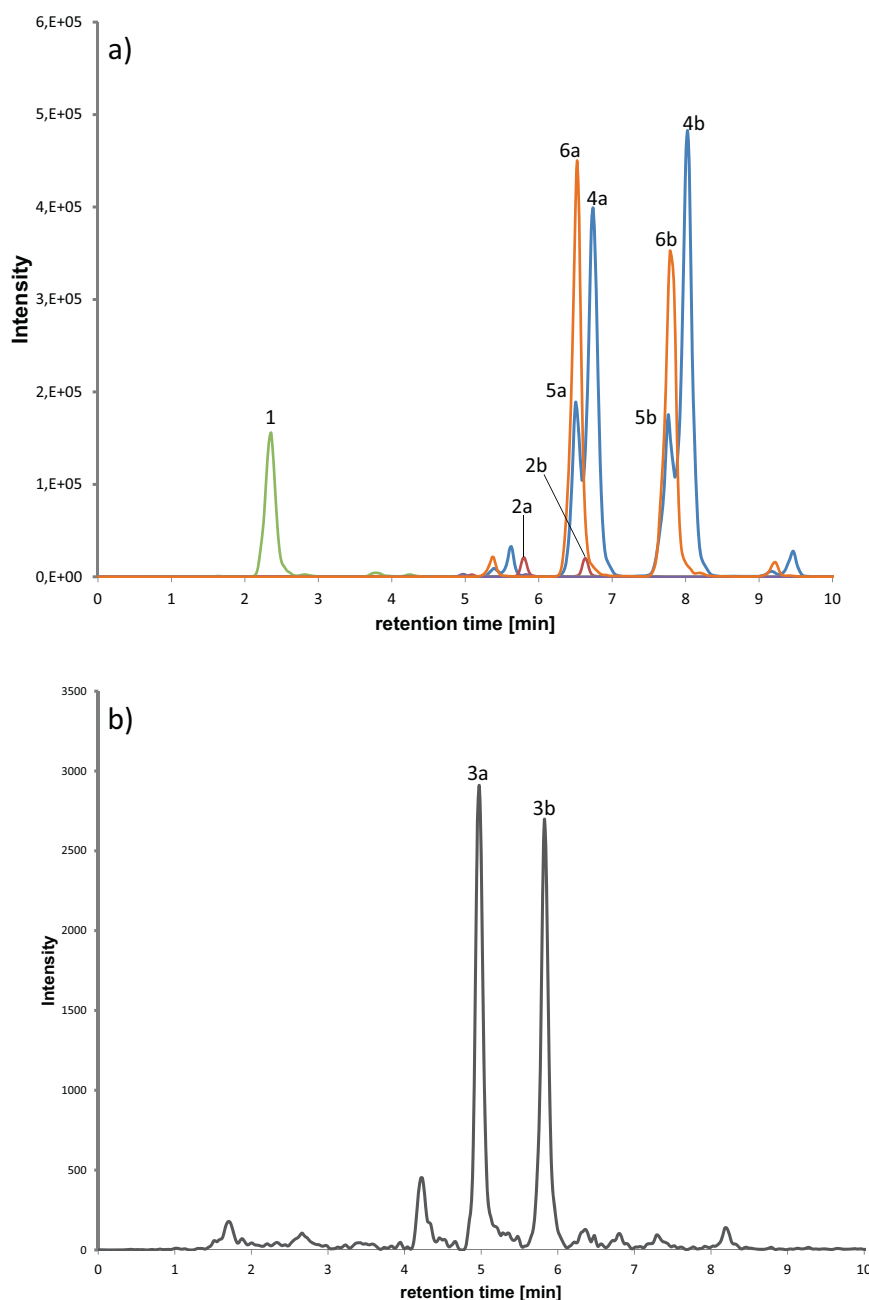
Fig. 2. Separation of triterpene monoesters via C18 chromatography. Overlay of XICs on in-source decay triterpene skeleton fragments ±10 mDa [M+H-2H$_2$O]$^+$ of faradiol and [M+H-fatty acid-nH$_2$O]$^+$ of monoesters detected in distinct samples. (**a**) faradiol (**1**) (*m/z* 407.3672), arnitriol A 3-myristate (**2a**) and -palmitate (**2b**) (*m/z* 405.3515, faradiol 3-myristate (**4a**) and –palmitate (**4b**) (*m/z* 407.3672), arnidiol 3-myristate (**5a**) -palmitate (**5b**) (*m/z* 407.3672), and maniladiol 3-myristate (**6a**) and -palmitate (**6b**) (*m/z* 407.3672). (**b**) lupane-3β,16β,20-triol 3-myristate (**3a**) and -palmitate (**3b**) (*m/z* 407.3672). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3. Identification of the novel faradiol fatty acid diesters by NMR

Re-investigation of the triterpenes of *C. officinalis* flowers also afforded a fraction with triterpene diesters (details are given in Experimental, 2.4. Extraction and isolation). ^{1}H NMR of the mixture showed the presence of six singlet methyl groups [δ$_H$ 0.86 (H-23), 0.86 (H-24), 0.90 (H-25), 1.08 (H-26), 1.10 (H-27), 0.84 (H-28)], an olefinic methyl group [δ$_H$ 1.67 (H-30)], a doublet methyl group [δ$_H$ 1.03 (H-29)], and an olefinic proton [δ$_H$ 5.28 (H-21)]. The occurrence of two oxymethine proton signals that were shifted downfield [δ$_H$ 4.50 (H-3), 4.88 (H-16)] indicated two esterified hydroxyl groups. The strong signal near δ$_H$ 1.28 for (CH$_2$)$_n$ in the ^{1}H spectrum and the long-range correlation between C-1′ at δ$_c$ 173.6

and H-2′ at δ$_H$ 2.33 indicated the occurrence of long chain fatty acids as acyl moieties. All ^{1}H NMR data were in agreement with a pentacyclic triterpene of the 20-taraxastene type with two esterified hydroxyl groups [26] which was confirmed by the ^{13}C NMR data (see Table 2) [26,34]. ^{1}H and ^{13}C NMR assignments were verified by HMBC and HSQC experiments.

Determination of the acyl moieties was not possible in the mixture. Preliminary studies by HR-APCI-MS (positive mode after direct infusion) only suggested the occurrence of myristic and palmitic acid as acyl moiety. However, a newly developed HPLC-QTOF-MS method enabled the identification of the complex mixture which is composed of faradiol 3,16-dimyristate

Table 2
^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) spectroscopic data for faradiol diesters (**7**–**10**) in CDCl$_3$ (δ in ppm, *J* in Hz).

Position	δ$_C$	δ$_H$ (*J* in Hz)
1	37.8	n.d.
2	22.6	1.63 m
3	80.4	4.50 dd (11.2, 5.6)
4	37.8	
5	55.3	0.84 m
6	18.0	1.42 m
7	34.1	1.41
8	41.1	
9	49.8	1.34 m
10	36.8	
11	21.4	1.59 m
12	27.0	2.08 m
13	38.7	n.d.
14	42.7	
15	33.1	1.35 m
		1.78 m
16	79.2	4.88 dd (11.6, 4.4)
17	37.8	
18	47.6	1.19 m
19	35.7	1.68 m
20	139.6	
21	117.9	5.28 brd (7.2)
22	38.7	1.65 m
		1.85 m
23	27.9	0.86 s
24	16.5	0.86 s
25	16.3	0.90 s
26	16.0	1.08 s
27	16.0	1.10 s
28	12.7	0,84 s
29	22.3	1.03 d (6.4)
30	21.4	1.67 s
acyl		
1′	173.6	
2′	34.8	2.33 t (7.2)
3′	25.1	1.33 m
4′–13′ (4′–11′)[a]	29.1–29.7	1.27 m
14′ (12′)[a]	31.9	1.29 m
15′ (13′)[a]	22.6	1.30 m
16′ (14′)[a]	14.1	0.90 t (7.2)

[a] In parenthesis: carbon for myristoyl, n.d. not determined.

(**7**), faradiol 3-myristate,16-palmitate (**8**) faradiol 3-palmitate,16-myristate (**9**) and faradiol 3,16-dipalmitate (**10**), (see next chapter).

3.4. HPLC-APCI-QTOF MS characterization of faradiol diesters

NMR analyses suggested the occurrence of faradiol diesters. To unambiguously identify the acyl moieties faradiol diesters were synthesized as standards and a new method was developed which allowed detection of the quasi-molecular ions and a (partial) separation of the diesters and mixed diesters, respectively. Thereby, liquid chromatography was based on an in-house developed C30 phase obtained by controlled surface modification of 5 μm (100 Å) silica particles with triacontyldimethylchlorosilane. This C30 stationary phase showed a better separation of the diesters (faradiol 3,16-dimyristate, FMM, and faradiol 3,16-dipalmitate, FPP), in particular it allowed partial resolution of the two mixed faradiol diesters (faradiol 3-myristate,16-palmitate, FMP, and faradiol 3-palmitate,16-myristate, FPM), which was not possible by non-aqueous RPLC on C18 phase, not even with highly efficient core-shell columns. It was used an in-house developed C30 phase instead of a commercial C30 column because it was hypothesized that selectivity could be tailored via the C30 ligand density which indeed was the case (data not shown). Elution conditions needed to be adjusted in order to be able to elute the highly lipophilic faradiol diesters in a reasonable time. Thus, acetonitrile was replaced by

MeOH as the solvent with weak elution strength in channel A and MeOH in channel B substituted by the stronger eluting 2-propanol. The corresponding gradient from 0 to 100% B over 25 min provided a reasonable separation (Fig. 3a). This LC method offers sufficient methylene selectivity to distinguish chromatographically between corresponding myristic and palmitic acid diesters of faradiol. Furthermore, some minor selectivity exists for the mixed diesters (FMP and FPM) which can be partially resolved (Rs ∼0.4). This, along with characteristic fragmentation spectra (*vide infra*), turned out to be sufficient for unequivocal identification and confirmation of the presence of FMP and FPM.

Due to strong in-source decay of the diesters, MS detection in this HPLC-APCI-QTOF MS run was based on monitoring the 3-monoester ion because of preferential loss of the fatty acid at 16-position (Fig. 4). This characteristic fragmentation gave further indication of the presence of both FMP and FPM because no peak would be detectable at the position of the mixed diesters in the chromatogram if the corresponding compound with the partial structure of the respective ester at C-3 is not present in the sample. Since a peak shows up at the retention time of the mixed diester in both traces, the XIC for faradiol 3-myristate fragment at *m/z* 635.5761 (blue trace) as well as the XIC for faradiol 3-palmitate fragment at *m/z* 663.6074 (red trace), we can safely conclude that both FMP and FPM are present in this sample (in Fig. 3a a mixture of faradiol mono- and diester standards is shown). Further, Fig. 3b shows an XIC of the TOF-MS run extracted for 407.3673 (±2.5 mDa) corresponding to the 20-taraxastene skeleton obtained for all compounds in the sample with this moiety obtained due to in-source fragmentation. In this chromatogram it becomes evident that also the faradiol 16-myristate and faradiol 16-palmitate monoesters are present in this standard mixture, eluting just before the corresponding 3-fatty acid monoesters.

The spectra in Fig. 4a and b confirm that loss of fatty acid in diesters occurs preferentially in position 16. Fig. 5 shows the fragmentation pattern concluded from the MS analysis of the faradiol esters. The MS/MS spectrum of faradiol 3-myristate,16-palmitate (FMP) documents that this component loses primarily palmitic acid and thus the fragment ion faradiol 3-myristate at *m/z* 635.5785 is more abundant than the corresponding faradiol 16-palmitate at *m/z* 663.6035 (Fig. 4a). Likewise, the fragment ion faradiol 3-palmitate at *m/z* 663.6057 is more abundant than the corresponding faradiol 16-myristate at *m/z* 635.5782 in the MS/MS spectrum of faradiol 3-palmitate,16-myristate (FPM) (Fig. 4b). These data clearly confirm the presence of both mixed diesters in the synthesized mixed standard and provide an additional means for the unequivocal assignment of the faradiol diesters.

Mass spectrometric detection of the intact faradiol diesters, however, is even more problematic and challenging than of the triterpene monoesters. The ester group at C-16 is particularly prone to neutral loss of fatty acid already in the ion-source, even in APCI mode. In a conventional APCI$^+$-TOF-MS survey scan, no [M + H]$^+$ ion can be detected due to insufficient signal intensity of this quasi-molecular ion and poor detection sensitivity in this mode. Thus, an MS/MS-high sensitivity scan mode with low CE of 5 V and multichannel averaging acquisition (MCA)-mode (1 Hz, time bins 80) has been used. With these settings it was possible to detect the [M + H]$^+$ quasi-molecular ions of the diesters by direct infusion experiments of the mixture containing faradiol diesters **7**–**10**. The results are summarized in Table 3. As can be seen, experimentally found accurate masses for the faradiol diesters are in good agreement with the theoretically calculated ones. Mass errors less than 12 mDa (<13 ppm) can, together with characteristic fragmentation patterns, be regarded as experimental evidence for a successful identification of novel faradiol diesters.

The detection of the diesters in the direct infusion experiments has been confirmed by the analysis of the standard mixture contain-
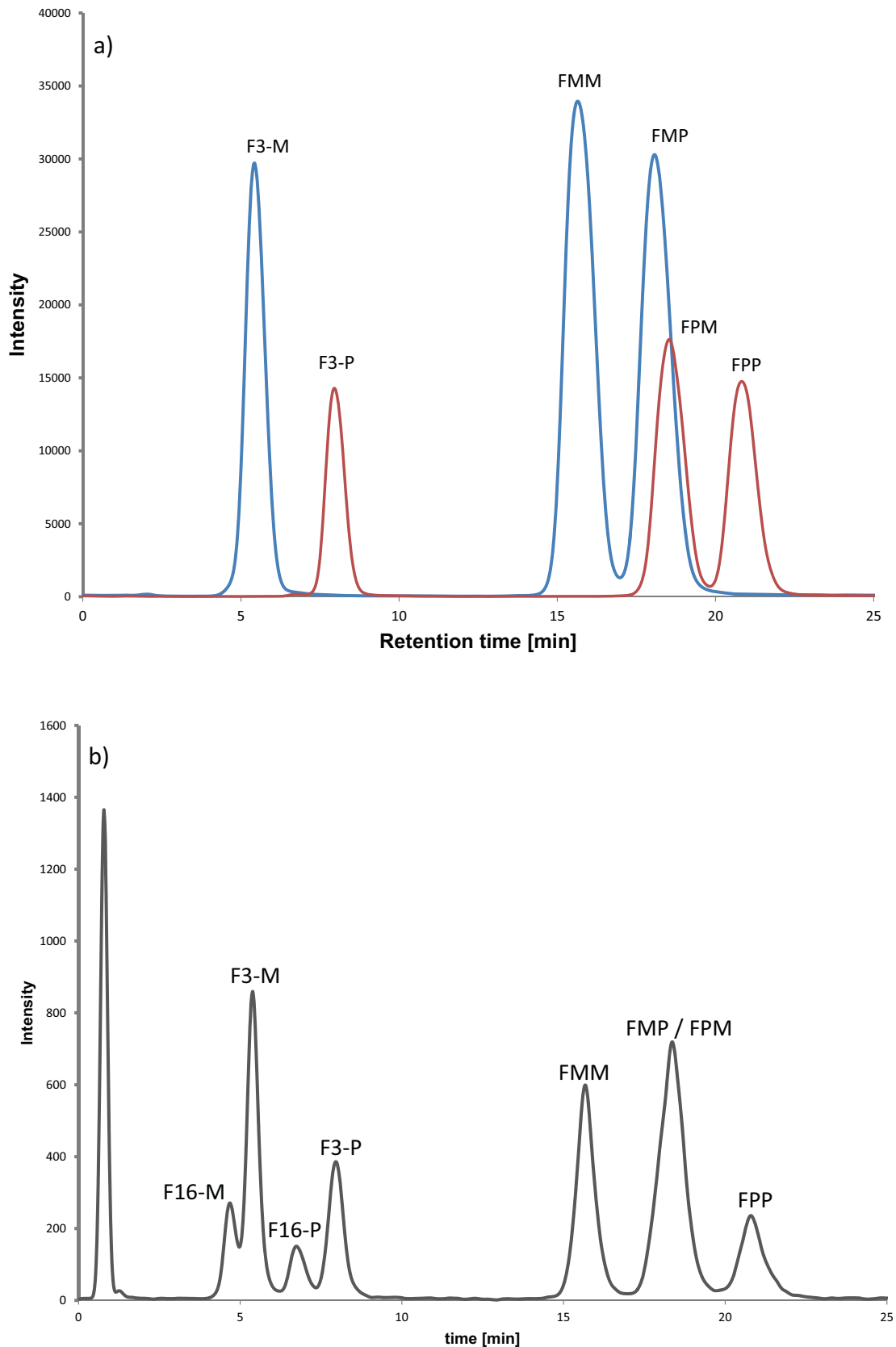
Fig. 3. Mixture of synthetic faradiol mono- and diester standards as measured by HPLC-APCI-QTOF MS on C30 column. (**a**) TOF-MS XICs extracted for m/z 635.5761 ($\pm$2.5 mDa) (faradiol 3-myristate − H$_2$O/faradiol 3,16-dimyristate − myristic acid/faradiol 3-myristate,16-palmitate − palmitic acid) (blue trace) and m/z 663.6074 ($\pm$2.5 mDa) (faradiol 3-palmitate − H$_2$O/faradiol 3,16-dipalmitate-palmitic acid/faradiol 3-palmitate,16-myristate − myristic acid) (red trace) (see also Fig. 5). (**b**) TOF-MS XIC on m/z 407.3673 ($\pm$2.5 mDa) corresponding to the 20-taraxastene skeleton in-source fragment. F3-M and F-3P: faradiol 3-myristate and -palmitate, F16-M and F16-P: faradiol 16-myristate and -palmitate. FMM: faradiol 3,16-dimyristate, FMP: faradiol 3-myristate,16-palmitate, FPM: faradiol 3-palmitate,16-myristate, FPP: faradiol 3,16-diplamitate, mixed diesters (FMP and FPM) appear as one peak with a slight shoulder. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
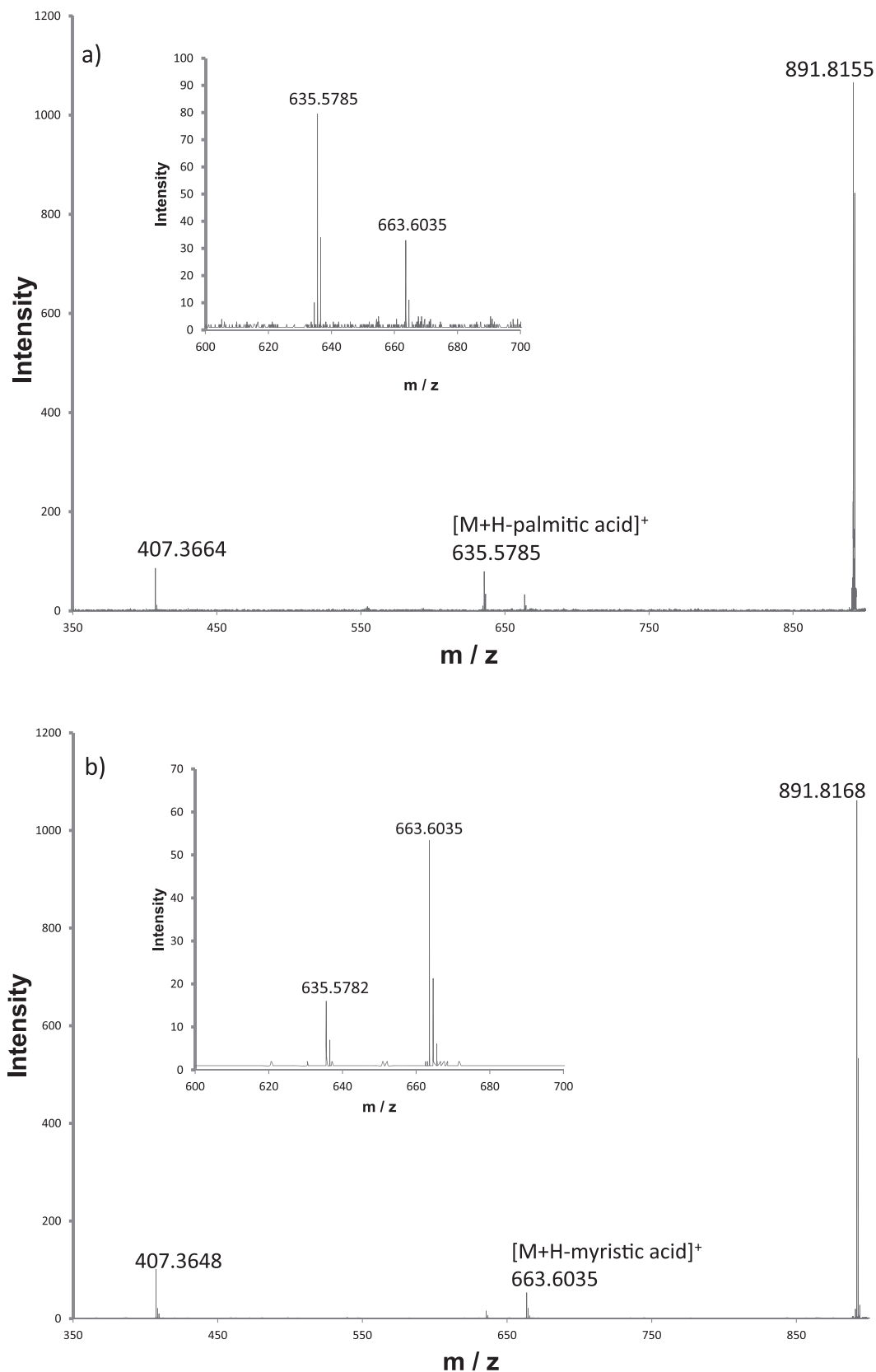
Fig. 4. Infusion MS spectra of standards show that at low CE (7 V) the fatty acid loss occurs preferentially in 16-position. (**a**) FMP standard selectively synthesized from faradiol 3-myristate. (**b**) FPM standard selectively synthesized from faradiol 3-palmitate. It can be observed from the insert that the corresponding 16-monoester fragment is detected with low abundance only.

Fig. 5. Fragmentation pattern of faradiol mono- and diesters concluded from the MS data. Decay of H_2O or fatty acid at C-16 leads to the faradiol 3-myristate fragment at m/z 635.5761 and decay of myristic acid at C-3 to the fragment at m/z 407.3672 of the 20-taraxastene skeleton (fragmentation similar for faradiol mono- and diesters with palmitic acid at C-3 leading to fragments at m/z 663.6074 and m/z 407.3672).

Fig. 6. Analysis of mixture **7**–**10** of *C officinalis* by HPLC-APCI-QTOF MS on C30 column. TOF-MS XICs extracted for m/z 635.5761($\pm$2.5 mDa) (blue trace) and m/z 663.6074 ($\pm$2.5 mDa) (red trace). Formation of the fragments similar as described for Fig. 3. (see also Fig. 5) Dotted lines represent overlay of standard traces as shown in Fig. 3a. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Identification of faradiol diesters in the mixture containing **7–10** in *C. officinalis* flowers. Note that accurate masses have been determined via direct infusion MS/MS in MCA mode, as the [M + H]$^+$ peak is unstable.

Compound	Formula	Calculated m/z	Found at m/z	Δ mDa	Δ ppm	t_R (min)
7	$C_{58}H_{102}O_4$	863.7851	863.7815	−3.60	−4.17	15.68
8/9	$C_{60}H_{106}O_4$	891.81639	891.8181	1.71	1.92	18.10/18.54
10	$C_{62}H_{110}O_4$	919.84769	919.8358	−11.89	−12.93	20.84

Faradiol 3,16-dimyristate (**7**); faradiol 3-myristate, 16-palmitate (**8**); faradiol 3-palmitate,16-myristate (**9**); faradiol 3,16-dipalmitate (**10**) t_R refers to C30 chromatography.

ing FMM, FMP, FPM and FPP by C30-HPLC-APCI$^+$-QTOF-MS (Fig. 6). The most abundant peak occurs from the mixed faradiol diester **8** (peak at ca. 18 min) followed by compound **7** (peak at ca 15.5 min). The diester **10** (peak at ca. 21 min) could be detected in this sample as well but mainly as a minor component. Last but not least, also traces of the mixed diester **9** could be identified in the fraction of the *n*-hexane extract of *C. officinalis*.

4. Conclusions

Here we show that triterpenes with long chain fatty acids can be separated by using an uncommon non-aqueous RP-HPLC mode with an in-house synthesized C30 phase. Moreover, the development of a set of new APCI-QTOF-MS methods using the TripleTOF 5600+ mass spectrometer proved to be a valuable tool to determine the quasi-molecular ions of this type of triterpene esters. Thus the combination of 1D and 2D NMR and direct infusion APCI-QTOF MS experiments in MS/MS high sensitivity scan mode with low collision energy and multichannel averaging acquisition proved to be a successful approach to elucidate the composition of complex mixtures of triterpene diesters. The proposed methods also show good potential for routine analysis of *C. officinalis* extracts to monitor monoesters and diesters of various triterpenes. For that purpose, however, modern highly efficient commercial C30 phases based on core–shell particles (e.g. Accucore C30) or sub-2 μm particles (which are expected to become also available in due time) are recommended and should be utilized.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jpba.2015.10.025.

References

[1] Committee on herbal medicinal products (HMPC), Community Herbal Monograph on *Calendula officinalis* L., flos, London, 2008.

[2] R. Della Loggia, A. Tubaro, S. Sosa, H. Becker, S. Saar, O. Isaac, The role of triterpenoids in the topical anti-inflammatory activity of *Calendula officinalis* flowers, Planta Med. 60 (1994) 516–520, http://dx.doi.org/10.1055/s-2006-959562.

[3] K. Zitterl-Eglseer, S. Sosa, J. Jurenitsch, M. Schubert-Zsilavecz, R. Della Loggia, A. Tubaro, M. Bertoldi, T. Franz, Anti-oedematous activities of the main triterpendiol esters of marigold (*Calendula officinalis* L.), J. Ethnopharmacol. 57 (1997) 139–144.

[4] H. Neukirch, M. D'Ambrosio, S. Sosa, G. Altinier, R. Della Loggia, A. Guerriero, Improved anti-inflammatory activity of three new terpenoids derived, by systematic chemical modifications, from the abundant triterpenes of the flowery plant *Calendula officinalis*, Chem. Biodivers. 2 (2005) 657–671, http://dx.doi.org/10.1002/cbdv.200590042.

[5] M. Hamburger, S. Adler, D. Baumann, A. Förg, B. Weinreich, Preparative purification of the major anti-inflammatory triterpenoid esters from marigold (*Calendula officinalis*), Fitoterapia 74 (2003) 328–338, http://dx.doi.org/10.1016/s0367-326x(03)00051-0.

[6] K. Zitterl-Eglseer, G. Reznicek, J. Jurenitsch, J. Novak, W. Zitterl, C. Franz, Morphogenetic variability of faradiol monoesters in marigold *Calendula officinalis* L, Phytochem. Anal. 12 (2001) 199–201, http://dx.doi.org/10.1002/pca.582.

[7] Teedrogen und Phytopharmaka, in: M. Wichtl (Ed.), 5th ed., Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart, 2009.

[8] Z. Wojciechowski, M. Bocheńska-Hryniewicz, B. Kucharczak, Z. Kasprzyk, Sterol and triterpene alcohol esters from *Calendula officinalis*, Phytochemistry 11 (1972) 1165–1168, http://dx.doi.org/10.1016/s0031-9422(00)88473-3.

[9] A.N. Assimopoulou, V.P. Papageorgiou, GC–MS analysis of penta- and tetra-cyclic triterpenes from resins of *Pistacia* species. Part I. *Pistacia lentiscus* var. Chia, Biomed. Chromatogr. 19 (2005) 285–311, http://dx.doi.org/10.1002/bmc.454.

[10] H. Budzikiewicz, J.M. Wilson, C. Djerassi, Mass spectrometry in structural and stereochemical problems. XXXII. Pentacyclic triterpenes, J. Am. Chem. Soc. 85 (1963) 3688–3699.

[11] C. Djerassi, H. Budzikiewicz, J. Wilson, Mass spectrometry in structural and stereochemical problems unsaturated pentacyclic triterpenoids, Tetrahedron Lett. 3 (1962) 263–270, http://dx.doi.org/10.1016/s0040-4039(00)70864-1.

[12] V. Elias, B. Simoneit, A. Pereira, J. Cardoso, Mass spectra of triterpenyl alkanoates, novel natural products, J. Mass Spectrom. 32 (1997) 1356–1361, http://dx.doi.org/10.1002/(SICI)1096-9888(199712)32:12<1356::AID-JMS602>3.0.CO;2-O.

[13] F. de L. Moreira, G.H.B. de Souza, I.V. Rodrigues, N.P. Lopes, A.R.M. de Oliveira, A non-michaelian behavior of the in vitro metabolism of the pentacyclic triterpene alfa and beta amyrins by employing rat liver microsomes, J. Pharm. Biomed. Anal. 84 (2013) 14–19, http://dx.doi.org/10.1016/j.jpba.2013.05.038.

[14] P.A.B. Ramos, Â.R. Guerra, O. Guerreiro, C.S.R. Freire, A.M.S. Silva, M.F. Duarte, A.J.D. Silvestre, Lipophilic extracts of *Cynara cardunculus* L. var. *altilis* (DC): a source of valuable bioactive terpenic compounds, J. Agric. Food Chem. 61 (2013) 8420–8429, http://dx.doi.org/10.1021/jf402253a.

[15] S. Wang, L. Chen, J. Leng, P. Chen, X. Fan, Y. Cheng, Fragment ion diagnostic strategies for the comprehensive identification of chemical profile of Gui-Zhi-Tang by integrating high-resolution MS, multiple-stage MS and UV information, J. Pharm. Biomed. Anal. 98 (2014) 22–35, http://dx.doi.org/10.1016/j.jpba.2014.05.013.

[16] S. Ji, Q. Wang, X. Qiao, H. Guo, Y. Yang, T. Bo, C. Xiang, D.A. Guo, M. Ye, New triterpene saponins from the roots of *Glycyrrhiza yunnanensis* and their rapid screening by LC/MS/MS, J. Pharm. Biomed. Anal. 90 (2014) 15–26, http://dx.doi.org/10.1016/j.jpba.2013.11.021.

[17] S. Wang, L. Qiao, X. Shi, C. Hu, H. Kong, G. Xu, On-line stop-flow two-dimensional liquid chromatography–mass spectrometry method for the separation and identification of triterpenoid saponins from ginseng extract, Anal. Bioanal. Chem. 407 (2015) 331–341, http://dx.doi.org/10.1007/s00216-014-8219-4.

[18] X. Zhang, Q. Li, C. Lv, H. Xu, X. Liu, Z. Sui, K. Bi, Characterization of multiple constituents in Kai-Xin-San prescription and rat plasma after oral administration by liquid chromatography with quadrupole time-of-flight tandem mass spectrometry, J. Sep. Sci. 38 (2015) 2068–2075, http://dx.doi.org/10.1002/jssc.201500123.

[19] G.A. van der Doelen, K.J. van den Berg, J.J. Boon, N. Shibayama, E. René de la Rie, W.J.L. Genuit, Analysis of fresh triterpenoid resins and aged triterpenoid varnishes by high-performance liquid chromatography-atmospheric pressure chemical ionisation (tandem) mass spectrometry, J. Chromatogr. A 809 (1998) 21–37, http://dx.doi.org/10.1016/s0021-9673(98)00186-1.

[20] C. Li, C. Huang, T. Lu, L. Wu, S. Deng, R. Yang, J. Li, Tandem mass spectrometric fragmentation behavior of lignans, flavonoids and triterpenoids in *Streblus asper*, Rapid Commun. Mass Spectrom. 28 (2014) 2363–2370, http://dx.doi.org/10.1002/rcm.7035.

[21] K. Naumoska, I. Vovk, Analysis of triterpenoids and phytosterols in vegetables by thin-layer chromatography coupled to tandem mass spectrometry, J. Chromatogr. A 1381 (2015) 229–238, http://dx.doi.org/10.1016/j.chroma.2015.01.001.

[22] C. Fanali, M. Beccaria, S. Salivo, P. Tranchida, G. Tripodo, S. Farnetti, L. Dugo, P. Dugo, L. Mondello, Non-polar lipids characterization of Quinoa (*Chenopodium*

quinoa) seed by comprehensive two-dimensional gas chromatography with flame ionization/mass spectrometry detection and non-aqueous reversed-phase liquid chromatography with atmospheric pressure chemical ionization mass spectrometry detection, J. Sep. Sci. (2015), http://dx.doi.org/10.1002/jssc.201500466.

[23] S. Vahur, A. Teearu, T. Haljasorg, P. Burk, I. Leito, I. Kaljurand, Analysis of dammar resin with MALDI-FT-ICR-MS and APCI-FT-ICR-MS, J. Mass Spectrom. 47 (2012) 392–409, http://dx.doi.org/10.1002/jms.2971.

[24] J.-L. Wolfender, G. Marti, A. Thomas, S. Bertrand, Current approaches and challenges for the metabolite profiling of complex natural extracts, J. Chromatogr. A 1382 (2015) 136–164, http://dx.doi.org/10.1016/j.chroma.2014.10.091.

[25] H. Neukirch, M. D'Ambrosio, J.D. Via, A. Guerriero, Simultaneous quantitative determination of eight triterpenoid monoesters from flowers of 10 varieties of *Calendula officinalis* L. and characterisation of a new triterpenoid monoester, Phytochem. Anal. 15 (2004) 30–35, http://dx.doi.org/10.1002/pca.739.

[26] T.J. Schmidt, J. von Raison, G. Willuhn, New triterpene esters from flowerheads of *Arnica lonchophylla*, Planta Med. 70 (2004) 967–977, http://dx.doi.org/10.1055/s-2004-832624.

[27] J.S. Pyrek, Terpenes of compositae plants. Part XI. Structures of heliantriols B0, B1 B2 and A1, new pentacyclic triterpenes from *Helianthus annuus* L. and *Calendula officinalis* L, Pol. J. Chem. 53 (1979) 2465–2490.

[28] J.S. Pyrek, E. Baranowska, Terpenes of Compositae plants. XII. Mass spectral fragmentation of 16-substituted taraxanes, taraxenes and psi-taraxenes, Pol. J. Chem. 52 (1978) 97–106.

[29] H. Nells, A. Deleenheer, Isocratic non-aqueous reversed-phase liquid-chromatography of carotenoids, Anal. Chem. 55 (1983) 270–275.

[30] M.D. Grynbaum, P. Hentschel, K. Putzbach, J. Rehbein, M. Krucker, G. Nicholson, K. Albert, Unambiguous detection of astaxanthin and astaxanthin fatty acid esters in krill (*Euphausia superba* Dana), J. Sep. Sci. 28 (2005) 1685–1693.

[31] K. Holtin, M. Kuehnle, J. Rehbein, P. Schuler, G. Nicholson, K. Albert, Determination of astaxanthin and astaxanthin esters in the microalgae *Haematococcus pluvialis* by LC-(APCI)MS and characterization of predominant carotenoid isomers by NMR spectroscopy, Anal. Bioanal. Chem. 395 (2009) 1613–1622, http://dx.doi.org/10.1007/s00216-009-2837-2.

[32] L.J.R. Barrón, G. Santa-María, Non-aqueous reverse-phase HPLC analysis of triglycerides, Chromatographia 23 (1987) 209–214, http://dx.doi.org/10.1007/bf02311483.

[33] G.W.K. Fong, R.N. Johnson, B.T. Kho, Non-aqueous reverse phase liquid-chromatographic determination of vitamin-D2 in multivitamin tablets, using vitamin-D3 as internal standard, J. Assoc. Off. Anal. Chem. 66 (1983) 939–945.

[34] N. Boutaghane, L. Voutquenne-Nazabadioko, A. Simon, D. Harakat, K. Benlabed, Z. Kabouche, A new triterpenic diester from the aerial parts of *Chrysanthemum macrocarpum*, Phytochem. Lett. 6 (2013) 519–525, http://dx.doi.org/10.1016/j.phytol.2013.06.009.

Mastering analytical challenges for the characterization of pentacyclic triterpene mono- and diesters of *Calendula officinalis* flowers by non-aqueous C30 HPLC and hyphenation with APCI-QTOF-MS

Christoph Nicolaus[a*], Adrian Sievers-Engler[b*], Renato Murillo[c], Michele D'Ambrosio[d], Michael Lämmerhofer[2#], Irmgard Merfort[1#]

[a] Institute of Pharmaceutical Sciences, Department of Pharmaceutical Biology and Biotechnology, University of Freiburg, Stefan-Meier-Strasse 19, D-79104 Freiburg, Germany

[b] Institute of Pharmaceutical Sciences, Department of Pharmaceutical Chemistry, University of Tübingen, Auf der Morgenstelle 8, D-72076 Tübingen, Germany

[c] Escuela de Quimica and CIPRONA, Universidad de Costa Rica, 2060 San Jose, Costa Rica

[d] Laboratorio di Chimica Bioorganica, Università degli Studi di Trento, Italy

*both authors contributed equally

 # Authors for correspondence:

Prof. Irmgard Merfort, e-mail: irmgard.merfort@pharmazie.uni-freiburg.de; Fax: ++49 761 203 8383

Prof. Michael Lämmerhofer, e-mail: michael.laemmerhofe@uni-tuebingen.de; Fax: ++49 7071 29 4565

UHPLC-APCI-QTOF MS characterization of isolated triperpene monoester samples

The following chromatograms, spectra and tables are in support of the data shown in Table 1 of the main document.
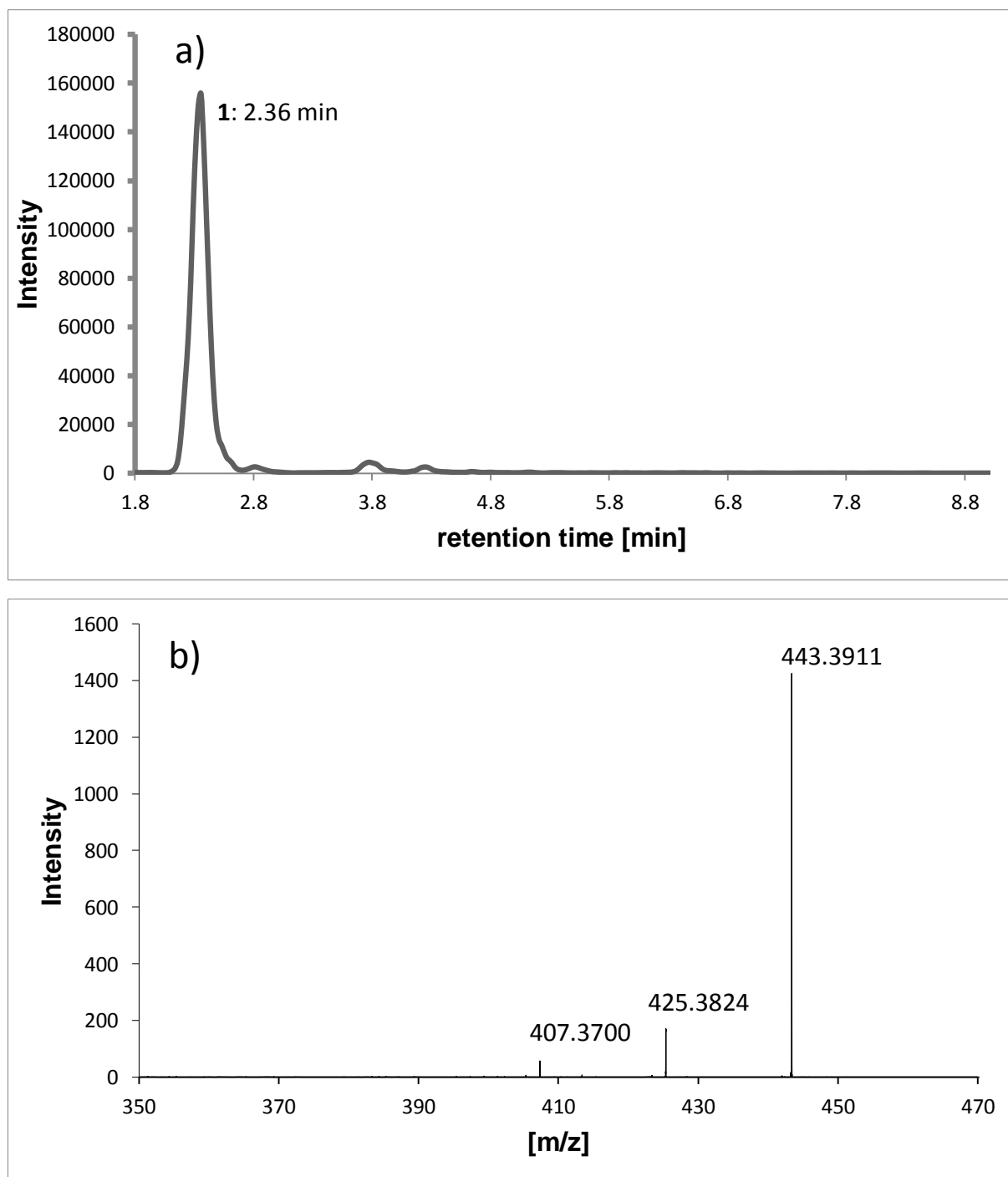
Fig. S1: Detection of faradiol (**1**) by UHPLC-APCI-QTOF-MS on Kinetex C18. **a)** XIC on in-source decay body fragment at 407.3672 (±10mDa) revealing a peak for faradiol (**1**) at

retention time of 2.36 min. **b)** Product ion scan on $[M+H]^+$ of **1** at t_R = 2.36 showing fragments characteristic for faradiol. The retention time and spectral data have been confirmed by an authentic standard.

Table S1: Summary of MS data from CID-MS/MS-spectra obtained for faradiol (**1**). Note that transition from #1 to #2 occurs as neutral loss.

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	443.3911	443.3884	2.8	6.9	$[M+H]^+$
2	425.3824	425.3778	4.6	10.8	$\#1-H_2O$
3	407.37	407.3672	2.7	6.1	$\#1-2H_2O$
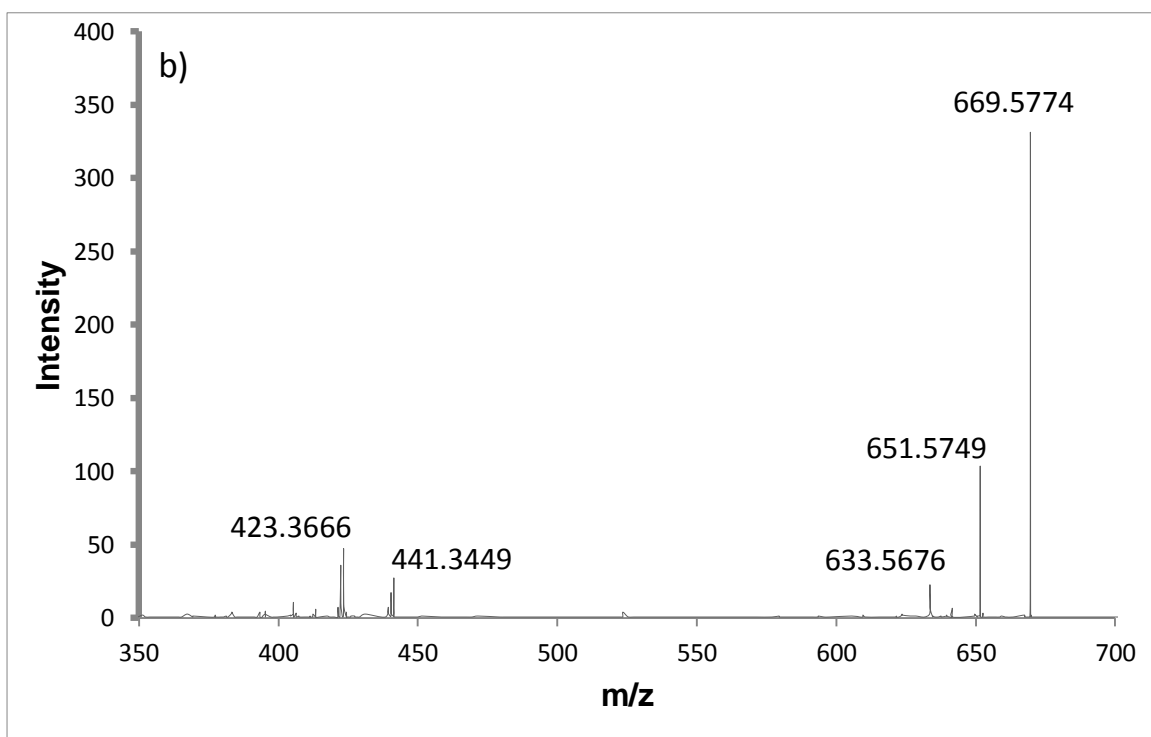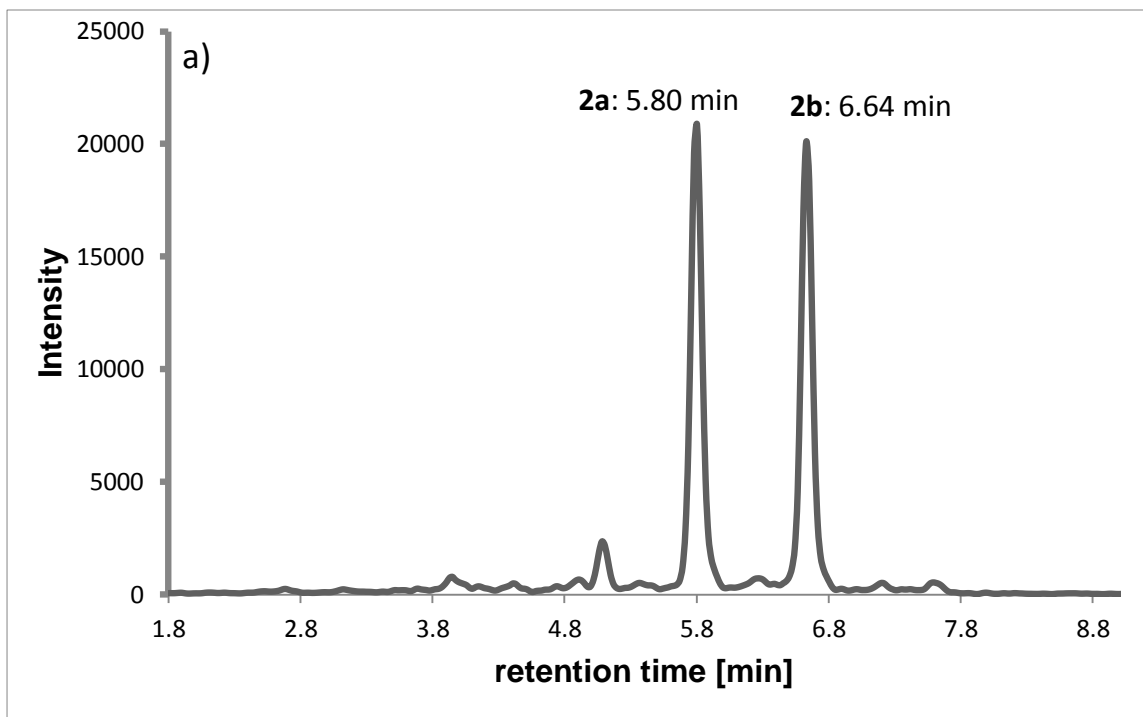
Fig. S2: Detection of arnitriol A 3-monoesters (**2a + b**) by UHPLC-APCI-QTOF-MS on Kinetex C18. **a)** XIC on in-source decay triterpene skeleton fragment at 405.3515 (±10 mDa), retention time 5.80 min for **2a** and 6.64 min for **2b**. **b)** Product ion scan on [M+H]$^+$ of **2a** showing characteristic fragments of arnitriol A 3-myristate eluted at t_R = 5.80 min. **c)** Product ion scan on [M+H]$^+$ of **2b** showing characteristic fragments of arnitriol A 3-palmitate eluted at t_R = 6.64 min.

Table S2: Summary of MS data from CID-MS/MS-spectra obtained for arnitriol A 3-myristate (**2a**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	669.5774	669.5816	13.8	20.6	$[M+H]^+$
2	651.5749	651.5710	15.8	24.3	#1-H_2O
3	633.5676	633.5604	4.9	7.7	#1-2H_2O
4	441.3449	441.3727	3.1	7.1	#1-myristic acid
5	423.3666	423.3621	4.4	10.5	#1-$H_2$0-myristic acid

Table S3: Summary of MS data from CID-MS/MS-spectra obtained for arnitriol A 3-palmitate (**2b**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	697.6195	697.6129	6.6	9.4	$[M+H]^+$
2	679.6081	679.6023	5.8	8.5	#1-H_2O
3	661.6058	661.5917	14.0	21.1	#1-2H_2O
4	441.3745	441.3727	1.8	4.1	#1-palmitic acid
5	423.3666	441.3727	4.4	10.5	#1-H_2O-palmitic acid

a)

3a: 4.97 min

3b: 5.83 min

b)

671.5430

653.5703

635.5575

425.3602

443.3623
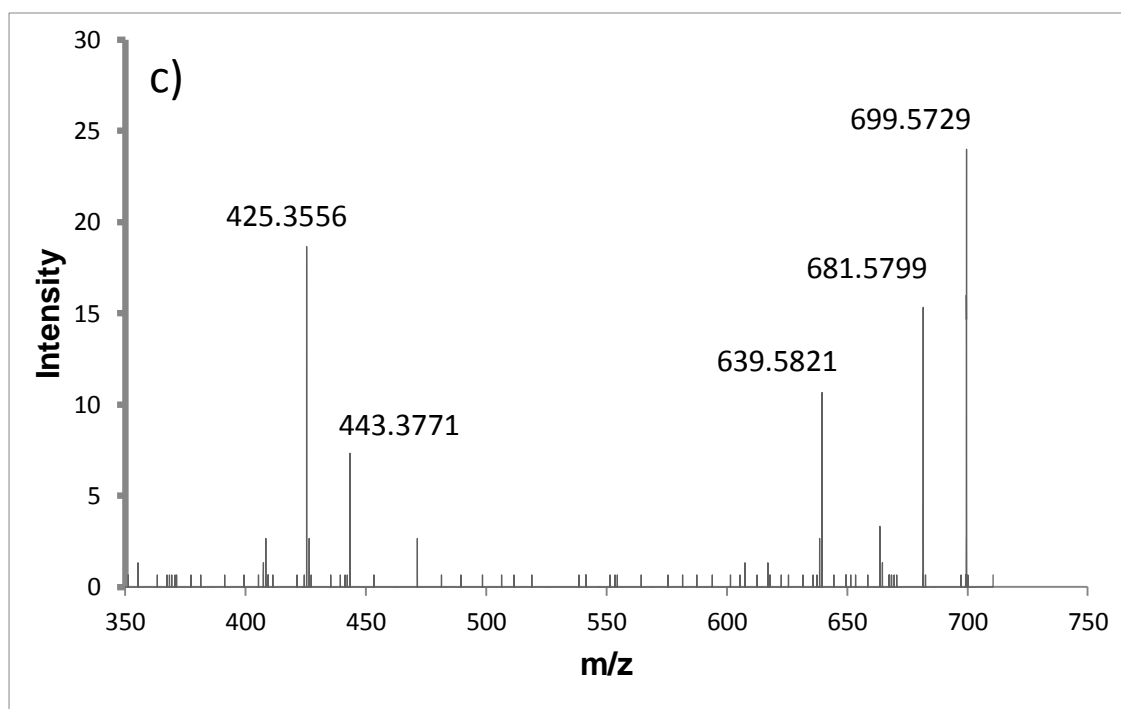
Fig. S3: Detection of lupane-3β,16β,20-triol 3-monoesters (**3a + b**) by UHPLC-APCI-QTOF-MS on Kinetex C18. **a)** XIC on in-source decay triterpene skeleton fragment at 407.3672 (±10 mDa), retention time 4.97 min for **3a** and 5.83 min for **3b**. **b)** Product ion scan on [M+H]$^+$ of **3a** showing characteristic fragments of lupane-3β,16β,20-triol 3-myristate eluted at t_R = 4.97 min. **c)** Product ion scan on [M+H]$^+$ of **3b** showing characteristic fragments of lupane-3β,16β,20-triol 3-palmitate eluted at t_R = 5.83 min.

Table S4: Summary of MS data from CID-MS/MS-spectra obtained for lupane-3β,16β,20-triol 3-myristate (**3a**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	671.543	671.5973	54.3	80.9	$[M+H]^+$
2	653.5703	653.5867	16.5	25.2	#1-H_2O
3	635.5575	635.5761	18.7	29.4	#1-2H_2O
4	443.3623	443.3878	26.1	58.9	#1-myristic acid
5	425.3602	425.3778	17.6	41.4	#1-H_2O-myristic acid

Table S5: Summary of MS data from CID-MS/MS-spectra obtained for lupane-3β,16β,20-triol 3-palmitate (**3b**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	699.5729	699.6286	55.7	79.6	$[M+H]^+$
2	681.5799	681.6180	38.2	56.0	#1-H_2O
3	639.5821	663.6074	11.1	17.3	#1-2H_2O
4	443.3771	443.3878	11.3	25.4	#1-palmitic acid
5	425.3556	425.3778	22.2	52.3	#1-H_2O-palmitic acid
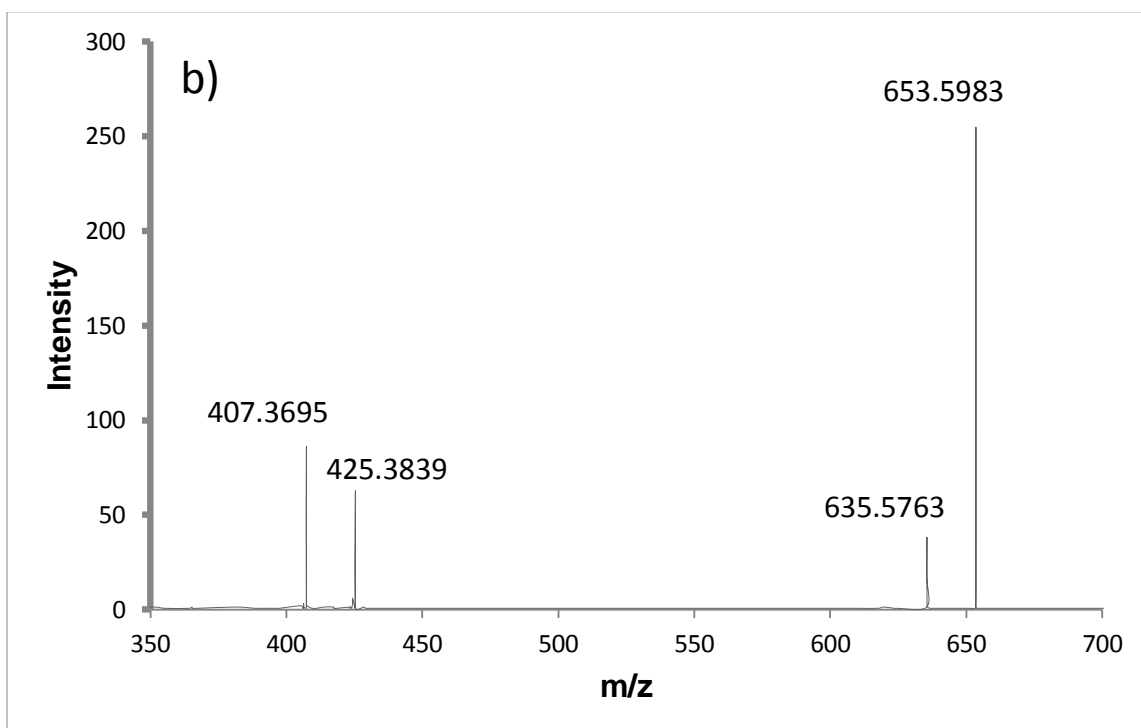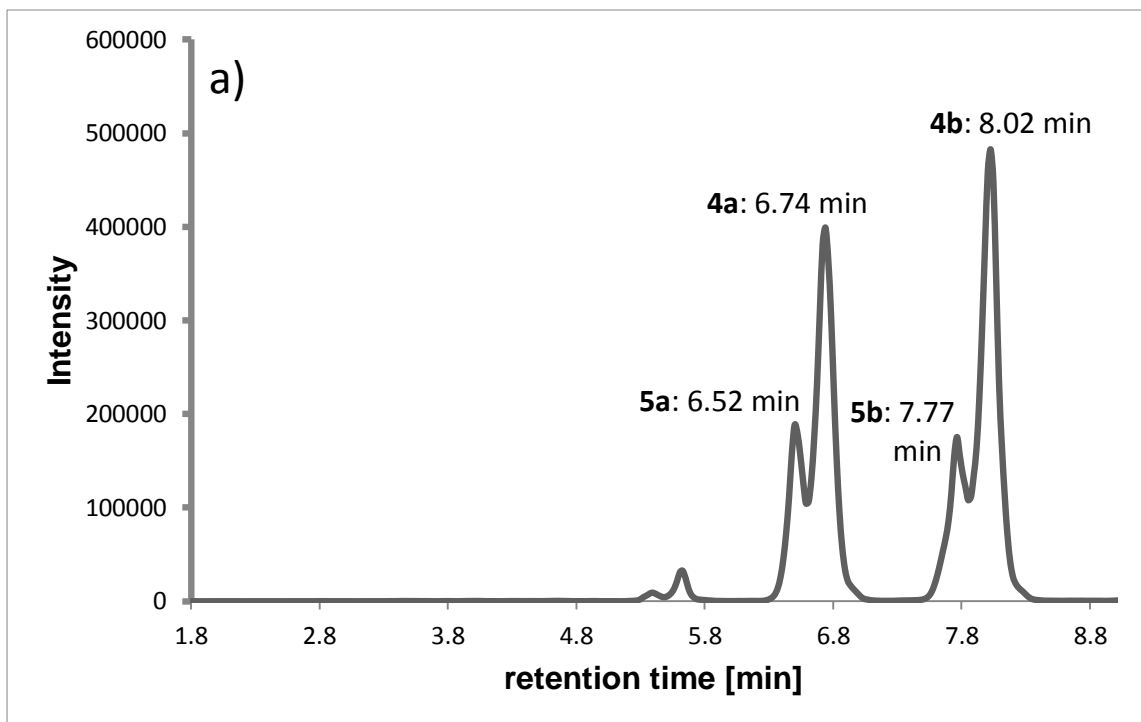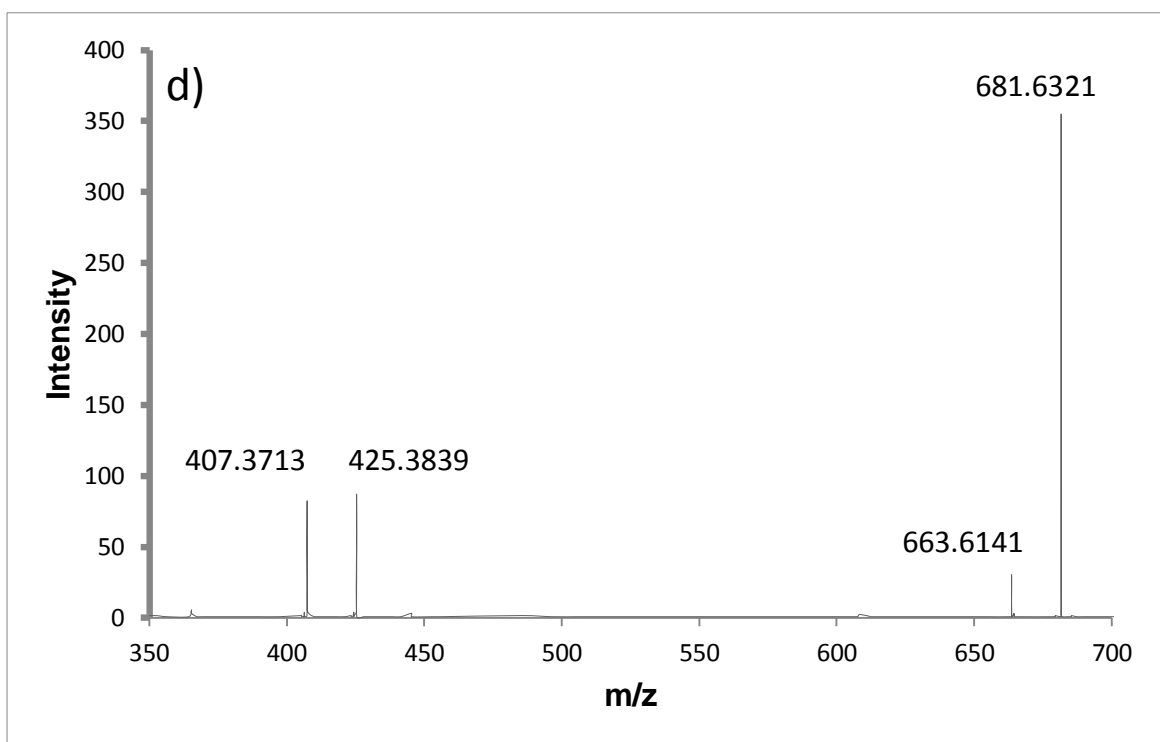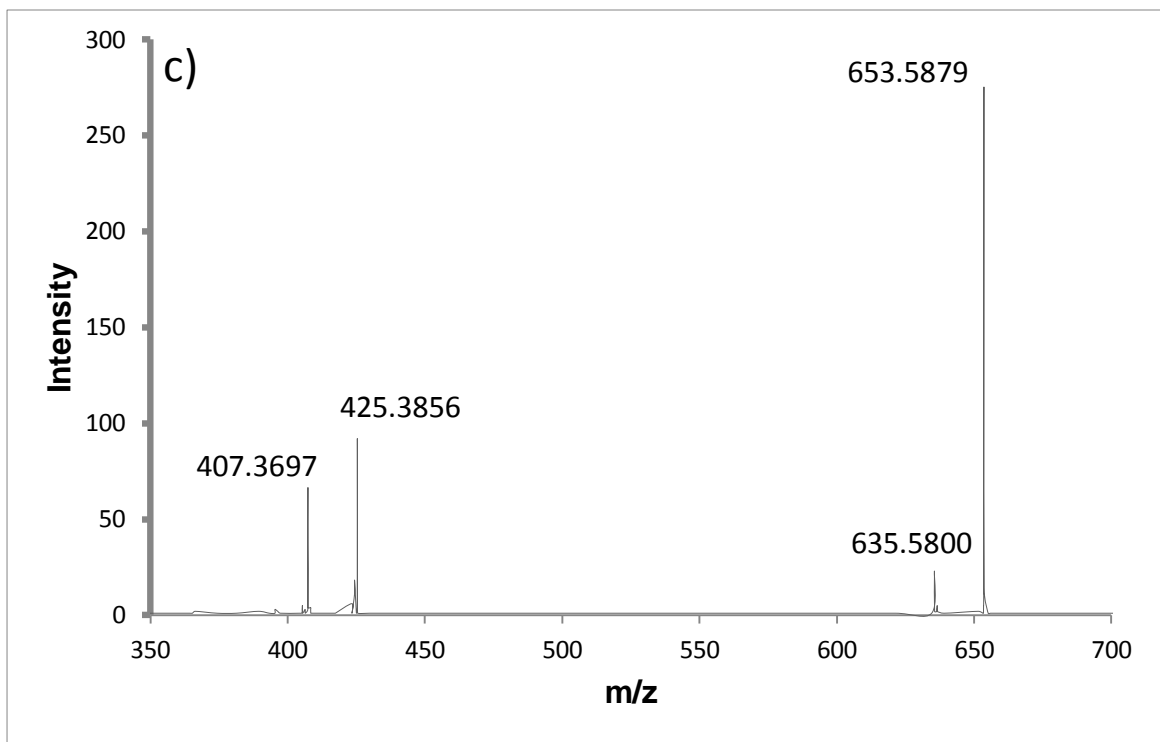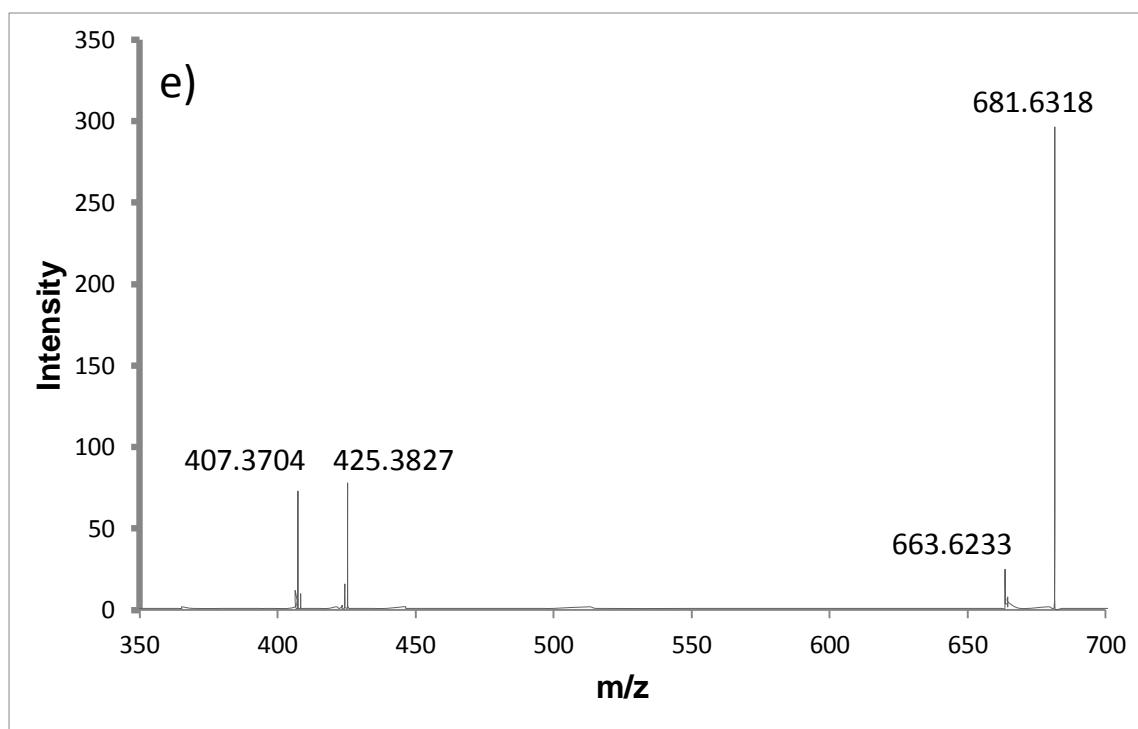
Fig. S4: Detection of faradiol 3- and arnidiol 3-monoesters (**4a + b** and **5a + b**, respectively) by UHPLC-APCI-QTOF-MS on Kinetex C18. **a)** XIC on in-source decay triterpene skeleton fragment at 407.3672 (±10 mDa), retention time 6.74 min for **4a** and 8.02 min for **4b**, 6.52 for **5a** and 7.77 for **5b**. **b)** Product ion scan on [M+H]$^+$ of **5a** showing characteristic fragments of arnidiol 3-myristate eluted at t_R = 6.52. **c)** Product ion scan on [M+H]$^+$ of **4a** showing characteristic fragments of faradiol 3-myristate eluted at t_R = 6.74. **d)** Product ion scan on [M+H]$^+$ of **5b** showing characteristic fragments of arnidiol 3-palmitate eluted at t_R = 7.77. **e)** Product ion scan on [M+H]$^+$ of **4b** showing characteristic fragments of faradiol 3- palmitate at t_R = 8.02.

Table S6: Summary of MS data from CID-MS/MS-spectra obtained for faradiol 3-myristate (**4a**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	653.5879	653.5867	1.2	1.8	[M+H]$^+$
2	635.58	635.5761	3.9	6.1	#1-H$_2$0
3	425.3856	425.3778	7.8	18.4	#1-myristic acid
4	407.3697	407.3672	2.5	6.0	#1-H$_2$O-myristic acid

Table S7: Summary of MS data from CID-MS/MS-spectra obtained for faradiol 3-palmitate (**4b**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	681.6318	681.6180	13.8	20.3	[M+H]$^+$
2	663.6233	663.6074	15.8	23.8	#1-H$_2$0
3	425.3827	425.3778	4.9	11.5	#1-palmitic acid
4	407.3704	407.3672	3.1	7.7	#1-H$_2$O-palmitic acid

Table S8: Summary of MS data from CID-MS/MS-spectra obtained for arnidiol 3-myristate (**5a**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	653.5983	653.5867	11.6	17.7	[M+H]$^+$
2	635.5763	635.5761	0.2	0.3	#1-H$_2$0
3	425.3839	425.3778	6.1	14.4	#1-myristic acid
4	407.3695	407.3672	2.2	5.5	#1-H$_2$O-myristic acid

Table S9: Summary of MS data from CID-MS/MS-spectra obtained for arnidiol 3-palmitate (**5b**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	681.6321	681.6180	14.1	20.7	$[M+H]^+$
2	663.6141	663.6074	6.6	10.0	#1-$H_2$0
3	425.3839	425.3778	6.1	14.4	#1-palmitic acid
4	407.3713	407.3672	4.1	10.1	#1-H_2O-palmitic acid
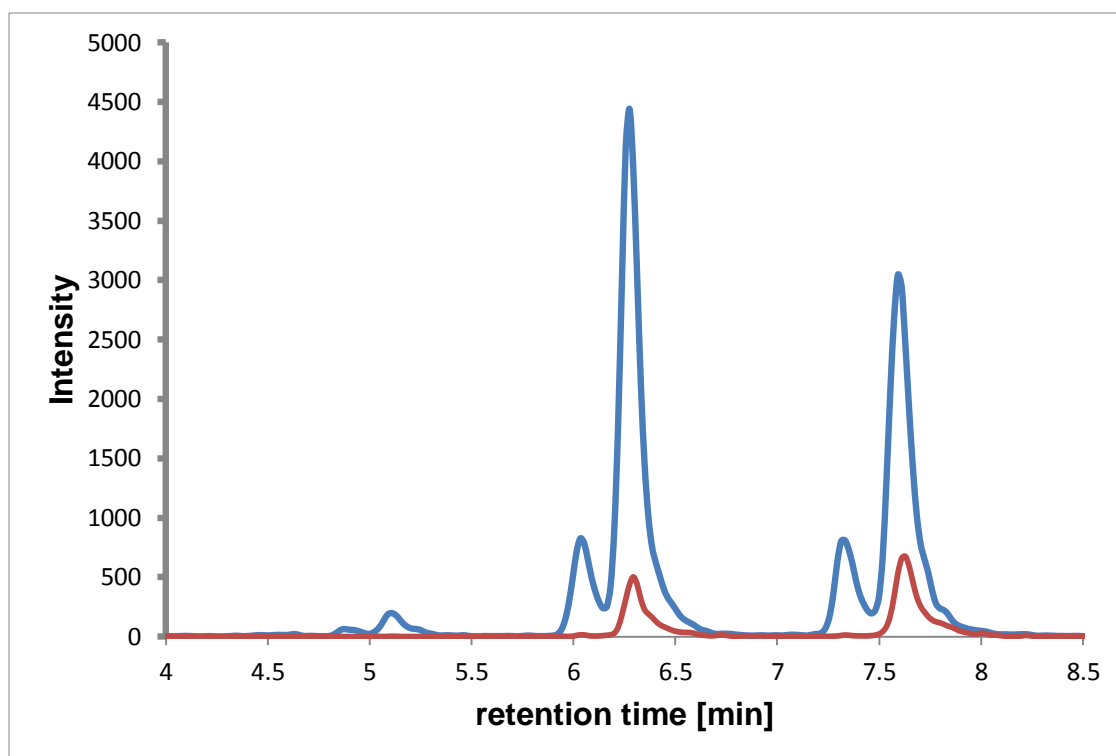
Fig. S5: Verification of the elution order of isobaric faradiol 3- (**4a + b**) and arnidiol 3-monoesters (**5a + b**). Blue: **4a + b** and **5a + b**. Red: faradiol 3-myristate and –palmitate standards mixture. Elution order: arnidiol 3-monoester are eluted before the corresponding faradiol 3-monoester.
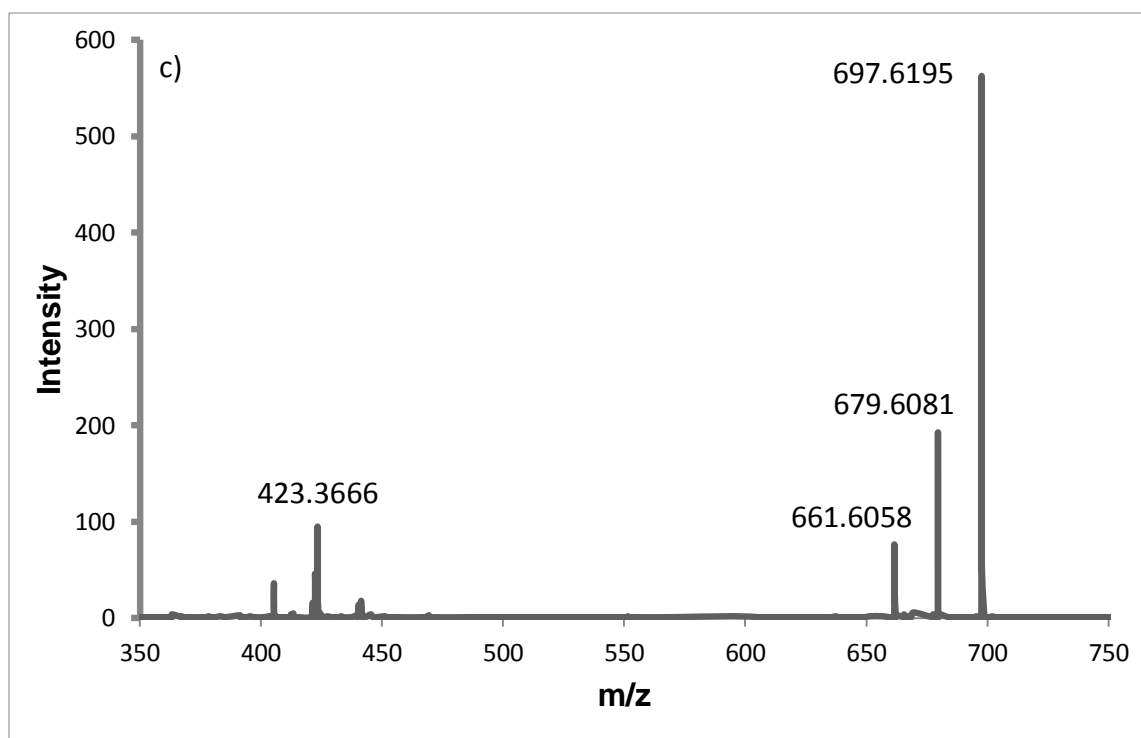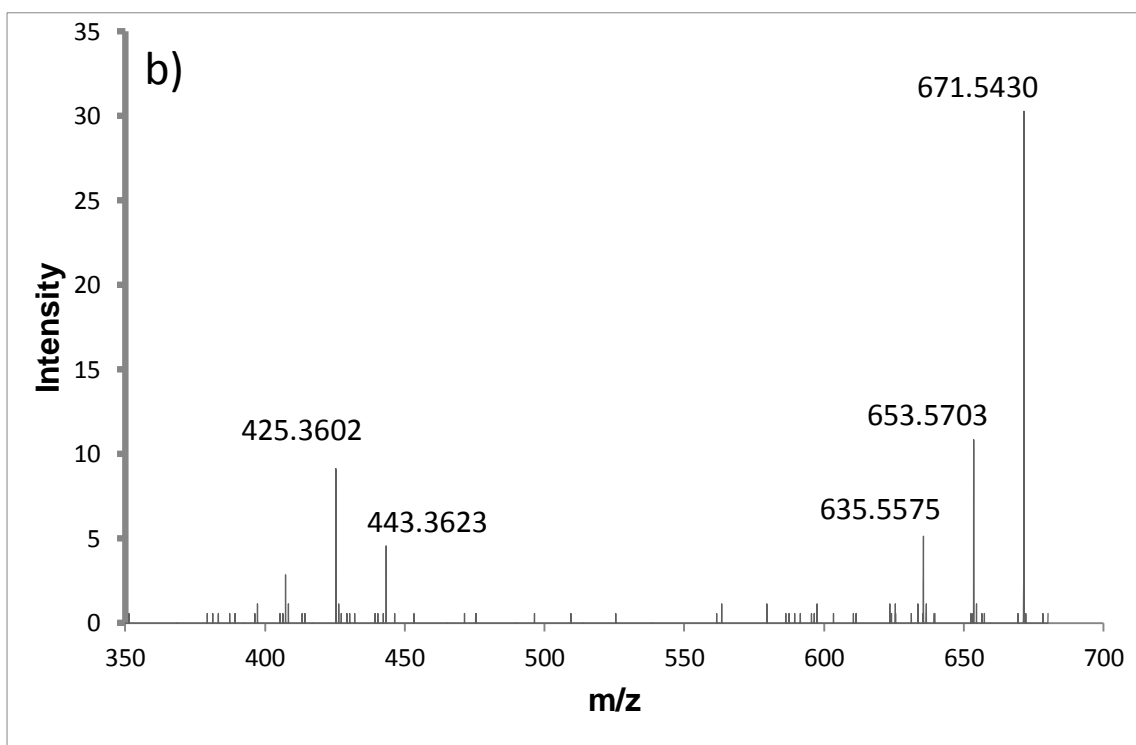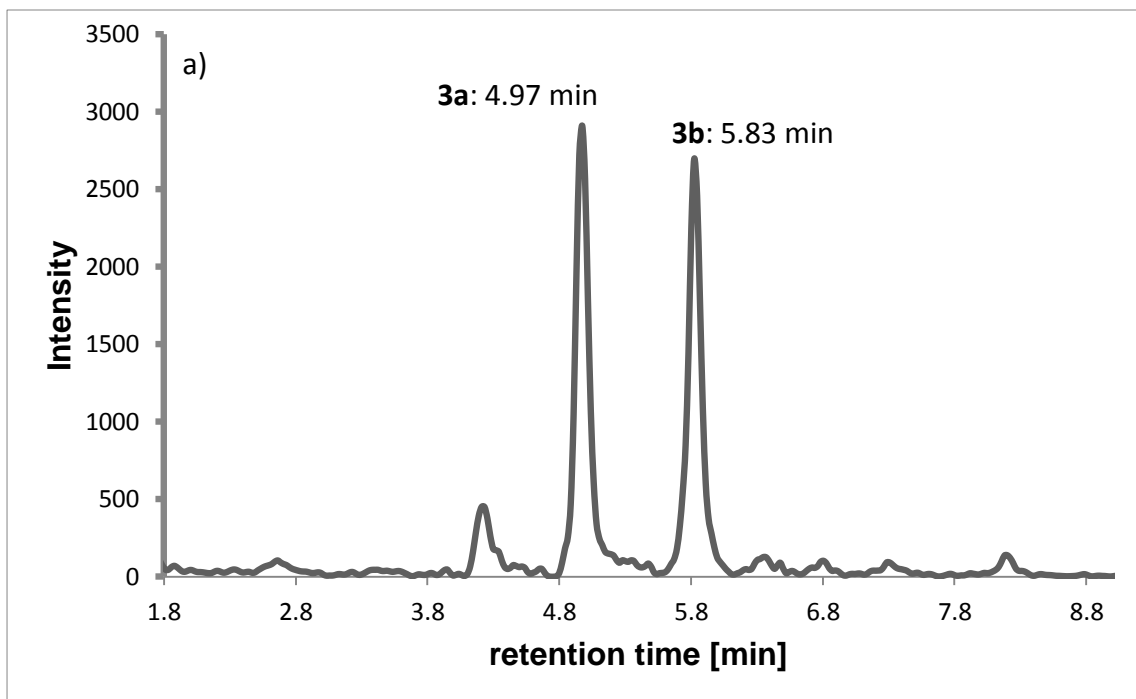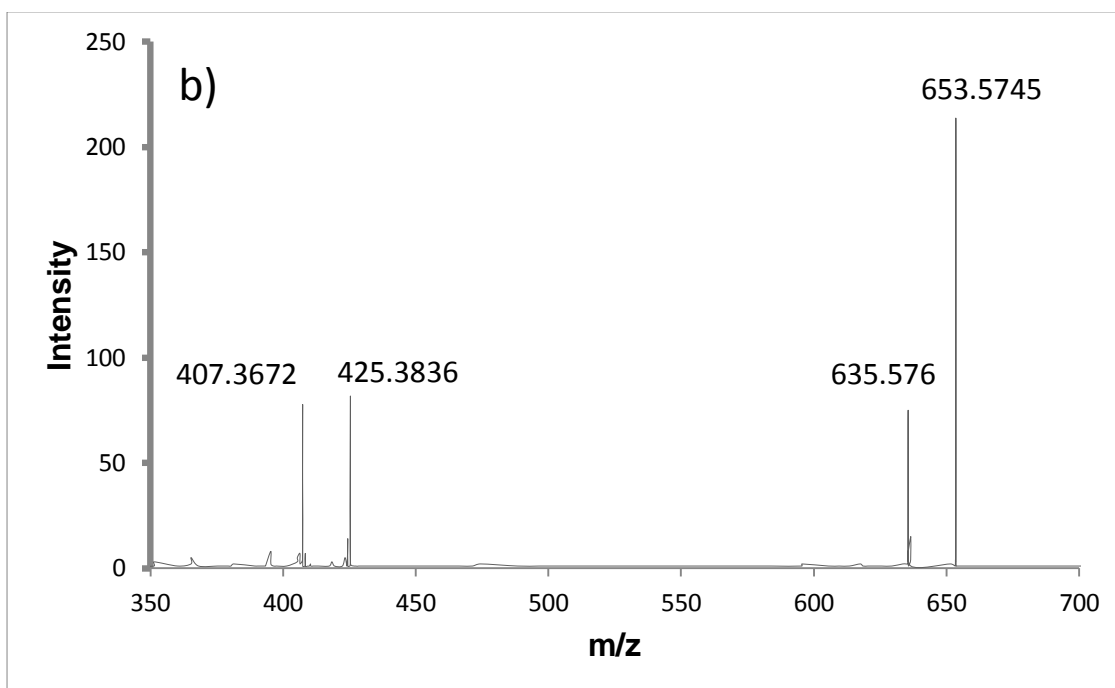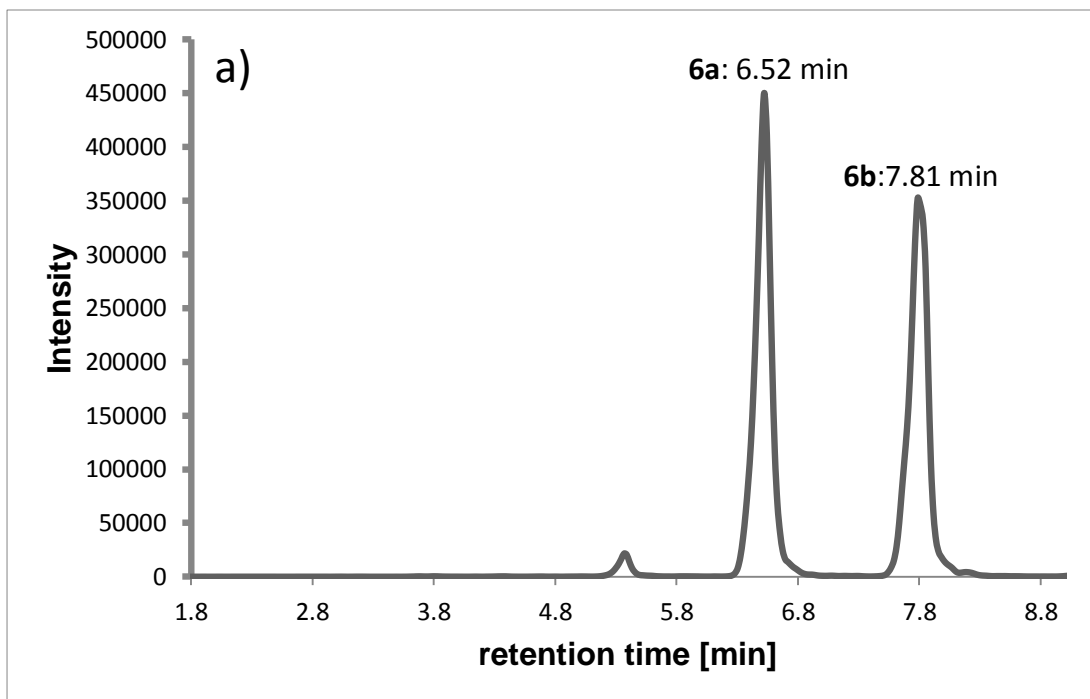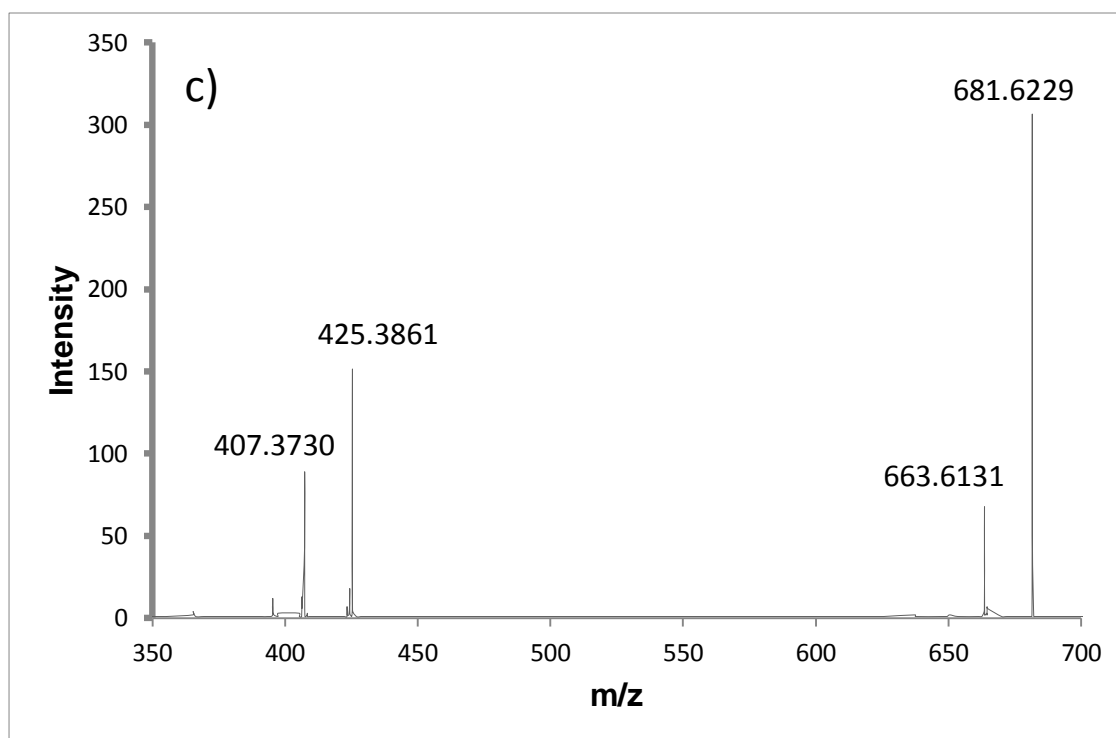
Fig. S6: Detection of maniladiol 3-monoesters (**6a + b**) by UHPLC-APCI-QTOF-MS on Kinetex C18. **a)** XIC on in-source decay triterpene skeleton fragment at 407.3672($\pm$10 mDa), retention time 6.52 min for **6a** and 7.81 min for **6b**. **b)** Product ion scan on $[M+H]^+$ of **6a** showing characteristic fragments of maniladiol 3-myristate eluted at t_R = 6.52. **c)** Product ion scan on $[M+H]^+$ of **6b** showing characteristic fragmentations of maniladiol 3-palmitate eluted at t_R = 7.81.

Table S10: Summary of MS data from CID-MS/MS-spectra obtained for maniladiol 3-myristate (**6a**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	653.5745	653.5867	12.19897	18.665	$[M+H]^+$
2	635.576	635.5761	0.19424	0.305613	#1-$H_2$0
3	425.3836	425.3778	5.804407	13.64511	#1-myristic acid
4	407.3672	407.3672	0.041887	0.102823	#1-H_2O-myristic acid

Table S11: Summary of MS data from CID-MS/MS-spectra obtained for maniladiol 3-palmitate (**6b**).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	681.6229	681.6180	4.9	7.1	$[M+H]^+$
2	663.6131	663.6074	5.6	8.4	#1-$H_2$0
3	425.3861	425.3778	8.3	19.6	#1-palmitic acid
4	407.373	407.3672	5.8	14.2	#1-H_2O-palmitic acid

APCI-QTOF MS characterization of faradiol diesters (7 – 10)

The following spectra are added in support of the MS data presented in Table 3.
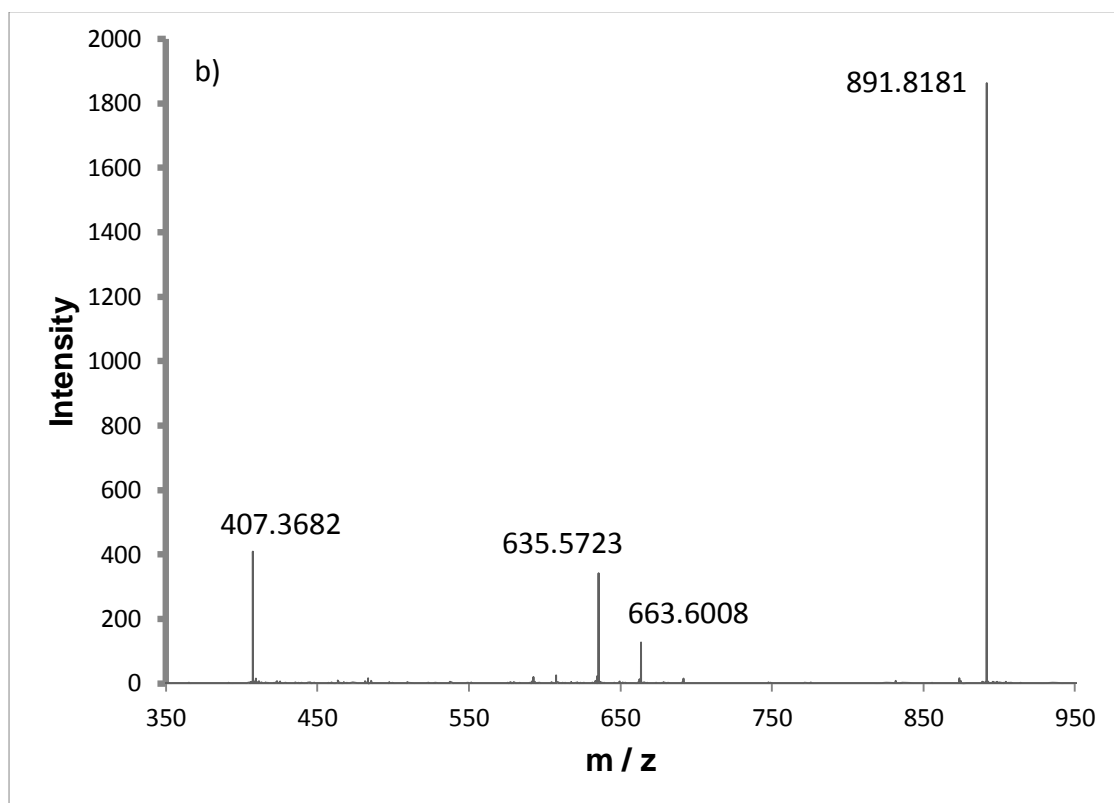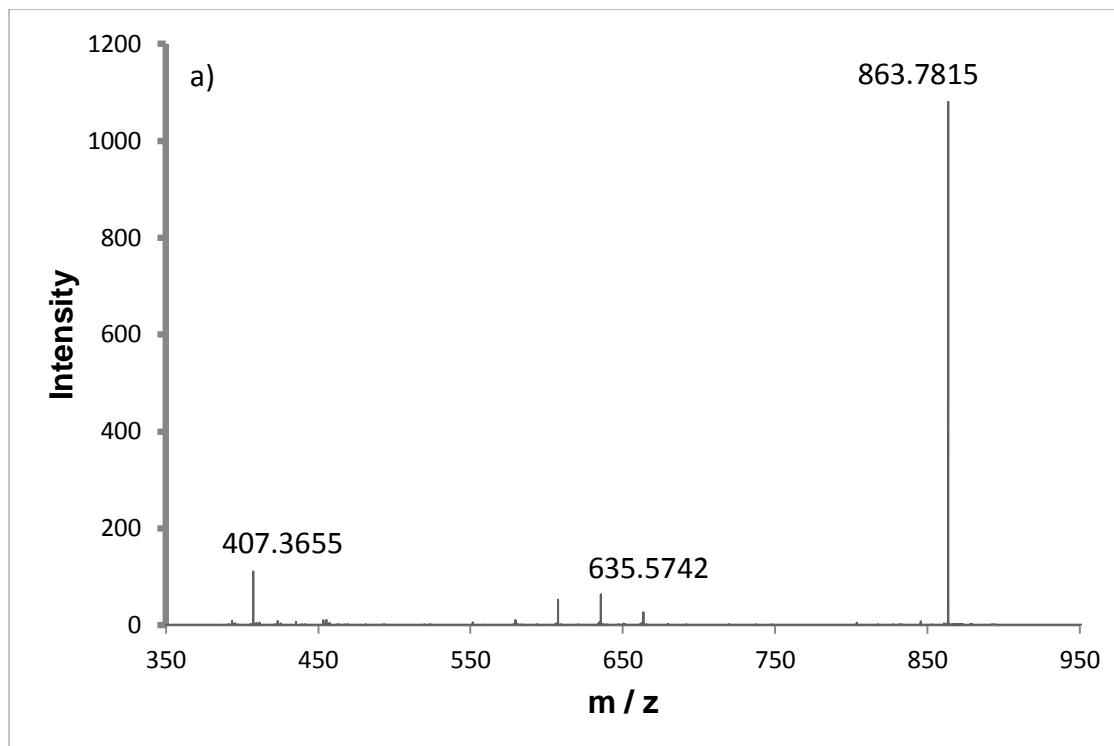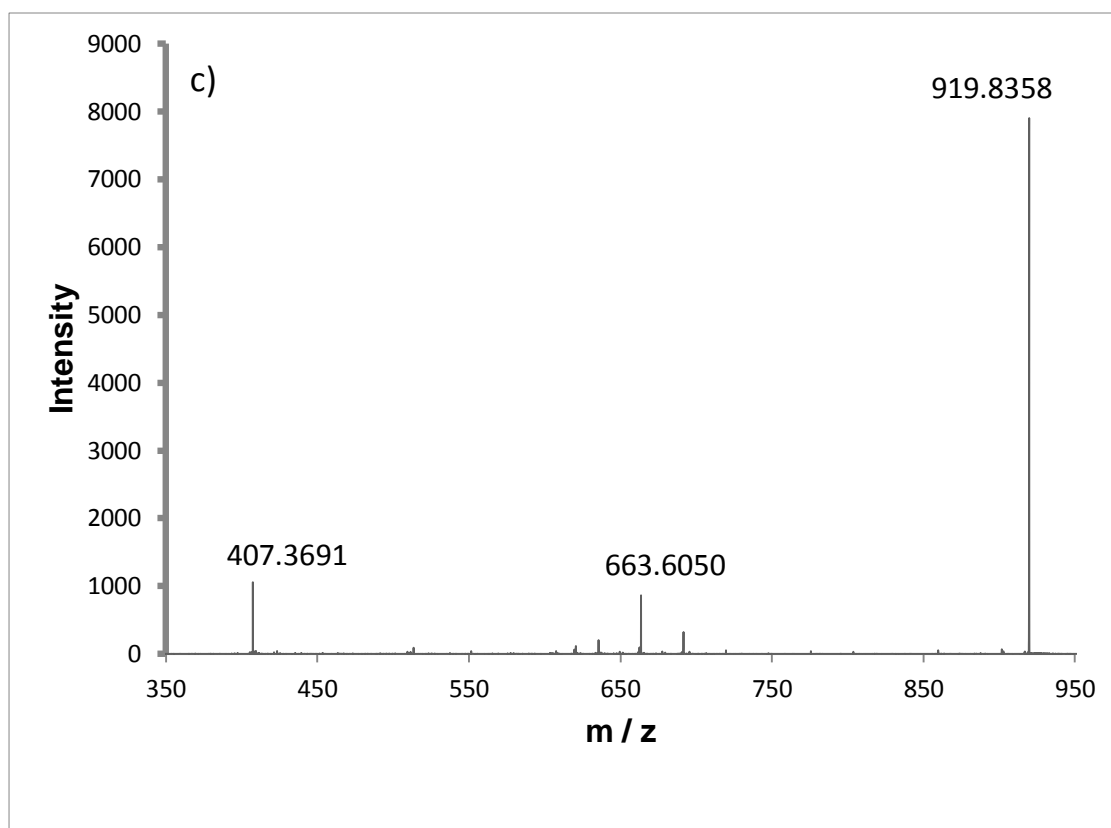
Fig. S7: Direct infusion MS/MS spectra (MCA mode) confirming the presence of faradiol diesters **7 – 10** in fraction of *C. officinalis* **a)** Faradiol 3,16-dimyristate (**7**). **b)** Faradiol 3-myristate,16-palmitate / faradiol 3-palmitate,16-myristate (**8 / 9**). **c)** Faradiol 3,16-dipalmitate (**10**)

Table S12: Summary of MS data obtained for faradiol 3,16-dimyristate (**7**) in direct infusion APCI-MS/MS experiments (MCA mode).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	863.7815	863.7851	-3.6	-4.17	[M+H]$^+$
2	635.5742	635.5761	1.9199	3.0207	#1-myristic acid
3	407.3655	407.3672	1.7211	4.2250	#1-2myristic acid

Table S13: Summary of MS data obtained for faradiol 3-myristate,16-palmitate / faradiol 3-palmitate,16-myristate (**8** / **9**) in direct infusion APCI-MS/MS experiments (MCA mode).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	891.8181	891.81639	1.71	1.92	[M+H]$^+$
2	663.6008	663.6074	6.619065	9.974467	#1-myristic acid
3	635.5723	635.5761	3.8641	6.079718	#1-palmitic acid
4	407.3682	407.3672	0.956636	2.348333	#1-myristic acid-palmitic acid

Table S14: Summary of MS data obtained for faradiol 3,16-dipalmitate (**10**) in direct infusion APCI-MS/MS experiments (MCA mode).

#	m/z found	m/z calculated	Δ mDa	Δ ppm	fragment
1	919.8358	919.84769	-11.89	-12.93	[M+H]$^+$
2	663.605	663.6074	2.491546	3.754562	#1-palmitic acid
3	407.3691	407.3672	1.83576	4.506379	#1-2palmitic acid

NMR spectra of the mixture of faradiol 3,16-diesters (7 – 10).
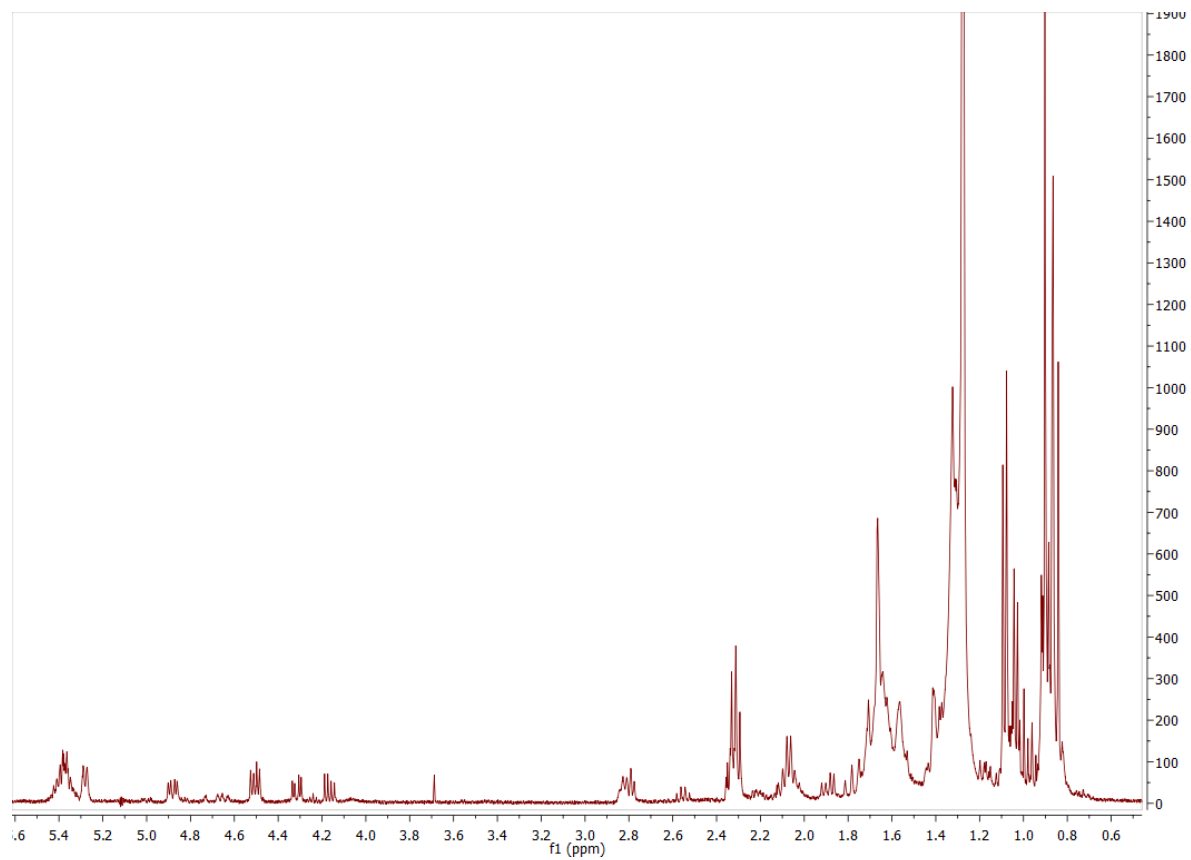
Fig. S8: ^{1}H NMR (400 MHz, CDCl$_3$) spectrum of faradiol 3,16-diesters (**7 – 10**)
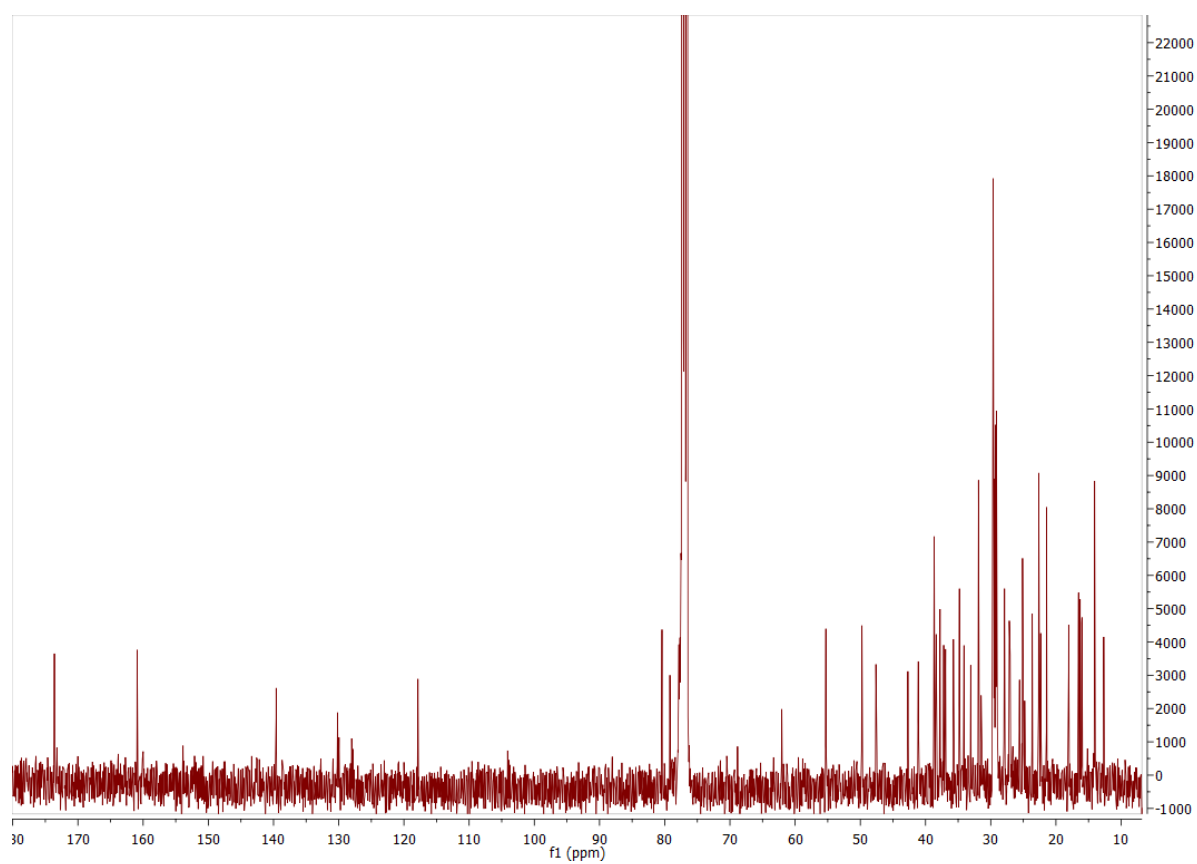
Fig. S9: ^{13}C NMR (100 MHz, CDCl$_3$) spectrum of faradiol 3,16-diesters (**7 – 10**)
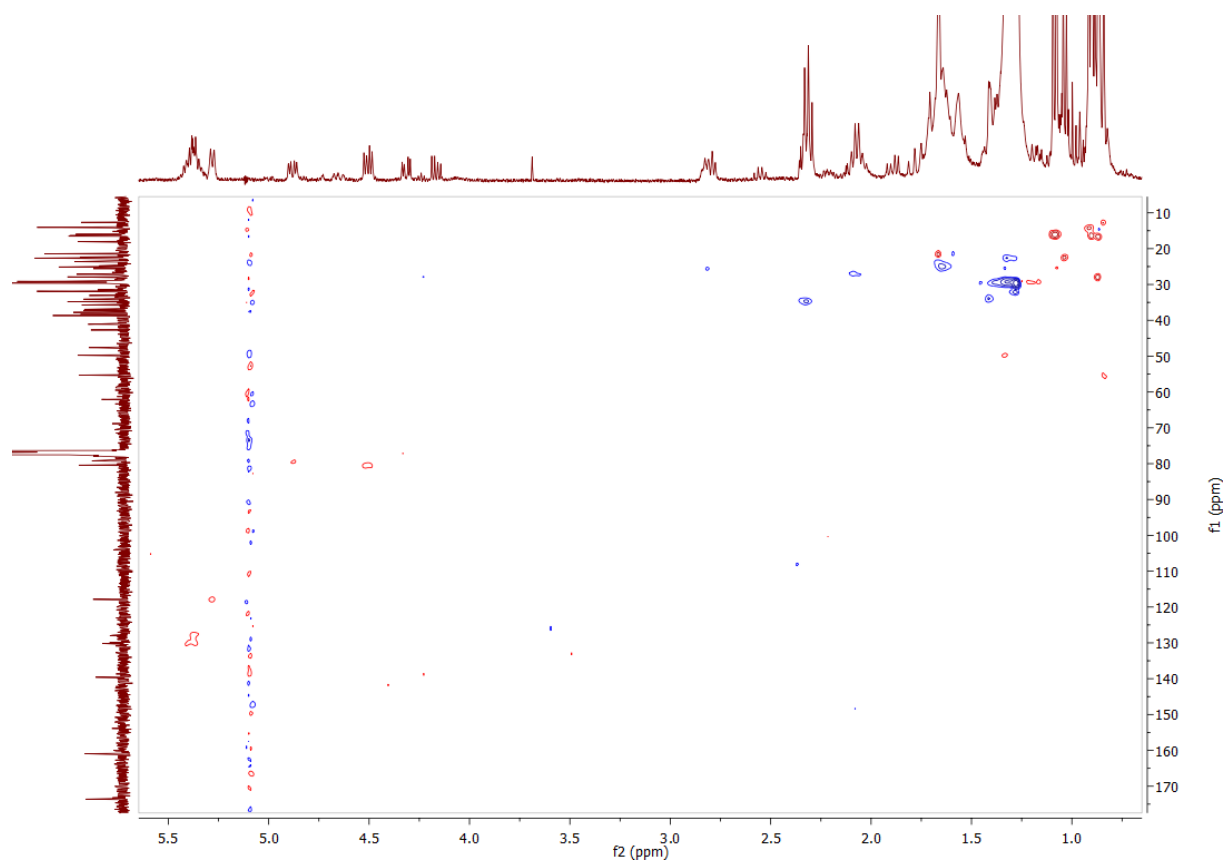
Fig. S10: HSQC spectrum of faradiol 3,16-diesters (**7 – 10**)
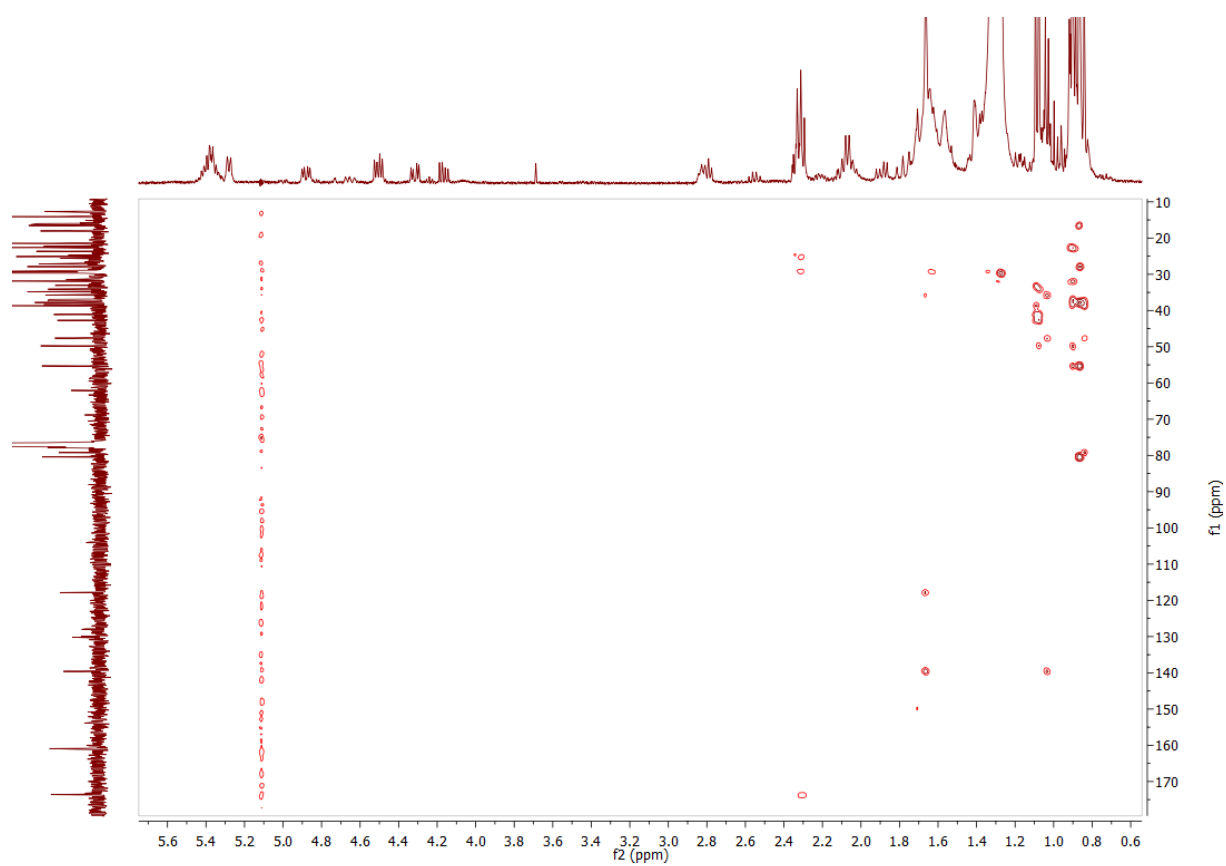
Fig. S11: HMBC spectrum of faradiol 3,16-diesters (**7 – 10**)
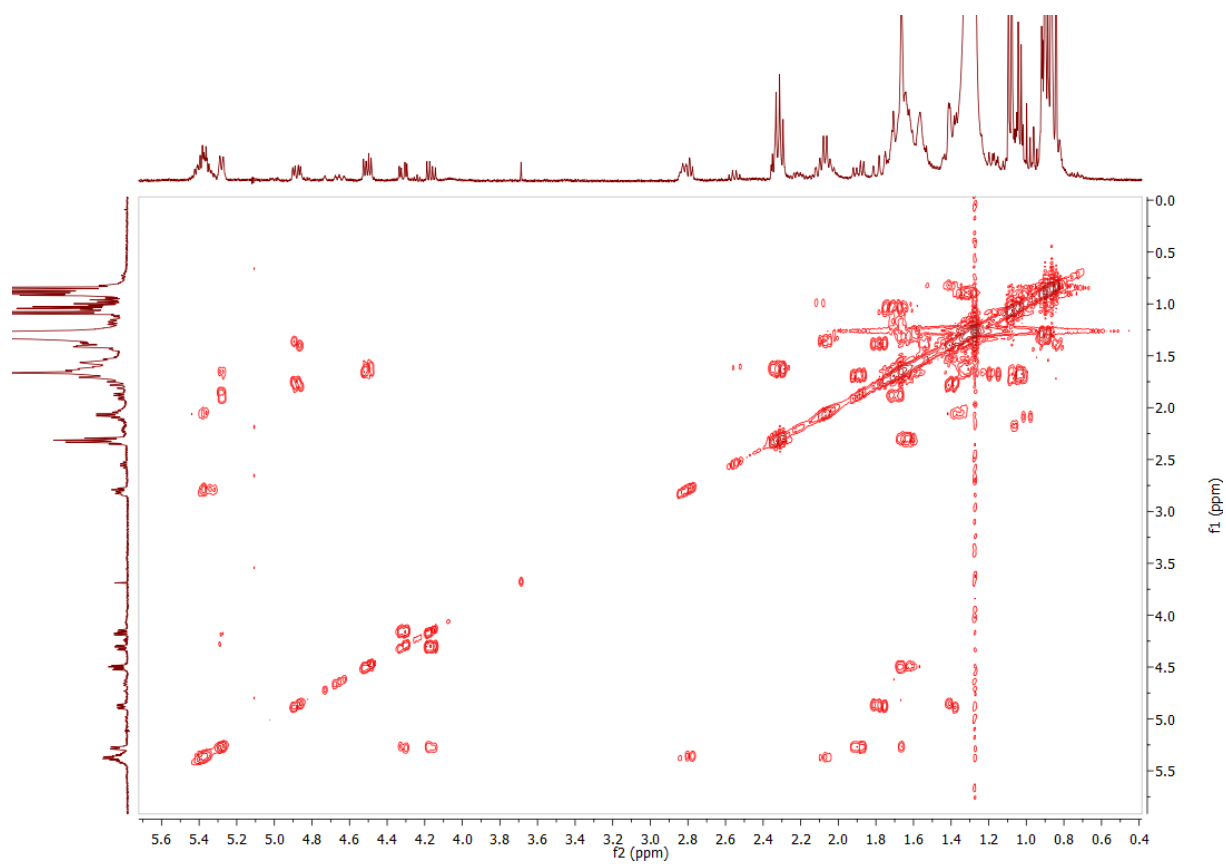
Fig. S12: ^{1}H/^{1}H-COSY spectrum of faradiol 3,16-diesters (**7 – 10**)

Appendix V

Accepted Manuscripts

Publication VII	
Title	Surface-crosslinked poly(3-mercaptopropyl)methylsiloxane-coatings on silica as new platform for low-bleed mass spectrometry-compatible functionalized stationary phases synthesized via thiol-ene click reaction
Authors	*Zimmermann, Aleksandra; Horak, Jeannie; **Sievers-Engler, Adrian**; Sanwald, Corinna; Lindner, Wolfgang; Kramer, Markus; Laemmerhofer, Michael*
Published in	Journal of Chromatography A (1436) 2016, p. 73-83
DOI	http://dx.doi.org/10.1016/j.chroma.2016.01.058

Reprint with permission from Elsevier

Surface-crosslinked poly(3-mercaptopropyl)methylsiloxane-coatings on silica as new platform for low-bleed mass spectrometry-compatible functionalized stationary phases synthesized via thiol-ene click reaction

Aleksandra Zimmermann[a], Jeannie Horak[a], Adrian Sievers-Engler[a], Corinna Sanwald[a], Wolfgang Lindner[b], Markus Kramer[c], Michael Lämmerhofer[a],*

[a] *Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany*
[b] *Lindner Consulting GmbH, 3400 Klosterneuburg, Austria*
[c] *Institute of Organic Chemistry, University of Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany*

ARTICLE INFO

ABSTRACT

A thin functional film of poly(3-mercaptopropyl)methylsiloxane was coated onto vinyl-modified silica particles (5 μm, 100 Å pore size) and chemically crosslinked to the surface. Excess of thiol functionalities allow bonding of alkene containing ligands by thiol-ene click reaction in a second step (QN-VII). Besides that a single step surface modification procedure was established in which alkene functional ligands were directly added to the polysiloxane coating solution and thus, after evaporation of the solvent, crosslinking to the vinylized surface and bonding of chromatographic ligand to the thiolated polysiloxane film occur simultaneously in one step (QN-VI). Successful bonding of the polysiloxane film was confirmed for both approaches by ^{29}Si cross-polarization/magic angle spinning NMR spectra. The new surface functionalization concept can be utilized as a new platform for the preparation of various low-bleed, mass spectrometry-compatible stationary phases with a variety of functional ligands. The concept was demonstrated by thiol-ene click reaction with quinine carbamate and its subsequent use for enantiomer separation by HPLC-UV and HPLC-ESI-QTOF-MS of acidic chiral analytes. Chromatographic enantioselectivities were similar to a comparable brush-type CSP (QN-V0). The greatly reduced background signal in LC-MS, however, comes at expense of somewhat lower chromatographic efficiencies (C-term by factor of 2 larger compared to brush-type CSP). For quantitative analysis in single reaction monitoring (MRM^{HR}) in high sensitivity mode, limit of detection and limit of quantification results are comparable for both surface-polymer modified CSPs, with only slightly higher values for the conventional brush-type CSP (QN-V0).

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Efficient surface functionalization is key in many disciplines of material sciences comprising sensor and microarray technologies, nanoscience, solid-phase synthesis, heterogeneous catalysis, and separation science to mention but a few. For liquid chromatographic separations of low molecular analytes, spherical porous silica particles have become the materials of first choice as supports and are commonly chemically modified in a brush-type manner for chromatographic separation modes such as reversed-phase [1],

polar bonded phases (amino phases, HILIC phases) [2–6], mixed-mode ion-exchangers [7–12], chiral stationary phases [13–17], and so forth. Along this line, surface silanols are chemically modified by silanization reaction with alkoxy or chlorosilanes via mono- or bifunctional siloxane bonds [18]. The resultant Si—O—Si bond is relatively stable under a broad range of conditions. However, it is susceptible to acidic and alkaline hydrolytic cleavage. Modern alkyl-bonded RP phases show only moderate bleeding and the eluted alkyl silanols do usually not or only minimally interfere with common detection concepts. Brush-type bonded HILIC silica phases are more critical, because the adsorbed water shell favours hydrolytic cleavage of Si—O—Si bonds. Most susceptible to hydrolytic cleavage amongst bonded silica phases are classical amino phases such as 3-aminopropyl silica [19]. Since those lig-

ands are commonly non-chromophoric, there is little interference with UV detection. However, nowadays mass spectrometric detection becomes more and more the method of choice for detection in LC and the column bleeding, especially of polar bonded phases, is no longer invisible to the detector. Alkyl silanols bleeding from RP phases exhibit poor ionization efficiencies in positive ion mode and thus are unproblematic. HILIC, amino and chiral ligands on contrary are polar with sometimes exceptional ionization efficiencies in electrospray ionization interfaces to mass spectrometry. In such cases column bleeding may interfere dramatically with detection and limit detection sensitivity.

To overcome this problem a number of strategies have been proposed. Polymeric siloxane bondings prepared by synthesis protocols which promote polymerization of silane precursor before condensation to the silica surface is one propagated concept but does not fully solve the problem, which may be due to imperfect condensation chemistries [20,21]. Hydrolytically more stable bonded chromatographic silica based stationary phases can be also obtained through hydrosilylation of olefins on a hydride-modified silica intermediate [22]. Hypercrosslinking of a monolayer of an aromatic network on top of silica was shown to create a very stable surface bonding which can be modified for various types of LC [23]. More stable stationary phases for RPLC were obtained by incorporation of organic bridges into the silica backbone and chemical modification of these molecules via ether bridges [24]. A frequently employed strategy is to use polymeric ligands which are coated onto the surface of silica and are crosslinked [25–27]. In a few works, also polysiloxane derivatives, e.g. with pendant alkyl groups, have been coated onto silica surface to prepare alkyl functionalized stationary phases, yet were not crosslinked to the surface [28,29]. While polysiloxane based chiral stationary phases are standard in GC enantiomer separation, such surface modification is rarely used in liquid chromatography with few examples demonstrated [30–36]. For instance, Schomburg and co-workers prepared chirally modified pre-polymers (viz. polymethylsiloxanes with SiH groups) onto which quinine was bonded by hydrosilylation reaction to the double bond of quinine and the resultant chiral pre-polymer then bonded to silica or vinyl silica i.e. the synthesis of the chiral polymer is done externally and not in-situ on the support [36]. We recently reported surface functionalization of an organic polymer monolith with 3-mercaptopropylmethylpolysiloxane and bonding of a chiral quinine carbamate selector for CEC enantiomer separation [37]. Crosslinking of polymeric selectors is also pursued as strategy to immobilize polysaccharide selectors [38–40] or by in-situ reaction of *N,N*'-diallyl-L-tartardiamide with a multifunctional hydrosilane, yielding a network polymer incorporating the the chiral selector that is simultaneously covalently bonded to vinyl silica [41,42].

The goal of this work was to develop a platform technology which allows the preparation of functionalized silica stationary phases for various HPLC separation modes via thiol-ene click reaction and is based on a bonding chemistry that does not suffer from ligand bleeding in UV, fluorescence, charged-aerosol detection and in particular mass spectrometric detection. The concept is based on the coating of a thin film of a polysiloxane with functional thiol groups onto vinylized silica particles and the subsequent crosslinking of the film to the surface by radical addition reaction (thiol-ene click reaction). Excess of accessible thiols can be used for immobilization of vinylized ligands, structurally tailored for the respective chromatographic mode. A one step coating/surface crosslinking/ligand immobilization approach with poly(3-mercaptopropyl) methylsiloxane, vinyl silica and vinyl-group bearing ligand is also proposed. The concept is illustrated by use of *tert*-butylcarbamoylquinine as vinyl-group containing ligand for enantioselective chromatography of chiral acidic analytes. The chromatographic performance of the new coated and crosslinked chiral stationary phases is compared to a corresponding brush-

type bonded CSP. The advantage of this new bonding chemistry for hyphenation with mass spectrometry is documented for total ion as well as selected ion monitoring mode.

2. Materials and methods

2.1. Materials

Poly(3-mercaptopropyl) methylsiloxane (PMPMS), 3-mercaptopropyl dimethoxymethylsilane and vinyltrichlorsilane were supplied by ABCR (Karlsruhe, Germany). The synthesis chemicals 4-(dimethylamino) pyridine (DMAP) and 2,2′-azobis(2-methylpropionitrile) (AIBN) as well as the chemicals used for the preparation of LC-mobile phases i.e. acetic acid, LCMS-grade ammonium acetate, LCMS-grade ammonium formate and HPLC grade methanol were purchased from Sigma Aldrich (Munich, Germany). Acetonitrile HPLC grade was purchased from J.T.Baker (Deventer, Netherlands). Ultra-LC–MS grade acetonitrile, methanol, formic acid and acetic acid were purchased from Carl Roth (Karlsruhe, Germany). The test compounds *N*-acetyl-D,L-phenylalanine (Ac-D,L-Phe), *N*-(carbobenzyloxy)-D,L-phenylalanine (Z-D,L-Phe), and dichloroprop were from Sigma–Aldrich. Spherical silica gel Kromasil 100 Å, 5 μm particle size, was supplied by Eka Chemicals (Bohus, Sweden).

2.2. Synthesis and characterization of stationary phases

2.2.1. Synthesis of vinyl silica gel

Vinyl silica gel was prepared by refluxing a slurry of 20 g Kromasil (100 Å, 5 μm) in 200 mL pyridine (dried over molecular sieves) with 52 mmol vinyltrichlorosilane in the presence of 58 mmol DMAP dissolved in 10 mL chloroform overnight at 80 °C using a mechanical stirrer. After washing repeatedly with hot toluene, chloroform, methanol and diethylether the modified silica gel was dried at 60 °C under vacuum and subjected to elemental analysis. The results are presented in Table 1.

2.2.2. Two-step polysiloxane-bonded tBuCQN CSP (QN-VII)

2.5 g of vinyl silica gel were suspended in 25 mL methanol and placed in a round bottom flask. 0.42 mL PMPMS (corresponding to 1.0 mmol thiol groups per g silica gel) and 0.6 mmol radical initiator AIBN were added. The slurry was slowly rotated and the solvent evaporated under vacuum at 35 °C. The dried silica gel was suspended again in 25 mL methanol and 2 mmol of tBuCQN ligand (synthesized as descried in [43–45]), together with 0.6 mmol AIBN, was added. The slurry was again slowly rotated and the solvent was evaporated under vacuum at 35 °C. The flask with dry and modified silica gel was filled with nitrogen, closed with a silicon rubber plug and placed in the drying cabinet at 60 °C overnight. The product was then washed several times with hot toluene and hot methanol, and was afterwards dried in a vacuum oven at 60 °C. The dried tBuCQN-modified silica gel (Fig. 1a) was subjected to elemental analysis to determine the selector coverage (based on nitrogen) (Table 1).

2.2.3. One-step polysiloxane-bonded tBuCQN CSP (QN-VI)

2.3 g of vinyl silica gel were suspended in 25 mL methanol and placed in a round bottom flask. Next, 1.9 mmol of tBuCQN ligand, together with 0.40 mL PMPMS (corresponding to 1.0 mmol thiol groups per g silica gel) and 0.6 mmol radical initiator AIBN, were added. The slurry was slowly rotated and the solvent evaporated under vacuum at 35 °C. The flask with dry tBuCQN-modified silica gel was filled with nitrogen, closed with a silicon rubber plug and placed in a drying cabinet at 60 °C overnight. The product was then washed several times with hot toluene and hot methanol, and afterwards dried in a vacuum oven at 60 °C. The dried modified silica gel

Table 1
Elemental analysis and surface coverage data for tBuCQN chiral stationary phases.

Stationary phases	%C	%H	%N	%S	S [mmol/g]	ligand coverage [mmol/g]	ligand coverage [μmol/m²]
Vinyl silica gel	5.32	1.09	0.03	0.01	–	–	–
QN-VII	15.27	2.82	1.1	2.19	0.68	0.25	0.8
QN-VI	14.05	2.35	1.11	1.64	0.51	0.25	0.8
QN-V0	15.14	2.38	1.59	1.84	0.73	0.38	1.3
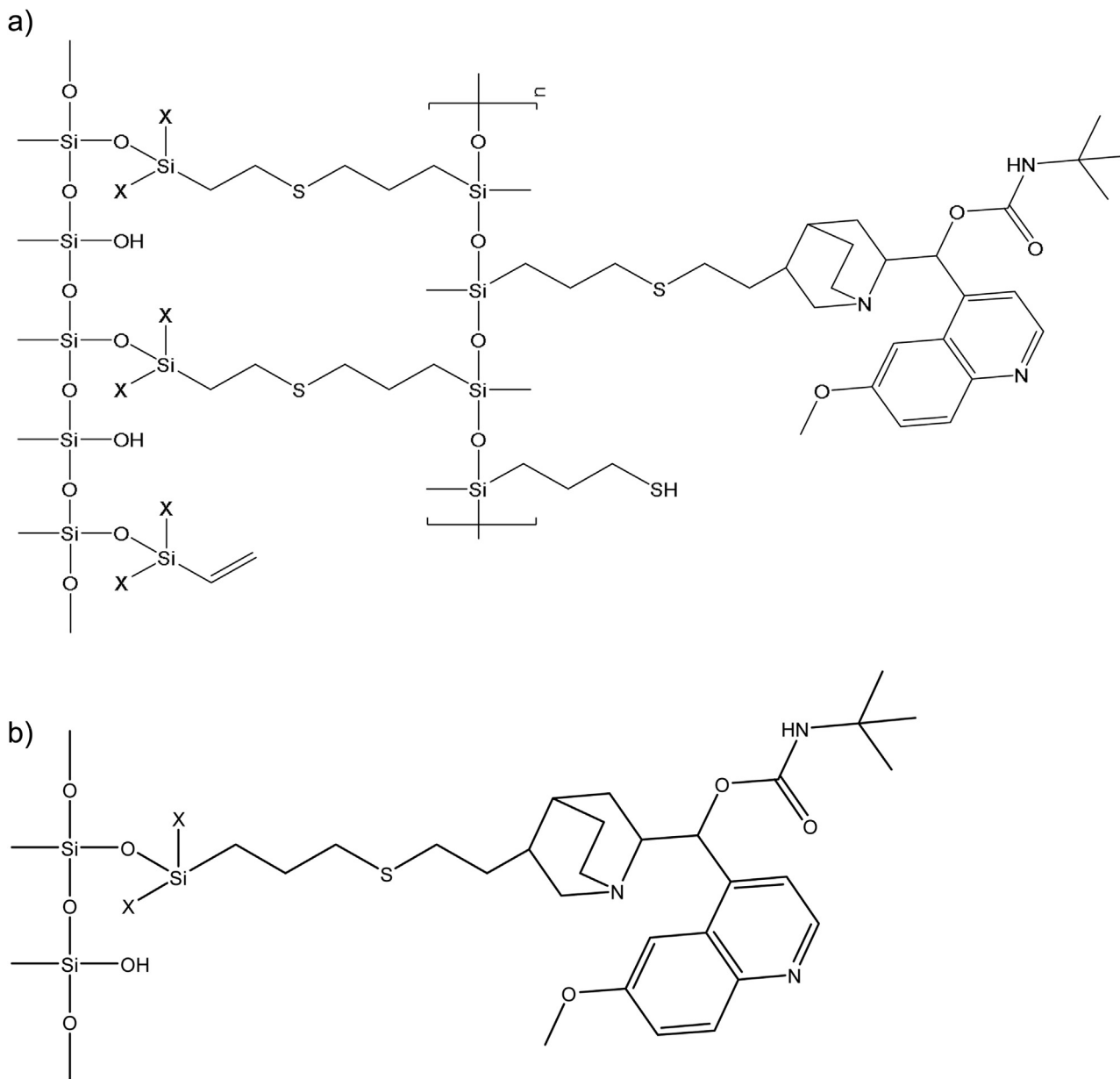
Fig. 1. Surface chemistries of crosslinked polysiloxane-coated (a) and brush-type (b) *tert*-butylcarbamoyl quinine (tBuCQN) modified CSPs.

was subjected to elemental analysis and the amount of tBuCQN ligand on the silica surface was determined by calculations based on nitrogen (Table 1).

2.2.4. Brush-type tBuCQN CSP (QN-V0)

The brush-type *tert*-butylcarbamoylquinine CSP (Fig. 1b), employed as benchmark for comparison, was synthesized using 3-mercaptopropyl dimethoxymethylsilane according to protocols described elsewhere [43–45].

2.2.5. 2 ²⁹Si Cross-polarization magic angle spinning (CP/MAS) NMR characterization

^{29}Si CP/MAS NMR spectra were acquired on a Bruker ASX 300 instrument (Bruker, Rheinstetten, Germany). Magic angle spinning was carried out with a 7 mm double bearing ZrO_2 rotor filled with about 250 mg modified silica sample at a spinning rate of 3 kHz. The 90° proton pulse length was 5.5 μs, and contact and relaxation delay times were 5 ms and 1 s, respectively. All chemical shifts were referenced externally to trimethylsilyl ester of octameric silica (Q_8M_8).
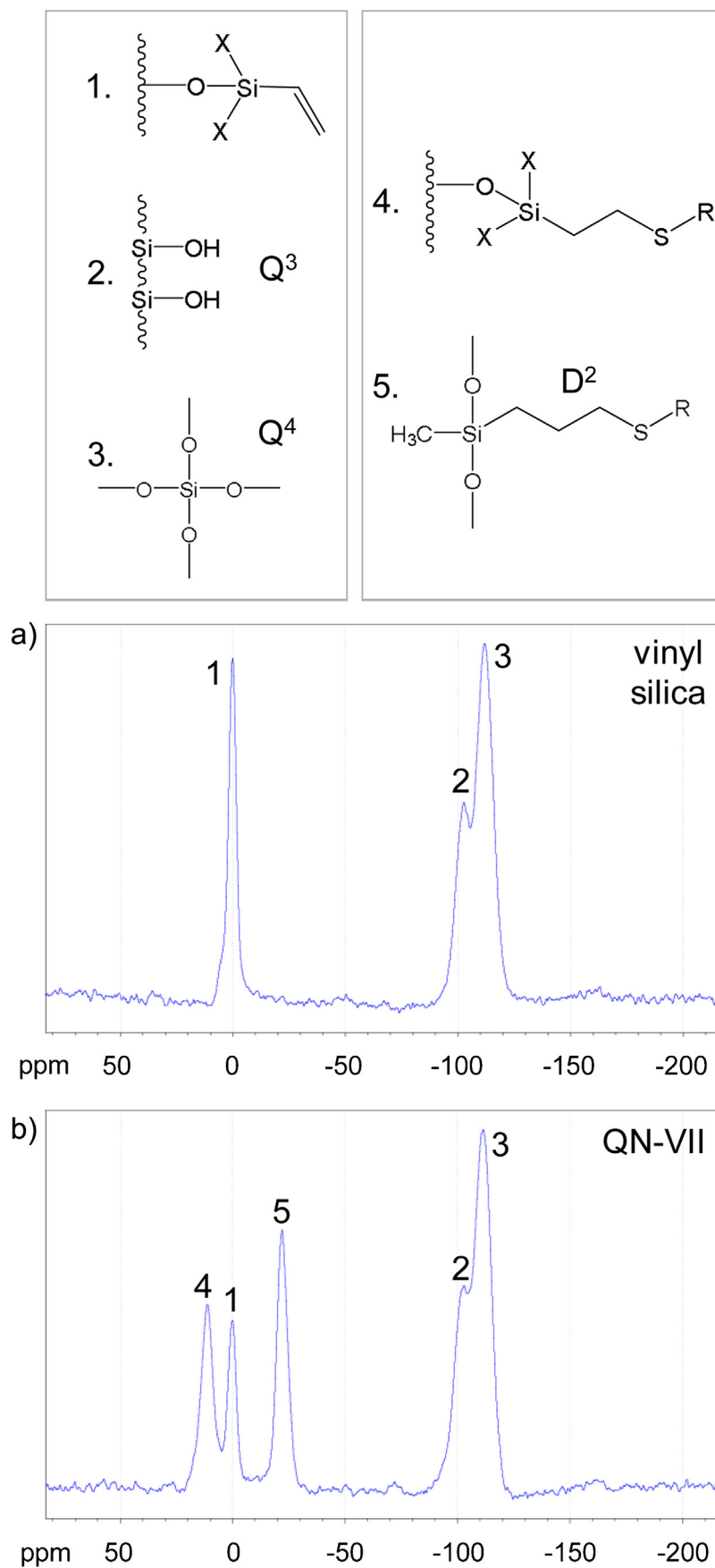
Fig. 2. ^{29}Si CP/MAS NMR spectra of (a) vinyl silica and (b) of two-step polysiloxane bonded tBuCQN CSP QN-VII.

A. Zimmermann et al. / J. Chromatogr. A 1436 (2016) 73–83

77

2.2.6. Column packing

All modified silica gels were packed into stainless steel columns of the dimension 150 mm × 4 mm ID at a pressure of 80 MPa using methanol as delivery solvent.

2.3. Liquid chromatographic experiments

2.3.1. Instrumentation

All samples were analyzed on an Agilent 1100 series LC system from Agilent Technologies (Waldbronn, Germany) equipped with an autosampler, degasser, quaternary pump, thermostated column compartment and diode array detector. The system was controlled by OpenLab CDS ChemStation—Edition for LC & LC/MS System (Rev. B.04.03.16). Data were analyzed using ChemStation software (Rev. B.04.03.16).

2.3.2. Stationary phase characterization

2.3.2.1. Chiral separation test conditions.
The mobile phase for the chiral separation test run consisted of methanol, acetic acid and ammonium acetate (98/2/0.5, v/v/w). The flow rate was 1 mL/min. The sample concentration was 1 mg/mL. All components were dissolved in the mobile phase. The injection volume was 10 μL. All samples were analyzed at a column temperature of 25 °C and at a wavelength of 254 nm. Acetone was used as a void volume marker.

2.3.2.2. Van Deemter analysis.
The column efficiency test was performed using a sample containing N-acetyl-D,L-phenylalanine (AcPhe) and acetone as void volume marker. The mobile phase consisted of methanol, acetic acid and ammonium acetate (98/2/0.5, v/v/w). The following flow rates were investigated: 3.0, 2.0, 1.0, 0.75, 0.5, 0.35, 0.25 mL/min. All samples were analyzed at a column temperature of 25 °C and at a wavelength of 254 nm. Theoretical plate numbers were calculated using peak width at half height, determined by ChemStation Software (Rev. B.04.03.16). Height equivalents of theoretical plates (HETP, H) were not corrected for extra-column contributions.

2.3.2.3. MS-compatibility tests.
Analyses were run on a Sciex TripleTOF 5600+ equipped with a Duospray ion source which was coupled to an Agilent 1290 Series UHPLC instrument (Agilent, Waldbronn, Germany) with a Pal HTC-XS autosampler from CTC (Zwingen, Switzerland). Experiments were carried out with ESI interface in TOF MS mode.

General MS settings were adjusted to the following optimized values: curtain gas (CG) 30 psi, nebulizer gas (GS1) 60 psi, drying gas (GS2) 50 psi, source voltage (ISVF) +5500 V, source temperature (T) 400 °C, scan window (SW) m/z = 100 to 2000, RF Transmission (RF) m/z = 80 50%, m/z = 230 50%, accumulation time (AT) 250 ms.

The columns were evaluated in non-aqueous as well as hydro-organic mobile phase mode and were examined for their propensity to cause ligand bleeding. For the non-aqueous conditions the mobile phase was composed of methanol and acetonitrile (1:1, v/v) with 25 mM ammonium formate and 6.6 mM of formic acid. The flow rate was 0.7 mL/min. For the hydro-organic condition, 100 mM ammonium acetate in methanol and water (4:1, v/v) was used with a flow rate of 1 mL/min. The sample concentration was 1 mg/mL. The test compound Ac-D,L-Phe was dissolved in the mobile phase. The injection volume was 10 μL. All samples were analyzed at a column temperature of 25 °C.

For the determination of the limit of detection (LOD) and limit of quantification (LOQ), on-column AcPhe concentrations 2.5, 5, 25, 50, 250, 500 and 2500 pg of R- and S-AcPhe in MilliQ water were analyzed in triplicates, while background noise was measured in quadruplets. UHPLC TripleTOF ESI-MS measurements were performed under non aqueous conditions in negative mode using high sensitivity settings for selected reaction monitoring

(MRM^{HR}) scan mode selecting the precursor ion m/z = 208.0974 using Unit (0.70 Da) Q1 resolution and monitoring the fragment (product) ion m/z = 164.0717. The optimum fragmentation conditions, collision energy (CE) of −16 V and DP of −151 V for negative mode as well as CE of +18 V and DP of +114 V for positive mode were determined by direct infusion. For data processing MultiQuant™ 3.0 was used, employing extracted ion chromatogram on m/z = 164.0717 ± 0.05 Da with a smoothing factor of 6. Peak integration was performed manually.

3. Results and discussion

3.1. Synthesis of coated chiral stationary phases

Brush-type bonded silica phases are commonly synthesized from mono-, di- or trialkoxy silanes and corresponding chlorosilanes, respectively. In any of these cases, there is a certain amount of monofunctionally bonded ligand on the surface which is relatively stable but may be released upon acidic or alkaline hydrolytic cleavage. Multi-functional bondings, on the contrary, are more stable because the ligand does not bleed from the surface upon hydrolysis of one siloxane bond. For this reason, a functional polymer was coated onto porous silica particles in this work and crosslinked to the vinylized surface. Vinyl-modified silica (100A, 5 μm) was used as substrate for the immobilization of the thiolated polysiloxane film. A coating-crosslinking surface functionalization strategy was selected in which the chiral selector was immobilized by either a two step or one step synthesis approach.

In the two step approach (QNVII), a thiol-modified silica gel is firstly prepared onto which the chiral vinyl selector is then immobilized in a second step. Thus, a quantity of poly(3-mercaptopropyl) methylsiloxane corresponding to about 1 mmol thiols per g silica was suspended in methanol and used to prepare a slurry with vinyl silica. A small amount of AIBN was also added before the suspension was evaporated to dryness to coat a thin film of the functional polysiloxane onto the vinyl silica surface. At this stage, the thiol-modified polysiloxane film can be crosslinked to the surface by briefly subjecting the coated material to heat treatment. As such, the material could be stored as thiol-modified silica to be used later for immobilization of vinyl ligands by thiol-ene click reaction. In the current study, no such heat treatment was performed at this stage. In a second step, the chiral selector was dissolved in methanol along with radical initiator AIBN and the mixture added to the slurry of thiolated silica. The solvent was evaporated again and the reaction mixture heated to 60 °C for initiating the radical addition reaction. Thereby, a part of the thiols from the polysiloxane film react with the vinyl groups of vinyl-modified silica while others react with the chiral selector.

In the one step approach (QNVI), poly(3-mercaptopropyl) methylsiloxane, corresponding to about 1 mmol thiols per g silica, vinyl silica, AIBN and chiral selector were all slurried together and the solvent was slowly evaporated. Subsequently, the thiol-ene click reaction was initiated by heat treatment and allowed to proceed at 60 °C overnight. Crosslinking of the polysiloxane film to the surface and bonding of selector to this thiol-modified polysiloxane film occurred simultaneously in a single step.

The data for elemental analysis shown in Table 1, clearly confirm the successful immobilization of both, the poly(3-mercaptopropyl) methylsiloxane and the quinine carbamate ligand on both of the two distinct coated-crosslinked tBuCQN CSPs, QN-VI and QN-VII. The thiol content was adjusted to be in the same order as in the brush type CSP by charging the reaction mixture for the coating process with an appropriate quantity of poly(3-mercaptopropyl) methylsiloxane. The quantity of chiral selector to be added into the reaction mixture, in order to obtain a similar surface concentration

of ligand as observed on the corresponding brush-type tBuCQN CSP (QN-V0), was pre-optimized with allylamine. It becomes evident from Table 1 that the sulfur loading is a bit lower for the one-step QN-VI as compared to the two-step polysiloxane-bonded tBuCQN CSP (QN-VII) (0.51 vs 0.68 mmol/g) while for the latter it is virtually identical to the brush-type CSP (0.73 mmol/g). On the other hand, the tBuCQN-ligand concentration is a good match between the two versions of coated CSP (0.25 mmol/g), yet is a little lower than on the QN-V0 (0.38 mmol/g). Further adjustments of surface concentrations of both thiol and chiral selector are readily possible, if necessary from a chromatographic viewpoint.

Structural information on the surface chemistry can be derived from solid-state NMR spectroscopy. Spectra from ^{29}Si CP/MAS NMR [46] provide a detailed picture about the type of silanol species present on the modified silica surface at distinct stages of surface modification [47–49]. Fig. 2 depicts ^{29}Si CP/MAS NMR spectra of vinyl silica and the QN-VII. Characteristic resonances for native silica can be seen at −110 ppm (Q^4, siloxane groups) and −102 ppm (Q^3, free silanol groups) in both of the modified silica materials. Vinyl silica (Fig. 2a) shows another resonance signal close to 0 ppm indicating successful bonding of vinyl silane via monofunctional Si-O-Si bonds. The spectrum of the polysiloxane bonded CSP (Fig. 2b) displays two new resonance signals at about −20 ppm and + 10 ppm, respectively, resulting from the coating and crosslinking of the polysiloxane film. The same resonance signals were also observed in the QN-VI (not shown). The signal at approximately − 20 ppm can be assigned as the methylsilane species in the polysiloxane chain. The signal at ca. +10 ppm appears upon reaction of thiols with the alkene group of vinyl silica altering the electronic environment of the surface-bound silane species due to transformation of the unsaturated alkene to a saturated alkyl group. Treatment of vinyl silica with AIBN (without polysiloxane and tBuCQN ligand) under otherwise identical conditions does not lead to formation of significant quantities of this silane species indicating successful crosslinking of the polysiloxane film to the silica surface.

These results clearly confirm successful bonding of a polysiloxane film and immobilization of chiral selector, and demonstrate the straightforward approach for the synthesis of coated-crosslinked CSPs via thiol-ene click reaction with controlled surface chemistry.

3.2. Comparison of chromatographic performance with brush type CSP

The cardinal question of this study was to clarify whether the new coated-crosslinked CSPs (Fig. 1a) can achieve the same chromatographic performance as the corresponding brush-type CSP (Fig. 1b) obtained by immobilization via monomeric 3-mercaptopropyl trialkoxysilane and representing state-of-art in the field. Thus, three chiral acids (Ac-Phe, Z-Phe and dichloroprop) were selected as test substances and analyzed comparatively on the two new coated-crosslinked CSPs, QN-VI and QN-VII as well as the brush-type tBuCQN-CSP QN-V0 using polar organic mode elution conditions composed of MeOH with acetic acid and ammonium acetate (98:2:0.5; v/v/w). The results are summarized in Table 2 and representative chromatograms are depicted in Fig. 3. In spite of lower selector coverage on the two coated-crosslinked CSPs, both retention factors as well as enantioselectivities are in the same range. α-Values of the two coated-crosslinked CSPs ranged between 96% (Z-Phe) and 102% (dichloroprop on single-step coated version) related to those of the corresponding brush-type CSP. For Ac-Phe, separation factors on polysiloxane-coated versions QNVI and QNVII were practically identical to those achieved on brush-type tBuCQN-CSP.

It may be striking that retention and separation factors are not much different although the brush-type CSP has higher selector loading. The reason may be a complex interplay between slight
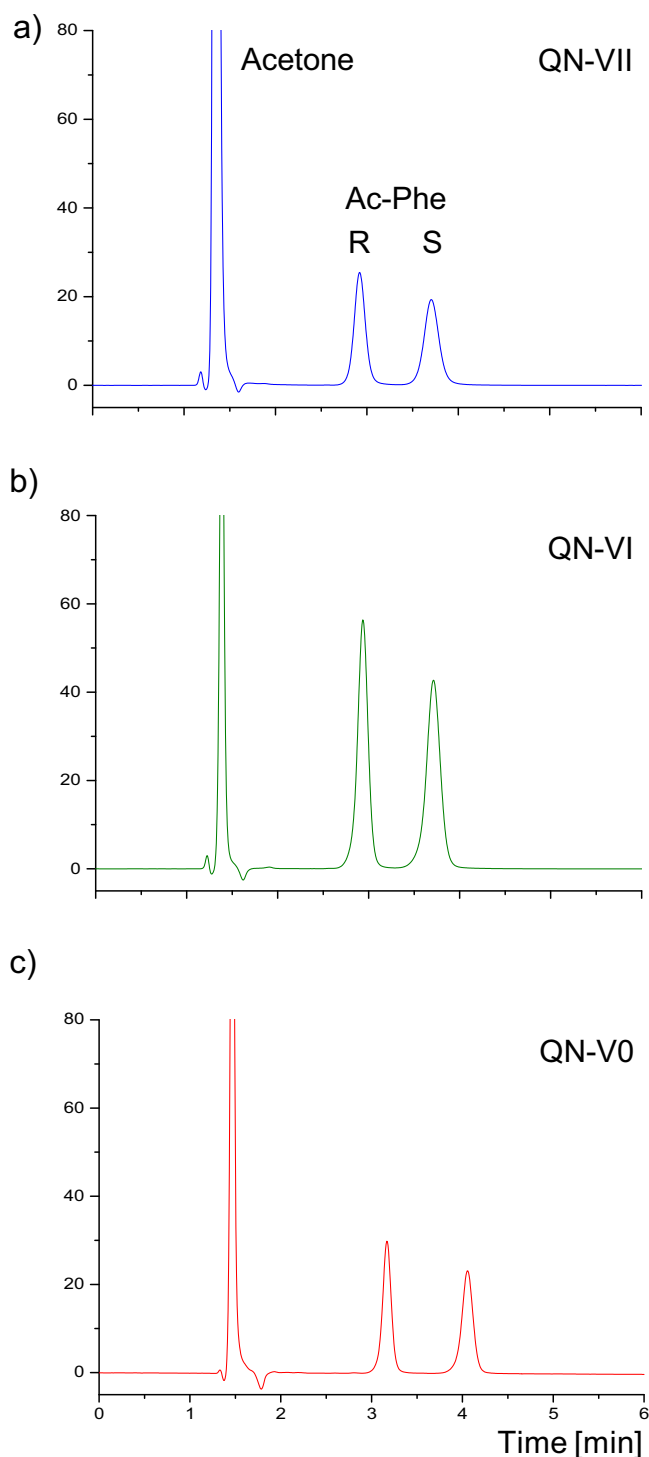
Fig. 3. Representative chromatograms of Ac-Phe obtained on tBuCQN-modified CSPs based on (a) crosslinked polysiloxane-coated silica via two-step approach (QN-VII) (b) crosslinked polysiloxane-coated silica via one-step approach (QN-VI), and (c) brush-type bonding (QN-V0). Conditions as specified in Experimental and Table 2, respectively.

differences in phase ratio, non-specific adsorption to polysiloxane backbone and support as well as specific binding to the selector. It could be argued that non-specific binding increments significantly add to retention, in particular under aqueous conditions, and thus reduced enantioselectivity could be expected in such a case. Thus, in order to complete the picture, chromatographic tests were also performed with hydroorganic mobile phases and it can be seen that
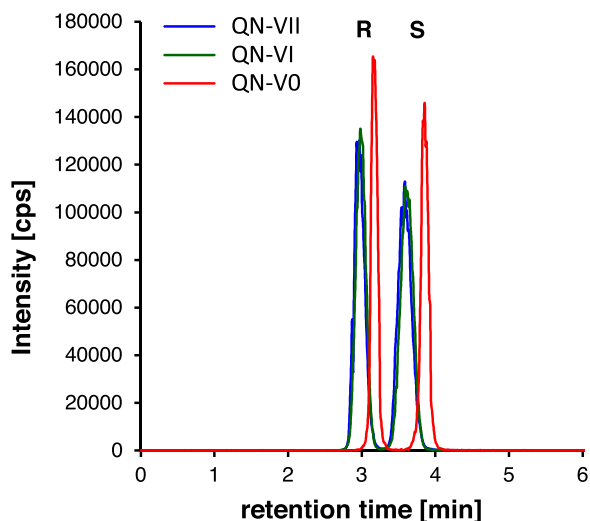
Fig. 4. Chromatogramms of Ac-R,S-Phe analyzed on QN-V0, QN-VI and QNVII analyzed under hydro-organic mobile phase conditions using ESI-TripleTOF-MS in positive mode. Mobile phase: 100 mM NH4Ac in MeOH:water (4:1 (v/v)), 25 °C, 1 mL/min.

Fig. 5. H/u curves for a) acetone (non-retained marker), b) Ac-R-Phe, c) Ac-S-Phe. For conditions see experimental section.

retention characteristics in this elution mode is also very similar, albeit retention is slightly higher on the brush-type CSP which has higher selector coverage (Fig. 4).

The trade-off of polymeric immobilization of the quinine carbamate selector via a coated and crosslinked polysiloxane film turns out to be a lower chromatographic efficiency. As can be seen from Table 2, plate numbers reach about 46% and 53% for QN-VII and QN-VI, respectively, as compared to the corresponding QN-V0. This translates also into somewhat lower resolution values (Table 2). In order to gain deeper insight into possible reasons for this efficiency loss, H/u curves were acquired on all three CSPs under comparable conditions (Fig. 5). The general trend is the same for all three CSPs. As common for enantioselective HPLC, the minimal theoretical plate height (H_{min}) is always observed at very low flow rate. On the other hand, the mass transfer resistance is significant. Multiple simultaneous interactions at the chiral selector cause typically a slower mass transfer than is common for RP-type chromatography on particles with identical morphology. In the comparison of the H/u-curves of the three CSPs it is most striking that the two coated-crosslinked CSPs reveal steeper H/u-curves i.e. the mass transfer resistance is more pronounced compared to the brush type CSP. Determination of C-terms indicates that they are by a factor of about 2 larger on both coated versions (e.g. 45 ms for (R)-Ac-Phe vs 19 ms on the brush type CSP). It might be related to diffusion limitations of analytes in the coated film, inhomogeneities in the coating, or too narrow pores due to a thick porous film rendering intra-particulate pore diffusion of analytes significantly more obstructed than in brush type CSPs without such coated film. Overall, it becomes evident that there is some room for optimization which will be objective of further studies.
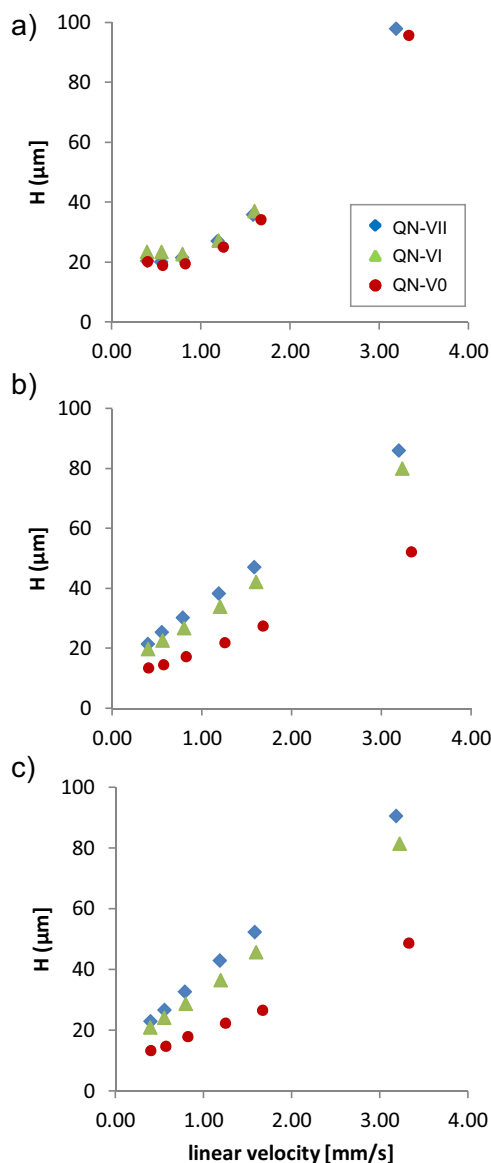
3.3. Minimized ligand bleeding and reduced MS background signals of new coated CSPs

Brush-type chiral stationary phases show stronger background signal in MS detection compared to what we are used from modern RP columns. This is due to a slight ligand bleeding which does typically not negatively affect enantioselectivity over a long time range, but originates in a higher noise and thus reduced signal-to-noise

Table 2
Comparison of chromatographic data of crosslinked polysiloxane bonded, QN-VI and QN-VII versus brush-type quinine carbamate CSPs, QNV0.

Analyte	QN-VII						QN-VI						QN-V0					
	k_1	k_2	α	R_S	N_1	N_2	k_1	k_2	α	R_S	N_1	N_2	k_1	k_2	α	R_S	N_1	N_2
Ac-Phe	1.17	1.75	1.50	5.04	2603	2442	1.12	1.68	1.50	5.36	3081	2831	1.16	1.77	1.52	7.60	5331	5291
Z-Phe	2.66	3.20	1.20	2.77	2279	2264	2.58	3.10	1.20	2.92	2657	2539	2.18	2.72	1.25	4.69	5063	4904
Dichloroprop	3.10	3.70	1.19	2.90	2507	2524	3.04	3.68	1.21	3.39	2966	2891	2.43	2.89	1.19	3.95	5344	5446

Experimental conditions: mobile phase, MeOH-AcOH-NH4Ac (98:2:0.5; v/v/w); flow rate: 1 mL/min; column temperature: 25 °C; detection: UV 254 nm; sample: 1 mg/mL; injection volume: 10 µL; column dimension: 150 × 4.0 mm ID.

ratios, consequently impaired detection sensitivities. This holds for non-commercial brush-type CSPs but also commonly utilized commercial brush type CSPs with low molecular mass selectors and macrocyclic intermediate sized selectors bonded to silica via monomeric silane linkers of the similar type as shown in Fig. 1b. In part, the stronger background signal emerges as a result of much better ionization efficiencies of polar selector molecules, in particular in positive ion mode, as compared to alkyl silanol ligands released as column bleed from RP phases.

Most commonly, trifunctional silanols are employed for immobilization of chiral selectors on silica surface. Silanization reactions with such derivatives give a distribution of mono-, di- and trifunctional bondings [48]. In particular, the mono-functionally bonded ligands are prone to ligand release by hydrolytic cleavage and significantly contribute to the background signal in MS. For better comparability, a methyl-silane linker was used herein for synthesis of the brush-type CSP, because this way the structural element by which the selector is bonded is exactly the same in both brush-type and polysiloxane-bonded tBuCQN CSPs. While the carbamate group is fully stable under mobile phase conditions, the most labile bond is the Si-O-Si linkage of the selector to the silica support.

Fig. 6a illustrates the total ion current (TIC) chromatograms of Ac-Phe measured in positive ion mode with a polar organic mobile phase as specified in the figure caption. As can be seen, the background signal abundance is significantly reduced for the polysiloxane-coated CSP versions, QNVI and QNVII as compared to the brush-type CSP QNV0. The TIC of the brush-type bonded CSP, QNV0 provides a background signal abundance of ca. 5×10^5. On the contrary, the TIC for the corresponding two-step synthesized polysiloxane-coated CSP QNVII reaches a signal abundance of only about 2.5×10^5, which is reduced by a factor of ca. 2. The noise level was further minimized for the single-step synthesized polysiloxane-coated CSP, QN-VI (signal intensity ca. 1.5×10^5) which was close to that of the mobile phase without any column (ca. 1.0×10^5). Since the background signal, like the analyte signal, is decisive for the detection sensitivity, the single-step synthesized polysiloxane-coated CSP provides lower detection limits due to improved signal-to-noise ratio. As a matter of fact, analyte peaks are significantly above the LOD for a 5 μg injection of Ac-Phe (2.5 μg of each enantiomer) in the TIC trace for the single-step synthesized polysiloxane-coated CSP, while no peaks were visible in the TIC for the same injected quantity in case of the brush-type CSP. The situation is virtually identical under hydroorganic elution conditions (Fig. 6b). This clearly confirms the advantage of the new polysiloxane-coated stable bond CSPs for LC–MS coupling (Fig. 6).

3.4. Determination of limit of detection and limit of quantification

In order to determine the influence of ligand bleeding on quantitative MS results, the limit of detection (LOD) and limit of quantification (LOQ) of AcPhe was determined for all three QN-based chiral columns for nonaqueous mobile phase condition in negative mode for the product ion m/z 164.0717 (± 0.025 Da). Since the stated conditions provided highest signal intensities, no other conditions are shown. Calibration curves were established by manual integration of 5 pg to 2.5 ng (on-column) of the R and S enantiomers of AcPhe within identical time segments using MultiQuant. Baseline correction was performed by subtraction of integrated background noise within said time segments of four subsequent blank injections. LOD and LOQ determination was performed according to ICH guidelines (LOD = $3.3 \times$ STD/slope and LOQ = $10 \times$ STD/slope) employing the residual standard deviation (STD) of the calibration curve. Table 3 shows the results along with calibration functions, linearities and range. It can be seen that a good linearity was achieved (even without the use of internal standards) in the concentration range between 5–2500 pg

Table 3
Limit of detection (LOD) and limit of quantification (LOQ) for AcPhe under non aqueous conditions, measured by ESI-QTOF-MS/MS in negative mode.

Column	Sample		Chromatographic data[a]				Calibration function information				LOD[b] [pg]	LOQ[b] [pg]
Type	Chirality [R/S]	t_0 [min]	RT [min]	k	α	Calibration range [pg]	Integration range [min]	Calibration function ($\pm$SE)		Determination coefficient R^2		
QN-VI	S	1.62	7.80	3.81	1.30	5–2500	7.00 to 9.50	$y = 3.319\ (\pm 0.010) \times - 11.125\ (\pm 10.331)$		0.99996	21.3	64.6
QN-VI	R	1.62	6.37	2.93	1.30	5–2500	5.40 to 7.00	$y = 3.676\ (\pm 0.020) \times - 23.708\ (\pm 20.483)$		0.99989	38.2	115.7
QN-VII	S	1.55	7.97	4.15	1.30	5–2500	7.25 to 9.40	$y = 2.087\ (\pm 0.014) \times - 46.540\ (\pm 15.252)$		0.99987	46.6	141.2
QN-VII	R	1.55	6.50	3.20	1.30	5–2500	5.30 to 7.25	$y = 1.962\ (\pm 0.017) \times - 41.366\ (\pm 18.654)$		0.9998	60.6	183.7
QN-V0	S	2.09	9.17	3.40	1.37	5–2500	8.00 to 9.60	$y = 1.478 \times (\pm 0.033) + 3.407\ (\pm 34.558)$		0.998	174.6	559.5
QN-V0	R	2.09	7.25	2.48	1.37	5–2500	6.00 to 8.00	$y = 1.684\ (\pm 0.028) \times - 59.623\ (\pm 29.223)$		0.9989	115.2	349.2

[a] Mobile phase: 25 mM NH4FA, 6.6 mM FA in ACN:MeOH (1:1 (v/v)), injection volume: 5 μL, temperature: 25 °C, flow: 0.7 mL/min.
[b] Calculated according to ICH guidelines: LOD = $3.3 \times$ STD/slope and LOQ = $10 \times$ STD/slope wherein STD is the residual standard deviation of the calibration function.
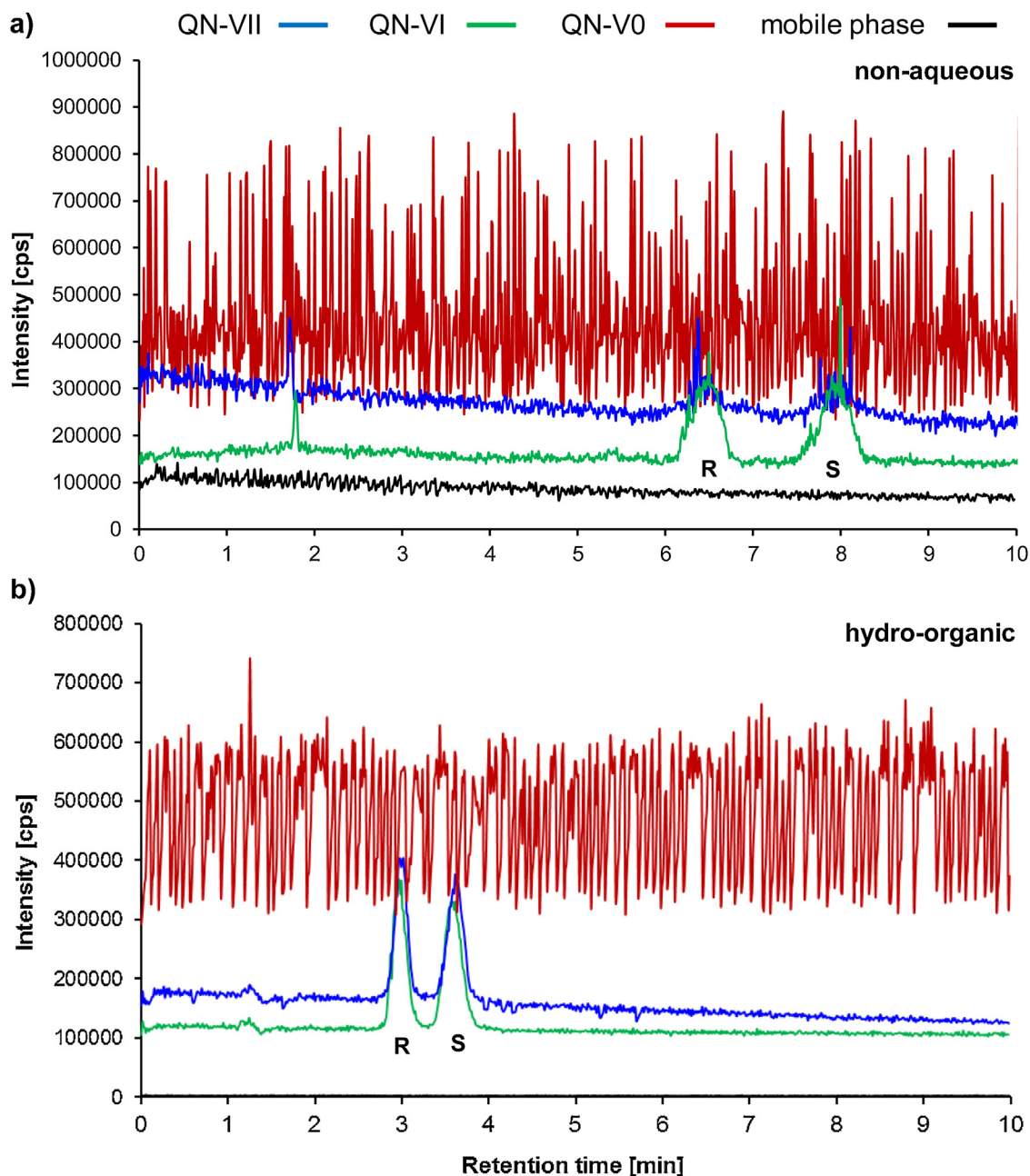
Fig. 6. Background signal for the polysiloxane-coated and brush-type tBuCQN CSPs as measured by total ion current (TIC) with a) non-aqueous and b) hydro-organic mobile phase conditions. TIC of injections of Ac-R,S-Phe onto tBuCQN-modified CSPs based on crosslinked polysiloxane-coated silica prepared via one-step approach (green trace) on crosslinked polysiloxane-coated silica prepared via two-step approach (blue trace), and based on brush-type bonding (red trace). Black trace corresponds to the background signal of the mobile phase (without column). Note that for hydro-organic condition the intensity equals to 3000 cps for the pure mobile phase and is therefore undistinguishable to the time axis.

Non-aqueous mobile phase: 25 mM NH₄FA and 6.6 mM FA in ACN/MeOH (1:1; v/v); hydro-organic mobile phase: 100 mM NH₄Ac in MeOH/water (4:1; v/v); column equilibration prior run 30 min each, without flushing into the source; MS: ESI(+). Other conditions as specified in Experimental section.

on-column both for polysiloxane-bonded and brush-type CSPs ($R^2 > 0.998$). It is noteworthy that slopes increase in the order from brush-type CSP over two-step to single-step polysiloxane-bonded CSP indicating that the best sensitivity is obtained with the latter one (presumably due to less ion suppression). The same trend is confirmed by the LODs and LOQs (Table 3). The results clearly confirm that enhanced detection sensitivity can be achieved with the polysiloxane-bonded CSPs, yet the brush-type tBuCQN CSP can be readily utilized for sensitive detection in MS/MS mode as well.

4. Conclusions

A new platform technology has been developed which enables the synthesis of low-bleed MS-compatible functional stationary phases via thiol-ene click reaction. It is based on either a two-step or a single-step coating of polysiloxane film with reactive thiol functionalities of which a small portion is utilized to crosslink the polymer film to the vinylized silica surface and the remainder of thiol groups is available for immobilization of any chromatographic ligand with vinyl group by radical addition reaction. The concept

has been demonstrated herein for the preparation of a chiral stationary phase based on a vinyl-group containing chiral selector, viz. *tert*-butylcarbamoylquinine. Enantioselecitivity could be easily adjusted to the same level as obtained with a corresponding brush-type CSP. While the brush-type CSP shows stable retention and separation factors over extended period, even years, the benefit of the new bonding chemistry was mainly related to fewer traces of bleeding of well ionizable polar chromatographic ligand and thus reduced background noise in enantioselective HPLC-ESI-QTOF-MS experiments. This renders the new immobilization chemistry better compatible with MS detection and provides lower detection limits. The technology is broadly applicable to any vinyl ligand. Besides current enantioselective chromatography, it has been already examined for the preparation of low-bleed MS-compatible mixed-mode reversed-phase/week anion-exchange stationary phases using *N*-undecenyl-3-aminoquinuclidine as well as amino phases based on allylamine. The results accomplished with these ligands will be presented elsewhere. Some synthesis parameters of the new polysiloxane-coated and crosslinked stationary phases deserve further optimization in order to achieve optimal chromatographic efficiencies. This will be the objective of future studies.

Acknowledgements

References

[1] L.R. Snyder, J.W. Dolan, D.H. Marchand, P.W. Carr, The hydrophobic-subtraction model of reversed-phase column selectivity, Adv. Chromatogr. (Boca Rato, FL, U.S.) 50 (2012) 297–376.

[2] P. Hemstroem, K. Irgum, Hydrophilic interaction chromatography, J. Sep. Sci. 29 (2006) 1784–1821.

[3] P. Jandera, Stationary and mobile phases in hydrophilic interaction chromatography, Anal. Chim. Acta 692 (2011) 1–25.

[4] A. Cavazzini, A. Felinger, Hydrophilic interaction liquid chromatography, in: S. Fanali, P.R. Haddad, C. Poole, P. Schoenmakers, D.K. Lloyd (Eds.), Liquid Chromatography: Fundamentals and Instrumentation, Elsevier, Amsterdam, 2013, pp. 105–119.

[5] M.E.A. Ibrahim, C.A. Lucy, Stationary phases for HILI, Chem. Anal. (Hoboken, NJ, U.S.) 177 (2013) 43–85.

[6] D.V. McCalley, Shell particle and UHPLC technologies for fast analysis of polar compounds in the HILIC mode, RSC Chromatogr. Monogr. 16 (2012) 164–185.

[7] C.T. Mant, R.S. Hodges, Mixed-mode hydrophilic interaction/cation-exchange chromatography (HILIC/CEX) of peptides and proteins, J. Sep. Sci. 31 (2008) 2754–2773.

[8] W. Bicker, M. Laemmerhofer, W. Lindner, Mixed-mode stationary phases as a complementary selectivity concept in liquid chromatography–tandem mass spectrometry-based bioanalytical assays, Anal. Bioanal. Chem. 390 (2008) 263–266.

[9] X. Liu, C. Pohl, A. Woodruff, J. Chen, Chromatographic evaluation of reversed-phase/anion-exchange/cation-exchange trimodal stationary phases prepared by electrostatically driven self-assembly process, J. Chromatogr. A 1218 (2011) 3407–3412.

[10] A.F.G. Gargano, T. Leek, W. Lindner, M. Laemmerhofer, Mixed-mode chromatography with zwitterionic phosphopeptidomimetic selectors from Ugi multicomponent reaction, J. Chromatogr. A 1317 (2013) 12–21.

[11] M. Laemmerhofer, R. Nogueira, W. Lindner, Multi-modal applicability of a reversed-phase/weak-anion exchange material in reversed-phase anion-exchange, ion-exclusion, hydrophilic interaction and hydrophobic interaction chromatography modes, Anal. Bioanal. Chem. 400 (2011) 2517–2530.

[12] A. Zimmermann, J. Horak, O.L. Sanchez-Munoz, M. Laemmerhofer, Surface charge fine tuning of reversed-phase/weak-anion-exchange type mixed-mode stationary phases for milder elution conditions, J. Chromatogr. A 1409 (2015) 189–200.

[13] M. Laemmerhofer, Chiral recognition by enantioselective liquid chromatography: mechanisms and modern chiral stationary phases, J. Chromatogr. A 1217 (2010) 814–856.

[14] A. Cavazzini, L. Pasti, A. Massi, N. Marchetti, F. Dondi, Recent applications in chiral high performance liquid chromatography, Anal. Chim. Acta 706 (2011) 205–222.

[15] W.J. Lough, HPLC chiral stationary phases for the stereochemical resolution of enantiomeric compounds: the current state of the art, Drugs Pharm. Sci. 211 (2012) 95–112.

[16] B. Natalini, R. Sardella, Chromatographic separations and analysis: chiral ion and ligand exchange stationary phases, in: J. Reedijk (Ed.), Reference Module in Chemistry, Molecular Sciences and Chemical Engineering, Elsevier, Amsterdam, 2012, pp. 115–152.

[17] M. Laemmerhofer, Liquid chromatographic enantiomer separation with special focus on zwitterionic chiral ion-exchangers, Anal. Bioanal. Chem. 406 (2014) 6095–6103.

[18] K.K. Unger, Journal of Chromatography Library, Vol. 16: Porous Silica, Its Properties and Use as Support in Column Liquid Chromatography, Elsevier, 1979.

[19] B. Porsch, J. Kratka, Chromatographic stability of silica-based aminopropyl-bonded stationary phases, J. Chromatogr. 543 (1991) 1–7.

[20] M. Hetem, L. Van de Ven, J. De Haan, C. Cramers, K. Albert, E. Bayer, Study of the changes in mono-, di- and trifunctional octadecyl-modified packings for reversed-phase high-performance liquid chromatography with different eluent compositions, J. Chromatogr. 479 (1989) 269–295.

[21] M. Pursch, L.C. Sander, K. Albert, Understanding reversed-phase LC with solid-state NMR, Anal. Chem. 71 (1999) 733A–741A.

[22] J.E. Sandoval, J.J. Pesek, Hydrolytically stable bonded chromatographic phases prepared through hydrosilylation of olefins on a hydride-modified silica intermediate, Anal. Chem. 63 (1991) 2634–2641.

[23] Y. Zhang, H. Luo, P.W. Carr, Silica-based, hyper-crosslinked acid stable stationary phases for high performance liquid chromatography, J. Chromatogr. A 1228 (2012) 110–124.

[24] H. Yu, C. Jia, H. Wu, G. Song, Y. Jin, Y. Ke, X. Liang, Highly stable high performance liquid chromatography stationary phase based on direct chemical modification of organic bridges in hybrid silica, J. Chromatogr. A 1247 (2012) 63–70.

[25] R.R. Drager, F.E. Regnier, High-performance anion-exchange chromatography of oligonucleotides, Anal. Biochem. 145 (1985) 47–56.

[26] J.L. Fausnaugh, E. Pfannkoch, S. Gupta, F.E. Regnier, High-performance hydrophobic interaction chromatography of proteins, Anal. Biochem. 137 (1984) 464–472.

[27] W. Kopaciewicz, F.E. Regnier, Synthesis of cation-exchange stationary phases using an adsorbed polymeric coating, J. Chromatogr. 358 (1986) 107–117.

[28] C.B.G. Bottoli, C.R.M. Vigna, G. Fischer, K. Albert, K.E. Collins, C.H. Collins, Self-immobilization and/or thermal treatment for preparing silica-poly(methyloctylsiloxane) stationary phases, J. Chromatogr. A 1030 (2004) 217–223.

[29] C.G.A. da Silva, C.H. Collins, E. Lesellier, C. West, Characterization of stationary phases based on polysiloxanes thermally immobilized onto silica and metalized silica using supercritical fluid chromatography with the solvation parameter model, J. Chromatogr. A 1315 (2013) 176–187.

[30] M. Schleimer, W.H. Pirkle, V. Schurig, Enantiomer separation by high-performance liquid chromatography on polysiloxane-based chiral stationary phases, J. Chromatogr. A 679 (1994) 23–34.

[31] V. Schurig, S. Negura, S. Mayer, S. Reich, Enantiomer separation on a Chirasil-Dex-polymer-coated stationary phase by conventional and micro-packed high-performance liquid chromatography, J. Chromatogr. A 755 (1996) 299–307.

[32] W.H. Pirkle, L.J. Brice, G.J. Terfloth, Liquid and subcritical CO₂ separations of enantiomers on a broadly applicable polysiloxane chiral stationary phase, J. Chromatogr. A 753 (1996) 109–119.

[33] W.H. Pirkle, G.J. Terfloth, Naproxen-derived segmented and sidechain-modified polysiloxanes as chiral stationary phases, J. Chromatogr. A 704 (1995) 269–277.

[34] G.J. Terfloth, W.H. Pirkle, K.G. Lynam, E.C. Nicolas, Broadly applicable polysiloxane-based chiral stationary phase for high-performance liquid chromatography and supercritical fluid chromatography, J. Chromatogr. A 705 (1995) 185–194.

[35] C. Wolf, W.H. Pirkle, Synthesis and evaluation of a copolymeric chiral stationary phase, J. Chromatogr. A 799 (1998) 177–184.

[36] H.W. Stuurman, J. Koehler, G. Schomburg, HPLC separation of enantiomers using quinine, covalently bonded to silica as stationary phase, Chromatographia 25 (1988) 265–271.

[37] E.J. Carrasco-Correa, G. Ramis-Ramos, J.M. Herrero-Martínez, M. Lämmerhofer, Polymethacrylate monoliths with immobilized poly-3-mercaptopropyl methylsiloxane film for high-coverage surface functionalization by thiol-ene click reaction, J. Chromatogr. A 1367 (2014) 123–130.

[38] E. Francotte, D. Huynh, Immobilized halogenophenylcarbamate derivatives of cellulose as novel stationary phases for enantioselective drug analysis, J. Pharm. Biomed. Anal. 27 (2001) 421–429.

[39] J. Shen, T. Ikai, Y. Okamoto, Synthesis and application of immobilized polysaccharide-based chiral stationary phases for enantioseparation by high-performance liquid chromatography, J. Chromatogr. A 1363 (2014) 51–61.

[40] P. Franco, T. Zhang, Common screening approaches for efficient analytical method development in LC and SFC on columns packed with immobilized polysaccharide-derived chiral stationary phase, Methods Mol. Biol. (NY, U.S.) 970 (2013) 113–126.

[41] S.G. Allenmark, S. Andersson, P. Moeller, D. Sancher, A new class of network-polymeric chiral stationary phases, Chirality 7 (1995) 248–256.

[42] S. Andersson, S. Allenmark, P. Moeller, B. Persson, D. Sanchez, Chromatographic separation of enantiomers on *N,N′*-diallyl-ʟ-tartardiamide-based network-polymeric chiral stationary phases, J. Chromatogr. A 741 (1996) 23–31.

[43] M. Laemmerhofer, W. Lindner, Quinine and quinidine derivatives as chiral selectors. I. Brush type chiral stationary phases for high-performance liquid chromatography based on cinchonan carbamates and their application as chiral anion exchangers, J. Chromatogr. A 741 (1996) 33–48.

[44] N.M. Maier, L. Nicoletti, M. Lammerhofer, W. Lindner, Enantioselective anion exchangers based on cinchona alkaloid-derived carbamates: influence of C8/C9 stereochemistry on chiral recognition, Chirality 11 (1999) 522–528.

[45] A. Mandl, L. Nicoletti, M. Lammerhofer, W. Lindner, Quinine versus carbamoylated quinine-based chiral anion exchangers. A comparison regarding enantioselectivity for N-protected amino acids and other chiral acids, J. Chromatogr. A 858 (1999) 1–11.

[46] P.V.D.V.E.F. Vansant, K.C. Vrancken, Appendix C ^{29}Si CP Mas NMR cross polarization magic angle spinning nuclear magnetic resonance, in: P.V.D.V.E.F. Vansant, K.C. Vrancken (Eds.), Studies in Surface Science and Catalysis, Elsevier, 1995, pp. 505–510.

[47] B.A. Ashu-Arrah, J.D. Glennon, K. Albert, Synthesis and characterization of bonded mercaptopropyl silica intermediate stationary phases prepared using multifunctional alkoxysilanes in supercritical carbon dioxide as a reaction solvent, J. Chromatogr. A 1222 (2012) 38–45.

[48] C. Hellriegel, U. Skogsberg, K. Albert, M. Laemmerhofer, N.M. Maier, W. Lindner, Characterization of a chiral stationary phase by HR/MAS NMR spectroscopy and investigation of enantioselective interaction with chiral ligates by transferred NOE, J. Am. Chem. Soc. 126 (2004) 3809–3816.

[49] N.M. Scully, G.P. O'Sullivan, L.O. Healy, J.D. Glennon, B. Dietrich, K. Albert, Preparation of a mercaptopropyl bonded silica intermediate in supercritical carbon dioxide: synthesis, characterisation and chromatography of a quinine based chiral stationary phase, J. Chromatogr. A 1156 (2007) 68–74.

Appendix VI

Accepted Manuscripts

Publication IX	
Title	Gold nanoparticle-conjugated pepsin for efficient solution-like heterogeneous biocatalysis in analytical sample preparation protocols
Authors	*Hoeldrich, Markus; **Sievers-Engler, Adrian**; Laemmerhofer, Michael*
Published in	Analytical and Bioanalytical Chemistry (408), 2016, issue 20, p. 5415-5427
DOI	http://dx.doi.org/10.1007/s00216-016-9657-y

Publication X	
Title	Tri- and Tetra-substituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3
Authors	*Muth, Felix; Ahmed El-Gokha, Ahmed; Ansideri, Francesco; Eitel, Michael; Döring, Eva; **Sievers-Engler,Adrian**; Lange, Andreas, Boeckler, Frank M.;Lämmerhofer, Michael; Koch, Pierre, Laufer, Stefan A.*
Published in	Journal of Medicinical Chemistry (60), 2017, issue 2, p. 594-607.
DOI	http://dx.doi.org/10.1021/acs.jmedchem.6b01180

Publication XI	
Title	Papain-functionalized gold nanoparticles as heterogeneous biocatalyst for bioanalysis and biopharmaceuticals analysis.
Authors	*Liu,Siyao,; Höldrich, Markus; **Sievers-Engler, Adrian**; Horak, Jeannie; Lämmerhofer, Michael*
Published in	Analytica Chimica Acta (963), 2017, p. 33-43
DOI	https://doi.org/10.1016/j.aca.2017.02.009

CrossMark

PAPER IN FOREFRONT

Gold nanoparticle-conjugated pepsin for efficient solution-like heterogeneous biocatalysis in analytical sample preparation protocols

Markus Höldrich[1] · Adrian Sievers-Engler[1] · Michael Lämmerhofer[1]

Abstract Immobilization of enzymes on mesoporous microparticulate carriers has traditionally been accompanied by reduction in enzyme activity. Herein, we document that immobilization of pepsin via amide coupling on gold nanoparticles (GNPs) with a carboxy-terminated hydrophilic PEG_7 shell resulted in a heterogeneous nanobiocatalyst with essentially equivalent turnover rates, k_{cat} (90 %), and enhanced catalytic efficiencies, k_{cat}/K_M (107 %), compared to homogeneous catalysis with pepsin in free solution for cytochrome C as model substrate. This heterogeneous catalyst showed further at least equivalent bioactivity in a digestion reaction of a protein mixture consisting of cytochrome C, bovine serum albumin, and myoglobin. UHPLC–ESI-QTOF-MS/MS analysis of the digests with subsequent Mascot database search allowed unequivocal identification of all proteins with high score and good sequence coverage. The functionalized nanoparticles were further characterized by Vis spectroscopy in terms of the surface plasmon resonance (SPR) band, by dynamic light scattering (DLS) with regard to hydrodynamic diameters, and in view of their ζ potentials at each step of synthesis and surface modification. These measurements also revealed that the pepsin-functionalized GNPs were sufficiently stable over at least 1 month; thus providing a satisfactory shelf life to the heterogeneous catalyst. Advantageously, the pepsin–GNP bioconjugate can be conveniently removed after reaction by simple centrifugation steps which makes them a useful tool for analysis of therapeutic peptides and proteins, including monoclonal antibodies. The practical utility of the nanobiocatalyst was documented by digestion of a monoclonal antibody which yielded the $F(ab')_2$ fragment with a mass of 97,619.4 Da.

ABC Highlights: authored by *Rising Stars and Top Experts.*

✉ Michael Lämmerhofer
  Michael.laemmerhofer@uni-tuebingen.de

[1] Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

Introduction

Workflows in protein analytics, e.g., in proteomics and analysis of protein therapeutics, commonly involve enzymatic digestion protocols prior to their LC–MS/MS analysis [1]. Trypsin is most popular for proteolytic cleavage into peptide fragments and is very efficient and yields short peptides with Arg or Lys at the C-terminal end. As a result of very short peptide fragments, MS identification often fails and leads to incomplete sequence information. Alternative enzymes are also often utilized, in proteomics primarily Lys-C, Glu-C, chymotrypsin, and pepsin [1] and in therapeutic protein (i.e., monoclonal antibody) analysis pepsin, papain, IdeS [2], or ideally a combination of these biocatalysts [1]. These enzymes can also be very helpful for stereochemical configuration-dependent sequence elucidations of therapeutic peptides and peptidic natural products [3]. To avoid detrimental interferences from enzyme and confusing results in such applications, immobilization of these biocatalysts to solid surfaces, e.g., beaded agarose or monolithic supports, has been suggested as a viable option to enable straightforward

removal of the enzyme after the proteolytic reaction [4]. There are a number of other advantages associated with such heterogeneous catalysts, amongst which ameliorated enzyme stability and reusability are prominent ones [5, 6]. Often, however, bonding to solid supports results in reduction of bioactivity of immobilized enzymes mainly because of hindered substrate access to the active site or unfavorably altered conformation of immobilized enzyme [5, 6]. In an attempt to overcome such limitations of heterogeneous catalysts, researchers have evaluated nanomaterials as carriers for enzyme immobilization [7, 8]. In fact, a number of studies reported increased activities and enhanced reaction rates as well as improved stabilities when enzymes were immobilized on nanocarriers [9, 10].

In general, according to IUPAC nanomaterials have a length scale of less than 100 nm in at least one dimension [11]. This gives them peculiar properties of which their large surface-to-volume ratio is perhaps the most significant one in terms of heterogeneous bionanocatalysis for providing a large surface for functionalization with enzyme and a large interface for enzyme–substrate reactions [12–14]. Frequently utilized nanoscaled supports for nanobiocatalysts comprise magnetic nanoparticles [8], silica nanoparticles, polymeric nanomaterials, metal oxides, nanofibers, and metallic nanoparticles onto which enzymes have been immobilized by adsorption, covalent bonding, ionic interactions, via affinity tags, by cross-linking, or physical entrapment (entanglement) in a polymeric network [7, 8]. Amongst the metallic nanoparticles, gold nanoparticles (GNPs) have received particular attention and belong to one of the most intensely studied nanomaterials [15]. They can be prepared by a straightforward and low-cost size-controlled synthesis by reduction and stabilization of gold(III) chloride with citrate according to the Turkevich–Frens method [16, 17]. Through a nucleation and growth-controlled mechanism [15, 18, 19], GNPs can be obtained in the size range between 1 and 50 nm. Furthermore, convenient quality control by Vis spectroscopy due to the surface plasmon resonance (SPR) band with absorbance maximum around 520 nm makes this material attractive as a carrier for biofunctionalized nanomaterials [20, 21]. Besides adsorptive bonding of enzymes by ionic and hydrophobic interactions, enzymes can be conveniently immobilized via bifunctional ligands containing a terminal thiol group for direct attachment to the GNP surface and dative Au–S bond through self-assembled monolayer (SAM) formation [22] and subsequent conjugation of enzyme [12, 23]. For this reason it is not further surprising that GNPs have become popular enzyme carriers for sample preparation in protein analysis, e.g., with trypsin [12, 24, 25], for enzyme-based sensor devices [26], for immunoassays [27, 28], and for theranostic applications, e.g., in vivo tumor targeting and detection by surface-enhanced Raman scattering [29].

In this work, we report on the synthesis and characterization of pepsin-modified heterogeneous nanobiocatalysts based on a

gold core with a hydrophilic PEG_7 shell onto which pepsin was covalently attached and their proteolytic performance for sample preparation in protein analytics. Pepsin is a digestive protease with a molecular weight around 36 kDa and an isoelectric point (pI) at approximately 3.2 [30, 31]. Pepsin cleaves peptide bonds of proteins preferentially at the C-terminal side of phenylalanine and leucine residues [32, 33]. It is sometimes used in proteomics as an alternative to trypsin (roughly estimated in about 0.4 % of proteomic studies [1]), but also for antibody digestion near the hinge region to produce specifically F(ab')$_2$ and Fc fragments [2]. For this reason, it was also a popular enzyme target for its immobilization to various supports such as microparticles [34, 35], magnetic beads [36], monoliths [37, 38], nanospray emitter [39], planar surface low-density polyethylene (PE) films, or on polycarbonate (PC) plates, on microscope glass slides [40], dextran-modified fused-silica capillaries [41], and on fibers [42]. In a few studies, pepsin was immobilized on nanoparticulate carriers such as alumina nanoparticles ([43]), anisotropic gold nanoclusters [44], and colloidal gold [45]. However, the weak physical binding in these studies with noncovalent immobilization is regarded as suboptimal in terms of chemical stability of these nanobiocatalysts as they have a tendency to undergo enzyme leaching from the surface. To alleviate such limitations we propose herein covalently linked pepsin–GNP conjugates with an optimized surface chemistry that provides both chemical as well as colloidal stability and does not compromise bioactivity

Materials and methods

Materials

Pepsin (from hog stomach, EC 3.4.23.1, 3651 U mg^{-1}) was received from Fluka (Steinheim, Germany). Albumin, from bovine serum (BSA, EC number 232-936-2, MW 66 kDa) was obtained from Sigma-Aldrich (Munich, Germany). Cytochrome C (CYC_HORSE, from horse heart, EC number 232-700-9, MW 12,270 Da) and myoglobin (MYG_HORSE, from horse skeletal muscle, EC number 309-705-0, MW 16, 900 Da) were supplied by Calbiochem (Darmstadt, Germany).

Gold(III) chloride trihydrate (HAuCl$_4$·3H$_2$O), trisodium citrate, tris(hydroxymethyl)aminomethane (Tris), N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide (EDC), O-(2-carboxyethyl)-O'-(2-mercaptoethyl)heptaethylene glycol (HS-PEG$_7$-COOH), Folin & Ciocalteu's phenol reagent glycerol, a silver staining kit, and monoclonal anti-HSA antibody for ELISA were obtained from Sigma-Aldrich (Munich, Germany).

Preparation of GNPs and immobilization of pepsin

The preparation of GNPs was based on the reduction and simultaneous stabilization of gold(III) chloride trihydrate with

trisodium citrate in accordance to the Turkevich–Frens method [12, 16, 17, 46]. For the synthesis of 10 mL solution of GNPs, 5.05 mg $HAuCl_4$ (1.28 mmol in 11.25 mL doubly deionized water (ddH_2O) corresponding to a final concentration of 1.14 mM) was used and heated at reflux under constant stirring. Afterwards 1.25 mL of 20.5 mM trisodium citrate solution (6.03 mg mL^{-1}) was added (final concentration of 2.28 mM in a total volume of 11.25 mL) and heated for a further 10 min, while the color turned red. The GNP solution was allowed to cool down and stirred at room temperature for another 60 min. The GNP solution was finally stored at 4 °C until usage [47].

To obtain covalently immobilized pepsin, first 1 µL of HS-PEG$_7$-COOH solution was added to 1 mL of aforementioned GNP solution and the bifunctional linker immobilized overnight under constant stirring at room temperature. Afterwards the solution was centrifuged at 12,000 rpm and resuspended in 20 mM MES buffer pH 4.5. This washing step was repeated twice. The 50 µL of EDC (12 mM in MeOH) to activate carboxylic acids and 250 µL pepsin (6 mg mL^{-1} in 20 mM MES buffer pH 4.5) were added and stirred for 2 h at room temperature. The nanoparticles were washed in 20 mM Tris-HCl buffer pH 7.5 twice to cap residual activated carboxylic groups. The supernatant of the first washing step was used for protein quantification by the Lowry assay (see below) [12].

The resultant nanoparticles were washed twice and resuspended in ddH_2O and characterized after each step of surface functionalization by Vis spectroscopy, DLS, and ζ potential measurements (see Fig. 1 and Electronic Supplementary Material (ESM) Fig. S1).

In an additional study, in one synthesis batch nanoparticles of each step of surface functionalization were tested for their stability over 35 days. For this purpose, all samples were diluted 1:5 with ddH_2O and characterized by Vis spectroscopy (see ESM Fig. S2).

Vis spectroscopic characterization of nanoparticles as well as size and zeta potential measurements

The SPR band was measured by Vis spectroscopy after each step of surface modification; spectra were acquired in the wavelength range between 350 and 800 nm. All measurements were performed with a VWR UV-1600PC spectrophotometer. The differently functionalized nanoparticles were further characterized by DLS and their ζ potentials, derived from electrophoretic mobility measurements, before and after each step of chemical modification using a Zetasizer Nano ZS instrument (Malvern Instruments, Herrenberg, Germany). The Zetasizer was equipped with a He–Ne laser, and the detection was performed at 173° backscatter detection mode. Samples were diluted in ddH_2O (1:5; v/v) and measured as triplicate. Each value was the mean of 15 subruns.

Chromatographic characterization of pepsin activity

A kinetic study was performed with pepsin in solution and with functionalized heterogeneous nanobiocatalysts with immobilized pepsin to determine the Michaelis–Menten parameters. Thus, several concentrations of CYC were digested in 20 mM sodium acetate pH 4.5 for 10 min at 37 °C. The digestion was stopped by adding 0.2 mL of 0.1 M NaOH to 0.8 mL reaction mixture (0.6 mL reaction buffer + 0.1 mL pepsin + 0.1 mL CYC stock solutions with different concentrations) to achieve neutral pH. For the study of pepsin in homogeneous solution as control reaction, 100 µL enzyme solution (2.5 mg mL^{-1} stock solution, 0.25 mg mL^{-1} final concentration) in the aforementioned buffer was used for the digestion. The approach with the immobilized pepsin was performed in 1 mL total volume of the nanoparticle solution (0.7 mL nanoparticle solution in reaction buffer + 0.1 mL CYC stock solutions with different concentrations). Therefore pepsin@GNP was centrifuged at 12,000 rpm for 10 min and resuspended in 0.7 mL 20 mM sodium acetate pH 4.5. The supernatant after the centrifugation step was discarded. The digestion was stopped by adding 0.2 mL 0.1 M NaOH (total volume reaction mixture 1 mL), the nanoparticles were centrifuged, and the supernatant was analyzed by LC–UV for quantification of undigested CYC.

Michaelis–Menten parameters K_M and v_{max} were calculated by linear and bilinear equations according to Lineweaver–Burk diagrams [48]. A calibration function was established and used to quantify undigested CYC in the reaction mixture. All samples were measured as triplicate. For the LC analysis method of undigested CYC, an Agilent 1100 series HPLC instrument with a binary gradient pump, vacuum degasser, autosampler, column oven, and a variable wavelength detector was used. The separation was performed with a 10-min gradient on a monolithic poly(styrene-*co*-divinylbenzene) Proswift™ RP-1S 50 × 4.6 mm ID column from Dionex (ThermoFisher
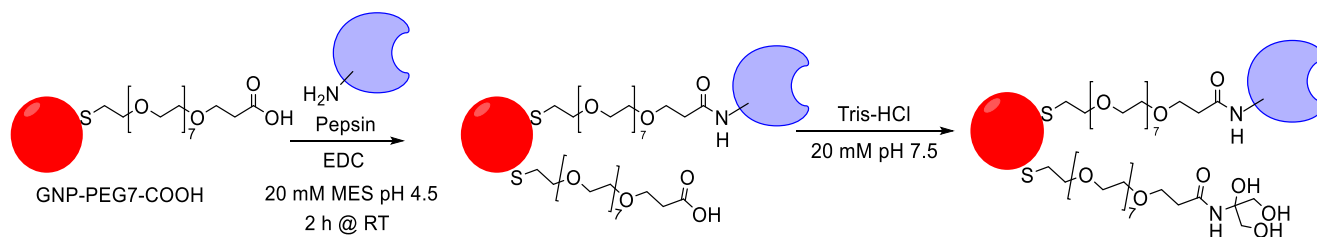

Fig. 1 Reaction scheme for functionalization of GNPs with bifunctional cross-linker and pepsin

Scientific) with 50 mM ammonium formate pH 3.5 containing 5 % ACN in channel A and ACN + 5 % 50 mM ammonium formate pH 3.5 in channel B (0–2 min, 20 % B; 2–6 min, 60 % B; 6.01–10 min, 20 % B). The flow rate of the mobile phase was 1 mL/min, and the column temperature was set to 40 °C. The injection volume was 5 μL, and detection was performed at 280 nm.

The method was validated in accordance with the International Council on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guidelines. Linearity, precision, and accuracy were evaluated. The LOD and LOQ were determined as the concentrations with a signal-to-noise ratio of 3:1 and 9:1, respectively (see ESM Figs. S3, S4, Table S1).

Protein determination by Lowry assay

The amount of immobilized pepsin was determined by the Lowry assay [49]. Briefly, the supernatant collected after the enzyme immobilization step was used for the quantification of nonimmobilized pepsin. The amount of immobilized enzyme could then be back calculated by applying the mass balance equation. A calibration function was set up with free pepsin in 20 mM MES buffer pH 4.5 in the range 0.05 to 1 mg/mL. A 25-μL aliquot of the sample/standard/blank was mixed with 125 μL of Lowry A solution and reacted for 20 min. Afterwards 12.5 μL Folin & Ciocalteu's phenol reagent (Lowry reagent B) was added and reacted for further 30 min before measurement. Lowry reagent A was freshly prepared every day (2.45 mL 4 % Na_2CO_3 anhydrous in ddH_2O, 2.45 mL 0.1 M NaOH, 0.1 mL 2 % Na-K-tartrate in ddH_2O, and 0.05 mL $CuSO_4 \cdot 5H_2O$ in ddH_2O). Folin & Ciocalteu's phenol reagent was commercially available from Sigma-Aldrich.

For the photometric determinations a Versa max microplate reader from Molecular Devices was used, and the wavelength was set to 650 nm. All measurements were carried out at room temperature. For more details see also ESM Figs. S5 and S6.

Protein digestion with pepsin-conjugated GNPs

BSA, CYC, and MYG (0.2 mg mL^{-1} of each protein) were utilized as model proteins to test the digestion efficiency of conjugated pepsin. For this purpose, 0.2 mL of the protein mixture was transferred to 1 mL of the nanoparticle suspension (pepsin@GNP in 1 mL 20 mM sodium acetate pH 4.5). The reaction with the nanoparticle approach was stopped by centrifugation (12,000 rpm for 10 min), which resulted in a clear supernatant for LC–MS/MS analysis. For the digestion protocol with free pepsin, 0.3 mg mL^{-1} pepsin was used (0.25 mg mL^{-1} final concentration in 1.2 mL total volume); 0.2 mL of the protein mixture was added to 0.8 mL pepsin in digestion buffer. Then 100-μL aliquots were taken, and the reaction was stopped

by adding 20 μL of 0.1 M NaOH before injecting the sample for LC–MS/MS analysis.

The digestion of the protein mixture was performed in 20 mM sodium acetate buffer pH 4.5 at 37 °C with a digestion time from 4 to 24 h. As control a sample at 0 h digestion time was also taken and analyzed. Blank samples as 20 mM sodium acetate pH 4.5 were prepared as well.

Recyclability (see ESM Fig. S7) and pepsin leaching from pepsin@GNP bioconjugate (see ESM Fig. S8) were also examined.

LC–MS/MS method

All samples were injected via the PAL HTC-xt autosampler (CTC Analytics, Zwingen, Switzerland). The separation was performed with an Agilent 1290 LC system (Agilent Technologies, Waldbronn, Germany) on an Aeris Peptide C18 (3.6 μm 150 × 2.1 mm ID) column from Phenomenex (Torrance, CA, USA) with a 90-min gradient containing ddH_2O with 0.1 % (v/v) formic acid in channel A and acetonitrile with 0.1 % (v/v) formic acid in channel B (0–5 min, 5 % B; 5–60 min, 5–60 % B; 60–65 min, 60–95 % B; 65–70 min, 95 % B; 70–72 min, to 5 % B; 72–90 min, 5 % B). The peptide masses were detected with a Triple TOF 5600+ quadrupole time-of-flight mass spectrometer from Sciex (Concord, Ontario, Canada) equipped with a DuoSpray source operated in ESI positive mode. The source temperature was set to 350 °C, and the ion spray voltage floating (ISVF) was 5500 V. The pressure of the ion source gases 1–3 was 50, 40, and 30 psi. Finally, the declustering potential was set to 100 V, and the collision energy to 10 V. Ions were measured in information-dependent data acquisition (IDA) mode for ions greater than 100 Da with an intensity threshold greater than 10 counts per second. Accumulation times for MS (survey scan) and MS/MS experiments were set to 250 ms. The dynamic background subtraction was activated, and the collision energy for all MS/MS experiments was set to 25 V.

A Mascot search was performed to detect the digested protein and to examine the occurrence of autodigestion of the immobilized pepsin. Mascot version 2.0.05 was used with the Swiss-Prot database. Pepsin A was selected as the enzyme. Precursor and fragment ion mass (monoisotopic) tolerance was set to ±0.2 Da. Peptide charge state was set to +1, +2, +3, and one missed cleavage was allowed. Taxonomy and modifications were not specified.

More detailed results on the UHPLC–ESI-QTOF-MS/MS analysis of the digested protein mix can be found in ESM Figs. S9–S12 and Table S2.

Antibody digestion with pepsin@GNP

For this experiment the intact mass of monoclonal anti-HSA from an ELISA kit was determined before and after the digestion with immobilized pepsin by UHPLC–μESI-QTof-MS.

For the digestion 100 μL pepsin@GNP in 20 mM sodium acetate pH 4.5 was mixed with 100 μL anti-HSA (1:100 dilution in 20 mM sodium acetate pH 4.5). The digestion was performed at 37 °C for 4 h. To obtain a clear supernatant for analysis the nanobiocatalysts were centrifuged (12,000 rpm for 10 min).

LC–MS method for determination of intact protein masses

A Zorbax column (SPE 300 Å C18 5 μm 35 × 0.3 mm ID) from Agilent (Santa Clara, CA, USA) was used for online SPE. Within the first 2 min the column was used for desalting and enrichment of the sample with a flow of 50 μL/min 100 % ddH$_2$O with 0.1 % (v/v) formic acid in channel A. From 2.01 to 8 min the adsorbed, desalted proteins were eluted from the SPE column and flushed into the mass spectrometer with 50 μL/min 100 % acetonitrile containing 0.1 % (v/v) formic acid in channel B. A 2-min pre-equilibration step was performed to regain the starting conditions of the method. The protein masses were detected with a Triple TOF 5600+ quadrupole time-of-flight mass spectrometer from Sciex (Concord, Ontario, Canada) equipped with a DuoSpray source operated in μESI positive mode, using a 50-μm-ID μESI-needle (Sciex). The source temperature was set to 400 °C, and the ISVF was 5100 V. Nebulizer gas (GS1) was set to 40 psi, drying gas (GS2) to 40 psi, and curtain gas to 30 psi. Finally, the declustering potential (DP) was set to 230 V, and the collision energy (CE) to 30 V. TOF-MS scan ranged from 500 to 4000 Da. For intact protein detection Sciex's Intact-Protein script was activated, Q1 transmission was set to 100 % at 1250 m/z, and sensitivity was increased by summing 60 time bins. Data processing and identification of the intact protein mass were performed with the Bio Tool Kit from Sciex.

Results and discussion

Synthesis and characterization of functionalized GNPs

Citrate-capped GNPs were prepared according to the Turkevich–Frens method [16, 17]. This approach is based on the reduction of HAuCl$_4$ by trisodium citrate and allows a straightforward size-controlled synthesis of GNPs in the range of 10 to 30 nm by variation of the molar citrate-to-gold tetrachloride ratio in the range between 6:1 and 2:1 [47]. Citrate ions attached to the GNP surface after synthesis are responsible for the good colloidal stability of the resultant nanoparticle suspensions. A convenient strategy for enzyme immobilization on nanoparticulate carriers represents adsorptive bonding driven by attractive electrostatic interaction forces. This approach towards heterogeneous nanobiocatalysts was examined herein as well, yet was found to be suboptimal in terms

of stability of resultant nanocolloids. Therefore, immobilized pepsin, with a pI between 2 and 3, may easily detach under reaction conditions (pH between 1 and 4), thereby contaminating the resultant reaction solution after removal of the nanoparticles with proteinogenic material which may interfere with final analytical determinations. Hence, a covalent attachment strategy of pepsin was adopted according to the reaction scheme in Fig. 1. The bifunctional linker HS-PEG$_7$-COOH was first self-assembled in a dense monolayer on the gold surface via stable dative thiol–gold bonds. It was previously shown that this linker provides stable self-assembled monolayers (SAMs) with surface coverages of 4.29 ± 0.45 ligands per nm^2 [50]. The PEG linker supports colloidal stability and prevents nonspecific binding of proteins. It was found that enzyme–GNP conjugates have better colloidal stability when the enzyme was bonded via a PEG linker as compared to corresponding alkyl linkers and compared to adsorptively bonded enzyme–GNP conjugates as well [12]. Longer spacers turned out to be advantageous in terms of bioactivity, and therefore the bifunctional linker HS-PEG$_7$-COOH was selected in the current work as first choice. The terminal carboxylic groups constitute reactive anchor functionalities for bonding of pepsin by direct amide coupling with EDC under weakly acidic conditions. As the last step in the synthesis, remaining EDC-activated carboxylic groups were capped with Tris moieties to yield the final pepsin-functionalized GNPs with residual carboxylic acid groups. The latter are supportive for the colloidal stability and shelf life of the resultant bionanocatalyst.

All nanoparticles were characterized after each step of synthesis and surface modification, respectively, by Vis spectroscopy for measurement of changes of the SPR band, by DLS analysis for nanoparticle size (i.e., hydrodynamic diameter) characterization, and by determination of the ζ potential. The results are illustrated in Fig. 2.

Citrate-stabilized GNPs obtained by synthesis with a 2:1 molar ratio of citrate/HAuCl$_4$ had a mean diameter of 31.4 ± 0.9 nm according to DLS measurements and a ζ potential of −34.7 ± 1.7 mV. Such nanoparticle sizes were recently found to be favorable over smaller ones for protein immobilization [12] and were therefore used as the starting carrier for all batches of modified GNPs in this study. The final concentration of the GNP solution of 1.5×10^{-6} mM was calculated with the results of the DLS measurements and the Lambert–Beer law [47].

Vis spectroscopy is the most straightforward and useful tool for quality control of surface modification. Alterations in size, size distribution, and shape as well as chemical changes at the surface lead to shifts of the characteristic SPR band which is found as an absorption band with a maximum at around 525 nm according to the Mie theory [51, 52]. As displayed in Fig. 2a, the SPR band as monitored by Vis spectroscopy is slightly shifted to higher wavelength upon each
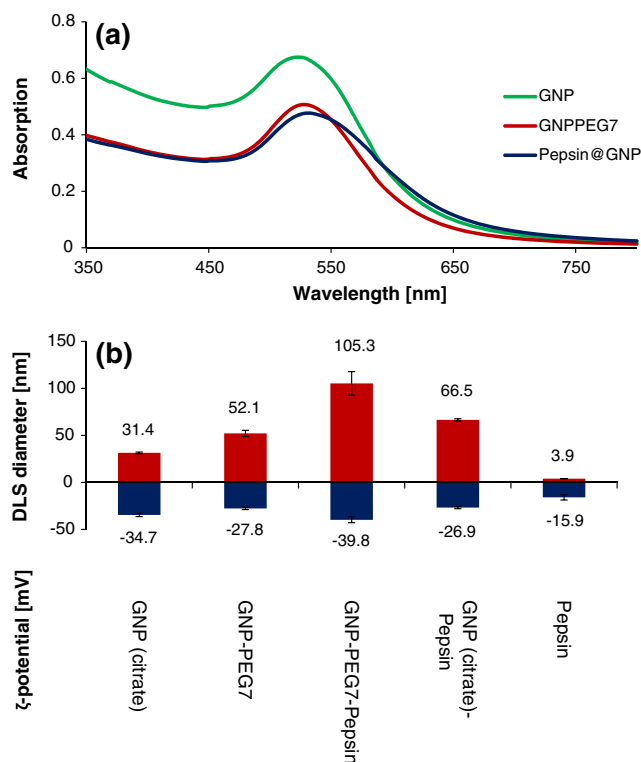
Fig. 2 Characterization of nanoparticles after each step of synthesis and surface functionalization, respectively. **a** Vis spectra showing slight shifts of the SPR band upon surface modification, **b** size (hydrodynamic diameters) as measured by DLS and ζ potentials of all nanoparticle stages, along with adsorptively immobilized pepsin-modified nanoparticles and free pepsin

step of surface modification (from $\lambda_{max} = 525 \pm 0.6$ nm for citrate-capped GNPs to $\lambda_{max} = 528 \pm 1$ nm after pegylation to $\lambda_{max} = 532 \pm 1.7$ nm after bonding of pepsin) and confirms successful bonding of the respective ligand at each step. It becomes further evident that no significant aggregation occurs upon pegylation which yielded carboxy-terminated GNPs nor after enzyme immobilization which afforded the final pepsin-conjugated GNPs. Aggregation typically leads to broad absorption bands with $\lambda_{max} > 550$ nm which are absent in the current Vis spectra (Fig. 2a) [46]. Narrow absorption bands are observed for all three studied particle types as a result of the absence of aggregation. The slightly broader absorption band for the pepsin-modified GNPs may be due to asymmetrically shaped particles and departure from strictly spherical shape after enzyme immobilization [47].

Sufficient colloidal stability is also confirmed by the ζ potentials of the modified particles. In any case, ζ potentials were less than −25 mV (Fig. 2b) which is indicative of sufficient electrostatic repulsion overcoming attractive van der Waals interactions that are the prime cause for aggregation of nanoparticles.

Nanoparticle diameters were measured by DLS. As expected, with each layer of surface modification the size of the resultant nanoparticles increased, thereby confirming successful ligand attachment. The bifunctional PEG_7 linker with

terminal carboxylic group has a length of about 3.5 nm [50]. After the PEG_7 linker bonded to the GNP surface by SAM formation, the diameter of the nanoparticle grew from 31 nm to 52 nm, which is in good agreement with expectations. The slight difference between theoretical and experimental values may be due to strong solvation of pegylated GNPs. The sphere equivalent particle diameter further increased to 105 nm after coupling of pepsin to the carboxy-terminated GNPs (Fig. 2b).

The surface coverage of pepsin on the GNPs was determined by protein quantification with the Lowry assay measuring free pepsin in the supernatant after the reaction and applying mass balance considerations. The pepsin concentration on the GNPs was determined to be 0.25 ± 0.03 mg mL^{-1} ($7.24 \pm 0.87 \times 10^{-6}$ mol/L). For sake of comparison, the data for nanoparticles with pepsin adsorptively bonded directly to citrate-capped GNPs are also included in Fig. 2b. It can be seen that the diameter was significantly smaller (66 nm) as expected because of the missing PEG shell. Most notably, the ζ potential of this adsorptively bonded bioconjugate was much lower compared to the covalently bonded pepsin–GNP conjugate. In fact, it turned out that such GNPs are less stable, may easily aggregate, and show reduced shelf life, which makes them less suitable for the intended analytical purpose. In contrast, covalent pepsin-modified GNPs with PEG linker showed excellent stability over extended periods (see also ESM Fig. S1). Hence, the covalent immobilization strategy proposed herein is favorable particularly from the viewpoint of shelf life of the functionalized nanoparticles and with regards to bleeding of enzyme into the reaction mixture in the course of sample preparation, which otherwise gives rise to interferences in the subsequent analysis.

Characterization of heterogeneous nanobiocatalyst by Michaelis–Menten enzyme kinetics

As a simple test to characterize the functionality of the new nanoparticulate biocatalyst, cytochrome C (CYC) was digested with pepsin–GNP bioconjugate in comparison to homogeneous catalysis with free pepsin, and the data were processed according to the Michaelis–Menten enzyme kinetics model. Three replicate batches of incubations at each substrate concentration level ($n=3$) were examined for both free and immobilized enzyme. To ensure a representative comparison, the enzyme concentration was kept constant in reactions with free and immobilized pepsin (concentrations of pepsin were 7.24×10^{-6} for free and $7.24 \pm 0.87 \times 10^{-6}$ mol/L incubation, respectively, for immobilized pepsin). For this purpose, an LC–UV method for the quantitative determination of nondigested CYC was developed and validated according to ICH guidelines (details of method validation and performance are given in ESM Figs. S3, S4, and Table S1). The assay, based on separation of the intact CYC protein from digested peptides which elute more or less unretained in the established RPLC method with poly(styrene-co-divinylbenzene) monolithic column, allowed accurate and precise quantification

of undigested CYC. For the quantitative evaluation of the enzyme kinetics the reaction rate was determined from the time-dependent decrease of educt (CYC). Incubations of variable concentrations of substrate (CYC) with free and immobilized pepsin were analyzed using the Michaelis–Menten equation (see Fig. 3). Both nonlinear and linear curve fitting to derive K_M and v_{max} gave insufficient quality of fit for both digestion with free and immobilized enzyme. The nonlinear behavior in the Lineweaver–Burk diagram in Fig. 2b is evident. Nevertheless, for a rough estimation of the kinetic parameters K_M and v_{max} were calculated from the data plotted according to the Lineweaver–Burk diagram. The results are summarized in Table 1. K_M values of 29 (±3) μM and 21 (±2) μM for free and immobilized pepsin, respectively, were calculated; these values indicated that immobilization did not negatively affect the affinity of the substrate for the enzyme. Apparently, access of CYC to the catalytic site is not hindered considerably by the linkage of pepsin to GNPs, and the chemical bonding does not seem to induce major conformational changes in the enzyme with negative effects on bioactivity. At this point it is, however, also quite clear that as a result of the intrinsic structural heterogeneity of the heterogeneous nanobiocatalyst originating from polydispersity in size, shape, protein coverage, and in particular orientation, each individual microscopic state of the enzyme conjugate may possess distinctly unique kinetic properties. As a consequence, the characteristic kinetic parameter K_M of the pepsin–GNP conjugate actually represents macroscopically a weighted average value of individual microscopically distinct nanobiocatalysts. Since the nondirected immobilization by the selected amide coupling leads to random protein orientation and hence presumably to hindered access of the active site in some orientations, the decrease of the macroscopic K_M value by 27 % relative to free pepsin is not surprising. The maximal rates of enzymatic reaction v_{max} values achieved by free pepsin and heterogeneous pepsin nanobiocatalyst at maximum (saturating) substrate concentrations are also very similar ($v_{max} = 4.4 \pm 0.3$ and $3.5 \pm 0.2 \times 10^{-8}$ M s^{-1}, respectively). Furthermore, knowledge of the total enzyme concentration in the reaction allows calculation of the turnover number k_{cat} which was 6.1×10^{-3} s^{-1} and 4.5×10^{-3} s^{-1} for free and immobilized pepsin. This corresponds to a catalytic constant for the heterogeneous nanobiocatalyst of 78 % of free pepsin which may be regarded as equivalent to homogeneous catalysis. Moreover, the ratio k_{cat}/K_M, which is a characteristic parameter to describe the catalytic efficiency, is even higher for the immobilized pepsin. Consequently, it can be concluded that immobilization of pepsin to pegylated GNPs affords a nanobiocatalyst which can benefit from properties of heterogeneous catalysis (like easy removal) while still exhibiting kinetic properties like a homogeneous biocatalyst.

As mentioned above, nonlinearity was found in Lineweaver–Burke diagrams for both free and immobilized pepsin. To more adequately characterize enzyme kinetics a biphasic behavior was assumed and bilinear Lineweaver–Burke diagrams exploited for deriving characteristic parameters. A summary of the results is given in Table 1. It can be seen that at higher CYC concentration levels, the K_M values ($K_{M,2}$) increase by a factor of around 10, thereby indicating lower affinity of the substrate to both free and immobilized enzyme. At the same time, theoretical v_{max} values increase by a factor of ca. 4 at high substrate concentrations. Overall, catalytic efficiencies are by a factor of ca. 2 lower in the high concentration regime. In any case this nonlinear behavior may indicate feedback inhibition or competitive binding of peptidic products. Such feedback inhibition has been reported previously for other pepsin-catalyzed reactions as well [53].

Protein digestion with free and GNP-conjugated pepsin

To test the wider scope of the performance with a more complex sample, the functionalized heterogeneous nanobiocatalyst with immobilized pepsin was evaluated for its digestion capability of a protein mixture composed of bovine serum albumin (BSA), CYC, and myoglobin (MYG) in comparison to its free analogue. Again, for comparison the pepsin content was adjusted to be similar for both incubations with free and immobilized pepsin. The resultant protein digests sampled at 0, 4, 12, and 24 h were analyzed by UHPLC–ESI-QTOF-MS/MS measurements in data-dependent IDA scan mode, and the exported peak list was subjected to Mascot search to elucidate the quality of digestion in a time-dependent manner by monitoring scores of protein
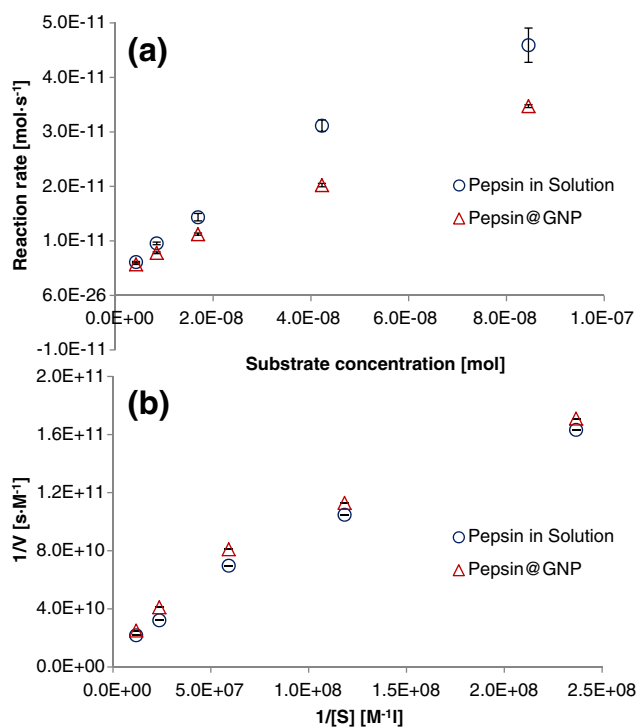

Fig. 3 Characterization of bioactivity of free and GNP-conjugated pepsin by Michaelis–Menten enzyme kinetics in nonlinear representation (**a**) and linearized form according to Lineweaver–Burk diagram (**b**)

Table 1 Kinetic data for digestion of cytochrome C as model protein by free pepsin and pepsin-conjugated GNPs

Parameter	Free pepsin	Pepsin@GNP	% (Pepsin@GNP vs free)
K_M [M]	$2.93 \pm 0.29 \times 10^{-5}$	$2.14 \pm 0.15 \times 10^{-5}$	73
V_{max} [M s^{-1}]	$4.44 \pm 0.30 \times 10^{-8}$	$3.47 \pm 0.22 \times 10^{-8}$	78
k_{cat} [s^{-1}]	$6.13 \pm 0.41 \times 10^{-3}$	$4.79 \pm 0.31 \times 10^{-3}$	78
k_{cat}/K_M [M^{-1} s^{-1}]	$2.09 \pm 0.12 \times 10^{2}$	$2.24 \pm 0.02 \times 10^{2}$	107
Biphasic model			
$K_{M, 1}$ [M]	$1.29 \pm 0.05 \times 10^{-5}$	$9.72 \pm 0.75 \times 10^{-6}$	75
$K_{M, 2}$ [M]	$1.15 \pm 0.17 \times 10^{-4}$	$9.72 \pm 0.89 \times 10^{-5}$	84
$V_{max, 1}$ [M s^{-1}]	$2.47 \pm 0.11 \times 10^{-8}$	$1.93 \pm 0.10 \times 10^{-8}$	78
$V_{max, 2}$ [M s^{-1}]	$1.13 \pm 0.17 \times 10^{-7}$	$8.32 \pm 0.96 \times 10^{-8}$	74
$k_{cat, 1}$ [s^{-1}]	$3.41 \pm 0.15 \times 10^{-3}$	$2.66 \pm 0.14 \times 10^{-3}$	78
$k_{cat, 2}$ [s^{-1}]	$1.56 \pm 0.24 \times 10^{-2}$	$1.15 \pm 0.13 \times 10^{-2}$	74
$k_{cat, 1}/K_{M, 1}$ [M^{-1} s^{-1}]	$2.64 \pm 0.02 \times 10^{2}$	$2.74 \pm 0.17 \times 10^{2}$	104
$k_{cat, 2}/K_{M, 2}$ [M^{-1} s^{-1}]	$1.35 \pm 0.01 \times 10^{2}$	$1.18 \pm 0.03 \times 10^{2}$	87

Subscript 1 refers to the low substrate concentration range and subscript 2 to the high concentration range

identification and corresponding sequence coverages. Figure 4a–d show chromatograms of the pepsin digests at different digestion times with pepsin–GNP conjugate, and Fig. 4e–h show the corresponding chromatograms obtained with free pepsin.

It can be seen in Fig. 4a that immediately after addition of pepsin–GNP conjugate to the protein mixture peptides appear in the chromatogram, while digested peptides are virtually absent in the free enzyme incubation (Fig. 4e) (see also Fig. 5a).
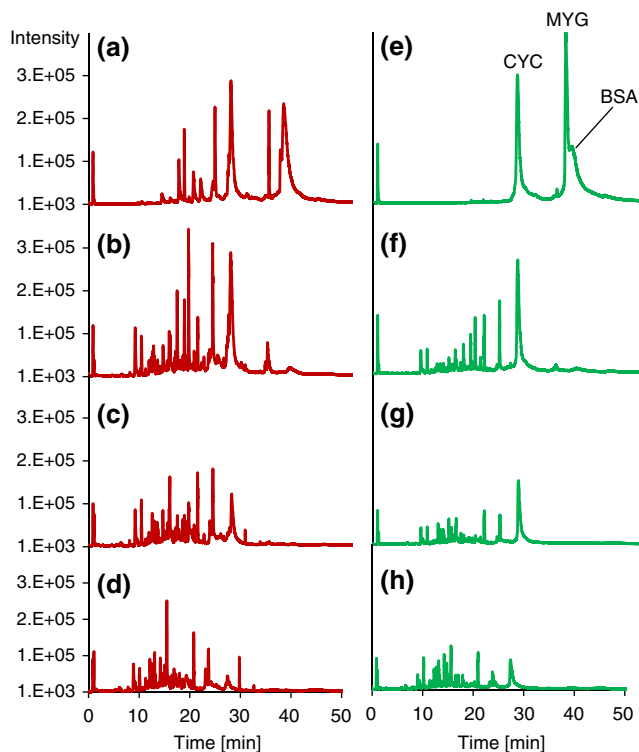
Fig. 4 UHPLC–ESI-QTOF-MS/MS measurements (TIC) of a protein mix digested with GNP-conjugated pepsin (**a**–**d**) and free pepsin (**e**–**h**) for 0 h (**a**, **e**), 4 h (**b**, **f**), 12 h (**c**, **g**), and 24 h (**d**, **h**)

This could indicate accelerated reaction with the immobilized pepsin, maybe owing to a higher local pepsin concentration on the enzyme–GNP conjugate as compared to that in free solution despite the approximately equal overall presence of pepsin in the system. Another explanation might be continued digestion during workup while spinning down the pepsin-modified GNPs. After 4 h, BSA and MYG are completely digested both with immobilized and free enzyme (Fig. 4b and f). A total of 39 and 42 peptides were detected with immobilized and free pepsin, respectively (Fig. 5b). The number of detected peptides was not significantly altered after 12 h (Figs. 4c, g and 5c), and only slightly increased after 24 h digestion (Figs. 4d, h and 5d). From the Venn diagrams [54] in Fig. 5 it is clearly evident that the heterogeneous pepsin bionanocatalyst shows a larger, or at least similar, number of digested peptides as free pepsin.

Mascot database search uses probability-based scoring to judge whether a result is significant and not a random event, and scores greater than 67 are deemed to be significant ($p < 0.05$) [55, 56]. Scores for the distinct digested samples are depicted in Fig. 6a. It can be seen that the Mascot search enables identification of the proteins with high scores (>100) in all digests after 4-h digestion time or more, not only for pepsin in solution but also the pepsin–GNP conjugate. In fact, the scores of the Mascot search were on average even higher for the immobilized pepsin. Moreover, sequence coverages greater than 20 % in the digests with 4-h and 12-h digestion times for the heterogeneous nanoparticle-based biocatalyst allows one to conclude that its bioactivity and catalytic performance is at least of equal quality as compared to pepsin in free solution which cannot be conveniently removed by brief spinning (Fig. 6).

A common problem of digestion using proteolytic enzymes in solution in sample preparation is autodigestion. This problem can be largely eliminated with immobilized enzymes. A dedicated search for signature peptides originating from pepsin in the protein digests was therefore carried out. However,
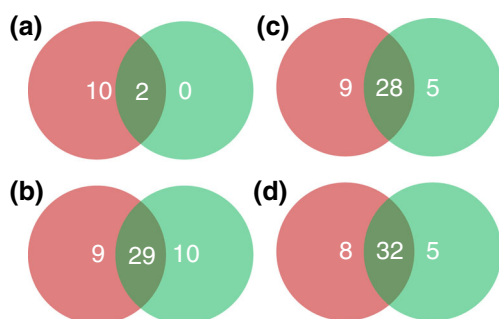
Fig. 5 Venn diagrams comparing numbers of peptides detected in samples digested with GNP-conjugated pepsin (*red circle*), free pepsin (*green circle*), and both (*intersection*) **a** immediately after admixing the enzyme to the incubation mixture, i.e., 0 h, **b** after 4 h, **c** after 12 h, and **d** after 24 h of digestion time

in the present case neither in solution nor in digests from immobilized pepsin could autodigestion be detected. However, this issue might get more important if reaction times need to be extended for proteins which are more resistant to digestion by pepsin.

On the other hand, no pepsin contamination was present in incubations with the immobilized pepsin nanobiocatalyst, whereas the soluble enzyme might represent an impurity and interference in the analyzed protein sample in conventional

reaction incubations. No pepsin bled from the GNP–pepsin conjugate in the course of digestion of CYC (see ESM Fig. S8).

Last but not least, the immobilized pepsin@GNP nanobiocatalyst can be reused. Recyclability was tested and it was found that the digestion performance (sequence coverage) was largely maintained at the same level for up to three cycles of reuse; however, after three reuse cycles its digestion performance started to decline (see ESM Fig. S7).

Comparative peptide mapping

To gain insight into the proteolytic specificity of the conjugated pepsin and to compare its performance with that of free pepsin, also in terms of missed cleavages and peptide homology, a closer look into the peptide fragments obtained by digestion of the protein mix (CYC, MYG, BSA) for 4 h with pepsin@GNP and free pepsin in solution was undertaken. Figures S9–S11 of the ESM show an overview of precursors generated by pepsin@GNP, overlays of extracted ion chromatograms (XICs) of protein-specific peptides for each CYC, MYG, and BSA, and an exemplary MS/MS spectrum of a model peptide. Figure 7 illustrates the protein sequences of CYC (Fig. 7a) and

Fig. 6 Mascot search results for protein identification in digests of a protein test mixture composed of BSA, CYC, and MYG obtained with GNP-PEG$_7$-pepsin and pepsin in solution which shows the score (**a**) and the sequence coverage (**b**) after 0, 4, 12, and 24 h digestion time (for corresponding chromatograms, see Fig. 4). The Mascot search using MS/MS data of the IDA scan mode was performed with the Swiss-Prot database

Fig. 7 Sequence of CYC (**a**) and MYG (**b**) with identified peptides found in digests with free pepsin in solution (*solid arrows*) and pepsin-modified GNPs with HS-PEG$_7$-COOH spacer (*dashed arrows*). Peptides were determined by Mascot search and compared with reference to the protein sequences
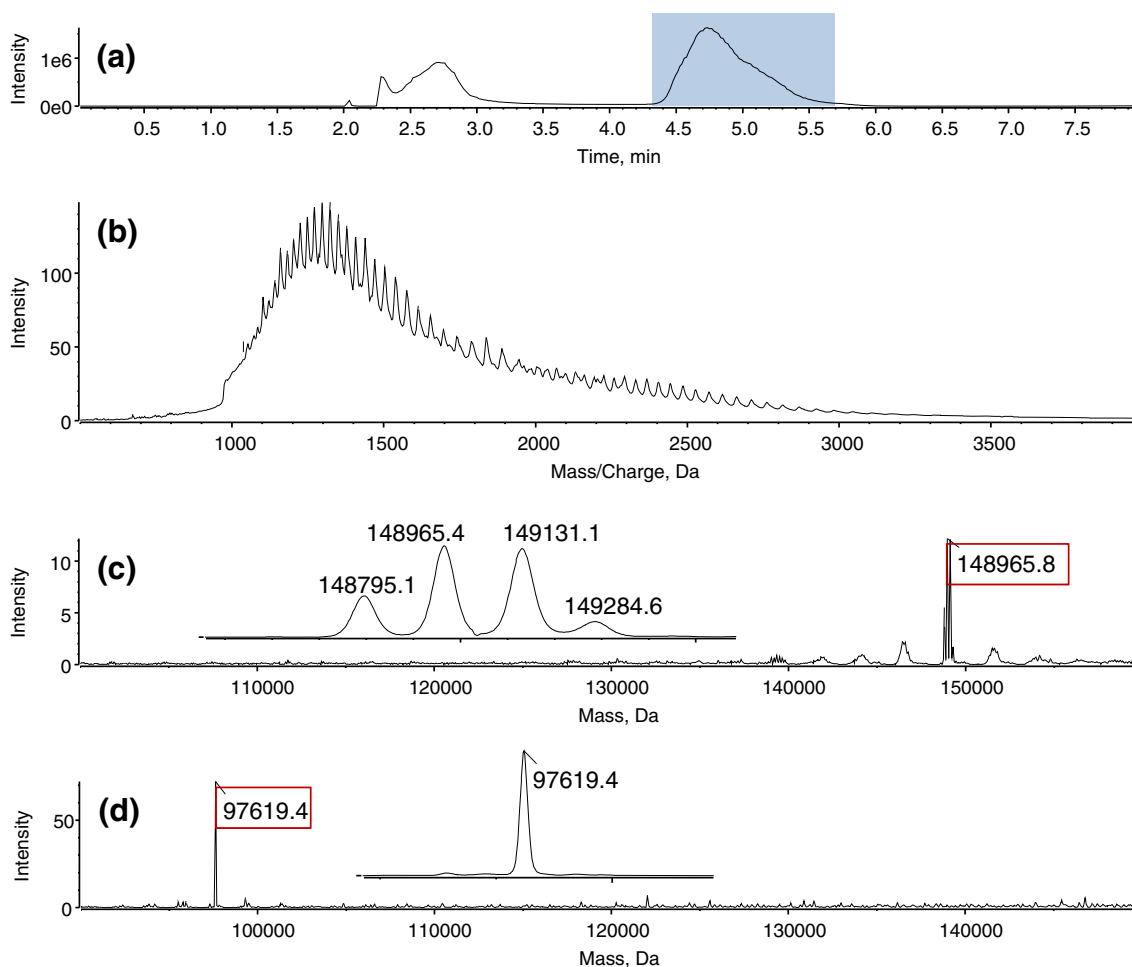
Fig. 8 Digestion of monoclonal antibody with pepsin@GNP. TIC (**a**) and TOF-MS spectrum (**b**) of intact anti-HSA as measured by online-SPE μESI-QTOF-MS. The TOF-MS spectrum was taken at the indicated area between t_R 4.5 and 5.5 min and shows the charge envelope of anti-HSA. The mass spectra **c** and **d** depict the intact protein mass of anti-HSA before (**c**) and after (**d**) digestion with pepsin@GNP

MYG (Fig. 7b), and the peptides identified in corresponding digests of free and immobilized pepsin are indicated by arrows covering the respective sequence (corresponding map for BSA can be found in ESM Fig. S10). Solid arrows show the fragments cut by pepsin in solution, dashed arrows the fragments cut by pepsin@GNP. A 100 % homology was found in the generated peptide fragments in the case of CYC, yielding a sequence coverage of 68 % for CYC both with free and gold-conjugated pepsin (Fig. 7a). Only two missed cleavages are found (F in position 11 and L in position 69). The former missed cleavage may be due to K in P3 while the latter could be a result of the negative effect on cleavage exerted by Pro at P3'. Moreover, the *N*-terminal tail was not identified by Mascot. To a large extent, good but slightly lower homology between heterogeneous and homogeneous catalysis was also found for MYG (Fig. 7b). While identified peptides were identical over large sequences at the *N*- and *C*-terminal tails, in the range between position 50 and 105 peptides and sequence coverage were different for the two approaches of digestion. Both failed to cleave at positions 3 (L), 44 (F), 47 (F), 73 (L), and 152 (F). Sequence coverage was

72 % and 62 % for MYG by free pepsin and pepsin@GNP, respectively. Overall, good homology in generated peptide fragments between immobilized and free pepsin was found also for BSA (see ESM Fig. S12).

Determination of antibody mass before and after digestion with pepsin@GNP

In addition to the characterization of the heterogeneous nanobiocatalysts by model proteins (BSA, CYC, and MYG), digestion of a real sample, a monoclonal antibody, in accordance with the intended purpose of use of the developed heterogeneous nanobiocatalyst should demonstrate its functionality. Pepsin digestion of IgG under nondenaturing conditions yields antibody fragments which are particularly useful for middle-up mass measurement for characterization of therapeutic antibodies by high-resolution MS. Since pepsin cleaves IgG at the C-terminal side of the inter-heavy-chain disulfides in the hinge region, it produces a bivalent antigen binding fragment, F(ab')2, with a molecular mass of about 100 kDa.

HR-MS using a QTOF equipped with a µESI sprayer and a C18 online trap column for desalting of the protein sample was used for characterization of the intact monoclonal antibody and the fragment obtained by digestion with pepsin@GNP. In the established assay, the protein elutes at around 5 min from the trap column, well separated from the salt plug at around 2.5 min (Fig. 8a). The MS spectrum of the intact monoclonal antibody is shown in Fig. 8b. It shows the characteristic charge envelope, as expected. The deconvoluted MS spectrum of the intact anti-HSA antibody revealed a mass of about 149 kDa with a number of isoforms (insert) (Fig. 8c). The same experiment with the antibody fragment obtained by digestion with pepsin@GNP resulted in a single peak with a mass of 97,619.4 Da for (Fab')$_2$. These results clearly document the functionality of the immobilized pepsin nanobiocatalyst and evidence its practical utility for therapeutic protein characterization.

Conclusions

Pepsin A was successfully immobilized on gold nanoparticles via a bifunctional PEG$_7$ linker resulting in efficient functionalized heterogeneous nanobiocatalysts with good colloidal stability and shelf life over at least a month. Characterization of the enzyme kinetics of the pepsin–GNP bioconjugate in comparison to free pepsin by digestion of the model protein cytochrome C clearly revealed that the enzyme fully retained its catalytic efficiency after immobilization on this nanoparticulate carrier. Digestions of protein mixtures followed by UHPLC–ESI-QTOF-MS/MS analysis with subsequent protein identification by Mascot database search documented the pertinent bioactivity of the pepsin–GNP bionanocatalyst and its utility for protein analysis and characterization, respectively. By measuring in intact protein mode it was possible to determine the (Fab')$_2$ fragment after digestion of anti-HSA with pepsin@GNP.

While the heterogeneous pepsin–GNP bionanocatalyst showed equivalent kinetic performance and catalytic efficiency, it offers significant additional benefits, thus combining favorable properties of homogeneous and heterogeneous catalysis. It can be pipetted like a solution owing to the stable colloidal suspension, which facilitates sample handling, automation, and miniaturization of sample preparation protocols, properties which are characteristic of homogeneous catalysis. Owing to the high density of the metallic gold nanoparticle core, the heterogeneous bionanocatalyst can be easily spinned down on a minispin in a few minutes at low rpm, equipment which is available in virtually any laboratory and does not involve complicated procedures. In the current enzyme–GNP bioconjugate, the enzyme is freely accessible on the surface of the nonporous gold particle and not deeply buried inside a pore channel such as in more common heterogeneous catalysts in which enzymes are immobilized on mesoporous microparticles, such as (cross-linked) agarose; therefore, diffusion limitations are thought to be less serious on enzyme–GNP bionanocatalysts. On the contrary, more common microparticulate heterogeneous catalysts may seriously suffer from diffusion limitations in intraparticulate pore spaces which may easily lead to substrate depletion as well as product accumulation at intraparticulate active enzyme centers. Overall, however, the removal of the functionalized heterogeneous nanobiocatalysts by centrifugation before the analytical determination is the striking advantage, both for single as well as coupled digestion or sample preparation protocols. This is of utmost importance in the intended primary application fields such as for therapeutic peptide and protein characterization as well as (sequence and stereoconfiguration) analysis of (lipo)peptides from natural pools where enzyme contaminations might easily interfere with the analysis and complicate the interpretation of results.

Compliance with ethical standards

Conflict of interest The authors declare no competing conflict of interest.

References

1. Tsiatsiani L, Heck AJR. Proteomics beyond trypsin. FEBS J. 2015;282(14):2612–26. doi:10.1111/febs.13287.

2. Beck A, Wagner-Rousset E, Ayoub D, Van Dorsselaer A, Sanglier-Cianferani S. Characterization of therapeutic antibodies and related products. Anal Chem. 2013;85(2):715–36. doi:10.1021/ac3032355.

3. Gerhardt H, Sievers-Engler A, Jahanshah G, Pataj Z, Ianni F, Gross H, et al. Methods for the comprehensive structural elucidation of constitution and stereochemistry of lipopeptides. J Chromatogr A. 2016;1428:280–91. doi:10.1016/j.chroma.2015.05.065.

4. Safdar M, Spross J, Janis J. Microscale immobilized enzyme reactors in proteomics: latest developments. J Chromatogr A. 2014;1324:1–10. doi:10.1016/j.chroma.2013.11.045.

5. Zhanga Y, Gea J, Zheng L. Enhanced activity of immobilized or chemically modified enzymes. ACS Catal. 2015;5(8):4503–13. doi:10.1021/acscatal.5b00996.

6. Liu J, Yang Q, Li C. Towards efficient chemical synthesis via engineering enzyme catalysis in biomimetic nanoreactors. Chem Commun. 2015;51(72):13731–9. doi:10.1039/C5CC04590H.

7. Misson M, Zhang H, Jin B. Nanobiocatalyst advancements and bioprocessing applications. J R Soc Interface. 2015;12(102):1–20. doi:10.1098/rsif.2014.0891.

8. Vaghari H, Jafarizadeh-Malmiri H, Mohammadlou M, Berenjian A, Anarjan N, Jafari N, et al. Application of magnetic nanoparticles in smart enzyme immobilization. Biotechnol Lett. 2015. doi:10.1007/s10529-015-1977-z.

9. Johnson BJ, Russ Algar W, Malanoski AP, Ancona MG, Medintz IL. Understanding enzymatic acceleration at nanoparticle

interfaces: approaches and challenges. Nano Today. 2014;9(1): 102–31. doi:10.1016/j.nantod.2014.02.005.

10. Ding S, Cargill AA, Medintz IL, Claussen JC. Increasing the activity of immobilized enzymes with nanoparticle conjugation. Curr Opin Biotechnol. 2015;34:242–50. doi:10.1016/j.copbio.2015.04.005.

11. Roco MC, Williams RS, Alivisatos P, editors. Nanotechnology research directions: IWGN workshop report; vision for nanotechnology R&D in the next decade. Washington DC: US National Science and Technology Council; 2000.

12. Hinterwirth H, Lindner W, Laemmerhofer M. Bioconjugation of trypsin onto gold nanoparticles: effect of surface chemistry on bioactivity. Anal Chim Acta. 2012;733:90–7. doi:10.1016/j.aca.2012.04.036.

13. Homaei A, Barkheh H, Sariri R, Stevanato R. Immobilized papain on gold nanorods as heterogeneous biocatalysts. Amino Acids. 2014. doi:10.1007/s00726-014-1724-0.

14. Bolivar JM, Eisl I, Nidetzky B. Advanced characterization of immobilized enzymes as heterogeneous biocatalysts. Catal Today. 2015. doi:10.1016/j.cattod.2015.05.004.

15. Polte J, Ahner TT, Delissen F, Sokolov S, Emmerling F, Thunemann AF, et al. Mechanism of gold nanoparticle formation in the classical citrate synthesis method derived from coupled in situ XANES and SAXS evaluation. J Am Chem Soc. 2010;132(4): 1296–301. doi:10.1021/ja906506j.

16. Turkevich J, Stevenson PC, Hillier J. The nucleation and growth processes in the synthesis of colloidal gold. Discuss Faraday Soc. 1951;11:55–75. doi:10.1039/df9511100055.

17. Frens G. Controlled nucleation for the regulation of the particle size in monodisperse gold suspensions. Nature Phys Sci. 1973;241:20–2.

18. Ojea-Jimenez I, Romero FM, Bastus NG, Puntes V. Small gold nanoparticles synthesized with sodium citrate and heavy water: insights into the reaction mechanism. J Phys Chem C. 2010;114(4):1800–4. doi:10.1021/jp9091305.

19. Kumar S, Gandhi KS, Kumar R. Modeling of formation of gold nanoparticles by citrate method. Ind Eng Chem Res. 2007;46(10): 3128–36. doi:10.1021/ie060672j.

20. Haiss W, Thanh NTK, Aveyard J, Fernig DG. Determination of size and concentration of gold nanoparticles from UV-Vis spectra. Anal Chem. 2007;79:4215–21. doi:10.1021/ac0702084.

21. Hendel T, Wuithschick M, Kettemann F, Birnbaum A, Rademann K, Polte J. In situ determination of colloidal gold concentrations with UV-Vis spectroscopy: limitations and perspectives. Anal Chem. 2014;86(22):11115–24. doi:10.1021/ac502053s.

22. Lin S-Y, Tsai Y-T, Chen C-C, Lin C-M, Chen C-H. Two-step functionalization of neutral and positively charged thiols onto citrate-stabilized Au nanoparticles. J Phys Chem B. 2004;108(7): 2134–9. doi:10.1021/jp036310w.

23. Hermanson GT, editor. Bioconjugate techniques. Amsterdam: Academic; 1995.

24. Cao Y, Wen L, Svec F, Tan T, Lv Y. Magnetic AuNP@Fe3O4 nanoparticles as reusable carriers for reversible enzyme immobilization. Chem Eng J. 2016;286:272–81. doi:10.1016/j.cej.2015.10.075.

25. Nidhin M, Ghosh D, Yadav H, Yadav N, Majumder S. Structural and functional aspects of trypsin-gold nanoparticle interactions: an experimental investigation. Mater Sci Eng B. 2015;202:46–53. doi:10.1016/j.mseb.2015.09.006.

26. Willner I, Baron R, Willner B. Integrated nanoparticle-biomolecule systems for biosensing and bioelectronics. Biosens Bioelectron. 2007;22(9-10):1841–52. doi:10.1016/j.bios.2006.09.018.

27. Ang SH, Yu CY, Ang GY, Chan YY, Yb A, Khor SM. A colloidal gold-based lateral flow immunoassay for direct determination of haemoglobin A1c in whole blood. Anal Methods. 2015;7(9): 3972–80. doi:10.1039/C5AY00518C.

28. Niu K, Zheng X, Huang C, Xu K, Zhi Y, Shen H, et al. A colloidal gold nanoparticle-based immunochromatographic test strip for rapid and convenient detection of Staphylococcus aureus. J Nanosci Nanotechnol. 2014;14(7):5151–6. doi:10.1166/jnn.2014.8703.

29. Qian X, Peng X-H, Ansari DO, Yin-Goen Q, Chen GZ, Shin DM, et al. In vivo tumor targeting and spectroscopic detection with surface-enhanced Raman nanoparticle tags. Nat Biotechnol. 2008;26(1):83–90. doi:10.1038/nbt1377.

30. Florkin M. Discovery of pepsin by Theodor Schwann. Rev Med Liege. 1957;12(5):139–44.

31. Northrop JH. Crystalline pepsin. Science. 1929;69(1796):580.

32. Schmelzer CEH, Schoeps R, Ulbrich-Hofmann R, Neubert RHH, Raith K. Mass spectrometric characterization of peptides derived by peptic cleavage of bovine β-casein. J Chromatogr A. 2004;1055(1-2):87–92. doi:10.1016/j.chroma.2004.09.003.

33. Goldstein J, Konigsberg W, Hill RJ. The structure of human hemoglobin. VII, The digestion of the β chain of human hemoglobin with pepsin. J Biol Chem. 1963;238:2028–33.

34. Huckel M, Wirth H-J, Hearn MTW. Porous zirconia: a new support material for enzyme immobilization. J Biochem Biophys Methods. 1996;31(3,4):165–79. doi:10.1016/0165-022X(95)00035-P.

35. Ahn J, Jung MC, Wyndham K, Yu YQ, Engen JR. Pepsin immobilized on high-strength hybrid particles for continuous flow online digestion at 10 000 psi. Anal Chem. 2012;84(16):7256–62. doi:10.1021/ac301749h.

36. Bandhu A, Sutradhar S, Mukherjee S, Greneche JM, Chakrabarti PK. Synthesis, characterization and magnetic property of maghemite (γ-Fe2O3) nanoparticles and their protective coating with pepsin for bio-functionalization. Mater Res Bull. 2015;70: 145–54. doi:10.1016/j.materresbull.2015.04.035.

37. Han W, Yamauchi M, Hasegawa U, Noda M, Fukui K, van der Vlies AJ, et al. Pepsin immobilization on an aldehyde-modified polymethacrylate monolith and its application for protein analysis. J Biosci Bioeng. 2015;119(5):505–10. doi:10.1016/j.jbiosc.2014.10.018.

38. Geiser L, Eeltink S, Svec F, Frechet JMJ. In-line system containing porous polymer monoliths for protein digestion with immobilized pepsin, peptide preconcentration and nano-liquid chromatography separation coupled to electrospray ionization mass spectroscopy. J Chromatogr A. 2008;1188(2):88–96. doi:10.1016/j.chroma.2008.02.075.

39. Long Y, Wood TD. Immobilized pepsin microreactor for rapid peptide mapping with nanoelectrospray ionization mass spectrometry. J Am Soc Mass Spectrom. 2015;26(1):194–7. doi:10.1007/s13361-014-1015-8.

40. Meridor D, Gedanken A. Enhanced activity of immobilized pepsin nanoparticles coated on solid substrates compared to free pepsin. Enzym Microb Technol. 2014;67:67–76. doi:10.1016/j.enzmictec.2014.09.004.

41. Stigter ECA, de Jong GJ, van Bennekom WP. Pepsin immobilized in dextran-modified fused-silica capillaries for on-line protein digestion and peptide mapping. Anal Chim Acta. 2008;619(2):231–8. doi:10.1016/j.aca.2008.04.060.

42. Innocent C, Bouhaine N, Djerourou A, Gherrou A, Seta P, Xu ZK. Preparation of catalytic materials by covalent immobilisation of pepsin on cotton fibers activated by tosyl chloride. Mater Sci Indian J. 2006;2(6):161–8.

43. Li J, Wang J, Gavalas VG, Atwood DA, Bachas LG. Alumina-pepsin hybrid nanoparticles with orientation-specific enzyme coupling. Nano Lett. 2003;3(1):55–8. doi:10.1021/nl025778s.

44. Sharma B, Mandani S, Sarma TK. Catalytic activity of various pepsin reduced Au nanostructures towards reduction of nitroarenes and resazurin. J Nanopart Res. 2015;17(1):1–12. doi:10.1007/s11051-014-2835-y.

45. Gole A, Dash C, Ramakrishnan V, Sainkar SR, Mandale AB, Rao M, et al. Pepsin-gold colloid conjugates: preparation, characterization, and enzymatic activity. Langmuir. 2001;17(5):1674–9. doi:10.1021/la001164w.

46. Haller E, Lindner W, Laemmerhofer M. Gold nanoparticle-antibody conjugates for specific extraction and subsequent analysis by liquid chromatography-tandem mass spectrometry of malondialdehyde-modified low density lipoprotein as biomarker for cardiovascular risk. Anal Chim Acta. 2015;857:53–63. doi:10.1016/j.aca.2014.12.024.

47. Hinterwirth H, Wiedmer SK, Moilanen M, Lehner A, Allmaier G, Waitz T, et al. Comparative method evaluation for size and size-distribution analysis of gold nanoparticles. J Sep Sci. 2013;36(17):2952–61. doi:10.1002/jssc.201300460.

48. Lineweaver H, Burk D. Determination of enzyme dissociation constants. J Am Chem Soc. 1934;56:658–66. doi:10.1021/ja01318a036.

49. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951;193(1):265–75.

50. Hinterwirth H, Kappel S, Waitz T, Prohaska T, Lindner W, Laemmerhofer M. Quantifying thiol ligand density of self-assembled monolayers on gold nanoparticles by inductively coupled plasma-mass spectrometry. ACS Nano. 2013;7:1129–36. doi:10.1021/nn306024a.

51. Mie G. Contributions to the optics of turbid media, especially colloidal metal solutions. Ann Phys. 1908;25:377–445.

52. Malola S, Lehtovaara L, Enkovaara J, Häkkinen H. Birth of the localized surface plasmon resonance in monolayer-protected gold nanoclusters. ACS Nano. 2013;7(11):10263–70. doi:10.1021/nn4046634.

53. Greenwell P, Knowles JR, Sharp H. Inhibition of pepsin-catalyzed reactions by products and product analogs. Kinetic evidence for ordered release of products. Biochem J. 1969;113(2):363–8. doi:10.1042/bj1130363.

54. Oliveros JC (2007) Venny. An interactive tool for comparing lists with Venn's diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html.

55. Koenig T, Menze BH, Kirchner M, Monigatti F, Parker KC, Patterson T, et al. Robust prediction of the MASCOT score for an improved quality assessment in mass spectrometric proteomics. J Proteome Res. 2008;7:3708–17. doi:10.1021/pr700859x.

56. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999;20:3551–67. doi:10.1002/(sici)1522-2683(19991201)20:18<3551::aid-elps3551>3.0.co;2-2.

Markus Höldrich is a PhD student in the Pharmaceutical (Bio-)Analysis group at the Institute of Pharmaceutical Sciences of the University of Tübingen, Germany. His research focuses on the synthesis and characterization of nanobiocatalysts mainly used for sample preparation in protein and antibody analysis.

Adrian Sievers-Engler is a PhD student in the Pharmaceutical (Bio-)Analysis group at the Pharmaceutical Sciences Institute of the University of Tübingen, Germany. He graduated in 2009 in Biochemistry at the University of Greifswald, Germany. His research interests include high-resolution mass spectrometry and separation techniques in bioanalysis and pharmaceutical bioanalysis.

Michael Lämmerhofer is Professor for Pharmaceutical (Bio-)Analysis at the Institute of Pharmaceutical Sciences of the University of Tübingen, Germany. He received his habilitation in Analytical Chemistry in 2002 from the University of Vienna where he was Associate Professor at the Department of Analytical Chemistry until 2011. His research interests include the development of functional separation materials and nanoparticles, lipidomics and metabolomics, pharmaceutical and biopharmaceutical analysis.

Gold nanoparticle-conjugated pepsin for efficient solution-like heterogeneous biocatalysis

in analytical sample preparation protocols

Markus Höldrich, Adrian Sievers-Engler, Michael Lämmerhofer

Aggregation-induced shift of SPR band

Aggregation in colloidal preparations of (functionalized) GNPs is detrimental for their functionality. For instance, aggregation of pepsin-conjugated GNPs may easily lead to loss of enzymatic activity because of precipitation of aggregates and preclusion of access of substrates to the active site of the immobilized enzyme. Thus, during synthesis and functionalization it is of importance to maintain stable colloidal suspension and monitor colloidal stability. A convenient and simple method to detect colloidal stability and aggregation, respectively, is monitoring the surface Plasmon resonance (SPR) band in the range between 400 and 800 nm of Vis spectra of the nanoparticle suspensions resulting after synthesis and surface modification, respectively. Upon aggregation, the intensity of the original SPR absorbance band at around 520 nm is reduced and a new red-shifted second absorbance band appears at higher wavelength (depending on size of aggregates and original particles e.g. 580 nm) indicating onset of aggregation. The conduction electrons near each particle surface become delocalized and are shared amongst neighboring particles in aggregates causing the surface Plasmon resonance shifting to lower energies which becomes evident as red-shifted SPR band. Fig. S1 shows an example of aggregated GNPs versus stable colloidal suspension. Eventually, aggregation can be observed even visually. The red nanoparticle solution turns to purple or dark, and finally a clear uncolored solution can be observed with black precipitate. Once aggregated, in nearly all cases it is difficult or even impossible to redisperse the aggregates into individual particles forming stable colloidal suspensions. During the current synthesis protocol, it was observed that activation of carboxy-terminated pegylated GNPs by OSu (*N*-hydroxysuccinimide) easily aggregate, but aggregates disappear upon protein binding. Nevertheless, a direct coupling without OSu activation was employed for this reason herein.
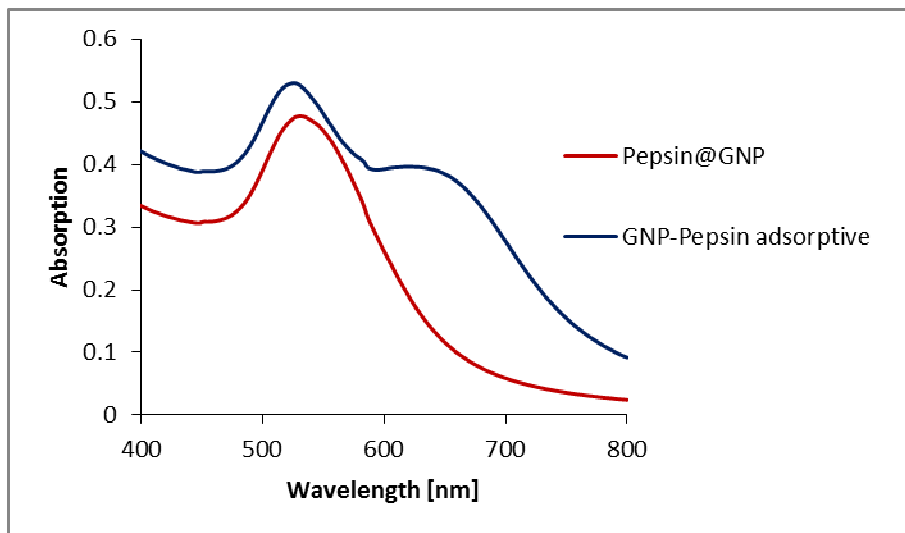
Fig. S1: Vis-spectra of Pepsin@GNP (pepsin-modified GNPs with PEG7 spacer) and adsorptively bound pepsin on GNPs. Because of a high colloidal stability the peak of

Pepsin@GNP shows a good shape. The undefined peak shape of adsorptively bound pepsin is due to particle aggregation.

Stability study of all nanoparticle stages for 35 days

One synthesis batch of pepsin-modified GNPs was used to study the long-term colloidal stability. Thus, an aliquot of each citrate-capped GNPs, pegylated GNPs (GNP-PEG7-COOH) and pepsin-modified GNPs with PEG7 spacer were stored in the refrigerator at 4°C in the dark for 35 days. Changes of the nanoparticles, in particular aggregation, were then monitored by Vis spectroscopy measuring Vis spectra between 350 and 800 nm in the first 9 days daily and then once a week and every second week, respectively.
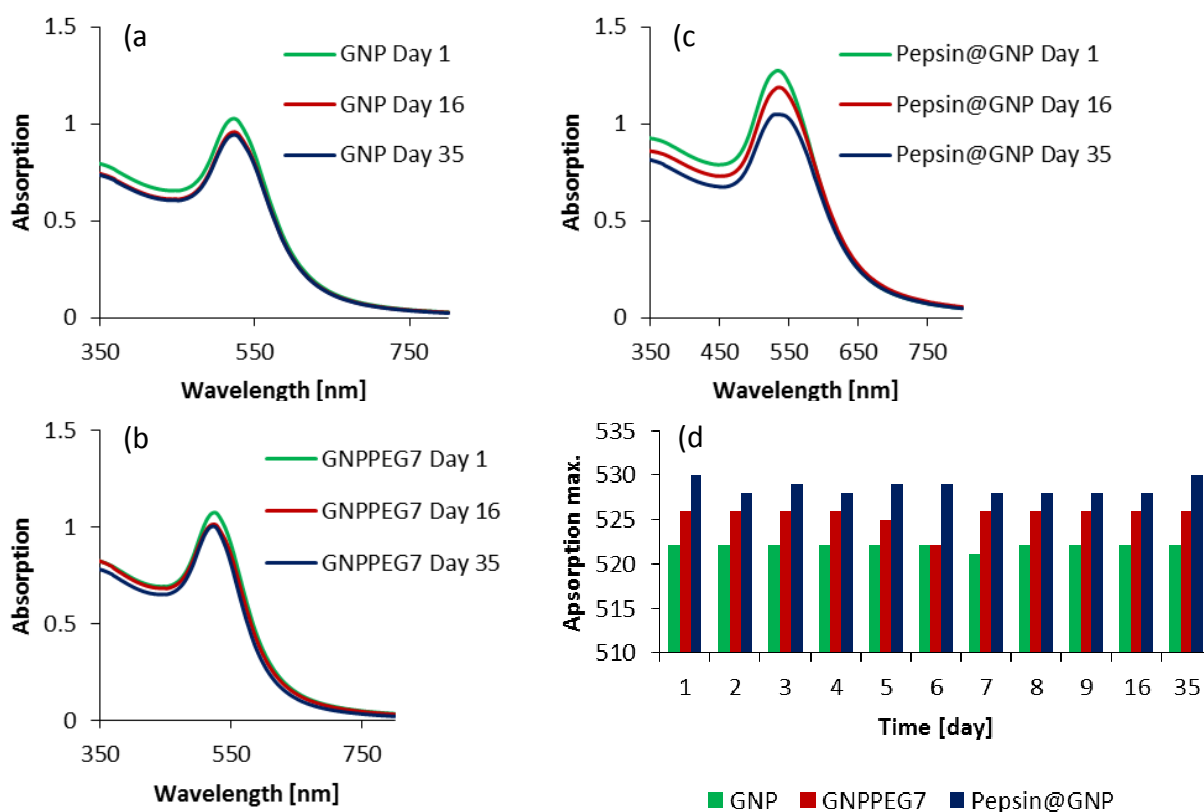
Fig. S2: Colloidal stability and shelf life of functionalized nanoparticles. (a) Citrate-stabilized GNPs, (b) GNPs coated with bifunctional PEG spacer having carboxy-terminated surface and (c) immobilized pepsin-GNP conjugate. (d) Absorption maxima of the SPR band of all nanoparticle stages measured over 35 days. All samples were diluted 1:5 in ddH$_2$O.

The results are illustrated in Fig. S2. As can be seen, essentially no changes in Vis spectra can be found for GNPs (Fig. S1a) and GNPs with PEG-spacer (Fig. S1b) over the entire period. For pepsin-modified GNPs wavelength shifts in absorbance maxima are virtually absent as well and

colloidal stability over 35 days is confirmed. However, a slight decline in absorbance of the SPR band between day 16 and 35 may indicate that some nanoparticles have precipitated leading to a slightly reduced concentration in the heterogeneous enzyme suspension. Such slightly lower concentration might, however, be easily and conveniently compensated for by pipetting slightly larger volumes to reaction mixtures in order to keep the total enzyme concentration in reaction incubations constant. This would guarantee satisfactory batch-to-batch reproducibility of enzymatic reactions when working with one pepsin-GNP conjugate batch over extended period.

Chromatographic assay for the determination of CYC in enzyme kinetics study
An accurate and precise validated HPLC-UV method was set up for the reliable quantitative analysis of undigested CYC in reaction batches of the enzyme kinetics study. For this chromatographic method a polystyrene-co-divinylbenzene monolith column with macropore diameter of 1 µm was selected. This type of columns has been shown to be favorable for fast protein analysis due to convective mass transfer. Due to a small mass transfer resistance term the column can be operated at high flow rates which allow fast analysis with high sample throughput. The mobile phase conditions and gradient profile were adjusted such that the protein was well retained while the digested peptides eluted at the front close to t_0, thereby avoiding interferences with CYC quantification. A representative chromatogram is given in Fig. S3.
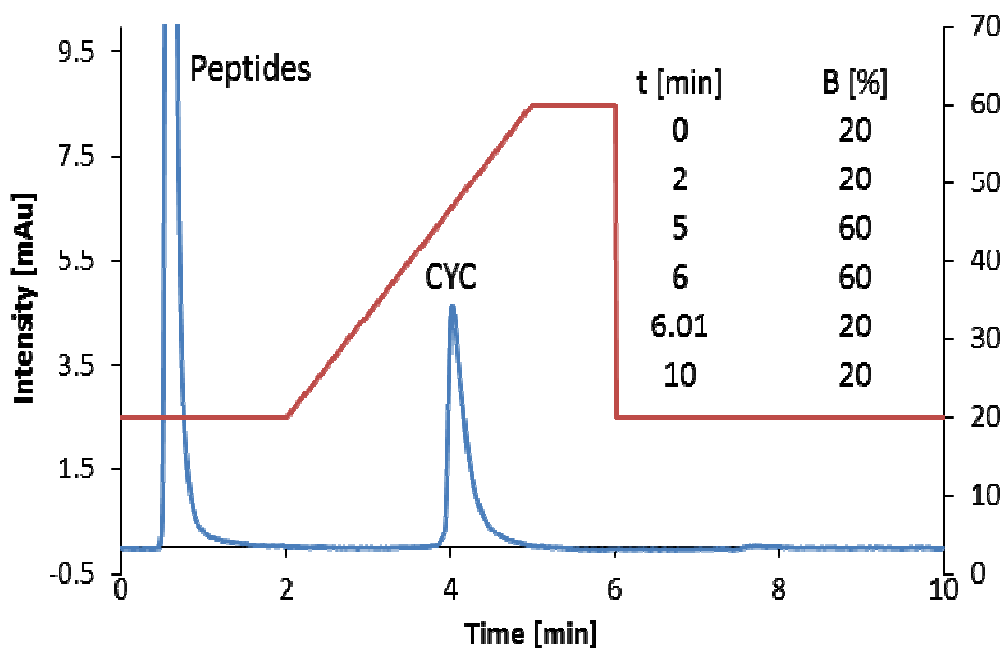
Fig. S3: Representative chromatogram of HPLC determination of undigested CYC in reaction batches from enzyme kinetics study. Experimental conditions: Column, Proswift[TM] RP-1S (4.6 x 50 mm, Dionex/ThermoFisher Scientific). Mobile phase, A: 50 mM ammonium formate pH 3.5 + 5 % ACN; B: ACN + 5 % 50 mM ammonium formate pH 3.5; gradient profile see insert; flow rate, 1 mL min[-1]; column temperature, 40 °C.

The method was validated according to ICH guidelines for method performance characterization in terms of detection limit (DL), quantitation limit (QL), intra-assay and inter-day precision as well as accuracy, and assay specificity.

Thus a calibration function was established by diluting a CYC stock solution of 2 mg/mL in 20 mM sodium acetate buffer pH 4.5 in the range of 0.05, 0.1, 0.2, 0.5, 1 and 2 mg/mL and analyzing the solutions by above HPLC-UV method. The calibration function is shown in Fig. S4. It can be seen that good linearity in the given concentration range was obtained with a determination coefficient R^2 of 0.9993. DL (S/N=3:1) and QL (S/N=10:1) were determined as 0.02 mg/ml and 0.08 mg/ml, respectively.

Calibration curve with y = 487.52x - 12.572, R^2 = 0.9993
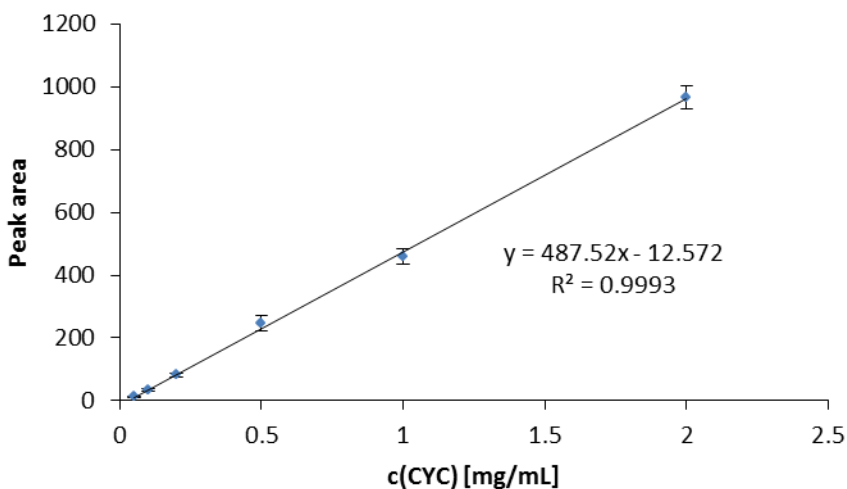
Fig. S4: Calibration function for CYC analyzed in the course of the enzyme kinetics study.

Table S1: Validation parameters for CYC

c(CYC)	0.05 mg/mL	0.1 mg/mL	1 mg/mL	2 mg/mL
accuracy	130.00%	114.67%	102.40%	104.87%
intraday precision	1.15%	0.79%	0.48%	0.31%
interday precision	1.21%	1.08%	0.60%	0.63%

For validating the assay performance intra-assay and inter-day precision and accuracy were validated across the relevant concentration range at three concentration levels (close to LOQ, intermediate and high concentration range) of quality control (QC) samples in three replicates. The results are given as %RSD for precision testing and as % recovery for accuracy testing. As can be seen from Table S1, RSD values < 1% in all cases as well as recoveries close to 100%

indicate excellent assay precision and accuracy. The assay was therefore suitable for measuring CYC concentrations in samples from enzyme kinetics study accurately and reliably.

Determination of surface coverage of GNPs with pepsin by Lowry Assay
For determination of surface coverage of GNPs with pepsin, common protein quantification assays were envisaged. Lowry assay was finally selected. In the Lowry assay, peptide bonds of the analyzed protein(s) are first complexed with Cu(II) ions in alkaline solution according to Biuret reaction. In a second step, Cu(II) is reduced to Cu(I) which in turn reduces the yellow Folin-Ciocalteu reagens (phosphomolybdate/phosphotungstic acid). The end product has a blue color. The protein quantity in the sample can be determined by absorbance measurements at 650 or 750 nm against a standard protein curve.
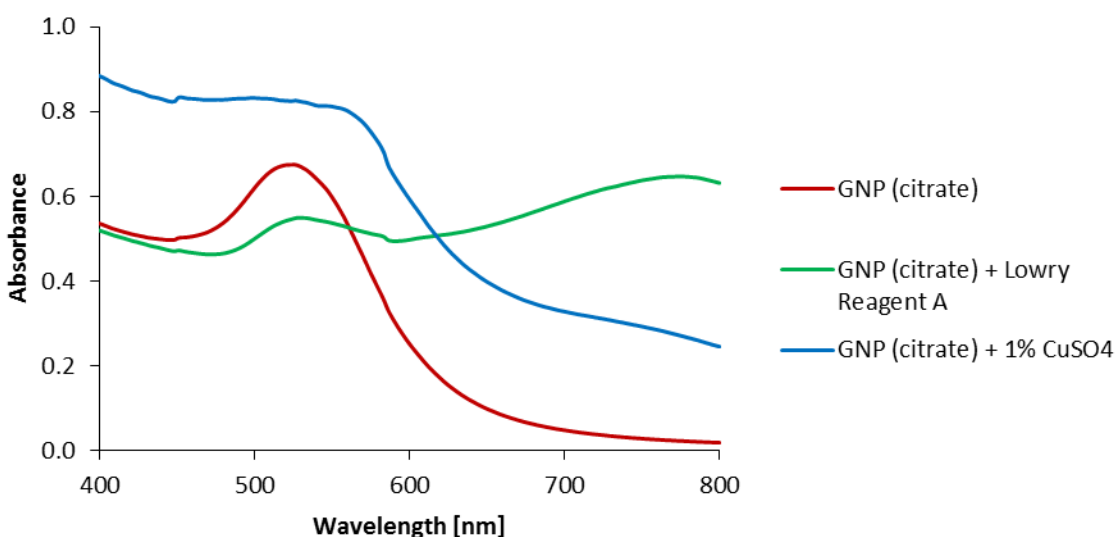
Fig. S5: The behavior of citrate stabilized GNPs in the presence of Lowry reagent A and 1 % $CuSO_4$.

Initially, on-bead protein quantitation using Lowry assay was intended. However, the citrate stabilized GNPs aggregated in the presence of Lowry reagent A (2.45 mL 4% Na_2CO_3 anhydrous in ddH_2O, 2.45 mL 0.1 M NaOH, 0.1 mL 2% Na-K-tartrate in ddH_2O and 0.05 mL $CuSO_4 \cdot 5H_2O$ in ddH_2O). In an additional experiment was discovered that the $CuSO_4$ has a considerable negative influence on the colloidal stability of the nanoparticles (Fig. S5). For this reason, the protein concentration of non-immobilized pepsin in the supernatant was finally determined and the immobilized protein calculated by mass balance. The total pepsin concentration in the reaction mixture was 1.2 mg/ml. Thus, a calibration function between 0.05 and 1.0 mg/ml pepsin in 20 mM MES buffer pH 4.5 (the reaction medium) was established (n=3) (Fig. S6). As can be seen, good linearity with a determination coefficient of $R^2 = 0.9947$ was obtained. Precision of the Lowry assay was good with average RSD values of 7.3 % over the relevant concentration range.

The mean pepsin content of three measured reaction mixtures in the supernatant after the immobilization by Lowry assay was determined to be 0.95 ± 0.11 mg/ml. Accordingly, the surface coverage of GNPs with pepsin was 0.25 ± 0.03 mg ml^{-1} of pepsin-GNP conjugate solution.
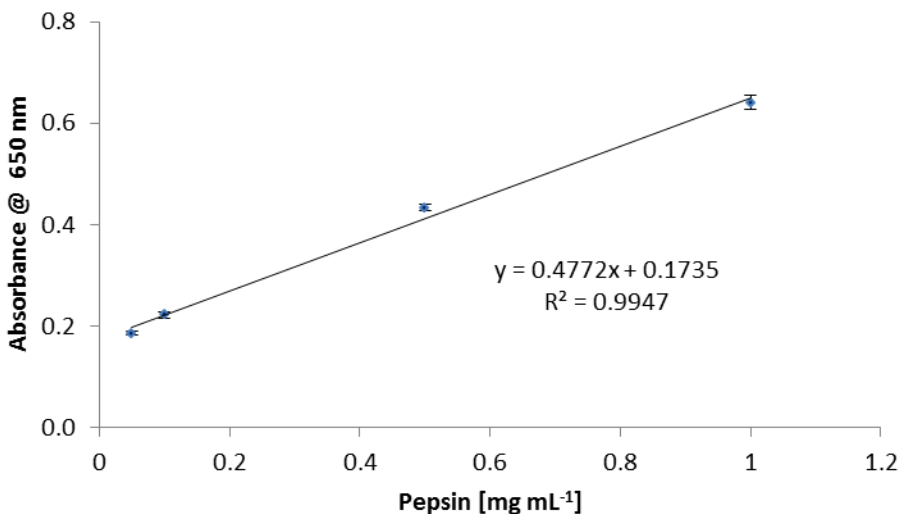
Fig. S6: Calibration function pepsin in 20 mM MES buffer pH 4.5 for Lowry assay.

Recyclability of Pepsin@GNP

CYC (1.2 mg mL^{-1} stock solution) was utilized as model protein to test the digestion efficiency of conjugated pepsin in the course of a re-usability study. For this purpose, 0.2 mL of the protein sample was transferred to 0.8 mL of the nanoparticle suspension (Pepsin@GNP in 1 mL 20 mM sodium acetate pH 4.5). The digestion was performed at 37 °C for 4 hours digestion time.

The reaction with the nanoparticle incubation was stopped by adding 0.2 mL of 0.1 M NaOH to reach neutral pH value. Pepsin@GNP was separated from the sample by centrifugation (12.000 rpm for 10 minutes) resulting in a clear supernatant for UHPLC-ESI-QTof-MS/MS analysis.

Five aliquots of aforementioned CYC sample were digested using iteratively the same Pepsin@GNP bioconjugate which was recovered by spinning and washed before re-use. Two washing steps in between were performed with 20 mM sodium acetate pH 4.5 (centrifugation at 12.000 rpm for 10 minutes).
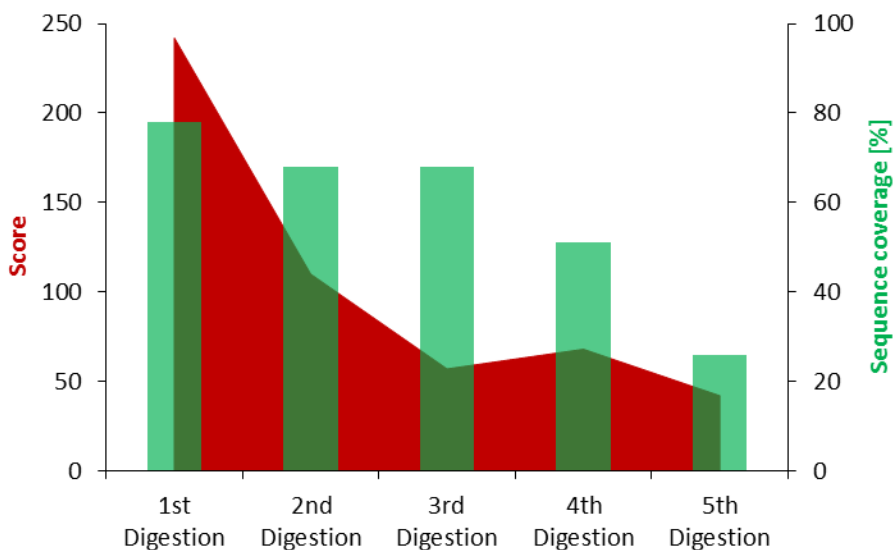
Fig. S7: Recyclability testing. Score and sequence coverage of CYC sample digested by the same Pepsin@GNP material after its re-cycling and washing.

As shown in Fig. S7 it was possible to identify CYC by Mascot search. The score value dropped in the first two cycles (from 242 to 110) but remained more or less constant thereafter (sore of around 50 in the 3rd to 5th re-use). Sequence coverage, on the other hand, remained essentially constant in the first three cycles of re-use (~70 %), dropped to 51 % in the 4th cycle and to 26 in the 5th cycle of re-use. Thus, recyclability of Pepsin@GNP could be demonstrated by this experiment, but the digestion performance was starting to decrease after three cycles of re-use.

Stability of the immobilized pepsin

In this study digestion was performed with immobilized pepsin and pepsin in free solution. For the digestion protocol with free pepsin 0.2 mg ml^{-1} pepsin was used (0.25 mg mL^{-1} final concentration in 1.2 mL total volume) and mixed with 0.8 mL pepsin in digestion buffer. The digestion was performed in 20 mM sodium acetate buffer pH 4.5 at 37 °C for 4 hours digestion time and stopped by adding 0.2 mL of 0.1 M NaOH. The digestion with immobilized pepsin was performed as described in the previous subchapter.

99 µL aliquots of the CYC samples digested by Pepsin@GNP and pepsin in solution were taken and reacted with 1 µL trypsin (1 mg mL^{-1} stock solution in 20 mM Tris-HCl pH 7.5) at 37 °C overnight, stopped by heat shock (95°C for 10 minutes) and analyzed by UHPLC-ESI-QTof-MS/MS.

Protein identification was performed by Mascot search to identify (leached) pepsin.
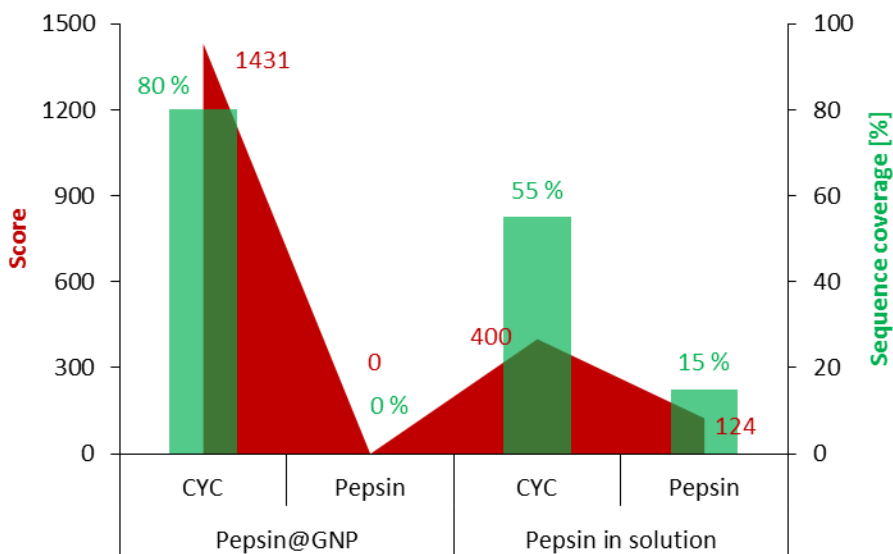
Fig. S8: Score and sequence coverage of the CYC samples digested by trypsin for the determination of (leached) pepsin.

After the coupled digestion protocol of CYC (first digestion by Pepsin@GNP/pepsin in solution, second digestion by trypsin) it was possible to identify CYC by Mascot search for both samples (Fig. S8). After the tryptic digestion pepsin could also be identified for the sample with pepsin in solution but it was not possible to detect pepsin in the Pepsin@GNP sample. That means pepsin is immobilized on the gold nanoparticle and will not bleed off from the particles within the experiments and/or the washing procedure.

Digestion of protein mix with GNP-conjugated pepsin
Fig. S9a shows the TIC of a digest obtained from the protein mix (CYC, MYG, BSA) with pepsin-conjugated GNPs after 4h, and the 2D-plot in Fig. S9b reveals an overview of the detected peptide precursor ions in dependence of the retention time. The numerous peptide peaks clearly confirm that the immobilized enzyme maintains its bioactivity. Exemplary an XIC of the peptide fragment VVSTQTALA, stemming from BSA, is shown in Fig. S9c and the +TOF MS-spectrum in Fig. S9d. Single and double charged peptide are found in the MS spectrum.

The peak list from IDA scan was loaded into Mascot and database search performed with pepsin as enzyme (other parameters are specified in Experimental part of main document). A number of peptides were identified by Mascot from the peak list. From these signature peptides, the corresponding XICs were extracted for each protein and overlaid (Fig. S10) to document the success of digestion with the new heterogeneous nanobiocatalyst. Fig. S9 illustrates exemplarily the MS/MS spectrum of a characteristic peptide fragment of BSA. As can be seen there is good coverage of the sequence due to b-ions.
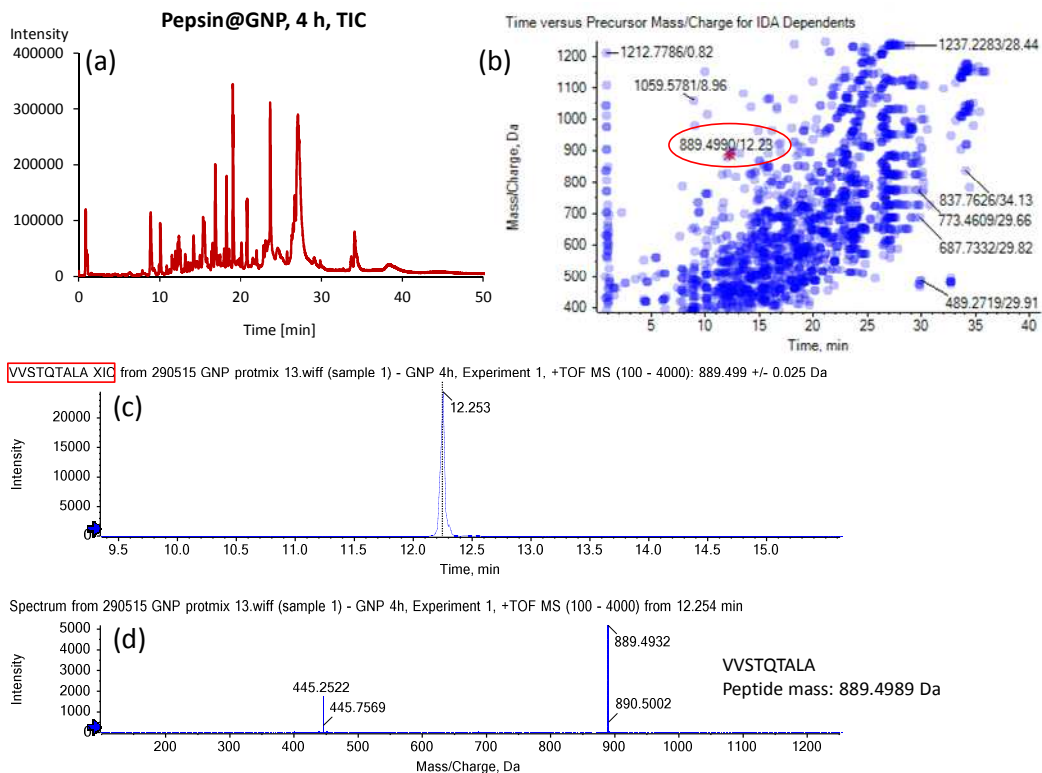
Fig. S9: Exemplary chromatogram of a peptide map obtained by digestion of the protein mix (CYC, MYG, BSA) with Pepsin@GNP conjugate and 4h digestion time (a), detected precursor ions in dependence of retention time (b), exemplary extracted ion chromatogram (XIC) of the peptide VVSTQTALA (c) and corresponding MS spectrum (d).

Fig. S10: Overlay of XICs of specific peptides of (a) CYC, (b) MYG, and (c) BSA identified by Mascot search. Digestion with Pepsin@GNP for 4h.

Spectrum from 290515 GNP protmix 13.wiff (sample 1) - GNP 4h, Experiment 2, +TOF MS^2 (100 - 4000) from 12.259 min
Precursor: 445.3 Da, CE: 25.0
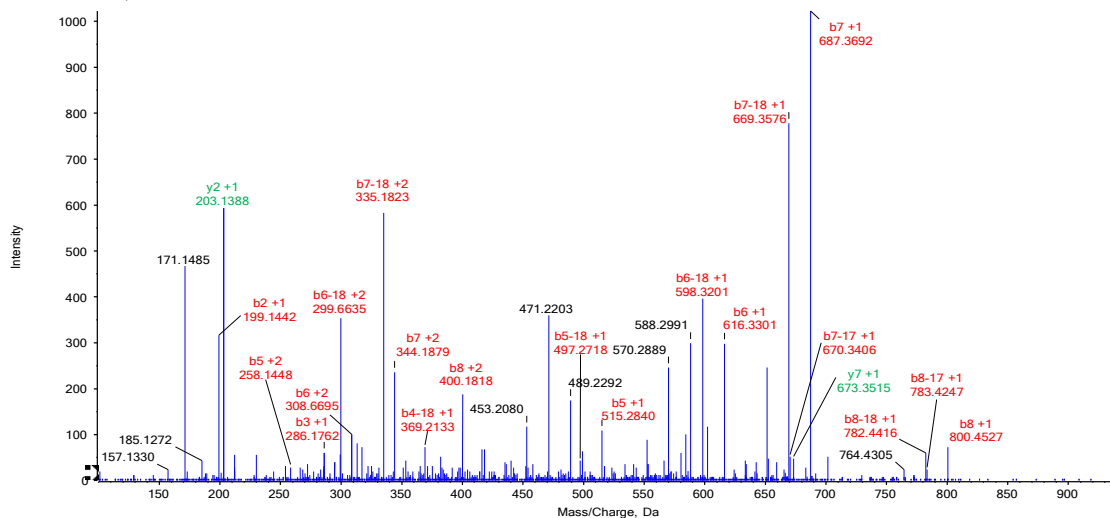
(b)

Symbol	Res. Mass	# (N)	b	b-17	b-18	y	y-17	y-18	# (C)
V	99.06841	1	100.0757	83.04914	82.06513	889.4989	872.4724	871.4884	9
V	99.06841	2	199.1441	182.1176	181.1335	790.4305	773.404	772.4199	8
S	87.03203	3	286.1761	269.1496	268.1656	691.3621	674.3355	673.3515	7
T	101.0477	4	387.2238	370.1973	369.2132	604.3301	587.3035	586.3195	6
Q	128.0586	5	515.2824	498.2558	497.2718	503.2824	486.2558	485.2718	5
T	101.0477	6	616.3301	599.3035	598.3195	375.2238	358.1973	357.2132	4
A	71.03711	7	687.3672	670.3406	669.3566	274.1761	257.1496	256.1656	3
L	113.0841	8	800.4512	783.4247	782.4407	203.139	186.1125	185.1285	2
A	71.03711	9	871.4884	854.4618	853.4778	90.05496	73.02841	72.04439	1

Fig. S11: Model MS/MS spectrum of peptide VVSTQTALA**.** Determined b-/y-ions with PeakView of the peptide VVSTQTALA after digestion of the protein mix with Pepsin@GNP for 4 h (mass spectra shown in Fig. 8d of main document). Precursor (+1): 889.4989 Da /(+2): 445.2531 Da.

Peptide mapping for assessment of digestion performance and proteolytic specificity

ExPasy Bioinformatics Resource Portal provides information on the proteolytic specificity of pepsin (http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html). It is stated that "Pepsin preferentially cleaves at Phe, Tyr, Trp and Leu in position P1 or P1'(Keil, 1992). Negative effects on cleavage are excerted by Arg, Lys and His in position P3 and Arg in position P1. Pro has favourable effects when being located in position P4 and position P3, but unfavourable ones when found in positions P2 to P3'. Cleavage is more specific at pH 1.3. Then pepsin preferentially cleaves at Phe and Leu in position P1 with negligible cleavage for all other amino acids in this position. This specificity is lost at pH ≥ 2"

Table S2: Proteolytic specificities of pepsin according to PeptideCutter tool.

Enzyme name	P4	P3	P2	P1	P1'	P2'
Pepsin (pH1.3)	-	not H,K, or R	not P	not R	F or L	not P
	-	not H,K, or R	not P	F or L	-	not P
Pepsin (pH>2)	-	not H,K or R	not P	not R	F,L,W or Y	not P
	-	not H,K or R	not P	F,L,W or Y	-	not P

Table reprinted from PeptideCutter publicly available under ExPasy Bioinformatics Resource Portal "http://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html"

Fig. 7 of the main document showed already the peptide fragments generated by pepsin in free solution and the Pepsin@GNP conjugate for cytochrome C and myoglobin. Good sequence homology was observed when digestion in free solution and with heterogeneous nanobiocatalyst were compared. Corresponding results for BSA are illustrated in Fig. S10. The majority of peptide fragments generated by pepsin are found in digests of homogeneous and heterogeneous catalysis, while a few were only detected in one of the two approaches. Overall good homology can be found for this larger protein as well.

ALBU_BOVIN:
Length: 582
IEP: 5.60

```
MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIA        50

FSQYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCK       100

VASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEF       150

KADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGAC       200

LLPKIETMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKAE       250

FVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKE       300

CCDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFL       350

GSFLYEYSRRHPEYAVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKL       400

KHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQVSTPTLVEVS       450

RSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCC       500

TESLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQT       550

ALVELLKHKPKATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVV       600

STQTALA                          Pepsin in solution       607
```

Pepsin@GNP

Fig. S12: Sequence of BSA with identified peptides found in digests with free pepsin in solution (solid arrows) and Pepsin@GNP (dashed arrows). Peptides were determined by Mascot search and compared with reference to the protein sequences.

14

Venn diagrams (Fig. S12) for identified peptides in protein mix (created with Venny 2.1; http://bioinfogp.cnb.csic.es/tools/venny/index.html)

In the control sample (0 hours) of the protein mixture (BSA, CYC and MYG) no peptides for BSA could be identified, neither for pepsin in solution nor for Pepsin@GNP. The number of identified peptides increase with increasing digestion times and were largely similar for homogeneous and heterogeneous catalysis.
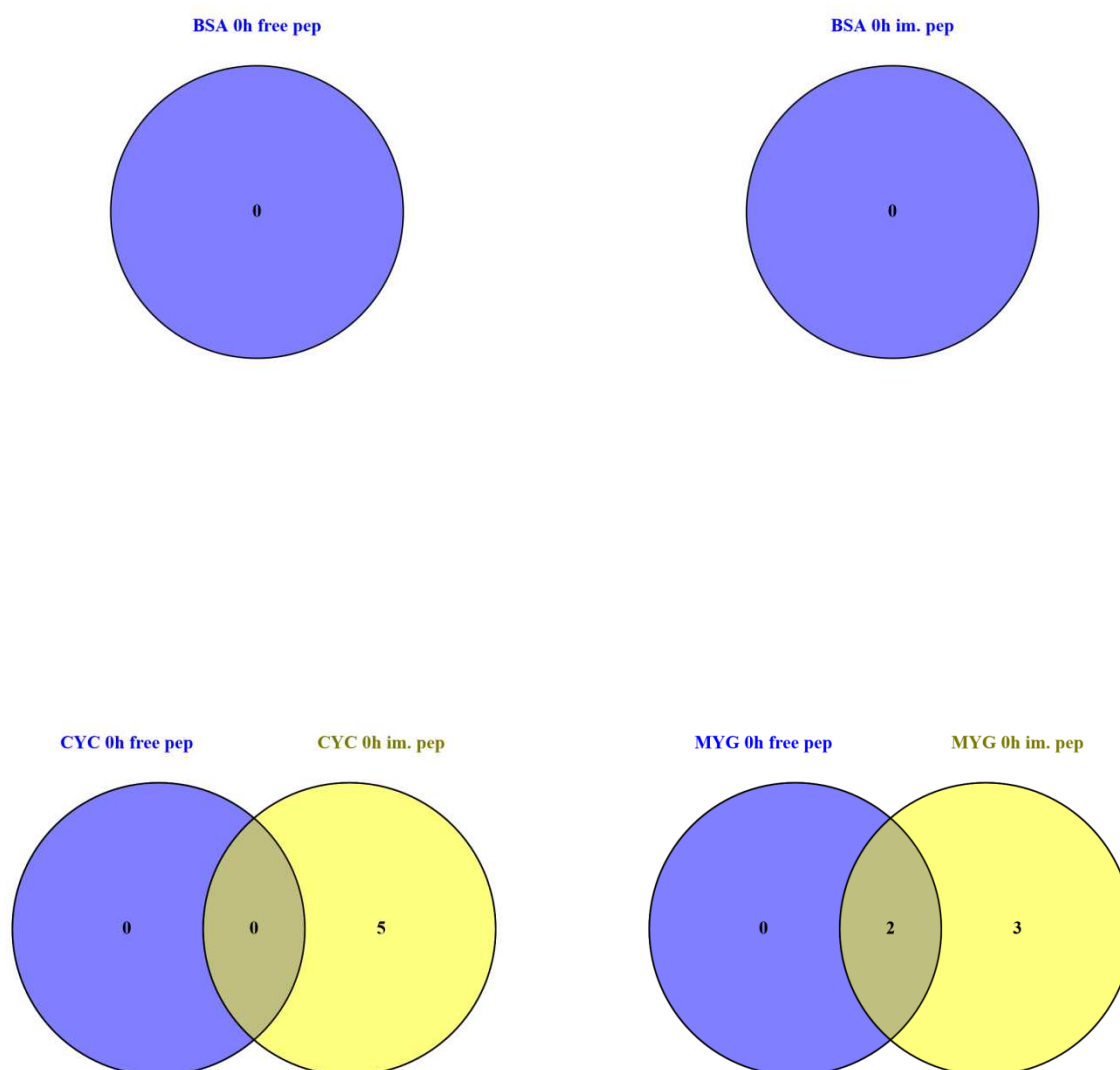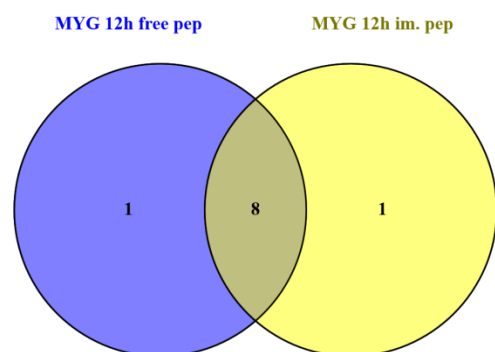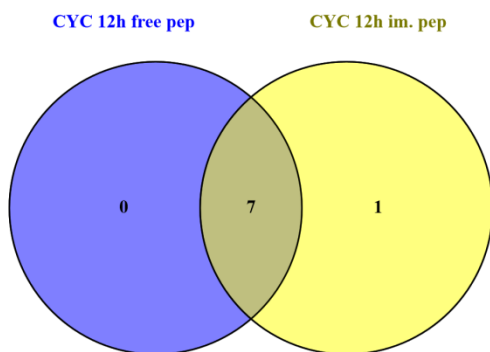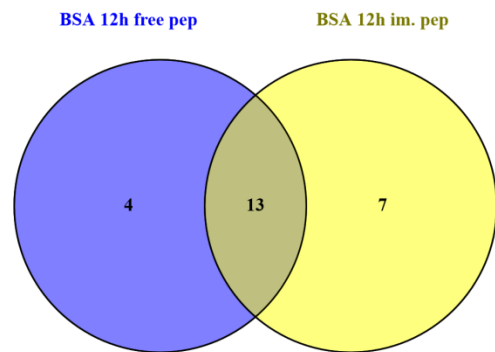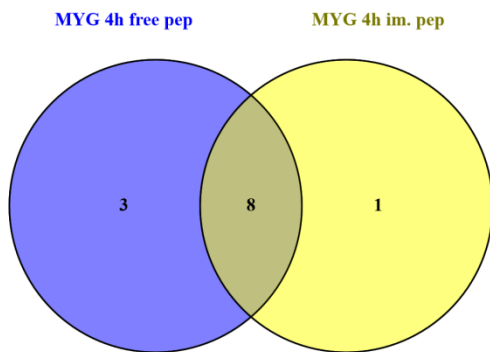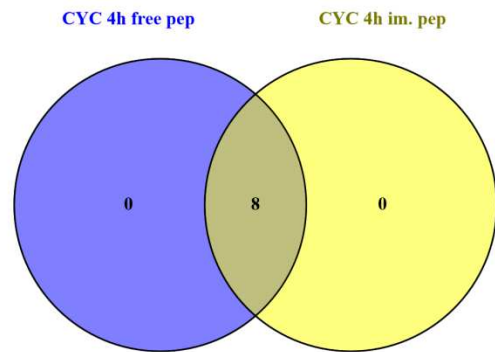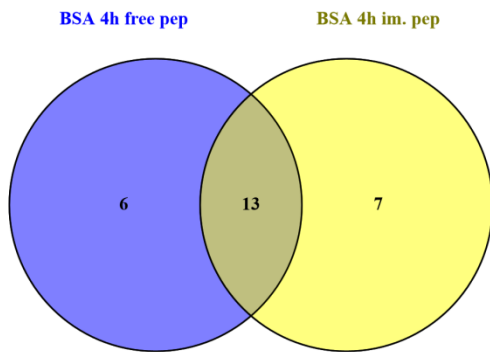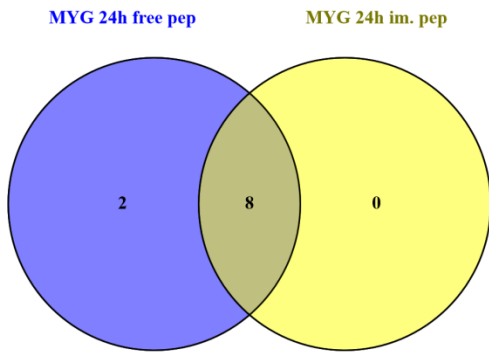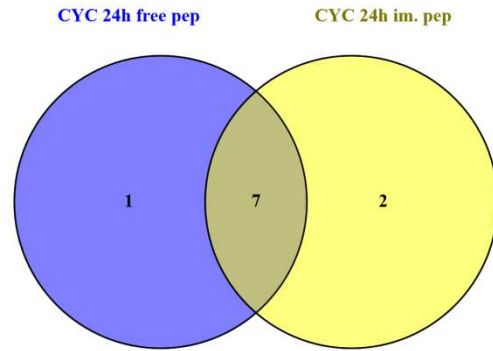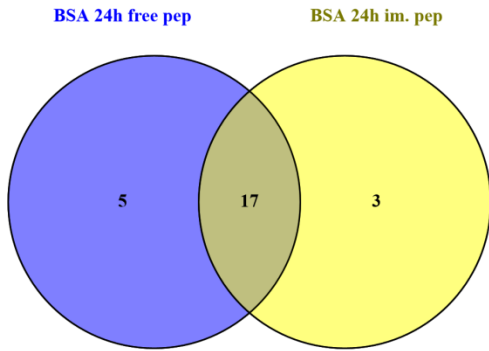
BSA 0h free pep

0

BSA 0h im. pep

0

CYC 0h free pep   CYC 0h im. pep

0   0   5

MYG 0h free pep   MYG 0h im. pep

0   2   3

BSA 4h free pep   **BSA 4h im. pep**

6   13   7

CYC 4h free pep   **CYC 4h im. pep**

0   8   0

MYG 4h free pep   **MYG 4h im. pep**

3   8   1

BSA 12h free pep   **BSA 12h im. pep**

4   13   7

CYC 12h free pep   **CYC 12h im. pep**

0   7   1

MYG 12h free pep   **MYG 12h im. pep**

1   8   1

BSA 24h free pep   **BSA 24h im. pep**

5   17   3

CYC 24h free pep   **CYC 24h im. pep**

1   7   2

MYG 24h free pep   **MYG 24h im. pep**

2   8   0

Tri- and Tetrasubstituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3

Felix Muth,[†,§] Ahmed El-Gokha,[†,‡,§] Francesco Ansideri,[†] Michael Eitel,[†] Eva Döring,[†] Adrian Sievers-Engler,[†] Andreas Lange,[†] Frank M. Boeckler,[†] Michael Lämmerhofer,[†] Pierre Koch,*[,†] and Stefan A. Laufer*[,†]

[†]Department of Pharmaceutical Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany
[‡]Chemistry Department, Faculty of Science, Menofia University, Menofia, Egypt

S *Supporting Information*
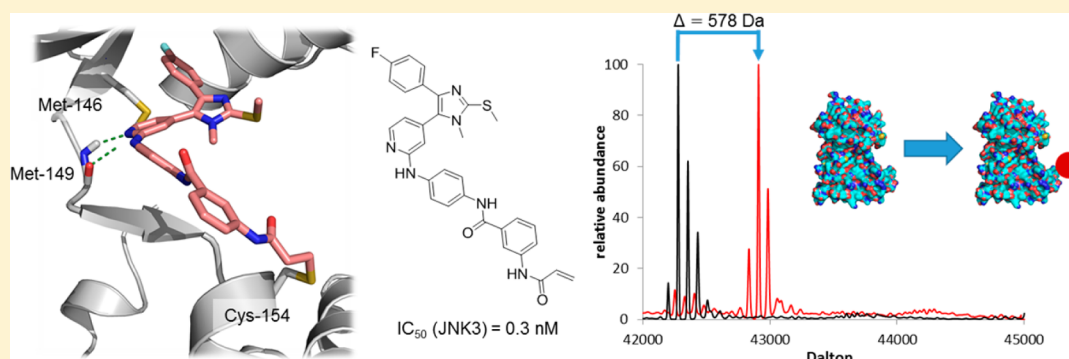
ABSTRACT: The concept of covalent inhibition of c-Jun N-terminal kinase 3 (JNK3) was successfully transferred to our well validated pyridinylimidazole scaffold varying several structural features in order to deduce crucial structure−activity relationships. Joint targeting of the hydrophobic region I and methylation of imidazole-N1 position increased the activity and reduced the number of off-targets. The most promising covalent inhibitor, the tetrasubstituted imidazole 3-acrylamido-*N*-(4-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (7) inhibits the JNK3 in the subnanomolar range (IC$_{50}$ = 0.3 nM), shows high metabolic stability in human liver microsomes, and displays excellent selectivity in a screening against a panel of 410 kinases. Covalent bond formation to Cys-154 was confirmed by incubation of the inhibitors with wild-type JNK3 and JNK3-C154A mutant followed by mass spectrometry.

INTRODUCTION

The c-Jun N-terminal kinase 3 (JNK3) belongs to the mitogen-activated protein (MAP) kinase family, which comprises 10 members: p38 MAP kinases (p38α, -β, -γ, -δ), JNKs (JNK1, -2, -3), and three extracellular-regulated kinases.[1] The three genes *jnk1, jnk2, jnk3* encode for a total of 10 alternatively spliced JNK isoforms with a molecular weight between 46 and 55 kDa.[2,3] Although structurally highly conserved, JNKs differ in tissue distribution fundamentally. JNK1 and JNK2 are ubiquitously expressed, whereas JNK3 expression is primarily restricted to the heart, testis, and brain.[4]

On the basis of results from several knock out experiments[5−7] and due to its specific tissue distribution, JNK3 has been identified as a promising target for potential treatment of neurodegenerative disorders such as Huntington's disease, Parkinson's disease, and Alzheimer's disease.[8]

Within the past years, covalent inhibition of enzymes has experienced a resurgence.[9] Carefully designed irreversible inhibitors have been proven to be even more selective than their reversibly binding counterparts.[10,11] With approval of the covalent inhibitor of the epidermal growth factor receptor afatinib in 2013,[12] the design of irreversible inhibitors raised the attention of the protein kinase research community.

The recently highlighted high quality kinase probe JNK-IN-8[13] (1, Figure 1) was discovered by Zhang et al.[14] and provided the first highly selective covalent pan-JNK inhibitor. The reported aminopyrimidine-based pan-JNK inhibitors possess an electrophilic moiety and target a cysteine (Cys-154, JNK3 numbering) conserved among all JNKs (in the hinge 7 position) that is not present in other protein kinases.[10]

In the search for new covalent inhibitors, researchers mostly follow two different paths: either a nonspecific weak reversible binder is used as a lead structure or a noncovalent, relatively potent inhibitor already optimized for the target kinase is chosen to be refined.[14] In continuing efforts to enhance both selectivity and activity of pyridinylimidazole-based JNK3 inhibitors, we applied the latter approach of covalent targeting
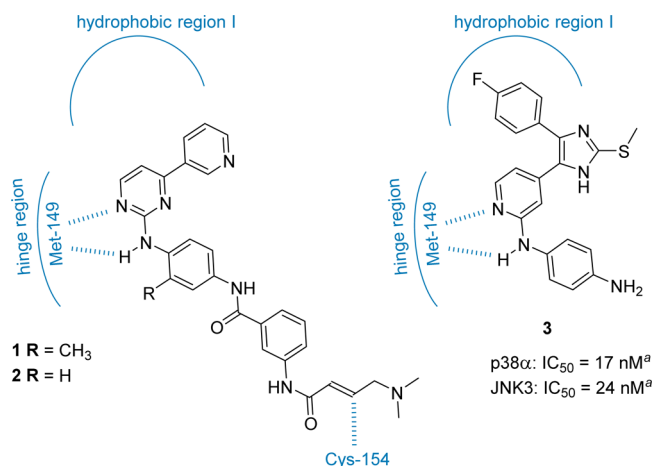
Figure 1. Pan-JNK inhibitors **1** and **2** and dual JNK3/p38α inhibitor **3**. [a]Data previously reported by Ansideri and co-workers.[15]

to our previously reported potent dual JNK3/p38α inhibitor **3** (Figure 1).[15] This scaffold was designed to align all moieties for optimal interaction with the active site of the enzymes. Computational docking studies with inhibitor **3** as well as comparison of known crystal structures of similar pyrimidinylimidazole inhibitors[16] (PDB codes 1PMN and 4Z9L) bound to JNK3 suggest that the 4-fluorophenyl ring of **3** is located in the hydrophobic region I of JNK3 and two hydrogen bond interactions are formed with Met-149 positioned in the hinge region. Therefore, it provides a suitable starting point to examine structure−activity relationships (SARs).

Due to the fact that the hinge binding motifs present in **1** and in our 2-aminopyridine scaffold are forming two hydrogen bonds to Met-149 (Figures 1 and 2), derivatives of **3** bearing an appropriate side chain at the pyridine-C2 position containing an electrophilic warhead should be able to target the unique cysteine (Cys-154).

This design hypothesis was supported by the comparing of the binding mode of JNK-IN-7 (**2**, Figure 1) in the crystal structure with JNK3 (PDB code 3V6S) and one of our
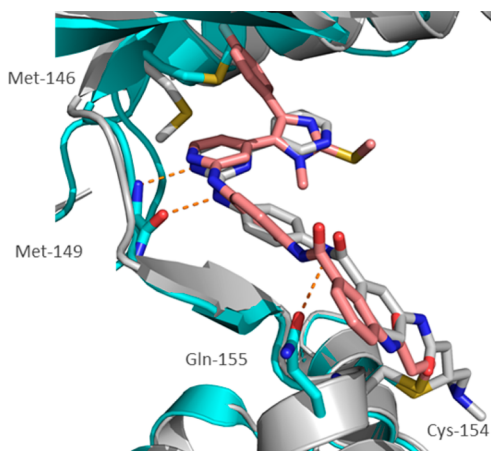
Figure 2. Superimposition of **2** (gray) in crystal structure with JNK3 (PDB code: 3V6S, cyan) and **7** (salmon) covalently docked into JNK3 (PDB code 1PMN, gray) using Schrödinger Glide.[17] Polar contacts of **2** and JNK3 are depicted as orange dashes. The image was generated with PyMOL.[18] Amino acid sequence was partially hidden for clarity.

designed inhibitors, compound **7**, computationally docked into the active site (Figure 2).

Herein, we report the synthesis of tri- and tetrasubstituted pyridinylimidazoles (compounds **4**−**23**) as well as their evaluation as covalent JNK3 inhibitors (Figure 3). The effects
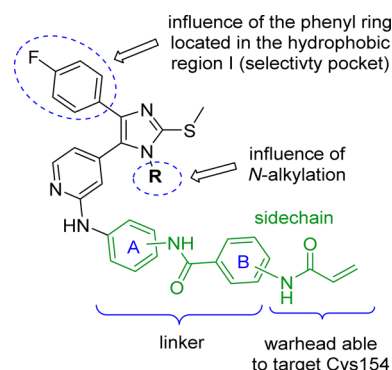
Figure 3. Structural modifications of imidazole scaffold and depiction of the terminology used throughout this article.

of the 4-fluorophenyl ring at the imidazole-C4 position as well as those of an imidazole N-alkylation were evaluated. Moreover, the influence of the substitution pattern of both phenyl rings A and B on the biological activity was investigated extensively. An acrylamide-type electrophilic warhead was chosen for the design of the herein presented compounds, since the α,β-unsaturated carbonyl group represents a soft electrophile, which reacts preferentially with the soft thiol present in cysteine side chains.[19] Furthermore, the presence of this Michael acceptor in several launched covalent kinase inhibitors (afatinib, ibrutinib, osimertinib)[20] as well as in drug candidates (neratinib, dacomitinib)[20,21] allows the assessment of a general safety profile for compounds bearing this moiety. However, in contrast to the N,N-dimethylaminomethyl acrylamide present in afatinib or neratinib as well as in compounds **1** and **2**, the unsubstituted acrylamide was preferred as electrophilic warhead for our series of inhibitors. In fact, although the tertiary amine has proven to improve properties as solubility and cellular half-life,[22] it also significantly increases the reactivity of the acrylamide by catalyzing the Michael addition[23] and therefore could be potentially responsible for off-target side effects.

■ RESULTS AND DISCUSSION

Chemistry. The synthesis of trisubstituted imidazole derivatives **4** and **5** bearing a 4-fluorophenyl moiety at the imidazole-C4 position was accomplished in three steps starting from reported inhibitor **3**[15] (Scheme 1). First, the primary amine was reacted with the respective nitrobenzoic acid using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) as a coupling reagent to yield amides **24** and **25**. Second, the nitro compounds **24** and **25** were reduced to the corresponding amines **26** and **27** using tin(II) chloride. Finally, the Michael acceptor system was introduced by reaction of the anilines **26** and **27** with acryloyl chloride in the presence of N,N-diisopropylethylamine (DIPEA) as a base. As a negative control for evaluation of covalent binding to the enzyme, compound **6** (the saturated analog of **4**) was synthesized by reaction of **3** with 3-propionamidobenzoic acid and PyBOP as coupling reagent (Scheme S1, Supporting Information).

Scheme 1. Synthesis of Trisubstituted Pyridinylimidazoles 4 and 5 Bearing a 4-Fluorophenyl Ring at the Imidazole-C4 Position[a]
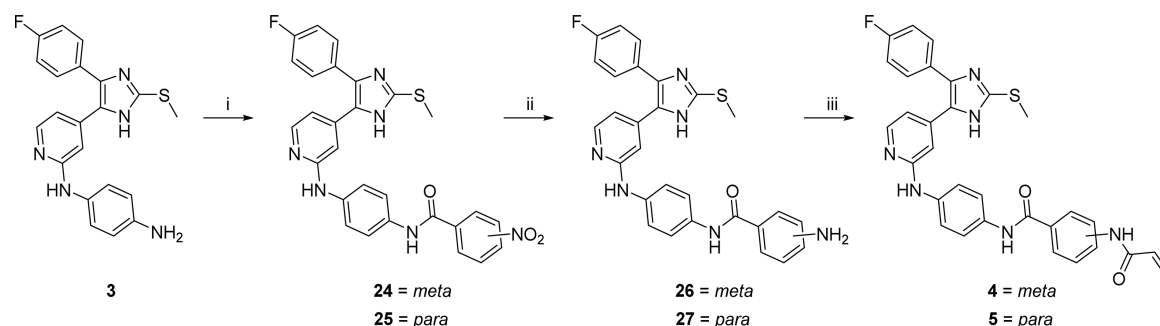
[a]Reagents and conditions: (i) *m*- or *p*-nitrobenzoic acid, PyBOP, DIPEA in dichloromethane, 5 h, rt; (ii) SnCl$_2$·2H$_2$O, EtOH, reflux, 6 h; (iii) acryloyl chloride, DIPEA, DMF, 0 °C to rt.

Scheme 2. Synthesis of Final Compounds 7, 8, and 10−20 Exemplified by Synthesis of 7[a]
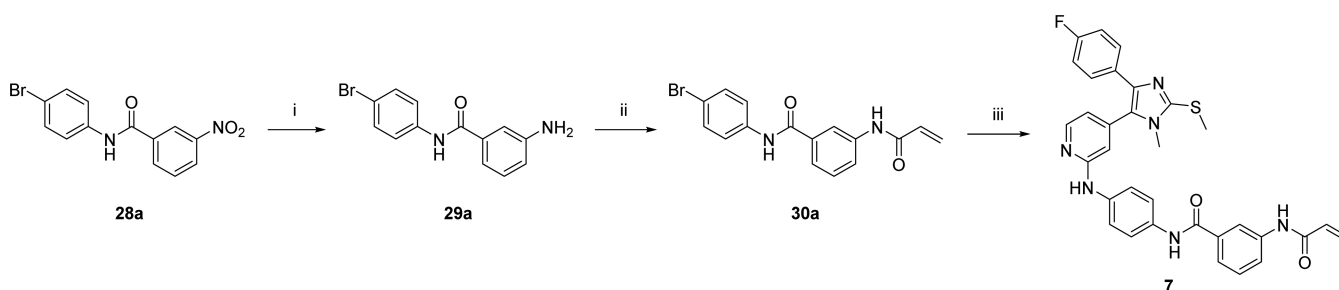
[a]Reagents and conditions: (i) SnCl$_2$·2H$_2$O, EtOH, 75 °C, 2.5 h; (ii) acryloyl chloride, DIPEA, 1,4-dioxane, rt, 2 h; (iii) **31**, Cs$_2$CO$_3$, BrettPhos precatalyst, 1,4-dioxane(abs)/*tert*-butanol 4:1, 125 °C, 5 h.

Scheme 3. Synthesis of Trisubstituted Pyridinylimidazoles 21 and 22 Bearing a Methyl Moiety at the Imidazole-C4 Position[a]
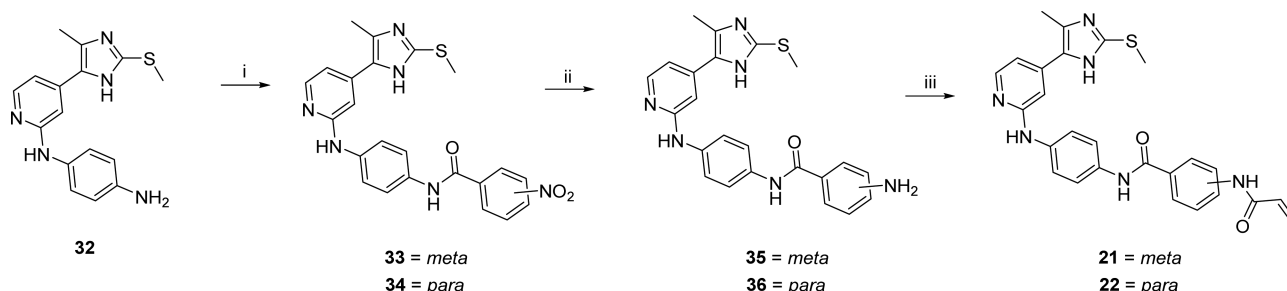
[a]Reagents and conditions: (i) *m*-nitrobenzoic acid, SOCl$_2$, pyridine, rt or *p*-nitrobenzoyl chloride, pyridine, rt; (ii) SnCl$_2$·2H$_2$O, EtOH, reflux, 6 h; (iii) acryloyl chloride, DIPEA, DMF, −20 °C to rt, 3 h.
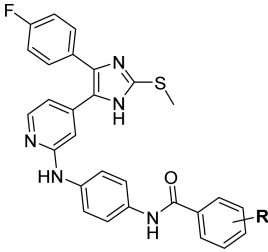
The tetrasubstituted imidazoles **7**, **8**, and **10−20** were prepared using a convergent synthetic strategy starting from the corresponding nitrobenzoic acid and bromoaniline (Scheme 2). Compounds **28a−h** were synthesized in good yields by adding *O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium tetra-fluoroborate (TBTU) and the corresponding bromoaniline to a suspension of nitrobenzoic acid in dichloromethane at mild conditions using DIPEA as base. In order to reduce the nitro group to a primary amine, compounds **28a−h** were treated with tin(II) chloride in ethanol at 75 °C. The acrylamide warhead and its saturated counterpart were introduced using the corresponding acid chlorides. Compounds **29a−h** were suspended in dry 1,4-dioxane and treated with the activated carboxylic acid at mild conditions using DIPEA as base. In the last step of this synthetic route, a Buchwald−Hartwig cross-coupling reaction was applied. Optimized conditions for the

introduction of substituted aryl moieties at the pyridine-C2 amino position have been previously described.[24] The scaffold 4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-amine (**31**)[25] and the side chains **30a−m** were coupled using BrettPhos Pd G1 precatalyst as the catalyst/ligand system and cesium carbonate as base.[19] Compound **9** was synthesized by performing the Buchwald−Hartwig aryl amination reaction with **28a** and **31** followed by a reduction step (Scheme S2, Supporting Information).

This synthetic strategy allows the introduction of the side chain using optimized Buchwald−Hartwig cross-coupling reaction conditions in the last step, thus providing a high degree of flexibility in terms of varying the scaffold or the side chain.

As further modifications, compounds **21** and **22** bearing a small methyl moiety at the imidazole-C4 position were

Table 1. Biological Activity of Trisubstituted Imidazoles 4−6

Cmp	R	JNK3[a]		p38α[b]		Q[d]
		IC$_{50}$ [nM][c]	SEM	IC$_{50}$ [nM][c]	SEM	
4	*meta*	0.6	0.04	0.8	0.11	~1
5	*para*	3	0.07	3	0.69	1
6	*meta*	2	0.07	4	0.27	2

[a]Incubation time, 50 min. [b]Incubation time, 60 min. [c]IC$_{50}$ values are the mean of three experiments. [d]Ratio: IC$_{50}$(p38α)/IC$_{50}$(JNK3).

synthesized in three steps (Scheme 3) starting from N^1-(4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)benzene-1,4-diamine (**32**; for synthesis of imidazole **32**, see Scheme S3, Supporting Information). The buildup of the linker bearing the electrophilic warhead at the aniline function was conducted similarly to the strategy described for the trisubstituted imidazoles bearing a 4-fluorophenyl ring at the imidazole core. The conditions used to synthesize compound **23**, the saturated analog of **21**, starting with precursor **32** are identical to those used for the synthesis of **6** (Scheme S4, Supporting Information).

Biological Evaluation. All synthesized inhibitors were evaluated in enzyme-linked immunosorbent assays (ELISA)[26,27] to determine the IC$_{50}$ values for JNK3 and p38α. Compounds **7** and **21** were examined for their metabolic stability and further screened in a panel against 410 kinases to determine their selectivity within the kinome. Results are presented in Tables 1−4 as well as in Tables S1 and S4 (Supporting Information).

Both examples of the trisubstituted pyridinylimidazoles series bearing a 4-fluorophenyl ring at the imidazole-C4 position (compounds **4** and **5**) are very potent dual JNK3/p38α inhibitors displaying IC$_{50}$ values down to the picomolar range (Table 1). While no preference for either tested kinase was detectable, compounds **4** and **5** deduced from compound **3** (Figure 1) revealed that the bulky side chain not only was well tolerated by both enzymes but also caused an increase in inhibitory activity that is distinct in compound **4** having the Michael acceptor in meta position. However, within the series of 4-fluorophenyl-substituted imidazoles, the influence of the Michael acceptor system in terms of JNK3 inhibition is rather low, since the saturated analog of **4**, compound **6**, shows no significant decrease in terms of JNK3 affinity.

We could not observe any beneficial effect of introducing electrophilic warheads in terms of selectivity regarding the trisubstituted imidazoles bearing a 4-fluorophenyl moiety. Therefore, as a parallel concept, we chose to alkylate the
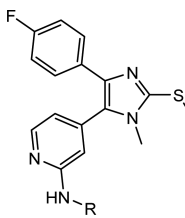
imidazole-N1 position with the utmost simple substituent as prior investigations of our group correlate the N-alkylation with a reduced p38α inhibition.[28]
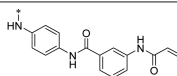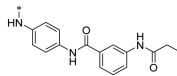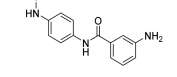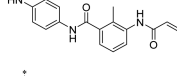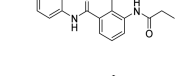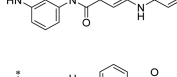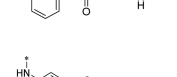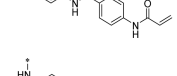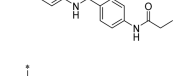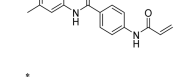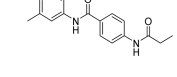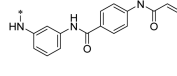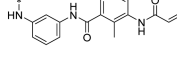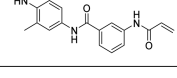
It is crucial for the linker between the scaffold and the warhead to comprise the optimal length and angle in order to orient the warhead ideally for the nucleophilic attack of the thiol to occur and simultaneously retain the original binding mode of the scaffold. Therefore, we synthesized a broad variety of linkers, altering the meta and para substitution patterns consequently. Furthermore, we introduced methyl groups at different positions of the A and B rings to slightly increase the torsion angle.

In order to elucidate the contribution of each moiety on ring B of **7** and **8**, we synthesized compound **9** (IC$_{50}$: JNK3, 30 nM; p38α, 149 nM). The para/meta substitution pattern (whereupon para refers to ring A and meta to ring B) with the unsubstituted aniline, presenting a 5-fold prevalence for JNK3 over p38α, was rather encouraging since the covalent tag was yet to be introduced. The para/meta substituted compound **7** (IC$_{50}$: JNK3, 0.3 nM; p38α, 36 nM), bearing an acrylamide warhead, features a picomolar inhibitory effect for JNK3 and 120-fold preference compared to the IC$_{50}$ value of p38α, emerging as the most potent inhibitor synthesized in this series. Surprisingly, we observed a lower IC$_{50}$ value for p38α alongside the JNK3 inhibition compared to compound **9**. The initial assumption that the carbonyl oxygen of the acrylamide forms an additional hydrogen bond to p38α and thereby increases its inhibitory effect was invalidated by compound **8**. It shows a 4-fold decreased inhibition for p38α compared to its unsaturated counterpart, compound **7**. More interestingly, JNK3 inhibition of compound **8** is significantly decreased compared to compound **7**. This can be considered as a strong indication for the formation of a covalent bond of **7** to the JNK3.

Compound **10** (IC$_{50}$: JNK3, 2 nM; p38α, 2 nM), which features the para/meta substitution pattern and a methyl group in ortho position at the B ring, inhibits both kinases in the low single digit nanomolar range. The methyl group at that position

Table 2. Biological Activities of Tetrasubstituted Imidazoles 7−20
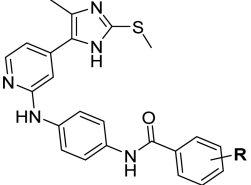
Cmp	R	JNK3[a]		p38α[b]		Q[d]
		IC$_{50}$ [nM][c]	SEM	IC$_{50}$ [nM][c]	SEM	
7	*(structure)*	0.3	0.03	36	4.38	120
8	*(structure)*	23	0.5	126	21.18	5
9	*(structure)*	30	0.54	149	2.71	5
10	*(structure)*	2	0.07	2	0.08	1
11	*(structure)*	17	0.92	38	3.25	2
12	*(structure)*	3	0.16	217	10.39	72
13	*(structure)*	175	17.56	241	6.28	1
14	*(structure)*	10	1.01	47	1.23	5
15	*(structure)*	27	1.08	17	0.83	1
16	*(structure)*	8	1.62	35	1.64	4
17	*(structure)*	25	0.68	22	0.92	1
18	*(structure)*	4	0.24	26	1.43	7
19	*(structure)*	3	0.07	100	8.60	33
20	*(structure)*	2	0.65	3	0.34	1

[a]Incubation time, 50 min. [b]Incubation time, 60 min. [c]IC$_{50}$ values are the mean of three experiments. [d]Ratio: IC$_{50}$(p38α)/IC$_{50}$(JNK3).

seems to increase the potency on p38α drastically. The saturated analog **11** (IC$_{50}$: JNK3, 17 nM; p38α, 38 nM) lacking the electrophilic warhead supports this assumption since JNK3 and p38α inhibition are alike. The inversion of the para/meta substitution pattern led to compound **18** (IC$_{50}$: JNK3, 4 nM; p38α, 26 nM) exhibiting a low preference for JNK3. The

Table 3. Biological Activity of Trisubstituted Imidazoles 21−23

Cmp	R	JNK3[a]		p38α[b]		Q[d]
		IC$_{50}$ [nM][c]	SEM	IC$_{50}$ [nM][c]	SEM	
21	*meta*	2	0.02	1,952	62	976
22	*para*	200	12	5,920	107	30
23	*meta*	253	8.7	>10,000	---	>39

[a]Incubation time, 50 min. [b]Incubation time, 60 min. [c]IC$_{50}$ values are the mean of three experiments. [d]Ratio: IC$_{50}$(p38α)/IC$_{50}$(JNK3).

introduction of an ortho methyl group at the A ring, which was described by Zhang et al.[14] to be the key determinant for selectivity, bred compound **20** (IC$_{50}$: JNK3, 2 nM; p38α, 3 nM). Since we could not observe a selectivity increase with respect to JNK3 and p38α inhibition, we introduced the ortho methyl group to the para/para substituted compound **16** (IC$_{50}$: JNK3, 8 nM; p38α, 35 nM) and its propionamide analog **17**. However, this did not yield the desired effect.

Judged by the small difference in inhibition between compounds **14** (IC$_{50}$: JNK3, 10 nM; p38α, 47 nM) and **15** (IC$_{50}$: JNK3, 27 nM; p38α, 17 nM) it appears that compound **14** does not establish a covalent bond with JNK3. Remarkably, compound **15** bearing the propionamide moiety is better tolerated by p38α compared to **14**, which stands in contrast to the prior observation. The meta/meta substituted **12** (IC$_{50}$: JNK3, 3 nM; p38α, 217 nM) features a 72-fold preference for JNK3 compared to the IC$_{50}$ value of p38α. Additionally, the almost 70 times lower JNK3 inhibition of **13** (IC$_{50}$: JNK3, 175 nM; p38α, 241 nM) and the almost constant p38α inhibition suggest the successful formation of a covalent bond between compound **12** and Cys-154. The unfavorable ortho methylation at the B ring seen with compounds **19** and **20** (increasing p38α inhibition) is in agreement with a similar observation in compounds **16** and **17**.

Since our efforts of optimizing the linker did not result in a JNK3 inhibitor exceeding 120-fold preference over p38α, we chose to replace the 4-fluorophenyl ring of compounds **4−6** by a small methyl group leading to compounds **21−23**.

The rationale behind this modification is the fact that numerous examples of known JNK selective inhibitors are not targeting the hydrophobic region I (for recent reviews on JNK inhibitors, see Koch et al.[29] and Gehringer et al.[30]). The most selective JNK3 inhibitor in this series (Table 3), compound **21**, shows an IC$_{50}$ value in the low nanomolar range and is 976-fold selective against p38α (IC$_{50}$: JNK3, 2 nM; p38α, 1,952 nM). Comparison of **21** and **22** demonstrates that the position of the electrophilic warhead on ring B is crucial for improving both JNK3 activity and selectivity. Compound **22** (IC$_{50}$: JNK3, 200 nM), having the acrylamide in para-position, is 100-fold less active than its isomer **21**, which has the same substituent in

meta-position. Compound **23** lacks the Michael acceptor system and shows a tremendous drop (>125-fold) in JNK3 inhibition compared to **21**. SARs of the trisubstituted pyridinylimidazoles series, which carry a small methyl moiety at the imidazole-C4 position, indicate the formation of covalent bond between JNK3 and the inhibitor **21**. Compound **22**, presenting a para/para substitution pattern, apparently does not establish a covalent bond to the cysteine in JNK3, since its inhibitory effect was similar to the saturated compound **23**, unable to form a covalent bond.

Inhibitors **7** and **21**, displaying the best inhibition profile within the series, were further investigated. In order to confirm the formation of a covalent bond to Cys-154, the enzyme−inhibitor complex was analyzed by liquid chromatography−mass spectrometry (LC−MS). Both compounds were incubated with JNK3, and subsequently subjected to LC−MS analysis. Unbound and reversibly bound ligand were removed both by chromatography and by application of a high declustering potential in the course of ionization (for further details, see Supporting Information). These two passages are aimed at disrupting reversible electrostatic interactions, allowing only the detection of covalently bound ligand. The expected mass shift of ∼578 Da and ∼484 Da in the case of incubation with inhibitor **7** and **21**, respectively, could be measured on the intact protein level using LC-μESI-QTOF mass spectrometry (for further details, see Table S2 and Figures S1−S3a,b,d, Supporting Information). Moreover, performing the same experiment with compounds **8** and **23**, the saturated analogs of **7** and **21**, no mass shift could be observed, since the noncovalent inhibitors do not sustain the binding to JNK3 during the LC−MS experiment (Figures S3c and S3e, Supporting Information). Interestingly, MS analysis carried out on compound **14** (Figure S4a, Supporting Information), which bears the electrophilic warhead in para position and had been assumed not capable of binding covalently, displayed the mass shift of ∼578 Da relative to the irreversible bond with the target. However, the shifted peak had a significantly decreased intensity compared to the mass shift of the aforementioned compounds **7** and **21**, revealing that the para/para substitution pattern of the linker can still permit the formation of the

covalent bond, although with a nonfavorable efficiency. As expected, incubation of compound **15**, the noncovalent counterpart of **14**, did not give rise to any enzyme mass increase (Figure S4b, Supporting Information). Finally, the analogous MS experiment was performed by incubating compounds **7**, **14**, and **21** with a JNK3 mutant, where the reactive Cys-154 was replaced by alanine (Figure S5, Supporting Information). Results for the three compounds tested showed absence of mass shift, proving our hypothesis both on the formation of the covalent bond and on the specific amino acid (Cys-154) targeted.

Covalent inhibitors **7** and **21** were further tested in a panel of 410 kinases, including 70 disease-relevant protein kinase mutants and 13 lipid kinases (Table S1, Supporting Information) to evaluate their selectivity within the kinome. At a testing concentration of 1 μM, which represents 500 times its IC$_{50}$ value, imidazole **21** inhibited only 15 kinases (including all JNK isoforms) out of the tested kinase panel, resulting in a low S (50) score of 0.037. This parameter is defined as the portion of kinases that are inhibited to more than 50%, in relation to all tested kinases.[31] Notably, the tetrasubstituted imidazole **7** reveals a remarkable selectivity profile. Besides the JNKs, **7** inhibits only three other protein kinases (Table 4) at a

Table 4. Off-Target Activity of Compounds 21 and 7 in a Panel of 410 Kinases

21 at 1 μM				7 at 0.5 μM
CK-1-δ	CLK4	MKK7	TAOK2	TIE2-wt
CK-1-γ2	MAP3K10	PKC-nu	TSF1	MAPKAP2
CK-1-γ3	MAP3K11	SGK1	ZAK	CK-1-δ
	10 – 29 %[a]		30 – 50 %[a]	

[a]Residual activity.

screening concentration of 0.5 μM (representing more than 1500-fold its IC$_{50}$ value) and possesses an excellent S (50) score of 0.015. The slightly changed selectivity is mainly attributed to the methylation of the imidazole-N1 position and specific interactions within the hydrophobic region I, the so-called selectivity pocket, which is not targeted by compound **21**. As expected, no selectivity was seen within the JNK family, since the targeted cysteine is conserved among all isoforms.

All final compounds were checked for being potential pan-assay interference compounds (PAINS) using the ZINC Patterns tool,[32] which led to no hits. Since acrylamides might interfere with intracellular proteins in a nondiscriminatory manner,[33] the most promising covalent inhibitor **7** (10 μM) was incubated with the nonspecific thiol glutathione (present in almost all cells, 5 mM) under similar conditions used in the kinase activity assay (maximum concentration 10 μM, 54 min, 37 °C). Subsequent high performance liquid chromatography (HPLC) analysis revealed only low depletion of **7** (Table S3; Supporting Information). This observation together with the rather clean profile of acrylamide compounds **7** and **21** in the selectivity screening shows no evidence of unpredictable interactions with other cysteine side chains.

Inhibitors **7** and **21** were further tested for their metabolic stability in male human liver microsomes (HLMs). Both covalent inhibitors undergo a partial biotransformation. Nevertheless, both tested compounds are exhibiting 13% and 31% degradation over a time span of 180 and 190 min, respectively,

and can therefore be considered to be metabolically stable (Table S4 and Figures S6 and S7, Supporting Information). In agreement with prior findings,[24] the metabolic degradation leads most likely to the oxidation of the imidazole-C2 sulfur.

CONCLUSION

We successfully transferred the concept of covalently inhibiting the JNKs to our pyridinylimidazole scaffold, which was elaborated for JNK interaction. Compounds **7** and **21** with their para/meta substitution pattern comprise the optimal linker length and angle for the electrophilic warhead to form a covalent bond to the Cys-154. Incubation of JNK3 with compounds **7** and **21** followed by mass experiments substantiated the assumption of a covalent bond formation. Both covalent inhibitors are metabolically relatively stable when exposed to HLM. Targeting the hydrophobic region I with a 4-fluorophenyl ring combined with a methylation of the imidazole-N1 position reduced the number of off targets, resulting in compound **7** with a remarkable selectivity profile.

EXPERIMENTAL SECTION

Chemistry. General. The purity of all tested compounds is ≥95% and was determined via reverse phase HPLC.

General Procedure for the Reduction of Nitro Group of Compounds 24, 25, 33, and 34 Using Tin(II) Chloride (General Procedure A). In 100 mL round-bottom flask, the nitro compound was dissolved in ethanol. Tin(II) chloride was added, and the reaction mixture was refluxed for 6 h. After cooling to rt, a saturated solution of sodium bicarbonate (20 mL) was added. The aqueous layer was extracted with ethyl acetate (3 × 10 mL), and the combined organic layer was dried over anhydrous Na$_2$SO$_4$. The solvent was evaporated under reduced pressure, and the crude product was purified by flash chromatography (SiO$_2$, dichloromethane/ethanol, 95:05 to 90:10).

General Procedure for the Preparation of Final Compounds 4, 5, 21, and 22 (General Procedure B). The respective primary amine **26**, **27**, **35**, or **36** was dissolved in DMF (2 mL) and placed in a round-bottom flask under argon atmosphere. DIPEA was added in one portion followed by dropwise addition of acryloyl chloride diluted in dry dichloromethane (1 mL) at 0 °C (in the cases of preparation of **4** and **5**) or in dry DMF (0.2 mL) at −20 °C (in the cases of preparation of **21** and **22**). The reaction mixture was further stirred for 2 h at rt, then poured over ice and stirred for 30 min. If the crude product precipitated, the precipitate was filtered off and purified by column chromatography. In the case of preparation of **21**, the crude product did not precipitate. Therefore, the aqueous phase was extracted with ethyl acetate (3 × 10 mL). The combined organic layers were dried over Na$_2$SO$_4$, the solvent was evaporated under reduced pressure, and the residue was purified with flash chromatography.

General Procedure for Synthesis of Amides 28a–h (General Procedure C). To a stirred suspension of corresponding nitrobenzoic acid (1.1 equiv) in dichloromethane (9 mL/mmol aniline), TBTU (1.1 equiv), the corresponding bromoaniline (1 equiv), and DIPEA (2 equiv) were added at 0 °C. The yellowish suspension cleared up as it was allowed to warm to room temperature and was subsequently stirred for 3 h until complete consumption of the respective bromoaniline was detected by TLC. The reaction mixture was poured on a pad of silica and eluted with a mixture of ethyl acetate and petrol ether 50/50.

General Procedure for Synthesis of 29a–h by Reduction of Nitrobenzene Analogs 28a–h (General Procedure D). To a suspension of tin(II) chloride (5 equiv) in ethanol (100 mL/mmol nitrobenzene) was added the corresponding nitrobenzene (1 equiv), and the reaction mixture was heated to 75 °C for 2.5 h until complete conversion was detected by TLC. The solvent was evaporated under reduced pressure. The white residue was treated with saturated NaHCO$_3$ solution and brine. The aqueous layer was extracted with ethyl acetate (3 × 60 mL). The combined organic phases were dried over Na$_2$SO$_4$ and evaporated under reduced pressure to yield the

desired title compound as a white solid, which was used without further purification.

General Procedure for Synthesis of Compounds 30a−m via Amide Coupling (General Procedure E). To a suspension of corresponding aniline **29** (1 equiv) in 1,4-dioxane (12 mL/mmol aniline), DIPEA (2.5 equiv) was added followed by the corresponding acid chloride (1.1 equiv) at 0 °C. The mixture was allowed to warm up to room temperature and was stirred for 2 h until complete conversion was detected by TLC. After removal of all volatiles under reduced pressure, methanol was added and the mixture was evaporated again. This process was repeated three times, and the solid obtained was used without further purification.

General Procedure for Synthesis of Final Compounds 7, 8, 10−20 via Buchwald−Hartwig Cross-Coupling Reaction (General Procedure F). BrettPhos precatalyst (2.5 mol %) was added to a dry mixture of *tert*-butanol and 1,4-dioxane (1:4; 54 mL/mmol **31**) in an argon flushed glass tube. Compound **31** (1 equiv) was added to the reaction mixture followed by cesium carbonate (2.5 equiv) and the respective aryl bromide **30** (1.1 equiv). The reaction mixture was heated to 125 °C and stirred for 2 h after sealing the glass tube tightly. Another 2.5 mol % BrettPhos precatalyst was added, and the reaction mixture was heated to 125 °C and stirred for an additional 3 h until HPLC confirmed complete conversion. The reaction mixture was allowed to cool to room temperature, filtered, and washed with ethyl acetate (20 mL). All volatiles were removed under reduced pressure, and the residue was purified by silica flash chromatography.

3-Acrylamido-N-(4-((4-(4-(4-fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (4). Compound **4** was synthesized according to general procedure B, using compound **26** (80 mg, 0.16 mmol), DIPEA (30 μL, 0.17 mmol), and acryloyl chloride (16 μL, 0.19 mmol). Purification by flash chromatography (SiO₂, dichloromethane/ethanol, 92:08) afforded 35 mg (40%) of a light yellow solid product. ¹H NMR (250 MHz, DMSO-d_6) δ 12.67 (br s, 1H), 10.35 (s, 1H), 10.14 (s, 1H), 8.97 (br s, 1H), 8.21−7.96 (m, 2H), 7.96−7.87 (m, 1H), 7.73−7.40 (m, 8H), 7.38−7.13 (m, 2H), 7.12−6.77 (m, 1H), 6.68 (d, J = 4.6 Hz, 1H), 6.48 (dd, J = 9.9, 17.0 Hz, 1H), 6.30 (dd, J = 2.3, 17.0 Hz, 1H), 5.88−5.72 (m, 1H), 2.62 (s, 3H). ¹³C NMR (101 MHz, DMSO-d_6) δ 164.9, 163.3, 161.9 (d, J = 247.0 Hz), 156.3, 147.2, 142.8, 141.9, 139.1, 137.9, 135.9, 134.7, 132.0, 131.7, 130.7 (d, J = 8.9 Hz), 129.5 (d, J = 3.5 Hz), 128.7, 127.1, 122.3, 122.0, 121.0, 118.8, 118.2, 115.7 (d, J = 23.0 Hz), 112.4, 111.8, 107.0, 15.0; HPLC t_R = 6.25 min. ESI-HRMS: calculated, 565.1817 [M + H]⁺; found, 565.1812.

4-Acrylamido-N-(4-((4-(4-(4-fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (5). Compound **5** was synthesized according to general procedure B, using compound **27** (78 mg, 0.15 mmol), DIPEA (30 μL, 0.17 mmol), and acryloyl chloride (15 μL, 0.18 mmol). Purification by flash chromatography (SiO₂, dichloromethane/ethanol, 92:08) afforded 22 mg (27%) of a light yellow solid product. ¹H NMR (400 MHz, DMSO-d_6) δ 12.68 (br s, 1H), 10.44 (s, 1H), 10.03 (s, 1H), 8.99 (br s, 1H), 8.12 (br s, 1H), 8.02−7.89 (m, 2H), 7.85−7.74 (m, 2H), 7.62 (d, J = 8.3 Hz, 2H), 7.57 (d, J = 8.6 Hz, 2H), 7.51 (br s, 2H), 7.31 (br s, 1H), 7.23 (br s, 1H), 7.12−6.74 (m, 1H), 6.74−6.60 (m, 1H), 6.49 (dd, J = 10.1, 17.0 Hz, 1H), 6.31 (d, J = 17.2 Hz, 1H), 5.81 (d, J = 10.1 Hz, 1H), 2.63 (s, 3H). ¹³C NMR (101 MHz, DMSO-d_6) δ 164.3, 163.4, 161.3 (d, J = 246.2 Hz), 156.3, 147.7, 147.1, 142.9, 142.7, 141.7, 137.6, 131.6, 130.6 (d, J = 5.5 Hz), 129.7, 128.4, 127.4, 126.8, 121.0, 118.5, 118.4, 118.3, 115.7 (d, J = 26.7 Hz), 111.8, 107.1, 99.9, 15.0. HPLC t_R = 6.11 min. ESI-HRMS: calculated, 565.1817 [M + H]⁺; found, 565.1804.

N-(4-((4-(4-(4-Fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-propionamidobenzamide (6). Under argon atmosphere, compound **3** (100 mg, 0.23 mmol) was suspended in dry dichloromethane (10 mL) and DIPEA (54 μL, 0.31 mmol) was added. Finally, a suspension of 3-propionamidobenzoic acid (54 mg, 0.28 mmol) and PyBOP (161 mg, 0.31 mmol) in dry dichloromethane (10 mL) was added and the reaction mixture was stirred at rt for 18 h. The reaction mixture was concentrated at reduced pressure and the crude was purified twice by flash column

chromatography (SiO₂, dichloromethane/ethanol, 97:03 to 80:20) yielding product **6** (30 mg, 21%). ¹H NMR (300 MHz, DMSO-d_6) δ 12.67 (br s, 1H), 10.11 (br s, 1H), 10.06 (s, 1H), 9.08−8.87 (m, 1H), 8.17−7.94 (m, 2H), 7.84 (d, J = 8.0 Hz, 1H), 7.67−7.15 (m, 10H), 7.10−6.76 (m, 1H), 6.73−6.64 (m, 1H), 2.63 (s, 3H), 2.36 (q, J = 7.5 Hz, 2H), 1.11 (t, J = 7.5 Hz, 3H); ¹³C NMR (101 MHz, DMSO-d_6) δ 172.2, 165.1, 163.1, 161.4 (d, J = 244.9 Hz), 156.3, 148.0, 147.2, 142.9, 142.0, 139.4, 138.9, 137.9, 137.4, 135.9, 134.7, 133.8, 132.5, 132.1, 130.7 (d, J = 8.0 Hz), 129.5 (d, J = 6.0 Hz), 128.6, 126.8, 126.2, 121.7, 121.0, 118.6, 118.5, 118.2, 115.8 (d, J = 22.0 Hz), 115.3 (d, J = 22.0 Hz), 112.4, 111.8, 111.4, 107.3, 107.0, 29.5, 15.1, 9.6. The peaks of the carbon are double due to the amide tautomerism. HPLC t_R = 5.47 min. ESI-HRMS: calculated, 567.1973 [M + H]⁺; found, 567.1957.

3-Acrylamido-N-(4-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-benzamide (7). According to general procedure F, **7** was synthesized from **31** and **30a**. Eluent ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 31% yield. ¹H NMR (200 MHz, DMSO-d_6) δ 10.34 (s, 1H), 10.15 (s, 1H), 9.12 (s, 1H), 8.26 (d, J = 5.4 Hz, 1H), 8.15 (s, 1H), 7.92 (d, J = 7.8 Hz, 1H), 7.70−7.59 (m, 5H), 7.53−7.40 (m, 3H), 7.14 (t, J = 8.9 Hz, 2H), 6.78−6.69 (m, 2H), 6.55−6.17 (m, 2H), 5.84−5.72 (m, 1H), 3.43 (s, 3H), 2.65 (s, 3H). ¹³C NMR (101 MHz, DMSO-d_6) δ 165, 163.3, 161.0 (d, J = 243.7 Hz), 156.5, 148.3, 143.5, 139.1, 138.9, 137.3, 136.8, 135.9, 132.5, 131.7, 130.6 (d, J = 2.9 Hz), 128.7, 128.3 (d, J = 8.0 Hz), 128.3, 127.1, 122.3, 122.0, 121.0, 118.8, 118.4, 115.1 (d, J = 21.5 Hz), 114.8, 111.1, 31.5, 15.3. ESI-HRMS: calculated, 578.1900 [M + H]⁺; found, 578.1876. HPLC t_R = 5.87 min.

N-(4-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-propionamido-benzamide (8). According to general procedure F, **8** was synthesized from **31** and **30j**. Eluent ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 62% yield. ¹H NMR (200 MHz, CDCl₃) δ 8.33−8.15 (m, 2H), 7.98 (d, J = 10.6 Hz, 2H), 7.72 (d, J = 8.6 Hz, 1H), 7.55−7.28 (m, 6H), 7.23−7.12 (m, 2H), 7.04 (s, 1H), 6.93 (t, J = 8.6 Hz, 2H), 6.67−6.58 (m, 2H), 3.44 (s, 3H), 2.65 (s, 3H), 2.33 (q, J = 7.7 Hz, 2H), 1.17 (t, J = 7.5 Hz, 3H). ¹³C NMR (101 MHz, DMSO-d_6) δ 172.2, 165.1, 161.0 (d, J = 243.4 Hz), 156.4, 148.2, 143.6, 139.5, 138.9, 137.2, 136.8, 135.9, 132.6, 130.5 (d, J = 2.9 Hz), 128.6, 128.3 (d, J = 7.9 Hz), 128.3, 121.8, 121.7, 121.0, 118.5, 115.2 (d, J = 21.4 Hz), 114.8, 111.1, 31.6, 29.5, 15.3, 9.6. ESI-HRMS: calculated, 580.2057 [M + H]⁺; found, 580.2054. HPLC t_R = 7.55 min.

3-Amino-N-(4-((4-(4-(4-fluorophenyl)-1-methyl-2-(methyl-thio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (9). Compound **9** was synthesized from **S1**, zinc dust (6 equiv), and ammonium formate (6 equiv) in ethanol (62 mL/mmol **S1**). The mixture was heated under reflux for 4 h until complete reduction of the nitro group was detected by TLC. The mixture was filtered through a pad of Celite, which was washed with ethanol. All volatiles were removed under reduced pressure, and the residue was purified using flash chromatography. Eluent ethyl acetate/n-hexane 45:55 to 75:25 afforded 75% yield. ¹H NMR (200 MHz, CDCl₃) δ 8.24 (d, J = 5.7 Hz, 1H), 7.83 (s, 1H), 7.57−7.37 (m, 5H), 7.24−7.10 (m, 6H), 7.05−6.74 (m, 4H), 6.69−6.59 (m, 2H), 3.46 (s, 3H), 2.68 (s, 4H). ¹³C NMR (50 MHz, CDCl₃) δ 165.8, 161.8 (d, J = 245.9 Hz), 156.7, 149.1, 146.9, 144.5, 140.1, 138.6, 136.2, 136, 133.4, 130.1 (d, J = 3.2 Hz), 129.5, 128.9 (d, J = 7.9 Hz), 128.1, 121.4, 121.3, 118.1, 116.2, 115.6, 115.1 (d, J = 21.4 Hz), 113.7, 109, 31.8, 16. ESI-HRMS: calculated, 524.1795 [M + H]⁺; found, 524.1761. HPLC t_R = 4.13 min.

3-Acrylamido-N-(4-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-2-methylbenzamide (10). According to general procedure F, **10** was synthesized from **31** and **30l**. Eluent: ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 12% yield. ¹H NMR (400 MHz, DMSO-d_6) δ 10.25 (s, 1H), 9.65 (s, 1H), 9.13 (s, 1H), 8.25 (d, J = 5.1 Hz, 1H), 7.67−7.52 (m, 5H), 7.51−

7.41 (m, 2H), 7.33−7.24 (m, 2H), 7.15 (t, J = 8.9 Hz, 2H), 6.77−6.71 (m, 2H), 6.61−6.51 (m, 1H), 6.31−6.22 (m, 1H), 5.81−5.74 (m, 1H), 3.43 (s, 3H), 2.65 (s, 3H), 2.24 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.2, 163.4, 161.0 (d, J = 243.3 Hz), 156.5, 148.3, 143.5, 138.9, 138.9, 137.2, 136.8, 136.6, 132.7, 131.6, 130.6 (d, J = 3.1 Hz), 129.4, 128.26 (d, J = 8.0 Hz), 128.3, 126.7, 126.4, 125.5, 124.1, 120.2, 118.5, 115.2 (d, J = 21.5 Hz), 114.8, 111.1, 31.5, 15.3, 14.6. ESI-HRMS: calculated, 592.2057 [M + H]$^+$; found, 592.2049. HPLC t_R = 6.60 min.

N-(4-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-2-methyl-3-propionamidobenzamide (11). According to general procedure F, **11** was synthesized from **31** and **30b**. Eluent ethyl acetate/n-hexane 45:55 to 75:25 afforded 63% yield. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.23 (s, 1H), 9.36 (s, 1H), 9.12 (s, 1H), 8.25 (d, J = 5.5 Hz, 1H), 7.71−7.52 (m, 4H), 7.52−7.39 (m, 3H), 7.30−7.06 (m, 4H), 6.80−6.66 (m, 2H), 3.43 (s, 3H), 2.65 (s, 3H), 2.37 (q, J = 7.3 Hz, 2H), 2.22 (s, 3H), 1.11 (t, J = 7.5 Hz, 3H). ^{13}C NMR (50 MHz, DMSO-d_6) δ 172.1, 167.3, 161.0 (d, J = 243.7 Hz), 156.5, 148.3, 143.5, 138.9 (d, J = 2.9 Hz), 137.1, 137.0, 136.76, 132.8, 130.6, 130.5, 129.5, 128.3, 128.2 (d, J = 8.0 Hz), 126.6, 125.4, 123.8, 120.2, 118.5, 115.1 (d, J = 21.5 Hz), 114.8, 111.0, 31.5, 28.8, 15.3, 14.5, 9.8. ESI-HRMS: calculated, 594.2213 [M + H]$^+$; found, 594.2231. HPLC t_R = 7.27 min.

3-Acrylamido-N-(3-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-benzamide (12). According to general procedure F, **12** was synthesized from **31** and **30c**. Eluent ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92/8 afforded 30% yield. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.37 (s, 1H), 10.24 (s, 1H), 9.26 (s, 1H), 8.26 (d, J = 5.2 Hz, 1H), 8.15 (s, 1H), 8.10 (s, 1H), 7.92 (d, J = 8.0 Hz, 1H), 7.65 (d, J = 7.4 Hz, 1H), 7.54−7.41 (m, 4H), 7.29−7.10 (m, 4H), 6.81 (s, 1H), 6.77 (d, J = 5.2 Hz, 1H), 6.52−6.41 (m, 1H), 6.30 (d, J = 16.9 Hz, 1H), 5.79 (d, J = 10.6 Hz, 1H), 3.44 (s, 3H), 2.65 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.5, 163.3, 161.1 (d, J = 243.6 Hz), 156.3, 148, 143.7, 141.4, 139.4, 139.1, 136.8, 135.9, 131.7, 130.4 (d, J = 2.6 Hz), 128.75, 128.62, 128.33 (d, J = 8.0 Hz), 128.2, 127.2, 122.5, 122.2, 118.9, 115.2 (d, J = 21.5 Hz), 115.2, 114.0, 113.5, 111.6, 110.8, 31.6, 15.3. ESI-HRMS: calculated, 578.1900 [M + H]$^+$; found, 578.1885. HPLC t_R = 6.38 min.

N-(3-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-propionamido-benzamide (13). According to general procedure F, **13** was synthesized from **31** and **30h**. Eluent ethyl acetate/n-hexane 45:55 to 75:25 afforded 62% yield. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.21 (s, 1H), 10.07 (s, 1H), 9.21 (s, 1H), 8.27 (d, J = 5.2 Hz, 1H), 8.09 (d, J = 9.2 Hz, 2H), 7.85−7.80 (m, 1H), 7.63−7.58 (m, 1H), 7.56−7.50 (m, 1H), 7.49−7.40 (m, 3H), 7.27−7.18 (m, 2H), 7.14 (t, J = 8.9 Hz, 2H), 6.81 (s, 1H), 6.78−6.73 (m, 1H), 3.44 (s, 3H), 2.65 (s, 3H), 2.35 (q, J = 7.5 Hz, 2H), 1.10 (t, J = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 172.1, 165.5, 161 (d, J = 243.5 Hz), 156.4, 148.2, 143.5, 141.5, 139.4, 139.3, 138.9, 136.8, 135.8, 130.5 (d, J = 3.0 Hz), 128.5, 128.2 (d, J = 7.8 Hz), 121.8, 118.6, 115.1, 115.1 (d, J = 21.4 Hz), 113.9, 113.4, 111.5, 110.6, 31.5, 29.5, 15.3, 9.5. ESI-HRMS: calculated, 580.2057 [M + H]$^+$; found, 580.2056. HPLC t_R = 7.19 min.

4-Acrylamido-N-(4-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-benzamide (14). According to general procedure F, **14** was synthesized from **31** and **30e**. Eluent ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 32% yield. ^{1}H NMR (400 MHz, CDCl$_3$/methanol-d_4) δ 8.13 (d, J = 5.2 Hz, 1H), 7.86 (d, J = 8.6 Hz, 2H), 7.72 (d, J = 8.6 Hz, 2H), 7.54 (d, J = 8.6 Hz, 2H), 7.40−7.34 (m, 2H), 7.20 (d, J = 8.7 Hz, 2H), 6.97 (t, J = 8.7 Hz, 2H), 6.66 (s, 1H), 6.60 (d, J = 5.3 Hz, 1H), 6.43−6.36 (m, 2H), 5.75 (dd, J = 7.5, 4.2 Hz, 1H), 3.53 (s, 3H), 2.60 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$/methanol-d_4) δ 166.9, 165.4, 162.6 (d, J = 246.4 Hz), 157.4, 148.9, 145.1, 142.0, 140.3, 139.2, 136.9, 134.1, 131.3, 130.6, 130.1 (d, J = 3.2 Hz), 129.8 (d, J = 8.0 Hz), 129, 128.7, 128.1, 122.4, 121.5, 120, 115.6, 115.6 (d, J = 21.6 Hz), 110.3, 32.4, 16.7. ESI-HRMS: calculated, 578.1900 [M + H]$^+$; found, 578.1894. HPLC t_R = 5.70 min.

N-(4-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-4-propionamido-benzamide (15). According to general procedure F, **15** was synthesized from **31** and **30k**. Eluent ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 48% yield. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.15 (s, 1H), 10.01 (s, 1H), 9.13 (s, 1H), 8.26 (d, J = 5.2 Hz, 1H), 7.92 (d, J = 8.7 Hz, 2H), 7.73 (d, J = 8.7 Hz, 2H), 7.67−7.57 (m, 4H), 7.51−7.43 (m, 2H), 7.15 (t, J = 8.9 Hz, 2H), 6.77−6.71 (m, 2H), 3.43 (s, 3H), 2.65 (s, 3H), 2.37 (q, J = 7.5 Hz, 2H), 1.10 (t, J = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 172.3, 164.4, 161.0 (d, J = 243.5 Hz), 156.5, 148.3, 143.5, 142.1, 138.9, 137.1, 136.8, 132.7, 130.6 (d, J = 3.1 Hz), 129.1, 128.4, 128.3, 128.2 (d, J = 7.9 Hz), 121.0, 118.4, 118.1, 115.1 (d, J = 21.5 Hz), 114.8, 111.1, 31.5, 29.5, 15.2, 9.4. ESI-HRMS: calculated, 580.2057 [M + H]$^+$; found, 580.2054. HPLC t_R = 7.04 min.

N-(4-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-propionamido-benzamide (16). According to general procedure F, **16** was synthesized from **31** and **30f**. Eluent: ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 37% yield. ^{1}H NMR (200 MHz, CDCl$_3$/methanol-d_4) δ 8.46 (s, 1H), 8.30−8.17 (m, 2H), 7.83−7.69 (m, 2H), 7.66−7.49 (m, 3H), 7.45−7.27 (m, 3H), 7.14−7.06 (m, 1H), 6.92 (t, J = 8.8 Hz, 2H), 6.62 (d, J = 5.1 Hz, 1H), 6.51−6.19 (m, 4H), 5.79−5.67 (m, 1H), 3.45 (s, 3H), 2.63 (s, 3H), 2.16 (s, 3H). ^{13}C NMR (50 MHz, CDCl$_3$/methanol-d_4) δ 166.4, 164.8, 162 (d, J = 246.3 Hz), 157.9, 148.4, 144.5, 141.5, 140.1, 138.7, 135.8, 133.8, 133.5, 130.8, 130, 129.5 (d, J = 3.2 Hz), 129.2 (d, J = 8.1 Hz), 128.4, 128.2, 127.8, 125.0, 123.2, 119.4, 119.2, 115.0 (d, J = 21.5 Hz), 114.7, 108.4, 32, 17.7, 16.2. ESI-HRMS: calculated, 592.2057 [M + H]$^+$; found, 592.2068. HPLC t_R = 6.90 min.

N-(4-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)-3-methylphenyl)-4-propionamidobenzamide (17). According to general procedure F, **17** was synthesized from **31** and **30i**. Eluent ethyl acetate/n-hexane 45:55 to 75:25 afforded 72% yield. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.23 (s, 1H), 9.36 (s, 1H), 9.11 (s, 1H), 8.25 (d, J = 5.2 Hz, 1H), 7.68−7.53 (m, 4H), 7.52−7.39 (m, 3H), 7.33−7.03 (m, 4H), 6.79−6.65 (m, 2H), 3.43 (s, 3H), 2.65 (s, 3H), 2.37 (q, J = 7.5 Hz, 2H), 1.11 (t, J = 7.5 Hz, 3H). ^{13}C NMR (50 MHz, CDCl$_3$/methanol-d_4) δ 174.5, 169, 162.1 (d, J = 246.4 Hz), 156.9, 148.6, 144.7, 140, 138.8, 138.6, 136.2, 135.9, 133.9, 130.9, 129.7 (d, J = 3.2 Hz), 129.3 (d, J = 8.0 Hz), 128.4 (d, J = 0.7 Hz), 127.6, 126.2, 125.1, 121.4, 121.0, 115.2, 115.2 (d, J = 21.5 Hz), 109.4, 32.1, 29.6, 16.3, 14.6, 9.8. ESI-HRMS: calculated, 594.2213 [M + H]$^+$; found, 594.2213. HPLC t_R = 6.68 min.

4-Acrylamido-N-(3-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-benzamide (18). According to general procedure F, **18** was synthesized from **31** and **30g**. Eluent: ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 26% yield. ^{1}H NMR (200 MHz, CDCl$_3$/methanol-d_4) δ 7.93 (d, J = 5.3 Hz, 1H), 7.70−7.43 (m, 5H), 7.21−7.10 (m, 2H), 7.07−6.89 (m, 2H), 6.89−6.79 (m, 1H), 6.72 (t, J = 8.8 Hz, 2H), 6.54 (s, 1H), 6.39 (d, J = 5.2 Hz, 1H), 6.17 (d, J = 5.6 Hz, 2H), 5.58−5.47 (m, 1H), 3.33 (s, 3H), 2.38 (s, 3H). ^{13}C NMR (50 MHz, CDCl$_3$/methanol-d_4) 166.4, 164.8, 161.9 (d, J = 246.3 Hz), 156.5, 148.1, 144.4, 141.5, 140.5, 139.6, 138.9, 138.5, 130.7, 129.8, 129.4 (d, J = 3.2 Hz), 129.1 (d, J = 8.0 Hz), 129.0, 128.3, 128.1, 127.5, 119.2, 115.7, 115.3, 115, 114.9 (d, J = 21.6 Hz), 112.5, 110.4, 31.8, 16.1. ESI-HRMS: calculated, 578.1900 [M + H]$^+$; found, 578.1900. HPLC t_R = 5.95 min.

3-Acrylamido-N-(3-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-2-methylbenzamide (19). According to general procedure F, **19** was synthesized from **31** and **30d**. Eluent ethyl acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 30% yield. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.35 (s, 1H), 9.65 (s, 1H), 9.21 (s, 1H), 8.25 (d, J = 5.2 Hz, 1H), 8.07 (s, 1H), 7.62−7.50

(m, 2H), 7.50−7.42 (m, 2H), 7.33−7.25 (m, 2H), 7.24−7.18 (m, 2H), 7.18−7.09 (m, 2H), 6.80 (s, 1H), 6.75 (d, J = 5.2 Hz, 1H), 6.62−6.50 (m, 1H), 6.26 (d, J = 17.1 Hz, 1H), 5.77 (d, J = 10.3 Hz, 1H), 3.43 (s, 3H), 2.69−2.60 (m, 3H), 2.23 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 167.7, 163.5, 161.0 (d, J = 243.6 Hz), 156.5, 148.2, 143.6, 141.6, 139.4, 138.9, 136.8, 136.6, 130.6 (d, J = 3.0 Hz), 130.5, 129.4, 128.7, 128.3 (d, J = 7.8 Hz), 128.2, 126.7, 126.5, 125.6, 124.1, 115.2, 115.1 (d, J = 21.4 Hz), 113.7, 112.6, 111.5, 109.8, 31.6, 15.3, 14.6. The missing signal is most probably overlaid by another signal. ESI-HRMS: calculated, 592.2057 [M + H]$^+$; found, 592.2045. HPLC t_R = 6.32 min.

3-Acrylamido-N-(4-((4-(4-(4-fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)-3-methylphenyl)benzamide (20). According to general procedure F, **20** was synthesized from **31** and **30m**. Eluent acetate/n-hexane 45:55 to 75:25. Another flash chromatography was performed to purify the product; eluent dichloromethane/methanol 100:0 to 92:8 afforded 43% yield. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.36 (s, 1H), 10.16 (s, 1H), 8.31 (s, 1H), 8.18−8.12 (m, 2H), 7.92 (d, J = 8.0 Hz, 1H), 7.67−7.58 (m, 2H), 7.52−7.37 (m, 5H), 7.18−7.10 (m, 2H), 6.69 (dd, J = 5.2, 1.2 Hz, 1H), 6.55−6.42 (m, 2H), 6.30 (dd, J = 17.0, 1.9 Hz, 1H), 5.82−5.74 (m, 1H), 3.42 (s, 3H), 2.63 (s, 3H), 2.17 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.1, 163.3, 161.0 (d, J = 243.6 Hz), 157.8, 148.5, 143.4, 139.1, 139.1, 136.8, 135.9, 135.1, 134.3, 131.9, 131.7, 130.6 (d, J = 3.0 Hz), 128.7, 128.4, 128.3 (d, J = 8.0 Hz), 127.1, 124.2, 122.5, 122.3, 122.1, 118.8, 118.4, 115.1 (d, J = 21.4 Hz), 114.3, 109.4, 31.5, 18.2, 15.4, 15.3. ESI-HRMS: calculated, 592.2057 [M + H]$^+$; found, 592.2064. HPLC t_R = 6.60 min.

3-Acrylamido-N-(4-((4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (21). Compound **21** was prepared according to general procedure B starting from **35** (87.5 mg, 0.20 mmol), DIPEA (46 μL, 0.26 mmol), and acryloyl chloride (22 μL, 0.24 mmol). Purification: flash column chromatography (SiO$_2$, dichloromethane/ethanol, 92:08). Yield: 40 mg (41%) as a light yellow solid product. ^{1}H NMR (250 MHz, DMSO-d_6) δ 12.28 (s, 1H), 10.35 (s, 1H), 10.13 (s, 1H), 8.99 (s, 1H), 8.14 (s, 1H), 8.08 (d, J = 5.4 Hz, 1H), 7.92 (d, J = 8.3 Hz, 1H), 7.71−7.59 (m, 5H), 7.52−7.38 (m, 1H), 7.15 (s, 1H), 6.98 (d, J = 5.6 Hz, 1H), 6.56−6.38 (m, 1H), 6.36−6.20 (m, 1H), 5.86−5.70 (m, 1H), 2.55 (s, 3H), 2.42 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 164.9, 163.3, 156.3, 147.2, 143.1, 139.1, 138.2, 137.0, 135.9, 134.0, 131.8, 131.7, 128.7, 127.1, 122.2, 122.0, 121.1, 118.8, 118.3, 117.9, 111.0, 105.9, 15.5, 11.4. HPLC t_R = 5.29 min. ESI-HRMS: calculated, 485.1754 [M + H]$^+$; found, 485.1738.

4-Acrylamido-N-(4-((4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (22). Compound **22** was synthesized according to general procedure B, using compound **36** (100 mg, 0.23 mmol), DIPEA (53 μL, 0.3 mmol), and acryloyl chloride (23 μL, 0.28 mmol). Purification by flash chromatography (SiO$_2$, dichloromethane/ethanol, 98:08) afforded 20 mg (22%) of a light yellow solid product. ^{1}H NMR (250 MHz, DMSO-d_6) δ 12.29 (br s, 1H), 10.42 (s, 1H), 10.01 (s, 1H), 8.98 (s, 1H), 8.09 (d, J = 5.4 Hz, 1H), 7.99−7.92 (m, J = 8.8 Hz, 2H), 7.84−7.76 (m, J = 9.0 Hz, 2H), 7.70−7.59 (m, 4H), 7.15 (s, 1H), 6.98 (d, J = 5.1 Hz, 1H), 6.48 (dd, J = 9.9, 17.0 Hz, 1H), 6.30 (dd, J = 2.2, 17.1 Hz, 1H), 5.76−5.85 (m, 1H), 2.55 (s, 3H), 2.42 (s, 3H). HPLC t_R = 2.66 min. ESI-HRMS: calculated, 485.1754 [M + H]$^+$; found, 485.1752.

N-(4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-propionamidobenzamide (23). Under argon atmosphere, compound **32** (80 mg, 0.26 mmol) was suspended in a mixture of dry dichloromethane (10 mL) and DIPEA (0.054 mL, 0.31 mmol) was added. Finally, a solution of 3-propionamidobenzoic acid (54 mg, 0.28 mmol) and PyBOP (162 mg, 0.31 mmol) in (5:1) mixture of dichloromethane/DMF (6 mL) was added. The reaction was stirred at rt for 6 h. The reaction mixture was poured in water, and the aqueous layer was extracted 3 times with ethyl acetate. The organic layers were washed with saturated NaCl solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. The residue was finally purified twice by flash column chromatography (SiO$_2$, dichloromethane/ethanol, 92:08 to 85:15) yielding 27 mg of product (21%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.15 (s, 1H),

10.06 (s, 1H), 9.17 (br s, 1H), 8.10 (br s, 1H), 8.07 (d, J = 5.6 Hz, 1H), 7.83 (d, J = 7.9 Hz, 1H), 7.73−7.58 (m, 5H), 7.49−7.40 (m, 1H), 7.15 (s, 1H), 7.01 (d, J = 5.4 Hz, 1H), 2.57 (s, 3H), 2.43 (s, 3H), 2.36 (q, J = 7.5 Hz, 2H), 1.11 (t, J = 7.4 Hz, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 172.2, 165.1, 155.8, 146.0, 143.2, 142.9, 139.8, 139.4, 137.3, 135.8, 132.6, 128.6, 121.7, 121.1, 118.8, 118.5, 113.8, 113.4, 111.0, 105.9, 29.5, 15.4, 9.6, 8.6. HPLC t_R = 4.12 min. ESI-HRMS: calculated, 487.1911 [M + H]$^+$; found, 487.1896.

N-(4-((4-(4-(4-Fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-nitrobenzamide (24). In 100 mL round-bottom flask, m-nitrobenzoic acid (341.5 mg, 2.00 mmol), PyBOP (1.28 g, 2.40 mmol), and DIPEA (463 μL, 2.60 mmol) were mixed together in 30 mL of dichloromethane. The mixture was stirred for 1 h at rt. Then, amine **3** (800 mg, 2.00 mmol) was added to the reaction mixture in one portion and the reaction stirred for 5 h at rt. The solvent was evaporated under vacuum, then ethyl acetate (50 mL) was added and washed with water (3 × 15 mL). The combined water layers were extracted with ethyl acetate one time. The combined organic layers were dried over anhydrous Na$_2$SO$_4$. The solvent was evaporated and the crude product was purified by flash chromatography (SiO$_2$, dichloromethane/ethanol, 95:05 to 90:10) to afford 462 mg (42%) of an orange solid product. ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.69 (br s, 1H), 10.48 (br s, 1H), 9.05 (br s, 1H), 8.80 (br s, 1H), 8.50−8.30 (m, 2H), 8.01 (br s, 1H), 7.89−7.75 (m, 1H), 7.64 (br s, 4H), 7.52 (br s, 2H), 7.30 (br s, 2H), 7.13−6.75 (m, 1H), 6.70 (br s, 1H), 2.63 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 162.7, 162.2 (d, J = 249.5 Hz), 156.2, 147.7, 147.1, 142.9, 142.1, 139.6, 138.2, 136.5, 134.0, 131.6 (d, J = 4.6 Hz), 130.6 (d, J = 7.3 Hz), 130.1, 129.6, 128.2, 125.9, 122.3, 121.3, 118.3, 115.8 (d, J = 24.8 Hz), 112.0, 107.3, 15.0. HPLC t_R = 6.87 min. FAB-MS: (m/z) calculated, 541.15 [M + H]$^+$; found: 541.2.

N-(4-((4-(4-(4-Fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-4-nitrobenzamide (25). Compound **25** was synthesized similarly to **24** using p-nitrobenzoic acid (310 mg, 1.85 mmol), PyBOP (1.1 g, 2.02 mmol) and DIPEA (353 μL, 2.02 mmol) and amine **3** (660 mg, 1.85 mmol) affording 637 mg (70%) of an orange solid product. ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.69 (br s, 1H), 10.46 (br s, 1H), 9.06 (br s, 1H), 8.46−8.29 (m, 2H), 8.26−8.13 (m, 2H), 8.03 (br s, 1H), 7.75−7.57 (m, 4H), 7.52 (br s, 2H), 7.28 (br s, 2H), 7.07 (br s, 1H), 6.76−6.62 (m, 1H), 2.63 (d, J = 2.3 Hz, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.2, 161.7 (d, J = 248.9 Hz), 156.2, 149.0, 142.2, 140.8, 138.2, 131.7 (d, J = 1.8 Hz), 129.1 (d, J = 2.4 Hz), 129.0, 123.6, 123.4, 122.5, 121.4, 121.2, 120.8, 119.0, 118.3, 115.7 (d, J = 22.2 Hz), 112.1, 107.3, 15.0. HPLC t_R = 6.80 min. FAB-MS: (m/z) calculated, 541.15 [M + H]$^+$; found: 541.2.

3-Amino-N-(4-((4-(4-(4-fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (26). Compound **26** was synthesized according to general procedure A, using compound **24** (160 mg, 0.30 mmol), EtOH (20 mL), and tin(II) chloride (167 mg, 0.74 mmol). Yield: 88 mg (58%) of a light yellow solid product. ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.67 (br s, 1H), 10.03−9.83 (m, 1H), 9.08−8.84 (m, 1H), 8.23−7.90 (m, 1H), 7.61 (d, J = 6.1 Hz, 2H), 7.58−7.44 (m, 4H), 7.37−7.17 (m, 2H), 7.17−6.99 (m, 4H), 6.84−6.72 (m, 1H), 6.72−6.63 (m, 1H), 5.29 (br s, 2H), 2.63 (br s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.8, 161.9 (d, J = 247.4 Hz), 161.3 (d, J = 241.8 Hz), 156.4, 148.6, 148.0, 147.2, 142.8, 142.0, 138.9, 137.9, 137.6, 137.2, 136.1, 134.7, 132.8, 132.4, 131.0, 130.7 (d, J = 8.1 Hz), 129.4 (d, J = 10.3 Hz), 128.6, 126.8, 126.8, 126.2, 120.9, 118.7, 118.3, 116.6, 115.9 (d, J = 22.3 Hz), 115.1 (d, J = 21.2 Hz), 114.7, 113.0, 112.3, 111.8, 107.2, 107.0, 15.1, 15.0. The peaks of the carbon are double due to the amide tautomerism. HPLC t_R = 5.32 min. FAB-MS: (m/z) calculated, 511.2 [M + H]$^+$; found: 511.3.

4-Amino-N-(4-((4-(4-(4-fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (27). Compound **27** was synthesized according to general procedure A, using compound **25** (352 mg, 0.65 mmol), EtOH (70 mL), tin(II) chloride (368 mg, 1.6 mmol). Yield: 183 mg (55%) of a light yellow solid product. ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.67 (br s, 1H), 9.65 (br s, 1H), 9.02−8.85 (m, 1H), 8.20−7.88 (m, 1H), 7.72 (d, J = 7.3 Hz,

2H), 7.56−7.55 (m, 1H), 7.64−7.41 (m, 6H), 7.39−7.11 (m, 2H), 7.11−6.73 (m, 1H), 6.67 (br s, 1H), 6.60 (d, J = 7.3 Hz, 2H), 5.72 (br s, 2H), 2.62 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 164.8, 161.9 (d, J = 241.0 Hz), 156.4, 151.8, 148.0, 147.1, 142.8, 141.9, 137.2, 132.8, 130.6 (d, J = 7.0 Hz), 129.1, 126.8 (d, J = 3.8 Hz), 121.4, 120.8, 118.7, 118.3, 115.8 (d, J = 21.3 Hz), 112.5, 111.7, 106.9, 15.0. HPLC t_R = 4.86 min. FAB-MS: (m/z) calculated, 511.2 [M + H]$^+$; found: 511.3.

N-(4-Bromophenyl)-3-nitrobenzamide (28a). According to general procedure C, **28a** was synthesized from 4-bromoaniline and 3-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 56% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.68 (s, 1H), 8.78 (s, 1H), 8.51−8.33 (m, 2H), 7.91−7.70 (m, 3H), 7.56 (d, J = 8.8 Hz, 2H). Exact mass: 320.0. ESI-MS (m/z): 319.4, 321.4 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromophenyl)-2-methyl-3-nitrobenzamide (28b). According to general procedure C, **28b** was synthesized from 4-bromoaniline and 2-methyl-3-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 89% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.71 (s, 1H), 8.01 (d, J = 7.5 Hz, 1H), 7.89−7.45 (m, 6H), 2.41 (s, 3H). Exact mass: 334.0. ESI-MS (m/z): 332.8, 334.8 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(3-Bromophenyl)-3-nitrobenzamide (28c). According to general procedure C, **28c** was synthesized from 3-bromoaniline and 3-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 79% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.70 (s, 1H), 8.79 (t, J = 1.9 Hz, 1H), 8.50−8.33 (m, 2H), 8.14−8.05 (m, 1H), 7.91−7.70 (m, 2H), 7.42−7.28 (m, 2H). Exact mass: 320.0. ESI-MS (m/z): 318.8, 320.8 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(3-Bromophenyl)-2-methyl-3-nitrobenzamide (28d). According to general procedure C, **28d** was synthesized from 3-bromoaniline and 2-methyl-3-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 67% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.73 (s, 1H), 8.11−7.97 (m, 2H), 7.85−7.75 (m, 1H), 7.70−7.50 (m, 2H), 7.40−7.26 (m, 2H), 2.44 (s, 3H). Exact mass: 334.0. ESI-MS (m/z): 333.4, 335.4 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromophenyl)-4-nitrobenzamide (28e). According to general procedure C, **28e** was synthesized from 4-bromoaniline and 4-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 89% yield as a yellow solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.66 (s, 1H), 8.37 (d, J = 8.9 Hz, 2H), 8.17 (d, J = 8.8 Hz, 2H), 7.77 (d, J = 8.9 Hz, 2H), 7.56 (d, J = 8.9 Hz, 2H). Exact mass: 320.0. ESI-MS (m/z): 319.4, 321.4 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromo-3-methylphenyl)-4-nitrobenzamide (28f). According to general procedure C, **28f** was synthesized from 4-bromo-3-methylaniline and 4-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 89% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.59 (s, 1H), 8.36 (d, J = 8.8 Hz, 2H), 8.17 (d, J = 8.8 Hz, 2H), 7.78 (s, 1H), 7.56 (s, 2H), 2.35 (s, 3H). Exact mass: 334.0. ESI-MS (m/z): 333.4, 335.4 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(3-Bromophenyl)-4-nitrobenzamide (28g). According to general procedure C, **28g** was synthesized from 3-bromoaniline and 4-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 70% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.68 (s, 1H), 8.37 (d, J = 8.8 Hz, 2H), 8.24−8.05 (m, 3H), 7.82−7.67 (m, 1H), 7.41−7.26 (m, 2H). Exact mass: 320.0. ESI-MS (m/z): 319.4, 321.4 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromo-3-methylphenyl)-3-nitrobenzamide (28h). According to general procedure C, **28h** was synthesized from 4-bromo-3-methylaniline and 3-nitrobenzoic acid. All volatiles were evaporated under reduced pressure to afford 72% yield as an off-white solid. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.60 (s, 1H), 8.82−8.74 (m, 1H),

8.47−8.33 (m, 2H), 7.89−7.73 (m, 2H), 7.64−7.49 (m, 2H), 2.35 (s, 3H). Exact mass: 334.0. ESI-MS (m/z): 333.3, 335.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Amino-N-(4-bromophenyl)benzamide (29a). According to general procedure D, **29a** was synthesized from **28a**. Yield: 96%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.18 (s, 1H), 7.75 (d, J = 8.9 Hz, 2H), 7.51 (d, J = 8.8 Hz, 2H), 7.23−6.97 (m, 3H), 6.75 (d, J = 7.1 Hz, 1H), 5.33 (s, 2H). Exact mass: 290.0. ESI-MS (m/z): 289.3, 291.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Amino-N-(4-bromophenyl)-2-methylbenzamide (29b). According to general procedure D, **29b** was synthesized from **28b**. Yield: 93%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.33 (s, 1H), 7.72 (d, J = 8.7 Hz, 1H), 7.50 (d, J = 8.6 Hz, 1H), 6.98 (t, J = 7.7 Hz, 1H), 6.73 (d, J = 7.9 Hz, 1H), 6.61 (d, J = 7.3 Hz, 1H), 5.07 (s, 1H), 2.05 (s, 2H). Exact mass: 304.0. ESI-MS (m/z): 303.3, 305.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Amino-N-(3-bromophenyl)benzamide (29c). According to general procedure D, **29c** was synthesized from **28c**. Yield: 91%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.21 (s, 1H), 8.10 (s, 1H), 7.82−7.65 (m, 1H), 7.37−7.00 (m, 5H), 6.76 (d, J = 8.0 Hz, 1H), 5.32 (s, 2H). Exact mass: 290.0. ESI-MS (m/z): 289.3, 291.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Amino-N-(3-bromophenyl)-2-methylbenzamide (29d). According to general procedure D, **29d** was synthesized from **28d**. Yield: 57%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.38 (s, 1H), 8.11 (s, 1H), 7.68 (d, J = 6.4 Hz, 1H), 7.35−7.18 (m, 2H), 6.99 (t, J = 7.7 Hz, 1H), 6.75 (d, J = 7.7 Hz, 1H), 6.63 (d, J = 7.3 Hz, 1H), 5.12 (s, 2H), 2.08 (s, 3H). Exact mass: 304.0. ESI-MS (m/z): 303.4, 304.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

4-Amino-N-(4-bromophenyl)benzamide (29e). According to general procedure D, **29e** was synthesized from **28e**. Yield: 34%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 9.88 (s, 1H), 7.79−7.65 (m, 4H), 7.48 (d, J = 8.9 Hz, 2H), 6.60 (d, J = 8.6 Hz, 2H), 5.79 (s, 2H). Exact mass: 290.0. ESI-MS (m/z): 289.4, 291.5 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

4-Amino-N-(4-bromo-3-methylphenyl)benzamide (29f). According to general procedure D, **29f** was synthesized from **28f**. Yield: 80%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 9.81 (s, 1H), 7.85−7.65 (m, 3H), 7.63−7.39 (m, 2H), 6.61 (d, J = 8.6 Hz, 2H), 5.78 (s, 2H), 2.32 (s, 3H). Exact mass: 304.0. ESI-MS (m/z): 303.3, 305.4 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

4-Amino-N-(3-bromophenyl)benzamide (29g). According to general procedure D, **29g** was synthesized from **28g**. Yield: 50%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 9.89 (s, 1H), 8.10 (s, 1H), 7.78−7.64 (m, 3H), 7.35−7.14 (m, 2H), 6.61 (d, J = 8.6 Hz, 2H), 5.81 (s, 2H). Exact mass: 290.0. ESI-MS (m/z): 289.3, 291.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Amino-N-(4-bromo-3-methylphenyl)benzamide (29h). According to general procedure D, **29h** was synthesized from **28h**. Yield: 78%. ^{1}H NMR (200 MHz, CDCl$_3$) δ 8.00 (s, 1H), 7.54 (d, J = 2.0 Hz, 1H), 7.44 (d, J = 8.6 Hz, 1H), 7.35−7.07 (m, 4H), 6.84−6.73 (m, 1H), 3.78 (s, 2H), 2.35 (s, 3H). Exact mass: 304.0. ESI-MS (m/z): 303.3, 305.3 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Acrylamido-N-(4-bromophenyl)benzamide (30a). According to general procedure E, **30a** was synthesized from **29a** and acryloyl chloride. Yield: 51%. ^{1}H NMR (200 MHz, CDCl$_3$) δ 8.06−7.96 (m, 1H), 7.90 (d, J = 7.4 Hz, 1H), 7.71−7.51 (m, 4H), 7.51−7.40 (m, 2H), 6.67−6.44 (m, 2H), 6.39−5.98 (m, 2H), 5.80 (dd, J = 10.3, 1.6 Hz, 1H). Exact mass: 344.0. ESI-MS (m/z): 343.5, 345.5 [M − H]$^-$ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromophenyl)-2-methyl-3-propionamidobenzamide (30b). According to general procedure E, **30b** was synthesized from **29b** and propionyl chloride. Yield: 77%. ^{1}H NMR (200 MHz, DMSO-d_6) δ 10.51 (s, 1H), 9.39 (s, 1H), 7.74 (d, J = 8.5 Hz, 2H), 7.63−7.38 (m, 3H), 7.26 (d, J = 4.4 Hz, 2H), 2.37 (q, J = 7.1 Hz, 2H), 2.21 (s,

3H), 1.11 (t, *J* = 7.5 Hz, 3H). Exact mass: 360.0. ESI-MS (*m/z*): 358.8, 360.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Acrylamido-*N*-(3-bromophenyl)benzamide (30c). According to general procedure E, **30c** was synthesized from **29c** and acryloyl chloride. Yield: 48%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.40 (d, *J* = 10.5 Hz, 2H), 8.20−8.08 (m, 2H), 7.99−7.90 (m, 1H), 7.81−7.61 (m, 2H), 7.50 (t, *J* = 7.8 Hz, 1H), 7.39−7.25 (m, 2H), 6.56−6.22 (m, 2H), 5.79 (dd, *J* = 9.6, 2.5 Hz, 1H). Exact mass: 344.0. ESI-MS (*m/z*): 343.4, 345.4 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Acrylamido-*N*-(3-bromophenyl)-2-methylbenzamide (30d). According to general procedure E, **30d** was synthesized from **29d** and acryloyl chloride. Yield: 77%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.57 (s, 1H), 9.67 (s, 1H), 8.11 (s, 1H), 7.79−7.47 (m, 2H), 7.40−7.21 (m, 4H), 6.68−6.44 (m, 1H), 6.27 (d, *J* = 16.9 Hz, 1H), 5.78 (d, *J* = 9.8 Hz, 1H), 2.23 (s, 3H). Exact mass: 358.0. ESI-MS (*m/z*): 357.5, 359.5 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

4-Acrylamido-*N*-(4-bromophenyl)benzamide (30e). According to general procedure E, **30e** was synthesized from **29e** and acryloyl chloride. Yield: 60%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.44 (s, 1H), 10.26 (s, 1H), 7.96 (d, *J* = 8.7 Hz, 2H), 7.87−7.68 (m, 4H), 7.53 (d, *J* = 8.9 Hz, 2H), 6.59−6.18 (m, 2H), 5.81 (dd, *J* = 9.7, 2.3 Hz, 1H). Exact mass: 344.0. ESI-MS (*m/z*): 343.4, 345.4 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

4-Acrylamido-*N*-(4-bromo-3-methylphenyl)benzamide (30f). According to general procedure E, **30f** was synthesized from **29f** and acryloyl chloride. Yield: 68%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.44 (s, 1H), 10.19 (s, 1H), 7.95 (d, *J* = 8.8 Hz, 2H), 7.81 (d, *J* = 8.7 Hz, 3H), 7.64−7.46 (m, 2H), 6.57−6.21 (m, 2H), 5.81 (dd, *J* = 9.7, 2.3 Hz, 1H), 2.34 (s, 3H). Exact mass: 358.0. ESI-MS (*m/z*): 356.8, 358.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

4-Acrylamido-*N*-(3-bromophenyl)benzamide (30g). According to general procedure E, **30g** was synthesized from **29g** and acryloyl chloride. Trituration with cold methanol afforded 74% yield. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.45 (s, 1H), 10.28 (s, 1H), 8.11 (s, 1H), 8.01−7.89 (m, 2H), 7.88−7.69 (m, 3H), 7.38−7.21 (m, 2H), 6.58−6.22 (m, 2H), 5.81 (dd, *J* = 9.6, 1.9 Hz, 1H). Exact mass: 344.0. ESI-MS (*m/z*): 342.8, 344.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(3-Bromophenyl)-3-propionamidobenzamide (30h). According to general procedure E, **30h** was synthesized from **29c** and propionyl chloride. Yield: 68%. ¹H NMR (400 MHz, DMSO-*d₆*) δ 10.40 (s, 1H), 10.08 (s, 1H), 8.10 (s, 2H), 7.84 (d, *J* = 8.0 Hz, 1H), 7.77−7.71 (m, 1H), 7.60 (d, *J* = 7.8 Hz, 1H), 7.45 (t, *J* = 7.9 Hz, 1H), 7.37−7.26 (m, 2H), 2.35 (q, *J* = 7.5 Hz, 2H), 1.10 (t, *J* = 7.5 Hz, 3H). Exact mass: 346.0. ESI-MS (*m/z*): 344.8, 346.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromo-3-methylphenyl)-4-propionamidobenzamide (30i). According to general procedure E, **30i** was synthesized from **29f** and propionyl chloride. Yield: 64%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.15 (s, 2H), 7.98−7.84 (m, 2H), 7.83−7.67 (m, 3H), 7.65−7.47 (m, 2H), 2.44−2.26 (m, 5H) (CH₂ and methyl group overlay), 1.09 (t, *J* = 7.5 Hz, 3H). Exact mass: 360.0. ESI-MS (*m/z*): 358.8, 360.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromophenyl)-3-propionamidobenzamide (30j). According to general procedure E, **30j** was synthesized from **29a** and propionyl chloride. Yield: 63%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.37 (s, 1H), 10.07 (s, 1H), 8.10 (s, 1H), 7.94−7.69 (m, 3H), 7.66−7.38 (m, 4H), 2.35 (q, *J* = 7.5 Hz, 2H), 1.10 (t, *J* = 7.5 Hz, 3H). Exact mass: 346.0. ESI-MS (*m/z*): 344.8, 346.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-Bromophenyl)-4-propionamidobenzamide (30k). According to general procedure E, **30k** was synthesized from **29e** and propionyl chloride. Yield: 90%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.19 (d, *J* = 13.5 Hz, 2H), 7.92 (d, *J* = 8.6 Hz, 2H), 7.82−7.68 (m, 4H), 7.52 (d, *J* = 8.5 Hz, 2H), 2.37 (q, *J* = 7.6 Hz, 2H), 1.09 (t, *J* = 7.5

Hz, 3H). Exact mass: 346.0. ESI-MS (*m/z*): 344.8, 346.8 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Acrylamido-*N*-(4-bromophenyl)-2-methylbenzamide (30l). According to general procedure E, **30l** was synthesized from **29b** and acryloyl chloride. Yield: 55%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.55 (s, 1H), 7.73 (d, *J* = 8.5 Hz, 2H), 7.52 (d, *J* = 8.5 Hz, 3H), 7.28 (d, *J* = 4.5 Hz, 2H), 6.76−6.53 (m, 1H), 6.25 (d, *J* = 16.0 Hz, 1H), 5.75 (d, *J* = 11.2 Hz, 1H), 2.22 (s, 3H). Exact mass: 358.0. ESI-MS (*m/z*): 358.9, 360.9 [M + H]⁺ (peaks were measured at half-intensity due to isotopic pattern of bromine).

3-Acrylamido-*N*-(4-bromo-3-methylphenyl)benzamide (30m). According to general procedure E, **30m** was synthesized from **29h** and acryloyl chloride. Yield: 61%. ¹H NMR (200 MHz, DMSO-*d₆*) δ 10.33 (d, *J* = 7.0 Hz, 1H), 8.15 (s, 1H), 7.91 (d, *J* = 7.5 Hz, 1H), 7.77 (s, 1H), 7.68−7.39 (m, 4H), 6.55−6.18 (m, 2H), 5.77 (dd, *J* = 9.8, 1.7 Hz, 1H), 2.32 (s, 3H). Exact mass: 358.0. ESI-MS (*m/z*): 357.5, 359.5 [M − H]⁻ (peaks were measured at half-intensity due to isotopic pattern of bromine).

N-(4-((4-(4-Methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-nitrobenzamide (33). In a round-bottom flask under dry conditions, *m*-nitrobenzoic acid (222.2 mg, 1.33 mmol) was dissolved in 3 mL of thionyl chloride. The reaction mixture was refluxed for 2 h. The solvent was removed under high vacuum and compound **32** (230 mg, 0.74 mmol) dissolved in 5 mL pyridine was added to the reaction mixture and the reaction was stirred at rt for 60 h. The reaction mixture was poured in water (10 mL) finding the formation of a precipitate. The precipitate was filtered off and purified by flash chromatography (SiO₂, dichloromethane/ethanol, 95:05 to 80:20) to afford 135 mg (40%) of a red solid. ¹H NMR (400 MHz, DMSO-*d₆*) δ 12.29 (br s, 1H), 10.48 (s, 1H), 9.05 (s, 1H), 8.80 (br s, 1H), 8.49−8.34 (m, 2H), 8.10 (d, *J* = 5.1 Hz, 1H), 7.91−7.78 (m, 1H), 7.72 (d, *J* = 8.6 Hz, 2H), 7.66 (d, *J* = 8.9 Hz, 2H), 7.18 (s, 1H), 6.99 (d, *J* = 5.3 Hz, 1H), 2.56 (s, 3H), 2.46−2.32 (m, 3H); ¹³C NMR (101 MHz, DMSO-*d₆*) δ 162.7, 156.3, 147.7, 147.2, 143.1, 139.1, 138.6, 136.5, 134.0, 131.2, 130.1, 127.2, 125.9, 122.3, 121.3, 118.3, 117.9, 111.1, 106.1, 15.5, 11.4. HPLC *t*ᴿ = 7.35 min. ESI-MS: (*m/z*) calculated, 461.14 [M + H]⁺; found: 461.1.

N-(4-((4-(4-Methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-4-nitrobenzamide (34). In a round-bottom flask under dry conditions, compound **34** (100 mg, 0.32 mmol) and *p*-nitrobenzoyl chloride (71.5 mg, 0.39 mmol) were stirred in pyridine (3 mL) at rt for 48 h. The reaction mixture was poured in water (10 mL) finding the formation of a precipitate. The precipitate was filtered off and purified by flash chromatography two times (SiO₂, dichloromethane/ethanol, 95:05 to 80:20) to afford 70 mg (47%) of a red solid. ¹H NMR (250 MHz, DMSO-*d₆*) δ 12.31 (br s, 1H), 10.44 (s, 1H), 9.04 (s, 1H), 8.47−8.31 (m, 2H), 8.27−8.14 (m, 2H), 8.10 (d, *J* = 5.4 Hz, 1H), 7.77−7.68 (m, 2H), 7.68−7.55 (m, 2H), 7.15 (br s, 1H), 6.98 (d, *J* = 3.9 Hz, 1H), 2.56 (s, 3H), 2.42 (s, 3H); ¹³C NMR (101 MHz, DMSO-*d₆*) δ 163.2, 156.3, 149.0, 147.2, 140.8, 138.6, 134.0, 131.3, 129.0, 127.5, 127.1, 123.4, 121.2, 117.9, 116.7, 111.1, 106.0, 15.4, 11.4. HPLC *t*ᴿ = 5.73 min. ESI-MS: (*m/z*) calculated, 461.14 [M + H]⁺; found: 461.1.

3-Amino-*N*-(4-((4-(4-methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (35). Compound **35** was synthesized according to general procedure A, using compound **33** (135 mg, 0.29 mmol), EtOH (20 mL), tin(II) chloride (661.5 mg, 2.9 mmol). Yield: 110 mg (87%) of a light yellow solid product. ¹H NMR (250 MHz, DMSO-*d₆*) δ 12.36 (br s, 1H), 9.93 (s, 1H), 9.00 (s, 1H), 8.08 (d, *J* = 5.6 Hz, 1H), 7.73−7.55 (m, 4H), 7.20−7.02 (m, 4H), 6.97 (d, *J* = 4.9 Hz, 1H), 6.79−6.63 (m, 1H), 5.29 (s, 2H), 2.55 (s, 3H), 2.42 (s, 3H); ¹³C NMR (101 MHz, DMSO-*d₆*) δ 165.7, 156.4, 148.6, 147.1, 143.0, 139.0, 137.9, 136.1, 134.0, 132.1, 128.6, 127.1, 120.9, 117.9, 116.5, 114.6, 112.9, 111.0, 105.9, 15.5, 11.4. HPLC *t*ᴿ = 3.99 min. ESI-MS: (*m/z*) calculated, 431.17 [M + H]⁺; found, 431.2.

4-Amino-*N*-(4-((4-(4-methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)amino)phenyl)benzamide (36). Compound **36** was synthesized according to general procedure A, using compound **34** (135 mg, 0.29 mmol), EtOH (20 mL), tin(II) chloride (661.5 mg,

2.9 mmol). Yield: 120 mg (95%) of a light yellow solid product. ^{1}H NMR (250 MHz, DMSO-d_6) δ 12.27 (br s, 1H), 9.62 (s, 1H), 8.92 (s, 1H), 8.07 (d, J = 5.4 Hz, 1H), 7.71 (d, J = 8.3 Hz, 2H), 7.61 (s, 4H), 7.14 (s, 1H), 6.96 (d, J = 5.6 Hz, 1H), 6.59 (d, J = 8.3 Hz, 2H), 5.70 (s, 2H), 2.55 (s, 3H), 2.42 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 164.8, 156.4, 151.8, 147.2, 143.0, 139.0, 137.5, 134.1, 132.6, 129.1, 127.0, 121.4, 120.8, 118.0, 112.5, 110.9, 105.8, 15.5, 11.4. HPLC t_R = 1.69 min. ESI-MS: (m/z) calculated, 431.17 [M + H]$^+$; found, 431.4.

Biochemistry. JNK3$^{39-402}$ was expressed and purified as described previously.[34] JNK3-C154A was prepared using the NEB Q5 site-directed mutagenesis kit. Expression and purification of the C154A mutant were performed exactly as for wild-type JNK3.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.6b01180.

> Details for the synthesis and characterization of compound **32**; structures of intermediates **28−30**; kinase profile; description of the ELISA activity assay; in vitro reactivity study of inhibitor **7** with glutathione; metabolism in human male liver microsomes; mass measurements (PDF)
> SMILES strings and some data (CSV)
> 3D coordinates of the JNK3 in complex with **7** (PDB)

■ AUTHOR INFORMATION

Corresponding Authors

*P.K.: phone, +49 7071 2974579; e-mail, pierre.koch@uni-tuebingen.de.
*S.A.L.: phone, +49 7071 2972459; e-mail, stefan.laufer@uni-tuebingen.de.

ORCID Ⓞ

Stefan A. Laufer: 0000-0001-6952-1486

Author Contributions

§F.M. and A. E.-G. contributed equally.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS USED

MAP, mitogen activated protein; JNK, c-Jun N-terminal kinase; PyBOP, benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate; TBTU, O-(benzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium tetrafluoroborate; ELISA, enzyme-linked immunosorbent assay; IC$_{50}$, half maximal inhibitory concentration; HLM, male human liver microsomes; PAINS, pan assay interference compounds; LC−MS, liquid chromatography−mass spectrometry

■ REFERENCES

(1) Cargnello, M.; Roux, P. P. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol. Mol. Biol. Rev.* **2011**, *75*, 50−83.

(2) Davis, R. J. Signal transduction by the JNK group of MAP kinases. *Cell* **2000**, *103*, 239−252.

(3) Gupta, S.; Barrett, T.; Whitmarsh, A. J.; Cavanagh, J.; Sluss, H. K.; Dérijard, B.; Davis, R. J. Selective interaction of JNK protein kinase isoforms with transcription factors. *EMBO J.* **1996**, *15*, 2760−2770.

(4) Bogoyevitch, M. A.; Kobe, B. Uses for JNK: the many and varied substrates of the c-Jun N-terminal kinases. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 1061−1095.

(5) Yoon, S. O.; Park, D. J.; Ryu, J. C.; Ozer, H. G.; Tep, C.; Shin, Y. J.; Lim, T. H.; Pastorino, L.; Kunwar, A. J.; Walton, J. C.; Nagahara, A. H.; Lu, K. P.; Nelson, R. J.; Tuszynski, M. H.; Huang, K. JNK3 perpetuates metabolic stress induced by Aβ peptides. *Neuron* **2012**, *75*, 824−837.

(6) Ungerstedt, U. 6-Hydroxy-dopamine induced degeneration of central monoamine neurons. *Eur. J. Pharmacol.* **1968**, *5*, 107−110.

(7) Ungerstedt, U.; Arbuthnott, G. W. Quantitative recording of rotational behavior in rats after 6-hydroxy-dopamine lesions of the nigrostriatal dopamine system. *Brain Res.* **1970**, *24*, 485−493.

(8) Antoniou, X.; Falconi, M.; Di Marino, D.; Borsello, T. JNK3 as a therapeutic target for neurodegenerative diseases. *J. Alzheimer's Dis.* **2011**, *24*, 633−642.

(9) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discovery* **2011**, *10*, 307−317.

(10) Liu, Q.; Sabnis, Y.; Zhao, Z.; Zhang, T.; Buhrlage; Sara, J.; Jones; Lyn, H.; Gray; Nathanael, S. Developing irreversible inhibitors of the protein kinase cysteinome. *Chem. Biol.* **2013**, *20*, 146−159.

(11) Sanderson, K. Irreversible kinase inhibitors gain traction. *Nat. Rev. Drug Discovery* **2013**, *12*, 649−651.

(12) Genova, C.; Rijavec, E.; Barletta, G.; Burrafato, G.; Biello, F.; Dal Bello, M. G.; Coco, S.; Truini, A.; Alama, A.; Boccardo, F.; Grossi, F. Afatinib for the treatment of advanced non-small-cell lung cancer. *Expert Opin. Pharmacother.* **2014**, *15*, 889−903.

(13) Knapp, S.; Arruda, P.; Blagg, J.; Burley, S.; Drewry, D. H.; Edwards, A.; Fabbro, D.; Gillespie, P.; Gray, N. S.; Kuster, B.; Lackey, K. E.; Mazzafera, P.; Tomkinson, N. C. O.; Willson, T. M.; Workman, P.; Zuercher, W. J. A public-private partnership to unlock the untargeted kinome. *Nat. Chem. Biol.* **2013**, *9*, 3−6.

(14) Zhang, T.; Inesta-Vaquera, F.; Niepel, M.; Zhang, J.; Ficarro, S. B.; Machleidt, T.; Xie, T.; Marto, J. A.; Kim, N.; Sim, T.; Laughlin, J. D.; Park, H.; LoGrasso, P. V.; Patricelli, M.; Nomanbhoy, T. K.; Sorger, P. K.; Alessi, D. R.; Gray, N. S. Discovery of potent and selective covalent inhibitors of JNK. *Chem. Biol.* **2012**, *19*, 140−154.

(15) Ansideri, F.; Lange, A.; El-Gokha, A.; Boeckler, F. M.; Koch, P. Fluorescence polarization-based assays for detecting compounds binding to inactive c-Jun N-terminal kinase 3 and p38α mitogen-activated protein kinase. *Anal. Biochem.* **2016**, *503*, 28−40.

(16) Scapin, G.; Patel, S. B.; Lisnock, J.; Becker, J. W.; LoGrasso, P. V. The structure of JNK3 in complex with small molecule inhibitors: Structural basis for potency and selectivity. *Chem. Biol.* **2003**, *10*, 705−712.

(17) *Small-Molecule Drug Discovery Suite 2015-4: Glide*, version 6.9; Schrödinger, LLC: New York, 2015.

(18) *The PyMOL Molecular Graphics System*, version 0.99; Schrodinger, LLC: New York, 2006.

(19) Kalgutkar, A. S.; Dalvie, D. K. Drug discovery for a new generation of covalent drugs. *Expert Opin. Drug Discovery* **2012**, *7*, 561−581.

(20) Baillie, T. A. Targeted covalent inhibitors for drug design. *Angew. Chem., Int. Ed.* **2016**, *55*, 13408−13421.

(21) Gonzales, A. J.; Hook, K. E.; Althaus, I. W.; Ellis, P. A.; Trachet, E.; Delaney, A. M.; Harvey, P. J.; Ellis, T. A.; Amato, D. M.; Nelson, J. M.; Fry, D. W.; Zhu, T.; Loi, C. M.; Fakhoury, S. A.; Schlosser, K. M.; Sexton, K. E.; Winters, R. T.; Reed, J. E.; Bridges, A. J.; Lettiere, D. J.; Baker, D. A.; Yang, J. X.; Lee, H. T.; Tecle, H.; Vincent, P. W. Antitumor activity and pharmacokinetic properties of PF-00299804, a

second-generation irreversible pan-erbB receptor tyrosine kinase inhibitor. *Mol. Cancer Ther.* **2008**, *7*, 1880−1889.

(22) Lanning, B. R.; Whitby, L. R.; Dix, M. M.; Douhan, J.; Gilbert, A. M.; Hett, E. C.; Johnson, T.; Joslyn, C.; Kath, J. C.; Niessen, S.; Roberts, L. R.; Schnute, M. E.; Wang, C.; Hulce, J. J.; Wei, B. X.; Whiteley, L. O.; Hayward, M. M.; Cravatt, B. F. A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat. Chem. Biol.* **2014**, *10*, 760−767.

(23) Tsou, H. R.; Mamuya, N.; Johnson, B. D.; Reich, M. F.; Gruber, B. C.; Ye, F.; Nilakantan, R.; Shen, R.; Discafani, C.; DeBlanc, R.; Davis, R.; Koehn, F. E.; Greenberger, L. M.; Wang, Y. F.; Wissner, A. 6-Substituted-4-(3-bromophenylamino)quinazolines as putative irreversible inhibitors of the epidermal growth factor receptor (EGFR) and human epidermal growth factor receptor (HER-2) tyrosine kinases with enhanced antitumor activity. *J. Med. Chem.* **2001**, *44*, 2719−2734.

(24) Muth, F.; Günther, M.; Bauer, S. M.; Döring, E.; Fischer, S.; Maier, J.; Drückes, P.; Köppler, J.; Trappe, J.; Rothbauer, U.; Koch, P.; Laufer, S. A. Tetra-substituted pyridinylimidazoles as dual inhibitors of p38α mitogen-activated protein kinase and c-Jun N-terminal kinase 3 for potential treatment of neurodegenerative diseases. *J. Med. Chem.* **2015**, *58*, 443−456.

(25) Laufer, S.; Hauser, D.; Stegmiller, T.; Bracht, C.; Ruff, K.; Schattel, V.; Albrecht, W.; Koch, P. Tri- and tetrasubstituted imidazoles as p38 alpha mitogen-activated protein kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6671−6675.

(26) Goettert, M.; Graeser, R.; Laufer, S. A. Optimization of a nonradioactive immunosorbent assay for p38 alpha mitogen-activated protein kinase activity. *Anal. Biochem.* **2010**, *406*, 233−234.

(27) Goettert, M.; Luik, S.; Graeser, R.; Laufer, S. A. A direct ELISA assay for quantitative determination of the inhibitory potency of small molecules inhibitors for JNK3. *J. Pharm. Biomed. Anal.* **2011**, *55*, 236−240.

(28) Selig, R.; Schattel, V.; Goettert, M.; Schollmeyer, D.; Albrecht, W.; Laufer, S. Conformational effects on potency of thioimidazoles and dihydrothiazolines. *MedChemComm* **2011**, *2*, 261−269.

(29) Koch, P.; Gehringer, M.; Laufer, S. A. Inhibitors of c-Jun N-terminal kinases: an update. *J. Med. Chem.* **2015**, *58*, 72−95.

(30) Gehringer, M.; Muth, F.; Koch, P.; Laufer, S. A. c-Jun N-terminal kinase inhibitors: a patent review (2010 − 2014). *Expert Opin. Ther. Pat.* **2015**, *25*, 849−872.

(31) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P. T.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127−132.

(32) Sterling, T.; Irwin, J. J. ZINC 15 − Ligand discovery for everyone. *J. Chem. Inf. Model.* **2015**, *55*, 2324−2337.

(33) Baell, J. B.; Ferrins, L.; Falk, H.; Nikolakopoulos, G. PAINS: Relevance to tool compound discovery and fragment-based screening. *Aust. J. Chem.* **2013**, *66*, 1483−1494.

(34) Lange, A.; Günther, M.; Büttner, F. M.; Zimmermann, M. O.; Heidrich, J.; Hennig, S.; Zahn, S.; Schall, C.; Sievers-Engler, A.; Ansideri, F.; Koch, P.; Laemmerhofer, M.; Stehle, T.; Laufer, S. A.; Boeckler, F. M. Targeting the gatekeeper MET146 of c-Jun N-terminal kinase 3 induces a bivalent halogen/chalcogen bond. *J. Am. Chem. Soc.* **2015**, *137*, 14640−14652.

Supporting Information

Tri- and Tetra-substituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3

Felix Muth,[a,‡] Ahmed El-Gokha,[a,b,‡] Francesco Ansideri,[a] Michael Eitel,[a] Eva Döring,[a] Adrian Sievers-Engler,[a] Andreas Lange,[a] Frank M. Boeckler,[a] Michael Laemmerhofer,[a] Pierre Koch[a,*] and Stefan A. Laufer[a,*]

[a]Department of Pharmaceutical Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany.

[b]Menofia University, Chemistry Department, Faculty of Science, Menofia, Egypt.

Table of Contents

General information

All reagents and (anhydrous) solvents are commercially available and were used without further purification. 4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-amine (**31**) was synthesized according to literature.[1] N^1-(4-(4-(4-Fluorophenyl)-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)benzene-1,4-diamine (**3**) was synthesized according to literature.[2]

NMR:
^{1}H- and ^{13}C-NMR spectra were obtained with Bruker 200 Avance, Bruker ARX 250, Bruker Avance III HD 300 NanoBay or with Bruker 400 Avance. The spectra were obtained in the indicated solvent and calibrated against the residual proton peak of the deuterated solvent. Chemical shifts (δ) are reported in parts per million (ppm).

Mass Spectrometry:
Mass spectra were recorded on Advion Expression S ESI-MS coupled with a TLC interface. High-resolution mass spectra (FT-ICR-MS) were obtained from the Institute of Pharmaceutical Sciences, Department of Pharmaceutical Analytics and Bioanalytics, Eberhard Karls Universität Tübingen. GC/MS analyses were carried out on a Hewlett Packard HP 6890 series GC-system equipped with a HP-5MS capillary column (0.25 µm film thickness, 30 m x 250 µm) and a HP 5973 mass selective detector (EI ionization). Helium was used as carrier gas in the following temperature program: start at 100 °C and hold for 1 min, then increase to 270 °C in 27.3 min, then increase to 300 °C in 9 min and hold for 2 min.

TLC:
Analyses were performed on fluorescent silica gel 60 F_{254} plates (Merck) and visualized under UV illumination at 254 nm and 366 nm.

Column chromatography:
Column chromatography was performed on Davisil LC60A 20-45 micron silica from Grace Davison and Geduran Si60 63-200 micron silica from Merck for the pre-column using an Interchim PuriFlash 430 automated flash chromatography system.

HPLC:
The purity of all tested compounds was determined via reverse phase high performance liquid chromatography. The purity of all tested compounds is ≥ 95 %.
In case of compounds **4,5** and **7-22**: Hewlett Packard HP 1090 Series II LC equipped with a UV diode array detector (DAD, detection at 230 nm and 254 nm); Phenomenex Luna 5u C8 column (150 mm x 4.6 mm, 5 µm); 35 °C oven temperature; injection volume: 5 µL; gradient (flow: 1.5 mL/min): 0.01 M KH_2PO_4, pH 2.3 (solvent A), methanol (solvent B): 40 % B to 85 % B in 8 min, 85 % B for 5 min, 85% to 40 % B in 1 min, 40 % B for 2 min, stop time 16 min.
In case of compounds **6** and **23**: Agilent 1100; XBridge™ C18 (150 mm x 4.6 mm, 5 µm); 30 °C oven temperature; injection volume: 10 µL; gradient (flow: 1.5 mL/min): 0.01 M KH_2PO_4, pH 2.3 (solvent A), methanol (solvent B): 45 % B to 85 % B in 9 min, 85 % B for 6 min, stop time 15 min.

Docking:
The docking studies were carried out using Schroedinger Glide Docking Tool.[3] The presented docking results were the best ranked ones. For visualization and generation of pharmacophore models, PyMOL Molecular Graphics System, Version 0.99 Schrödinger, LLC[4] was used.

Schemes S1-S4

Scheme S1. Synthesis of *tri*-substituted pyridinylimidazole 6.[a]
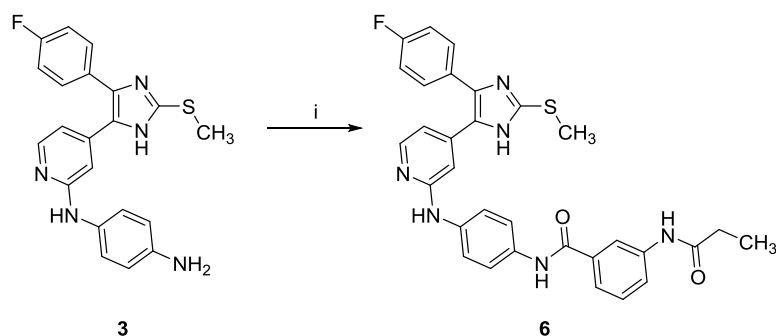
[a]Reagents and conditions: (i) 3-propionamidobenzoic acid, PyBOP, DIPEA, DMF, DCM, rt, 18 h.

Scheme S2. Synthesis of *tetra*-substituted pyridinylimidazole 9.[a]
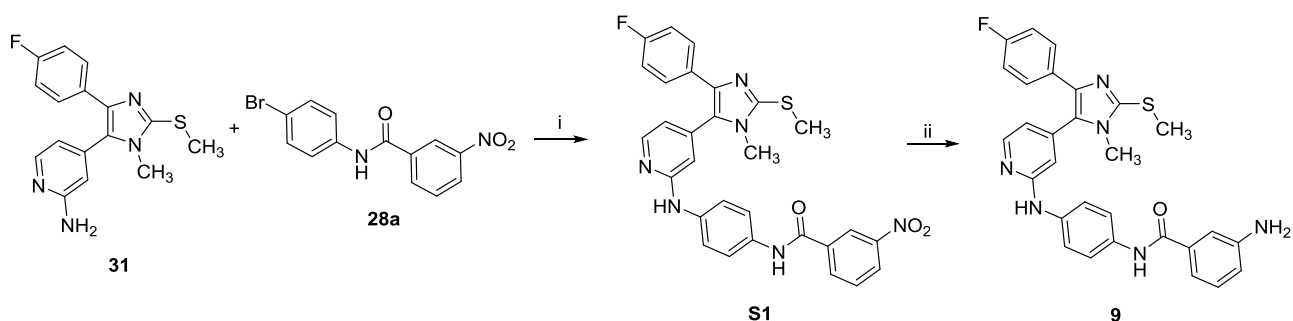
[a]Reagents and conditions: (i) Cs_2CO_3, BrettPhos precatalyst, 1,4-dioxane(abs)/tert-butanol 4:1, 125 °C, 5 h; (ii) Zn, ammonium formate, EtOH, reflux temperature, 4 h.

Scheme S3. Synthesis of *N*1-(4-(4-Methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)benzene-1,4-diamine (32).[a]
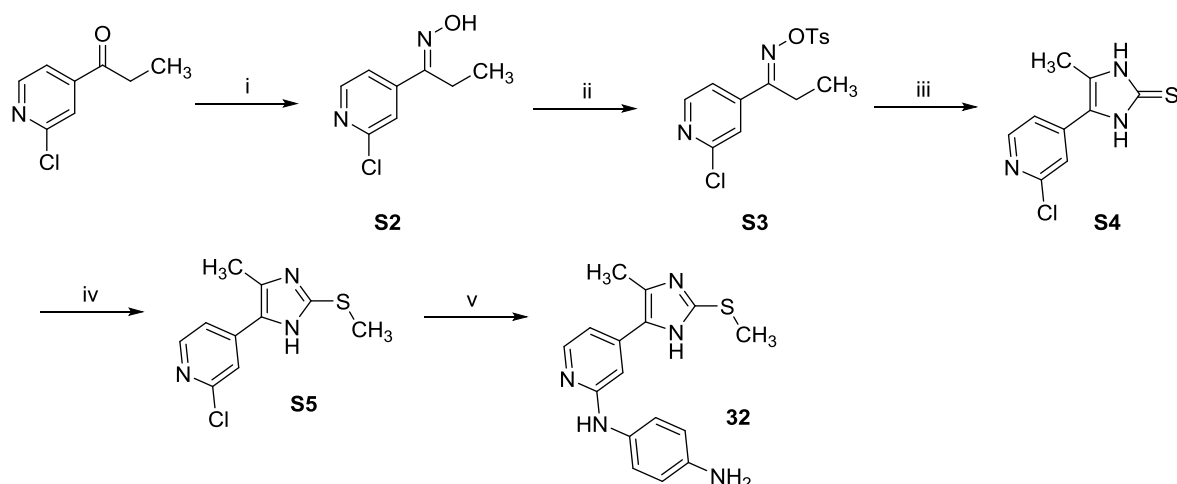
[a]Reagents and conditions: (i) $NH_2OH \cdot HCl$, NaOH, CH_3OH, 0 °C, 2 h; (ii) *p*-toluenesulfonylchloride, pyridine, 24 h; (iii) a) K, EtOH, 0 °C to rt, Et_2O, 16h; b) HCl, 50 °C, 4 h; c) KSCN, MeOH, reflux temperature, 4 h; (iv) NaO*t*Bu, MeI, MeOH, 0 °C to 55 °C; (v) *p*-phenylenediamine, *n*-butanol, 1.25 M HCl in ethanol, 160 °C, 18 h.

Scheme S4. Synthesis of *tri*-substituted pyridinylimidazole 23.[a]

32 → **23**
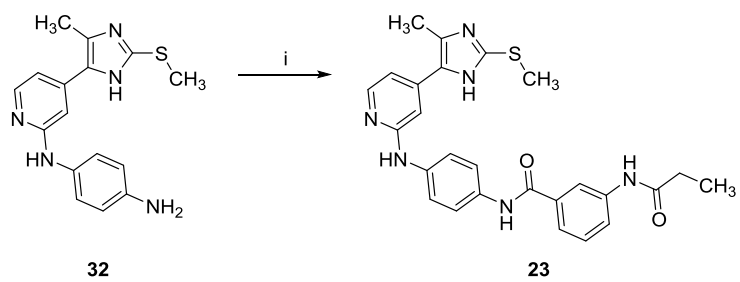
[a]Reagents and conditions: (i) 3-propionamidobenzoic acid, PyBOP, DIPEA, DMF, DCM, rt, 18 h.

Experimental procedures

N-(4-((4-(4-(4-Fluorophenyl)-1-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)phenyl)-3-nitrobenzamide (S1)
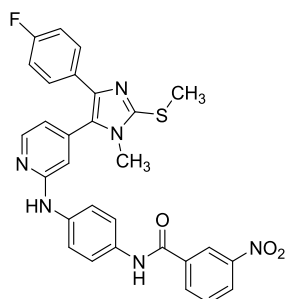
According to general procedure **F**, **S1** was synthesized from **31** and **28a**. Eluent: ethyl acetate/n-hexane 30/70 → ethyl acetate/n-hexane 60/40. The product containing fractions were unified and all volatiles were evaporated under reduced pressure. Yield: 28 %. **^{1}H NMR** (200 MHz, DMSO-d_6) δ 10.49 (s, 1H), 9.18 (s, 1H), 8.82 – 8.75 (m, 1H), 8.47 – 8.36 (m, 2H), 8.27 (d, J = 5.6 Hz, 1H), 7.83 (t, J = 8.0 Hz, 1H), 7.66 (s, 4H), 7.53 – 7.41 (m, 2H), 7.21 – 7.08 (m, 2H), 6.79 – 6.71 (m, 2H), 3.43 (s, 3H), 2.65 (s, 3H). **Exact Mass:** 554.2; **ESI-MS** (m/z): 353.2 [M-H]$^-$

1-(2-Chloropyridin-4-yl)propan-1-one oxime (S2).
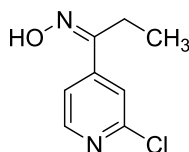
A solution of sodium hydroxide (20%; 10 mL) and hydroxylamine hydrochloride (1.79 g, 25.69 mmol) in water (10 mL) was added to a solution of 2-chloro-4-propionylpyridine (3.96 g, 23.25 mmol) in methanol (20 mL) at 0 °C. After the reaction was stirred for 2 h at 0 °C, the product was extracted by ethyl acetate and the solvent was evaporated under reduced pressure to obtain the product as a white-yellowish solid (90 %; 3.56 g). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.88 (s, 1H), 8.42 (d, J = 5.1 Hz, 1H), 7.65 (dd, J = 4.4, 5.9 Hz, 2H), 2.72 (q, J = 7.3 Hz, 2H), 1.02 (t, J = 7.5 Hz, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 155.5, 151.0, 150.2, 146.6, 120.2, 119.5, 17.6, 10.4. **GC-MS** (m/z) calculated: 184.04 [M]$^+$; found: 184.0.

1-(2-Chloropyridin-4-yl)propan-1-one O-tosyl oxime (S3)
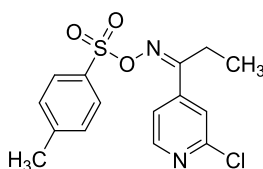
p-Toluenesulfonylchloride (10.58 g, 55.50 mmol) was added to a solution of 2-chloro-4-propionylpyridine oxime (**S2**) (6.81 g, 37.00 mmol) in dry pyridine (40 mL) under argon atmosphere. After the reaction was stirred for 24 h at rt, the solution was diluted with ice water (300 mL) and stirred for further 3 h. The water layer was extracted with ethyl acetate, washed with water and purified by flash chromatography (SiO$_2$, n-hexane/ethyl acetate, 100:00 to 90:10) to get a viscous yellow oil (yield: 83 %; 10.47 g). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.50 (d, J = 5.1 Hz, 1H), 7.91 (d, J = 7.8 Hz, 2H), 7.64 (s, 1H), 7.56 (d, J = 5.1 Hz, 1H), 7.49 (d, J = 7.8 Hz, 2H), 2.81 (d, J = 7.5 Hz, 2H), 2.41 (s, 3H), 0.99 (s, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 166.6, 151.2, 150.8, 145.8, 143.0, 131.4, 130.1, 128.5, 121.7, 120.5, 21.1, 20.6, 10.4. **FABMS**: (m/z) calculated: 339.06 [M+H]$^+$; found: 339.1.

4-(2-Chloropyridin-4-yl)-5-methyl-1,3-dihydro-2*H*-imidazole-2-thione (S4)
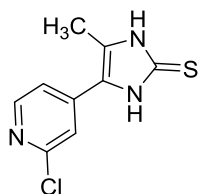
The preparation of **S4** was performed in three steps:

1) Solid potassium (1.21 g, 30.92 mmol) was added in pieces to absolute ethanol (50 mL). After the potassium was totally dissolved the reaction mixture was cooled to 0 °C and 1-(2-chloropyridin-4-yl)propan-1-one O-tosyl oxime (**S3**) (10.47 g, 30.92 mmol) in absolute ethanol (350 mL) was added dropwise to the solution. After completion of the addition the reaction was warmed to rt and stirred for 1 h. The solution was diluted with diethyl ether (100 mL) and stirred for 16 h. The precipitate was removed by filtration and washed with diethyl ether and the filtrate was concentrated at reduced pressure to obtain a yellow viscous liquid (7.80 g).

2) Concentrated hydrochloric acid (60 mL) was added to the crude product (7.80 g) obtained in step 1 and stirred for 4 h at 50 °C. Then, the solvent was evaporated under reduced pressure and a yellow solid (8.15 g) was obtained.

3) Potassium thiocyanate (15.02 g; 154.60 mmol) was added to the crude product (8.15 g) obtained in step 2 in methanol (160 mL) and stirred for 4 h at 90 °C. The formed yellow precipitate was filtered, washed with water and dried under vacuum to get the product as a yellow solid (4.56 g; 65 % over 3 steps). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.50 (br. s, 1H), 12.43 (br. s, 1H), 8.43 - 8.27 (m, 1H), 7.59 (s, 1H), 7.51 - 7.40 (m, 1H), 2.27 (s, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 161.5, 151.1, 150.1, 139.1, 126.9, 120.1, 118.8, 118.3, 10.8. **HPLC tR** = 2.67 min, purity: 100% (λ = 254 nm). **FABMS**: (m/z) calculated: 226.02 [M+H]$^+$; found: 226.1.

2-Chloro-4-(4-methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridine (S5)
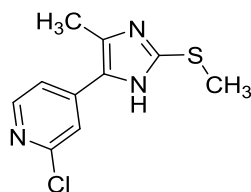
Sodium *tert*-butoxide (427 mg, 4.44 mmol) was added to a solution of **40** (500 mg, 2.22 mmol) in absolute methanol (20 mL). The reaction mixture was cooled to 0 °C and methyl iodide (147.5 µL, 2.22 mmol) was added slowly under argon atmosphere. The reaction mixture was stirred for 30 min at 0 °C and 3 h at 55 °C. After cooling to rt, the solvent was removed under vacuum, the residue was dissolved in water, extracted with ethyl acetate and dried over Na$_2$SO$_4$. The solvent was evaporated under reduced pressure and the residue was purified by flash chromatography (SiO$_2$, dichloromethane/ethanol, 100:0 to 90:10) to obtain the product as a yellowish solid (76 %; 404 mg). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.46 (br. s, 1H), 8.32 (d, *J* = 4.8 Hz, 1H), 7.62 (br. s, 1H), 7.59 (d, *J* = 4.8 Hz, 1H), 2.57 (s, 3H), 2.44 (s, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 150.8, 149.8, 145.6, 140.2, 132.2, 129.3, 118.9, 118.6, 15.2, 11.4. **HPLC: tR** = 4.6 min. **ESIMS**: (m/z) calculated: 240.04 [M+H]$^+$; found: 240.0.

***N*1-(4-(4-Methyl-2-(methylthio)-1*H*-imidazol-5-yl)pyridin-2-yl)benzene-1,4-diamine (32)**
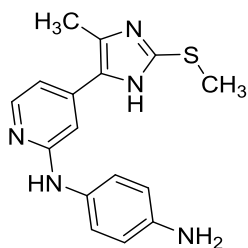
In a pressure glass tube, compound **41** (100 mg, 0.42 mmol) and *p*-phenylenediamine (67.7 mg, 0.63 mmol) were suspended in *n*-butanol (5 mL). 1.25 M HCl in ethanol (345 μL, 0.42 mmol) was added and the reaction was heated overnight at 160 °C. After cooling to rt, the solvent was removed and the residue was purified twice by flash chromatography (SiO$_2$, dichloromethane/ethanol, 90:10) to afford 80 mg (62%) as a brown solid.
^{1}H NMR (250 MHz, DMSO-*d$_6$*) δ 12.23 (br. s, 1H), 8.36 (s, 1H), 7.97 (d, *J* = 5.4 Hz, 1H), 7.33 - 7.13 (m, 2H), 6.97 (s, 1H), 6.91 - 6.64 (m, 1H), 6.59 - 6.43 (m, 2H), 5.01 - 4.26 (m, 2H), 2.53 (s, 3H), 2.37 (s, 3H); **^{13}C NMR** (101 MHz, DMSO-*d$_6$*) δ 157.4, 147.3, 143.1, 142.8, 138.8, 134.3, 131.1, 126.8, 121.2, 114.2, 110.0, 104.4, 15.5, 11.3.
HPLC tR = 1.89 min. **ESIMS**: (m/z) calculated: 312.42 [M+H]$^+$; found: 312.2.

Structure of intermediates 28-30

N-(4-Bromophenyl)-3-nitrobenzamide (28a)
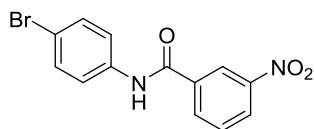
N-(4-Bromophenyl)-2-methyl-3-nitrobenzamide (28b)
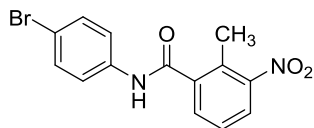
N-(3-Bromophenyl)-3-nitrobenzamide (28c)
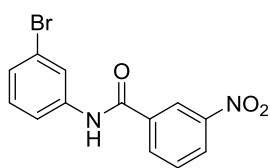
N-(3-Bromophenyl)-2-methyl-3-nitrobenzamide (28d)
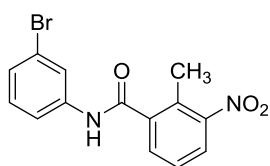
N-(4-Bromophenyl)-4-nitrobenzamide (28e)
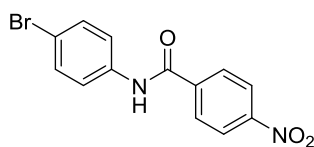
N-(4-Bromo-3-methylphenyl)-4-nitrobenzamide (28f)
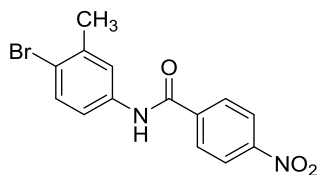
N-(3-Bromophenyl)-4-nitrobenzamide (28g)
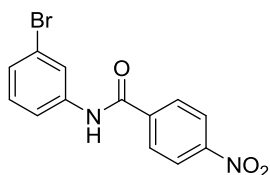
_N_-(4-Bromo-3-methylphenyl)-3-nitrobenzamide (28h)
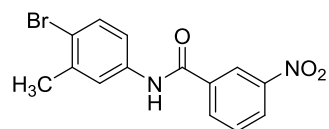
3-Amino-_N_-(4-bromophenyl)benzamide (29a)
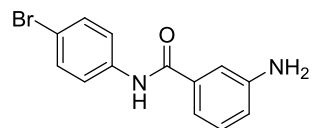
3-Amino-_N_-(4-bromophenyl)-2-methylbenzamide (29b)
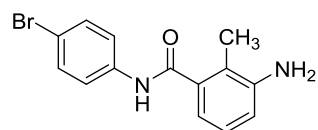
3-Amino-_N_-(3-bromophenyl)benzamide (29c)
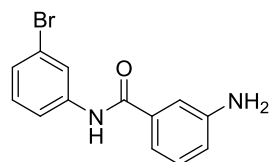
3-Amino-_N_-(3-bromophenyl)-2-methylbenzamide (29d)
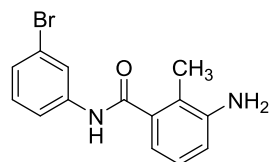
4-Amino-_N_-(4-bromophenyl)benzamide (29e)
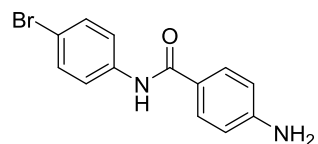
4-Amino-_N_-(4-bromo-3-methylphenyl)benzamide (29f)
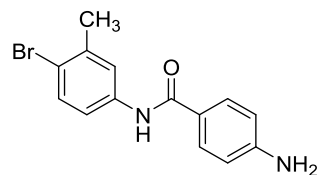
4-Amino-*N*-(3-bromophenyl)benzamide (29g)
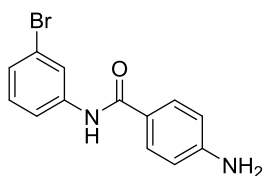
3-Amino-*N*-(4-bromo-3-methylphenyl)benzamide (29h)
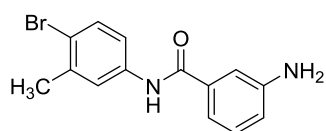
3-Acrylamido-*N*-(4-bromophenyl)benzamide (30a)
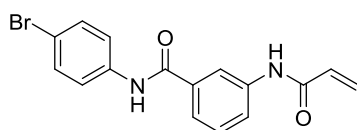
_N_-(4-Bromophenyl)-2-methyl-3-propionamidobenzamide (30b)
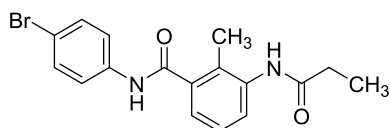
3-Acrylamido-*N*-(3-bromophenyl)benzamide (30c)
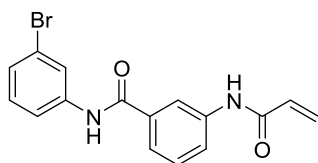
3-Acrylamido-*N*-(3-bromophenyl)-2-methylbenzamide (30d)
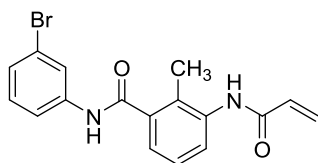
4-Acrylamido-*N*-(4-bromophenyl)benzamide (30e)
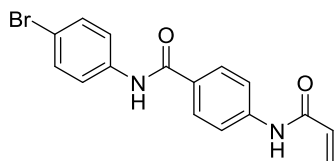
4-Acrylamido-*N*-(4-bromo-3-methylphenyl)benzamide (30f)
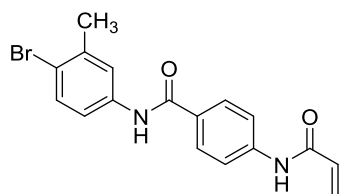
4-Acrylamido-*N*-(3-bromophenyl)benzamide (30g)
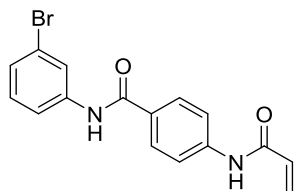
***N*-(3-Bromophenyl)-3-propionamidobenzamide (30h)**
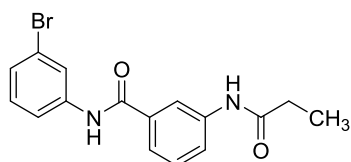
***N*-(4-Bromo-3-methylphenyl)-4-propionamidobenzamide (30i)**
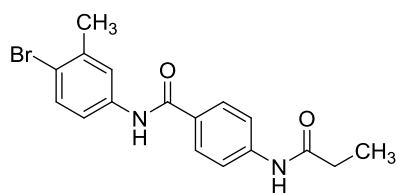
***N*-(4-Bromophenyl)-3-propionamidobenzamide (30j)**
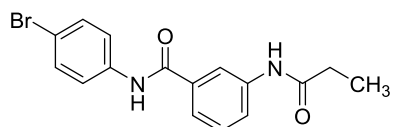
***N*-(4-Bromophenyl)-4-propionamidobenzamide (30k)**
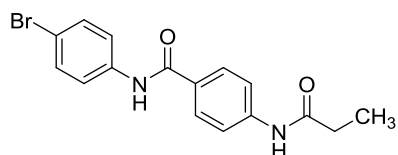
3-Acrylamido-*N*-(4-bromophenyl)-2-methylbenzamide (30l)
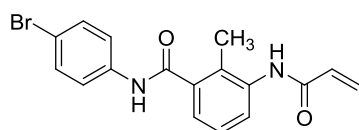
3-Acrylamido-*N*-(4-bromo-3-methylphenyl)benzamide (30m)
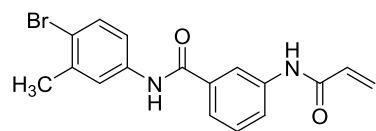
Selectivity Screen

Compounds **7** and **21** were tested at ProQinase GmbH (Freiburg, Germany) against 410 kinases at concentrations of 1 µM, 5 µM and 0.1 µM, 0.5 µM respectively. Activities are specified as residual activities (% of control) and values under 50 % are highlighted in orange.

Table S1

#	Kinase Name	Kinase Family*	7 c = 0.1 µM	7 c = 0.5 µM	21 c = 1 µM	21 c = 5 µM
1	ABL1 E255K	TK	106	114	94	88
2	ABL1 F317I	TK	102	102	102	82
3	ABL1 G250E	TK	96	75	88	89
4	ABL1 H396P	TK	99	91	93	71
5	ABL1 M351T	TK	101	80	87	61
6	ABL1 Q252H	TK	101	93	96	65
7	ABL1 T315I	TK	94	93	91	82
8	ABL1 wt	TK	108	109	96	71
9	ABL1 Y253F	TK	105	108	85	64
10	ABL2	TK	134	113	74	84
11	ACK1	TK	130	128	100	95
12	ACV-R1	TKL	107	93	85	54
13	ACV-R1B	TKL	92	91	94	94
14	ACV-R2A	TKL	97	92	81	80
15	ACV-R2B	TKL	88	82	88	68
16	ACV-RL1	TKL	93	86	74	52
17	AKT1 aa106-480	AGC	104	87	84	69
18	AKT2 aa107-481	AGC	108	105	92	78
19	AKT3 aa106-479	AGC	108	122	95	90
20	ALK C1156Y (GST-HIS-tag)	TK	122	117	86	69
21	ALK F1174L (GST-HIS-tag)	TK	113	106	84	74
22	ALK F1174S (GST-HIS-tag)	TK	132	119	92	72
23	ALK L1196M (GST-HIS-tag)	TK	114	102	72	55
24	ALK R1275Q (GST-HIS-tag)	TK	119	123	85	73
25	ALK wt (GST-HIS-tag)	TK	111	105	98	72
26	AMPK-alpha1 aa1-550	CAMK	70	62	100	93
27	ARK5	CAMK	85	91	97	88
28	ASK1	STE	108	104	92	95
29	Aurora-A	OTHER	97	98	74	34
30	Aurora-B	OTHER	102	103	73	40
31	Aurora-C	OTHER	84	81	86	57
32	AXL	TK	105	102	99	76
33	BLK	TK	101	97	77	81
34	BMPR1A	TKL	96	84	83	88
35	BMX	TK	108	111	106	85
36	B-RAF V600E	TKL	91	103	110	113
37	B-RAF wt	TKL	107	107	96	82
38	BRK	TK	109	125	101	105
39	BRSK1	CAMK	101	109	107	95
40	BRSK2	CAMK	98	95	103	85
41	BTK	TK	107	107	101	87
42	BUB1B	OTHER	129	112	95	80
43	CAMK1D	CAMK	99	97	90	75
44	CAMK2A	CAMK	88	81	90	90
45	CAMK2B	CAMK	103	89	97	79
46	CAMK2D	CAMK	101	106	95	92
47	CAMK2G	CAML	68	58	81	54

#	Name	Group				
48	CAMK4	CAMK	130	115	110	91
49	CAMKK1	OTHER	64	66	98	108
50	CAMKK2	OTHER	123	122	83	86
51	CCDC6-RET	TK	106	98	84	84
52	CDC42BPA	AGC	93	89	102	91
53	CDC42BPB	AGC	90	85	89	82
54	CDC7/DBF4	OTHER	105	109	94	89
55	CDK1/CycA2	CMGC	96	116	89	79
56	CDK1/CycB1	CMGC	98	100	94	77
57	CDK1/CycE1	CMGC	98	88	90	74
58	CDK12 R722C/CycK	CMGC	90	99	93	66
59	CDK12/CycK wt	CMGC	71	65	82	57
60	CDK16/CycY	CMGC	103	108	88	87
61	CDK19/CycC	CMGC	93	105	72	65
62	CDK2/CycA2	CMGC	109	98	78	56
63	CDK2/CycE1	CMGC	105	106	88	62
64	CDK3/CycC	CMGC	105	108	91	74
65	CDK3/CycE1	CMGC	103	97	98	79
66	CDK4/CycD1	CMGC	107	115	94	76
67	CDK4/CycD3	CMGC	124	117	92	83
68	CDK5/p25NCK	CMGC	100	95	77	54
69	CDK5/p35NCK	CMGC	96	98	74	36
70	CDK6/CycD1	CMGC	106	119	91	87
71	CDK6/CycD3	CMGC	106	109	93	86
72	CDK7/CycH/MAT1	CMGC	112	114	89	86
73	CDK8/CycC	CMGC	109	100	67	29
74	CDK9/CycK	CMGC	99	107	85	39
75	CDK9/CycT1	CMGC	98	93	63	38
76	CHK1	CAMK	110	100	94	100
77	CHK2	CAMK	91	93	70	36
78	CK1-alpha1	CK1	100	85	70	38
79	CK1-delta	CK1	59	20	13	3
80	CK1-epsilon	CK1	87	54	57	21
81	CK1-gamma1	CK1	97	96	65	31
82	CK1-gamma2	CK1	81	75	22	5
83	CK1-gamma3	CK1	92	77	38	12
84	CK2-alpha1	OTHER	91	91	92	75
85	CK2-alpha2	OTHER	92	94	80	52
86	CLK1	CMGC	91	98	67	37
87	CLK2	CMGC	94	92	81	51
88	CLK3	CMGC	83	68	98	89
89	CLK4	CMGC	83	86	46	20
90	COT	STE	92	99	95	103
91	CSF1-R	TK	102	89	106	69
92	CSK	TK	102	96	87	81
93	DAPK1	CAMK	103	98	100	94
94	DAPK2	CAMK	76	83	92	78
95	DAPK3	CAMK	88	84	88	88
96	DCAMKL2	CAMK	97	101	95	94
97	DDR2 N456S	TK	76	81	88	67
98	DDR2 T654M	TK	79	82	89	70
99	DDR2 wt	TK	92	91	90	76
100	DMPK	AGC	107	97	97	92
101	DNA-PK	ATYPICAL	86	98	92	103
102	DYRK1A	CMGC	122	103	92	70
103	DYRK1B	CMGC	106	102	85	53
104	DYRK2	CMGC	96	105	72	30
105	DYRK3	CMGC	101	101	82	39

106	DYRK4	CMGC	90	89	90	84
107	EEF2K	ATYPICAL	100	98	91	92
108	EGF-R d746-750	TK	126	109	115	88
109	EGF-R d747-749/A750P	TK	133	132	101	101
110	EGF-R d747-752/P753S	TK	111	104	85	86
111	EGF-R d752-759	TK	107	101	88	96
112	EGF-R G719C	TK	68	100	96	75
113	EGF-R G719S	TK	107	104	88	97
114	EGF-R L858R	TK	111	119	95	83
115	EGF-R L861Q	TK	114	98	104	91
116	EGF-R T790M	TK	106	108	91	84
117	EGF-R T790M/L858R	TK	114	110	96	96
118	EGF-R wt	TK	138	125	108	101
119	EIF2AK2	OTHER	123	129	84	55
120	EIF2AK3	OTHER	100	94	92	97
121	EML-ALK	TKL	108	107	91	89
122	EML-ALK F117L	TKL	105	106	93	80
123	EPHA1	TK	105	98	90	63
124	EPHA2	TK	142	128	92	81
125	EPHA3	TK	103	92	98	89
126	EPHA4	TK	99	99	93	78
127	EPHA5	TK	118	113	94	87
128	EPHA6	TK	114	105	99	100
129	EPHA7	TK	106	96	93	89
130	EPHA8	TK	104	101	96	99
131	EPHB1	TK	122	116	83	82
132	EPHB2	TK	108	97	93	78
133	EPHB3	TK	129	115	106	108
134	EPHB4	TK	143	131	110	111
135	ERBB2	TK	111	104	92	94
136	ERBB4	TK	121	108	93	90
137	ERK1	CMGC	108	107	101	99
138	ERK2	CMGC	95	95	96	100
139	ERK5	CMGC	93	94	83	73
140	ERK7	CMGC	93	84	66	31
141	FAK aa2-1052	TK	136	133	106	92
142	FER	TK	114	114	104	93
143	FES	TK	117	110	88	91
144	FGF-R1 V561M	TK	115	123	99	65
145	FGF-R1 wt	TK	116	103	105	87
146	FGF-R2	TK	136	128	105	87
147	FGF-R3 G697C	TK	141	130	103	83
148	FGF-R3 K650E	TK	128	116	104	79
149	FGF-R3 K650M	TK	116	121	97	72
150	FGF-R3 wt	TK	118	112	98	78
151	FGF-R4	TK	125	128	88	91
152	FGR	TK	100	113	87	79
153	FLT3 D835Y	TK	90	90	65	32
154	FLT3 ITD	TK	107	107	78	49
155	FLT3 wt	TK	96	105	94	62
156	FRK	TK	104	112	99	75
157	FYN wt	TK	140	132	101	100
158	GRK2	AGC	94	105	101	88
159	GRK3	AGC	123	110	94	95
160	GRK4	AGC	92	88	101	97
161	GRK5	AGC	87	84	97	94
162	GRK6	AGC	61	76	88	90
163	GRK7	AGC	87	84	100	99

164	GSG2	OTHER	137	146	84	56
165	GSK3-alpha	CMGC	103	88	90	61
166	GSK3-beta	CMGC	112	111	91	85
167	HCK	TK	117	115	105	88
168	HIPK1	CMGC	91	84	103	143
169	HIPK2	CMGC	111	106	100	87
170	HIPK3	CMGC	91	94	93	90
171	HIPK4	CMGC	112	104	76	37
172	HRI	OTHER	100	87	85	66
173	IGF1-R	TK	101	95	92	86
174	IKK-alpha	OTHER	109	111	78	57
175	IKK-beta	OTHER	104	93	77	63
176	IKK-epsilon	OTHER	99	94	88	70
177	INS-R	TK	103	97	91	83
178	INSR-R	TK	137	131	81	72
179	IRAK1	TKL	103	100	50	11
180	IRAK4 (untagged)	TKL	117	123	85	56
181	ITK	TK	120	118	95	81
182	JAK1 aa583-1154 wt	TK	86	101	98	96
183	JAK2	TK	96	97	80	72
184	JAK3	TK	96	104	79	54
185	JNK1	CMGC	84	16	5	1
186	JNK2	CMGC	27	5	8	1
187	JNK3	CMGC	20	4	3	0
188	KIT A829P	TK	118	120	68	75
189	KIT D816H	TK	124	121	91	66
190	KIT D816V	TK	129	120	87	55
191	KIT T670I	TK	100	104	107	98
192	KIT V559D	TK	144	138	102	94
193	KIT V559D/T670I	TK	103	109	96	77
194	KIT V559D/V654A	TK	148	145	94	85
195	KIT V560G	TK	137	135	95	67
196	KIT V654A	TK	144	113	103	80
197	KIT wt	TK	138	119	98	101
198	LCK	TK	120	107	85	89
199	LIMK1	TKL	104	102	72	46
200	LIMK2	TKL	102	102	90	80
201	LRRK2 G2019S	TKL	102	100	72	52
202	LRRK2 I2020T	TKL	108	96	71	49
203	LRRK2 R1441C	TKL	103	96	80	51
204	LRRK2 wt	TKL	84	92	77	47
205	LTK	TK	131	123	108	99
206	LYN	TK	118	110	99	94
207	MAP3K1	STE	86	91	100	96
208	MAP3K10	STE	95	87	27	4
209	MAP3K11	STE	112	140	28	13
210	MAP3K7/MAP3K7IP1	STE	96	86	101	77
211	MAP3K9	STE	89	89	80	45
212	MAP4K2	STE	100	89	89	50
213	MAP4K4	STE	100	104	79	55
214	MAP4K5	STE	74	72	71	49
215	MAPKAPK2	CAMK	43	38	71	42
216	MAPKAPK3	CAMK	101	100	100	65
217	MAPKAPK5	CAMK	87	84	96	97
218	MARK1	CAMK	107	101	100	88
219	MARK2	CAMK	109	108	93	84
220	MARK3	CAMK	105	100	82	92
221	MARK4	CAMK	92	87	92	99

222	MATK	TK	146	119	92	111
223	MEK1 wt	STE	98	90	96	113
224	MEK2	STE	99	98	76	36
225	MEK5	STE	92	92	87	60
226	MEKK2	STE	88	88	99	88
227	MEKK3	STE	92	89	102	103
228	MELK	CAMK	102	95	95	70
229	MERTK	TK	125	118	76	41
230	MET D1228H	TK	113	117	76	65
231	MET D1228N	TK	104	96	89	74
232	MET F1200I	TK	111	111	81	59
233	MET M1250T	TK	117	115	88	65
234	MET wt	TK	106	103	93	78
235	MET Y1230A	TK	108	111	97	85
236	MET Y1230C	TK	108	114	90	76
237	MET Y1230D	TK	112	114	88	73
238	MET Y1230H	TK	97	97	92	72
239	MET Y1235D	TK	113	106	105	86
240	MINK1	STE	116	107	75	44
241	MKK4	STE	113	134	71	90
242	MKK6 S207D/T211D	STE	103	85	103	126
243	MKK7	STE	100	84	45	39
244	MKNK1	CAMK	120	118	90	55
245	MKNK2	CAMK	96	93	80	31
246	MLK4	TKL	105	131	81	59
247	MST1	STE	119	112	96	80
248	MST2	STE	106	104	83	69
249	MST3	STE	85	82	85	89
250	MST4	STE	111	124	94	97
251	mTOR	ATYPICAL	117	107	88	91
252	MUSK	TK	118	120	74	31
253	MYLK	CAMK	107	96	84	60
254	MYLK2	CAMK	111	98	83	59
255	MYLK3	CAMK	104	108	89	68
256	NEK1	OTHER	85	87	80	64
257	NEK11	OTHER	103	96	103	87
258	NEK2	OTHER	102	124	92	78
259	NEK3	OTHER	111	123	73	50
260	NEK4	OTHER	111	101	79	51
261	NEK6	OTHER	109	108	91	92
262	NEK7	OTHER	76	81	91	95
263	NEK9	OTHER	107	104	97	73
264	NIK	STE	107	96	91	87
265	NLK	CMGC	107	98	86	37
266	NPM1ALK	TK	108	105	94	96
267	NPM1ALK F1174L	TK	127	105	104	106
268	p38-alpha	CMGC	98	50	90	67
269	p38-beta	CMGC	104	93	93	75
270	p38-delta	CMGC	87	77	91	80
271	p38-gamma	CMGC	114	103	94	87
272	PAK1	STE	97	100	90	84
273	PAK2	STE	105	90	99	86
274	PAK3	STE	98	88	98	83
275	PAK4	STE	98	98	92	76
276	PAK6	STE	101	98	92	76
277	PAK7	STE	98	91	90	67
278	PASK	CAMK	107	99	81	55
279	PBK	OTHER	112	107	141	89

280	PDGFR-alpha D842V	TK	107	101	68	41
281	PDGFR-alpha T674I	TK	100	118	96	84
282	PDGFR-alpha V561D	TK	104	91	91	68
283	PDGFR-alpha wt	TK	96	104	97	72
284	PDGFR-beta	TK	109	101	74	46
285	PDK1	AGC	117	91	74	39
286	PHKG1	CAMK	110	127	80	44
287	PHKG2	CAMK	104	94	94	96
288	PIM1	CAMK	103	107	87	67
289	PIM2	CAMK	74	64	101	69
290	PIM3	CAMK	97	92	93	82
291	PKA	AGC	102	103	84	66
292	PKC-alpha	AGC	116	114	102	80
293	PKC-beta1	AGC	121	116	111	106
294	PKC-beta2	AGC	101	100	103	81
295	PKC-delta	AGC	110	117	84	94
296	PKC-epsilon	AGC	113	95	97	108
297	PKC-eta	AGC	107	106	91	78
298	PKC-gamma	AGC	105	113	88	100
299	PKC-iota	AGC	123	102	80	52
300	PKC-mu	AGC	100	93	68	31
301	PKC-nu	AGC	87	76	44	15
302	PKC-theta	AGC	114	115	97	96
303	PKC-zeta	AGC	102	85	87	86
304	PKC-zeta wt aa 184-592 (PKM-zeta)	AGC	109	108	90	68
305	PKMYT1	OTHER	95	101	74	56
306	PKN3	AGC	111	122	79	54
307	PLK1	OTHER	106	106	98	96
308	PLK3	OTHER	111	106	103	75
309	PRK1	AGC	104	104	100	91
310	PRK2	AGC	110	114	90	76
311	PRKD2	CAMK	92	92	69	43
312	PRKG1	AGC	100	91	92	87
313	PRKG2	AGC	105	97	84	81
314	PRKX	AGC	108	67	86	61
315	PYK2	TK	107	119	91	96
316	RAF1 Y340D/Y341D (untagged)	TKL	97	97	96	92
317	RET E762Q	TK	121	114	95	78
318	RET G691S	TK	112	119	90	77
319	RET M918T	TK	154	124	97	82
320	RET R749T	TK	121	113	96	76
321	RET R813Q	TK	126	117	98	75
322	RET S891A	TK	114	112	95	74
323	RET V804L	TK	107	105	88	80
324	RET V804M	TK	141	131	98	90
325	RET wt	TK	128	161	87	72
326	RET Y791F	TK	145	123	96	71
327	RIPK2	TKL	99	97	77	43
328	RIPK5	TKL	93	95	90	67
329	ROCK1	AGC	93	100	104	80
330	ROCK2	AGC	100	111	95	75
331	RON	TK	156	145	92	104
332	ROS	TK	117	109	87	74
333	RPS6KA1	AGC	85	86	82	64
334	RPS6KA2	AGC	107	90	88	69
335	RPS6KA3	AGC	86	77	81	58
336	RPS6KA4	AGC	84	104	97	94
337	RPS6KA5	AGC	103	100	94	87

338	RPS6KA6	AGC	103	115	88	68
339	S6K	AGC	124	119	105	87
340	S6K-beta	AGC	116	102	87	79
341	SAK	OTHER	103	79	77	57
342	SGK1	AGC	92	109	49	15
343	SGK2	AGC	85	84	93	71
344	SGK3	AGC	122	102	82	76
345	SIK1	CAMK	107	107	107	102
346	SIK2	CAMK	98	95	108	90
347	SIK3	CAMK	109	107	65	47
348	SLK	STE	114	133	102	76
349	SNARK	CAMK	98	99	99	87
350	SNK	OTHER	100	103	86	81
351	SRC (GST-HIS-tag)	TK	123	123	98	86
352	SRMS	TK	107	104	95	92
353	SRPK1	CMGC	92	106	99	86
354	SRPK2	CMGC	104	87	82	75
355	STK17A	CAMK	84	78	78	35
356	STK23	CAMK	115	107	97	101
357	STK25	STE	86	81	88	95
358	STK33	CAMK	99	98	76	71
359	STK39	STE	105	77	100	73
360	SYK aa1-635	TK	114	110	94	86
361	TAOK2	STE	107	92	38	16
362	TAOK3	STE	95	102	95	72
363	TBK1	OTHER	88	95	84	69
364	TEC	TK	98	96	99	81
365	TGFB-R1	TKL	96	94	106	92
366	TGFB-R2	TKL	111	128	53	4
367	TIE2 R849W	TK	87	96	92	107
368	TIE2 wt	TK	136	18	85	81
369	TIE2 Y1108F	TK	128	128	95	96
370	TIE2 Y897S	TK	118	110	85	85
371	TLK1	AGC	105	112	97	90
372	TLK2	AGC	92	91	92	83
373	TNK1	TK	103	111	91	61
374	TRK-A	TK	115	111	81	58
375	TRK-B	TK	96	105	89	81
376	TRK-C	TK	116	110	99	79
377	TSF1	OTHER	96	93	23	6
378	TSK2	CAMK	91	87	90	97
379	TSSK1	CAMK	86	77	76	72
380	TTBK1	CK1	96	102	102	106
381	TTBK2	CK1	96	84	94	93
382	TTK	OTHER	107	111	72	34
383	TXK	TK	98	98	87	77
384	TYK2	TK	97	94	77	46
385	TYRO3	TK	114	104	100	86
386	VEGF-R1	TK	106	93	84	66
387	VEGF-R2	TK	121	110	78	50
388	VEGF-R3	TK	112	93	73	68
389	VRK1	CK1	72	70	95	97
390	VRK2	CK1	101	73	97	83
391	WEE1	OTHER	113	120	73	79
392	WNK1	OTHER	127	106	95	81
393	WNK2	OTHER	74	85	98	78
394	WNK3	OTHER	120	99	79	67
395	YES	TK	113	100	89	75

396	ZAK	TKL	98	95	48	15
397	ZAP70	TK	127	114	88	93
398	PI4K2A	Lipid Kinase	113	113	117	133
399	PI4K2B	Lipid Kinase	113	112	128	120
400	PI4KB	Lipid Kinase	91	92	101	79
401	PIK3C2A	Lipid Kinase	138	152	189	183
402	PIK3C2B	Lipid Kinase	109	139	132	133
403	PIK3C2G	Lipid Kinase	145	155	129	108
404	PIK3C3	Lipid Kinase	98	99	108	103
405	PIK3CA/PIK3R1	Lipid Kinase	108	108	118	106
406	PIK3CB/PIK3R1	Lipid Kinase	132	131	138	166
407	PIK3CD/PIK3R1	Lipid Kinase	110	108	119	113
408	PIK3CG	Lipid Kinase	109	115	176	156
409	PIP5K1A	Lipid Kinase	94	100	102	44
410	PIP5K1C	Lipid Kinase	114	111	107	102
Selectivity Score (<50 % residual activity):			**0.007**	**0.015**	**0.037**	**0.146**

*Classification of protein kinase families:[5]
AGC: containing PKA, PKG and PKC families
CAMK: Calcium/Calmoduline-dependent protein kinases
CK1: Casein kinase 1-like
CMGC: containing CDK, MAPK, GSK3 and CLK families
TK: Tyrosine Kinase
TKL: Tyrosine Kinase-like
STE: Homologs of Yeast Sterile 7, Sterile 11, Sterile 20 Kinases

Direct enzyme- linked immunosorbent assay (ELISA) for routine screening of p38α MAPK and JNK3 inhibitors

Being a natural substrate of both p38α MAP kinase and JNK3, activation transcription factor 2 (ATF-2) purchased from ProQinase, Freiburg, Germany (# 0594-0000-2) as full-length protein is adsorbed to the 96 well assay plates (Nunc Maxisorp®) yielding a concentration of 10 µg/mL.

Dilution rows of candidate inhibitor are prepared in a kinase buffer containing active p38α MAP kinase or activate JNK3 enzyme. The active p38α MAP kinase was obtained from Prof. Dr. J. Schultz (University of Tübingen, Germany), whereas the active JNK3 enzyme was purchased from ProQinase, Freiburg, Germany (#0900-0000-1). The ATP concentrations used in the respective kinase buffers are adjusted to twice the K_m value, depending on the kinase.

The activity of p38α MAP kinase or JNK3 kinase after one hour of incubation at 37°C with the candidate inhibitors is measured by the phosphorylation degree of ATF-2, which is directly detected by a monoclonal peroxidase-conjugated antibody purchased from Sigma Aldrich (#A6228). The phosphorylation degree achieved with the respective kinase in absence of inhibitor is taken as positive control (STIM). Pure kinase buffer without kinase serves for detection of non-specific binding (NSB). After staining with 3,3',5,5'-tetramethylbenzidine reagent (BD Biosciences Europe) and termination of colour development using 1M sulphuric acid, the optical density is read out in an ELISA reader at 450 nm. As phosphorylation is inversely correlated with the inhibitor potency, the calculation of inhibition is carried out according to the formula.[6, 7]

$$Inhibition[\%] = 100 - \left(\frac{OD_{450}Sample - OD_{450}NSB}{OD_{450}STIM - OD_{450}NSB} \right) \times 100$$

Investigation of the covalent bond formation using LC-µESI-QTOF mass spectrometry

Covalent bonding has been verified by LC-µESI-QTOF mass spectrometry of intact JNK3. Potential noncovalent interactions have been eliminated by chromatography prior to ESI-ionisation and the employment of a high declustering potential (DP) of 230V. This is substantiated by the charge series of intact JNK3 consisting only of $[M+nH]^{n+}$ ions, i.e. only H^+ -adduct series has been found. Further evidence is provided by **15**, an analogue of **7** with inactive, non-Michael-reactive warhead. Presumably noncovalent bonding to the ATP-binding site is disrupted during chromatography and/or the ESI-process due to application of high declustering potential. Results show (s. Figures S3a, S3c and S3e), that inactive species produce the same spectrum like the control (unmodified JNK3), confirming that only active ligands result in a mass shift corresponding to the mass of the inhibitor (Figures S3b and S3d) by covalent modification of JNK3.
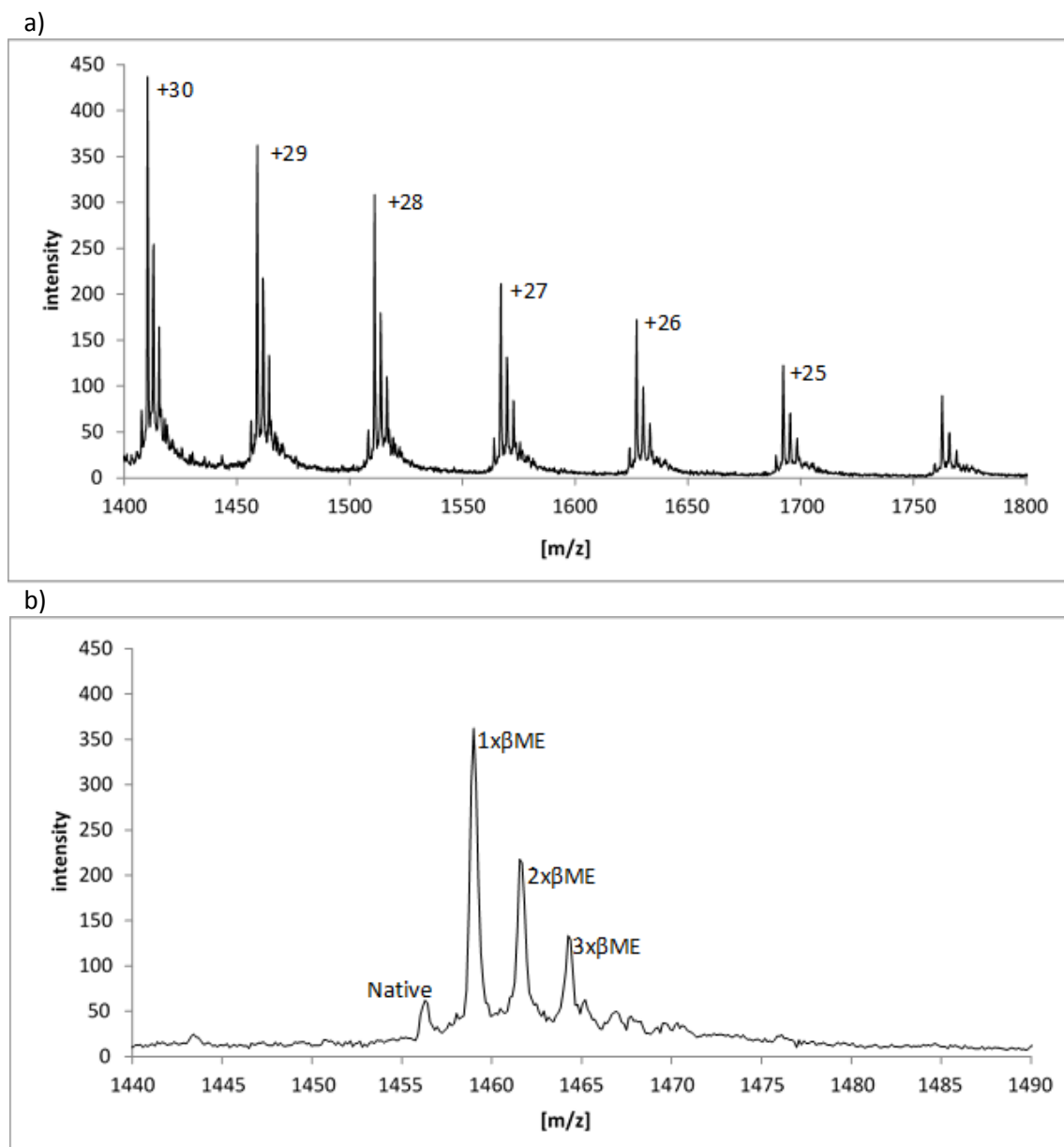
Figures S1. Charge series of JNK3. a) Charge series consists exclusively of H^+-adducts. Charge states from +30 to +25 have been used for intact protein mass deconvolution. Charge series deconvolutes to M = 42281.9073 (1xβME-JNK3). Calculated M for 1xβME-JNK3 is 42281.8941, mass accuracy 0.3ppm; b) Triplet structure of charge series is caused by 1-3 times β-mercaptoethanoylation.
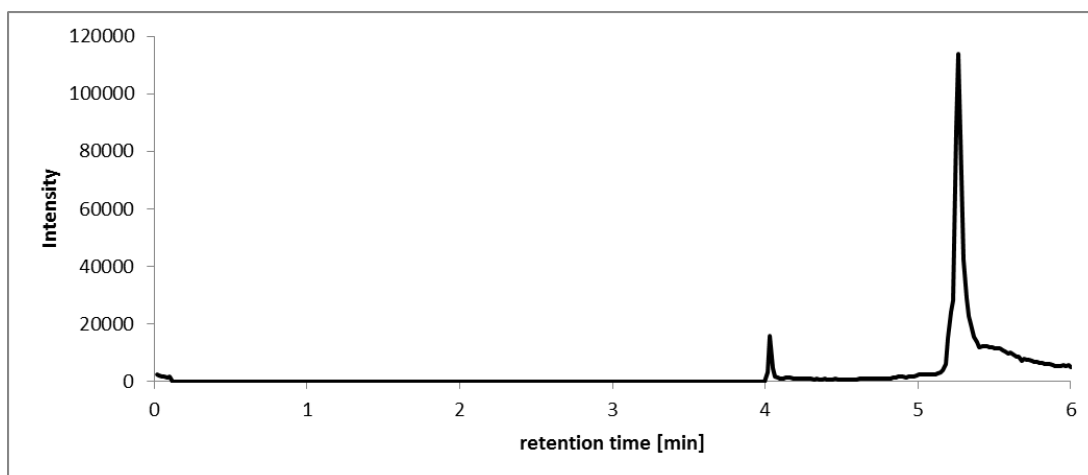
Figure S2. Chromatogram of JNK3. First 4 minutes have been switched to waste via valve, effectively preventing the reaction buffer from entering the ESI-source. JNK3 eluting at t = 5.2 min.

a)

b)
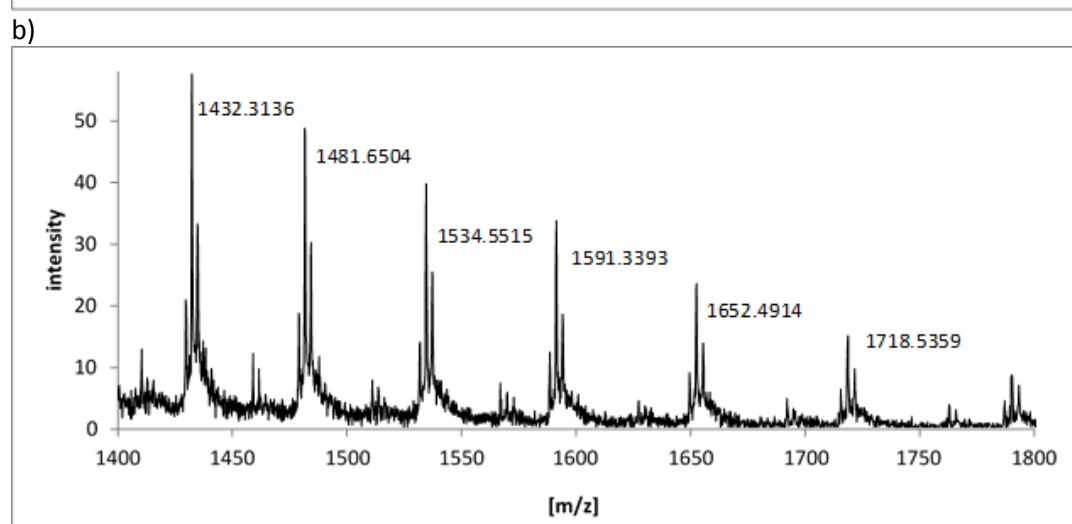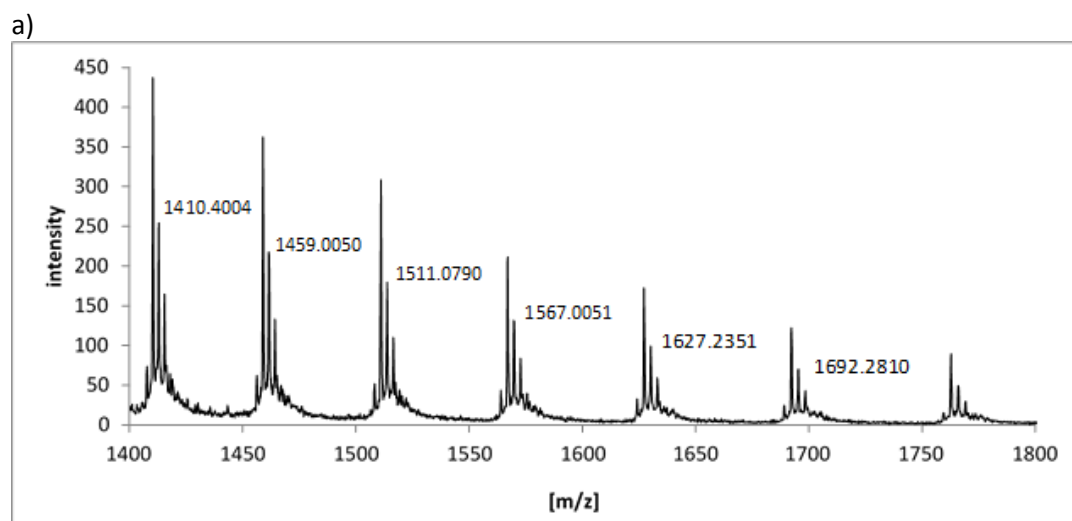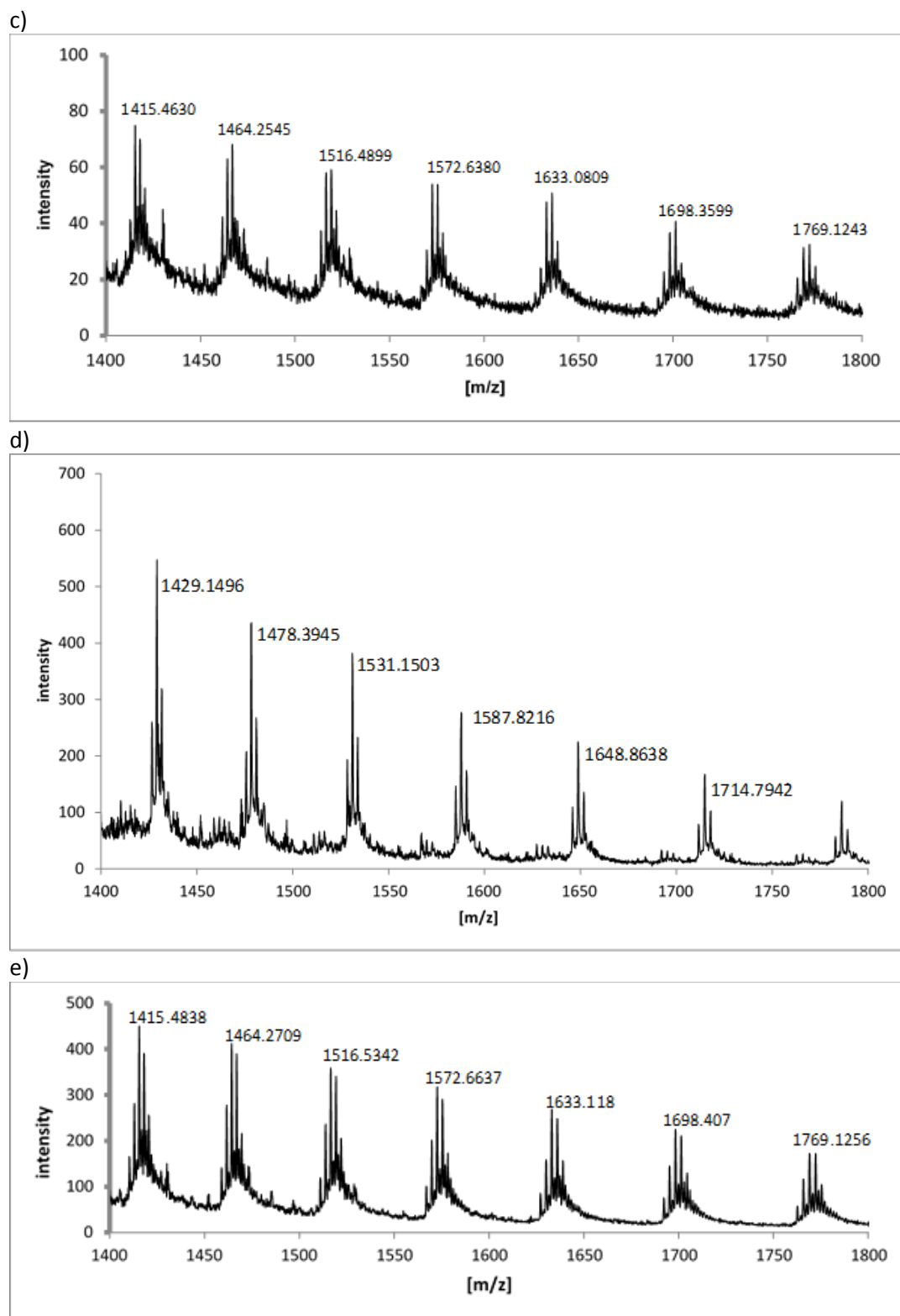
Figures S3. Charge series after 1 h of incubation of JNK3 wt with 5 molar excess of inhibitor. a) Control, only JNK3; b) **7**, reactive warhead (para/meta substituent), mass shift according to mass of ligand. Small residual of unmodified JNK3 present; c) **8**, no reactive warhead (saturated counterpart of **7**), spectrum identical to control, no covalent bonding to JNK3; d) **21**, reactive warhead, mass shift according to mass of ligand; e) **23**, no reactive warhead (saturated counterpart of **21**), spectrum identical to control, no covalent bonding to JNK3.
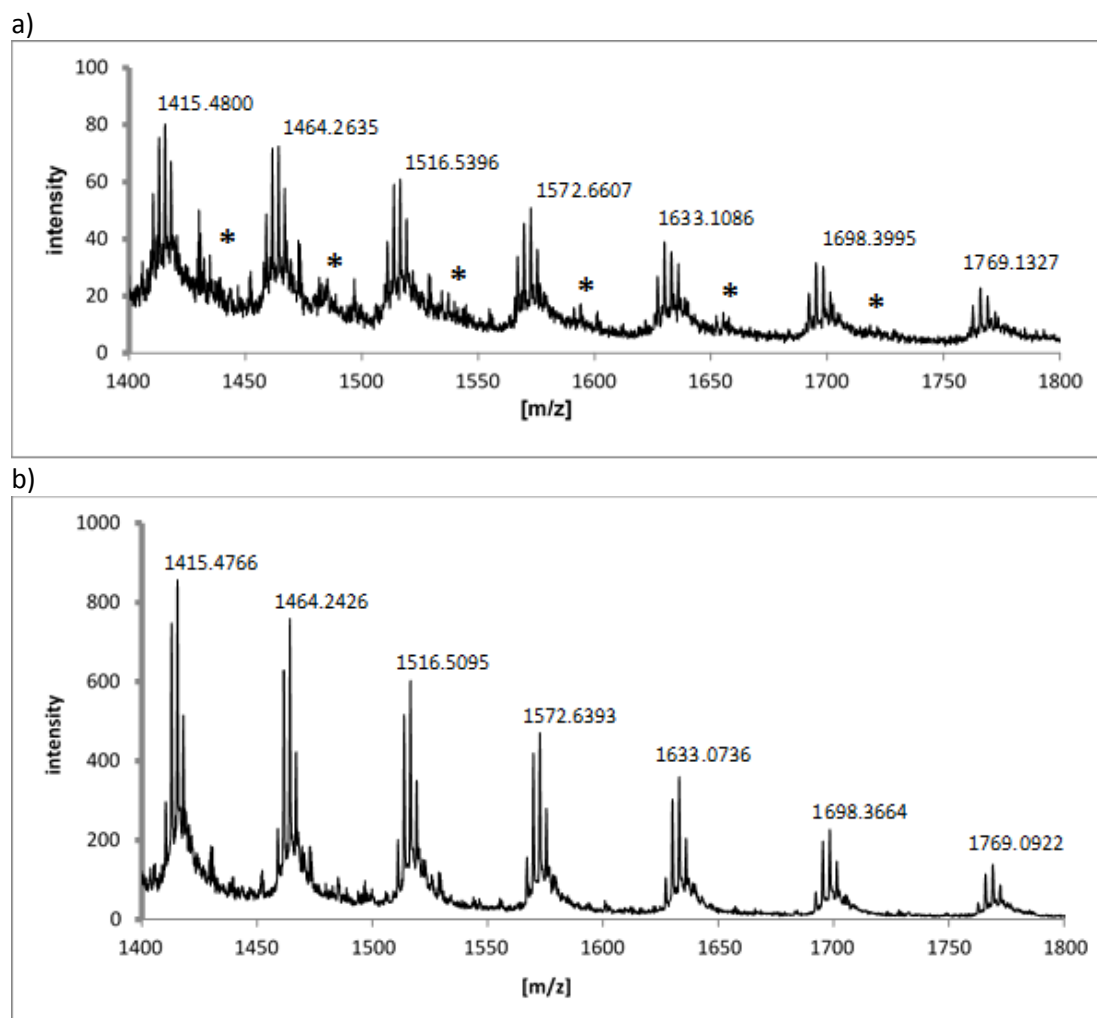
Figures S4. Charge series after 1 h of incubation of JNK3 wt with 5 molar excess of inhibitor. a) Binding assay with **14**. For **14**, that exhibits some activity in IC$_{50}$-Assay, binding to some extend was observed, as minor charge series denoted by (*). For identification, most intense charge series of [JNK3-3xβME-1x410] was chosen, with m/z-series of 1434.8416; 1484.2622; 1537.2659; 1594.1314; 1655.3635; 1721.6486; 1793.2949; for deconvolution. Binding ratio was calculated comparing area in 50mDa corridor around target mass after baseline subtraction to 29% bound and 71 unbound. Unmodified [JNK3+3xβME] charge series: 1415.4822; 1464.2572; 1516.5161; 1572.6461; 1633.0937; 1698.3772; 1769.1009; b) **15**, no reactive warhead (saturated counterpart of **14**), spectrum identical to control, no covalent bonding to JNK3.
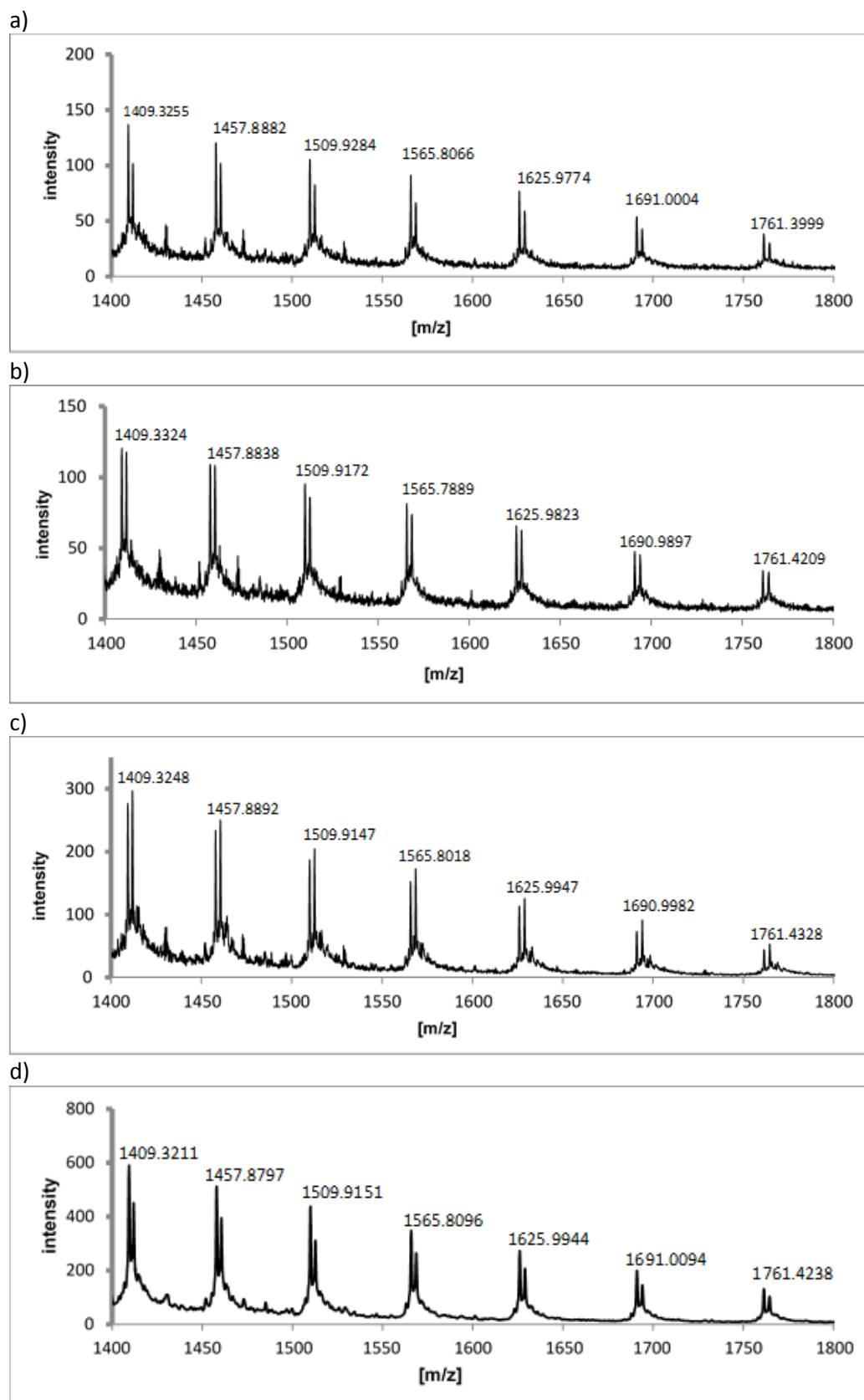
Figure S5. Charge series after 1h of incubation of with Cys154 → Ala154-mutant of JNK3 with 5 molar excess of inhibitor. a) control; b) **7**; c) **21**; d) **14**; No covalent modification by inhibitors could be observed, precluding reactivity with other than target Cys154. Protein was visible as β-ME-adduct-series ranging from 0 to 4 mercaptoethanoylations. For mass accuracy, the intense 1x β-mercaptoethanol-adduct was chosen (left one of each peak doublets) for evaluation. Theoretical charge series of unmodified [Cys→Ala-JNK3-1xβ-ME]: 1409.3247; 1457.8874; 1509.9188; 1565.8044; 1625.9889; 1690.9882; 1761.4041.

Table S2. Calculations of ligand masses from charge series deconvolution.

Cmp #	Protein Mass			Ligand Mass shift	
	Calculated	Measured	error ppm	Expected	Found
7	42784.5576	42784.2239	-7.8	578.3280	578.6617
8	42205.8959	42205.51046	9.1	0	0
15	42205.8959	42205.9047	0.2	0	0
21	42690.4710	42690.1030	-8.6	484.2071	484.2071
23	42205.8959	42206.342	10.5	0	0

All calculations show the covalent binding of the inhibitors could be proved by sub Dalton mass accuracy. Inhibitor models without reactive warhead do not show any binding to JNK3.

Note: Ligand masses are denoted as average masses, as protein mass spectrometry in the mass range of the target is performed on average rather than monoisotopic peaks. For easy comparison the deconvoluted masses of the protein-ligand adducts have been corrected for mercaptoethanoylation.

Materials and Methods

LC-QTOF-µESI-MS of intact JNK3-Ligand adducts

Analysis was performed on a Sciex (Concord, Ontario, Canada) 5600+ TripleTOF mass spectrometer equipped with an Agilent (Waldbronn, Germany) 1290 Series UHPLC and a PAL-xts (CTC, Zwingen, Switzerland) autosampler for injection. Chromatographic separation was conducted using a ProsWIFT (Thermo Scientific, Waltham, USA) 100 x 0.5 mm monolithic polystyrene-dinvinylbenzene copolymer capillary column. Mobile phases consisted of A: H_2O, 0.1 % formic acid and B: acetonitrile, 0.1 % formic acid. Gradient profile was 0 - 1.5 min 20 % B, 1.5 - 4.5 min: 20 – 80 % B, 4.5 - 6 min 95 % B. Flow rate: 50 µl/min. First 4 minutes were switched to waste to prevent buffer components from entering the ESI-source. Column temperature was kept at 22 °C. Injection volume was 5 µl, corresponding to 5 pmol JNK3 absolute per injection. Acetonitrile and formic acid were of Ultra-MS-grade, purchased from Carl Roth (Karlsruhe, Germany). Ultra-MS-grade water was produced using an Elga (Veolia Water Technologies, Germany) Purelab Ultra system.

For µflow-electrospray ionization (µESI) the Sciex Duospray Ionsource was outfitted with a 50 µm I.D. PEEKsil-stainles steel tip hybrid micro-electrode. Mass spectrometry was done using SCIEX "Intact Protein Mode"-Script with a detector voltage (CEM) lowered by 100V. MS parameters were as follows: gas1 (nebulizer) 50 psi, gas2 (drying gas) 40 psi, curtain gas 30 psi, source temperature 400 °C, ion source floating voltage (ISFV) 5200V. Collision energy (CE) was set to protein mode typical 30V facilitating clearance of Q2 (collision cell). Declustering potential (DP) was set to 230V for removal of non-covalent species. Q1 transmission: 100 % at 1250 m/z. time bins: 100. Positive-TOF-mode mass scan: 500-4000 m/z.

Sequence of recombinant JNK3 (average mass was 42205.8959 Da)

GGSMSKSKVDNQFYSVEVGDSTFTVLKRYQNLKPIGSGAQGIVCAAYDAVLDRNVAIKKLSRPFQNQTHAKRAYRELVLMK
CVNHKNIISLLNVFTPQKTLEEFQDVYLVMELMDANLCQVIQMELDHERMSYLLYQMLCGIKHLHSAGIIHRDLKPSNIVVKS
DCTLKILDFGLARTAGTSFMMTPYVVTRYYRAPEVILGMGYKENVDIWSVGCIMGEMVRHKILFPGRDYIDQWNKVIEQLG
TPCPEFMKKLQPTVRNYVENRPKYAGLTFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISVDDALQHPYINVWY
DPAEVEAPPPQIYDKQLDEREHTIEEWKELIYKEVMNSE

Sequence of mutant JNK3 (average mass 42173.8359 Da)

GGSMSKSKVDNQFYSVEVGDSTFTVLKRYQNLKPIGSGAQGIVCAAYDAVLDRNVAIKKLSRPFQNQTHAKRAYRELVLMK
CVNHKNIISLLNVFTPQKTLEEFQDVYLVMELMDANL(A)QVIQMELDHERMSYLLYQMLCGIKHLHSAGIIHRDLKPSNIVV
KSDCTLKILDFGLARTAGTSFMMTPYVVTRYYRAPEVILGMGYKENVDIWSVGCIMGEMVRHKILFPGRDYIDQWNKVIEQ
LGTPCPEFMKKLQPTVRNYVENRPKYAGLTFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISVDDALQHPYINV
WYDPAEVEAPPPQIYDKQLDEREHTIEEWKELIYKEVMNSE

Binding assay
Binding assay was performed in 60 mM ammonium acetate. Reaction volume was 200 µl, containing 200 pmol of JNK3 and 1 nmol of inhibitor and 0.5 % (v/v) DMSO. Binding reaction was performed at 37 °C for 1h. Samples were analyzed immediately afterwards, stored at 4 °C in the autosampler during analysis.

In vitro reactivity study of inhibitor 7 with glutathione

The *in vitro* reactivity of covalent inhibitor **7** was determined similar to a protocol published by Schmidt *et al.*[8] This protocol was adapted by using the conditions of our in house kinase assay (buffer, reaction time and temperature). *N*-Phenylacrylamide (PAA) served as a positive control since it is known that PAA forms adducts with glutathione under physiological conditions.[9]

Buffer: 50 mM Tris [pH 7.5], 10 mM $MgCl_2$, 10mM β-glycerolphosphate, 100 µg/ml BSA, 1 mM dithiothreitol, 0.1 mM Na_3VO_4

Table S3. Incubation of compound **7** (10 µM) and PAA (10 µM) with 5 mM glutathione

time [min]	cpd **7** [%]	PAA [%]
0	99.9	100
9	99.2	91.6
18	98.5	86.1
27	97.6	81.0
36	96.8	76.7
45	96.1	70.5
54	93.7	66.5

In Vitro Metabolism Studies

Pooled human male liver microsomes (HLM) were purchased from Sigma-Aldrich (Steinheim, Germany). These microsomes were characterized in protein and cytochrome P-450 content. All incubations were made in the presence of an NADPH-regenerating system, consisting of 5 mM Glucose-6-phosphate, 5 U/mL Glucose-6-phosphate dehydrogenase and 1 mM NADP⁺. The substrate (100 µM), the NADPH regenerating system and 3.8 mM $MgCl_2$ x 6 H_2O in 0.1 M Tris-buffer (pH = 7.4 at 37 °C) were preincubated for 5 min in a shaking heating block at 37 °C and 550 rpm.[10] The incubation mix was split into 75 µL aliquots and the reaction was started by addition of the HLM. Thereby the microsomal protein content was standardized to 1 mg/mL. To follow the course of metabolism, the reaction tubes were quenched at selected time points (0, 10, 20, 30, 45, 60, 130/135 and 180/190 min; analyte **7/21**) by adding 225 µL icecooled internal standard at a concentration of 100 µM for **7** and 20 µM for **21** in acetonitrile (ACN). The samples were vortexed for 30 s and centrifuged (19800 relative centrifugal force/4°C/20 min). The supernatant was directly used for LC-MS analysis. All incubations were conducted in triplicates and incubations with heat-inactivated HLM were used to proof that analyte reduction results from metabolic degradation only. In all incubations a limit of 1 % organic solvent was not exceeded.[11]

Screening of Metabolites by LC-MS Analysis

Metabolite formation was analyzed with an Alliance 2695 Separations Module (Waters GmbH, Eschborn). Samples were maintained at 4 °C, the column temperature was set to 40 °C and injection volume was 10 µL. The chromatographic separation of analyte **7** was performed on a Waters Symmetry®C18 (150 x 4.6 mm; 5 µm); **21** on a Phenomenex Synergi Max RP column (150 x 4.6 mm; 5 µm) with a precolumn of the same material, respectively. An isocratic elution of 11.0 min for **7** and 8.0 min for **21** with 63 % solvent A (90 % H_2O, 10 % ACN, 0.1 % formic acid) and 37 % solvent B (ACN, 0.1 % formic acid) at a flow rate of 400 µL/min was used. The detection was performed on a Micromass Quattro micro triple quadrupole mass spectrometer (Waters GmbH, Eschborn) using the electrospray ionization in the positive-mode. Correspondent to the analyte the spray voltage was set to 3.5 - 4.25 kV. The heated capillary operated at 250 °C and the desolvation gas flow worked at 500 L/h.
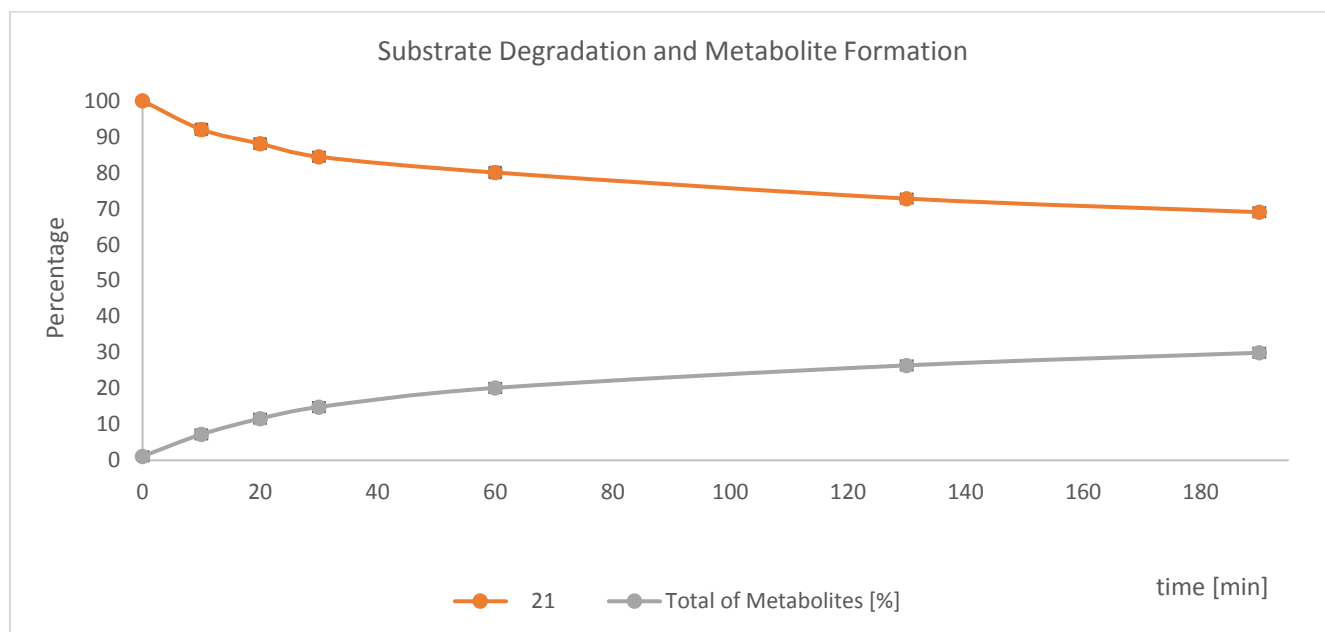
Figure S6. Degradation scheme of compound **21** while incubating in HLM for 190 min.

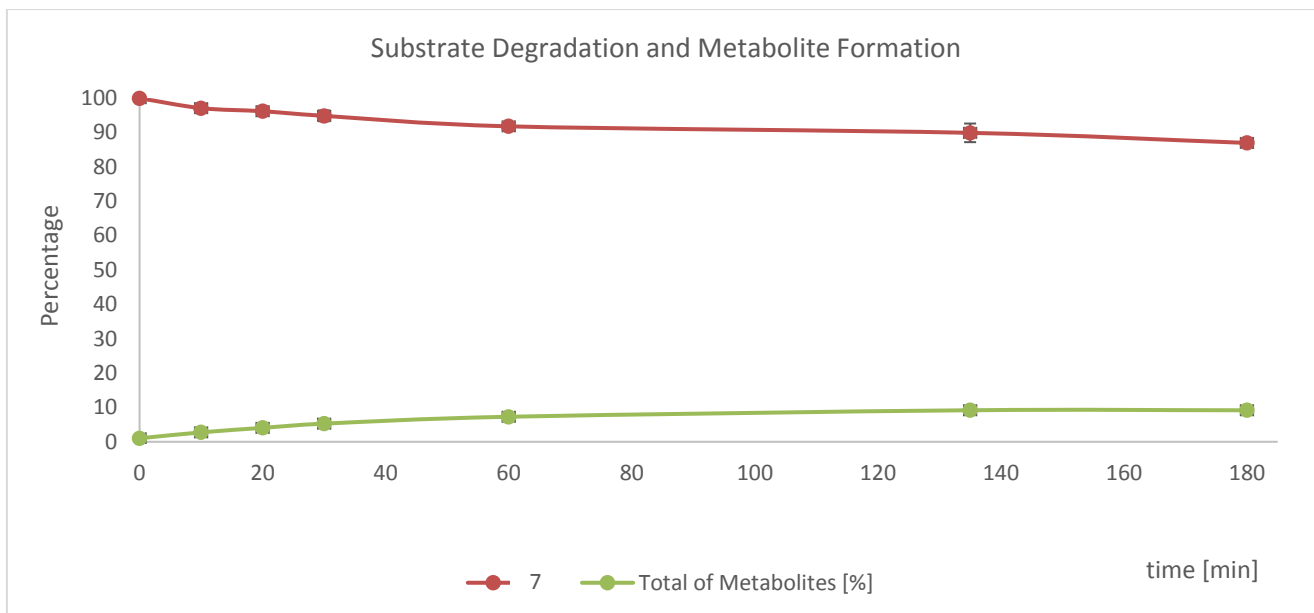**Figure S7.** Degradation scheme of compound **7** while incubating in HLM for 180 min.

Table S4. Metabolic stability

| Cmp | IC$_{50}$ [nM] | | OxMet[a] |
	JNK3	p38α	
21	2	1,952	69[b]
7	<1	36	87[c]

[a]Percent remaining after incubation with HLM; [b]incubation time: 190 min; [c]incubation time: 180 min.

References

1.   Laufer, S.; Hauser, D.; Stegmiller, T.; Bracht, C.; Ruff, K.; Schattel, V.; Albrecht, W.; Koch, P. Tri- and tetrasubstituted imidazoles as p38 alpha mitogen-activated protein kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6671-6675.

2.   Ansideri, F.; Lange, A.; El-Gokha, A.; Boeckler, F. M.; Koch, P. Fluorescence polarization-based assays for detecting compounds binding to inactive c-Jun N-terminal kinase 3 and p38α mitogen-activated protein kinase. *Anal. Biochem.* **2016**, *503*, 28-40.

3.   Schrödinger, LLC. Small-Molecule Drug Discovery Suite 2015-4: Glide, version 6.9. In New York, 2015.

4.   Schrodinger, LLC. The PyMOL molecular graphics system, version 0.99. In New York, 2006.

5.   Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298*, 1912-1934.

6.   Goettert, M.; Graeser, R.; Laufer, S. A. Optimization of a nonradioactive immunosorbent assay for p38 alpha mitogen-activated protein kinase activity. *Anal. Biochem.* **2010**, *406*, 233-234.

7.   Goettert, M.; Luik, S.; Graeser, R.; Laufer, S. A. A direct ELISA assay for quantitative determination of the inhibitory potency of small molecules inhibitors for JNK3. *J. Pharm. Biomed. Anal.* **2011**, *55*, 236-240.

8.   Schmidt, T. J.; Ak, M.; Mrowietz, U. Reactivity of dimethyl fumarate and methylhydrogen fumarate towards glutathione and N-acetyl-L-cysteine - Preparation of S-substituted thiosuccinic acid esters. *Bioorg. Med. Chem.* **2007**, *15*, 333-342.

9.   Cee, V. J.; Volak, L. P.; Chen, Y. P.; Bartberger, M. D.; Tegley, C.; Arvedson, T.; McCarter, J.; Tasker, A. S.; Fotsch, C. Systematic study of the glutathione (GSH) reactivity of N-arylacrylamides: 1. Effects of aryl substitution. *J. Med. Chem.* **2015**, *58*, 9171-9178.

10.   Baur, B.; Storch, K.; Martz, K. E.; Goettert, M. I.; Richters, A.; Rauh, D.; Laufer, S. A. Metabolically stable dibenzo[b,e]oxepin-11(6H)-ones as highly selective p38 MAP kinase inhibitors: optimizing anti-cytokine activity in human whole blood. *J. Med. Chem.* **2013**, *56*, 8561-8578.

11.   Chauret, N.; Gauthier, A.; Nicoll-Griffith, D. A. Effect of common organic solvents on in vitro cytochrome P450-mediated metabolic activities in human liver microsomes. *Drug Metab. Dispos.* **1998**, *26*, 1-4.

Contents lists available at ScienceDirect

Analytica Chimica Acta

Papain-functionalized gold nanoparticles as heterogeneous biocatalyst for bioanalysis and biopharmaceuticals analysis

Siyao Liu, Markus Höldrich, Adrian Sievers-Engler, Jeannie Horak, Michael Lämmerhofer[*]

Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Auf der Morgenstelle 8, 72076, Tübingen, Germany

HIGHLIGHTS

- An immobilized enzyme reactor was synthesized with gold nanoparticles (GNPs) as support.
- Polyelectrolyte layer-by-layer GNP modification with subsequent papain coupling.
- Resonant mass measurement technology for determination of surface coverage.
- Enhanced catalytic efficiency of immobilized papain compared to free papain.
- Successful digestion of human IgG with papain-GNP verified by HPLC-μESI-QTOF-MS in intact protein mode.

GRAPHICAL ABSTRACT
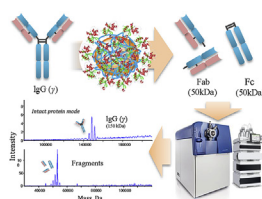
ARTICLE INFO

ABSTRACT

Surface-modified gold nanoparticles (GNPs) were synthesized via layer-by-layer process with alternating cationic polyallylamine and anionic poly(acrylic acid) polyelectrolyte layers leading to a highly hydrophilic biocompatible shell supporting colloidal stability. Afterwards, papain was covalently immobilized on the modified GNPs via amide coupling between the amino groups on papain and the terminal carboxylic groups of the modified GNPs by using N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide and N-hydroxysulfosuccinimide sodium as coupling agents. The resultant papain-functionalized gold nanoparticles were characterized by surface plasmon resonance, dynamic light scattering and zeta potential measurements. The new technology resonant mass measurement was applied for determining the average number of papain molecules immobilized per GNP by measurement of the single nanoparticle buoyant mass in the range of femtograms. The activity of the immobilized enzyme was estimated by determination of the kinetic parameters (K_m, V_{max} and k_{cat}) with the standard chromogenic substrate N_α-benzoyl-DL-arginine-4-nitroanilide hydrochloride. It was found that K_m of immobilized and free enzyme are in the same order of magnitude. On contrary, turnover numbers k_{cat} were significantly higher for GNP-conjugated papain. Further, the gold nanobiocatalyst was applied for digestion of polyclonal human immunoglobulin G to yield protein fragments. The resultant fragment mixture was further analyzed by high-performance liquid chromatography-microelectrospray ionization-quadrupole-time-of-flight mass spectrometry, which demonstrated the applicability of the bioreactor based on papain functionalized

GNPs. The immobilized papain not only has higher catalytic activity and better stability, but also can be easily isolated from the reaction medium by straightforward centrifugation steps for reuse.

1. Introduction

Enzymes with a number of excellent properties (high activity, selectivity and specificity) are extensively used to catalyze a wide range of commercially important processes [1]. However, some limitations exist for the applications of enzymes, such as low stability and high sensitivity to the employed conditions. Therefore, immobilization techniques of enzymes (leading to immobilized enzyme reactors, IMERs) emerged as a powerful strategy to overcome some of these limitations [2—10]. They received particular interest due to some advanced properties such as high catalytic efficiency, improved stability, elimination of self-digestion, flexible control of the reaction, easy removal after reaction, no contamination of the product with enzyme and repetitive usage [4,11—14]. Enzymes immobilized on nanomaterials possess considerable prospect in various fields, because the catalytic properties of enzymes can be flexibly combined with the unique features of nanomaterial structures [12,15—19].

Silica nanoparticles [20], polystyrene [4] and magnetic nanoparticles [21] have been frequently utilized for the enzyme immobilization by covalent binding, entrapment, adsorption, ionic binding, affinity binding and so on [22—24]. A growing number of studies suggested that immobilizing enzyme on nanomaterials can enhance reaction rates while improving enzyme stability [25—28]. Amongst the nanomaterials, gold nanoparticles (GNPs) have received great attention in biology, biochemistry and biomedical research areas due to controlled geometrical, excellent optical and flexible surface chemical properties [29]. GNPs can be synthesized in a straightforward and low-cost method by reduction of gold (III) chloride. Due to the chemical inertness of gold they possess excellent chemical stability and due to charged capping groups on their surface also a high colloidal stability. They can be easily further functionalized with appropriate organic or biological ligands which form the basis for their extremely broad applications. Like for other metal nanoparticles, a surface plasmon resonance (SPR) band can be observed in the visible spectrum which is the result of the collective oscillation of electrons in the conduction band of gold nanoparticles in resonance with a specific wavelength of incident light [30]. For gold nanoparticles it results in a strong absorbance band in the visible region (around 525 nm). The exact wavelength maximum as well as width of the absorption band depends on the nanoparticle size and shape, size distribution, and morphological uniformity of GNPs [31,32]. SPR measurement is therefore a straightforward and useful tool to characterize GNPs and monitor the success of surface modification as well as to evaluate the aggregation status of GNP suspensions. If a second peak can be observed in the red-shifted region of the UV-VIS spectrum, it implies that aggregation of GNPs occurred. The large surface-to-volume ratio of GNPs provides considerable promise for the bioconjugation with various receptor molecules (e.g. proteins, DNAs, aptamers, antibodies, or lectins) [33]. The bioconjugation chemistries used for the immobilization are generally derived from the protein labeling chemistries by using various commercial crosslinkers based on maleimides, succinimidyl esters and so on. In addition, carboxylic and amino groups of proteins are widely used as the reactive sites for conjugation through amide bond formation [34].

Papain, a cysteine protease with broad specificity present in the latex of *Carica papaya*, catalyzes the hydrolysis of peptide, ester and amide bonds. Therefore, it is extensively employed in food, pharmaceutical, biology and biomedical researches [35—40]. In recent years, preparing Fab fragments with papain via specific digestion above the hinge region (the site of the disulfide bonds which connect the two heavy chains) of the whole antibody attracted some attention. The Fab fragment is a monovalent antibody structure harboring the CDR (complementarity-determining region) without Fc portion. It can still specifically recognize and bind to antigens. In biopharmaceuticals research and quality control, Fab fragments are prepared to characterize the protein on an intermediate level instead of the whole intact immunoglobulin (middle up and middle down). Middle-up refers to the analysis of an antibody after its cleavage into large fragments, e.g. by reduction or limited proteolytic cleavage. Proteins of such smaller size are easier to analyze than larger proteins like intact antibody and can be detected with better sensitivity. Middle-down refers to the mass measurement of the gas phase fragmentation of antibody fragments, in analogy to the mass measurement of the gas phase fragmentation of intact protein in top-down analysis [41]. An antibody digested by papain generates two Fab fragments (about 50 kDa) and one Fc fragment (also about 50 kDa). Pure Fab fragment can be obtained after purification of digests by affinity chromatography (with protein A to remove the Fc part), ion exchange, and size exclusion chromatography [39,40]. The main goal of this study was to prepare stable immobilized papain on GNPs for use in sample preparation protocols of antibody analytics. For this purpose, it is necessary that the obtained immobilized papain nano-biocatalyst has sufficient colloidal stability and satisfactory bioactivity to yield useful Fab fragments for analytical characterization by mass spectrometry via enzyme digestion. In this research, surface-modified GNPs with highly hydrophilic and biocompatible shell were first prepared from oppositely charged polyelectrolytes via their alternating deposition by a layer-by-layer process (LBL) onto citrate stabilized GNPs [42—44]. The immobilization process was carried out by using cross-linker agents to link the carboxylic group on the surface of carboxy-modified GNPs with the amino group on papain [42—44]. The covalent (instead of adsorptive) bonding was selected for irreversible immobilization to avoid enzyme leaching from the support surface [45]. The number of immobilized papain molecules on GNPs is an important figure for the determination of the immobilization efficiency. Resonant mass measurement (RMM) was applied herein to measure the mass increase of the nanoparticles upon immobilization of papain and determine the surface coverage [46]. It is well-known that papain, like other proteases, is prone to auto-digestion. The resultant peptides will cause the contamination of the protein digestion products. Therefore, immobilization of proteases like papain on a solid support is a good solution to eliminate this kind of contamination produced by auto-digestion [47]. In addition, the papain-functionalized GNPs can be easily removed from the reaction solution by a simple centrifugation process, which also terminates the digestion. Therefore, using immobilized papain makes it easy to control the reaction time without using any stop reagents which normally introduce contamination to the product solution.

2. Experimental

2.1. Materials

Papain (EC 3.4.22.2) was purchased from Acros Organics (Geel, Belgium). Human immunoglobulin G (Gammanorm) (h-IgG) was obtained from Octapharma (Heidelberg, Germany). The lot employed in this study was already expired. Gold (III) chloride trihydrate (HAuCl$_4$·3H$_2$O), trisodium citrate, poly(acrylic acid, sodium salt) solution (PAA, Mw. ~ 15,000, 35% in H$_2$O), poly(allylamine hydrochloride) (PAH, Mw.~17,500), N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide (EDC), N-hydroxysulfosuccinimide sodium (sulfo-NHS), sodium phosphate monobasic dihydrate, sodium carbonate anhydrous, sodium hydroxide, K-Na-tartrate (potassium sodium tartrate), N-benzoyl-DL-arginine-4-nitroanilide hydrochloride (BApNA) and p-nitroaniline, copper (II) sulfate, Folin & Ciocalteu's phenol reagent glycerol, 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris base), acrylamide/bis-acrylamide (30% solution), glycerol, and a ProteoSilver TM silver staining kit were obtained from Sigma-Aldrich (Munich, Germany). Ammonium persulfate was from AppliChem (Darmstadt, Germany). HPLC-grade acetonitrile was purchased from JT Baker Chemical Co. (Deventer, The Netherlands). Double deionized water (produced by Elga Purelab Ultra ELGA, System LabWater, Celle, Germany) was used throughout synthesis of functionalized GNPs and analytical procedures including LC-MS.

2.2. Preparation of papain-functionalized GNPs

Gold nanoparticles were prepared according to the Turkevich-Frens method with some modifications [48]. In brief, 25.25 mg HAuCl$_4$ were dissolved in 50 mL water (with a final concentration of 1.14 mM) and then heated at 170 °C under reflux and constant stirring for 10 min. Afterwards 6.25 mL trisodium citrate (2.28 mM) were added and heated for another 10 min under reflux and constant stirring. Then the colloidal solution was kept stirring without heating for an additional 60 min to cool down to room temperature. The obtained citrate-GNP solution was stored at 4 °C for further usage.

The surface modification of GNPs was realized via the layer-by-layer process (LDL) reported by Schneider and Decher [49]. In brief, citrate-stabilized GNP suspensions were first centrifuged (12,000 rpm, 10 min), the supernatant discarded and the pellet washed with water to remove citrate from the supernatant. Afterwards 10 mL citrate-GNP suspension were added dropwise to the same volume of PAH solution (20 mg mL^{-1}, 10 mL) under continuous vigorous stirring, and then kept gently stirring for 30 min at room temperature in the dark. To remove excess PAH, the reaction mixture was centrifuged (12,000 rpm, 15 min), the supernatant discarded and the pellet washed twice with 10 mL water. The second layer coating process was performed according to the first coating step but PAH was replaced by PAA (20 mg mL^{-1}). After removal of the supernatant and two washing steps performed as above, the modified GNPs (GNP/PAH/PAA) were resuspended in 20 mM phosphate buffer (pH 6.8). Subsequently, 20 μL EDC (12 mM) and 20 μL sulfo-NHS (60 mM) solutions in water were added into 760 μL GNP/PAH/PAA suspension to yield activated carboxylic groups in the form of sulfo-NHS esters which were further coupled with amino groups on the surface of papain. For this purpose, 800 mL activated GNP/PAH/PAA suspension was added dropwise into 200 μL papain solution with four different concentrations (1 mg mL^{-1}, 2 mg mL^{-1}, 5 mg mL^{-1}, 10 mg mL^{-1}) under continuous vigorous stirring. The reaction mixture was then stirred gently for 2 h, and allowed to stand for 12 h at room temperature. To remove excess papain, the obtained papain-GNP

suspension was centrifuged (12,000 rpm, 15 min), the supernatant discarded and the pellet washed with water. The wash procedure was repeated two times and finally the pellet was resuspended in 20 mM phosphate buffer (pH 6.8). The papain-GNP solution was stored at 4 °C prior to use.

2.3. Characterization

After each step of surface modification, SPR bands of GNPs were measured in the wavelength range between 300 nm and 800 nm with a UV-1600PC spectrophotometer (from VWR, Darmstadt, Germany) to determine the size changes as well as nanoparticle dispersion and aggregation status. Size and size distribution of GNPs were studied by dynamic light scattering (DLS) based on the Brownian motion of the particles using a Zetasizer Nano ZS from Malvern Instruments (Herrenberg, Germany). The stability of the GNP dispersions was determined by zeta potential (ζ-potential) measurements before and after each modification step with the Zetasizer Nano ZS. The Zetasizer Nano ZS was equipped with a He−Ne laser and each measurement was performed at 173 °C backscatter detection mode. Each sample was diluted with water (1:5; v/v) before detection and measured in triplicate. Each reported value represents the mean of at least 15 subruns.

2.4. Quantitative determination of the immobilized papain

RMM using the Archimedes instrumentation from Malvern Instruments was employed to determine the amount of immobilized papain on nanoparticles. All the samples were diluted with water (1:5; v/v) before each measurement. Before measurements, (mass) calibration of the microchip was performed with 1 μm latex standard. The microchannel was then flushed with purified water before analysis, and the impurities in the system were subsequently removed by sneezing steps. The limit of detection was set to 0.030 Hz, which was well above the noise level and optimal for detection of nanoparticles. Each measurement was carried out at room temperature, and the determination numbers of particles were set to 2000 to ensure satisfactory statistics of the nanoparticle distributions. The density of gold was 19.320 g mL^{-1}, and the density of fluid was 1 g mL^{-1}. With these conditions, the average buoyant masses of particles before and after immobilization of papain were measured. The average real masses (drymass) of papain-GNPs were obtained from the conversion of buoyant mass.

2.5. Assays of papain activity

In order to determine the activity of both free papain and immobilized papain, N-benzoyl-DL-arginine-4-nitroanilide hydrochloride (BApNA) was selected as a substrate for the kinetic assays. In principle, BApNA was cleaved by papain to produce p-nitroaniline which was measured with the Versa max microplate reader (from Molecular Devices LLC, Biberach, Germany) at an absorbance of 410 nm. A concentration series of BApNA was digested with papain (free papain and immobilized papain, respectively) in 20 mM phosphate buffer (pH 6.8) at 37 °C. For the digestion with papain, different volumes of BApNA stock solutions in phosphate buffer (2 mM) and 50 μL enzyme solution (2 mg mL^{-1}) were added to the 96 wells, and different volumes of phosphate buffer were filled to obtain the final volume of 100 μL. Digestion was performed at 37 °C controlled by the microplate reader, and the absorbance of all the samples were recorded every two minutes with kinetics scan mode. For the digestion with papain-GNP, 50 μL papain-GNP solution in 20 mM phosphate buffer (pH 6.8) were added and digestion was performed as described above for 10 min. After stopping the digestion by immediate centrifugation, the

supernatant was analyzed with microplate reader in endpoint mode. The Michaelis-Menten plot and Lineweaver-Burk plot were built based on the calculated initial velocity. Finally, the kinetic parameters (K_m, V_{max} and k_{cat}) were calculated by using the Lineweaver-Burk equation. All experiments were performed in triplicate.

2.6. Antibody digestion with papain-functionalized GNPs

For the antibody digestion, immunoglobulin G solutions were first digested with papain-functionalized GNPs. For digestion, 10 mg mL^{-1} IgG solution was diluted with water to 0.5 mg mL^{-1} and used as stock solution. Subsequently, 100 μL of above solution of immobilized papain on GNPs was mixed with 100 μL IgG solution (0.5 mg mL^{-1}), and incubated at 50 °C with gentle shaking at 450 rpm with a Thermoshaker (Peqlab, Erlangen, Germany) for 4 h, 8 h and 24 h, respectively. After digestion, the GNP-based nano-biocatalyst was easily separated from the crude digest by a centrifugation step (12,000 rpm, 10 min). Finally, the clear digestion solution (supernatant) was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and high-performance liquid chromatography-microelectrospray ionization-quadrupole-time-of-flight mass spectrometry (HPLC-μESI-QTOF-MS) in intact protein mode (*vide infra*).

2.7. SDS-PAGE

A Mini-PROTEAN 3Cell with Mini-PROTEAN 3 PowerPac 300 system (200/240 V) obtained from Bio-Rad Laboratories (Munich, Germany) was applied for the SDS-PAGE. The Precision plus protein mix (from Bio-Rad Laboratories) was used as the standard protein marker. A ten percent Tris-HCl gel with a thickness of 0.75 mm and 10 sample wells was prepared according to the product technical information from Bio-Rad Laboratories. All the samples were mixed with non-reducing sample buffer in a ratio of 1:1 (v/v) before application of 7.5 μL into each well. 2 μL of the protein standard were applied. After development of the gel, the protein bands were stained with a Proteosilver Plus Silver Stain Kit from Sigma-Aldrich (Munich, Germany) according to the product technical information.

2.8. HPLC-μESI-QTOF-MS method for determination of IgG masses before and after digestion

An Agilent Technologies (Waldbronn, Germany) 1290 Series UHPLC and a PAL-xts (CTC, Zwingen, Switzerland) autosampler were used for chromatographic separation and injection, respectively. A ProSwift RP-4H monolithic capillary column (500 μm ID × 10 cm) from Thermo Scientific (Waltham, Massachusetts, USA) was applied for protein separation. The LC gradient is shown in Table 1. Solvent A consisted of ultrapure water (from Elga Purelab Ultra ELGA) with 0.1% (v/v) formic acid, and solvent B consisted of

Table 1
The LC gradient profile for HPLC-μESI-TOF-MS. Solvent A consisted of water with 0.1% (v/v) formic acid, and solvent B consisted of acetonitrile with 0.1% (v/v) formic acid.

Step	Total Time (min)	Flow Rate (μL/min)	%A	%B
0	0.0	50	80.0	20.0
1	2.0	50	80.0	20.0
2	12.0	50	65.0	35.0
3	13.0	50	5.0	95.0
4	17.0	50	5.0	95.0
5	17.5	50	80.0	20.0
6	20.0	50	80.0	20.0

acetonitrile with 0.1% (v/v) formic acid. The flow rate was set to 50 μL/min and injection volume was 3 μl.

Detection was carried out with a Triple TOF 5600+ quadrupole time-of-flight mass spectrometer from AB Sciex (Concord, Ontario, Canada) equipped with a DuoSpray source operated in positive mode using a 50 μm ID microelectrospray ionization (μESI) needle (from Sciex). The source temperature was set to 400 °C and the ionspray voltage floating was 5100 V. Nebulizer gas (GS1) was set to 50 psi, drying gas (GS2) to 40 psi, and curtain gas to 30 psi. Finally, the declustering potential (DP) was set to 230 V, and the collision energy (CE) to 30 V. TOF-MS scan ranged from 500 to 4000 Da. For intact protein detection Sciex's Intact-Protein script was activated, Q1 transmission was set to 100% at 1250 m/z, and sensitivity was increased by summing 60 time bins. PeakView from Sciex containing the Bio Tool Kit software package was used for data processing and identification of the proteins by a deconvolution procedure.

3. Result and discussion

3.1. Preparation of papain functionalized GNPs

For analytical and bioapplications, the GNPs should support the property of colloidal stabilization in physiological medium as well as provide a large surface for a high bioconjugation capacity with receptor molecules or, in present case, enzymes [50]. To cope with this goal a LBL activation of the GNP surface was employed before bioconjugation of the enzyme. The schematic presentation for the preparation process of the functionalized GNPs is shown in Fig. 1. First, citrate stabilized GNPs were prepared according to the Turkevich-Frens method [48]. In this reaction, citrate played the role as both reducing and stabilizing reagent. Capping citrate anions on the GNP surface provide them with negative charges and give them colloidal stability due to electrostatic repulsion of the particles precluding their van der Waals driven aggregation. Modification with charged polymer reagents (i.e. polyelectrolytes) was supposed to further increase the colloidal stability of GNPs by formation of soft shells on the GNP surface based on electrostatic repulsion and steric hindrance. Moreover, the coated polyelectrolytes provide sufficient anchor groups for covalent bonding of protein. This biocompatible and highly hydrophilic shell was obtained by an alternating deposition of oppositely charged polyelectrolytes on the nanoparticle surface. For the first layer, PAH with multiple positive charges was adsorbed onto the surface of negatively charged GNPs based on electrostatic interactions. Afterwards, addition of the positively charged PAH/GNPs to the negatively charged PAA solutions caused the formation of a second stabilized layer on the surface of GNPs with negative charge. After each modification step, two washing steps were performed to remove unbound polyelectrolyte as well as coexisting particles formed by the nanoaggregation of PAH with PAA [51]. Immobilization of papain on the LBL-modified GNPs was performed by using EDC and sulfo-NHS as activation reagents for amide coupling between carboxylic groups on the surface of the LBL-modified GNPs and amino groups of papain. Immobilization of papain on the first cationic PAH polyelectrolyte layer failed, because the single-layered and positively charged PAH/GNPs were easily aggregated in phosphate buffer (20 mM, pH 6.8) which was commonly used for the immobilization and digestion with papain.

3.2. Characterization of functionalized GNPs

In order to monitor the quality of the modified GNPs, various physicochemical characterizations were performed after each surface modification step with analytical techniques such as visible
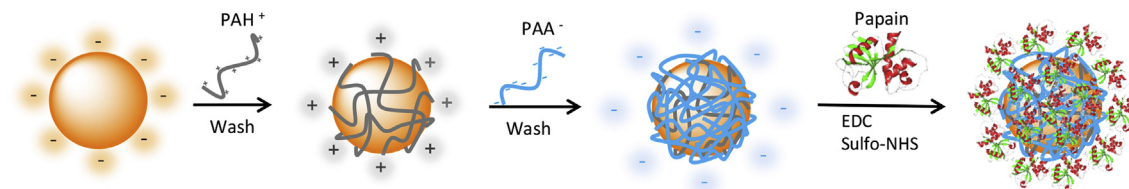
Fig. 1. Schematic presentation of the GNP functionalization process with papain. The surface modification of the GNPs was performed by a layer-by-layer (LBL) polyelectrolyte deposition strategy, and finally carboxylic acid moieties stemming from PAA were functionalized with papain by activation of carboxylic acid groups of the LBL-modified GNPs with EDC and sulfo-NHS as coupling agents and subsequent amide coupling with amino groups of papain. (PAH$^+$, polyallylamine hydrochloride; PAA$^-$, polyacrylic acid sodium).

spectroscopy measuring the absorbance maximum (λ_{max}) of the SPR band, DLS for measuring particle size distributions (i.e. hydrodynamic diameters) and electrophoretic light scattering for ζ-potential determinations [52].

The characterization results of the various synthesized nanoparticles with regard to their SPR band are shown in Fig. 2. It becomes evident that λ_{max} of the SPR band of citrate-stabilized GNPs (527 nm) and the LBL-modified GNP/PAH/PAAs are essentially identical while it is obviously shifted for the papain-functionalized GNPs (papain-GNPs) towards longer wavelength (543 nm). Furthermore, the broader peak width of papain-GNPs in comparison to GNP/PAH/PAAs and citrate-GNPs reveals that papain has

Fig. 2. Characterization of nanoparticles via surface plasmon resonance band. (a) Vis spectra of GNPs after each step of synthesis and surface functionalization. (b) Vis spectra of GNPs after functionalization with four different concentrations of papain in the reaction mixture (0.2, 0.4, 1 and 2 mg mL^{-1}, respectively).

been successfully immobilized. Due to absence of a strong red shifted absorption band at $\lambda > 600$ nm, it can be concluded that no significant aggregation of papain-GNPs occurs. During the functionalization process, four different concentrations of papain (0.2, 0.4, 1, and 2 mg mL^{-1}, respectively) in the reaction mixture were used to prepare four batches of papain-GNPs with different sizes and protein-coating thicknesses. Even though a longer wavelength and broader peak shape has been found after immobilization of papain from 0.2 mg mL^{-1} papain reaction solution, it can be interestingly seen from Fig. 2b that negligible SPR band shifts result among the four batches of papain-GNP solutions prepared from differently concentrated papain solutions. Therefore, these results imply that SPR is not sensitive enough for detecting these further size and size distribution changes. Therefore, another characterization method, DLS, has been used to reveal the size differences between the four batches of papain-GNPs.

Hydrodynamic diameters (d_h) and ζ-potentials of the different stages of modification are shown separately in Fig. 3a and b. The citrate-capped GNPs were first prepared with mean diameter of 36.4 ± 0.3 nm and a ζ-potential value of -35.7 ± 0.3 mV. After first layer modification, the size greatly increased to 127.6 ± 5.7 nm with the ζ-potential changing from the negative value of -35.7 ± 0.3 mV to the positive value of 69.6 ± 3.6 mV. The results confirmed successful modification with PAH. After second layer modification with PAA, the ζ-potential of the nanoparticles changed to a negative value of -44.4 ± 1.0 mV with an average size of 126.3 ± 6.3 nm according to DLS measurements. It is striking that the size does not change significantly although effective deposition of the anionic polyelectrolyte PAA is evident from the ζ-potential. The internal repulsive electrostatic interactions between PAH polymer chains which lead to an increase of the shell thickness after first layer deposition appear to be effectively attenuated when the oppositely charged PAA polyelectrolyte is applied on the surface. This might lead to a condensation of the shell thickness so that there is actually not much difference in d_h compared to the stage before. After functionalization with four different concentrations of papain (0.2, 0.4, 1, and 2 mg mL^{-1}, respectively), the hydrodynamic radius d_h grew significantly from 176.2 ± 5.2 nm (for 0.2 mg mL^{-1}) to 1315 ± 136 nm (for 2 mg mL^{-1}). Since severe aggregation was not found in the visible spectra (Fig. 2b), the increase in the hydrodynamic diameters with more papain in the reaction mixture was (largely) attributed to the increased amount of papain on the GNPs.

The ζ-potentials of all these four batches of papain-GNPs still remained highly negatively charged, which indicated that a good colloidal stability persisted. Moreover, it is worth noting that absolute values of the ζ-potential decreased with increasing amounts of immobilized papain. Papain ($pI \sim 8.75$) which carries a net positive charge in the employed phosphate buffer (pH 6.8) partly compensates negative charges of the 2-step LBL-modified GNPs which might compromise their colloidal stability.

In conclusion, DLS measurements confirm the successful immobilization of papain and that the amount of papain covalently
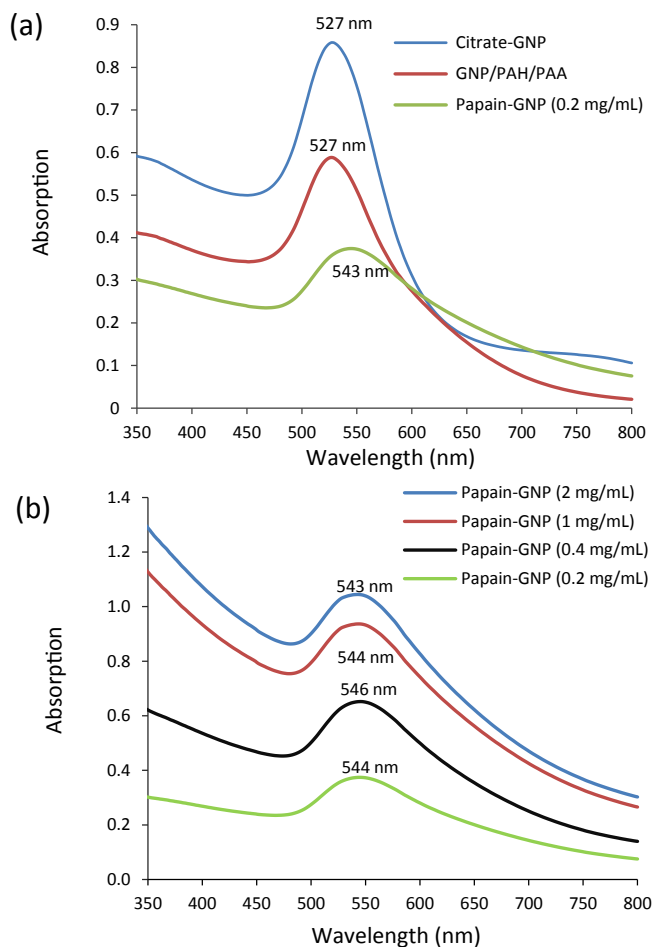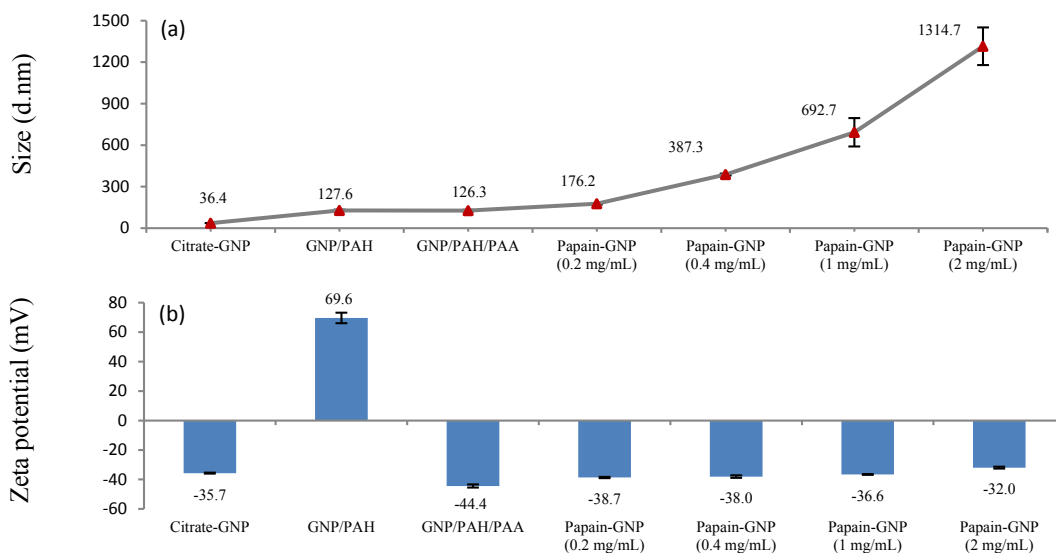
Fig. 3. Characterization of nanoparticles after each step of modification by (a) DLS and (b) ζ-potential measurement (n = 3) (note, the indicated concentrations refer to the final papain concentrations in the reaction mixture).

bonded on the GNPs can be readily increased by higher concentrations of papain used during protein coupling reaction. ELS measurements, on the other hand, reveal that ζ-potentials of these nanobioconjugates are large enough to support their colloidal stability.

3.3. Loading amount of immobilized papain

RMM was recently introduced as an innovative technology for accurately measuring the mass of nano- and microparticles with detection limits in the range of femtograms to attograms [46,53—55]. It allows convenient particle classification in the size range of 50 nm to 5 μm if a sufficient number of particles is counted. Therefore, RMM was envisaged as a useful tool in our study for the determination of the average number of papain molecules immobilized on GNPs by using an Archimedes instrument. This instrument accommodates a sensor chip with a microfluidic channel (having a dimension of 8×8 μm^2) in which a resonant structure is embedded inside. When a nanoparticle passes through this microchannel, a shift of the resonant frequency occurs because the resonant structure senses the density difference between the particle ($\rho_{particle}$) and the transport fluid (ρ_{fluid}). Information on the buoyant mass of the nanoparticles can be derived from the resonant frequency shift [46].

The number of particles measured was set to 2000. The average buoyant mass (M_B [fg]) of this nanoparticle population was calculated by conversion from the resonant frequency shift (Δf [s^{-1}]) according to Eq. (1).

$$M_B = \Delta f \times S \tag{1}$$

wherein S [fg s] represents the microchannel resonator's sensitivity. It is a fixed value for each resonator, which reflects a simple linear relationship over the entire range of measurable particles, and has to be determined by a simple calibration procedure. From the buoyant mass the dry mass M [fg] can be calculated for a defined chemically homogeneous particle according to Eq. (2) [46,56].

$$M = \frac{M_B}{\left(1 - \rho_{fluid} / \rho_{particle}\right)} \tag{2}$$

However, in the present case we are dealing with a composite material i.e. a shell particle in which gold core and protein shell have distinct densities which needs to be considered in the calculations. Hence, a modified version of Eq. (2) has been utilized. By comparing the average buoyant mass changes of GNPs before and after functionalization (ΔM_B), the average mass of immobilized papain on a single GNP ($M_{(papain\ shell)}$) can be approximated by Eq. 3

$$M_{(papain\ shell)} = \frac{\Delta M_B}{\left(1 - \rho_{fluid} / \rho_{shell}\right)} \tag{3}$$

wherein ρ_{shell} is the density of the papain shell. The density of the fluid (water) used was 1 g mL^{-1} in all calculations, and the density of the papain shell on the GNP surface was calculated to be 1.43 g cm^{-3} from the exponential function reported by H. Fischer [57]. Hence, the average mass of immobilized papain was obtained. More importantly, with the calculated average mass of a papain molecule from the known molecular weight of papain (23,406 Da), the average number of papain molecules immobilized on each nanoparticle was finally obtained [46]. The particle buoyant mass distributions of LBL-modified GNPs (GNP/PAH/PAA) and papain-GNPs (0.2 mg mL^{-1} and 1 mg mL^{-1}) are shown in Fig. 4. The graph in Fig. 4a shows a relatively narrow buoyant mass distribution of GNP/PAH/PAA nanoparticles. Surface functionalization with 0.2 mg mL^{-1} papain yielded a slight, but insignificant shift of the average buoyant mass from 3.14 ± 0.06 to 3.42 ± 0.31 fg (Fig. 4b and Table 2). On the other hand, for GNPs functionalized with 1 mg mL^{-1} papain solution, a significant shift of the distribution towards a larger average buoyant mass (27.30 ± 2.25 fg) with a broader width indicative for increased polydispersity can be observed (Fig. 4c and Table 2).

Therefore, with this technology the average mass of papain per GNP, the concentration of nanoparticles, the concentration of immobilized papain and the number of papain molecules per GNP can be calculated (shown in Table 2). Such knowledge is of importance for the calculation of the turnover numbers (k_{cat}) of the immobilized enzyme nanobioreactors. It can be seen that the average mass increase due to papain immobilization (with 0.2 mg mL^{-1}) on GNP/PAH/PAA corresponds to the mass of $2.39 \pm 2.14 \times 10^4$ papain molecules per single GNP. By using the
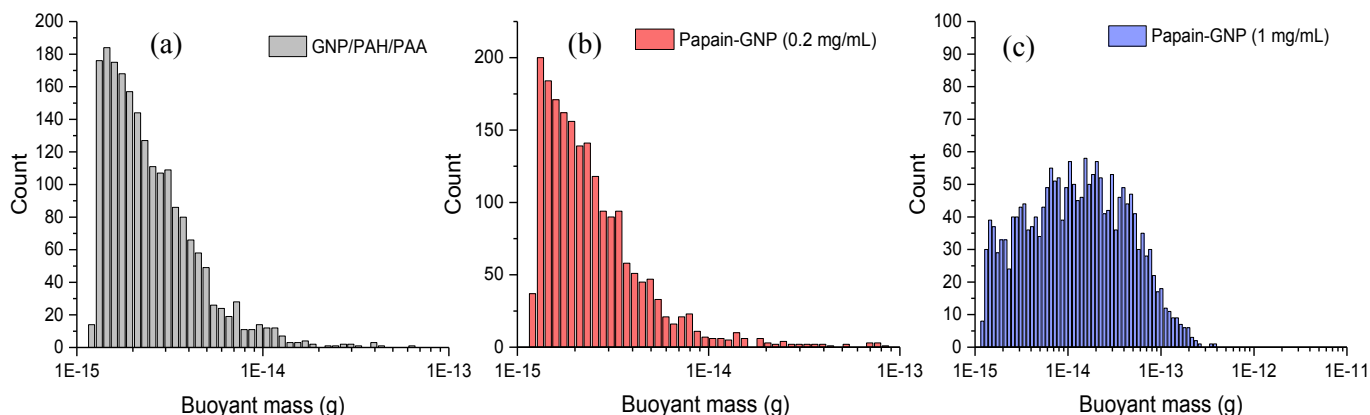
Fig. 4. Distribution of the buoyant mass of particles in 20 mM phosphate buffer (pH 6.8) of LBL-modified GNPs (GNP/PAH/PAA) (a), GNPs functionalized with papain (0.2 mg mL^{-1}) (b) and (1 mg mL^{-1}) (c) (note, the indicated concentrations refer to the final papain concentrations in the reaction mixture).

Table 2
Summary of the experimental results from RMM measurements comprising mean buoyant mass, concentration of nanobeads, mean mass of papain per GNP, concentration of immobilized papain and mean number of papain molecules per GNP (n = 3).

Nanobeads	Mean buoyant mass [fg]	Concentration of nanobeads [mol L^{-1}]	Mean mass of papain per GNP [fg]	Concentration of immobilized papain [mol L^{-1}]	Mean number of papain per GNP
GNP/PAH/PAA	3.14 ± 0.06	1.06 × 10^{-11}	–	–	–
Papain-GNPs [0.2 mg mL^{-1}]	3.42 ± 0.31	7.74 × 10^{-12}	0.93 ± 0.83	1.85 ± 1.65 × 10^{-7}	2.39 ± 2.14 × 10^4
Papain-GNPs [1 mg mL^{-1}]	27.30 ± 2.25	1.43 × 10^{-12}	80.2 ± 7.3	2.95 ± 0.27 × 10^{-6}	2.06 ± 0.19 × 10^6

same calculation for the papain-GNPs (from 1 mg mL^{-1}), the average mass increase corresponds to 2.06 ± 0.19 × 10^6 molecules of papain immobilized per GNP. The factor 100 increase in surface coverage upon a factor 5 increase of the papain concentration in the reaction mixture is unexpected. However, it is confirmed by the more than proportional size increase as measured by DLS and may have to do with the specific conformational arrangement and morphology, respectively, of the polyelectrolyte coating. Calculations based on the approximately available surface area of GNP/PAH/PAA with a diameter of 126 nm for papain adsorption and the hydrodynamic diameter of papain as estimated by DLS leads to the conclusion that the papain shell is formed by a multilayer bonding on the GNPs. One could imagine bonding of multiple papain molecules on loosely adsorbed loops of the outer PAA polymer chains yielding elevated protein surface coverage.

Furthermore, it can be clearly seen from Table 2 that the nanoparticle concentrations in the suspension continuously decline. The GNP/PAH/PAA solution showed the highest concentration of 1.06 × 10^{-11} mol L^{-1}, but the concentration decreased after functionalization with papain, which is due to the loss of nanoparticles during reaction. In conclusion, the RMM technology is an effective tool for measuring the average mass of immobilized protein on each nanoparticle.

3.4. Determination of kinetic parameters of immobilized and free papain

The enzyme activity of free papain and of the four distinct papain-GNP bioconjugates was evaluated at 37 °C and pH 6.8 by investigating the hydrolysis of BApNA as the substrate [48]. In principle, BApNA is cleaved in the presence of papain to produce chromogenic *p*-nitroaniline (pNA) which exhibits strong absorbance at the wavelength of 410 nm (Suppl. Fig. S1). Thus, the activity of papain can be conveniently measured by the released pNA

with a photometric assay. To obtain the Michaelis-Menten parameters (K_m, V_{max}, k_{cat}; see also Suppl. Material), the activities of free and immobilized papain for various concentrations of BApNA were measured and the obtained data were plotted in the form of Lineweaver-Burk plots, as shown in Fig. 5a. These plots document a good linear relationship between 1/[S] and 1/V$_0$ for both free and immobilized papain. K_m and V_{max} values of free and immobilized papain were calculated from the intercepts on x- and y-axes, respectively, and the results are summarized in Table 3.

The Michaelis constant K_m is the substrate concentration at which half of the maximal reaction rate (V_{max}/2) can be achieved. In free solution, it is independent of the enzyme concentration. Upon immobilization, the enzyme conformation might change, thus negatively (or even positively) influencing enzyme-substrate complex formation. Furthermore, the access of the substrate to the active site might be altered upon immobilization. Both of these properties might change in dependence of the surface coverage. For instance, at low surface concentrations the active site might be better accessible than at high coverage for which sterically hindered and limited access to the enzyme's catalytic site might compromise the rate constant for association and thus affect or alter K_m. The data in Table 3, however, reveal that Michaelis constants K_m remain in the same order of magnitude when different concentrations of papain are immobilized on GNPs, ranging between 1.4 and 2.6 mM (mean = 2.0 ± 0.6 mM; the slight fluctuations represent experimental uncertainties which are in a similar order as for the repetitive experiments with free papain) (see also Suppl. Fig. S2). A comparison of this mean with K_m of free solution reaction reveals that there is no statistically significant difference in K_m. Hence, binding affinity is not compromised upon immobilization of papain [58].

V_{max} of enzyme reactions reflects how fast the biocatalyst can catalyze the biotransformation. V_{max} rises linearly with the enzyme concentration in free solution reactions. This is also observed for
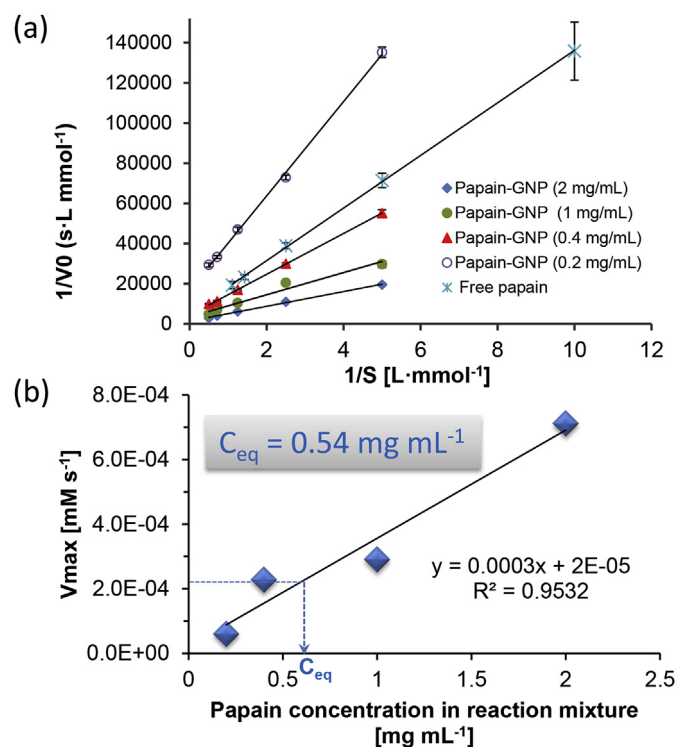
Fig. 5. a) Lineweaver-Burk diagrams for BApNA hydrolysis at pH 6.8 catalyzed by papain-GNP bioconjugates obtained from reaction mixtures with different papain concentrations, and for comparison by free (1 mg mL^{-1}) papain, and (b) effect of different surface coverages (as obtained from different papain concentrations in the reaction buffer) on maximal reaction rates V_{max} (indicated values are mean values of three experiments, n = 3, and error bars represent the standard deviations; note, the indicated concentrations refer to the final papain concentrations in the reaction mixture; C_{eq} represents the equivalent papain concentration needed in the reaction buffer for immobilization to obtain GNP-conjugated papain with the same Vmax as a 1 mg mL^{-1} free papain solution).

the papain-GNP conjugates varying in papain surface concentrations (Fig. 5b). V_{max} values of immobilized papain range from $6.00 \pm 0.27 \times 10^{-5}$ mM s^{-1} (0.2 mg mL^{-1} papain in reaction mixture) to $7.12 \pm 0.17 \times 10^{-4}$ mM s^{-1} (2.0 mg mL^{-1}) (Table 3). The corresponding reaction with free papain reached a V_{max} of $2.03 \pm 0.57 \times 10^{-4}$ mM s^{-1}, respectively. A faster reaction with immobilized enzyme is clearly evident. This, in turn, has also the advantage that less enzyme could be used to reach the same maximum reaction rate if it is immobilized to the GNP carrier. This is highly relevant in biocatalytic industrial applications with precious enzymes and those of limited availability. For instance, the

same V_{max} as with 1 mg mL^{-1} papain in free solution can be achieved with papain-GNP conjugated biocatalyst prepared from only 0.54 mg mL^{-1} in the reaction mixture (equivalent concentration C_{eq}) (Fig. 2b). Thus, a factor of almost 2 less enzyme would be consumed if the supernatant after reaction is discarded and disposed. Actually, since only a small fraction of the enzyme is bonded to the GNPs, the unbound enzyme can be easily recycled from the reaction mixture after centrifugation. For instance, in the reaction with papain-GNP conjugate prepared from 1 mg mL^{-1} the acutal papain concentration in the digestion is only 21.3 μg mL^{-1} (2.9 μg mL^{-1} in the digestion with bioconjugate prepared from 0.2 mg mL^{-1}). Thus, the savings of enzyme are even more pronounced (about factor 100), making this approach with nanoparticulate heterogeneous biocatalysts most attractive for precious enzymes.

A better figure to define and compare the reaction rate of a series of biocatalysts is the turnover number k_{cat}. It is the number of substrate molecules converted to product per enzyme molecule per second. By using the known concentrations of free papain and immobilized papain calculated from RMM, the k_{cat} values were finally calculated to be $4.75 \pm 1.33 \times 10^{-3}$ s^{-1} (free), 0.20 ± 0.02 s^{-1} (conjugated from 1 mg mL^{-1}) and 0.65 ± 0.03 s^{-1} (conjugated from 0.2 mg mL^{-1}), respectively. It becomes striking that the reaction is accelerated by a factor of 137 when free solution enzyme reaction and heterogeneous nanobiocatalysis with papain-GNP conjugate prepared from 1 mg mL^{-1} are compared. Local accumulation of substrate on the surface of the nanobiocatalyst or close to it, thus forming a concentration gradient to the adjacent solution (like in a double layer model) might be invoked as an explanation for this increased reaction rates. Furthermore, the ratio k_{cat}/K_m, a characteristic parameter, was calculated to describe the catalytic efficiency of the immobilized papain in comparison to free papain. It was found that immobilized papain had a much higher ratio k_{cat}/K_m compared to the free papain, which means that the immobilized papain exhibited significantly higher catalytic efficiency. Therefore, from above results it can be concluded that the immobilized enzyme with many additional advantages provides the higher catalytic efficiency compared to the free papain. Besides acceleration of reaction rates, immobilization on nanoparticulate carriers might be a viable strategy to safe precious enzymes in industrial applications.

3.5. SDS-PAGE for monitoring antibody digestion

The primarily intended application of the papain-modified nanoparticles is fragmentation of IgG into smaller fragments which facilitate their MS analysis (middle-down and middle-up) [41]. In the present case, a therapeutic protein, human IgG, was utilized to verify the bioactivity and digestion performance of

Table 3
Summary of the kinetic parameters (K_m, V_{max}, k_{cat}) calculated by Lineweaver-Burk plots for both free papain and immobilized papain (0.2 mg mL^{-1} and 1 mg mL^{-1}, respectively) (Values represent mean ± standard deviation of 3 replicate experiments; n.d., not determined).

Parameter	Papain-GNP (0.2 mg mL^{-1})[a]	Papain-GNP (0.4 mg mL^{-1})[a]	Papain-GNP (1 mg mL^{-1})[a]	Papain-GNP (2 mg mL^{-1})[a]	Free papain (1 mg mL^{-1})[b]
K_m [mM]	1.41 ± 0.05	2.32 ± 0.04	1.61 ± 0.03	2.62 ± 0.15	2.71 ± 0.98
V_{max} [mM s^{-1}]	$6.00 \pm 0.27 \times 10^{-5}$	$2.28 \pm 0.06 \times 10^{-4}$	$2.91 \pm 0.03 \times 10^{-4}$	$7.12 \pm 0.17 \times 10^{-4}$	$2.03 \pm 0.57 \times 10^{-4}$
k_{cat} [s^{-1}]	0.65 ± 0.03	n.d.	0.20 ± 0.02	n.d.	$4.75 \pm 1.33 \times 10^{-3}$
k_{cat}/K_m [mM^{-1} s^{-1}]	0.46 ± 0.01	n.d.	0.12 ± 0.03	n.d.	$1.8 \pm 0.2 \times 10^{-3}$
Regression equation	Y = 23499X + 16,693	Y = 10155X + 4382.1	Y = 5530.9X + 3432.9	Y = 3671.5X + 1404.7	Y = 13096X + 5248.1
R^2	0.9988	0.9997	0.9667	0.9980	0.9998

[a] The concentrations refer to the papain concentration used for the immobilization, the final concentration in the digestion solution are 9.25×10^{-8} mol L^{-1} and 1.48×10^{-6} mol L^{-1} for papain-GNP (0.2 mg mL^{-1} and 1 mg mL^{-1}), respectively (see Table 2).
[b] The concentration refers to the final papain concentration in the digestion solution.

papain-functionalized GNPs (prepared from 0.2 mg mL^{-1} papain in the reaction mixture) in comparison to in-solution digestion (1 mg mL^{-1} papain). By taking advantage of the heterogeneous biocatalyst, the papain-GNPs can be readily removed after digestion, avoiding the contamination of enzyme into the fragment products. IgG digests obtained from GNP-conjugated and free papain were characterized by SDS-PAGE under non-reducing conditions using silver staining for detection (providing detection limits around 0.25–0.5 ng of protein). The resultant gel is shown in Fig. 6. It is evident that new bands (at around 50 kDa) appeared in the samples (lanes B,C,D and F,G,H) after digestion. The molecular weight of these protein bands corresponds to the expected IgG fragments. Papain cleaves the intact IgG above the Hinge region to yield three fragments of similar size, two Fab fragments (around 50 kDa each) and one Fc fragment (around 50 kDa as well). In addition, bands around 100 kDa and 25 kDa were also found in the digested samples. It seems that the employed papain does not have enough specificity for the cleavage of IgG just above the Hinge region, but also the F (ab')$_2$ and Fc/2 fragments are obtained. In sharp contrast, in the non-digested IgG solution (lane A), only intact IgG (around 150 kDa) and aggregated forms (ca. 300 kDa) rather than the fragments were found (the large quantities of aggregates can be explained by the fact that the h-IgG used in this study was already expired for a long time). These results clearly indicate that IgG was efficiently digested with papain-functionalized GNPs.

To test for possible bleeding of papain from the nanoparticulate carrier, a suspension of papain-functionalized GNPs was allowed to stand in 10 mM phosphate buffer pH 6.8 overnight. Afterwards the suspension was spinned down and the supernatant was also applied as sample to the gel (lane E). It can be clearly seen that there is no band (E) at the position where papain is expected (ca. 23 kDa region). This implies that, after their removal by centrifugation, papain-GNP conjugates do not cause any background interferences in the gel and that papain is not bleeding from the nanoparticulate carrier due to covalent attachment. From these results it can be derived that the new nanobioreactor is suitable for practical applications.

3.6. HPLC-µESI-QTOF-MS method for characterization of IgG fragments after digestion

The generated fragments of human IgG were finally analyzed by HPLC-µESI-QTOF-MS to document the practical applicability of the

nanoparticulate enzyme reactor. High resolution-MS using a QTOF equipped with a µESI sprayer and a ProSwift column for separation of proteins were used for the characterization of intact IgG and the fragments digested with papain-GNPs. Detection limits for smaller proteins like IgG fragments are around 500 fg on-column with this µESI setup. Fig. 7a shows the total ion chromatogram of the intact IgG before digestion, and Fig. 7b shows the deconvoluted mass spectrum of the peak with t$_R$ 14.66 min corresponding to IgG with a mass of around 150 kDa (note, this is not a monoclonal antibody but human IgG isolated from plasma). After digestion, the deconvoluted mass of IgG (150 kDa) disappeared, and meanwhile the fragments (Fab and Fc), eluting at ca. 9.6 min could be detected (Fig. 7c). The deconvoluted TOF-MS spectrum in Fig. 7d for the peak at 9.6 min shows several peak groups in the range of 50–55 kDa corresponding to Fab and Fc, respectively. It can be explained by the fact that human IgG is a collection of immunoglobulin molecules, and therefore a mixture of similar fragments with slightly different masses can be expected for the digested IgG as well.

The TIC chromatogram in Fig. 7c, like the gel, shows some additional peaks. The first two peaks (between t$_R$ 7.7 and 8.5 min) could not be assigned to a reasonable mass. However, the deconvoluted peak (t$_R$ between 8.8 and 9.5 min) showed fragment masses of around 25 kDa, which correspond to the Fc/2 fragments (Fig. 7e). In addition, a peak at t$_R$ of 14.63 min was detected in the digest and it was initially assumed that this peak corresponds to intact human IgG due to incomplete digestion. However, the deconvoluted TOF-MS spectrum of the peak at 14.63 min actually corresponds to a 100 kDa F(ab')$_2$ fragment of human IgG (Fig. 7f) while the intact IgG (150 kDa) was not detected at all. This finding clearly confirms the successful and complete digestion. It seems that the employed papain does not have enough specificity for cleaving the IgG just above the Hinge region, but produces also a F(ab')$_2$ fragment in significant amounts besides Fab. Overall, the LC-MS results clearly document that the new nanoparticulate enzyme reactor has adequate bioactivity and catalytic efficiency for enzymatic IgG fragmentation. The limited specificity of papain both in free and immobilized forms, however, explains why nowadays IdeS (immunoglobulin G-degrading enzyme of *Streptococcus pyogenes,* a bacterial cysteine protease which specifically cleaves IgGs under their hinge region [41]) is mainly used for middle-down characterization of IgGs by LC-MS. It is obvious that this highly specific enzyme should be immobilized in the same way to afford a highly specific, highly efficient immobilized enzyme reactor which can be easily removed and reused for IgG fragmentation.

4. Conclusion

In the present study, GNPs were functionalized with papain by a layer-by-layer strategy producing an efficient heterogeneous biocatalyst. Characterization with SPR, DLS and ζ-potential measurements revealed the successful immobilization of papain. The higher amount of papain used in the reaction mixture for immobilization has led to bigger particle sizes than expected. With the RMM technology, the concentration and average buoyant mass of the GNP biocatalyst was obtained, and the number of papain molecules immobilized on one GNP was calculated. The results implied that GNPs activated with polyelectrolyte layers provided a high loading capacity for papain due to its large surface to volume ratio and the specific surface modification with polyelectrolyte chains. In comparison with free papain, the immobilized papain provides the higher catalytic efficiency with the advantages of easy removal and flexible control of reaction. SDS-PAGE and HPLC-µESI-QTOF-MS characterization proved the successful digestion of IgG with papain-modified GNPs as heterogeneous biocatalyst which indeed shows great potential in bioanalysis. The limited specificity of
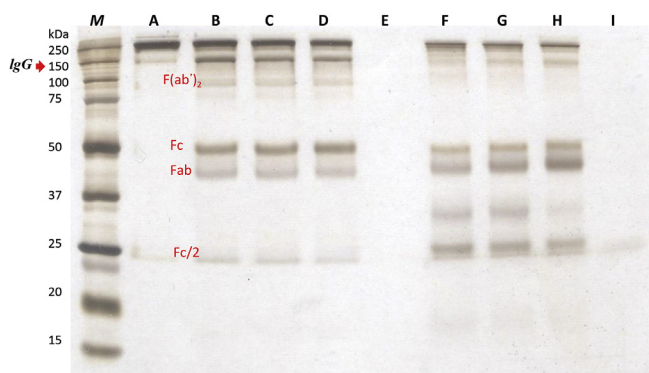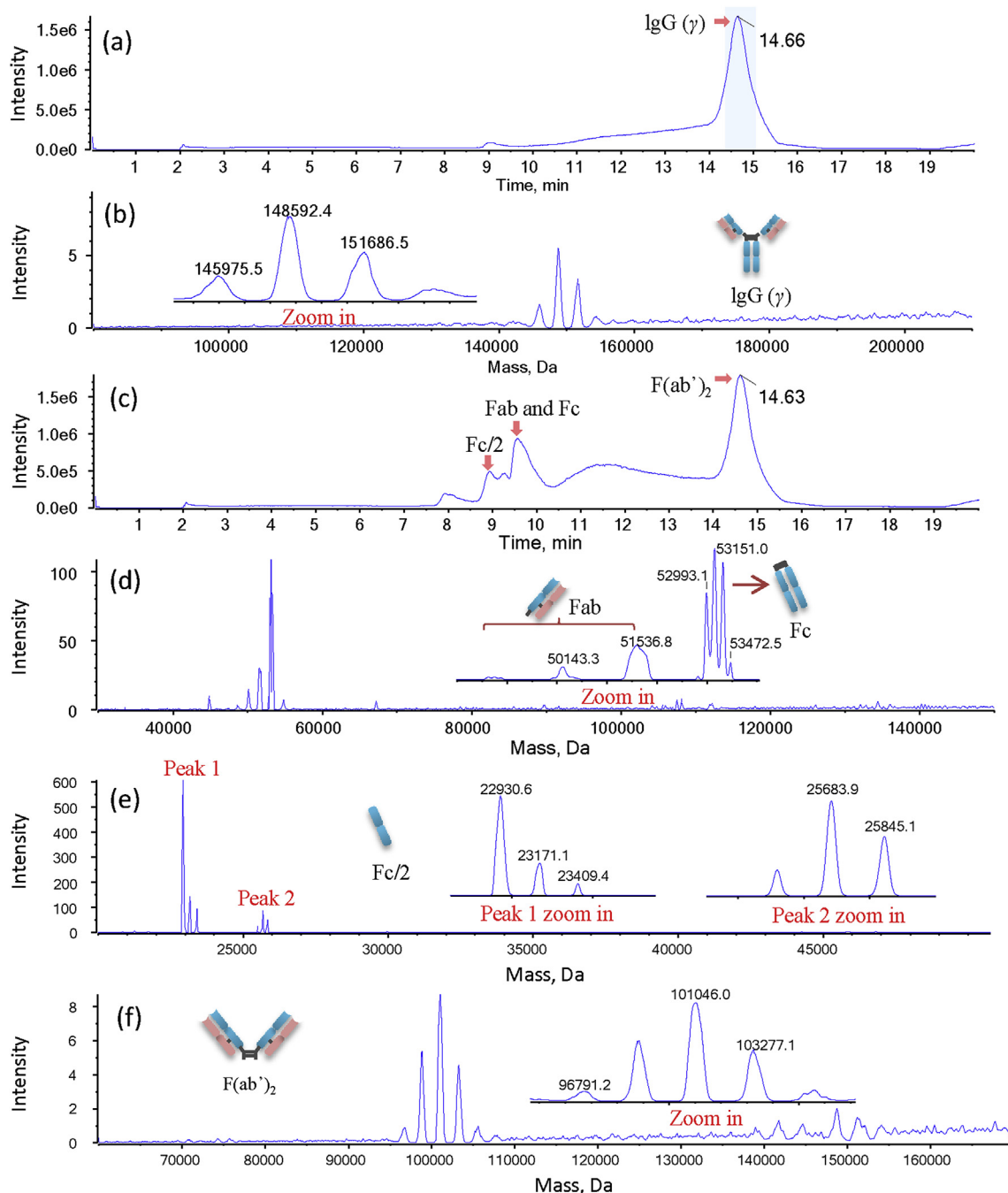

Fig. 6. SDS-PAGE (gel: 10%-Tris-HCl) for non-reduced human IgG digestion samples with silver staining for protein visualization. (M) protein marker; (A) IgG in phosphate buffer (10 mM, pH 6.8); (B–D) IgG digested with papain-GNPs for 4 h (B); for 8 h (C); for 24 h (D); (E) suspension of papain-GNPs after centrifugation; (F–H) IgG digested with free papain for 4 h (F); for 8 h (G); for 24 h (H); (I) free papain in phosphate buffer (10 mM, pH 6.8) (note, usually 0.25–0.5 ng of protein are considered as limit of detection for silver staining).

Fig. 7. HPLC-μESI-TOF-MS analysis of human IgG digested with papain-GNP bioconjugate. (a) TIC chromatogram of the IgG before digestion; (b) the deconvoluted TOF-MS spectrum of the peak with t_R 14.66 showing the intact IgG (ca. 150 kDa); (c) TIC chromatogram of human IgG after digestion with papain-GNPs; (d) the deconvoluted TOF-MS spectrum of the peak with t_R 9.62 showing the IgG fragments (Fab and Fc, ca. 50 kDa); (e) the deconvoluted TOF-MS spectrum of Fc/2 fragments (ca. 25 kDa) of the peak at t_R between 8.8 and 9.5 min; (f) the deconvoluted TOF-MS spectrum of the 100 kDa F(ab')$_2$ IgG fragment of the peak at t_R 14.63 min.

papain could be overcome by use of IdeS as an enzyme for immobilization.

resonant mass measurements.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.aca.2017.02.009.

References

[1] P.K. Robinson, Enzymes: principles and biotechnological applications, Essays Biochem. 59 (2015) 1—41.
[2] D. Meridor, A. Gedanken, Enhanced activity of immobilized pepsin

nanoparticles coated on solid substrates compared to free pepsin, Enzyme Microb. Technol. 67 (2014) 67—76.

[3] K. Meller, P. Pomastowski, D. Grzywiński, M. Szumski, B. Buszewski, Preparation and evaluation of dual-enzyme microreactor with co-immobilized trypsin and chymotrypsin, J. Chromatogr. A 1440 (2016) 45—54.

[4] S.A. Ansari, Q. Husain, Potential applications of enzymes immobilized on/in nano materials: a review, Biotechnol. Adv. 30 (2012) 512—523.

[5] Y. Lv, Z. Lin, T. Tan, F. Svec, Preparation of reusable bioreactors using reversible immobilization of enzyme on monolithic porous polymer support with attached gold nanoparticles, Biotechnol. Bioeng. 111 (2014) 50—58.

[6] Y. Liang, C. Wu, Q. Zhao, Q. Wu, B. Jiang, Y. Weng, Z. Liang, L. Zhang, Y. Zhang, Gold nanoparticles immobilized hydrophilic monoliths with variable functional modification for highly selective enrichment and on-line deglycosylation of glycopeptides, Anal. Chim. Acta 900 (2015) 83—89.

[7] J.R. Freije, P.P.M.F.A. Mulder, W. Werkman, L. Rieux, H.A.G. Niederlander, E. Verpoorte, R. Bischoff, Chemically modified, immobilized trypsin reactor with improved digestion efficiency, J. Proteome Res. 4 (2005) 1805—1813.

[8] Y. Cao, L. Wen, F. Svec, T. Tan, Y. Lv, Magnetic AuNP@Fe3O4 nanoparticles as reusable carriers for reversible enzyme immobilization, Chem. Eng. J. 286 (2016) 272—281.

[9] E. Calleri, S. Ambrosini, C. Temporini, G. Massolini, New monolithic chromatographic supports for macromolecules immobilization: challenges and opportunities, J. Pharm. Biomed. Anal. 69 (2012) 64—76.

[10] M. Bonichon, A. Combès, C. Desoubries, A. Bossée, V. Pichon, Development of immobilized-pepsin microreactors coupled to nano liquid chromatography and tandem mass spectrometry for the quantitative analysis of human butyrylcholinesterase, J. Chromatogr. A 1461 (2016) 84—91.

[11] N. Miletić, A. Nastasović, K. Loos, Immobilization of biocatalysts for enzymatic polymerizations: possibilities, advantages, applications, Bioresour. Technol. 115 (2012) 126—135.

[12] B.J. Johnson, W. Russ Algar, A.P. Malanoski, M.G. Ancona, I.L. Medintz, Understanding enzymatic acceleration at nanoparticle interfaces: approaches and challenges, Nano Today 9 (2014) 102—131.

[13] C. Garcia-Galan, Á. Berenguer-Murcia, R. Fernandez-Lafuente, R.C. Rodrigues, Potential of different enzyme immobilization strategies to improve enzyme performance, Adv. Synth. Catal. 353 (2011) 2885—2904.

[14] R. DiCosimo, J. McAuliffe, A.J. Poulose, G. Bohlmann, Industrial use of immobilized enzymes, Chem. Soc. Rev. 42 (2013) 6437—6474.

[15] Y. Yu, Z. Chen, S. He, B. Zhang, X. Li, M. Yao, Direct electron transfer of glucose oxidase and biosensing for glucose based on PDDA-capped gold nanoparticle modified graphene/multi-walled carbon nanotubes electrode, Biosens. Bioelectron. 52 (2014) 147—152.

[16] K. Saha, S.S. Agasti, C. Kim, X. Li, V.M. Rotello, Gold nanoparticles in chemical and biological sensing, Chem. Rev. 112 (2012) 2739—2779.

[17] M. Höldrich, A. Sievers-Engler, M. Lämmerhofer, Gold nanoparticle-conjugated pepsin for efficient solution-like heterogeneous biocatalysis in analytical sample preparation protocols, Anal. Bioanal. Chem. 408 (2016) 5415—5427.

[18] W. She, N. Li, K. Luo, C. Guo, G. Wang, Y. Geng, Z. Gu, Dendronized heparin—doxorubicin conjugate based nanoparticle as pH-responsive drug delivery system for cancer therapy, Biomaterials 34 (2013) 2252—2264.

[19] E. Blanco, H. Shen, M. Ferrari, Principles of nanoparticle design for overcoming biological barriers to drug delivery, Nat. Biotechnol. 33 (2015) 941—951.

[20] L. Sun, H. Liang, Q. Yuan, T. Wang, H. Zhang, Study on a carboxyl-activated carrier and its properties for papain immobilization, J. Chem. Technol. Biotechnol. 87 (2012) 1083—1088.

[21] G.K. Kouassi, J. Irudayaraj, G. McCarty, Activity of glucose oxidase functionalized onto magnetic nanoparticles, Biomagn. Res. Technol. 3 (2005) 1—10.

[22] L.S. Wong, J. Thirlway, J. Micklefield, Direct site-selective covalent protein immobilization catalyzed by a phosphopantetheinyl transferase, J. Am. Chem. Soc. 130 (2008) 12456—12464.

[23] T. Tan, J. Lu, K. Nie, L. Deng, F. Wang, Biodiesel production with immobilized lipase: a review, Biotechnol. Adv. 28 (2010) 628—634.

[24] Y. Yücel, Biodiesel production from pomace oil by using lipase immobilized onto olive pomace, Bioresour. Technol. 102 (2011) 3977—3980.

[25] G.A. Petkova, K. Záruba, P. Žvátora, V. Král, Gold and silver nanoparticles for biomolecule immobilization and enzymatic catalysis, Nanoscale Res. Lett. 7 (2012) 287.

[26] S.F. Oliveira, G. Bisker, N.A. Bakh, S.L. Gibbs, M.P. Landry, M.S. Strano, Protein functionalized carbon nanomaterials for biomedical applications, Carbon 95 (2015) 767—779.

[27] N.R. Mohamad, N.H.C. Marzuki, N.A. Buang, F. Huyop, R.A. Wahab, An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes, Biotechnol. Biotec. Eq. 29 (2015) 205—220.

[28] A. Homaei, R. Etemadipour, Improving the activity and stability of actinidin by immobilization on gold nanorods, Int. J. Biol. Macromol. 72 (2015) 1176—1181.

[29] L. Dykman, N. Khlebtsov, Gold nanoparticles in biomedical applications: recent advances and perspectives, Chem. Soc. Rev. 41 (2012) 2256—2282.

[30] N.G. Khlebtsov, Optics and biophotonics of nanoparticles with a plasmon resonance, Quantum Electron+ 38 (2008) 504.

[31] K. Saha, S.S. Agasti, C. Kim, X. Li, V.M. Rotello, Gold nanoparticles in chemical

[32] and biological sensing, Chem. Rev. 112 (2012) 2739—2779.

[32] M.A. Martins, S. Fateixa, A.V. Girão, S.S. Pereira, T. Trindade, Shaping gold nanocomposites with tunable optical properties, Langmuir 26 (2010) 11407—11412.

[33] M.-C. Daniel, D. Astruc, Gold Nanoparticles: assembly, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology, Chem. Rev. 104 (2004) 293—346.

[34] K.E. Sapsford, W.R. Algar, L. Berti, K.B. Gemmill, B.J. Casey, E. Oh, M.H. Stewart, I.L. Medintz, Functionalizing nanoparticles with biological molecules: developing chemistries that facilitate nanotechnology, Chem. Rev. 113 (2013) 1904—2074.

[35] B. Sahoo, S.K. Sahu, D. Bhattacharya, D. Dhara, P. Pramanik, A novel approach for efficient immobilization and stabilization of papain on magnetic gold nanocomposites, Colloid. Surf. B 101 (2013) 280—289.

[36] K. Kinoshita, K. Sato, M. Hori, H. Ozaki, H. Karaki, Decrease in activity of smooth muscle L-type Ca2+ channels and its reversal by NF-κB inhibitors in Crohn's colitis model, Am. J. Physiol. 285 (2003) G483—G493.

[37] K. Shiozaki, M. Yanagida, A functional 125-kDa core polypeptide of fission yeast DNA topoisomerase II, Mol. Cell. Biol. 11 (1991) 6093—6102.

[38] H. Xiang, Y.G. Xiang, M.Z. Lu, L. Guo, H. Eckstein, Total enzymatic synthesis of cholecystokinin CCK-5, Amino Acids 27 (2004) 101—105.

[39] T.A. Seldon, K.E. Hughes, D.J. Munster, D.Y. Chin, M.L. Jones, Improved protein-a separation of V(H)3 Fab from Fc after papain digestion of antibodies, J. Biomol. Tech. 22 (2011) 50—52.

[40] Y. Zhao, L. Gutshall, H. Jiang, A. Baker, E. Beil, G. Obmolova, J. Carton, S. Taudte, B. Amegadzie, Two routes for production and purification of Fab fragments in biopharmaceutical discovery research: papain digestion of mAb and transient expression in mammalian cells, Protein Expr. Purif. 67 (2009) 182—189.

[41] A. Beck, E. Wagner-Rousset, D. Ayoub, A. Van Dorsselaer, S. Sanglier-Cianférani, Characterization of therapeutic antibodies and related products, Anal. Chem. 85 (2013) 715—736.

[42] S. Mandal, A. Bonifacio, F. Zanuttin, V. Sergo, S. Krol, Synthesis and multidisciplinary characterization of polyelectrolyte multilayer-coated nanogold with improved stability toward aggregation, Colloid Polym. Sci. 289 (2011) 269—280.

[43] C. Boyer, A. Bousquet, J. Rondolo, M.R. Whittaker, M.H. Stenzel, T.P. Davis, Glycopolymer decoration of gold nanoparticles using a LbL approach, Macromolecules 43 (2010) 3775—3784.

[44] S. Pereira, A. Barros-Timmons, T. Trindade, Biofunctionalisation of colloidal gold nanoparticles via polyelectrolytes assemblies, Colloid Polym. Sci. 292 (2014) 33—50.

[45] J. Kim, J.W. Grate, P. Wang, Nanostructures for enzyme stabilization, Chem. Eng. Sci. 61 (2006) 1017—1026.

[46] M.R. Nejadnik, W. Jiskoot, Measurement of the average mass of proteins adsorbed to a nanoparticle by using a suspended microchannel resonator, J. Pharm. Sci. 104 (2015) 698—704.

[47] J. Sproß, A. Sinz, Immobilized monolithic enzyme reactors for application in proteomics and pharmaceutics, Anal. Bioanal. Chem. 395 (2009) 1583—1588.

[48] H. Hinterwirth, W. Lindner, M. Lämmerhofer, Bioconjugation of trypsin onto gold nanoparticles: effect of surface chemistry on bioactivity, Anal. Chim. Acta 733 (2012) 90—97.

[49] G. Schneider, G. Decher, From functional core/shell nanoparticles prepared via layer-by-layer deposition to empty nanospheres, Nano Lett. 4 (2004) 1833—1839.

[50] Y.-C. Yeh, B. Creran, V.M. Rotello, Gold nanoparticles: preparation, properties, and applications in bionanotechnology, Nanoscale 4 (2012) 1871—1880.

[51] A. Elbakry, A. Zaky, R. Liebl, R. Rachel, A. Goepferich, M. Breunig, Layer-by-Layer assembled gold nanoparticles for siRNA delivery, Nano Lett. 9 (2009) 2059—2064.

[52] E. Haller, W. Lindner, M. Lämmerhofer, Gold nanoparticle—antibody conjugates for specific extraction and subsequent analysis by liquid chromatography—tandem mass spectrometry of malondialdehyde-modified low density lipoprotein as biomarker for cardiovascular risk, Anal. Chim. Acta 857 (2015) 53—63.

[53] A.R. Patel, D. Lau, J. Liu, Quantification and characterization of micrometer and submicrometer subvisible particles in protein therapeutics by use of a suspended microchannel resonator, Anal. Chem. 84 (2012) 6833—6840.

[54] T.P. Burg, M. Godin, S.M. Knudsen, W. Shen, G. Carlson, J.S. Foster, K. Babcock, S.R. Manalis, Weighing of biomolecules, single cells and single nanoparticles in fluid, Nature 446 (2007) 1066—1069.

[55] D. Weinbuch, S. Zölls, M. Wiggenhorn, W. Friess, G. Winter, W. Jiskoot, A. Hawe, Micro—flow imaging and resonant mass measurement (archimedes) — complementary methods to quantitatively differentiate protein particles and silicone oil droplets, J. Pharm. Sci. 102 (2013) 2152—2165.

[56] M.G. von Muhlen, N.D. Brault, S.M. Knudsen, S. Jiang, S.R. Manalis, Label-free biomarker sensing in undiluted serum with suspended microchannel resonators, Anal. Chem. 82 (2010) 1905—1910.

[57] H. Fischer, I. Polikarpov, A.F. Craievich, Average protein density is a molecular-weight-dependent function, Protein Sci. 13 (2004) 2825—2828.

[58] Y.-Y. Liang, L.-M. Zhang, Bioconjugation of papain on superparamagnetic nanoparticles decorated with carboxymethylated chitosan, Biomacromolecules 8 (2007) 1480—1486.

Supplementary Material

Papain-functionalized Gold Nanoparticles as Heterogeneous Biocatalyst for Bioanalysis and Biopharmaceuticals Analysis

Siyao Liu, Markus Höldrich, Adrian Sievers-Engler, Jeannie Horak, Michael Lämmerhofer*

Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen,

Auf der Morgenstelle 8, 72076 Tübingen, Germany

*Corresponding author:

Prof. Dr. Michael Lämmerhofer

Phone: +49 70712978793

E-mail: michael.laemmerhofer@uni-tuebingen.de

Postal address: Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

Determination of kinetic parameters of immobilized and free papain

The enzyme activity of free and immobilized papain was evaluated at 37 ˚C at pH 6.8 by investigating the hydrolysis of BApNA as the substrate (Figure S1). To determine the kinetic parameters (K_m, V_{max}, k_{cat}) the reaction was carried out at variable substrate concentrations with free papain and immobilized papain-GNP conjugates obtained from different concentrations of papain in the reaction mixture. The kinetic parameters were derived from the Lineweaver Burk plots and the results are given in Figure 5 and Table 3 of the main document along with corresponding data of a reaction with 1 mg mL^{-1} free papain.
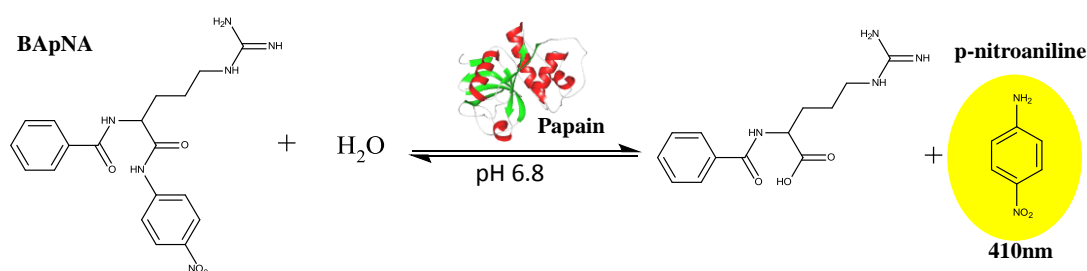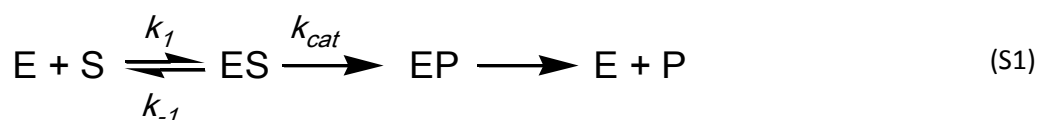

Figure S1. Reaction scheme for the hydrolysis of the substrate BApNA by papain. The activity of papain can be measured by the released *p*-nitroaniline (pNA) at an absorbance wavelength of 410 nm.

This single-substrate enzyme reaction can be described by the following equation (eq. S1)

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} EP \longrightarrow E + P \tag{S1}$$

Wherein E, S, ES, P and EP represent the enzyme, substrate, enzyme-substrate complex, product and enzyme-product complex concentrations, k_1 and k_{-1} are the rate constants for enzyme-substrate association and dissociation, respectively, and k_{cat} is the rate constant for product formation (also termed turnover number). The kinetics of this reaction is usually investigated by measuring the initial

velocities v_0 in dependence of the substrate concentration. The results are then evaluated in terms of the Michaelis-Menten equation (eq. S2)

$$v_0 = \frac{V_{max} \cdot [S]}{K_m + [S]}$$ (S2)

wherein V_{max} is the maximal reaction rate achieved when all enzyme is saturated with substrate ($[ES]=[E]_t$; with $[E]_t$ being the total enzyme concentration in the reaction i.e. $[E] + [ES]$). K_m is the Michaelis constant which is defined by (eq. S3)

$$K_m = \frac{k_{-1} + k_{cat}}{k_1}$$ (S3)

As mentioned above, the data have been evaluated in the linearized form, i.e. the Lineweaver-Burk plot (eq. S4)

$$\frac{1}{v_0} = \frac{K_m}{V_{max} \cdot [S]} + \frac{1}{V_{max}}$$ (S4)

In a reaction with free enzyme K_m does not change with enzyme concentration in the reaction mixture. Upon immobilization of enzyme, K_m might change with enzyme concentration immobilized on the carrier because the conformation of the enzyme and steric access of the binding site might be altered in dependence of the enzyme's surface concentration. For instance, at low surface concentrations the active site might be better accessible than at high coverage for which limited access to the enzyme's catalytic site might compromise the rate constant for association.

Figure S2 shows a plot of K_m values in dependence of distinct papain concentrations in the reaction mixture during bioconjugation. It becomes evident that K_m is in the same order of magnitude for all 4 bioreactors (Figure S2). Hence, it can be concluded that access to active sites is not compromised at higher papain surface concentrations.
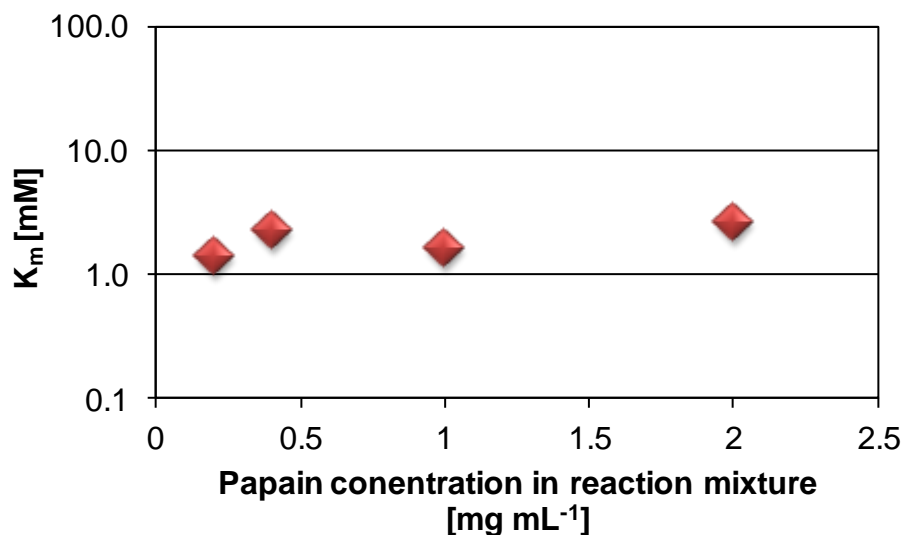
Figure S2: Michaelis-Menten constant versus papain concentration in the reaction mixture in the course of preparation of papain-GNP bioconjugate.

In sharp contrast, V_{max} changes linearly with the enzyme concentration in accordance to eq. S5

$$V_{max} = k_{cat} \cdot [E]_t \qquad \qquad \text{(S5)}$$
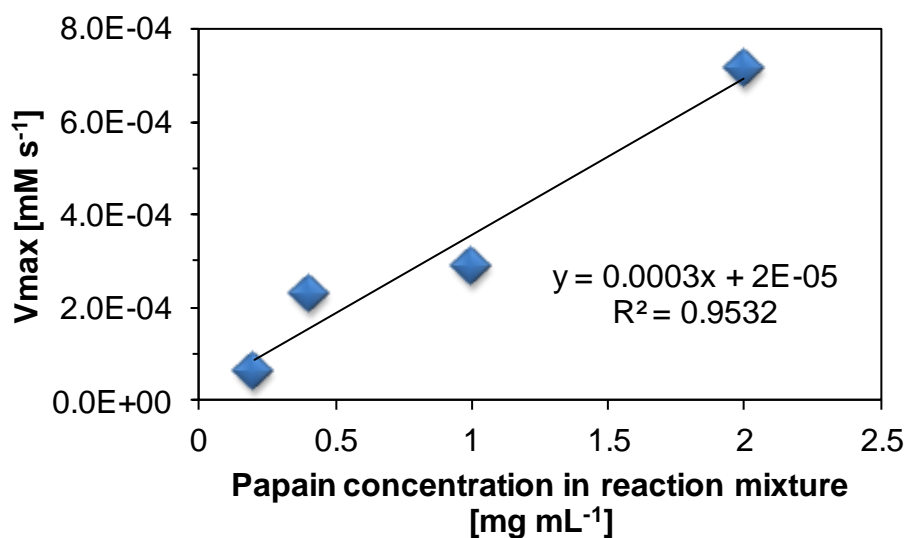

y = 0.0003x + 2E-05
R² = 0.9532

Figure S3: V_{max} in dependence of the papain concentration in the reaction mixture in the course of preparation of papain-GNP bioconjugate (Standard Error of Est. = 0.0000734; P-value intercept = 0.7548; P-value slope = 0.0237).

This is also found for a series of enzyme reactions performed with the bioreactors having distinct papain coverages (Figure S3). Higher enzyme concentrations in the incubation mixture due to use of papain-GNP conjugates with higher surface coverages (which result from higher papain concentrations in the reaction mixtures during protein coupling step) accelerate the conversion.

Eq. S5 can be used to calculate k_{cat}, the turnover number, which is a better figure to define the reaction rate. In this work k_{cat} has been calculated for free papain and two bioconjugates with distinct surface coverage of papain (see Table 3 of main document). Finally, the ratio of k_{cat}/K_m, termed specificity constant or catalytic efficiency, is a measure of how efficiently an enzyme converts substrates into products and is given in Table 3 of the main document as well.