Receptor mediated delivery of gRNAs to induce RNA editing

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Oliver Heß

aus Nürtingen

Tübingen

2021

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

Tag der mündlichen Qualifikation:	22.03.2022
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Thorsten Stafforst
2. Berichterstatter:	Prof. Dr. Dirk Schwarzer

DANKSAGUNG

An dieser Stelle möchte ich mich bei allen Personen bedanken, dich mich bei der Erstellung dieser Arbeit begleitet und unterstützt haben.

Mein Dank gilt meinem Doktorvater Herrn Prof. Dr. Thorsten Stafforst, der mir stets mit zahlreichen Ideen und Ratschlägen zur Seite stand und mir die akademische Welt nähergebracht hat. Des Weiteren möchte ich mich dafür bedanken, dass mir sowohl für meine wissenschaftliche Verwirklichung als auch für meine persönliche Entwicklung stets alle Freiheiten zur Verfügung standen.

Ich möchte mich ebenfalls bei meinem Zweitbetreuer Herrn Prof. Dr. Dirk Schwarzer bedanken, der sich freundlicher Weise bereit erklärt hat, auch als Zweitgutachter dieser Arbeit zu fungieren.

Mein Dankeschön gilt auch meinen ehemaligen Kollegen der Stafforst Gruppe für die gegenseitige Unterstützung und die sehr angenehme und bereichernde Arbeitsatmosphäre. Im Besonderen möchte ich mich jedoch bei meinen ehemaligen Bürokollegen Anna Stroppel, Alfred Hanswillemenke und Ngadhnjim Latifi für die konstruktiven Diskussionen, den aufgeschlossenen Gedankenaustausch sowie die unterhaltsame und angenehme Zeit bedanken.

Vielen Dank auch meinen ehemaligen Kommilitonen Anna Staerz und Torben Saatkamp, sowie allen meinen Freunden, welche mir stets mit einem offenen Ohr zur Seite standen.

Zusätzlich möchte ich mich bei Torben Saatkamp und Ngadhnjim Latifi für das Korrekturlesen dieser Arbeit bedanken.

Ein ganz besonderer Dank gilt meiner gesamten Familie, im Besonderen jedoch meiner Mutter Beate, meinem Vater Hartmut, sowie meinem Bruder Benjamin und seiner Frau Marina, welche mich stets in allen Bereichen unterstützten und mir so vieles in meinem Leben ermöglicht haben. Dies gilt ebenso für meine Großeltern, bei denen ich mich leider nicht mehr in einer angemessenen Art und Weise bedanken kann.

Schlussendlich möchte ich mich ganz besonders herzlich bei meiner Lebensgefährtin Pia bedanken, die mir stets bedingungslosen Rückhalt gegeben hat und die mir bei allem was ich mache zur Seite steht.

TABLE OF CONTENTS

List of abbreviations I

List of Figures VI

List of Schemes VII

List of Tables VII

Abstract IX

Zusammenfassung XI

1. Introduction 1

1.1. RNA editing 1

1.2. ADARs 2

1.2.1. The protein family and localization of ADARs 2

1.2.2. Mechanism of RNA binding and adenosine deamination 3

1.2.3. Physiological role and regulation of ADARs 6

1.3. Therapeutic approaches using DNA and RNA editing 9

1.3.1. Site-directed DNA editing 10

1.3.2. Site-directed RNA editing 11

1.4. Antisense- and siRNA-based therapeutic drug systems 17

1.4.1. RNAi and RISC 17

1.4.2. ASOs and RNase H1 18

1.5. Chemical modifications of oligonucleotides 20

1.5.1. Ribose modifications of oligonucleotides 21

1.5.2. Phosphate modifications of oligonucleotides 22

1.6. The asialoglycoprotein receptor and a targeted delivery 23

1.6.1. The asialoglycoprotein receptor 23

1.6.2. Triantennary *N*-acetyl galactosamine 25

1.6.3. The targeted delivery of oligonucleotides 26

2. Aims and approach of the thesis 28

3. Results and discussion 29

3.1. Synthesis of triantennary GalNAc and its conjugation to gRNAs 29

3.1.1. Synthesis of triantennary GalNAc 29

3.1.2. Chemical modification of gRNAs 37

3.1.3. GalNAc modification of SNAP®-ADAR gRNAs 41

3.1.4. GalNAc modification of RESTORE gRNAs 48

3.2. Molecular cloning of the ASGPR and gen. of ASGPR expressing cell lines 52
3.2.1. Isolation, molecular cloning and proof of concept of the ASGPR 52
3.2.2. The Generation of ASGPR expressing cell lines 56
3.3. gRNA mediated A-to-I editing of ASGPR expressing cell lines 69
3.3.1. RNA Editing of ASGPR and SNAP®-ADAR expressing FlpIn™ T-REx™ 293 cells targeting STAT1 Y701C 69
3.3.2. RNA Editing of ASGPR expressing HeLa cells targeting GAPDH L157L 79
4. Summary and Conclusion 92
5. Outlook 95
6. Methods and Materials 96
6.1. General approach 96
6.2. Consumables 96
6.2.1. Antibodies 96
6.2.2. Solvents and Chemicals 96
6.2.3. Media, buffers, solutions and additives 99
6.2.4. Commercially available kits 102
6.2.5. Enzymes and restriction enzymes 102
6.2.6. Primers 103
6.2.7. gRNAs 105
6.2.8. Cell lines 106
6.2.9. Plasmids 107
6.3. Analytics and Equipment 108
6.3.1. Nuclear magnetic resonance spectroscopy 108
6.3.2. LCMS analysis 108
6.3.3. HPLC analysis 108
6.3.4. Preparative HPLC separation 109
6.3.5. HRMS analysis 109
6.3.6. LCMS analysis of oligonucleotides 109
6.3.7. UV spectroscopy 109
6.3.8. Thin layer and column chromatography 109
6.3.9. Microscopy 110
6.3.10. Western Blot imaging 110
6.3.11. Nucleic acid and protein quantification 110
6.3.12. Fluorescent imaging of Urea-PAGE 111

6.4. Synthesis of triantennary GalNAc 111
6.4.1. Synthesis of compound 2 111
6.4.2. Synthesis of compound 3 112
6.4.3. Synthesis of compound 4 113
6.4.4. Synthesis of compound 5 114
6.4.5. Synthesis of Compound 6 114
6.4.6. Synthesis of Compound 7 115
6.4.7. Synthesis of Compound 9 116
6.4.8. Synthesis of Compound 10 117
6.4.9. Synthesis of Compound 11 in a small scale approach 118
6.4.10. Synthesis of Compound 11 in a large scale approach 119
6.4.11. Synthesis of Compound 11 using active ester 12 121
6.4.12. Synthesis of Compound 12 121
6.4.13. Synthesis of Compound 13 122
6.4.14. Synthesis of Compound 13 123
6.4.15. Synthesis of Compound 14 124
6.4.16. Synthesis of Compound 15 125
6.4.17. Synthesis of Compound 16 126
6.4.18. Synthesis of Compound 16 127
6.4.19. Synthesis of compound 17 128
6.4.20. Synthesis of compound 17-NHS 128
6.4.21. Synthesis of Compound 18 129
6.4.22. Synthesis of Compound 18-NHS 129
6.4.23. Synthesis of Compound 19 130
6.5. Molecular biology methods 131
6.5.1. General procedures 131
6.5.2. DNAse digest with DNAse I 131
6.5.3. Amplification of template DNA using Phusion DNA polymerase 131
6.5.4. Amplification of template DNA using *Taq* DNA polymerase 132
6.5.5. Agarose gel electrophoresis and gel extraction 133
6.5.6. Restriction digest and ligation for subcloning 133
6.5.7. Heat-shock transformation into $CaCl_2$ competent *E.Coli* XL1-blue, overnight culture and plasmid isolation 134
6.5.8. Colony-PCR using *Taq* DNA polymerase 135

6.5.9. RNA isolation using Monarch® RNA Cleanup Kit (10 µg) 135
6.5.10. DNAse digest and RT-PCR using One Step RT-PCR Kit 136
6.5.11. Nested-PCR 137
6.5.12. Sequencing of plasmid DNA and PCR amplicons 137
6.6. gRNA synthesis and purification 138
6.6.1. Conjugation of BisBG-COONHS to NH_2-functionalized oligonucleotides 138
6.6.2. Purification of conjugated oligonucleotides using preparative Urea-PAGE 138
6.6.3. Quantification of DTT using Ellman's reagent 139
6.6.4. Deprotection of Thiol functionalized gRNAs 257 and 258 140
6.6.5. Conjugation of GalNAc-Maleimide to thiol funct. gRNAs 257 and 258 140
6.6.6. Deprotection of Thiol functionalized gRNAs 471-473 140
6.6.7. Conjugation of GalNAc-Maleimide to thiol funct. gRNAs 471-473 141
6.6.8. Conjugation of GalNAc-COONHS to NH_2-funct. oligonucleotides 141
6.6.9. Analytical Urea-PAGE of conjugated oligonucleotides 141
6.6.10. Labeling of oligonucleotides with ATTO 594-NHS ester 142
6.7. RNA isolation and cDNA synthesis for molecular cloning of ASGPR 142
6.7.1. RNA isolation of HepG2 cells using TRI Reagent® 142
6.7.2. DNAse digestion and reverse transcription 143
6.7.3. Amplification of cDNA and cloning of ASGPR subunits 143
6.8. Cell culture techniques for human cell lines 144
6.8.1. General procedures for cell culture 144
6.8.2. Generation of stable expressing cell lines 146
6.8.3. Characterizations of stable expressing cell lines 149
6.8.4. Preparation of whole cell lysates, protein separation and western blotting 151
6.8.5. Transfections and receptor mediated uptakes of gRNAs 154
7. References 162
8. Appendix 198
8.1. Supplementary Information 198
8.1.1. Supplementary Figures 198
8.1.2. Supplementary schemes 211
8.1.3. Supplementary tables 213
8.1.4. NMR spectra 214
8.1.5. High resolution mass spectra 227

8.1.6. Mass spectrometry reports 233
8.1.7. Mass spectra 237
8.1.8. HPLC chromatograms 244
8.1.9. UV/VIS spectrometry 248
8.1.10. Vector sequences 249
8.1.11. Overview of experiments 268
8.2. Conference and retreat contributions 270

List of abbreviations

°C	Degree Celsius
AAT	α-1-antitrypsin deficiency
AAV	Adeno-associated viruses
Ac	Acetyl
ACN	Acetonitrile
AcOH	Acetic acid
ADAR	Adenosine deaminase that act on RNA
AGS	Aicardi-Goutières syndrome
AHP	Acute hepatic porphyria
ALS	Amyotrophic lateral sclerosis
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
APS	Ammonium persulfate
aq	Aqueous
ASGPR	Asialoglycoprotein receptor
ASO	Antisense oligonucleotide
ATTR	Hereditary transthyretin-mediated amyloidosis
Boc	*tert*-butyloxycarbonyl
bp	Base pair
brine	Saturated NaCl aq
BSA	Bovine serum albumin
bw	Backward
Cas	CRISPR-associated genes
Cbz	Carboxybenzyl (benzyloxycarbonyl)
cDNA	Complementary deoxyribonucleic acid
CFTR	Cystic fibrosis transmembrane conductance regulator
$CHCl_3$	Chloroform
conc.	Concentrated
COSY	Correlation spectroscopy
CRD	Carbohydrate recognition domain
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
CVD	Cardiovascular disease

DCM	Dichloromethane
DIC	*N,N'*-Diisopropylcarbodiimide
DIPEA	*N,N*-Diisopropylethylamine
DMAP	4-(Dimethylamino)pyridin
DMD	Duchenne muscular dystrophy
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
ds	Double-stranded
DSB	Double-strand break
dsRBD	Double-stranded RNA binding domain
E.Coli	*Escherichia coli*
EBV	Epstein-Barr virus
EDCI	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimid
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
eGFP	Enhanced green fluorescence protein
EndoV	Endonuclease V
ER	Endoplasmic reticulum
ESI	Electron spray ionisation
Et	Ethyl
FA	Formic acid
FACS	Fluorescence-activated cell sorting
FDA	U.S. Food and Drug Administration
FITC	Fluorescein 5(6)-isothiocyanate
fw	Forward
g	Gram
g	gravitational acceleration
GalNAc	triantennary *N*-acetyl galactosamine
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GLP-1	Glucagon-like peptide 1
GOI	Gene of interest

gRNA	guide ribonucleic acid
GTP	Guanosine-5'-triphosphate
h	Hour
h	*human*
HBTU	3-[Bis(dimethylamino)methyliumyl]-3*H*-benzotriazol-1-oxide hexafluorophosphate
HCMV	*Human cytomegalovirus*
HDR	Homology-directed repair
HEK	Human embryonal kindey
HEPN	Higher eukaryotes and prokaryotes nucleotide-binding
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HOBt	1-Hydroxybenzotriazol
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HRP	Horseradish peroxidase
HSQC	Heteronuclear single-quantum correlation spectroscopy
IFN	Interferon
IHP	Inositol hexakisphosphate
kb	Kilo base
L	Litre
LCMS	Liquid chromatography mass spectrometry
LEAPER	Leveraging Endogenous ADAR for Programmable Editing of RNA
LINE	Long interspersed elements
Lit.	Literature
M	Molar
MAVS	Mitochondrial antiviral-signaling adaptor protein
MCP	MS2 bacteriophage coat protein
MDA5	Melanoma differentiation-associated protein 5
Me	Methyl
MECP2	Methyl CpG binding protein 2
MHz	Megahertz
min	Minutes

mol	Mole
MS	Mass sprectrometry
NES	Nuclear export signal
NGS	Next generation sequencing
NHEJ	Nonhomologous end joining
NHS	*N*-Hydroxysuccinimide
NLS	Nuclear localization signal
NMR	Nuclear magnetic Resonance spectrometry
Npom	6-nitropiperonyloxy-methyl
nt	Nucleotide
Oligo	Oligonucleotide
ON	overnight
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAM	Protospacer adjacent motifs
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
Pd/C	Palladium on carbon
PfpOH	Pentafluorophenol
Ph-PCR	Phusion polymerase chain reaction
PKR	Protein kinase R
PMO	Phosphorodiamidate morpholino oligomer
PO	Phosphate
ppm	Parts per Million
PS	Phoshorothioate
PVDF	Polyvinylidene fluoride
RAN	RAs-related Nuclear protein
RESTORE	Recruiting Endogenous ADAR to Specific Transcripts for Oligo-nucleotide-mediated RNA editing
R_f	Retention factor
RIG	Retinoic acid-inducible gene
RISC	RNA-induced silencing complex

RLR	Retinoic acid-inducible gene-I-like receptor
RNA	Ribonucleic acid
RNAi	RNA-interference
rpm	Revolutions per minute
RT	Room temperatur
RT-PCR	Reverse transcription polymerase chain reaction
sASGPR	Soluble ASGPR
SDS	Sodium dodecyl sulfate
sec	Second
seq	Sequence
SINE	Short interspersed elements
SMA	Spinal muscular atrophy
SNP	Single nucleotide polymorphism
ss	Single-stranded
STAT1	Signal transducer and activator of transcription 1
TadA	tRNA adenosine deaminase
TAE	TRIS-Acetate-EDTA
TALEN	Transcription activator-like effector nuclease
Taq	*Thermus aquaticus*
TAR	Trans-activation response element
TBE	Tris/Borate/EDTA
TBP	TAR hairpin binding protein
TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween20
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofurane
TIE	Translation inhibitory elements
TMD	Transmembrane domain
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
Tris	Tris(hydroxymethyl)aminomethane

TRN1	Transportin-1
UHPLC	Ultra high performance liquid chromatography
UTR	Untranslated region
UV	Ultraviolet
VIS	Visible
XPO1	Exportin-1
XPO5	Exportin-5
ZFN	Zinc-finger nuclease

List of Figures

Figure 1: Molecular insights into RNA editing 1
Figure 2: Illustration of the three types of ADARs 2
Figure 3: Molecular insights to the RNA bound ADAR2 E488Q deaminase domain 4
Figure 4: Physiological role and regulation of ADAR mediated RNA editing 8
Figure 5: DNA base editor systems 11
Figure 6: Different site-directed RNA editing approaches to introduce A-to-I substitutions 15
Figure 7: Impacts of ssASO, siRNA or miRNA to RNA stability and translation modulation 19
Figure 8: Chemical modifications of nucleotides 21
Figure 9: Triantennary GalNAc and ASGPR mediated uptake 24
Figure 10: Fluorescence imaging of GalNAc conjugated gRNA 257 and 258 45
Figure 11: Fluorescence imaging of GalNAc conjugated gRNA 471-473 46
Figure 12: Fluorescence imaging of GalNAc conjugated gRNAs 324 and TMR189 51
Figure 13: Sequence alignment of the different ASGPR isoforms the the vector sequences 54
Figure 14: Live cell imaging of the transient transfection of the ASGPR into HEK 293T 55
Figure 15: Generation of a SA1Q and H1a stably expressing FlpIn™ T-REx™ 293 cell line 58
Figure 16: Charazerization of a SA1Q and H1a stably expressing FlpIn™ T-REx™ 293 cell line 61
Figure 17: Generation of a H1a stably expressing HeLa cell line 64
Figure 18: Characterization of a H1a stably expressing HeLa cell line 65
Figure 19: FACS sorting of the H1a stably expressing HeLa cell line 67
Figure 20: Transfection of BisBG gRNAs into FlpIn™ T-REx™ 293 71
Figure 21: Receptor mediated endocytosis of SNAP®-ADAR gRNAs 324 and 507 72
Figure 22: Transfection and receptor mediated uptake of ATTO 594 labeled BisBG gRNAs 76

Figure 23: Transfection of RESTORE v2 gRNas into HeLa cells 80
Figure 24: Receptor mediated uptake of RESTORE v2 into unsorted HeLa cells 82
Figure 25: Receptor mediated uptake of RESTORE v2 into sorted HeLa cells 84
Figure 26: Transfection and receptor mediated endocytosis of ATTO 594 labeled RESTORE v2 gRNAs into HeLa cells 89

List of Schemes

Scheme 1: Triantennary N-acetyl galactosamine (14). 29
Scheme 2: Synthetic route of the tris-based triantennary core 30
Scheme 3: Synthetic route of the terminal N-acetyl galactosamine residues 31
Scheme 4: Conjugation of the N-acetyl galactosamine residue to the triantennary core 33
Scheme 5: Activation of the terminal carboxylic acid and conjugation to the triant. core 34
Scheme 6: Deprotection of compound 11 35
Scheme 7: Conjugation and functionalization of compound 14 36
Scheme 8: General workflow of the conjugation of BisBG-COOH (18) to NH2- gRNAs 40
Scheme 9: Conjugation of 4-methoxybenzyl mercaptan to compound 16 42
Scheme 10: General workflow of the conjugation of compound 16 to BisBG-gRNA 43
Scheme 11: General workflow of the conjugation of NH_2-terminal gRNAs to compound 17 50

List of Tables

Table 1: Observed and reported ^{1}H-NMR signals of the anomeric proton of compound 9 32
Table 2: Expected and found ^{1}H-NMR signals of compound 11 33
Table 3: Overview of the BisBG-COOH (18) conjugated gRNAs 38
Table 4: Different approaches to conjugate GalNAc-COONHS (17-NHS) to NH_2-gRNAs 48
Table 5: NCBI reference sequences of the different ASGPR subunits and isoforms 53
Table 6: Antibodies used for immunofluorescence and western blotting 96
Table 7: Solvents used for chemical synthesis 96
Table 8: Chemicals used for synthesis 97
Table 9: Commercially available media, buffers, solutions and additives 99
Table 10: Prepared buffers with corresponding components and concentrations 100
Table 11: Commercially available kits 102
Table 12: Commercially available enzymes and restriction enzymes 102
Table 13: Primers used for molecular cloning, PCR amplification and sequencing. 103
Table 14: Overview of used gRNAs 105
Table 15: Overview of generated plasmids 107

Table 16: Excitation and emission wavelengths λ 110
Table 17: General approach for DNAse digests using DNAse I 131
Table 18: General procedure for template DNA amplifications using Ph-DNA pol. 131
Table 19: Thermocycler program for PCRs using Phuison DNA polymerase 132
Table 20: General procedure for template DNA amplifications using Taq DNA pol. 132
Table 21: Thermocycler program for PCRs using Taq DNA polymerase 133
Table 22: Overview over the conditions used for agarose gel electrophoresis 133
Table 23: Exemplary restriction digest 134
Table 24: General procedure for colony-PCRs using Taq DNA polymerase 135
Table 25: General procedure for combined reverse transcription of isolated RNA and amplification of cDNA targeting GAPDH 136
Table 26: General procedure for combined reverse transcription of isolated RNA and amplification of cDNA targeting STAT1 136
Table 27: Thermocycler program of the reverse transcription and amplification of cDNA 137
Table 28: Used sequencing primers of PCR amplicons for evaluation of editing yields 138
Table 29: Polyacrylamide gels (15 % and 20 %) used for oligonucleotide purification 139
Table 30: General procedure for cDNA syntheses of the ASGPR subunits using reverse transcription 143
Table 31: Thermocycler program for the reverse transcription 143
Table 32: Primer set used for amplification of ASGPR cDNA 144
Table 33: Information about sample preparation for western blotting 152
Table 34: Information about antibody treatment used for western blotting 154
Table 35: Overview of the performed experiments and assignments to the described sections and laboratory journal numbers 268

ABSTRACT

About a quarter of all human pathogenic genetic variants are related to a G-to-A single nucleotide polymorphism. As a consequence, the deamination of adenosines by adenosine deaminases that act on RNA (ADARs) provides a promising approach to recode the incorrect genetic information on the RNA level. Because of the biochemical interpretation of inosine as guanosine, site-directed RNA editing arose as a highly promising post-transcriptional modification to introduce wide-ranging consequences in RNA function by A-to-I base substitutions. While ADARs are highly selective on double-stranded RNAs, so-called gRNAs are applied to hybridize with the target site mRNA and to form the necessary dsRNAs. However, besides viral transductions, especially cell toxic transfection reagents are nowadays most commonly used within cell culture approaches to internalize gRNAs, which are mandatory for the hybridization with the target mRNA and the formation of double-stranded RNA. For this reason, it is the aim of this doctoral thesis to transfer the advantages of a receptor mediated endocytosis of other RNA targeting therapeutic applications to two different RNA editing systems: The SNAP®-ADAR system, using an artificial deaminase fusion protein, and the RESTORE system, recruiting endogenous ADAR. Especially the asialoglycoprotein receptor (ASGPR) and its synthetic triantennary *N*-acetyl galactosamine ligand (GalNAc) were thought to add to the toolbox of site-directed editing systems and their potential in future therapeutic applications. Therefore, a novel wet chemical synthetic route was established to modify disulfide terminal gRNAs with a prior synthesized GalNAc-maleimide derivative. For each editing system the applicability of GalNAc modified gRNAs was demonstrated in appropriate and specifically engineered cell lines. Due to the application of fluorescence imaging, as well as Chloroquine as an additive with endosomal disruptive properties, it was possible to conclude that a sufficient amount of gRNA was internalized by the receptor mediated endocytosis but did not become available within the cytoplasm or the nucleus. Using Chloroquine, it was possible to obtain editing yields up to 51 % in the open reading frame (ORF) of an endogenous transcript (GAPDH) utilizing the receptor mediated endocytosis of RESTORE v2 gRNAs and low gRNA concentrations of 0.2 µM. In this context, and regarding the reduced lysosomal stability of the used gRNAs, an endosomal entrapment and the limited endosomal release of gRNAs was elaborated as major bottleneck for future applications. In addition, the still uncharacterized, but not negligible gymnotic uptake or receptor mediated endocytosis of unconjugated phosphorothioate (PS) oligonucleotides provided still highly competing pathways for the targeted delivery. However, the receptor mediated delivery of GalNAc

conjugated gRNAs into ASGPR expressing cells is providing a promising technique to prevent the necessity of cell toxic transfections reagents, and especially for the use of the RESTORE system, another step towards *in vivo* approaches and a potential future therapeutic application was made.

ZUSAMMENFASSUNG

Etwa ein Viertel aller humanpathogenen genetischen Varianten stehen im Zusammenhang mit einer G-zu-A Punktmutation und insbesondere RNA spezifische Adenosin-Desaminasen (ADARs) boten eine vielversprechende Möglichkeit, die fehlerhaften Informationen auf RNA Ebene zu korrigieren. Da Inosin biochemisch als Guanosin interpretiert wird, entwickelte sich daraus die zielgerichtete A-nach-I Editierung, als vorteilhafte und posttranskriptionelle Modifikation, welche weitreichende Konsequenzen für die RNA und deren Funktion mit sich bringt. Da ADARs hochselektiv an doppelsträngige RNA binden, werden zur Hybridisierung mit der Ziel mRNA so genannte gRNAs verwendet. Neben der viralen Transduktion werden für diese heutzutage in Zellkulturversuchen jedoch immer noch zelltoxische Transfektionsreagenzien als häufigstes Transportmittel verwendet. Daher was es das Ziel dieser Doktorarbeit, die Vorteile einer rezeptorgesteuerten Endozytose anderer RNA Therapeutika auf zwei unterschiedliche Editierungssysteme zu übertragen: Auf das SNAP®-ADAR System, welches ein künstliches Desaminase Fusionsprotein verwendet, sowie auf das RESTORE System, welches in der Lage ist endogenes ADAR zu rekrutieren. Insbesondere der Asialoglycoproteinrezeptor (ASGPR) und sein synthetischer, dreiarmiger *N*-Acetylgalaktosaminligand (GalNAc), sollten sich dafür eignen, um das generelle Repertoire der zielgerichteten RNA Editierung zu erweitern und deren Potential als zukünftiges Therapeutikum zu steigern. Mit diesem Ziel wurde eine neue und nasschemische Methode etabliert, um Disulfid terminale gRNAs an ein zuvor synthetisiertes GalNAc-Maleimidderivat zu konjugieren. Zusätzlich wurde die Anwendbarkeit von GalNAc konjugierten gRNAs in entsprechenden und eigens dafür erstellten Ziellinien demonstriert. Durch die Anwendung fluoreszenzbildgebender Methoden, sowie der Verwendung von Chloroquin, welches einen destabilisierenden Effekt auf Endosomen aufweist, wurde jedoch klar, dass zwar eine ausreichende Menge an gRNAs mittels der rezeptorgesteuerten Endozytose in die Zellen transportiert wurde, diese jedoch weder im Zytoplasma noch im Zellkern zur Verfügung stand. Des Weiteren wurden unter der Verwendung von Chloroquin und der rezeptorgesteuerten Aufnahme einer RESTORE v2 gRNA, sowie einer niedrigen gRNA Konzentration (0.2 µM), Editierungsausbeuten von bis zu 51 % in der kodierenden Sequenz eines endogenen Transkripts (GAPDH) erzielt. In diesem Zusammenhang und wegen der geringeren lysosomalen Stabilität der verwendeten gRNAs, wurde ein endosomaler Einschluss und die damit verbundene, begrenzte endosomale Freisetzung der gRNAs als limitierender Faktor für zukünftige Anwendungen herausgearbeitet. Darüber hinaus boten die noch nicht vollständig

charakterisierte, aber nicht zu vernachlässigende gymnotische Aufnahme, sowie die rezeptorgesteuerte Endozytose von unkonjugierten, aber Phosphorothioat modifizierten Oligonukleotiden, erhebliche und konkurrierende Transportwege in die Zelle. Nichts desto trotz bietet der rezeptorgesteuerte Transport von GalNAc konjugierten gRNAs in ASGPR exprimierende Zellen eine vielversprechende Möglichkeit, um die Verwendung zelltoxischer Transfektionsreagenzien zu vermeiden. Des Weiteren und insbesondere für die Verwendung des RESTORE Systems, ist es ein weiterer Schritt in Richtung einer *in vivo* Anwendung und dadurch einer potenziellen und zukünftigen Applikation als weiteres RNA Therapeutikum.

1. Introduction

1.1. RNA editing

With the discovery of an unknown enzymatic activity to unwind double-stranded RNA (dsRNA) in *Xenopus laevis* oocytes and embryos by Bass and Weintraub in 1987[1] followed by the discovery of its activity to deaminase adenosines[2,3], the field of RNA editing was born. An adenosine deaminase that acts on RNA (ADAR) converts adenosines to inosines (A-to-I) by hydrolytic deamination of the C^6 position (Figure 1a)[2,3]. Next to methylation or isomerization, A-to-I editing provides another post-transcriptional modification to introduce wide-ranging consequences in RNA function. Because inosine is able to pair stably with cytidine by Watson-Crick base pairing (Figure 1b), it is biochemically interpreted as guanosine by cellular processes[4,5]. Inosine can also pair weakly with uridine or adenosine and is therefore read as adenosine or uridine, respectively (Figure 1b)[6]. Introducing or deleting splice sites, altering microRNA recognition sites, or changing the meaning of specific amino acid codons during translation provide only a few post-transcriptional modifications, that are in relation to the biochemically interpretation of inosine as guanosine, adenosine, or uridine in mRNA transcripts[7–9].

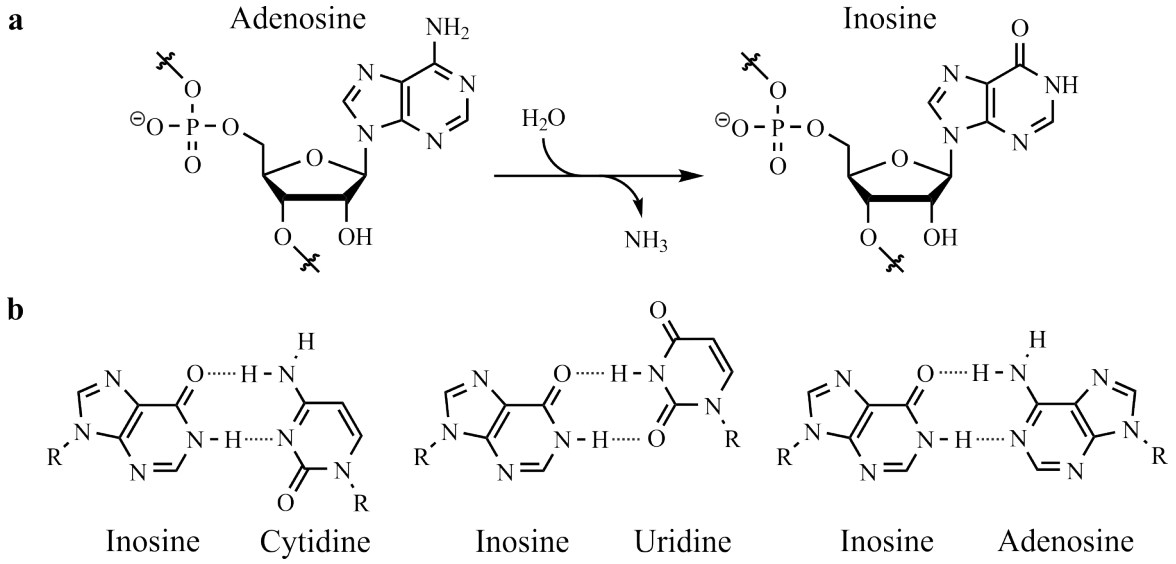


Figure 1: Molecular insights into RNA editing. *(a) ADAR catalyzed hydrolytic deamination of adenosine to inosine at the C^6 position. (b) Watson-Crick base paring of inosine with cytidine, uridine, and adenosine. Residue R is indicating the ribose unit of the corresponding nucleotide and Watson-Crick base pairing is indicated as dotted lines. Structures are adapted from ref.* 10.

Next to the family of ADAR's, there are two more families of enzymes that perform a post-transcriptional deamination: the family of an apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC), a cytidine deaminase, and the family of an adenosine deaminase that acts on tRNA (ADAT)[11]. By deamination, APOBEC enzymes perform a base substitution

of cytidine to uridine (C-to-U) of single stranded RNA, enabling also several post-transcriptional modifications as described before[12]. Besides the post-transcriptional modification of mRNA, inosines are also common modifications within tRNAs. Especially position 34, the first nucleotide of the anticodon (wobble position) is often a deaminized adenosine (inosine), which allows decoding of multiple cognate codons using a single tRNA[13,14]. A-to-I, as well as C-to-U RNA editing of mRNAs, ncRNAs, and tRNAs highly increases the regulatory and coding capacity of the genome and therefore the diversity of the proteome.

However, within this thesis, only ADAR mediated A-to-I editing is of importance and will be further discussed in the following sections.

1.2. ADARs

1.2.1. The protein family and localization of ADARs

In most vertebrates, there are three different types of ADARs: ADAR1[15], ADAR2[16] and ADAR3[17,18]. All three types of ADARs have functional domains in common, but there are also differences (Figure 2). All ADARs share a *C*-terminal deaminase domain, and a nuclear localization signal (NLS). However, only ADAR1 and ADAR2 seem to be catalytically active[17,18]. While ADAR1 contains three *N*-terminal double-stranded RNA binding domains (dsRBDs), ADAR2 and ADAR3 share only two and ADAR3 is the only type containing an arginine-rich single stranded RNA binding domain (R) in its *N*-terminal region[17].

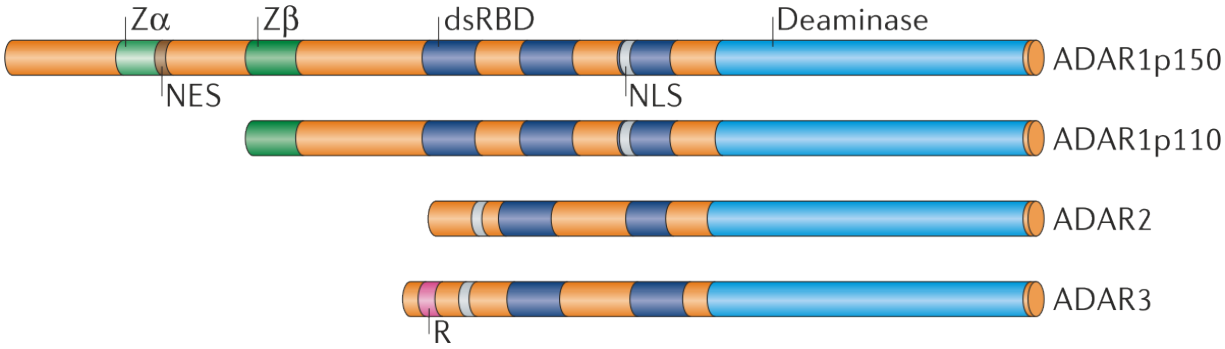


Figure 2: Illustration of the three types of ADARs. *All types have similar but also different functional domains: A deaminase domain (light blue), two to three double-stranded RNA binding domains (dsRBDs) (dark blue), a nuclear localization signal (NLS) (grey), a nuclear export signal (NES) (brown), a Z_α (light green) and a Z_β (dark green) Z-RNA binding domain and an arginine-rich single stranded RNA binding domain (R) (red). Figure is adopted from ref.* 10.

Besides the three different types of ADARs, ADAR1 is ubiquitously expressed in two different isoforms from different promotors, ADAR1p110 and p150[19]. ADAR1p150, an interferon inducibly expressed isoform of 150 kDa, contains both, a Z_α and a Z_β Z-RNA binding domain[20] and an additional nuclear export signal (NES). ADAR1p110, a constitutively

expressed and shorter isoform of 110 kDa, lacks the Z_α Z-RNA binding domain, including an *N*-terminal stretch and the NES. While both isoforms are known to shuttle between cytoplasm and nucleus, ADAR1p150 is mainly localized in the cytoplasm whereby ADAR1p110 accumulates in the nucleus[21]. The transport to the nucleus of both, ADAR1p150 and p110, is mediated by transportin-1 (TRN1), which is binding to dsRBD3 in the absence of dsRNA[22]. While exportin-5 (XPO5) and RAN-GTP are regulating the nuclear export of isoform p110 by binding to dsRBDs[23], the nuclear export of isoform p150 is mediated by Exportin-1 (XPO1) and RAN-GTP, binding to the NES[24]. ADAR2 is primarily expressed as a single isoform and contains a NLS, but no NES – homologous to ADAR1p110. ADAR2 is prominently localized in the nucleus and nucleolus[25] and is mostly expressed in lung and brain[16], whereby the expression of ADAR3 is restricted to the brain and post-mitotic neurons[17,18,26].

1.2.2. Mechanism of RNA binding and adenosine deamination

The molecular mechanism of the A-to-I conversion was unclear for a long period of time. However, researchers were able to provide the crystal structure of the *human* ADAR2 (*h*ADAR2) E488Q deaminase domain, bound to different dsRNA substrates[27]. For the hydrolytic deamination of adenosine, several compounds were found to be necessary. While proper protein folding is dependent on inositol hexakisphosphate (IHP), which is located within the enzyme core environed by several arginine and lysine residues, a zinc ion and a glutamate residue (E396) are responsible for the activation of water, which is used for the hydrolytic deamination (Figure 3)[28]. Mutational analysis of the glutamate to alanine has been shown to abolish the editing activity of ADAR1 (E912A) and ADAR2 (E396A). Therefore, the glutamate is assumed to mediate the proton transfer to and from the target adenosine[29,30]. Using a dsRNA substrate with 8-azanebularine, which is mimicking the tetrahedral intermediate at the target site, showed that the deaminase domain of *h*ADAR2 binds the dsRNA by interacting with the phosphodiester backbone of approximately 20 nucleotides (nt) (Figure 3)[27].

The accessibility of the target adenosine for deamination is provided by a base flipping mechanism and while the minor groove of the incorporated dsRNA is penetrated by a loop of the *h*ADAR2's deaminase domain, the position of the flipped out base is occupied by residue E488[27]. By conformational changes of the deaminase domain, induced by binding to dsRNA, a previously disordered loop (amino acid 454-477) is contacting the minor groove and is inserting into the adjacent major groove of the bound dsRNA. Furthermore, a change in conformation of the RNA substrate's A-form further exposes the flipped out base and increases the access of the adenosines C^6 position to the active site[27]. This loop sequence is different in *h*ADAR1[31] and a

mutational high-throughput approach of *h*ADAR2 revealed the importance of this region. This loop might also be responsible for the different substrate specifications of *h*ADAR1 and *h*ADAR2[31].

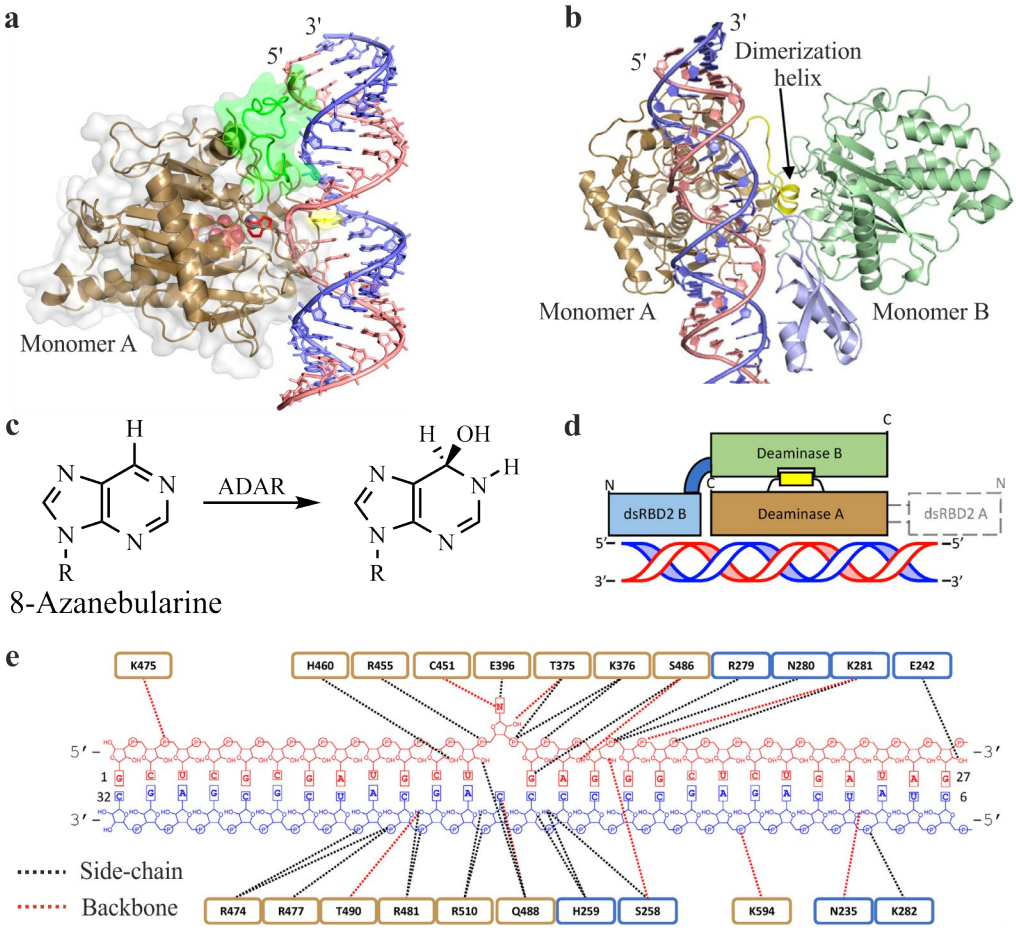


Figure 3: Molecular insights to the RNA bound ADAR2 E488Q deaminase domain. *(a) Structural view of a dsRNA bound ADAR2 E488Q deaminase domain monomer orthogonal to the dsRNA helical axis. (b) Structural view of a dsRNA bound asymmetric ADAR2 E488Q deaminase domain dimer orthogonal to the dsRNA helical axis. Color coding of (a) and (b): Monomer A (brown), Monomer B (dark green), flipped out base (red in a), zinc (yellow sphere in a), "disordered loop" (light green in a), dimerization helix (yellow in b), IHP (red, grey and orange spheres in a), dsRBD2 (light blue in b). (c) Reaction mechanism of ADAR2 showing 8-azanebularine and the hydrated product. R is indicating the ribose unit of the corresponding nucleotide. (d) Schematic illustration of the asymmetric protein dimer bound to dsRNA. Color coding is similar in (a) and (b). (e) Overview of interactions between a dsRNA substrate and amino acid residues of a deaminase domain. Structures are adapted from ref. 27 and* 32.

In combination with the different substrate specifications, *h*ADAR1 and *h*ADAR2 have also differences concerning the nearest neighbor preferences of the target adenosines. While the extent of adenosine deamination is increased for both, *h*ADAR1 and *h*ADAR2, having a 5'-nearest neighbor preference of U > A > C > G, their 3'-nearest neighbor preference is G > C ~ A > U and G > C > U ~ A, respectively[33]. Besides the 5'-nearest neighbor preferences, ADARs have also a common counter base preference. While a purine at the orphan base position is sterically more demanding than a pyrimidine base, A·C and A·U mismatches are more preferred

than A·A and A·G mismatches[34]. In *h*ADAR2, mutation of the glutamate residue 488 to glutamine (E488Q) provides a hyperactive variant with an enhanced catalytic activity[35]. The hyperactive mutant (E488Q) shows a similar binding affinity to dsRNA, but an increased base flipping opportunity. This leads to a higher catalytic rate but also a decreased specificity[35].

However, the detailed mechanism of substrate binding and target site recognition is not completely understood. While homodimerization of ADARs was shown in *Drosophila* or using FRET-based experiments with *h*ADAR1 and hADAR2 in an RNA dependent manner, ADAR1 showed also heterodimerization with RNA processing enzymes, such as Dicers, in an RNA independent manner[30,36–42]. Recently, researches provided also a crystal structure of two *h*ADAR2 deaminase domains bound to a 91 base pair RNA duplex, whereby one deaminase also contained its dsRBD2[32]. While the dsRBD is bound to the phosphodiester backbone in a similar fashion as previously described for RNA bound *rat* ADAR2 dsRBDs[43–46], the naked deaminase domain (Figure 3b, Monomer A) was involved in direct binding to the dsRNA substrate at the target site. The deaminase domain containing its dsRBD2 (Figure 3b, Monomer B) interacts with the dsRNA bound deaminase using protein-protein interactions. The RNA bound deaminase exposes a short α-helix (Figure 3b, dimerization helix, amino acids 501-509), which is recognized by the catalytic site of the second deaminase to participate a protein-protein binding. Interestingly, similar residues were involved in both, comprising the catalytic pocket for A-to-I substitutions as well as forming the protein-protein interactions. Furthermore, a sequence alignment of ADAR proteins from multiple species revealed a high degree of conservation of residues mediating the protein-protein binding[32]. Therefore, the family of ADARs provide a particularly and highly conserved interface which is able to perform protein-protein as well as protein-RNA interactions. Mutagenesis of both the catalytic site as well as the dimerization helix disrupts dimerization of the deaminase domains and showed adverse effects for the deamination reaction for most RNA substrates.

For ADAR1, there is no crystal structure reported yet. However, a comparable base flipping mechanism can be assumed, which was indicated by mutational analysis of the glutamate to alanine residue of *h*ADAR1 (E912A) to abolish the editing activity[29,30]. As mentioned before, this was also observed for *h*ADAR2 (E396A) and while a further mutational approach of a corresponding dimerization helix residue of *h*ADAR1p110 (D1023A) and *h*ADAR2 (D503A) lead to a substantial decrease of editing, the substrate specification of both mutants became different[32]. This is indicating a more complex substrate recognition and more data are necessary to fully understand the different substrate specifications of ADAR1 and ADAR2. For ADAR3,

no catalytic activity is reported yet. However, a regulatory function is indicated by the inhibition of RNA editing of other ADARs[17,47].

1.2.3. Physiological role and regulation of ADARs

The physiological role of ADAR mediated A-to-I editing can be distinguished into two major groups: editing of coding sequences (e.g. alteration of amino acid codons of the coding sequence) and editing of non-coding sequences (e. g. alteration of splice sites, miRNAs, long ncRNAs or tRNAs)[10]. Merging the nearest neighbor preferences mentioned before, 12 out of the 20 canonical amino acid codons are potentially targetable and recodible[48]. However, only a small fraction of physiological editing sites is well understood and forming homodimers as well as heterodimers, multiple regulatory mechanisms are probably involved in controlling A-to-I editing[32,39]. For instance, interactions of ADARs with other proteins of the cellular machinery, such as Dicers, are also increasing the impact of ADARs to their physiological role in an editing independent manner[49].

A nowadays well-known example where A-to-I editing mediates an alteration of protein function provides the transcript encoding the glutamate ionotropic receptor AMPA type subunit 2 (GRIA2 or GLUR2)[50–52]. While intron 11 and exon 11 of the GRIA2 pre-mRNA transcript are forming an imperfect dsRNA duplex, ADAR2 mediated editing results in an alteration of a glutamine (Q) codon (CAG) to an arginine (R) codon (CIG). A Q/R substitution of the AMPA receptor site highly reduces the Ca^{2+} permeability, which is essential for a postnatal survival of mice. *ADAR2* knockout mice exhibit onset epilepsy and postnatal lethality[51] and the level of editing at the Q/R site is supposed to be necessary for neural integrity, neural cell development and cell differentiation[53,54]. In patients with sporadic amyotrophic lateral sclerosis (ALS), deficient editing of the Q/R site seems also to cause the loss of function of motor neurons, which was also observed in *ADAR2* knockout mice[53,55]. Additionally, the formation of an imperfect hairpin of exon 13 and the adjacent intron, the substitution of an arginine (R) codon (AGG) to a glycine (G) codon (IGG) is catalyzed by ADAR1 and ADAR2[16]. The described R/G substitution was also found in GRIA3 and GRIA4, two further subunits of the AMPA receptor originating form a common pre-mRNA transcript, and changes the receptors desensitization[50]. In Glioblastoma, hypo-editing of GRIA2 is accompanied by an increased protein expression level of ADAR3. The higher expression level of ADAR3 is indicating a possible negative regulation of RNA editing by ADAR3, but a reason for the increased expression level needs to be explored[47].

Further well-characterized editing sites of protein coding sequences are located within the transcripts of the serotonine-2C receptor (5-$HT_{2C}R$)[56], the voltage-gated K^+ channel subfamily A member 1 ($K_v1.1$)[57], and the $GABA_A$ receptors subunit α3 (GABRA3)[58,59]. For 5-$HT_{2C}R$, the substitution of five adenosines of the second intracellular loop reduces the receptor's coupling efficiency with G proteins[56,60]. While recoding of the ion conducting pore of $K_v1.1$ enables a faster recovery and inactivation rate[57], editing of GABRA3 reduces trafficking as well as proper localization[58]. However, the protein alterations mentioned before are all related to neurotransmitter receptors and are located in the brain. While ADAR2, which is predominantly expressed in the brain, is responsible primarily for editing of coding sites[16], ubiquitously expressed ADAR1 is mainly responsible for editing of non-coding sequences such as miRNAs, and short as well as long interspersed elements (SINEs and LINEs)[61].

Alu elements comprise SINEs of about 300 nucleotides and are the most common SINEs in humans. More than a million copies were found throughout the *human* genome and the majority of editing sites are located within these repetitive elements[61,62]. While *Alu* repeats are involved in many cellular processes such as circular RNA biogenesis, transcriptional elongation and splicing, editing of these inverted repeats is also very crucial for their intracellular recognition[49,63]. Therefore, RNA editing of inverted repeats prevents interferon response and recognition by the innate immune system[64–66]. As primary editase of *Alu* repeats[61,67], ADAR1 deficiencies in mice cause embryonic death related to interferon overproduction, defective hematopoiesis and apoptosis[65,68–70]. However, lethality of mice embryos can be rescued by mutating the melanoma differentiation-associated protein 5 (MDA5) or the mitochondrial antiviral signaling adaptor protein (MAVS)[65,70]. Catalytically inactive ADAR1 mutations were also found in *human* patients with Aicardi-Goutières syndrome (AGS). AGS is an autosomal-recessive disorder, which is also related to interferon upregulation and an aberrant immune response induced by decreased editing levels within *Alu* repeats[71]. Editing of endogenous dsRNAs prevents MDA5 and protein kinase R (PKR) sensing and activation of an innate immune signaling such as the retinoic acid-inducible gene-I-like receptor pathway (RIG-1-like receptor or RLR pathway). Therefore, ADAR1 mediated editing avoids a translational shutdown of endogenous RNAs during interferon response (Figure 4b)[64].

However, editing of *Alu* repeats has much further implications than regulating the innate immune signaling of endogenous dsRNAs. A-to-I conversions are also able to generate splice donor (GU) or acceptor (AG) sites (Figure 4a). Prominent examples for RNA editing mediated exonizations are nuclear prelamin A[72] or the G protein coupled receptor 107[73]. Additionally, a

negative autoregulatory mechanism is reported for ADAR2[8,74]. Self-editing of an intronic sequence of its pre-mRNA transcript forms an alternative splice acceptor site and results in the suppression of ADAR2 expression. Furthermore, an endonuclease V (EndoV) induced degradation of hyper-edited *Alu* repeats might also serve as a regulatory mechanism of gene expression[75,76].

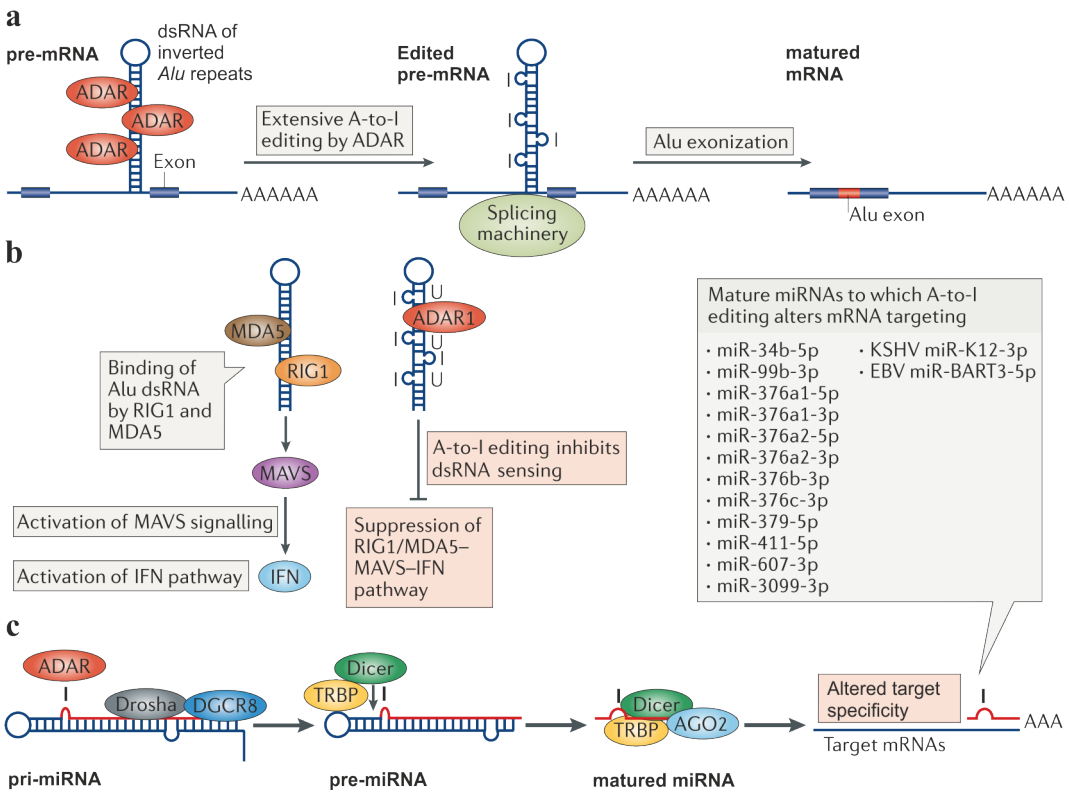


Figure 4: Physiological role and regulation of ADAR mediated RNA editing. *(a) Inverted Alu repeats of introns and 3'-untranslated regions (UTRs) are edited by ADARs to alter splice sites and induce exonization. (b) Extensive editing of double-stranded inverted Alu repeats prevents recognition by MDA5 and PKR, and inhibits an IFN induced immune signaling such as the RIG1-like receptor pathway. Unedited, long dsRNAs, such as viral dsRNAs, are recognized by MDA5 and PKR and IFN upregulation is inducing innate immune responses. (c) ADAR mediated RNA editing affects miRNA biogenesis, Drosha and/or Dicer cleavage, RISC loading as well as miRNA target specificity. Figures and captions are adapted from ref. 10.*

Next to regulations based on ADAR mediated RNA editing of *Alu* repeats such as innate immune responses or splice site alterations, RNA editing is providing a much more incisive impact to the cellular regulatory machinery. Thus, ADARs are also reported to edit miRNAs, which are particularly important for tissue differentiation, cell proliferation, viral defense and apoptosis[77–79]. While A-to-I editing within a matured miRNAs seed region alters its base pairing properties and the target specificity of the RNA-induced silencing complex (RISC)[80,81], editing of a pri-miRNA inhibits or stimulates RNA hairpin recognition and subsequent maturation by the Drosha-DGCR8 complex (Figure 4c)[82,83]. For instance, Epstein-Barr virus (EBV) is expressing miR-BART6-5p, a miRNA targeting the human Dicer mRNA when not edited[84].

Therefore, RNA editing is providing a human antogonizing strategy against RNAi suppression by EBV.

As already mentioned before, ADARs are also able to influence cellular processes in an RNA independent manner. While a competitive binding between the RNAi machinery and ADARs to dsRNA substrates is very speculative[85], especially ADAR1p110 is reported to form heterodimers with Dicer proteins to promote their activity[36,86]. Therefore, neither the ADARs dsRBDs nor its deaminase domain is directly responsible for the increased activity of Dicers. A DEAD-box RNA helicase domain is auto-inhibiting the catalytic activity of Dicers, whereby ADAR1 is assumed to enhance their activity by binding to the helicase domain and preventing its inhibitory effect[36,87]. Mass spectrometry screenings identified also ADAR1 enrichment in H3K27me3 marked heterochromatin as well as EED and PRC1, which are involved in chromatin formation and gene expression[88–91]. Vigilin, a RNA binding protein, which is involved in heterochromatin and chromosomal segregation, is interacting with SUV39H1, a histone methyltransferase, and also binding to inosine containing RNAs, such as highly edited inverted *Alu* repeats[92–94]. ADAR1 is reported to form complexes with vigilin, KH68-70, ATP-dependent RNA helicase A (RHA), as well as heterochromatin protein 1 (HP1), which are all involved in heterochromatin formation and gene silencing[92,94]. Next, a competitive binding of ADAR1 and ADAR2 to RNA was also reported. While ADAR1 binding to inverted *Alu* repeats in the 3'-UTR of several transcripts was found to inhibit Staufen1 mediated decay[95], ADAR2 binding is competitive to RNA decay proteins such as PARN in an editing independent manner[96].

Therefore, the physiological role of ADARs and its impact to the regulatory mechanism of cellular processes is much more complex than the "simple" A-to-I conversion, which is translationally interpreted as guanosine. However, for a detailed understanding of ADARs and their role within the cellular machinery, further research is necessary and many interactions in an editing dependent or independent manner remain speculative.

1.3. Therapeutic approaches using DNA and RNA editing

In recent years, efforts for the elucidation of the human genome as well as novel sequencing techniques, such as next generation sequencing (NGS) highly increased the knowledge about human pathogenic disease variants and novel approaches for a targeted treatment of genetic disorders arose[97]. The majority of about 58 % of genetic diseases are in relation with a single nucleotide polymorphism (SNP, point mutation), and compared to other base pair changes, G-to-A point mutations are highly overrepresented (47 %) in humans[97]. Therefore, a targeted

exchange of bases within nucleic acids offers a tremendous therapeutic potential for a variety of genetic disorders.

1.3.1. Site-directed DNA editing

A promising but also very challenging possibility to treat genetic disorders and to manipulate genes is based on genome editing, which is mediated by clustered regularly interspaced short palindromic repeat-CRISPR-associated genes (CRISPR-Cas) systems[98]. The development of the CRISPR-Cas9 system revolutionized genome engineering techniques for gene manipulation including silencing, repair, and insertion or deletion[98]. The Cas9 endonuclease is part of the type II CRISPR-Cas immune system of bacteria, which is responsible for the protection against invading genetic elements such as plasmids or viruses[99,100]. Nucleases, such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), or Cas9 endonucleases introduce DNA double-strand breaks (DSBs), which are repaired by either non-homologous end joining (NHEJ) or homology directed repair (HDR)[101]. While NHEJ leads to the formation of indels (random insertions or deletions), which can cause a translational frameshift, HDR enables the insertion of single nucleotides or whole transgenes that originate from an exogenous DNA template. However, ZFNs or TALENs require protein engineering to alter the target site specificity, while Cas9 endonucleases are steered to the target site using a modular and exchangeable guideRNA (gRNA). This gRNAs contains a CRISPR RNA (crRNA) and *trans*-activating CRISPR RNA (tracrRNA) duplex, which is responsible for the target site direction by recognition of specific protospacer adjacent motifs (PAM)[102]. PAM directed Cas9 endonucleases locally denature the DNA duplex, while the gRNA is hybridizing with its complementary ssDNA and the unpaired single strand is forming a disordered loop (R loop)[102,103]. Furthermore, NHEJ and HDR are competitive processes, whereby NHEJ is more efficient than HDR[104,105]. While the high efficiency of NHEJ is providing a useful tool for gene disruptions, the use of HDRs is very challenging to introduce gene repair. In 2020, Emmanuelle Charpentier as well as Jennifer Doudna were awarded with the Nobel Prize in chemistry for the development of the CRISPR-Cas9 system and nowadays it is a common and powerful technique within laboratory as well as clinical research. Further developments of the system utilizing a mutated Cas9 nuclease (D10A) resulted in a Cas9 nickase, which prevents a DSB and facilitates nicking of the gRNA bound DNA strand[106]. Fusion of the Cas9 (D10A) nickase to one wild-type and catalytically inactive tRNA adenosine deaminase (TadA = ADAT) as well as one mutated, and on ssDNA catalytically active tRNA adenosine deaminase (TadA*), provided an A-to-I DNA base editor[107]. Natively in *E.Coli*, TadA acts as homodimers, whereby the

conjugation of two TadA units was necessary for sufficient editing (Figure 5b)[108]. The conjugation of a ssDNA selective APOBEC1 deaminase to the Cas9 (D10A) nickase provided additionally a C-to-U DNA base editor (Figure 5a)[106]. Subsequent DNA repair and replication resulted in an A-to-G or C-to-T conversion of the target base, respectively. However, severe off-target editing is reported for those DNA base editor systems, and the requirement of a PAM sequence in a defined distance to the target site is limiting the target versatility[107,109].

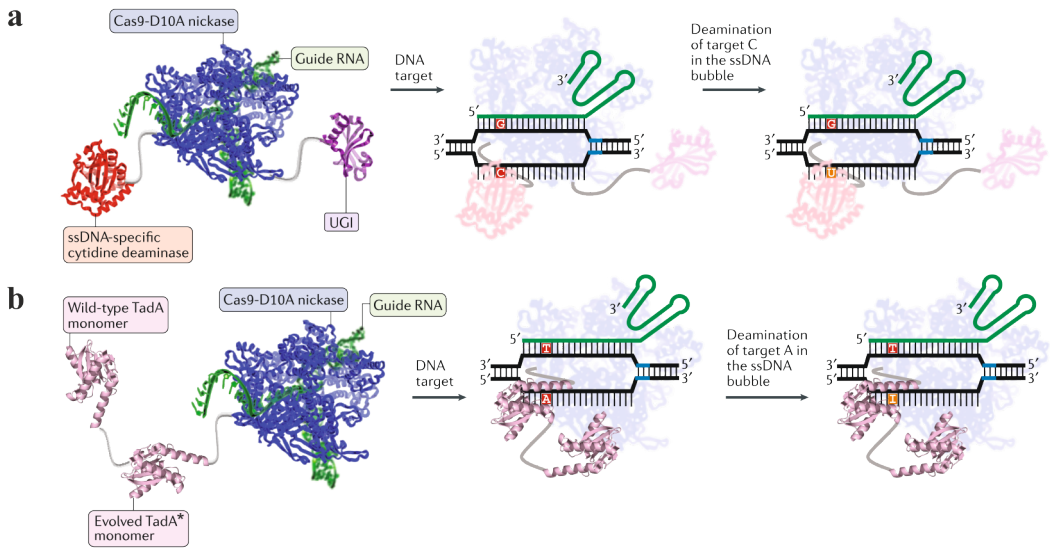


Figure 5: DNA base editor systems. *(a) APOBEC1-based mediated DNA base editing to introduce C-to-U substitutions. R loop (green) exposes a region of ssDNA to the cytosine deaminase domain. Uracil N-glycosylase (UGI) is necessary to inhibit U•G mismatch recognition and subsequent cleavage of the glyosidic bonds by uracil N-glycosylase (UNG). (b) TadA-based mediated DNA base editing to introduce A-to-I substitutions. R-loop exposes a region of ssDNA to the adenosine deaminase domain. Wild-type TadA was mutationally evolved (TadA*) to provide deaminase activity to ssDNA. Structures are adapted from ref.* 97.

In theory, DNA base editing is providing a promising possibility to cure genetic disorders, which are based on certain point mutations. However, to this point of time, a high off-target editing on the genome level is abolishing any therapeutic approach and further research is necessary. Besides DNA base substitution, editing on the RNA level is providing a more promising therapeutic approach. While many disease-relevant phenotypes based on amino acid codon alterations, as well as post-translational modifications such as phosphorylations (of e.g. serine, threonine, or tyrosine) or glycosylations (of e.g. asparagine), can be corrected using RNA editing, a therapy on the RNA level, which is reversible, flexible and therefore highly beneficial for any adverse and secondary effects.

1.3.2. Site-directed RNA editing

In 1995, Woolf and colleagues proposed to use the A-to-I editing activity of ADARs for site-directed editing[110]. The hybridization of a 52 nt oligonucleotide with a luciferase reporter

mRNA, prior to microinjection into *Xenopus* embryos, restored luciferase activity and indicated the direction of endogenous ADAR to catalyze RNA editing. Since then, different approaches to enable site-directed A-to-I editing on the RNA level have been developed. Nowadays, the use of three different approaches are persistent but still evolving: 1) The use of artificial and engineered ADAR deaminase domains, 2) the use of wild-type, but overexpressed ADARs, and 3) the use of endogenous ADARs[111]. While all of the site-directed RNA editing approaches differ in their use of engineered, wild-type or even endogenous ADARs, all of the systems require a certain RNA to direct the deaminases to the target sites. Those certain RNAs are complementary to the target mRNA, but differ in length and design for each system. Because of their guiding capacity, those complementary RNAs are so-called guideRNAs (gRNAs). While gRNAs, used for Cas9-based systems, are involved in target site recognition and stabilization of the unwound DNA/RNA duplex (Figure 5), gRNAs, used for site-directed RNA editing, are responsible for the formation of dsRNAs, which are recognized by dsRBDs or deaminase domains. Additionally, the gRNAs of all of the mentioned systems described below contain an A·C mismatch at the target site, which is not mentioned within each description.

1.3.2.1. The SNAP®-ADAR system

The SNAP®-ADAR system is the first published approach using an artificial editase and to this point of time the best characterized editing system (Figure 6a)[112,113]. This artificial editase is a combinatorial protein comprising a wild-type or hyperactive E/Q mutant deaminase domain of ADAR1 or ADAR2 and a *N*-terminal SNAP-tag® protein. The SNAP-tag® is a mutationally evolved self-labeling protein, originated from *human* O^6-alkylguanine DNA alkyl transferases (*h*AGT)[114,115]. While *h*AGTs are responsible for O^6-alkylguanine repair in DNA, SNAP-tags® are engineered to bind to O^6-benzylguanine (BG) in a highly specific manner. Preliminary conjugation of a BG moiety to a NH_2-terminal and approximately 22 nt chemically stabilized gRNA, facilitates covalent binding with SNAP®-ADAR fusion proteins in a 1:1 stochiometric ratio *in vitro* and *in vivo*[112,116,117]. Transfection of these gRNAs into SNAP®-ADAR stable expressing HEK 293 cells enables editing of several endogenous targets such as GAPDH, GUSB, KRAS, and STAT1 within coding (ORF) and non-coding (3'-UTR and 5'-UTR) regions and editing efficiencies up to 90 %[118]. While 11 out of the 16 5'-NAN amino acid codons were editable with an efficiency of > 50 % using both wild-type and E/Q variants, hyperactive E/Q mutants showed higher off-target events than the wild-type deaminase domains. Applying chemical modified gRNAs, including 2'-O-methylation (2'-OMe), 2'-deoxy-2'-fluoro (2'-F) and phoshorothioate (PS) linkages, was beneficial concerning several

aspects. Chemical modified gRNAs showed a higher stability and resistance towards nucleases such as ribonucleases (RNases), and while the editing level of the target site was further increased the number of off-target events was decreased. Furthermore, the co-transfection of gRNAs, targeting different endogenous transcripts, showed sufficient editing levels without a loss of efficiency. Bisfunctional gRNAs were also reported recently, which are able to recruit two similar SNAP®-ADAR fusion proteins at once or to co-recruit either ADAR1/ADAR2 or ADAR1/APOBEC1 heteromeric fusion proteins in an orthogonal and concurrent manner[119]. Besides so-called multiplexing, 6-nitropiperonyloxy-methyl (Npom) protected BG gRNAs enabled also light inducible editing of eGFP constructs in *Platynereis dumerilii*[117] as well as the alteration of localization signals of membrane proteins[120]. However, chemically modified and BG conjugated gRNAs can not be genetically encoded, which is limiting the choice of delivery systems such as adeno-associated viruses (AAVs), and the use of artificial but not endogenous ADARs is a major hurdle for any therapeutic application.

1.3.2.2. The λN- and MS2-ADAR systems

Following the publication of the SNAP®-ADAR system, the λN editing system was reported (Figure 6b)[121]. An ADAR2 deaminase domain, similar to the SNAP®-ADAR system, is fused to a 22 nt λN peptide, derived from *Escherichia virus lambda*, which is naturally binding to short RNA hairpin structures so-called BoxB motifs[122]. Preliminary experiments, co-injecting the λN fusion protein, a gRNA (containing one 17 nt BoxB hairpin), and the transcript mRNA demonstrated the restoration of a premature stop codon of a cystic fibrosis transmembrane conductance regulator (CFTR) in *Xenopus* oocytes[121]. In HEK 293 cells, applying four λN peptides, fused to a hyperactive E/Q mutant of ADAR2, and a gRNA containing two BoxB motifs increased the correction of a W58X GFP reporter from 20-70 %[123]. However, high off-target events were detected within both approaches[121,123]. Furthermore, a virus based delivery into primary murine neurons, using AAVs as cargos, was employed to correct an endogenous methyl CpG binding protein 2 (MECP2) transcript, which is in relation to the Rett syndrome[124]. Joint transduction of the fusion protein containing four λN peptides, the hyperactive E/Q mutant of ADAR2, and six copies of the two BoxB motif containing gRNA restored 72 % of the mutant mRNA and resulted in a functional protein repair.

Another system is reported to use hairpin motifs for the target site direction, similar to the λN system. This approach utilizes a fusion protein combining an ADAR1 deaminase domain with a MS2 bacteriophage coat protein (MCP) derived from *Escherichia virus MS2*, which is naturally binding to MS2 stem loops (Figure 6c)[125]. Co-transfection and therefore

overexpression in HEK 293 cells of the fusion protein, an eGFP reporter, and the gRNA containing a 21 nt complementary part as well as six 5'-MS2 stem loops restored 5 % of a 5'-UAG premature stop codon of the eGFP reporter transcript.

1.3.2.3. The CRISPR-Cas13b-ADAR system

Furthermore, an RNA editing system based on the CRISPR-Cas system was reported as well. In bacteria a ribonuclease activity was discovered for Type IV CRISPR-Cas13a-d nucleases, whereby no PAM sequence is necessary for the target site recognition (Figure 6d)[126–135]. The mutation of two higher eukaryotes and prokaryotes nucleotide binding (HEPN) RNase domains of Cas13b from *Prevotella sp. P5-125* provided a catalytically inactive, but still programmable RNA binding nuclease (dCas13b)[127,134–136]. The catalytically inactive dCas13b nuclease was fused to hyperactive E/Q deaminase domains of ADAR1 or ADAR2, and co-transfection and thus overexpression of the artificial dCas13b-ADAR editase, gRNA, and target transcripts enabled site-directed RNA editing. The ~ 85 nt gRNA contains typically a 50 nt 5'-terminal complementary region to the target sequence and a 35 nt hairpin sequence at the 3'-end, which is supposed to recruit dCas13b binding. Applying the dCas13b-ADAR2 E488Q editase, typical editing levels between 12-35 % were observed for all 16 possible 5'-NAN amino acid codons using a luciferase reporter from *cypridina* (*C*Luc). Furthermore, 34 disease relevant targets (mostly preferred 5'-UAG codons) showed comparable editing levels between 5-30 %. For 5'-UAG sites of two endogenous transcripts, KRAS and PPIB, editing efficiencies of 15-45 % were obtained as well. However, massive off-target events were observed, which were attributed to the use of hyperactive E/Q variants. A further mutational design provided a ADAR2 E488Q/T375G double mutant, which showed much lower off-target editing on the endogenous level (typically < 20 %), but also a halved on-target editing level on an *C*Luc reporter. However, the dCas13b-ADAR2 E488Q fusion protein is exceeding the packing capability of AAVs, which is limiting the choice of delivery. Furthermore, originating from bacteria, dCas13 may provoke adverse immune responses, similar to λN or MS2 proteins, and missing toxicity assays for all mentioned systems serves a major hurdle regarding any clinical settings.

1.3.2.4. The CIRTS-ADAR system

To overcome any adverse effects, such as immune responses based on bacterial origins, a CRISPR-Cas-inspired RNA targeting system (CIRTS) was developed, with its major objective to provide a completely genetically encodable system utilizing only *human* proteins (Figure

6e)[137]. Therefore, the *h*ADAR2 deaminase domain or its hyperactive E/Q variant was fused to a TAR hairpin binding protein (TBP), and the ssRNA binding protein β-defensin 3. The gRNA contains a 5'-terminal 31 nt TAR hairpin motif and a complementary guide sequence to bind the target mRNA transcript. While the TBP is responsible for hairpin binding and target site direction, β-defensin 3 is necessary to protect unbound single-stranded gRNA from degradation. A premature 5'-UAG stop codon of a dual luciferase reporter transcript was corrected with an efficiency of 15 % or 50 % using the wild-type *h*ADAR2 deaminase domain or its hyperactive E/Q variant, respectively.

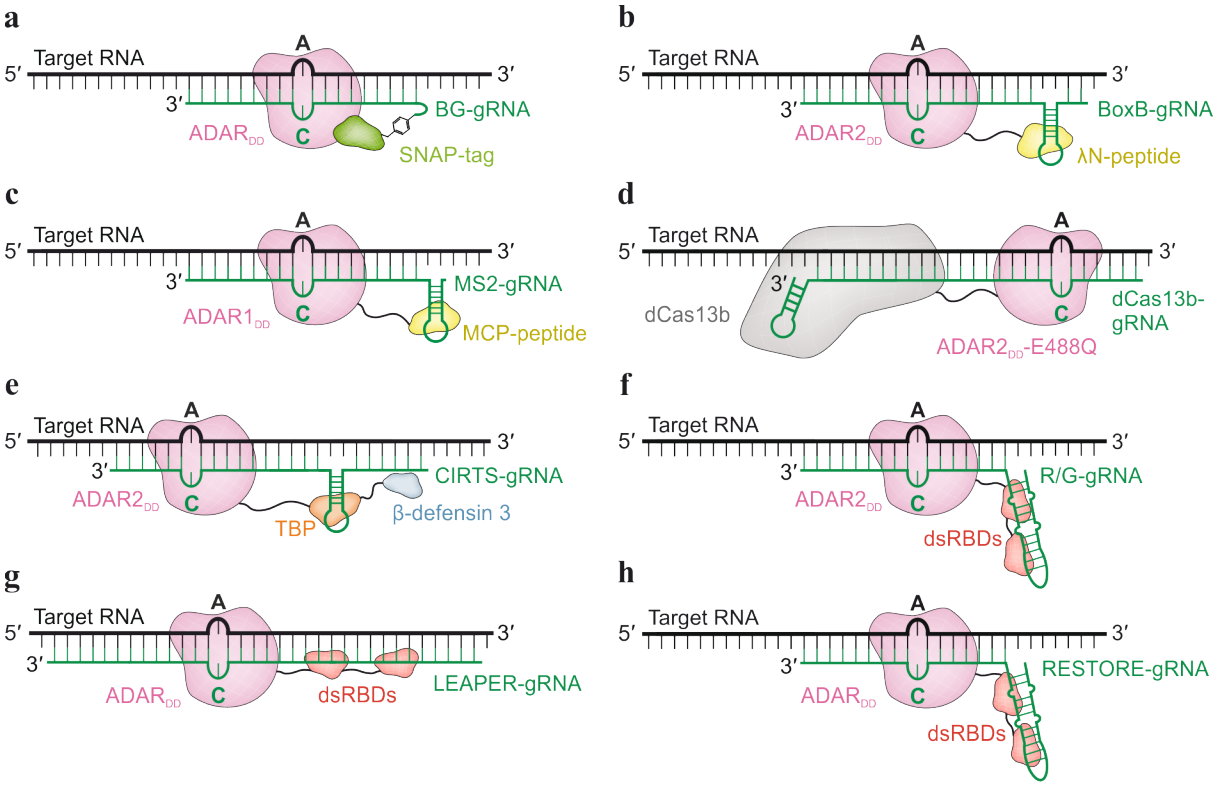


Figure 6: Different site-directed RNA editing approaches to introduce A-to-I substitutions. *(a) SNAP®-ADAR mediated A-to-I editing using artificial SNAP®-ADAR constructs and BG gRNAs. (b) The λN-ADAR system using λN peptides and BoxB gRNAs. (c) The MS2-ADAR approach using MCP-MS2 peptides and MS2 stem loop gRNAs. (d) The CRISPR-dCas13b editing system using dCas13b hairpin gRNAs. (e) The CIRTS-approach using exclusively encodable gRNAs and proteins from human origin. (f) The R/G motif inspired editing system using wild-type hADAR2 and genetically encodable gRNAs. (g) The LEAPER system using chemically modified or unmodified gRNAs to recruit endogenous ADARs. (h) RESTORE system using R/G motif inspired, chemically modified gRNAs to recruit endogenous ADARs. ADAR deaminase domains (ADAR$_{DD}$) are illustrated in light pink, gRNAs in dark green, SNAP-tag in light green, λN and MCP peptides in yellow, dCas13b in grey, TBP in orange, β-defensin 3 in light blue and dsRBDs in red.*

1.3.2.5. The R/G system

A further attempt to overcome any adverse effects, based on using artificial editases or enzymes originating from bacteria, was the approach to harness wild-type *h*ADAR2[138]. Therefore, the gRNA design is inspired by mimicking the naturally occurring imperfect hairpin, which is

inducing the R/G substitution within the GLUR2 mRNA transcript. This hairpin motif is providing a strongly recognized substrate of the dsRBDs of ADAR2. The termed R/G gRNAs contain a 16-29 nt part complementary to the target transcript and a 45 nt imperfect hairpin motif. Co-transfection of plasmids encoding gRNAs and reporter transcripts into *h*ADAR2 stable expressing HEK 293 cells enables editing of a premature 5'-UAG stop codon of an eGFP reporter with an efficiency of 65 %. However, ectopic overexpression of *h*ADAR2 and R/G gRNAs, targeting 5'-UAG codons of endogenous transcripts, such as GAPDH, GUSB, and RAB7A resulted in lower editing levels up to 38 %. As a promising therapeutic approach, a premature 5'-UAG codon within PINK1 (a Parkinson disease related transcript[139]) was corrected with an efficiency of 10 %, overexpressing all components in Hela cells. While low editing yields of only 10 % were observed, the PINK/Parkin mediated mitophagy phenotype was rescued in 85 % of the cells, expressing all components[138]. Further gRNA designs showed that it was also possible to harness *h*ADAR1p110 as well as p150 utilizing R/G gRNAs[140].

1.3.2.6. Recruiting endogenous ADARs

One step further to any therapeutic application is the recruitment of endogenous ADARs. Up to this point there are two promising editing systems reported to recruit endogenous ADARs: The RESTORE system (Recruiting Endogenous ADAR to Specific Transcripts for Oligonucleotide mediated RNA editing)[141] and the LEAPER system (Leveraging Endogenous ADAR for Programmable Editing of RNA)[142].

The RESTORE system uses, same as the R/G system, gRNAs which are mimicking the naturally occurring hairpin within the GLUR2 mRNA transcript, which is inducing the R/G substitution[141]. While R/G gRNAs are completely genetically encodable, RESTORE gRNAs are highly chemically modified to prevent ribonuclease degradation. They also contain a 20-40 nt region complementary to the target sequence and a characteristic hairpin motif to enable dsRBD binding. While editing levels of 4-34 % within the 3'-UTR and ORF of GAPDH were observed for a wide range of immortalized cell lines, editing was increased for all cell lines to 11-74 % under interferon α induction. The same pattern was observable for primary cell lines, whereby the editing levels were increased from 10-63 % to 35-77 % after interferon α treatment. The transfection of RESTORE gRNAs enables a correction of 10-20 % of the PiZZ mutation (E342K) of *SERPINA1* transcripts – an α-1-antitrypsin deficiency[141,143] – within Hela cells, and within primary fibroblasts and RPE cells editing levels of 7-21 % are observed for tyrosine 701 of *STAT1* transcripts.

In contrast, the LEAPER system is using a different approach regarding the gRNA design. LEAPER gRNAs are typically 71-191 nt long complementary RNAs containing additional G-A mismatches at non-targeted adenosines to reduce off-target editing[142]. Ectopic expression of a 151 nt gRNA from a plasmid within HEK 293 cells enables editing levels up to 50 % within the 5'-UTR as well as 20 % within the ORF of endogenous transcripts. In primary cells, editing levels of 30-80 % were observed within the 5'-UTR of *PPIB* transcripts. Furthermore, electroporation of chemically modified, 111 nt long gRNAs into fibroblasts, obtained from patients with the Hurler syndrome, restored 30 % of the deficient IDUA enzyme targeting the pre-mRNA. The Hurler syndrome is a severe form of mucopolysaccharidosis type 1, originating from a W402X point mutation.

As demonstrated, recruiting of endogenous ADARs, and thereby overcoming the need of artificial deaminases, serves a promising approach for therapeutic applications. However, the use of chemical modified, and therefore genetically not encodable gRNAs is still limiting the choice of delivery, which causes a major hurdle for any therapeutic application.

1.4. Antisense- and siRNA-based therapeutic drug systems

While RNA editing systems provide a promising approach for future therapeutic applications, other RNA targeted drug systems are one step ahead. Especially antisense oligonucleotide (ASO)- and short-interfering RNA (siRNA)-based drug systems are very promising state-of-the-art RNA targeted drugs. Several ASO- or siRNA-based drugs are currently participating clinical trials or are already FDA approved drugs. Famous examples are Fomivirsen, the first approved ASO-based drug and a *human cytomegalovirus (HCMV)* targeting phosphorothioate (PS) ASO, Eteplirsen, a phosphorodiamidate morpholino oligomer (PMO) which induces a splice site alteration of exon 51 within patients with Duchenne muscular dystrophy (DMD) or Patisiran, a siRNA-based drug targeting hereditary transthyretin mediated amyloidosis (ATTR)[144]. ASO- or siRNA-based drug systems are able to regulate protein levels by altering mRNA levels or the translation of proteins, gene activation or silencing, splice site alterations, transcript degradation or antigen synthesis[144–146]. These possibilities provide only a few cellular processes, in which RNA targeted drug systems can interfere.

1.4.1. RNAi and RISC

RNA interference (RNAi) is an endogenous mechanism to induce translational suppression or repression within infections or genetic abnormalities[147] using siRNAs, but is also involved in cancer development[148–150], infectious diseases[151,152], immunity[153], cell-cycle progression[154], and

metabolism[155] using miRNAs. Both, siRNAs (19-22 nt) and miRNAs (21-22 nt) are short dsRNAs processed from longer dsRNAs by Drosha and Dicer, and provide prodrug-like molecules consisting of a complementary duplex of sense (passenger) and antisense (guide) strands[156–159]. Matured siRNA as well as miRNA duplexes are loaded to Ago2, a cytoplasmic RNA endonuclease which is highly selective towards RNA-RNA cleavage (Figure 7b and c)[160]. Subsequent degradation of the passenger strand enables the pharmacological activity of the antisense strand, which is responsible for guiding the RNA-induced silencing complex (RISC) to its complementary and target RNA[156,159]. While siRNA loaded RISC complexes are responsible for target mRNA degradation and translational repression[159,161–163], miRNA loaded RISC complexes modulate gene expression of the target transcripts via binding to the 3'-UTR in a slicer independent pathway, which is highly dependent on the miRNA's complementary degree (Figure 7d)[147,157,158]. Furthermore, miRNAs are able to regulate epigenetic modifications, such as methylations of CpG islands within the promoter region of different genes[164–167]. However, only a seven nucleotide seed region within siRNAs and miRNAs is responsible for the target site discrimination of RISC complexes, whereby off-target effects were observed in cell culture as well as animal models[168,169]. But the ongoing research concerning chemical modifications and thereto related stabilities, protein interactions, or distributions of siRNAs, is providing promising results regarding off-target reduction, specificity and activity[146]. Additionally, due to the cytoplasmatic localization of Ago2 and RISC complexes, Ago2 bound siRNA or miRNA modulations are restricted to cytoplasmatic RNAs[159,170]. Hence, mimicking the natural structure of processed siRNAs or miRNAs by introducing chemical modified RNA duplexes to cells or tissues is providing a platform with high potential for RNA targeted drugs to induce gene silencing modulations or several post-transcriptional modifications.

1.4.2. ASOs and RNase H1

Besides RNAi induced modulations of cellular processes, ASOs are also able to influence the fate of an RNA within multiple pathways. Generally, ASOs are 16-22 nt long chemically modified single-stranded oligonucleotides, antisense to the target transcript. While occupancy-only pathways are able to introduce splice site alterations, increase the translation of target proteins or cause a translational arrest, occupancy-dependent pathways are highly selective for RNase H1 mediated degradation of RNAs (Figure 7a and d)[144]. Occupancy-only pathways that alter translations in both directions (upregulations and downregulations), are based on competitive binding of ASOs to either sequence or structural elements, such as 5'-UTR located

upstream ORFs (uORFs) or translation inhibitory elements (TIEs), which are responsible for mRNA suppression[171,172]. Additionally, binding to heterogeneous nuclear ribonucleoproteins (hnRNPs), a protein family which is involved in several post-translational modifications, such as splicing, also induces splice site alterations. For example, in patients with spinal muscular atrophy (SMA), exon 7 is excluded during splicing due to a genetic mutation. Nusinersen, a 18 nt fully 2'-methoxyethyl (2'-MOE) modified PS ASO, binds competitively to a splicing site of splicing receptor hnRNP A1/A2 and results in an efficient integration of exon 7 during SNM2 transcript maturation[173].

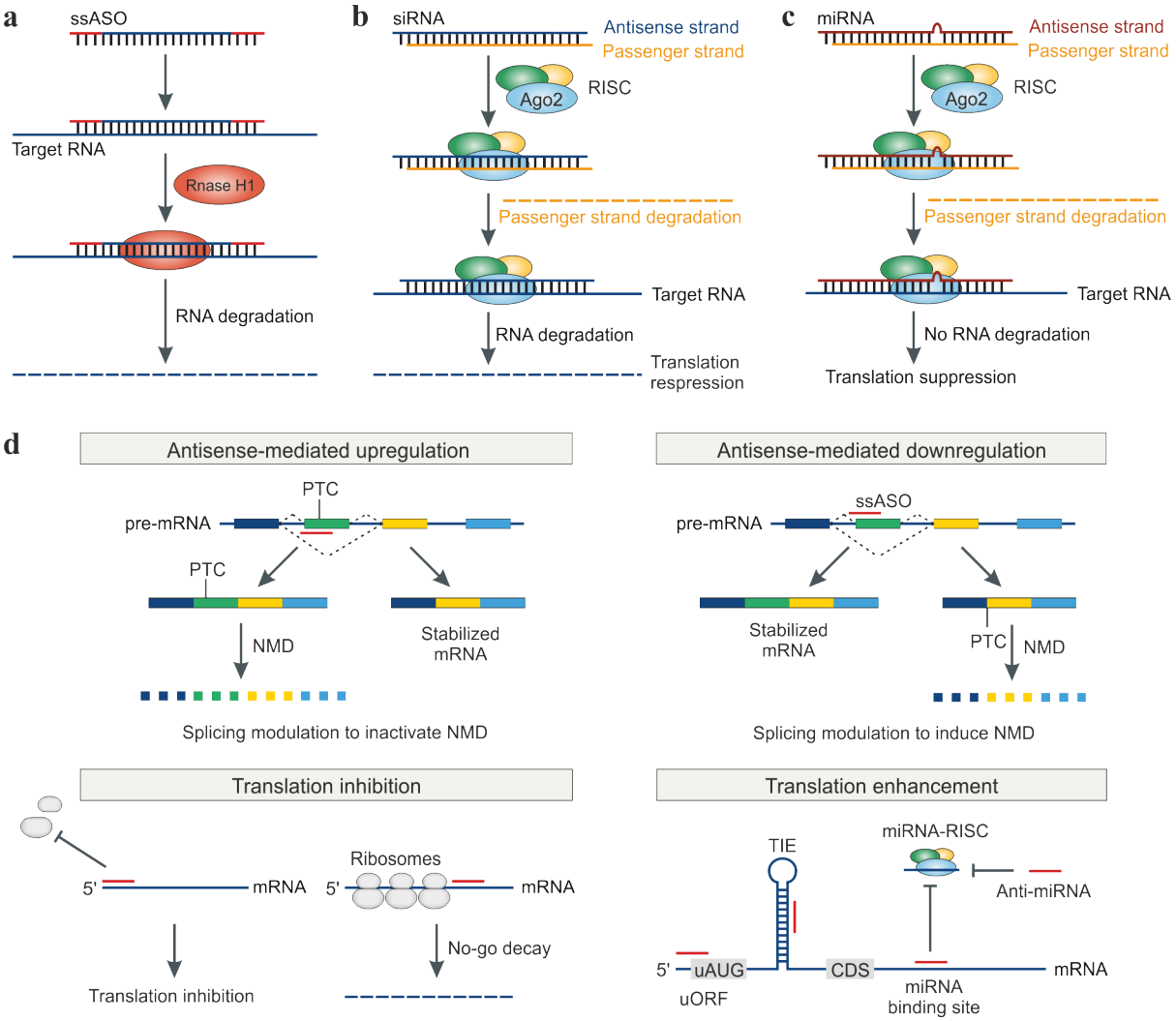


Figure 7: Impacts of ssASO, siRNA or miRNA to RNA stability and translation modulation. *(a) ssASO mediated RNase H1 recruitment and RNA degradation. (b) siRNA mediated RISC recruitment and RNA degradation. (c) miRNA mediated RISC recruitment and translation suppression. (d) ssASO (red) mediated occupancy-only mechanisms to modulate gene expression by splice site alterations, translation inhibition or translation enhancement. CDS = coding region sequence, NMD = non-sense mediated decay, PTC = premature termination codon, RISC = RNAi-induced silencing complex, uORF = upstream open reading frame, TIE = translation inhibitory element. Structures are adapted from ref.* 144.

In contrast to occupancy-only pathways, ASOs are also able to induce RNase H1 mediated RNA degradation. RNase H1 is an ubiquitously expressed double-stranded endonuclease that cleaves RNA only in RNA/DNA hybrid duplexes[174,175]. The catalytic site of RNase H1 is highly sensitive to the sequence and the helical geometry of the heteroduplex[174], and due to its localization within the nucleus, cytoplasm and mitochondria, RNase H1 is able to alter the fate of several precursor as well as matured RNAs. This includes pre-mRNA, mRNA[176,177], pre-rRNA, rRNA[178], tRNA, lncRNA[179], snRNA, snoRNA[180], antisense transcripts[181], as well as toxic RNAs[182]. While RNase H1 is able to cleave RNA only in DNA/RNA hybrids, most ASOs which are thought to recruit RNase H1 also contain deoxyribonucleotides. So-called "gapmers" are ASOs which contain a deoxyribonucleotide gap in the center and 2'-chemically modified ribonucleotides at both ends[144]. In contrast to Ago2 mediated degradations, the entire sequence information of bound ASOs is recognized by RNase H1, which is limiting the propensity of incorrect hybridization and off-target cleavage[174]. Therefore, ASO-based RNA targeted drugs provide a very promising approach for therapeutic applications targeting a variety of mechanisms. While ASOs can affect all amenable RNAs, ASOs are also able to interact with the translation regulatory mechanism or stimulate the innate immune system. Nowadays, eight single-stranded ASO-based therapeutics are commercially available and more than 50 are currently enrolled in clinical trials[144].

1.5. Chemical modifications of oligonucleotides

Since RNA targeted drug systems are emerging aspects of research and furthermore of therapeutic applications, the field of chemical modifications within (oligo-) nucleotides became more important as well. Chemical modifications within oligonucleotides, especially siRNAs, ASOs and gRNAs, are highly beneficial concerning stability and nuclease resistance, oligonucleotide-protein interactions and therefore cellular uptake and distribution (cellular and systemic), as well as immunogenicity and toxicological aspects of off-target events[118,144,146,183–185]. Within oligonucleotides, there are three major possibilities to introduce chemical modifications: phosphate, ribose, and nucleobase modifications. Due to an altered Watson-Crick base pairing capability, nucleobase modifications are less important for oligonucleotides and only phosphate and ribose modifications will be part of the following discussion. Furthermore, 3'- or 5'-conjugation-based modifications were found to impact the cellular uptake and systemic distribution of oligonucleotides.

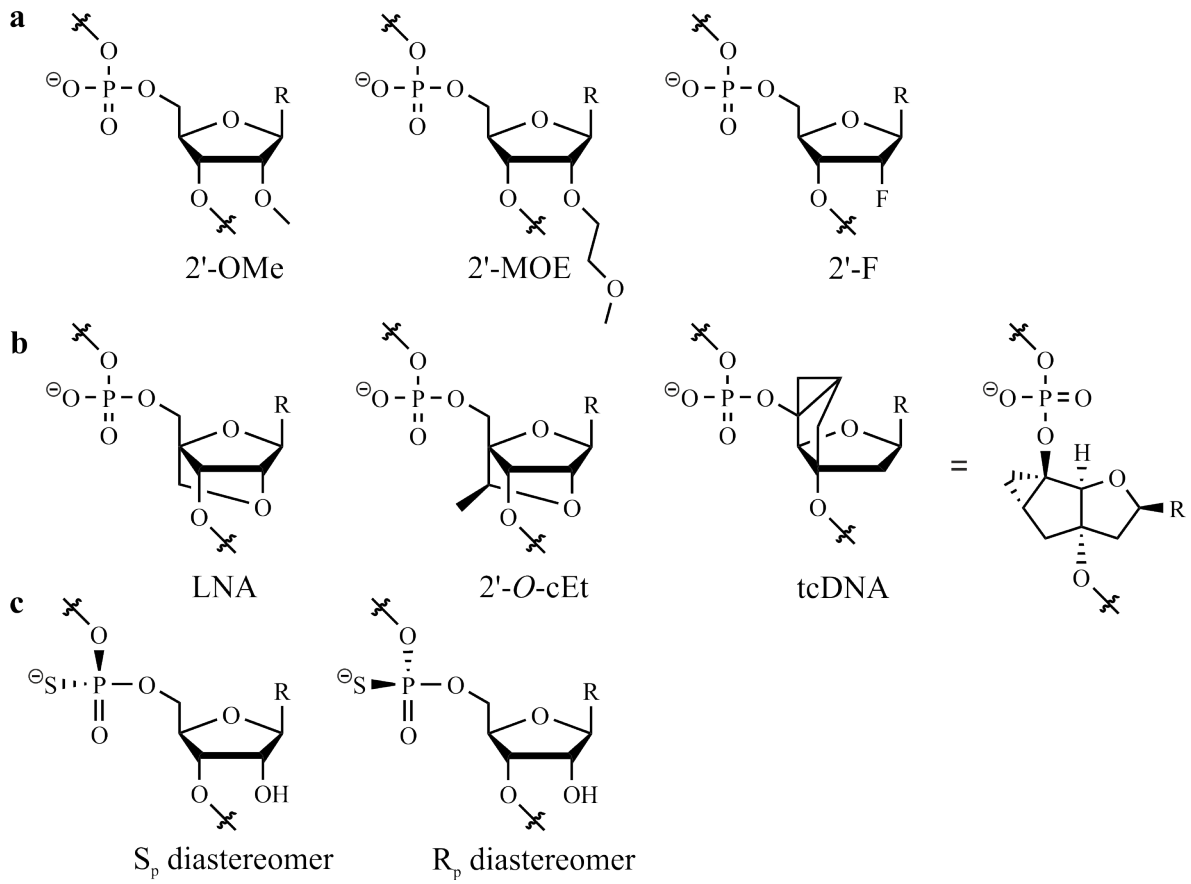


Figure 8: Chemical modifications of nucleotides. *(a) Different 2'-ribose modifications. 2'-OMe = 2'-O-methyl, 2'-MOE = 2'-Methoxyethyl, 2'-F = 2'-deoxy-fluoro. (b) Bridged modifications of nucleotides. LNA = Locked nucleic acids, 2'-O-cEt = 2'-S-constrained ethyl, tcDNA = tricyclo-DNA. (c) Phosphorothioate modifications of nucleotides. S_p and R_p diastereomers are shown for RNA. Residue R = Nucleobase.*

1.5.1. Ribose modifications of oligonucleotides

While several thousands of 2'-ribose modifications are nowadays evaluated, comprising different architectural approaches and features, "simple" 2'-*O*-metyhl (2'-OMe) and 2'-methoxyethyl (2'-MOE) substitutions are the most abundant and best characterized modifications (Figure 8a)[144,146,186]. Furthermore, 2'-deoxy-fluoro (2'-F), bicyclic riboses such as locked nucleic acids (LNAs) or *S*-constrained ethyl bridged nucleic acids (2'-cEt BNAs), or even tricyclic derivatives such as tricyclo-DNA (tcDNA), provide only a few exemplary ribose modifications, whereby each modification comprises different physico-chemical and biological properties (Figure 8a and b)[185]. A major benefit of 2'-ribose modifications is the increased resistance of oligonucleotides against ribonucleases. Furthermore, several 2'-modifications showed an increased binding affinity to RNA or DNA and therefore a higher potency, as well as a reduced immunogenicity compared to unmodified oligonucleotides. However, especially bridged modifications like LNAs or cEts showed also adverse effects, such as cytotoxicity in multiple cell lines, hepatotoxicity in mice, rats and non-human primates (NHPs), as well as

hepatotoxicity, thrombocytopenia, and nephrotoxicity in humans[187,188]. These adverse effects are attributed to the significantly increased binding affinity and therefore to an off-target hybridization and RNase H1 cleavage.

1.5.2. Phosphate modifications of oligonucleotides

Another milestone within the development of oligonucleotide-based therapeutics was the discovery of phosphorothioate (PS) modifications (Figure 8c)[184]. Within phosphorothioates, a non-bridging oxygen of a phosphate (PO) linkage is exchanged with a sulfur, generating a chiral center at every PS linkage[189,190]. Creating a chiral center causes the formation of two diastereomers (S_p and R_p), and different physic-ochemical and biological properties are observed for the use of stereopure as well as stereoimpure PS ASOs. While stereopure S_p PS ASOs showed a higher nuclease resistance, stereopure R_p PS ASOs showed an increased binding affinity to the cognate RNA[191]. However, chirally impure PS ASOs exhibit a lower binding affinity to RNA as chirally pure ones but also a higher nuclease resistance compared to natural DNA. Dependent on the desired target and properties, both stereopure and stereoimpure configurations are used for therapeutic applications of PS ASOs[144]. While siRNAs typically contain one or two PS linkages at the 5'- as well as one to three at its 3'-end[146], a general rule is not reported for ASOs, but a PS content of more than 50 % in not unusual[192,193]. Furthermore, a higher binding affinity to proteins is observed for PS ASOs. In accordance with the concept of "hard and soft acids and bases" (HSAB theory), sulfur is more polarizable than oxygen. Therefore, a higher charge distribution within the sulfur causes a more lipophilic character of the PS, which results in a higher protein binding affinity of PS ASOs[192–195]. In fact, utilizing NanoBRET™ and BioID analysis, many cellular and plasma proteins were identified to interact with PS ASOs. By this, 90 % of a subcutaneous (SQ) or intramuscular (IM) administered therapeutic dose of a PS 2'-MOE ASO is bound to plasma proteins, with albumin as its major binding partner[196,197]. Besides an increased stability and RNA binding affinity of PS ASOs, the capability of protein binding was also highly beneficial regarding pharmacokinetics, circulation and biodistribution as well as cellular uptake. Thus, plasma protein bound PS ASOs can be distributed to several peripheral tissues, whereby the liver, kidneys, or spleen are accumulating the majority of the administered dose. While protein bound PS ASOs were distributed systemically, applying subcutaneous (SQ)[170,196,198–205], intramuscular (IM)[196,200–205], intravenous (IV)[200,201], or even oral[205] administrations, the targeted delivery to desired organs or tissues remains challenging without using local administrations such as intravitreal (IVT)[206], intradermal (ID)[207], intrathecal (IT)[200,208–210], or rectal[211–213]. Furthermore, binding of PS ASOs

to membrane proteins is inducing a cellular uptake using different entries. Nowadays, different pathways, such as several clathrin- or caveolin-dependent endocytoses as well as macropintocytosis are reported to internalize PS ASOs in a productive or non-productive manner[184]. However, less than 0.1 % of the internalized PS ASO is released from the late endosomes and is intracellularly available within the cytosol or the nucleus (Figure 9c)[184,200,214,215]. While transfected or electroporated PS ASOs accumulate in PS bodies, paraspeckles or paraspeckle-like structures within the nucleus[216–218], no accumulation of PS ASOs was observed utilizing productive or non-productive uptake pathways[181]. However, a nuclear specific activity is reported for PS ASOs internalized via free uptake, but detailed mechanisms of endosomal release and re-localization between the nucleus and cytoplasm remain unknown[184].

1.6. The asialoglycoprotein receptor and a targeted delivery

1.6.1. The asialoglycoprotein receptor

Decades ago, the asialoglycoprotein receptor (ASGPR) was first discovered by Ashwell and Morell during studies on the metabolism of plasma glycoproteins[219]. The ASGPR is a mammalian Ca^{2+}-dependent C-type lectin and is predominantly expressed on the basolateral surface of liver parenchymal cells with an abundance of $1\text{-}5 \cdot 10^5$ copies per cell[220–222]. It is primary responsible for maintaining the serum glycoprotein homeostasis, which is regulated by the internalization of desialylated glycoproteins into hepatocytes[222]. Desialylated glycoproteins, carrying several terminal carbohydrate residues, such as glucoses, galactoses or *N*-acetyl galactosamines, are internalized by endocytosis via clathrin-coated pits and are degraded within the lysosomes[222]. During endosomal maturing, an acidification of the endosomes interior causes a glycoprotein-receptor dissociation, whereby the receptor is constitutively recycled to the basolateral surface of the cell and the glycoproteins are further processed to the late endosomes and lysosomes[223–226]. For receptor recycling, a typical turnover of about 20 minutes is reported.

In general, mammalian ASGPRs are composed of two different homologous type-II single-spanning membrane proteins, a 46 kDa major subunit H1 and a 50 kDa minor subunit H2[227,228]. Membrane bound receptors are functional heterooligomers bearing varying combinations of different receptor subunits[229–237]. However, the most abundant configuration is a trimeric combination containing two H1 and one H2 subunits (Figure 9b). While both subunits are encoded by distinct genes, a sequence identity of 58 % is observed and an 18 amino acid (aa) insert within the cytoplasmic domain of H2 is the most relevant difference[238]. Besides a ~ 40 aa *N*-terminal cytoplasmic domain, H1 and H2 have further domains in common. In detail, both

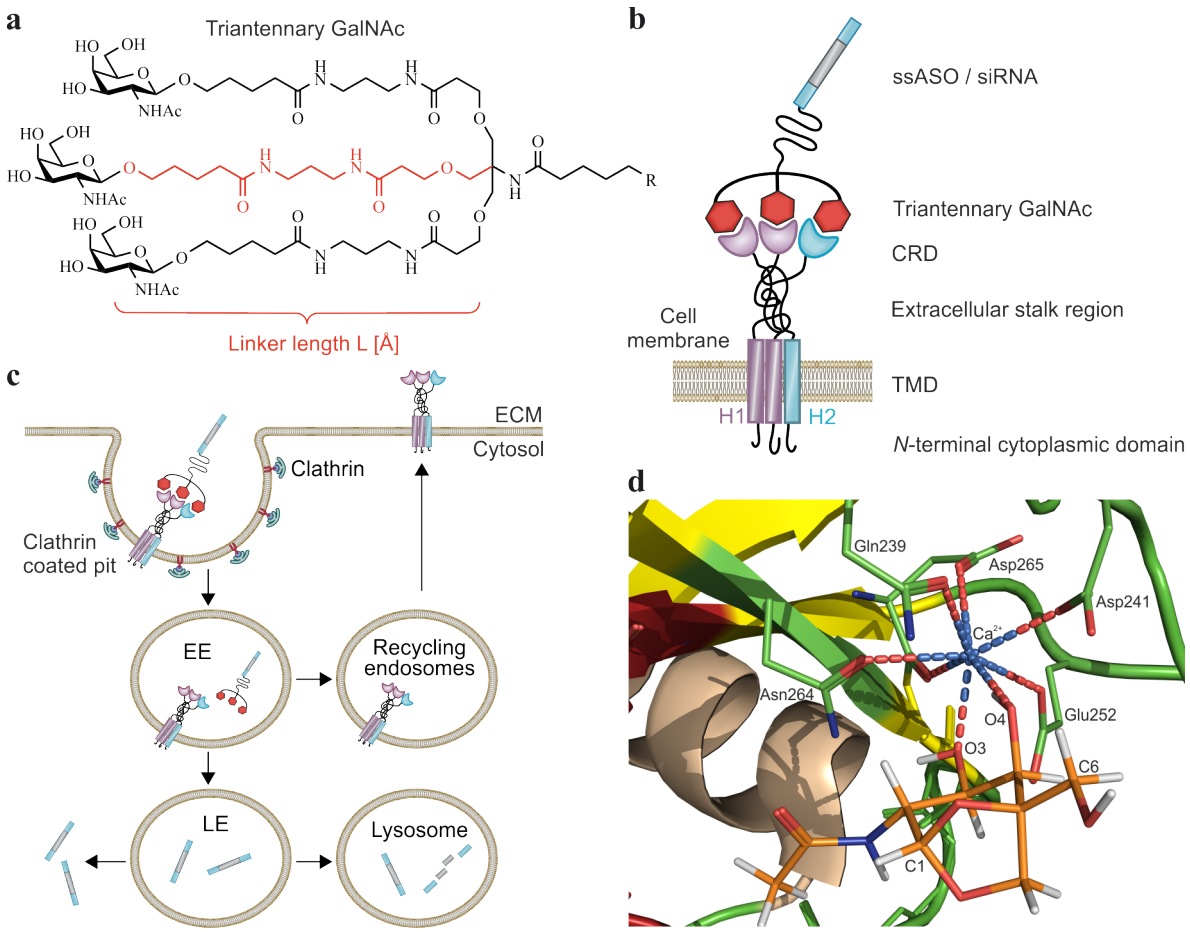


Figure 9: Triantennary GalNAc and ASGPR mediated uptake. *(a) Chemical structure of triantennary GalNAc. (b) Illustration of membrane bound ASGPR. CRD = carbohydrate recognition domain, TMD = transmembrane domain. (c) Illustration of ASGPR mediated endocytosis with endosomal maturation, receptor recycling and lysosomal degradation of oligonucleotides. ECM = extracellular matrix, EE = early endosome, LE = late endosome. Structures are adapted from ref. 184. (d) Crystal structure of a bicyclic N-acetyl galactosamine (orange) bound to the CRD. Galactosamine-receptor binding is Ca^{2+}-dependent and interacting aa residues are assigned. The crystal structure (PDB: 5JQ1) was published in ref. 245.*

subunits exhibit a ~ 140 aa carbohydrate recognition domain (CRD), a ~ 80 aa extracellular stalk region, and a ~ 20 aa single-pass transmembrane domain (TMD) (Figure 9b)[239,240]. Different isoforms are observed for each subunit as well[241,242]. For subunit H1, two different isoforms (H1a and H1b) are reported, whereas subunit H2 consists of three different isoforms (H2a, H2b, and H2c). While H1a is translated from the full-length mRNA, the transcript of isoform H1b is lacking the TMD, and is therefore assumed to serve as a soluble and secreted fraction of the ASGPR (sASGPR). sASGPRs are thought to bind to glycoproteins within the blood to prevent their interaction with other cells and tissues[241]. The transcript of H2a comprises also the full-length transcript, including a five aa insert between the TMD and the ectodomain. This five aa insert acts as a cleavage signal, which is responsible for proteolysis and the secretion of the H2a ectodomain[243]. While H2b is lacking the five aa insert, it is not proteolytically active and oligomerizes with H1a to form functional ASGPRs. As H1b, H2c is

lacking a 19 aa insert within the TMD as well as the five aa insert. However, the function of H2c remains unknown[244].

1.6.2. Triantennary *N*-acetyl galactosamine

Since the discovery of the ASGPR, substantial efforts have been made to investigate the natural binding motif of its CRD. Subsequent to the publication of a synthetic glycopeptide ligand containing three terminal *N*-acetyl galactosamine moieties, the field of the ASGPR mediated uptake gained increased attention[246]. Researchers developed different synthetic ligands mimicking the natural binding partner of the CRD, and nowadays, several multivalent ligands are well-characterized[245,247,248]. Utilizing different techniques such as surface plasmon resonance spectroscopy, fluorescence microscopy, flow cytometry, or X-ray crystallography, impacts of carbohydrate configurations and modifications as well as ligand geometries are nowadays well understood. For instance, the binding affinity of the CRD to *N*-acetyl galactosamines is 60-fold higher than to galactoses and can be further increased by using bicyclic *N*-acetyl galactosamine derivatives (Figure 9d)[245,247,249]. Furthermore, while the impact of the anomeric configuration is not as significant as expected, the distance between the carbohydrate moieties, and thus the linker length (L, Figure 9a) as well as its hydrophilic-hydrophobic balance, were found to be important parameters[247,250].

With the development of promising synthetic ligands, the use of the natural function of the ASGPR became an attractive strategy for a targeted delivery of RNA therapeutics into hepatocytes and the treatment of liver diseases such as Hepatitis B, acute hepatic porphyria (AHP), or cardiovascular disease (CVD)[251]. Especially a tris-(hydroxyl-methyl)-aminomethane (tris)-based triantennary ligand with three terminal *N*-acetyl galactosamine moieties (GalNAc) showed promising uptake efficiencies of ASOs and siRNAs *in vitro* and *in vivo* (Figure 9a)[252–254]. While unconjugated ASOs targeting *SRB1* transcripts are primarily delivered to non-parenchymal liver cells, GalNAc conjugated ASOs showed a predominant uptake into hepatocytes and a six to seven-fold increased drug level[253]. Furthermore, comparing PS ASOs with PO ASOs and their uptake efficiency within different hepatic cell lines and primary murine hepatocytes, the effect of an increased uptake capability of GalNAc conjugated ASOs became even more pronounced[252]. As described before, PS ASOs are able to bind to several membrane proteins and can be internalized using different productive pathways. In contrast, low protein binding affinities and marginal internalizations are observed for unconjugated PO ASOs. Consequently, the difference between GalNAc conjugated PO ASOs compared to unconjugated

ones is higher than between PS ASOs. Therefore, the conjugation of GalNAc provides a promising approach to target hepatocytes in a highly specific manner.

1.6.3. The targeted delivery of oligonucleotides

Nowadays, transfection, electroporation, or viral transduction are common cell culture techniques to introduce genetically active elements into immortalized cell lines or even primary cells originating from different hosts and tissues. While such techniques cause a rapid accumulation of the internalized material within the cytoplasm or nucleus, protein-induced or free uptake is more time consuming and correlates with the kinetics of cellular uptake and release[200]. Especially a tissue specific and therefore targeted delivery remains challenging and in the past years, different approaches arose to internalize ASOs, siRNAs, or other genetically active elements[144,146]. For instance, the addition of an epidermal growth factor (EGF) to PS ASOs increases the productive delivery of ASOs to EGF receptor (EGFR) expressing cells[255]. However, EGFRs are ubiquitously expressed receptors within most tissues and are therefore unsuitable for targeted delivery approaches. Another attempt for the targeted delivery is the addition of glucagon-like peptide 1 (GLP-1) to PS ASOs[256]. GLP-1 binds to G-protein-coupled receptors of pancreatic β-cells and guanosine triphosphate (GTP) induced signaling results in an effective and productive internalization of PS ASOs in a specific manner.

As indicated before, the most promising and best characterized opportunity for the targeted delivery is the conjugation of triantennary *N*-acetyl galactosamine (GalNAc) to ASOs or siRNAs. GalNAc is mimicking the natural motif of the ASGPR and provides a significantly increased productive delivery into hepatocytes[253,254]. Thus, GalNAc conjugated PS 2'-MOE or 2'-cET ASOs showed a 15- to 30-fold increased potency within clinical trials[257]. Therefore, a 15- to 30-fold lower dose is required for a functional target knockdown, which is highly beneficial regarding any immunogenic indications. While unconjugated siRNAs do not distribute to any peripheral tissues in a sufficient extent[258–260], GalNAc conjugated siRNAs accumulate in the liver in a highly significant manner[261]. Fluorescence-based assays using labeled siRNAs showed also highly increased uptake efficiencies of GalNAc conjugated siRNAs into primary *mouse* hepatocytes compared to unconjugated ones[254,262]. Due to the high tissue selectivity and internalization efficiency, several GalNAc conjugated ASO- and siRNA-based drugs are involved in clinical trials targeting different liver diseases such as hereditary transthyretin mediated amyloidosis (ATTR)[261], Hepatitis B[263], or α-1-antitrypsin deficiency (AAT)[264]. Givosiran (GILVAARI™), a GalNAc-siRNA conjugate treating AHP, provides the first FDA-approved GalNAc conjugated RNA targeting therapeutic[265,266]. Additionally, the

technology of an ASGPR mediated uptake is broadening its potential to other RNA targeting therapeutic approaches, such as anti-miRNAs or small activating RNAs (saRNAs)[267]. Approaches to integrate the ASGPR into several non-hepatic and therefore non-ASGPR expressing cells, became also an interesting tool within cell culture to further investigate selectivities, pharmacokinetics and potencies of different RNA targeting drug systems[268–270]. Thereby, it was found, that the integration of subunit and isoform H1a is sufficient for an effective internalization and ASGPR function. As for other endocytosis-based delivery pathways, the endosomal release into the cytoplasm is still providing the major bottleneck and further research is necessary to increase the potency of GalNAc conjugates (Figure 9c)[271]. However, the approach of an ASGPR mediated delivery is providing the fundamental idea of this thesis, the targeted delivery of GalNAc conjugated gRNAs into ASGPR expressing cell lines to induce RNA editing.

2. Aims and approach of the thesis

RNA-based therapeutics, such as siRNAs, miRNAs, or RNase H1 induced therapies are very promising applications to introduce manipulations of genetic information without changing the genome. Since the development of site-directed RNA editing, especially the recruitment of endogenous ADAR, the approach of A-to-I editing is also providing a promising strategy to treat several diseases based on G-to-A SNPs. Additionally, the SNAP®-ADAR-based system is providing a highly effective tool to further investigate codon specificities, the tolerance of ADAR deaminase domains regarding chemical modifications of gRNAs, toxicity assays, as well as the impact of the dimerization of deaminase domains. While the SNAP®-ADAR system is using an artificial, but genetically encodable deaminase domain, its ectopic expression via transient transfection or stable integration is necessary. In contrast, recruiting endogenous ADAR is avoiding an ectopic expression of artificial fusion proteins. However, utilizing chemical modified and therefore genetically not encodable gRNAs, a transfection-based strategy is the method of choice to deliver gRNAs into cells within both approaches. For other RNA-based therapeutic systems mentioned before, the ASGPR mediated endocytosis of *N*-acetyl galactosamine (GalNAc) modified oligonucleotides arose as a favorable technique to circumvent such transfection-based strategies and to enable a targeted delivery into hepatocytes. Therefore, the aim of this thesis was to investigate the receptor mediated uptake and targeted delivery of gRNAs into ASGPR expressing cell lines, to enable the cellular delivery of gRNAs without the need of cell-toxic transfection reagents. The wet chemical synthesis of the tris-based triantennary *N*-acetyl galactosamine as well as the molecular cloning of the ASGPR from isolated HepG2 mRNA, will be the first two milestones to enable a targeted delivery. Next, it will be necessary to generate different cell lines, stably expressing the ASGPR as well as the ASGPR and the SNAP®-ADAR fusion protein. These cell lines should be characterized in an informative manner using immunofluorescence-based assays as well as western blot analysis. A wet chemical approach to conjugate the synthesized, triantennary GalNAc to chemical modified gRNAs, which are previous reported or designed by colleagues needs to be established and characterized using state-of-the-art analytics. Lastly, a combination of both approaches needs to be investigated to fathom the possibility of a receptor mediated internalization of gRNAs to induce RNA editing of two different endogenous transcripts, GAPDH and STAT1. The goal of this work, to show the viability of the receptor mediated endocytosis of gRNAs, may enable future work towards the implementation of RNA editing in a broad context of a potential therapeutic application.

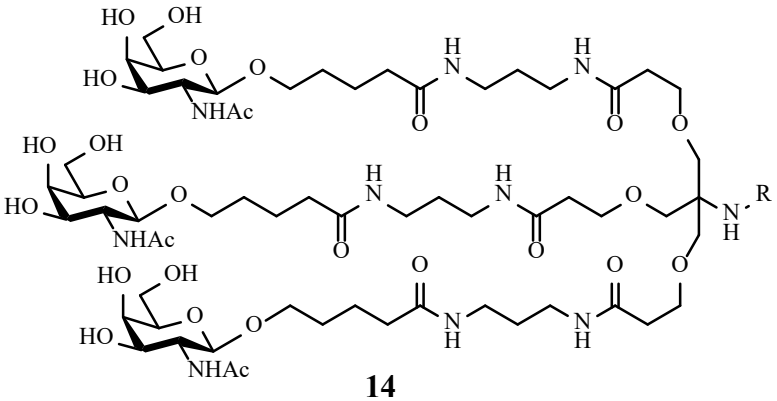
3. Results and discussion

This section describes the progress and the results of the synthesis of a triantennary GalNAc and its conjugation to gRNAs, the targeted delivery and therefore, the ASGPR mediated endocytosis of gRNAs, achieved during this thesis. Where noted, some of the following results are also reported within the manuscript of the bachelor's thesis of Yannis Stahl, who was involved in establishing the conjugation of triantennary GalNAc and maleimide derivative to SNAP-ADAR® gRNAs. All utilized SNAP®-ADAR gRNAs were designed by Ngadhnjim Latifi and Paul Vogel and all utilized RESOTRE v2 gRNAs were designed by Tobias Merkle. The stability assays of gRNA 324 and 507 were performed by Ngadhnjim Latifi and the stability assays of gRNA TMR189 and TMR236 were performed by Laura Pfeiffer. The reported stability assays are not part of this thesis.

3.1. Synthesis of triantennary GalNAc and its conjugation to gRNAs

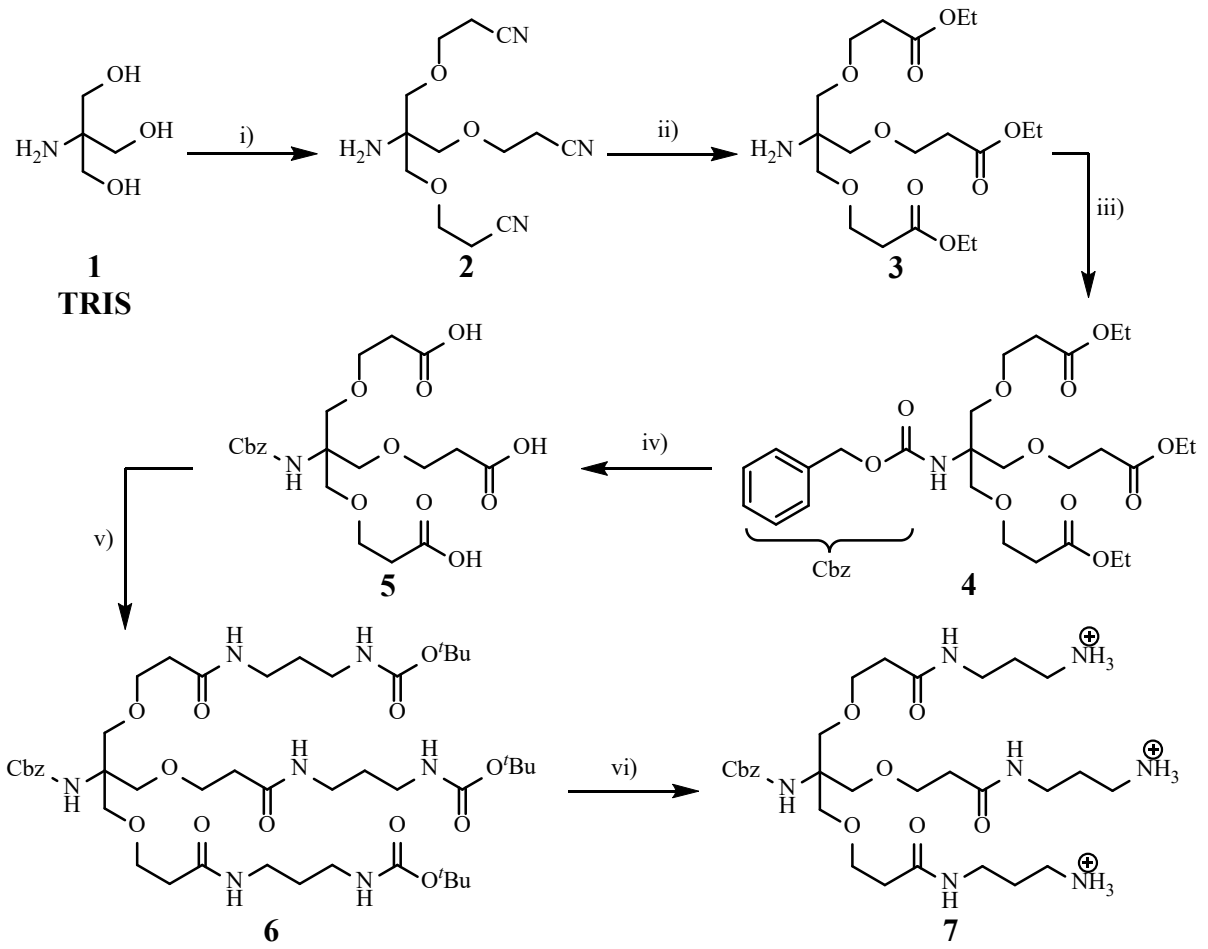
3.1.1. Synthesis of triantennary GalNAc

Based on the idea to enable a targeted delivery of gRNAs, the primary challenge was to synthesize an ASGPR ligand to facilitate the cellular endocytosis of receptor expressing cells. As described before, a tris-based triantennary *N*-acetyl galactosamine (GalNAc, **14**, Scheme 1) showed promising uptake efficiencies of ASOs and siRNAs within several cell culture experiments using primary hepatocytes as well as *in vivo* experiments in mice or even clinical trials[144,253,254]. Therefore, the reported structure was chosen as a blueprint for the beforehand synthesis of the artificial ASGPR ligand to induce a targeted delivery. Starting from commercially available substances, tris and β-D-galactosamine penta acetate, the synthesis was separated in three different parts: 1) the triantennary core (Scheme 2), 2) the terminal *N*-acetyl galactosamine residues (Scheme 3), and 3) the conjugation of both (Scheme 4 and Scheme 5). The different syntheses will be discussed in detail, within the following section.



14

Scheme 1: Triantennary N-acetyl galactosamine (14).

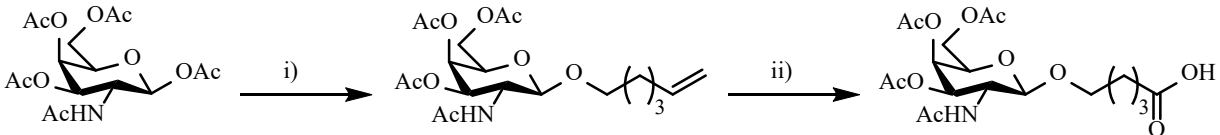


Scheme 2: Synthetic route of the tris-based triantennary core. *i) 1) KOH_{aq} (40 %), 1,4-dioxane, 2) acrylonitrile, RT / ON, 51 %; ii) H_2SO_4 (96 %), EtOH, 1) reflux / 7 h, 2) RT / ON, 47 %; iii) 1) $Na_2CO_{3\,aq}$ 2) benzyl chloroformate, 1,4-dioxane, RT / 12 h, 83 %; iv) LiOH, MeOH, RT / 5 h, 95 %; v) Mono-boc 1,3-propane diamine, HBTU, DIPEA, DMF, RT / 24 h, 90 %; vi) TFA, $CHCl_3$, RT / 1 h, 103 %.*

As described before, the triantennary core is a tris-based derivative and the trimeric structure enabled the conjugation of three *N*-acetyl galactosamine residues, which are necessary for a sufficient receptor-ligand binding. Starting from commercially available tris (**1**), the addition of acrylonitrile to the three hydroxyl groups under basic conditions (KOH_{aq}), serves as a primary elongation and the terminal nitriles provide an increased functionality compared to the hydroxyls. A simultaneous conjugation at three positions is very challenging and in contrast to subsequent reactions, the triple conjugated product (**2**) was observed as major product during the addition of acrylonitrile (51 %). Next, the functionalization of the nitrile was changed to a carboxylic acid, applying a Pinner reaction. In principle, the Pinner reaction is an acid catalyzed reaction of a nitrile and an alcohol, forming an imino ester intermediate (Pinner salt), which is converted to amidines, carboxylate esters, or thio esters using amines, water or thiols, respectively[272]. Generally, the acid catalyzed addition of alcohols to nitriles is conducted using gaseous HCl, but regarding safety concerns and as reported in literature, the use of conc. $HCl_{aq.}$

(37 %) should also yield the desired product[272,273]. However, the use of conc. aq. HCl (37 %) was not beneficial with respect to the simultaneous addition to the three different nitriles and resulted in a mixture of single, double and triple substituted products (see supplementary information, Figure S1a). The use of conc. H_2SO_4 (96 %) instead of conc. aq. HCl (37 %), in EtOH, shifted the reaction towards the higher substituted products and the subsequent purification by chromatography (silica) provided the desired and fully conjugated product (**3**) in a sufficient yield (47 %) (see supplementary information, Figure S1b)

Following the formation of the carboxylate esters from nitriles, the free amino group was protected using benzyl chloroformate (Cbz) under basic conditions ($Na_2CO_{3\ aq}$), yielding the fully protected product (**4**) in a high yield (83 %). The subsequent alkaline hydrolysis of the carboxylate esters to provide the product with free carboxylic acids (**5**) was performed with LiOH in an almost quantitative manner (95 %). Afterwards, the trimeric conjugate was further elongated with mono-boc 1,3-propane diamine using standard peptide synthesis conditions (HBTU, DIPEA). Unexpectedly, the conjugation of mono-boc 1,3-propane diamine to the three free carboxylic acids, yielded almost exclusively the triple conjugated product (**6**) in a high yield (90 %). Lastly, the deprotection of the three *tert*-butyloxycarbonyl (Boc) protected amines was performed using TFA and provided the desired product with free terminal amines as TFA salts (**7**) in a quantitative manner (103 %). The exceeding yield of more than 100 % is most probably related to a residual amount of solvent or TFA. However, the complete deprotection was confirmed applying ^{1}H-NMR spectroscopy (see supplementary information, Figure S28 and Figure S29).



8
β-D-Galactosamine penta acetate

9

10

Scheme 3: Synthetic route of the terminal N-acetyl galactosamine residues. *i) 1) TMSOTf, $CHCl_3$, molecular sieve (3 Å), RT / 70 h, 2) 5-Hexen-1-ol, RT / 4 h, 59 %, ii) 1) $NaIO_4$, DCM/ACN/H_2O (2:2:3), < 10 °C / 15 min, 2) $RuCl_3$, RT / 1 h, 62 %.*

Following the synthesis of the triantennary core, the other part was the synthesis of the *N*-acetyl galactosamine residues. Therefore, commercially available β-D-galactosamine penta acetate (**8**) was activated with TMSOTf *in situ*, followed by the substitution of 5-Hexen-1-ol to the anomeric carbon in a one-pot synthesis to yield the β-configurated product (**9**) in a good

manner (59 %). In contrast to the previously reported synthesis[253], it was not possible to isolate the deacetylated and bicyclic product, activated by TMSOTf). However, the reaction mechanism (see supplementary information, Scheme S1), indicated the selective formation of the β-anomer, which was confirmed by ^{1}H-NMR spectroscopy. The chemical shifts δ and the coupling constant *J* of the ^{1}H-NMR signals differ between α- and β-anomers and are in accordance with previously reported results by Nair *et al.*[254]. The coupling constant *J* is also coinciding with the Karplus equation, which is indicating the formation of the correct anomer. The observed and reported ^{1}H-NMR signals are compared in Table 1 and the formation of the β-configurated product was confirmed. Subsequent to the conjugation of 5-Hexen-1-ol, an oxidative cleavage of the terminal alkene double-bond using $RuCl_3$ and $NaIO_4$ lead to the formation of a terminal carboxylic acid (**10**) in a good yield (62 %).

Table 1: Observed and reported ^{1}H-NMR signals of the anomeric proton of compound 9. *All measurements were performed in DMSO-d_6. d = doublet.*

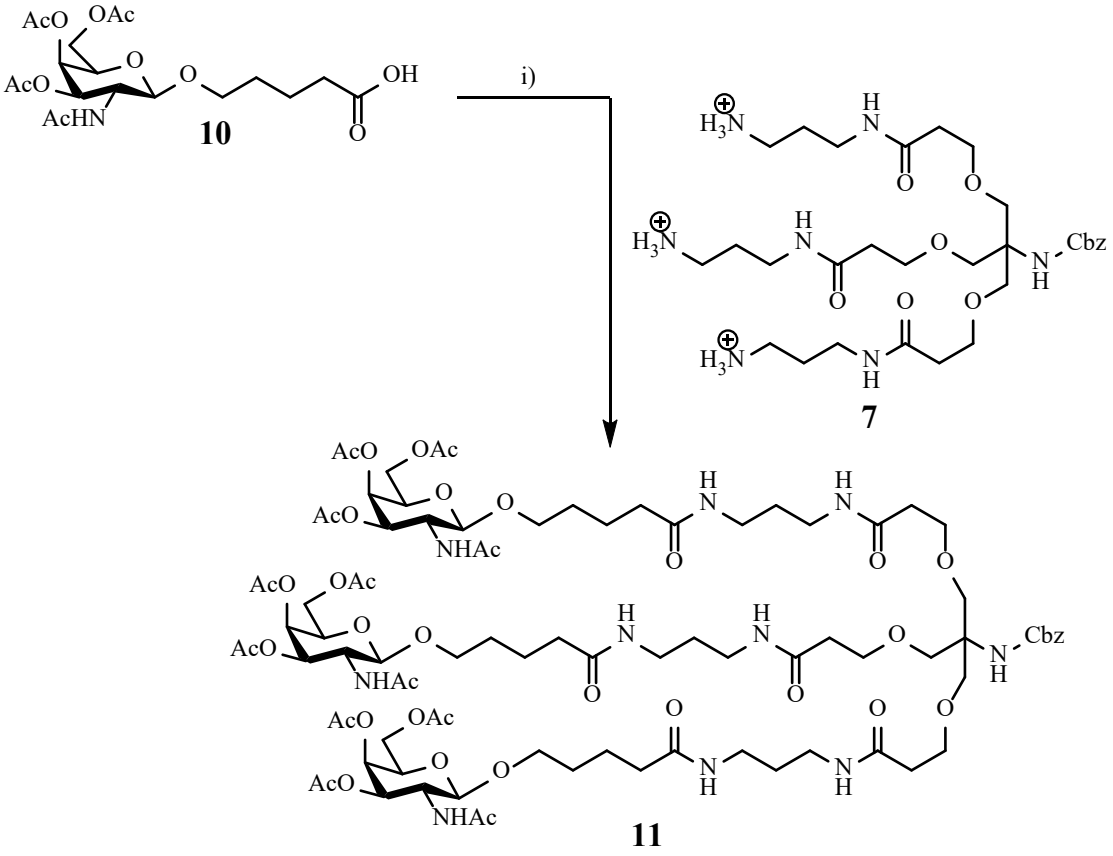
Compound	Confi-guration	Position	Chemcial shift δ [ppm]	Coupling constant *J* [Hz]	Multi-plicity	Number of Protons
9	β	H1	4.48	8.4	d	1
36[254]	β	H1	4.47	8.5	d	1
36a[254]	α	H1	4.83	4.8	d	1

To conjugate the modified *N*-acetyl galactosamines to the deprotected amines of the triantennary core (**7**), the terminal carboxylic acid (**10**) was activated using HBTU and HOBt under basic conditions (DIPEA) in a small-scale reaction (0.58 mmol) (Scheme 4). The subsequent addition of the trimeric amine derivative (**7**) lead to the formation of the desired and fully conjugated product (**11**) in a high yield (71 %), whereby only minor amounts of the single and double conjugated products were observed. Furthermore, the degree of conjugation was confirmed by ^{1}H-NMR spectroscopy. Anomeric H1- as well as acetyl-protons of the three terminal *N*-acetyl galactosamines were related to CH_2-protons of the triantennary core and aromatic protons of the Cbz-protecting group, and a correct proton ratio was observed (see Table 2). However, upscaling of the reaction (9.48 mmol) was not successful and the single and double conjugates were observed as major products within LCMS analysis (see supplementary information, Figure S2). The purification via chromatography (silica) provided the desired and fully conjugated product (**11**) in a very low yield (9 %) and the incomplete conjugation of the

side-products was confirmed using ^{1}H-NMR spectroscopy (see methods and materials, NMR data of compound **11**).

Table 2: Expected and found ^{1}H-NMR signals of compound 11. *All measurements were performed in DMSO-d_6. S = singlet, d = doublet, br s = broad singlet, m = multiplet (multiplicity is not specifiable).*

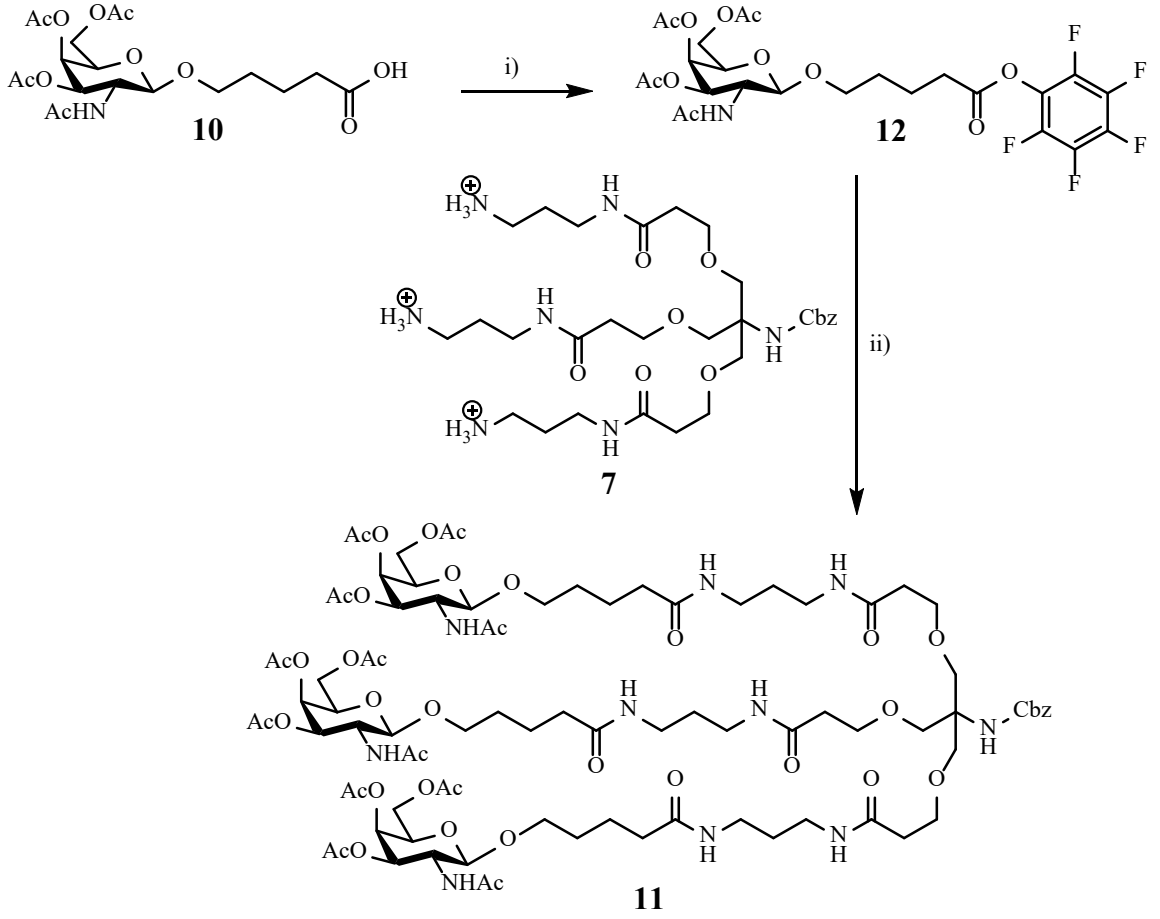
Position	Chemcial shift δ [ppm]	Coupling constant *J* [Hz]	Multi-plicity	Number of protons (expected)	Number of protons (found)
-NHOCH_3	1.77	-	s	9	9
-$H1_{\text{anomeric}}$	4.49	8.5	d	3	3
-C_q-CH_2-O (tris)	3.48	-	br s	6	6
-CH_{Aromatic} (Cbz)	7.72-7.38	-	m	5	5



Scheme 4: Conjugation of the N-acetyl galactosamine residue to the triantennary core. *i) 1) HBTU, HOBt, DIPEA, DMF, RT / 15 min 2) Compound 7, DMF, RT / ON, 71 % (9 %).*

To further increase the reactivity of the carboxylic acid, a preliminary activation using pentafluorophenol (PfpOH) and EDCI (Scheme 5) was conducted and provided a highly reactive and isolatable active ester (**12**) in a good yield (65 %). The subsequent conjugation of the unprotected amines of the triantennary core to the active ester provided the fully conjugated

product (**11**) in a quantitative yield (99 %) and the formation of the single and double conjugated products was not observed (see supplementary information, Figure S3).

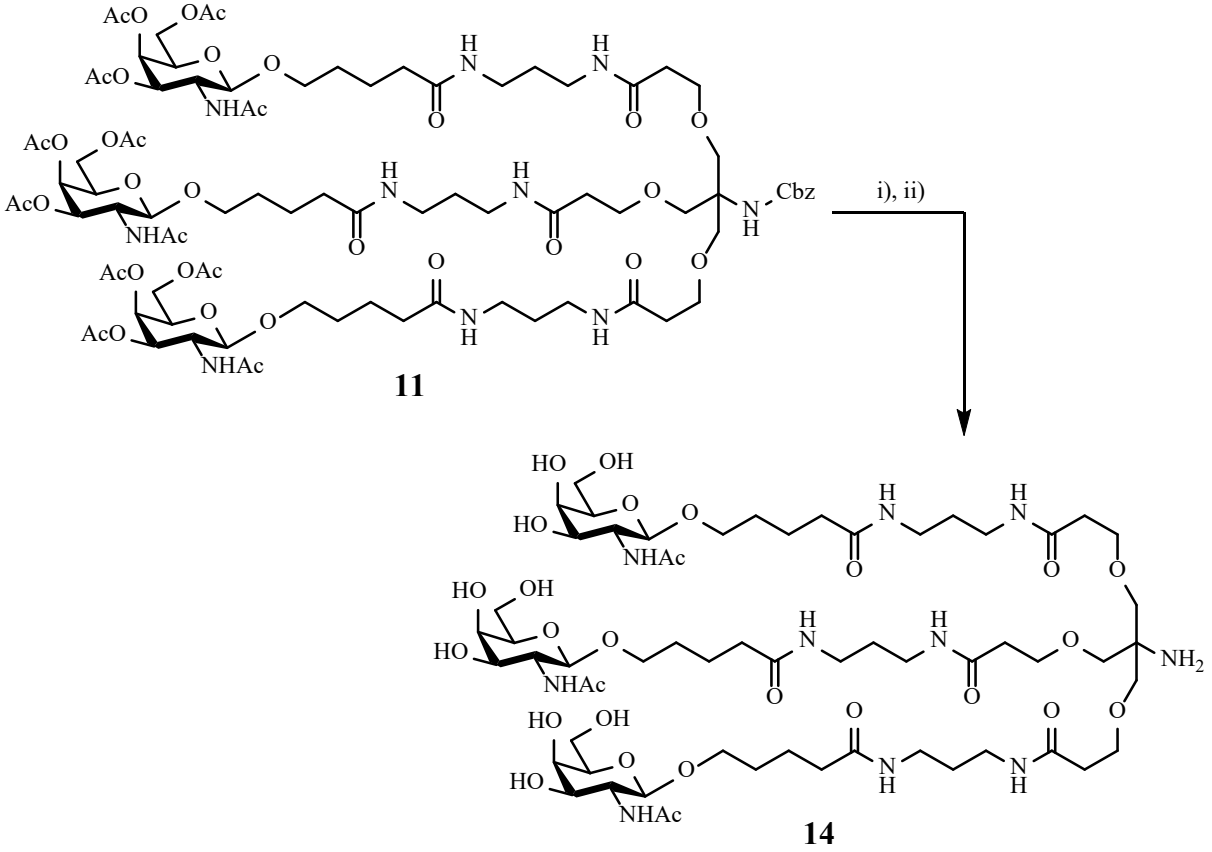


Scheme 5: Activation of the terminal carboxylic acid and conjugation to the triantennary core. *i) PfpOH, EDCI, $CHCl_3$, 1) 0 °C / 1 h, 2) RT / ON, 65 %; ii) Compound 7, DIPEA, DMF, RT / ON, 99 %.*

To provide the fully deprotected triantennary *N*-acetyl galactosamine (Scheme 6), the Cbz-protecting group was primarily hydrogenated using Palladium on carbon (Pd/C) and either molecular hydrogen (H_2) or ammonium formiate (HCO_2NH_4) was used as hydrogen source. While the use of molecular hydrogen provided the deprotected amine (**13**) in a high yield (83 %), the use of ammonium formiate yielded the hydrogenated product in a quantitative manner (105 %). As before, the exceeding yield of more than 100 % is in relation to a residual amount of solvent. However, the crude product was used without further purification and the subsequent deacetylation of the *O*-acetyls using $MeNH_2$ in EtOH provided the fully deprotected product (**14**) in a high yield (85 %).

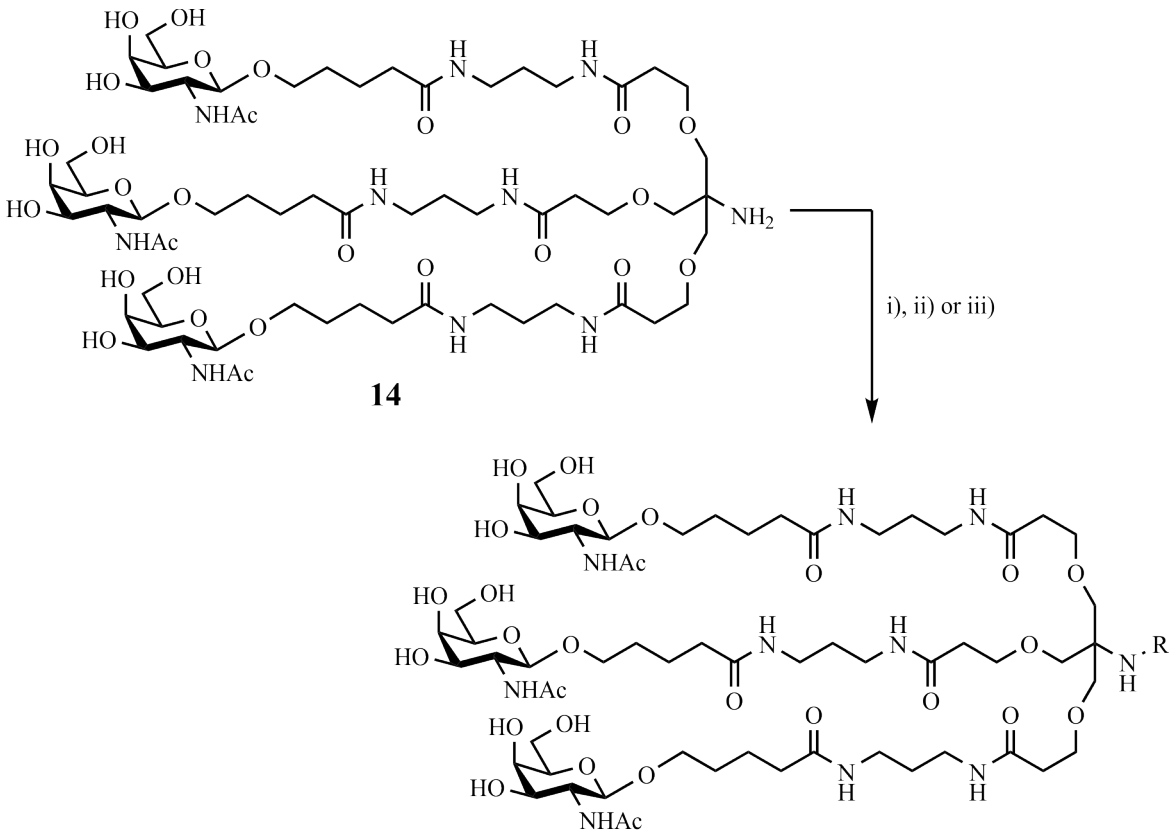
The fully deprotected *N*-acetyl galactosamine (GalNAc, **14**) was either conjugated to fluorescein 5(6)-isothiocyanate (FITC) to provide a fluorescent active conjugate or was further functionalized with 4-maleimidobutyrate or glutaric acid to enable a reactivity towards thiols

or amines, respectively (Scheme 7). The conjugation of FITC to the unprotected amine originated from the tris-based core and was performed under basic conditions (Et_3N) from which the fluorescein conjugate (**15**) was obtained in a tolerable yield (26 %). While no reaction progress was observed under neutral conditions, the addition of Et_3N was highly beneficial for the product formation. The purified product was dissolved in deuterated DMSO for ^{1}H-HMR analysis and the exact yield and concentration were determined applying UV/VIS spectrometry. For the calculations using the law of Lambert-Beer, a similar molar attenuation coefficient ε_λ at a wavelength of λ = 494 nm was assumed for FITC and the FITC conjugated GalNAc (**15**). As reference, the molar attenuation coefficient ε_λ was determined using unconjugated FITC dissolved in DMSO (see supplementary information, section 8.1.9).



Scheme 6: Deprotection of compound 11. *i) HCO_2NH_4, Pd/C (dry), MeOH, RT / ON, 105 %, or $H_{2\,g}$, Pd/C (dry), MeOH, RT / ON (83 %); ii) $MeNH_2$ (33 wt% in EtOH), RT / ON, 85 %.*

To further functionalize the unprotected GalNAc (**14**), 4-maleimidobutyrate and glutaric acid were also conjugated to the unprotected primary amine originating from the tris-based core. The conjugation of 4-maleimidobutyrate was performed using commercially available *N*-succinimidyl 4-maleimidobutyrate as active ester under basic conditions (Et_3N) and provided



Scheme 7: Conjugation and functionalization of compound 14. *i) 1) Fluorescein 5(6)-isothiocyanate (FITC), DMF, RT / ON, 2) Et_3N, RT / 5h, 26 %; ii) N-succinimidyl 4-maleimidobutyrate, Et_3N, DMF, RT / 72 h, 25 %, or 35 °C / ON, 27 %; iii) Glutaric anhydride, DMAP, DIPEA, DMF, 1) 50 °C / 5 h, 2) RT / ON, 20 %.*

the thiol reactive product (**16**) in an acceptable yield (25-27 %). In addition, the conjugation of glutaric acid was performed using glutaric anhydride and DMAP under basic conditions (DIPEA) to provide the terminal carboxylic acid and therefore amine reactive conjugate (**17**) in a tolerable yield (20 %). Obviously, a lower yield was observed for all three reactions, using different reactive conjugates (isothiocyanates, NHS-esters and anhydrides), which was addressed to a decreased reactivity of the primary amine. This is in accordance with structural properties of the triantennary core, whereby oxygen (O) is more electronegative than carbon

(C), nitrogen (N) or hydrogen (H). The -I effects (inductive effects) of the three surrounding ether groups, are lowering the electron density of the neighboring primary amine (tris), which resulted in a decreased reactivity. Therefore, higher molar excesses of the reactive conjugates were necessary to provide a sufficient product formation, which was at least very crucial for the use of glutaric anhydride. The formation of a terminal carboxylic acid under basic conditions is providing a carboxylate, which is also reactive to glutaric anhydrides. Therefore, a higher molar excess of glutaric anhydride is able to cause the formation of an asymmetric anhydride, which was also observed as a side product within LCMS analysis (data not shown). In addition, a reaction of the reactive conjugates and the unprotected hydroxyl groups of the *N*-acetyl galactosamine can not be completely excluded. However, especially maleimides are also reactive towards amines under basic conditions. A conjugation of amines to maleimides would inhibit the maleimide species and would prevent its subsequent accessibility and reactivity towards thiols. Therefore, and due to the high molar excess as well as its basic and nucleophile character, a deacetylation using $MeNH_2$, subsequent to the conjugation of 4-maleimidobutyrate, is counterproductive. Additionally, the use of other deacetylation approaches, such as hydroxides or alcoholates, are also very challenging regarding their reactivity towards Michael acceptors (maleimides). Thus, the described synthetic route was chosen to provide different functionalizations originated from a single molecule, compound **14**.

In summary, the synthesis of the triantennary *N*-acetyl galactosamine (GalNAc, **14**) and its further functionalization was successful and besides an already published and amine reactive conjugate (**17**)[274], it was possible to provide novel fluorescent active (**15**) as well as thiol reactive (**16**) GalNAc derivatives. Due to a different technical setup and equipment, it was necessary to deviate from the reported syntheses[253,254] within several reactions.

3.1.2. Chemical modification of gRNAs

Following the successful synthesis of the triantennary GalNAc (**14**) and its derivatives, the next step was to conjugate the synthesized GalNAc to chemically stabilized gRNAs. To provide different reactivities, the synthesized GalNAc (**14**) was either conjugated to 4-maleimidobutyrate (**16**) or glutaric acid (**17**) to enable selectivity towards thiols or amines, respectively. Amine reactive conjugates (**17**) are thought to enable a reactivity to NH_2-terminal RESOTRE gRNAs, whereby the thiol reactive derivative (**16**) is providing a possibility to further modify BG conjugated SNAP®-ADAR guides. BG moieties are generally conjugated to NH_2-terminal chemically stabilized gRNAs, which is preventing an additional GalNAc conjugation via an amine reactive functionalization. Detailed information about the chemically

stabilized gRNAs, their terminal functional groups and the conjugation of BG moieties as well as GalNAc conjugates are described within the following sections.

3.1.2.1. BisBG modification of SNAP®-ADAR gRNAs

As already mentioned, the SNAP®-ADAR system is to this point of time the best characterized RNA editing system to induce site-directed A-to-I base substitutions. An artificial SNAP®-tag and ADAR deaminase domain containing fusion protein is covalently bound (*in situ*) to an O^6-benzylguanine (BG) modified and chemically stabilized gRNA, which is complementary to the target site. Besides the described mono-BG containing gRNAs, which comprises the BG moiety and therefore the SNAP®-ADAR fusion protein, as well as the gRNA in a 1:1 stoichiometric ratio, recent research provided also BisBG modified gRNAs. BisBG is a dimeric linker which enables the binding of two SNAP®-ADAR fusion protein to one gRNA at its 3'- or 5'-end. The synthesis of BisBG-COOH (**18**) and the conjugation to gRNAs was established by Alfred Hanswillemenke and is reported in detail by Stroppel *et al.*[119]. During this thesis, only BisBG conjugated gRNAs were used for experiments utilizing the SNAP®-ADAR editing system.

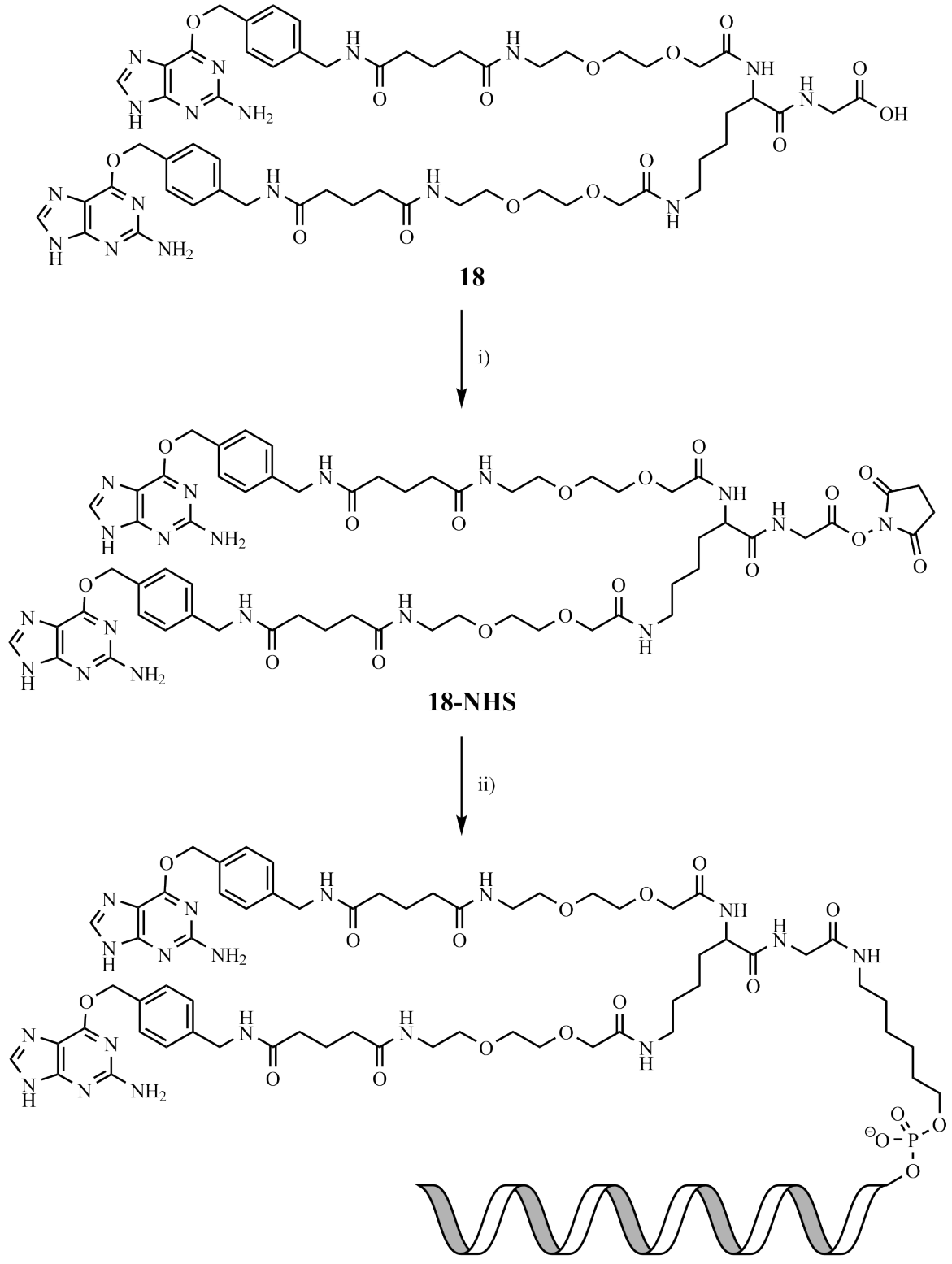
Table 3: Overview of the BisBG-COOH (18) conjugated gRNAs. *The following chemical modification are used: * = PS linkage,* N *= RNA, N = 2'-OMe,* **N** *= DNA,* {N} *= LNA,* N *= Target site missmatch. The structures of the terminal modifications (-C_6-NH_2, -C_6-Disulfide, and -GalNAc) are illustrated in supplementary information Scheme S2. The used SNAP®-ADAR gRNAs were designed by colleagues. gRNAs 257 and 258 were designed by Paul Vogel and gRNAs 324, 507, 471, 472 and 473 were designed by Ngadhnjim Latifi. Nt's = nucleotides. To provide a linker between the BisBG moiety and the gRNA, the first three nucleotides of the 5'-end of gRNA 257, 258, 324 and 507 are not complementary to the target site.*

No. of gRNA	Target	Sequences and base modifications (5' to 3')	No. of Nt's	Terminal modifications
257	GAPDH 3'UTR	*G*A*ACAAGGGGU***C**N**C**A*CAUGGCA* A*C*U*G*	25	5': C_6-NH_2 3': C_6-Disulfide
258	GAPDH ORF2	*C*C*GAGGUUUUU***C**N**C**A*GACGGCA* G*G*U*C*	25	5': C_6-NH_2 3': C_6-Disulfide
324	STAT1 Y701C	*A*G*U*{G}*U*{C}*UUGAU***A**N**C**A*UCCAG UU*C**{C}**U**{T}	25	5': C_6-NH_2 3': none
507	STAT1 Y701C	*A*G*U*{G}*U*{C}*UUGAU***A**N**C**A*UCCAG UU*C**{C}**U**{T}	25	5': C_6-NH_2 3': GalNAc
471	STAT1 Y701C	{G}**U**{C}**U*U*G*A*U****A*****C*****A***U *C*C*A*G*U*U*C**{C}**U**{T}	22	5': C_6-NH_2 3': C_6-Disulfide
472	STAT1 Y701C	{T}**U**{G}**A*U****A*****C*****A***U*C*C*A* G**{T}**U**{C}	16	5': C_6-Disulfide 3': C_6-NH_2
473	STAT1 Y701C	{G}**U**{C}**U*U*G*A*U****A*****C*****A***U *C**{C}**A**{G}	16	5': C_6-Disulfide 3': C_6-NH_2

In general, BisBG-COOH (18) is synthesized using solid-phase-synthesis from a *C*-terminal and resin bound glycine, followed by the subsequent conjugation of Fmoc-lysine, Fmoc-PEG, glutaric anhydride and O^6-(4-aminomethyl-benzyl) guanine. After resin cleavage, the terminal carboxylic acid of the glycine residue is providing an amine reactive functionalization, which enables the conjugation to NH_2-terminal gRNAs. The 3'- or 5'-terminal amine of gRNAs is provided by the manufacturer and comprises a C_6 linker in between the gRNA and the primary amine. However, native carboxylic acids are very unreactive towards amines, and a preliminary activation is necessary to provide a sufficient reactivity. Therefore, the terminal carboxylic acid of BisBG-COOH (**18**) was activated using DIC and NHS under basic conditions (DIPEA) (Scheme 8) and the reaction was monitored using HPLC analysis. With an activation progress of > 50%, the reaction was lyophilized and used for the subsequent conjugation to NH_2-terminal gRNAs without further purification and an overview of the gRNAs conjugated to BisBG-COOH (**18**) is listed in Table 3. The design of the gRNAs was established by colleagues (Ngadhnjim Latifi, Paul Vogel) and further details about the target (STAT1 Y701C), their designs and the results of their application within transfection or passive uptake experiments will be discussed in a subsequent section (see section 3.3.1).

The lyophilized, and activated BisBG-COONHS (**18-NHS**) was conjugated to the desired gRNA under basic conditions (DIPEA) (Scheme 8) and separated from unconjugated oligonucleotide as well as excessive BisBG-COOH (**18**) utilizing a denaturing TBE-7 M Urea-PAGE (20 %). The desired product was sliced out, extracted into water, and further purified using either precipitation (NaOAc/EtOH) or C18 cartridges (Sep-Pak® Plus C18 cartridges). No precipitation was applicable for the gRNAs 471, 472, and 473. These gRNAs are shorter (16-22 nucleotides compared to 25 nucleotides) and comprise PS linkages between all nucleotides. As described before, phosphorothioates (PS) are softer bases than phosphates (PO) and while a higher protein binding capability is attributed to the softer base character of the PS linkage, the same property is counterproductive for ion-ion interactions, which are necessary for precipitation. Therefore, the shorter and fully PS modified design of gRNAs 471-473 is most probably the reason for the ineffective precipitation. Because of this a Sep-Pak® Plus C18 cartridge purification was necessary after Urea-PAGE separation. In contrast, gRNAs with a length of ~25 nucleotides and primarily PO linkages, such as 257, 258, 324, or 507 are purifiable via precipitation as expected. Unexpectedly, the additional hydrophilic GalNAc modification of gRNA 507 did not show any adverse effects concerning the precipitation. Purified BisBG-471 was analyzed by mass spectrometry (ESI-) in cooperation with BioSpring

GmbH (Frankfurt, Germany), and the successful conjugation of the preliminary activated BisBG-COONHS (**18-NHS**) to the NH_2-terminal gRNA 471 was confirmed.



Scheme 8: General workflow of the conjugation of BisBG-COOH (18) to NH2- gRNAs. *The primary amine is either located at the 5'- or at the 3'-end of the gRNA. (i) NHS, DIC, DIPEA, DMSO, 45 °C / 4 h; (ii) NH_2-terminal gRNA in H_2O, DIPEA, DMSO, 37 °C / 2h.*

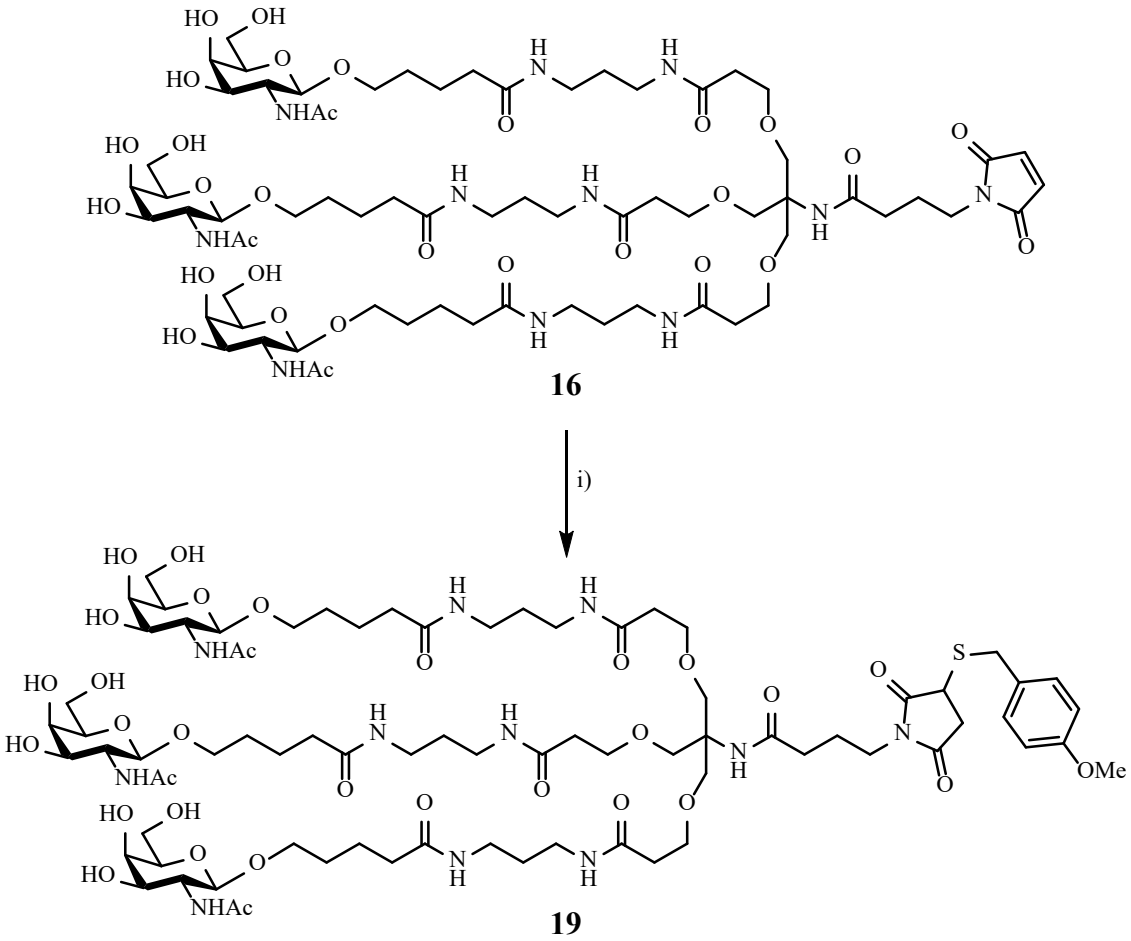
In summary, the conjugation of BisBG-COOH (**18**) to NH_2-terminal gRNAs is a well-established protocol within our research group and yields between 10 % and 40 % are commonly observed. While gRNAs with a classical design (~25 nucleotides, minor PS content) are generally purified by precipitation, the purification of gRNA 471, 472, and 473 was more challenging due to the shorter and fully PS modified design. However, the BisBG-COOH (**18**) conjugation to all desired SNAP®-ADAR gRNAs was successful, which was confirmed via mass spectrometry for 471, and the provided gRNAs were used for the further GalNAc conjugation or subsequent cell culture experiments.

3.1.3. GalNAc modification of SNAP®-ADAR gRNAs

After the successful conjugation of BisBG to gRNAs, the next step was to modify the provided BisBG gRNA with the synthesized GalNAc-maleimide derivative (**16**) to induce an ASGPR mediated endocytosis. As described before, maleimides are thiol reactive functionalizations and to prevent an oxidative dimerization of thiols, they are provided by the manufacturer as a disulfide derivative. Therefore, a deprotection (or cleavage) of the terminal disulfide was necessary to enable the accessibility of the free thiol prior to the conjugation to the maleimide modified GalNAc (**16)** (Scheme 10). The design of the utilized gRNAs was inspired by passive uptake results targeting MECP2, which was investigated by Clemens Lochmann under the supervision of Ngadhnjim Latifi[275]. While PS modified oligonucleotides are able to enter the cell within several pathways, editing yields of about 40 % were observed for shorter (16-22 nt's) and fully PS modified gRNAs. With the idea to further increase the uptake capability inducing an ASGPR mediated endocytosis, the design of the investigated gRNAs was applied to target STAT1 Y701, including a terminal disulfide linker at the 5'- or 3'-end (Table 3, gRNAs 471, 472 and 473). STAT1 (Signal transducer and activator of transcription 1) is a transcription activation factor and is responsible for cell growth and apoptosis, T_H1 cell-specific cytokine production, as well as antimicrobial defense[276]. The phosphorylation of the key residue Y701 is induced by Janus kinases (JAKs), SRC family kinases or tyrosine kinases and is crucial for the STAT protein activity. Therefore, the 5'-UAU codon of the phosphorylation site Y701 is providing an attractive and endogenous target to manipulate signaling cascades via SNAP®-ADAR mediated A-to-I editing[118].

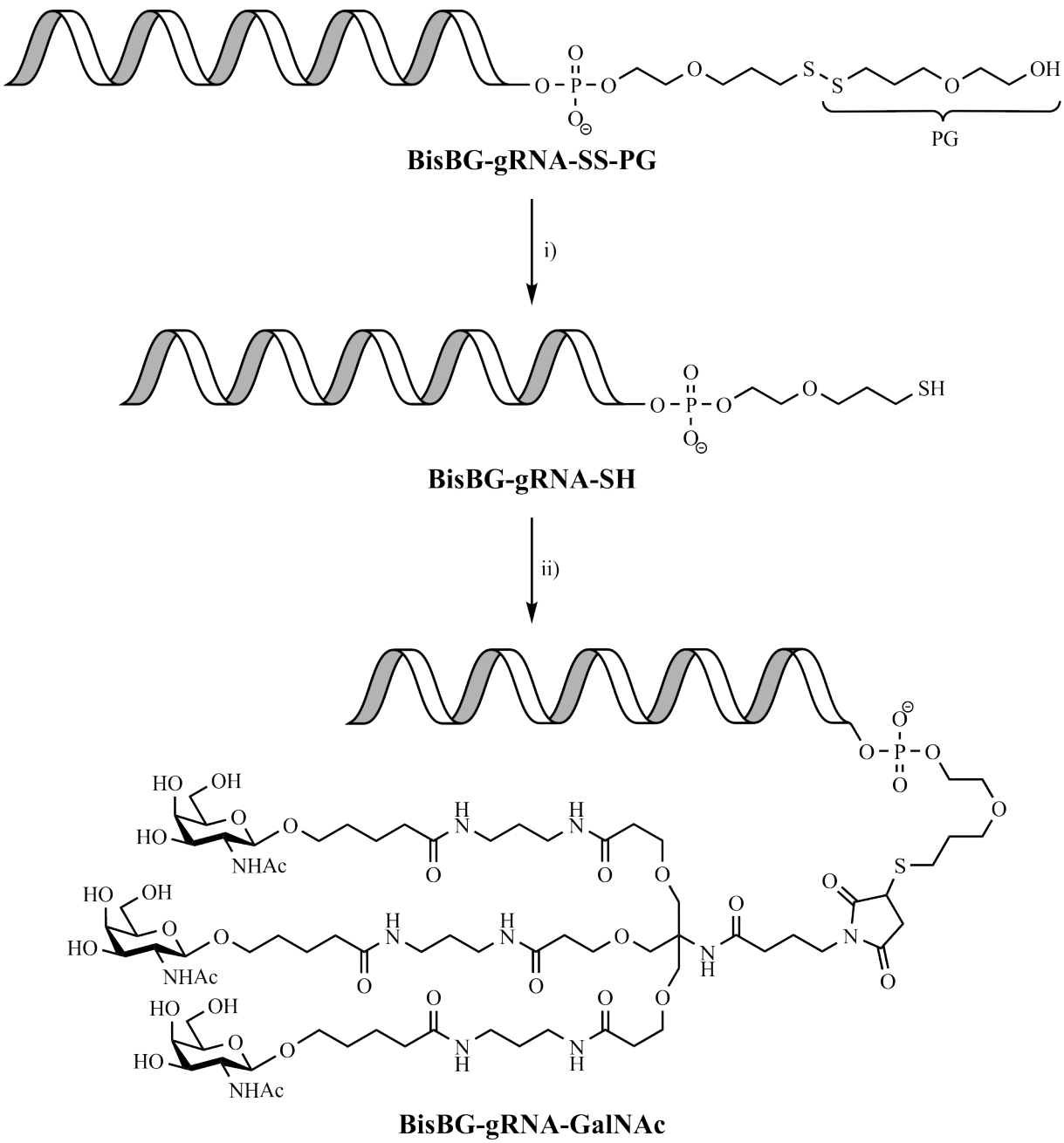
Preliminary to the deprotection of gRNAs and the subsequent conjugation to compound **16**, the reactivity of the synthesized GalNAc-maleimide derivative (**16**) towards thiols was confirmed within an exemplary reaction in a small-scale approach. Therefore, the GalNAc-maleimide derivative (**16**) and 4-methoxybenzyl mercaptan were incubated in PBS (pH 7.4) at

RT for two hours and the reaction progress was monitored using LCMS analysis (Scheme 9). After two hours, no educt was observable and the selectivity as well as reactivity of compound **16** was confirmed by the formation of a single product (see supplementary information, Figure S4). Additionally, any side reactions of the unprotected hydroxyl groups of the *N*-acetyl galactosamines were excluded, due to missing formation of any side products. Therefore, the synthesized GalNAc-maleimide derivative (**16**) was used for further approaches to modify thiol-terminal gRNAs.



Scheme 9: Conjugation of 4-methoxybenzyl mercaptan to compound 16. *i) 4-methoxybenzyl mercaptan, phosphate buffer (0.1 M, pH 7), RT / 2 h, yield: n. d.*

To gain access to the free thiols, the primary approach to deprotect disulfide modified gRNAs was performed with the previously designed gRNAs 257 and 258 (Table 3) targeting a 5'-UAG codon of the 3'-UTR or the ORF of GAPDH, respectively. The described gRNAs were used exclusively for establishing the deprotection of disulfide modified gRNAs and were not used for any cell culture experiments. The deprotection of the described BisBG gRNAs was performed within a small-scale approach (35 pmol) using DTT in phosphate buffer (pH 8.4).



Scheme 10: General workflow of the conjugation of compound 16 to BisBG-gRNAs. *The disulfide is either located at the 5'- or at the 3'-end of the gRNA. (i) DTT, phosphate buffer (0.1 M, pH 8.4), RT / 1 h / 300 rpm; (ii) Compound* ***16****, phosphate buffer (0.1 M, pH 7), RT / 3 h / 300 rpm. The BisBG moiety is not illustrated and is located at the opposite terminus of the disulfide. PG = protecting group.*

The excess of DTT was removed by precipitation (NaOAc/EtOH) and the amount of DTT within the supernatant was quantified for each purification step using Ellman's reagent. The results are described in Table S1 and Table S2 and under the applied conditions, no DTT was detectable within the supernatant of the second washing step. Unfortunately, the determined amount of DTT within the supernatant of the primary precipitation is exceeding the inserted amount of DTT (1 µmol). This can be attributed to the following aspects: 1) inaccuracies during the preparation of the stock solutions (weighing errors), 2) inaccuracies during pipetting, and

3) the assumption of a cylindrical shape for the calculation of the path length d, which is not in accordance with the reality of a concave surface. However, the removal of excessive DTT via precipitation was confirmed and the deprotected gRNAs were further used for the conjugation to the synthesized GalNAc-maleimide derivative (**16**). The subsequent conjugation of the deprotected gRNAs to compound **16** was performed in nuclease free water overnight at 4 °C and the conjugation was separated utilizing a denaturing TBE-7 M Urea-PAGE (15 %) and visualized using SYBR™ Gold Nucleic Acid Gel Stain and fluorescence imaging. The results of the conjugation and the subsequent separation are shown in Figure 10, and the deprotection of the gRNAs as well as the conjugation to compound **16** enabled the formation of additional bands (Figure 10a and b). The cleavable disulfide comprised an uncharged PEG derivative (see Scheme 10, ΔM = 135.1 Da) and compared to the total molecular mass and volume of oligonucleotides, its impact regarding the capability to migrate within a TBE-7 M Urea-PAGE was negligible. Therefore, the band of the unprotected gRNAs is expected at a comparable, but due to the reduced molecular mass and size, slightly different shift than the disulfide containing oligonucleotides. This is indeed indicated as a faint and less intense band within the samples of the deprotected gRNAs (Figure 10c). Furthermore, an additional band is observed within the deprotected controls and the conjugation reactions (Figure 10a), and a further band appeared when compound **16** was added (Figure 10b). The band which appeared after the addition of compound **16** (Figure 10b) can be attributed to the formation of the desired products, whereby the bands which is observed within the deprotected controls as well as the conjugations (Figure 10a) remains unknown. However, and as described before, free thiols tent to dimerize under oxidative conditions. While non-degassed nuclease free water was used for the precipitation, conjugation and gel loading, the observed and unknown band might be related to dimerized gRNAs as side products. Molecular oxygen is a strong oxidizing agent, and due to the presence of solute oxygen within non-degassed solvents, the formation of dimerized side products is most reasonable. However, the deprotection and subsequent conjugation of BisBG gRNAs 257 and 258 to compound **16** showed promising results and the approach to provide GalNAc modified BisBG gRNAs was further pursued using the STAT1 targeting gRNAs 471, 472 and 473. To avoid the formation of undesired side products, the use of degassed solvents was used for the further approaches to deprotect disulfides and the subsequent conjugation to compound **16**. However, the applicability of the described method is highly dependent on the gRNAs capability to precipitate and as described before, no precipitation is utilizable for the shorter and fully PS modified gRNAs 471-473. To ensure a quantitative removal of the reducing agent,

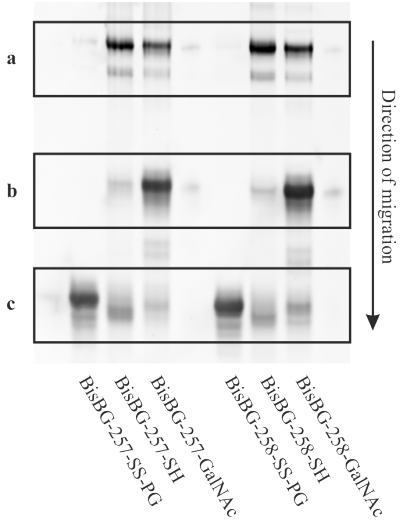


Figure 10: Fluorescence imaging of GalNAc conjugated gRNA 257 and 258. *The gRNAs were conjugated to the GalNAc-maleimide derivative (**16**) and separated using a denaturing TBE-7 M Urea-PAGE in an analytical scale. As references, protected and deprotected gRNAs were cast to the gel. The oligonucleotides were stained using SYBR™ Gold Nucleic Acid Gel Stain. PG = protecting group.*

another approach was necessary to prevent any side reactions of compound **16** with the free thiols of the excessive or residual DTT and to provide the deprotected BisBG gRNAs with a sufficient amount and purity. A promising opportunity to remove the excessive reducing agent would be provided by the use of Sep-Pak® Plus C18 cartridges. However, several different solvents, which are generally prepared prior to use, are necessary for the cartridge purification, and the use of degassed solvents is highly recommended. Therefore, and due to the time consuming procedure of solvent degassing, the applicability of Sep-Pak® Plus C18 cartridges is not as beneficial as expected to remove the excessive amount of reducing agent. However, another promising possibility was the use of size exclusion and desalting spin columns (Zeba™ Spin Desalting Columns 7 kDa MWCO). In contrast to Sep-Pak® Plus C18 cartridges, a single solvent is used, which is not necessarily prepared prior to use. This was highly beneficial concerning the use of degassed solvents and in combination with a fast procedure, the application of the described desalting spin columns became an attractive opportunity. Additionally, deprotections using DTT and conjugations of free thiols to maleimide derivatives are in general performed under different conditions. While the deprotection with DTT is conducted under a pH of 8.4, the conjugation of free thiols to maleimide derivatives requires a pH of 7.0 to prevent any side reactions of amines with maleimide derivatives at basic conditions. Therefore, the buffer exchange procedure of the described spin columns, provided by the manufacturer, was also a promising approach to remove the excess of reducing agent as well as to exchange the buffer from pH 8.4 to pH 7.0 at once. The capability to remove DTT in a sufficient degree was additionally investigated, and as described before, the amount of DTT

within the eluent of the mentioned spin columns was quantified using Ellman's reagent. The results are described in Table S3 and under the used conditions, no DTT was detectable within the primary eluent of the desalting spin columns. While minor amounts of the initially used DTT (< 1 %) were detected within the second eluent of the desalting spin columns, the capability to remove DTT in a sufficient degree was confirmed and the described columns were used for the purification of the deprotected gRNAs 471-473. The deprotection of gRNAs 471-473 was performed as described before for gRNA 257 and 258, using DTT in phosphate buffer (pH 8.4) in a larger scale approach (2 nmol) and the deprotection was purified with the described spin columns according to the manufacturer's protocol and the buffer exchange procedure.

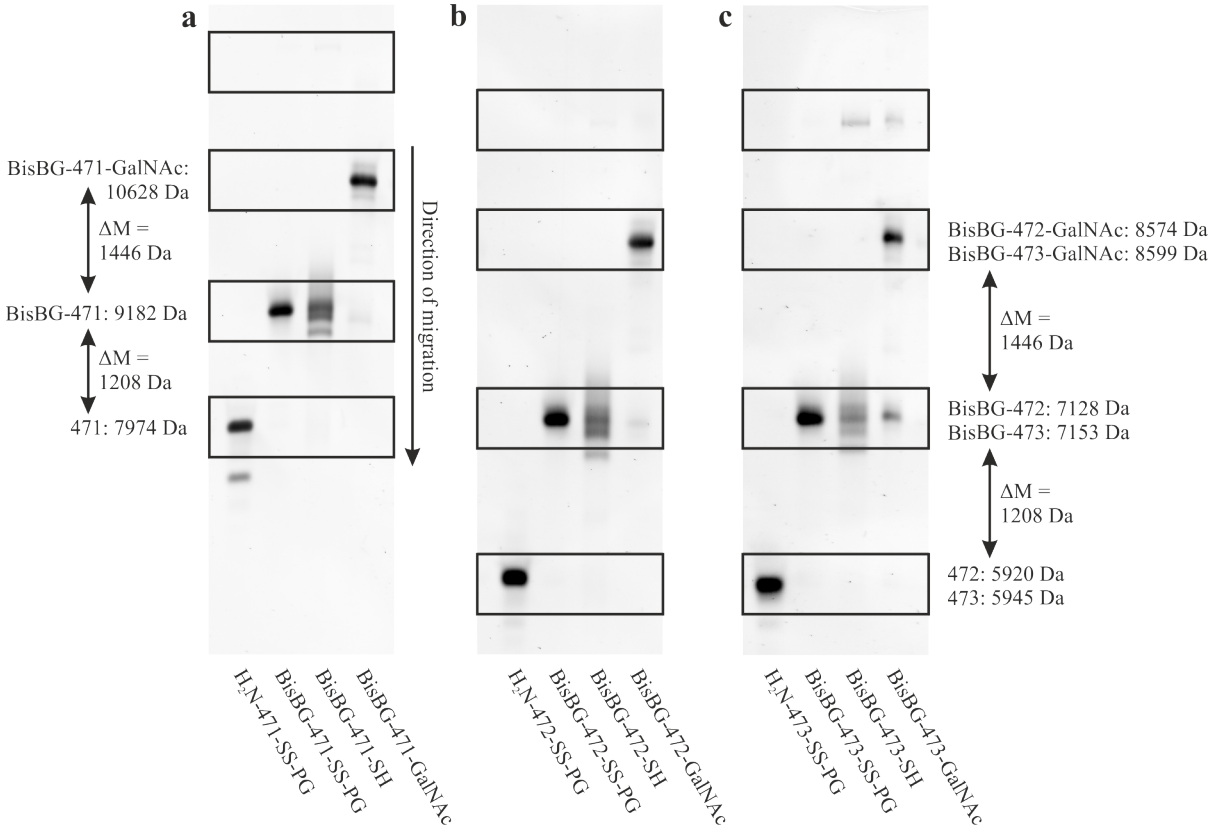


Figure 11: Fluorescence imaging of GalNAc conjugated gRNA 471-473. *The gRNAs were conjugated to the GalNAc-maleimide derivative (**16**) and separated using a denaturing TBE-7 M Urea-PAGE in an analytical scale. (a) Conjugation of compound 16 to gRNA 471 (2 nmol, M = 7974 Da, 22 nt's). (b) Conjugation of compound 16 to gRNA 472 (2 nmol, M = 5920 Da, 16 nt's). (c) Conjugation of compound 16 to gRNA 473 (2 nmol, M = 5945 Da, 16 nt's). As references, protected and NH_2-terminal, protected, as well as deprotected BisBG gRNAs (-SH) were cast to the gel, whereby a single band is observed for each sample. PG = protecting group. For visualization, the contrast of each gRNA setting was adjusted differently. The oligonucleotides were stained using SYBR™ Gold Nucleic Acid Gel Stain and for a uniform contrast adjustment, see supplementary information Figure S5.*

The subsequent conjugation of compound **16** was performed in phosphate buffer (pH 7.0) and the conjugation was separated utilizing a denaturing TBE-7 M Urea-PAGE (15 %) and visualized using SYBR™ Gold Nucleic Acid Gel Stain and fluorescence imaging. The results

of the conjugation and the subsequent separation are shown in Figure 11, and as observed before, the deprotection of the gRNAs as well as the conjugation to compound **16** enabled the formation of additional bands. As expected, a different shift is also observable for gRNAs 471 (Figure 11a) than for 472 (Figure 11b) and 473 (Figure 11c). This is in accordance with their length and therefore their molecular mass and volume, whereby a further migration is expected and observed for shorter oligonucleotides (472 and 473). As before, the band of the unprotected gRNAs is expected at a comparable, but due to the reduced molecular mass and size, slightly different shift as the disulfide containing oligonucleotides (ΔM (protecting group) = 135.1 Da). This is indeed indicated as a faint and blurred band within the samples of the deprotected gRNAs. In contrast to the deprotection and the conjugation of compound **16** to gRNAs 257 and 258, only a faint and unknown band was observed for gRNA 473 (Figure 11c). For gRNAs 471 and 472, a similar band was exclusively observable utilizing high contrast adjustments (see supplementary information, Figure S5). Additionally, a single band was observed after the addition of compound **16**, which is most probable attributed to the formation of the desired products and only negligible amounts of unreacted educts or undesired side products were observed. As reference, unmodified NH_2-terminal gRNAs were cast to the gel, and the shift of the gRNA was changed in a similar way for both, the conjugation of BisBG-COOH (**18**) as well as the GalNAc-maleimide derivative (**16**). While the charge of the modified oligonucleotide is maintained during each chemical modification, the molecular mass has changed in a comparable degree (see Figure 11). Therefore, the observed shift of the different gRNAs and their conjugates is in accordance with the already described expectations regarding the impact of the molecular mass and volume of oligonucleotides and their capability to migrate. As for the conjugation of BisBG, the assumed product was sliced out, extracted into water, and further purified using Sep-Pak® Plus C18 cartridges. In cooperation with BioSpring GmbH (Frankfurt, Germany), purified BisBG-471-GalNAc was additionally analyzed by mass spectrometry (ESI) and the successful conjugation of GalNAc-maleimde (**16**) to BisBG-471 was confirmed.

In summary, the conjugation of the synthesized GalNAc-maleimide derivative (**16**) to disulfide and preliminary BisBG modified SNAP®-ADAR gRNAs using a wet chemical approach was established successfully and the identity of the desired product was confirmed via mass spectrometry. While no precipitation was applicable for gRNAs 471-473, the primary challenge was the removal of the excessive reducing agent, which was utilized using size exclusion and desalting spin columns. The application of the described spin columns in combination with the use of degassed solvents was also highly beneficial regarding the

suppression of an undesired but unknown side product. However, low yields of around 10 % were observed for all gRNAs, which is most probably due to the use of the described spin columns and the subsequent purification of the conjugation by denaturing TBE-7 M Urea-PAGE. While a complete retention of DTT was observed using the described spin columns, a total recovery of the used gRNAs was desired but only about half of the amount of the inserted gRNA was obtained. This is most probably in relation with the shorter design of the gRNAs, which is in conflict with the molecular weight cut-off (MWCO) of 7 kDa of the used spin columns. In addition, and to prevent a competitive internalization of unconjugated GalNAc, which was previously investigated by Nair *et al*.[254], the further purification of the product is highly recommended to remove unconjugated gRNA as well as excess amounts of compound **16**. However, and despite the low yield, it was possible to establish a wet chemical approach to conjugate the synthesized GalNAc derivative to disulfide and BisBG modified SNAP®-ADAR gRNAs.

3.1.4. GalNAc modification of RESTORE gRNAs

In addition to the capability to modify disulfide containing gRNAs with the synthesized GalNAc maleimide derivative (**16**), another approach was the possibility to provide GalNAc conjugated gRNAs via an amine reactive functionalization. The idea was to modify NH_2-terminal RESTORE gRNAs with the carboxylic acid terminal GalNAc conjugate (**17**) utilizing a similar approach as for the BisBG-COOH (**18**) modification of SNAP®-ADAR gRNAs.

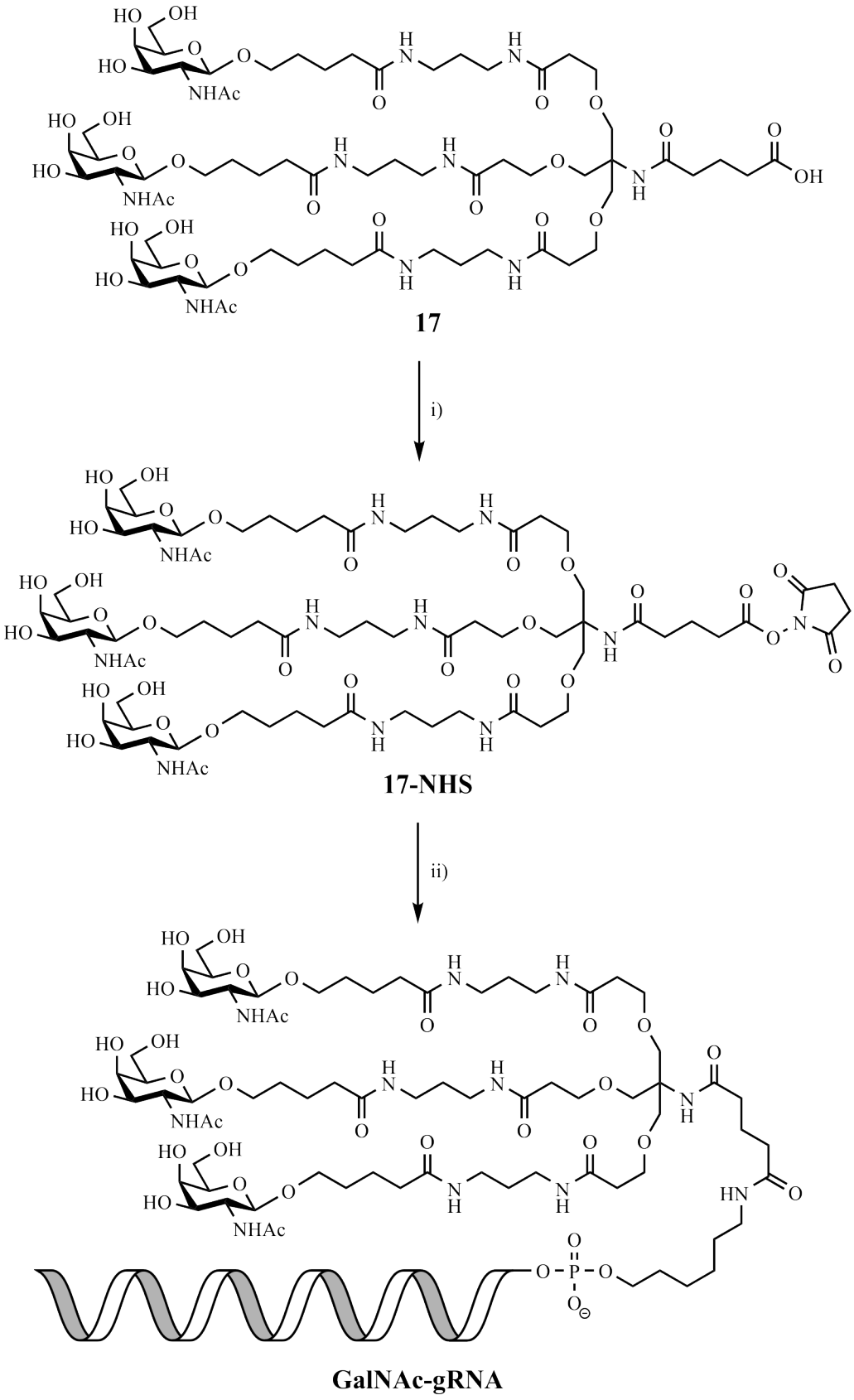
Table 4: Different approaches to conjugate GalNAc-COONHS (17-NHS) to NH_2-gRNAs. *For detailed information about the used gRNAs, see Table 14. eq. = equivalents.*

gRNA	n (gRNA) [pmol]	n (17-NHS) [nmol]	Estimated eq.
324	30	6	200
	30	54	1800
TMR189	30	6	200
	30	54	1800

As mentioned before, native carboxylic acids are very unreactive towards amines, and a preliminary activation is necessary to provide a sufficient reactivity. Therefore, the terminal carboxylic acid of GalNAc-COOH (**17**) was activated using DIC and NHS under basic conditions (DIPEA) (Scheme 11) and the reaction was monitored using HPLC analysis. The complete reaction was lyophilized and the crude product was purified by preparative HPLC to provide the NHS ester (**17-NHS**) as colorless solid. However, due to a small scale reaction and

a missing absorbance at a suitable wavelength, it was not possible to determine the exact yield of the purified product **17-NHS**, and while no educt was detectable via HPLC analysis, a yield of 50 % was assumed after purification. For the subsequent conjugation of NH_2-terminal gRNAs, different approaches were performed using two different gRNAs (324 and TMR189) and two different GalNAc/gRNA ratios for each gRNA (see Table 4). While both gRNAs contain 5'-terminal primary amines, gRNA 324 is a 25 nt long SNAP®-ADAR gRNA targeting STAT1 Y701 and TMR189 is a 59 nt long RESTORE gRNA targeting the ORF of GAPDH L157L. Both gRNAs were additionally obtained from the manufacturer as 3'- and GalNAc modified oligonucleotides (507 and TMR236, respectively), which provided a suitable reference for the analysis of the GalNAc-COOH conjugated gRNAs via Urea-PAGE.

The conjugation of the desired gRNAs to the GalNAc-COONHS ester (**17-NHS**) was performed in gRNA labeling buffer (phosphate buffer with $NaHCO_3$, pH 8.3) and the conjugation was separated utilizing a denaturing TBE-7 M Urea-PAGE (15 %) and visualized using SYBR™ Gold Nucleic Acid Gel Stain and fluorescence imaging (see Figure 12). Both commercially available and GalNAc conjugated gRNAs (507 and TMR236) showed a decreased migration than the unconjugated ones (324 and TMR189). This is in accordance with the previously discussed results and is most probably related to the increased molecular mass, whereby the charge of the oligonucleotide remained unchanged. Additionally, a different migrational change was observed for the commercially available and GalNAc conjugated SNAP®-ADAR gRNA 507 compared with the conjugation of gRNA 324 and the synthesized GalNAc-COONHS ester (17-NHS) (Figure 12a). In contrast, a similar shift was observed for the commercially available RESTORE gRNAs TMR236 and the conjugated gRNA TMR189 (Figure 12b). The GalNAc derivative of the manufacturer comprises a different chemical structure than the synthesized GalNAc-COOH (17) (see supplementary information Scheme S2, ΔM (GalNAc, manufacturer) = 1784.9 Da; ΔM (GalNAc-COOH, **17**) = 1510.8 Da) and the impact of the differing molecular masses is taking an increased effect on the SNAP®-ADAR gRNAs than on the RESTORE gRNAs (see molecular masses, Figure 12). Therefore, the observed migrations are in accordance with the expectation and furthermore, while an increased product formation is generally observed for the conjugations of the SNAP®-ADAR gRNAs, only a very high excess of compound **17-NHS** (~1800 eq.) lead to a sufficient formation (50 %) of the desired GalNAc RESTORE gRNA conjugate. However, a high excess of compound **17-NHS** (~200 eq.) is also necessary to provide an acceptable product formation (~50 %) of the desired GalNAc SNAP®-ADAR gRNA conjugate. Therefore, and according to the observed



Scheme 11: General workflow of the conjugation of NH_2-terminal gRNAs to compound 17. *The primary amine is located at the 5'-end of the gRNAs. (i) NHS, DIC, DIPEA, DMSO, 45 °C / overnight; (ii) NH_2-terminal gRNA, gRNA labeling buffer (pH 8.3), DMSO, RT / overnight.*

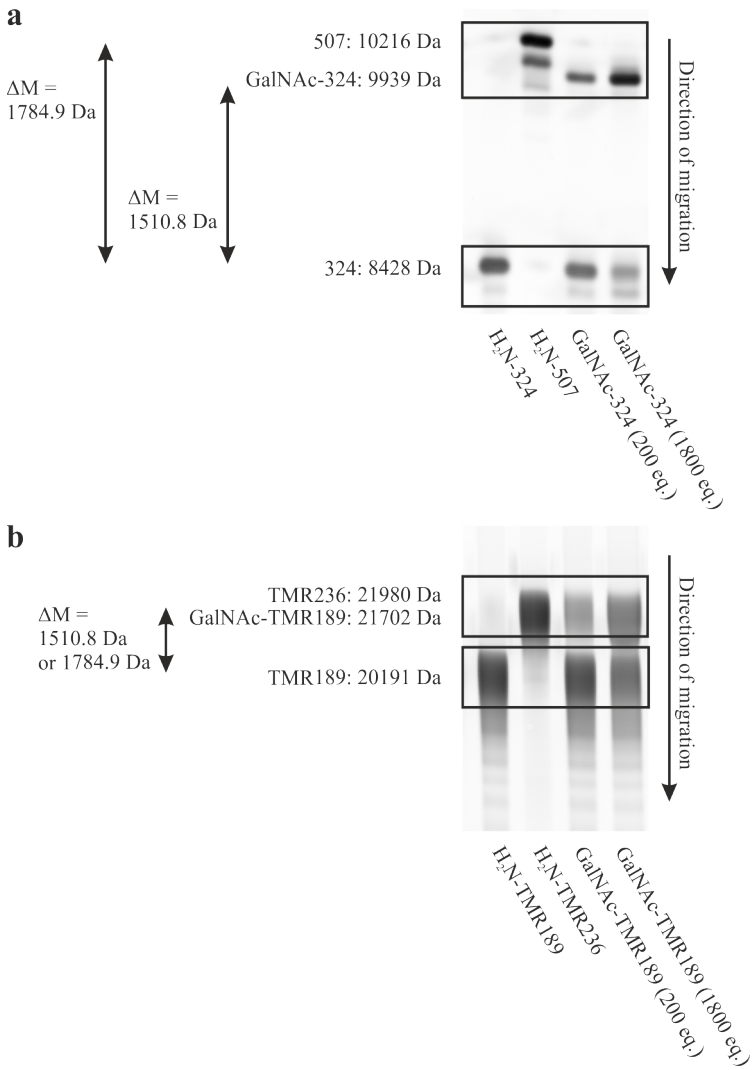


Figure 12: Fluorescence imaging of GalNAc conjugated gRNAs 324 and TMR189. *The gRNAs were conjugated to the GalNAc-COONHS ester (**17-NHS**) and separated using a denaturing TBE-7 M Urea-PAGE in an analytical scale. (a) Conjugation of compound 17-NHS to gRNA 324 (30 pmol, M = 8427.9 Da, 25 nt's). (b) Conjugation of compound **17-NHS** to gRNA TMR189 (30 pmol M = 20191 Da, 59 nt's). As references, commercially available 3'-GalNAc modified (507, TMR236) were cast to the gel. The conjugation of compound 17-NHS to the gRNAs resulted in the formation of a single band. For visualization, the contrast of each gRNA setting was adjusted differently. The changes of the molecular mass of the oligonucleotides differs between the commercially available and the synthesized GalNAc compound (ΔM (GalNAc, manufacturer) = 1784.9 Da; ΔM (GalNAc, **17**) = 1510.8 Da), due to a different chemical structure. The oligonucleotides were stained using SYBR™ Gold Nucleic Acid Gel Stain and for a uniform contrast adjustment and uncropped image sections, see supplementary information Figure S6.*

migration, a successful conjugation of the GalNAc-COONHS ester (17-NHS) to the described NH_2-terminal gRNAs was assumed. Especially for the use of RESTORE gRNAs a high excess of the active ester (**17-NHS**) is required for a sufficient product formation, which is a major disadvantage of the performed modification. This is in accordance with the previously described results by Østergaard *et al.* about the synthesis and evaluation of 5'-GalNAc modified antisense oligonucleotides using a NHS activated and therefore carboxylic acid terminal GalNAc derivative[274]. Additionally, the described yields are qualitatively assumed based on the

signal intensity and are not in any relation to an isolated yield and due to the small scale approach, the product was also not characterized via mass spectrometry and a further investigation is necessary.

In summary, it was possible to provide a novel and highly selective, as well as suitable and wet chemical method to conjugate disulfide containing SNAP®-ADAR gRNAs to the synthesized GalNAc-maleimide derivative (**16**) and the product formation was confirmed via mass spectrometry. The low yield (~10 %) is most probably attributed to a combination of the short design, the use of the described desalting spin columns and the application of a denaturing TBE-7 M Urea-PAGE for separation. Moreover, previously BisBG conjugated SNAP®-ADAR gRNAs were also purified via Urea-PAGE with an average yield of 10 % to 40 %. Therefore, considerable expenses are accountable to the overall and very low yield of 1 % to 4 %, which are related to the successive modification of BisBG and GalNAc conjugates. Furthermore, and due to the need of a high excess, the conjugation of a GalNAc-COONHS ester (**17-NHS**) to NH_2-terminal gRNAs was additionally demonstrated as less beneficial. Thus, the preparation of a sufficient amount of GalNAc conjugated gRNA for cell culture experiments is providing the major challenge and as described before, GalNAc modified oligonucleotides became commercially available during this thesis. Therefore, commercially available and the 3'-conjugated GalNAc gRNAs 507 and TMR236 were obtained from the manufacturer and used for further investigations of a receptor mediated uptake of gRNAs into ASGPR expressing cell lines.

3.2. Molecular cloning of the ASGPR and generation of ASGPR expressing cell lines

3.2.1. Isolation, molecular cloning and proof of concept of the ASGPR

Besides the synthesis of the triantennary GalNAc, its functionalization and therefore its conjugation to gRNAs, the isolation and molecular cloning of the ASGPR was necessary to establish a receptor mediated uptake of gRNAs into different non-hepatic cell lines. Starting from HepG2 cells, a *human* hepatocellular carcinoma cell line expressing the ASGPR, the total RNA was isolated according to a standard procedure of our laboratory using TRI Reagent®. The subsequent DNAse I digestion and reverse transcription (RT) provided the complementary DNA (cDNA) of the receptor's mRNA transcripts, which were amplified applying Phusion PCR (Ph-PCR). As described before, the ASGPR is expressed in two different subunits (H1 and H2), whereby each subunit is spliced into different isoforms (H1a, H1b, H2a, H2b, and H2c). The primer set, used for the RT and Ph-PCR, are designed based on the full-length subunits (H1a and H2a) and the sequences of the different subunits (spliced and unspliced) were

obtained from the National Center for Biotechnology Information (NCBI) and are referred to the NCBI reference sequences as listed in Table 5.

Table 5: NCBI reference sequences of the different ASGPR subunits and isoforms.

Isoform	NCBI reference sequence	Isoform	NCBI reference sequence
H1a	NM_001671.4	H2a	NM_001181.4
H1b	NM_001197216.2	H2b	NM_001201352.1
		H2c	NM_080913.3

The primer design was also used to introduce the desired restriction sites, which were necessary for the molecular cloning, Kozak consensus sequences as well as stop codons. The PCR amplicons were cloned into pcDNA™3.1(+) vectors using standard molecular cloning techniques (restriction digestion and ligation), followed by the transformation into $CaCl_2$ competent *E.Coli* XL1-blue. The subsequent plasmid isolation provided the desired vectors, which were analyzed using Sanger sequencing to confirm the type of receptor subunit and to determine the obtained isoform. For a detailed procedure or an overview about RNA isolation, RT, amplification, restriction enzymes, primer sequences (cloning and sequencing), or resistances see Methods and Materials, Table 15. The sequencing results were aligned to the corresponding NCBI reference sequences to compare the inserts with the sequences of the different receptor isoforms (Figure 13). Within the isolated vectors, three different isoforms were observed, H1a, H2b, and H2c. The insert sequences of pTS689, pTS690, and pTS691 are similar to the sequences of isoforms H1a, H2b and H2c, respectively. While the natural function of H2c remains unknown, H1a and H2b are reported as the main functional isoforms of the ASGPR. As described before, a heterooligomer of the two isoforms H1a and H2b, with a ratio of 2:1 (H1a:H2b), is reported as the most abundant form of membrane bound ASGPRs. Therefore, it was possible to provide the two main functional isoforms of the ASGPR via RNA isolation, molecular cloning and plasmid isolation.

To verify the capability of the ASGPR to internalize GalNAc conjugated derivatives and to confirm the functionality of the previously synthesized triantennary GalNAc, the main functional receptor subunits were ectopically expressed within a standard cell culture cell line. Therefore, the subunits H1a and H2b were either single transfected or co-transfected transiently into a wild-type HEK 293T cell line. After transfection, the cells were incubated with 1 µM GalNAc-FITC (**15**) and the cells were analyzed using live cell imaging and fluorescence microscopy. For a detailed procedure, see methods and materials, section 6.8.3.3.

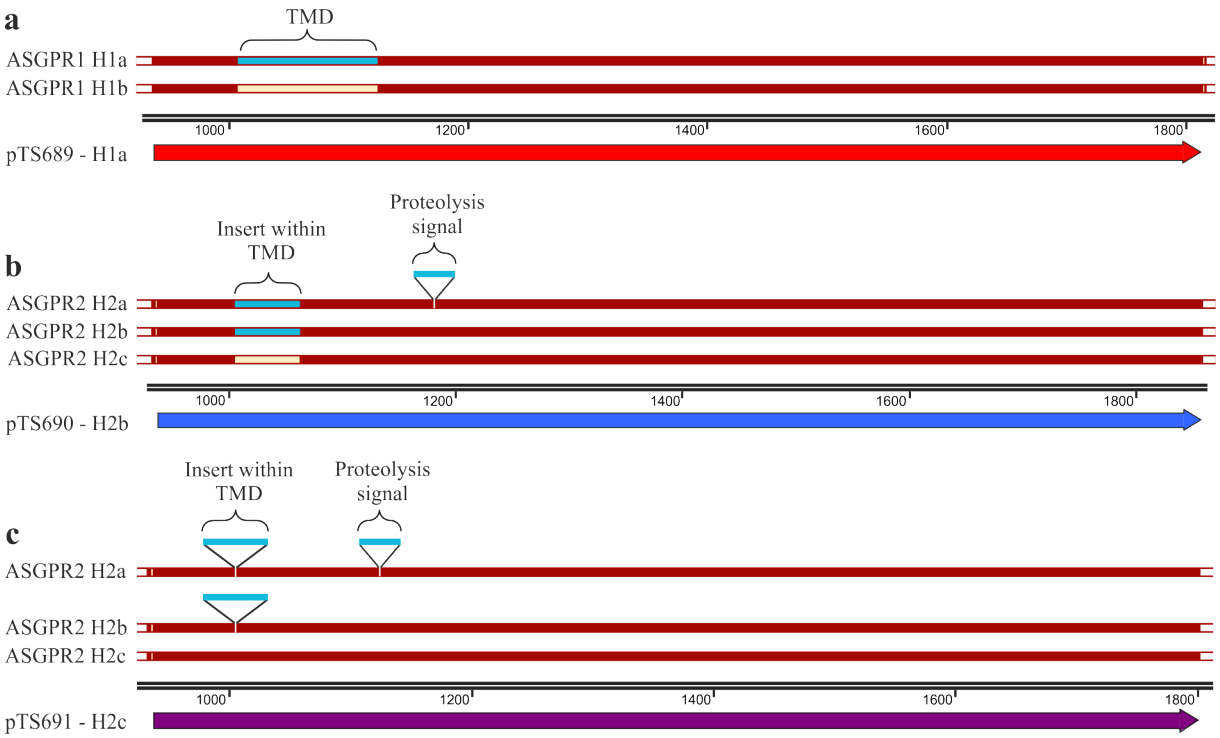


Figure 13: Sequence alignment of the different ASGPR isoforms the the vector sequences. *(a) Sequence alignment of the H1 isoforms to pTS689. H1a is translated from the full-length mRNA, whereby isoform H1b is lacking the TMD. The sequence of pTS689 is corresponding to the sequence of subunit H1a. (b) and (c) Sequence alignments of the H2 isoforms to pTS690 and pTS691, respectively. H2a comprises the full-length transcript, including 19 aa insert within the TMD and a five aa insert between the TMD and the ectodomain, which serves as a proteolytic cleavage signal. H2b is lacking the five aa insert and is not proteolytically active. H2c is lacking a 19 aa insert within the TMD as well as the five aa insert. The sequence of pTS690 is corresponding to the sequence of subunit H2b (b) and the sequence of pTS691 is corresponding to the sequence of subunit H2c (c). TMD = Transmembrane domain. Sequence conformity is illustrated in red-brown, missing inserts in beige, and additional inserts in light blue.*

The results of the fluorescence microscopy are shown in Figure 14b, and an interesting pattern was observable. While the co-transfection of both receptor subunits showed an expected and successful internalization of GalNAc-FITC (**15**), the transfection of single receptor subunits revealed a heterogeneous internalization capability. The single and ectopic expression of receptor subunit H1a was able to internalize GalNAc-FITC (**15**) in a similar fashion as the co-transfected subunits, whereby the ectopic expression of H2b alone showed no internalization. The results indicated the importance of the receptor subunit H1a for a successful internalization. However, due to the capability of subunit H1a to internalize GalNAc-FITC (**15**) without the presence of subunit H2b, it is not possible to confirm a successful co-transfection of both receptor subunits without isolating the RNA of the subunit transcripts or using immunofluorescence-based assays, such as fluorescence microscopy or western blot analysis.

Within the framework of this thesis the potential of the single receptor subunit H1a to internalize GalNAc conjugated ASOs was also reported by Scharner *et al.*[269]. A transduced and

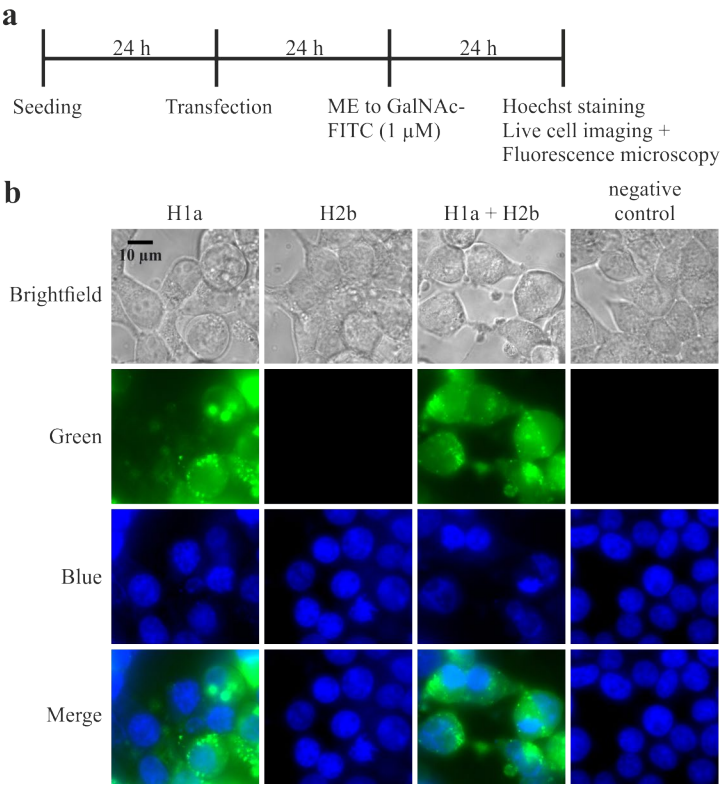


Figure 14: Live cell imaging of the transient transfection of the ASGPR into HEK 293T. *(a) General workflow of the experiment. $3{\cdot}10^4$ cells/well were seeded into poly-D-lysine HBr_{aq} coated 96-well imaging plates and incubated for 24 h. The adherent cells were transfected for 24 h with ASGPR H1a (100 ng), H2b (100 ng) or H1a + H2b (50 ng + 50 ng) using Lipofectamine™ 2000 (4 µL/µg) in OptiMEM™, followed by an incubation for 24 h with 1 µM GalNAc-FITC (**15**). The nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed using fluorescence microscopy. As negative control, no vector was used for the transfection. (b) Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC (**15**) using a 63x magnification. Within a single light channel, similar exposure times and intensities are applied for the different transfection conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree. ME = Media exchange.*

therefore receptor subunit expressing U87 glioblastoma cell line was investigated regarding the internalization of GalNAc conjugated and non-conjugated splice modulating ASOs targeting *SMN2* (Survival of motor neuron 2) transcripts. In addition to the capability of subunit H1a to internalize GalNAc conjugated substances, and next to expected very bright and concentrated FITC signals within fluorescence microscopy, which are most reasonably endosomal-like structures, a homogenous and uniform intracellular distribution was observed. In general, GalNAc-FITC (**15**) is a molecule with very hydrophilic properties and a reduced membrane permeability is thus expected. This is also in accordance with the missing FITC signal within the negative control, whereby non-transfected cells were incubated with GalNAc-FITC (**15**) similar to the transfected ones. Therefore, the intracellular availability was also very promising concerning the endosomal release of internalized GalNAc conjugates. However, oligonucleotides are molecules with highly different physico-chemical properties and only minor amounts of siRNAs and ASOs are reported to undergo an endosomal release to provide

a sufficient intracellular availability. Compared to the internalization of GalNAc-FITC (**15**), a different cellular availability is thus expected for the ASGPR mediated uptake of GalNAc conjugated gRNAs.

However, the RNA isolation of HepG2 cells, the molecular cloning of the desired inserts and the subsequent plasmid isolations, provided three different isoforms of the ASGPR, including the two main functional isoforms H1a and H2b. A fluorescence microscopy-based assay and an ectopic expression of the receptor subunits confirmed the ability to internalize the previously synthesized GalNAc-FITC (**15**) conjugate and a heterogeneous uptake capability of the different receptor subunits was discovered, which was simultaneously reported in literature. Based on these results, receptor subunit and isoform H1a was used exclusively for further experiments and the generation of ASGPR expressing cell lines.

3.2.2. The Generation of ASGPR expressing cell lines

To investigate an ASGPR mediated uptake of gRNAs into non-hepatic cell lines, it was necessary to create stable ASGPR expressing cell lines and to provide GalNAc conjugated gRNAs. For the use of the described SNAP®-ADAR editing system, a FlpIn™ T-REx™ 293 cell line was generated, which is inducibly expressing the artificial SNAP®-ADAR1 E406Q (SA1Q) fusion protein as well as the main functional receptor subunit and isoform H1a. In contrast, and for the recruitment of endogenous ADARs utilizing the described RESTORE gRNAs, it is not necessary to express the SNAP®-ADAR fusion protein. For this, the PiggyBac transposon system was used to integrate a stable and inducible receptor isoform H1a into a wild-type Hela cell line.

3.2.2.1. The Generation of a SNAP®-ADAR1 E406Q and ASGPR expressing FlpIn™ T-REx™ 293 cell line

For the generation of a SNAP®-ADAR and ASGPR expressing FlpIn™ T-REx™ 293 cell line, it was primarily necessary to provide a pcDNA™5/FRT vector, which is containing both, the SA1Q fusion protein as well as the receptor isoform H1a. Therefore, a bidirectional pcDNA™5/FRT expression vector was designed for the stable integration of the desired constructs (Figure 15a). The used vector backbone as well as the bidirectional promotor construct were provided by Anna Stroppel and its generation and characterization is reported in detail by Stroppel *et al.*[119]. The SA1Q and H1a expressing vector was generated using standard molecular cloning techniques (restriction digestion and ligation) and the three desired constructs (SA1Q, H1a and promotor region) were introduced within a single ligation. SA1Q,

ASGPR H1a as well as the bidirectional promotor construct were previously amplified using Ph-PCRs and the desired restriction sites (NotI/PacI, AvrII/ClaI, AvrII/NotI, respectively) were integrated simultaneously. The ligation was transformed into $CaCl_2$ competent *E.Coli* XL1-blue using heat-shock transformation, and the subsequent plasmid isolation provided the desired vector. The bidirectional and inducible constructs were confirmed utilizing Sanger sequencing and a detailed overview about the insert origins and the used primer sets is described in Table 15. The design of the bidirectional expression vector is also based on results of Stroppel *et al.*[119], whereby a slight leaky expression was observed for the uninduced EF1α-core promoter. While a long-term expression of the receptor isoform H1a appeared to be noxious for a least FlpIn™ T-REx™ 293 cells (see supplementary information Figure S7), the receptor subunit H1a is therefore expressed under the control of the inducible CMV promoter and the SA1Q construct is expressed under the control of the inducible, but leaky EF1α-core promoter.

The obtained bidirectional pcDNA™5/FRT expressing vector was stable transfected into a FlpIn™ T-REx™ 293 host cell line, which is containing the integrated Flp Recombination Target (FRT) site, using Lipofectamine™ 2000 according to the manufacturer's protocol (Figure 15b). The transfected cells were selected with Blasticidin S and Hygromycin B for 14 days and characterized utilizing receptor mediated uptake of GalNAc-FITC (**15**), immunofluorescence imaging, as well as Western Blot analysis. The results of the different characterizations are showed in Figure 15d as well as Figure 16b and c.

First, the generated cell line was analyzed concerning their capability to internalize GalNAc-FITC (**15**). The cells were seeded into coated 96-well imaging plates and incubated with and without doxycycline induction for 24 h. Afterwards, the cells were further incubated with 1 μM GalNAc-FITC (**15**) and the cells were analyzed using live cell imaging and fluorescence microscopy (Figure 15). For a detailed procedure, see methods and materials, section 6.8.3.3. In accordance with previous results (Figure 14), a uniform and intracellular distribution of GalNAc-FITC (**15**) was only observed under doxycycline induction and therefore under ASGPR H1a expression. Furthermore, very bright and concentrated FITC signals were observed as well, which are most reasonably attributed to endosomal-like structures (Figure 15d). In contrast to the expectations, and especially within the use of a lower magnification (10x), it was shown that not all cells were transfected, and therefore integrated successfully. In general, by utilizing the FlpIn™ system a resistance gene (*HygR*) is co-integrated next to the gene(s) of interest (GOI), which is protecting the integrated cells from cellular death during selection. Therefore, the presence of non-integrated cells is not in

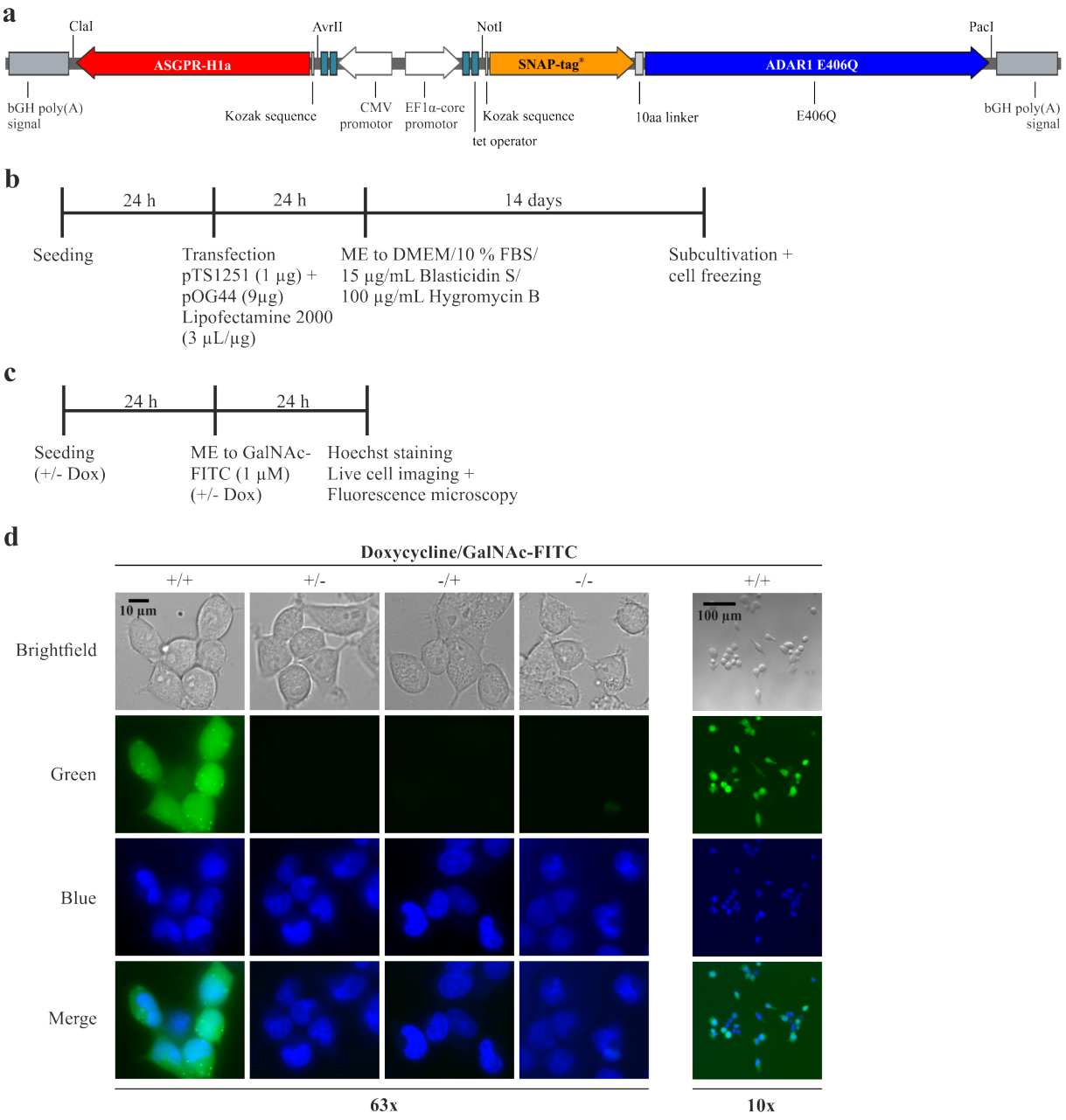


Figure 15: Generation of a SA1Q and H1a stably expressing FlpIn™ T-REx™ 293 cell line. *(a) Illustration of the bidirectional construct used for the stable integration into a FlpIn™ T-REx™ 293 host cell line. (b) General workflow of the stable transfection of FlpIn™ T-REx™ 293 cells. 4•10^6 cells were seeded into a 10 cm cell culture dish and incubated for 24 h. The adherent cells were co-transfected for 24 h with the pcDNA™5/FRT expression vector containing the GOI (1 µg) and the Flp-Recombinase expression vector (pOG44) (9 µg) using Lipofectamine™ 2000 (3 µL/µg) in OptiMEM™. After transfection, the cells were selected with Blasticidin S (15 µg/mL) and Hygromycin B (100 µg/mL) for 14 days. Afterwards, the cells were subcultured, prepared for long-time storage in liquid nitrogen or used for the desired cell culture experiments and characterizations. (c) General workflow of the internalization of GalNAc-FITC (**15**) of FlpIn™ T-REx™ 293 cells. 1.5·10^4 cells/well were seeded into poly-D-lysine HBr_{aq} coated 96-well imaging plates and incubated for 24 h with and without doxycycline induction (10 ng/mL). The adherent cells were further incubated for 24 h with 1 µM GalNAc-FITC (**15**). The nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed using fluorescence microscopy. (d) Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC (**15**) using a 63x magnification (left). Additionally, doxycycline induced (+) and GalNAc-FITC treated cells (+) are shown using a 10x magnification (right). An overview of all samples using the 10x magnification is shown in supplementary information Figure S8. Within a single light channel and magnification, similar exposure times and intensities are applied for the different*

conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree. ME = Media exchange, aa = amino acid.

accordance with the basic mechanism and another, but unknown factor is reasonable to assume. While bGH poly(A) signals are involved in the termination of both integrated genes (SA1Q and H1a, see Figure 15a), biological processes, such as a homologous recombination could only be one possibility for the heterogeneous integration. However, the internalization of the synthesized GalNAc-FITC (**15**) provided promising results and the generated cell line was further characterized utilizing immunofluorescence imaging as well as Western Blot analysis.

For immunofluorescence imaging, the cells were seeded onto coated cover glasses and incubated with and without doxycycline induction for 24 h. The expressed SA1Q was stained using monoacetylated BG-FITC followed by fixation and immunofluorescent staining using an antibody against the receptor subunit H1 (*Mouse* α-ASGPR1) and an Alexa Fluor™ 647 conjugated secondary antibody (*Goat* α-*Mouse* Alexa Fluor™ 647). The mounted cells were analyzed by fluorescence microscopy and the results are shown in Figure 16. Monoacetylated BG-FITC is a cell membrane permeable conjugate of monoacetylated FITC and O^6-benzylguanine (BG), which is covalently binding to SNAP®-tag proteins, similar to BG or BisBG modified gRNAs. The use of monoacetylated BG-FITC is a well-established procedure within our research group to detect artificial SNAP®-tag fusion proteins.

As expected, and in accordance with the results of the receptor mediated uptake of GalNAc-FITC (**15**), a heterogeneous expression of the desired SA1Q and H1a is observed within the results of the immunofluorescence imaging. In contrast, a BG-FITC stained FlpIn™ T-REx™ 293 cell line, which is exclusively expressing SA1Q, showed a ubiquitously localized and homogenously distributed signal of the FITC conjugated SA1Q fusion protein (Figure 16c) within almost all cells. A ubiquitous localization of the FITC conjugated SA1Q signal was also observed within positively integrated cells, which showed an additional mCherry (red) signal of the Alexa Fluor™ 647 stained receptor subunit H1. The cells were further compared to a naturally ASGPR expressing HepG2 cell line, using the same antibody against the receptor subunit H1 (*mouse* α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (*goat* α-*mouse* Alexa Fluor™ 488). Here, a comparable distribution of the integrated receptor subunit was observed via immunofluorescence imaging (Figure 16b and d). Compared to the background fluorescence, which is related to BG-FITC binding to random proteins or cell compartments, a more intense FITC signal is observed for the uninduced SA1Q and H1a expressing cells (Figure 16b green) than for the uninduced SA1Q cells (Figure 16c, green). This

is in accordance with the previously reported results of a leaky expression of the uninduced EF1α-core promoter[119] and is further confirming the successful integration of the desired and bidirectional construct.

For Western Blot analysis, the cells were seeded into 24-well cell culture plates and incubated as before with and without doxycycline induction for 24 h. The induced and non-induced cells were lysed using RIPA Lysis and extraction buffer (supplemented with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail) overnight at -80 °C and the total amount of protein was determined using a Pierce™ BCA Protein Assay Kit with BSA as reference (0-1.5 mg/mL). For SDS-PAGE, 30 µg whole cell lysates were cast onto Novex™ WedgeWell™ 8-16 % (tris-glycine) Mini Protein Gels and the separated proteins were transferred onto PVDF membranes. The desired proteins were stained using *rabbit* α-SNAP (1:1000), *rabbit* α-GAPDH (1:1000) and *mouse* α-ASGPR1 (1:500) antibodies and for detection, *goat* α-*mouse* HRP (1:5000) or *goat* α-*rabbit* HRP (1:5000) antibodies were used. For detailed information about sample preparation, protein separation, blotting, and visualization, see section 6.8.4.

The results of the Western Blot analysis are shown within Figure 16e and while the signals and intensities of the SA1Q signals are in accordance with the expectations, the detection of the ASGPR H1 subunit was not successful. As described before, and as seen via the internalization of GalNAc-FITC (**15**) and immunofluorescence imaging, an incomplete integration of the bidirectional construct was observed. This is reflected within the SA1Q signal intensities of the doxycycline induced samples (+ dox), whereby a less intense signal is observed for the SA1Q and H1a expressing cell line than for the exclusively SA1Q expressing ones. In addition, a weak signal is observed for the SA1Q and H1a expressing cell line without doxycycline induction (- dox), which is in accordance with the leaky expression of SA1Q under the control of the uninduced EF1α-core promoter. As expected, no SA1Q signal is observed for the uninduced SA1Q expressing cell line under the control of an inducible CMV promotor. However, the expression of the ASGPR subunit H1a of the generated SA1Q and H1a expressing cell line can not be confirmed via Western Blot analysis. While a protein size of 46 kDa is reported for the matured and glycosylated receptor subunit H1a[277], a weak signal is observed at an appropriate position within all samples. This is not in accordance with the expectations at least for the exclusively SA1Q expressing cell line. Therefore, the observed signal is likely not related to the expression of the ASGPR subunit H1a and can be most probably attributed to an unkown or random antibody binding.

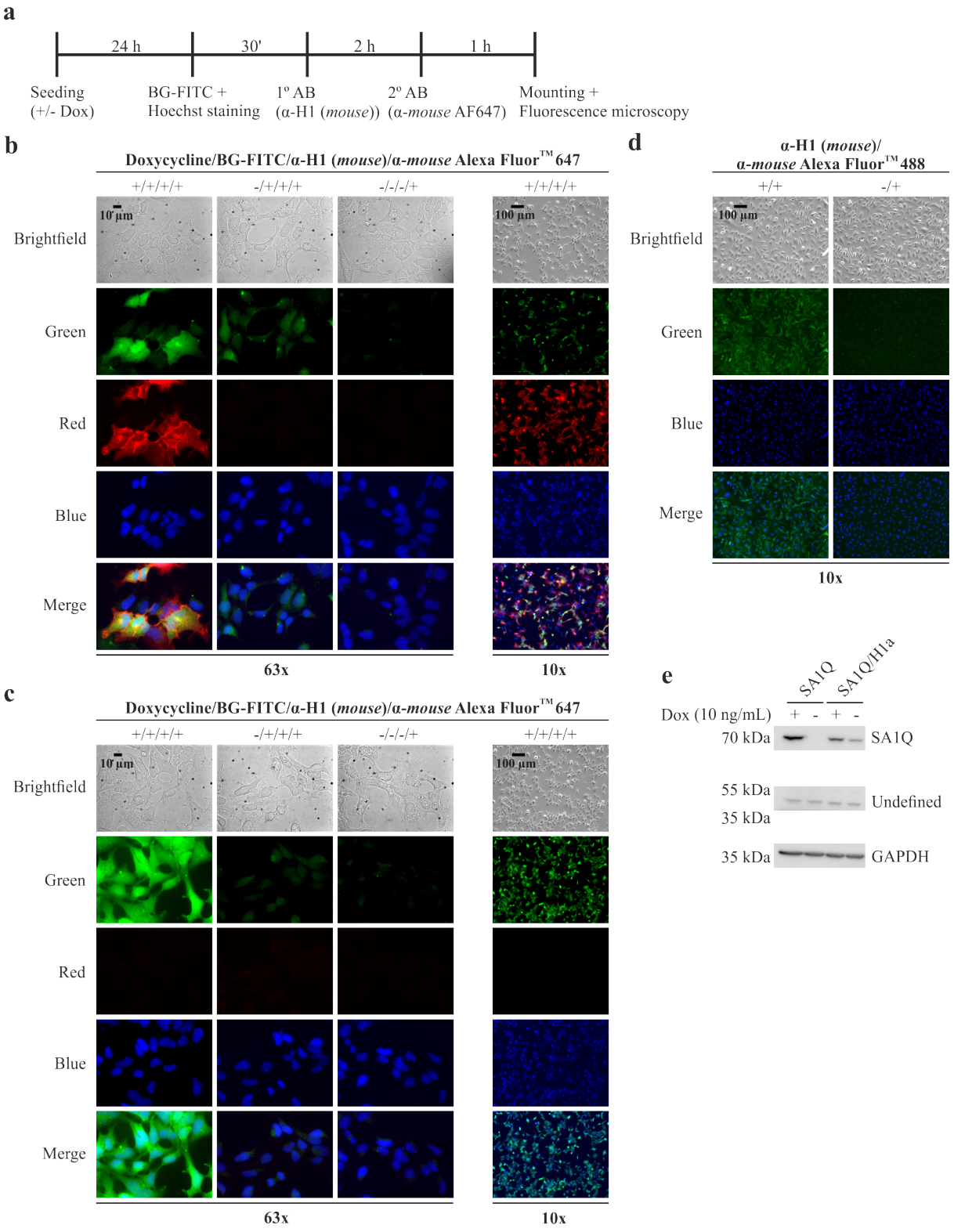


Figure 16: Charazerization of a SA1Q and H1a stably expressing FlpIn™ T-REx™ 293 cell line *(a) General workflow of the immunofluorescence imaging of FlpIn™ T-REx™ 293 cells. 1.5·10^5 cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h with and without doxycycline induction (10 ng/mL). The expressed SA1Q of the adherent cells was stained with monoacetylated BG-FITC and the nuclei were stained with NucBlue™ Live ReadyProbes, simultaneously. The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 647 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 647). The cells were mounted and analyzed by fluorescence microscopy. (b) Fluorescence imaging of the immunofluorescence staining*

of the SA1Q and H1a expressing FlpIn™ T-REx™ 293 cells using a 63x magnification (left). Additionally, doxycycline induced stained cells are shown using a 10x magnification (right). (c) Fluorescence imaging of the immunofluorescence staining of the SA1Q expressing FlpIn™ T-REx™ 293 cells using a 63x magnification (left). Additionally, doxycycline induced stained cells are shown using a 10x magnification (right). An overview of all samples using the 10x magnification is shown in supplementary information Figure S9. (d) Fluorescence imaging of the immunofluorescence staining of HepG2 cells using a 10x magnification. $1 \cdot 10^5$ cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h. The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed by fluorescence microscopy. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and within a single experiment, the contrasts of all FITC signals (green) are adjusted to a similar degree. (e) Western Blot analysis of the of SA1Q and H1a expressing FlpIn™ T-REx™ 293 compared to the SA1Q expressing FlpIn™ T-REx™ 293. $3 \cdot 10^5$ cells/well were seeded into a 24-well plate and incubated for 24 h with and without doxycycline induction (10 ng/mL). The cells were lysed using RIPA Lysis and extraction buffer (supplemented with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail) overnight at -80 °C and the total amount of protein was determined using a Pierce™ BCA Protein Assay Kit with BSA as reference (0-1.5 mg/mL). 30 µg whole cell lysates were used for western blotting. Subsequently, SNAP®-ADAR, GAPDH and ASGPR H1 were stained and detected using rabbit α-SNAP (1:1000), rabbit α-GAPDH (1:1000) and mouse α-ASGPR1 (1:500), respectively. For detection, goat α-mouse HRP (1:5000) or goat α-rabbit HRP (1:5000) were used. SNAP®-ADAR and GAPDH were detected simultaneously, while ASGPR was stained and detected separately and after the detection of SNAP®-ADAR and GAPDH. Full images are shown within supplementary information Figure S10. Dox = Doxycycline, HRP = horseradish peroxidase.

In summary, based on the described results, the generation of a novel SNAP®-ADAR1 E406Q and ASGPR H1a expressing FlpIn™ T-REx™ 293 cell line was successful and the results of the different performed characterizations are consistent with each other. Thus, the generated cell line is able to internalize GalNAc derivatives (GalNAc-FITC (**15**)) in a promising degree. While the immunofluorescence staining of the artificial SNAP®-ADAR1 fusion protein as well as ASGPR H1a was successful, it was not able to detect the receptor subunit H1 within Western Blot analysis using the same antibody. However, compared to an exclusively SA1Q expressing FlpIn™ T-REx™ 293 cell line, a lower degree of integration is observed within multiple fluorescence microscopies. Only a 10x magnification allows for the observation of a predominant expression of both, SA1Q and receptor subunit H1a, within the different setups. The distribution of the receptor subunit was additionally compared to naturally ASGPR expressing cells and comparable and very promising results were obtained. As described before, the generation of a novel SNAP®-ADAR1 E406Q and ASGPR H1a expressing FlpIn™ T-REx™ 293 cell line was thereby confirmed. However, quantitative evaluations are very challenging utilizing optical techniques, such as fluorescence imaging, and while no reasonable results were obtained via Western Blot analysis, a further evaluation of the generated FlpIn™ T-REx™ 293 cell line is necessary. Besides the transfection of gRNAs to further evaluate the functionality of the generated cell line, a particular interest of the novel SNAP®-ADAR1 E406Q and ASGPR H1a expressing cell line is to investigate the ASGPR mediated endocytosis of BisBG gRNAs to induce SNAP®-ADAR mediated RNA editing.

3.2.2.2. The Generation of an ASGPR expressing Hela cell line

For the generation of an ASGPR H1a expressing HeLa cell line, utilizing the PiggyBac transposon system, and to induce the receptor mediated uptake of RESTORE gRNAs, it was primarily necessary to provide an inducible XLone PiggyBac vector[278] that contains the receptor isoform H1a. The used XLone PiggyBac transposon vector backbone is commercially available from Addgene (Watertown (MA), USA) and the desired H1a expression vector was generated using standard molecular cloning techniques (restriction digestion and ligation) with two successive cloning steps. Prior to the integration of the receptor isoform H1a, it was necessary to exchange the resistance gene from Blasticidin (*BSD*) to Neomycin (*NeoR*) to enable the selection with Geniticin™ (G418). As before, the constructs of the ASGPR subunit H1a as well as the *NeoR* resistance gene were previously amplified using Ph-PCRs and the desired restriction sites (NotI/PacI and AgeI/AvrII, respectively) were integrated simultaneously. Each ligation was transformed into $CaCl_2$ competent *E.Coli* XL1-blue using heat-shock transformation, where the subsequent plasmid isolation provided the desired vector. The integration of the desired construct was confirmed utilizing Sanger sequencing and a detailed overview of the insert origins and the used primer sets is described in Table 15. The final XLone PiggyBac transposon vector containing both, the receptor subunit H1a and the *NeoR* resistance gene, as well as the transposase containing vector (pTS687), was co-transfected into a wild-type HeLa cell line, using FuGENE® 6 according to the manufacturer's protocol (Figure 17a). The transfected cells were selected with Geniticin™ (G418) for 8 days and characterized utilizing receptor mediated uptake of GalNAc-FITC (**15**), immunofluorescence imaging, as well as Western Blot analysis. The results of the different characterizations are shown in Figure 17c, Figure 18b as well as Figure 19c and d.

Similar to the generated FlpIn™ T-REx™ 293 cell line, the generated HeLa cell line was analyzed concerning their capability to internalize GalNAc-FITC (**15**). The cells were seeded into coated 96-well imaging plates and incubated with and without doxycycline induction (200 ng/mL) for 24 h. Afterwards, the cells were further incubated for 24 h with 1 µM GalNAc-FITC (**15**) and the cells were analyzed using live cell imaging and fluorescence microscopy (Figure 17c). For a detailed procedure, see methods and materials, section 6.8.3.3. In contrast, and while a doxycycline concentration of 10 ng/mL is sufficient to induce FlpIn™ T-REx™ 293 cells, a higher concentration of >100 ng/mL is necessary to provide an adequate expression level of the GOI using the Tet-ON® 3G induction system[278]. This was confirmed via Western Blot analysis of two SA1Q and SA2Q expressing HepG2 cell lines, whereby the GOIs were

also integrated using the PiggyBac transposon system and the inducible XLone PiggyBac expression vector (see supplementary information Figure S12).

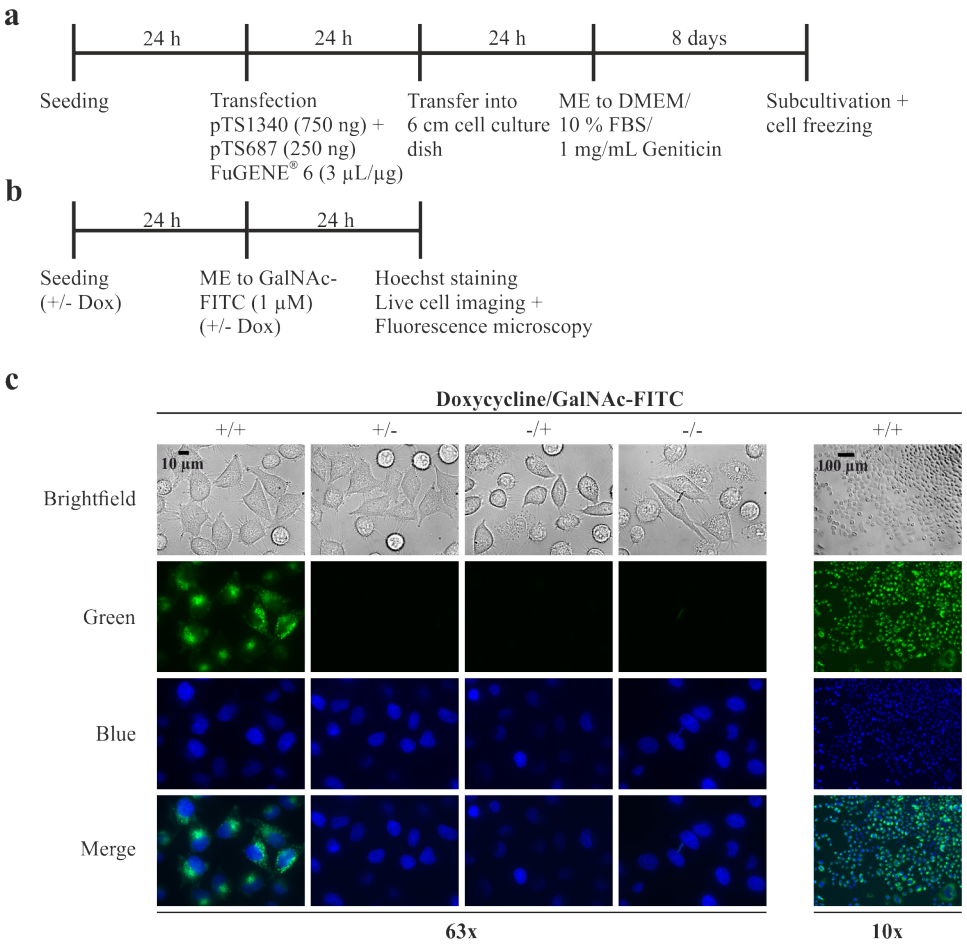


Figure 17: Generation of a H1a stably expressing HeLa cell line. *(a) General workflow of the stable transfection of HeLa cells. 1.5•10^5 cells were seeded into a 24-well cell culture plate and incubated for 24 h. The adherent cells were co-transfected for 24 h with the XLone PiggyBac transposon vector containing the GOI (750 ng) and the transposase containing expression vector (pTS687) (250 ng) using FuGENE® 6 (3 µL/µg) in OptiMEM™. After transfection, the cells were transferred into a 6 cm cell culture dish and incubated for 24 h. The adherent cells were selected with Geniticin™ (G418) (1000 µg/mL) for 8 days, followed by subcultivation, preparation for long-time storage under cryogenic conditions or the use for the desired cell culture experiments and characterizations. (b) General workflow of the internalization of GalNAc-FITC (**15**) into H1a expressing Hela cells. 1•10^4 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well imaging plates and incubated for 24 h with and without doxycycline induction (200 ng/mL). The adherent cells were further incubated for 24 h with 1 µM GalNAc-FITC (**15**). The nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed using fluorescence microscopy and under live cell imaging conditions. (c) Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC (**15**) using a 63x magnification (left). Additionally, doxycycline induced (+) and GalNAc-FITC treated cells (+) are shown using a 10x magnification (right). An overview of all samples using the 10x magnification is shown in supplementary information Figure S11. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree. ME = Media exchange.*

In accordance with the previously described results, a cellular uptake of GalNAc-FITC (**15**) was only observed under doxycycline induction and therefore under ASGPR H1a expression. Additionally, and compared to the integrated FlpIn™ T-REx™ 293 cells, a higher degree of integration was observed as well, with varying signal intensities between different cells. This

is most likely related to a heterogeneous integration of the receptor, which is in accordance with the basic mechanism of the PiggyBac transposon system. While a single copy of the GOI is integrated within each cell using the FlpIn™ T-REx™ system, random numbers of copies are integrated using PiggyBac transposon system[278]. Therefore, a heterogeneous expression level of the receptor subunit H1a and a related heterogeneous internalization capability is most reasonable. Furthermore, no homogenous and intracellular distribution of the FITC signal was observed for the generated HeLa cell line, and exclusively endosomal-like structures were observed.

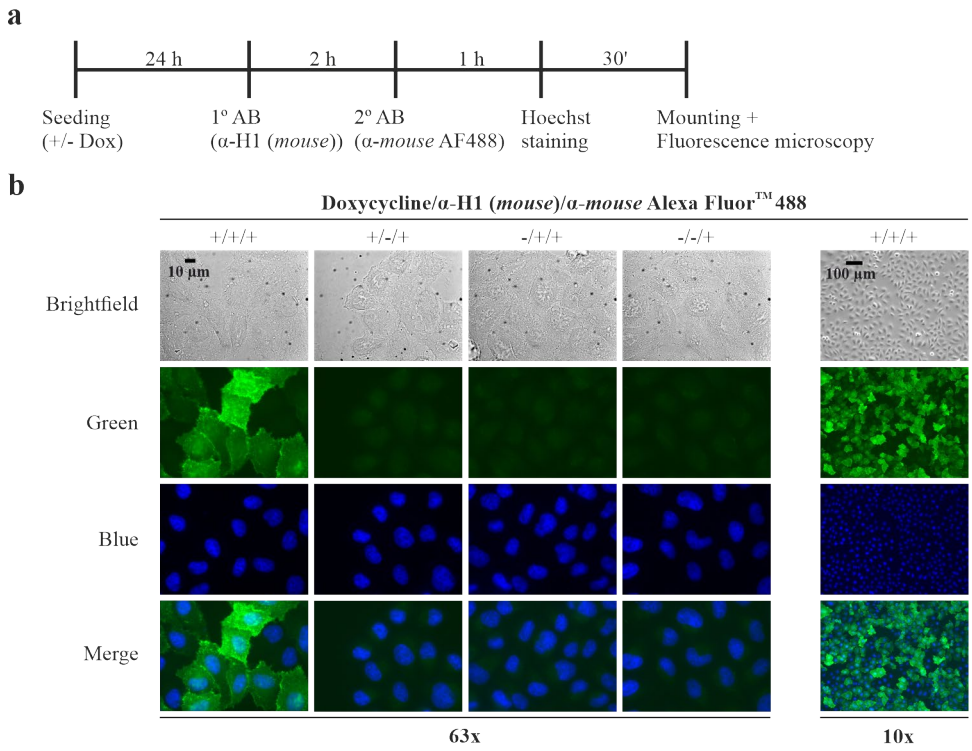


Figure 18: Characterization of a H1a stably expressing HeLa cell line. *(a) General workflow of the immunofluorescence imaging of HeLa cells. $1 \cdot 10^5$ cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h with and without doxycycline induction (200 ng/mL). The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed by fluorescence microscopy. (b) Fluorescence imaging of the immunofluorescence staining of the H1a expressing HeLa cells using a 63x magnification (left). Additionally, doxycycline induced and stained cells are shown using a 10x magnification (right). An overview of all samples using the 10x magnification is shown in supplementary information Figure S13. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all Alexa Fluor™ 488 (green) signals are adjusted to a similar degree.*

A heterogeneous expression level was also observed via immunofluorescence imaging. For this, the cells were seeded onto coated cover glasses and incubated as before with and without doxycycline induction (200 ng/mL) for 24 h. The adherent cells were fixated using *p*-formaldehyde and immunofluorescently labeled using an antibody against the receptor subunit

H1 (*Mouse* α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (*Goat* α-*Mouse* Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed using fluorescence microscopy.

The results are shown in Figure 18b and as expected, a heterogeneous expression level of the receptor subunit H1 was observed within the results of the immunofluorescence imaging. This is in accordance with the previous results of the receptor mediated uptake of GalNAc-FITC (**15**), and thus, the successful integration of an inducible receptor is reasonable to assume. However, and with the objective to generate a cell line which is homogeneously expressing the ASGPR variant H1a, similar to ASGPR expressing hepatoma cells, the cells were sorted utilizing fluorescence activated cell sorting (FACS). Therefore, induced HeLa cells containing the receptor subunit H1a were incubated with GalNAc-FITC (**15**) for 1.5 h and sorted into two different fractions according their FITC signal intensity, a high intensity (FITC pos) and a low intensity (FITC dim) fraction. The corresponding fluorescence imaging of the precursory GalNAc-FITC internalization, as well as the gating strategy and sorting results are shown in Figure 19a and b. The sorted cells were characterized via immunofluorescence imaging in accordance with the previously described setting (Figure 18a). In contrast to the cells before sorting and the low intensity fraction, a homogenous expression level and membrane integration of the receptor was observed for the sorted cells with a high intensity FITC signal (Figure 19d). As before, the cells were further compared to a naturally ASGPR expressing HepG2 cell line, and a comparable expression and distribution of the integrated receptor subunit was observed via immunofluorescence imaging (Figure 19d and e).

Similar to the SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell line, the generated as well as the sorted HeLa cells were further characterized via Western Blot analysis and the results are shown in Figure 19c Additionally, human primary hepatocytes (Lonza Group Ltd, Basel, Switzerland) were used as a reference to interpret the expression level of the generated cell lines. For this, the cells were seeded into 24-well cell culture plates and incubated as before with and without doxycycline induction for 24 h. The induced and non-induced cells were lysed using RIPA Lysis and extraction buffer (supplemented with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail) overnight at -80 °C and the total amount of protein was determined using a Pierce™ BCA Protein Assay Kit with BSA as reference (0-1.5 mg/mL). For SDS-PAGE, 30 µg whole cell lysates were cast onto Novex™ WedgeWell™ 8-16 % (tris-glycine) Mini Protein Gels and the separated proteins were transferred onto PVDF membranes. The desired proteins were stained using *rabbit* α-GAPDH (1:1000) and *mouse* α-ASGPR1 (1:500)

antibodies and for detection, *goat* α-*rabbit* HRP (1:5000) and *goat* α-*mouse* HRP (1:5000) antibodies were used, respectively. For detailed information about sample preparation, protein separation, blotting and visualization, see section 6.8.4.

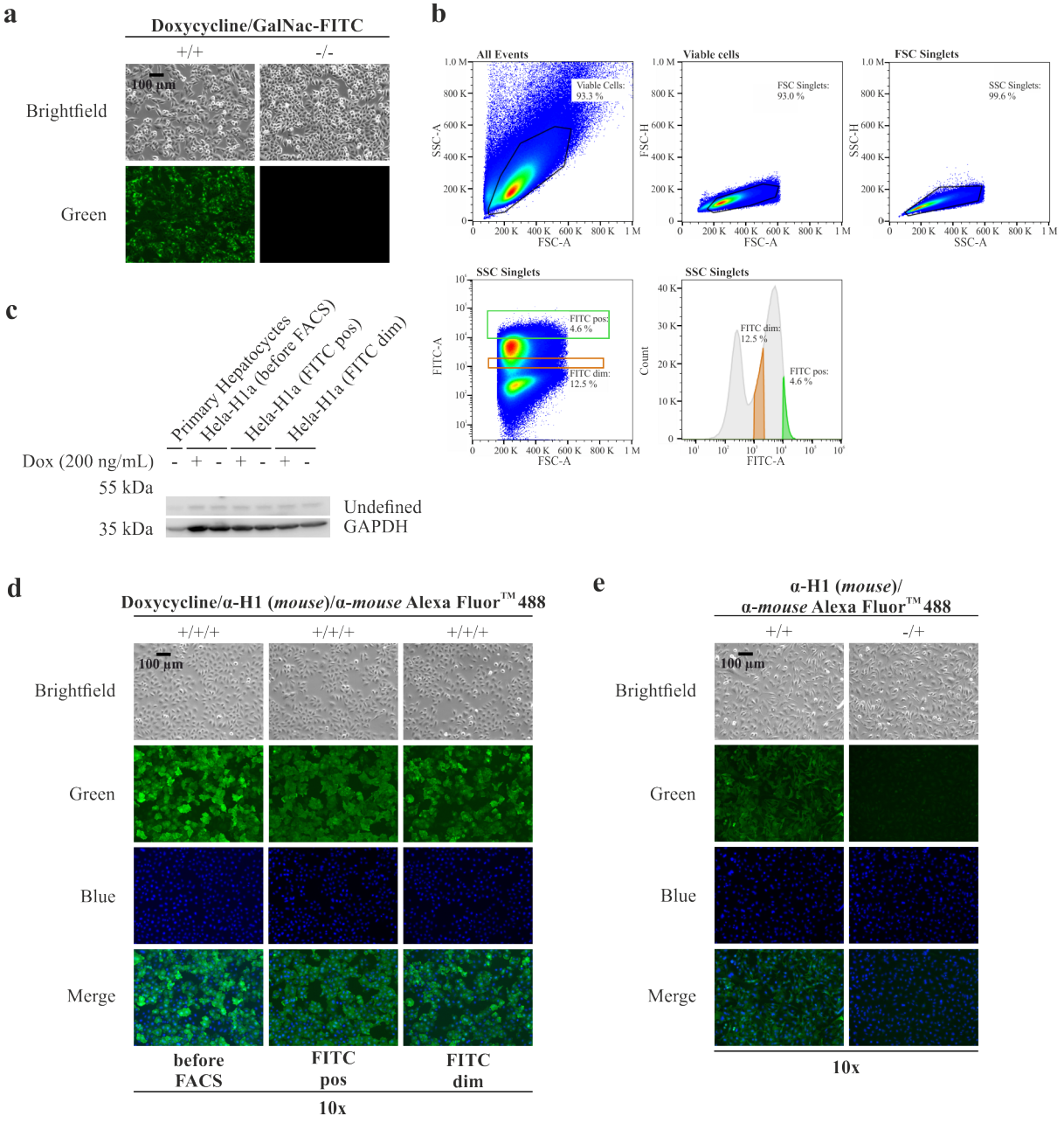


Figure 19: FACS sorting of the H1a stably expressing HeLa cell line. *(a) Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC (**15**) using a 10x magnification before FACS sorting. 4·10^6 cells were seeded into a 10 cm cell culture dish and incubated for 24 h with and without doxycycline induction (200 ng/mL). The induced cells were further incubated for 1.5 h with 1 µM GalNAc-FITC (**15**) and the cells were analyzed using fluorescence microscopy and under live cell imaging conditions. (b) Gating strategy and results of the FACS sorting. Height and area correlations of the forward (FSC) and side scatter (SSC) signals were used to sort the cells concerning their size and granularity to ensure sorting of only viable cells and the cells were further sorted according their FITC signal intensity into two different fractions (FITC pos and FITC dim). For FITC compensation and gating, untreated HeLa cells were used, which were prepared in a similar way. (c) Western Blot analysis of the of the sorted and unsorted ASGPR H1a expressing HeLa cells.7.5•10^5 cells/well were seeded into a 6-well plate and incubated for 24 h with and without doxycycline induction (200 ng/mL). As reference primary hepatocytes were used. 4.5•10^5 cells/well were seeded into a 24-well plate and incubated for 24 h. The*

cells were lysed using RIPA Lysis and extraction buffer (supplemented with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail) overnight at -80 °C and the total amount of protein was determined using a Pierce™ BCA Protein Assay Kit with BSA as reference (0-1.5 mg/mL). 30 µg whole cell lysates were used for western blotting. For protein conjugation, rabbit α-GAPDH (1:1000) and mouse α-ASGPR1 (1:500) were used, and for detection, goat α-rabbit HRP (1:5000) and goat α-mouse HRP (1:5000) were used, respectively. GAPDH and ASGPR were stained and detected separately and full images are shown within supplementary information Figure S10. Dox = Doxycycline, HRP = horseradish peroxidase. (d) Fluorescence imaging of the immunofluorescence staining of the sorted and unsorted H1a expressing HeLa cells using a 10x magnification. An overview of all samples using the 10x magnification is shown in supplementary information Figure S14. The results of the unsorted cells were also previously reported within Figure 18b. $1 \cdot 10^5$ cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h with and without doxycycline induction (200 ng/mL). The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed by fluorescence microscopy. (e) Fluorescence imaging of the immunofluorescence staining of HepG2 cells using a 10x magnification. $1 \cdot 10^5$ cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h. The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed by fluorescence microscopy. These data were also previously shown within Figure 16. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and within a single experiment, the contrasts of all FITC signals (green) are adjusted to a similar degree.

However, no reasonable signal of the receptor subunit was observed via Western Blot analysis. As described before during the Western Blot analysis of the FlpIn™ T-REx™ 293 cells, a faint signal was observed at an appropriate position at around 46 kDa[277] within all samples, including the uninduced samples. While uninduced cells were not able to internalize GalNAc-FITC (**15**), and no signal of the receptor subunit was observed via immunofluorescence imaging, the observed signal of the uninduced samples is not in accordance with the expectations and previous results. A random antibody binding, which was also observed for the SA1Q as well as the SA1Q and H1a expressing FlpIn™ T-REx™ 293 cells, may also be assumed in this case. All samples of FlpIn™ T-REx™ 293 cells, primary hepatocytes as well as all HeLa cells were conducted within a single Western Blot, and preparative issues can not be fully excluded. While reasonable signals were observed for the detection of SA1Q and GAPDH as reference, as well as immunofluorescence imaging using the same antibody, the missing signal of the receptor is most probably related to any incompatibility of the antibody with Western Blot analysis or any unknown preparative issues during lysate preparation or antibody treatment.

In summary, and based on the described results, the generation of a homogeneously ASGPR H1a expressing HeLa cell line was successful and the results of the different performed characterizations (except Western Blot analysis) are consistent with each other. Thus, the generated cell line is able to internalize GalNAc derivatives (GalNAc-FITC (**15**)), and while promising results were obtained via immunofluorescence imaging of the receptor, it was not

able to detect the receptor subunit H1 via Western Blot analysis using the same antibody. FACS sorting further provided a HeLa cell line with a homogeneous expression level of the receptor subunit and a high degree of integration, which was also confirmed via immunofluorescence imaging. The expression level and degree of membrane integration of the receptor subunit was additionally compared to naturally ASGPR expressing HepG2 cells and comparable and very promising results were obtained as well. Therefore, the generation of a homogeneously ASGPR H1a expressing HeLa cell line was confirmed. However, quantitative evaluations are very challenging utilizing optical techniques, such as fluorescence imaging, and while no reasonable results were obtained via Western Blot analysis, a further evaluation of the generated HeLa cell lines is necessary. Besides the transfection of gRNAs to further evaluate the functionality of the generated cell lines, a particular interest of the ASGPR H1a expressing HeLa cell line is to investigate the ASGPR mediated endocytosis of RESTORE gRNAs to induce RNA editing recruiting endogenous ADAR.

3.3. gRNA mediated A-to-I editing of ASGPR expressing cell lines

To induce an ASGPR mediated uptake of gRNAs, it was necessary to first investigate the generated cell lines concerning their capability to perform A-to-I-editing. Therefore, all used gRNAs were first off transfected into the desired cell lines to investigate the functionality of the used gRNAs itself, as well as the fundamental functionality of the cell lines to perform RNA editing. After transfection, RNA editing as well as fluorescence microscopy of ATTO 594 labeled gRNAs were used as a read out to investigate the receptor mediated uptake of the generated cell lines. For both cell lines, different endogenous targets and therefore, different gRNAs were used and the results are described within the following sections.

3.3.1. RNA Editing of ASGPR and SNAP®-ADAR expressing FlpIn™ T-REx™ 293 cells targeting STAT1 Y701C

To investigate the generated ASGPR H1a and SNAP®-ADAR1 E406Q expressing FlpIn™ T-REx™ 293 cell line concerning their capability to perform A-to-I-editing, STAT1 Y701 was chosen as endogenous target and gRNAs 324, 471-473, as well as 507 were used. As described before, STAT1 is a transcription activation factor[276] and the 5'-UAU codon of the key phosphorylation site Y701 is providing an attractive and endogenous target to manipulate signaling cascades via SNAP®-ADAR mediated A-to-I editing[118]. The capability of the synthesized BisBG gRNAs 471-473 to edit the desired STAT1 Y701 target was investigated by Yannis Stahl within the scope of his bachelor's thesis[279]. While editing yields between 20 %

and 40 % were observed using 1 pmol of gRNA and the well-characterized and SA1Q stably expressing FlpIn™ T-REx™ 293 cell line[118], the use of the FlpIn™ T-REx™ 293 cells stably expressing SA1Q as well as the receptor isoform H1a, obtained editing yields between 5 % and 20 %. The decreased editing yields are most probably related to the issue that several cells showed no integration of the bidirectional construct, as before. While a successful conjugation of both, BisBG (**18**) and GalNAc (**16**) provided the desired double conjugated gRNAs, the very low and total yields of 1-4 % are a major challenge. Regarding monetary aspects, the commercially available and 3'-conjugated GalNAc gRNA 507 was used for further investigations of a receptor mediated uptake into FlpIn™ T-REx™ 293 cells. gRNA 507 is similar in length and sequence to gRNA 324, but contains an additional 3'-conjugated triantennary *N*-acetyl galactosamine.

3.3.1.1. Transfection of SNAP®-ADAR gRNAs 324 and 507

For the transfection of the BisBG conjugated gRNAs 324 and 507 (conjugation is described before, see sections 3.1.2.1) into exclusively SA1Q as well as SA1Q and H1a expressing FlpIn™ T-REx™ 293 cells, the cells were seeded 24 h before transfection under doxycycline induction (10 ng/mL). The induced cells were reverse transfected for 24 h with 2 pmol of gRNA using Lipofectamine™ 2000 according to the manufacturer's protocol. The transfected cells were harvested and the total RNA was isolated using a Monarch® RNA Cleanup Kit (10 µg), followed by a reverse transcription and amplification of the desired target using a One Step RT-PCR Kit according to the manufacturer's protocols. The purified cDNA amplicons were analyzed by Sanger sequencing and the editing yield was determined. Detailed information about procedures and primers are described within methods and materials, section 6.5.12, and the results of the described experiments are shown in Figure 20.

In general, the functionality of the used gRNAs 324 and 507 was confirmed and high editing yields of > 70 % were observed for the exclusively SA1Q expressing cell line utilizing both gRNAs. Editing yields of 40-50 % were observed for the SA1Q and H1a expressing cell line, which is in accordance with the expectations and the previous described results of Yannis Stahl using gRNAs 471-473. All editing yields are in consistence with the results of the previous described BG-FITC staining of both cell lines, and the heterogeneous integration of the bidirectional GOI is also reflected by the decreased editing yields of the SA1Q and H1a expressing cell line. Therefore, a direct comparison of the two cell lines is very challenging and it is not possible to conclude to any impact of the receptor isoform H1a. However, the basic functionality of the used gRNAs and of the generated cell line to mediate A-to-I editing, was

confirmed. Additionally, only minor differences were observed regarding the editing yields of gRNA 324 compared to 507 using the same cell line. While gRNA 507 is containing a 3'-conjugated triantennary *N*-acetyl galactosamine, any impact of the additional GalNAc modification, such as sterical hindrances or conformational influences to the artificial editase or the gRNA itself can be ruled out. Therefore, both gRNAs and the generated cell line were used for the further investigation of the receptor mediated uptake of gRNAs into H1a expressing cells.

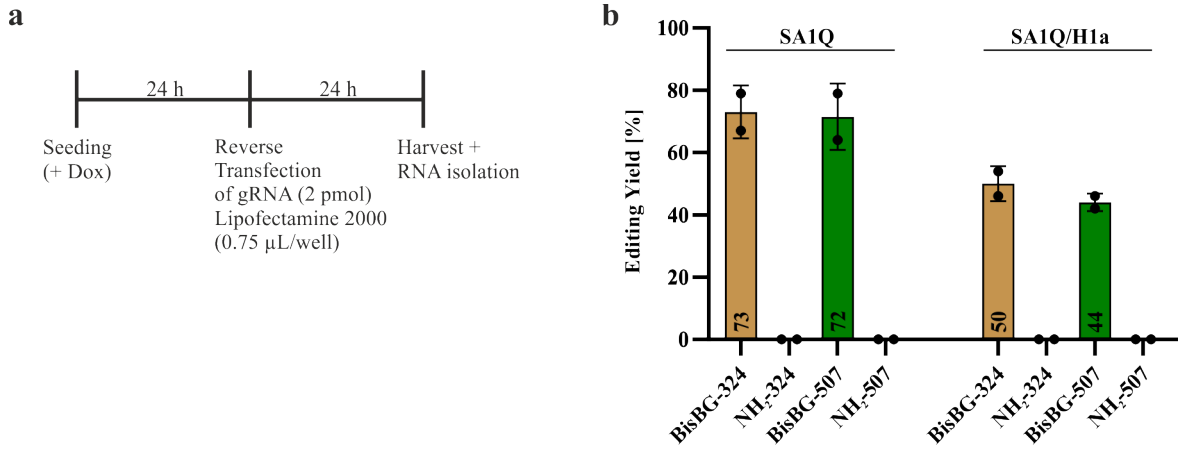


Figure 20: Transfection of BisBG gRNAs into FlpIn™ T-REx™ 293. *(a) General workflow of the experiment. Within a 24-well scale, 300.000 cells/well of each cell line were induced with doxycycline (10 ng/mL) for 24 h prior to use. $8 \cdot 10^4$ cells/well of each cell line were reverse transfected with 2 pmol of gRNA 324 and 507 using a 96-well cell culture plate and Lipofectamine™ 2000 (0.75 µL/well) in OptiMEM™ for 24 h under doxycycline induction (10 ng/mL). The transfected cells were harvested and the total RNA was isolated. The desired target RNA was further processed using reverse transcription and Sanger sequencing to evaluate the editing yield. Each experimental condition was performed in duplicates (n=2). (b) Mean editing yields in [%] of the STAT1 Y701C codon. Standard deviations are indicated as error bars and measurement data are indicated as dots.*

3.3.1.2. Receptor mediated endocytosis of SNAP®-ADAR gRNAs 324 and 507

Based on the promising results of the transfection, the next step was to investigate the receptor mediated endocytosis of the used gRNAs. The passive, and therefore undirected uptake of gRNAs with a similar design than gRNA 324 (25 nt, 2'-OMe ribose modifications, DNA at the target triplet, two PS modifications at the 5'-end and four PS modifications at the 3'-end) was previously investigated by Clemens Lochmann during his bachelor's thesis and under the supervision of Ngadhnjim Latifi. With the described design, passive uptake and therefore RNA editing was only observed utilizing BisBG conjugated gRNAs with an additional 3'-conjugated cholesterol[275]. Therefore, an additional GalNAc modification instead of cholesterol was thought to be highly beneficial to direct the internalization into a highly selective and productive pathway, the ASGPR mediated endocytosis.

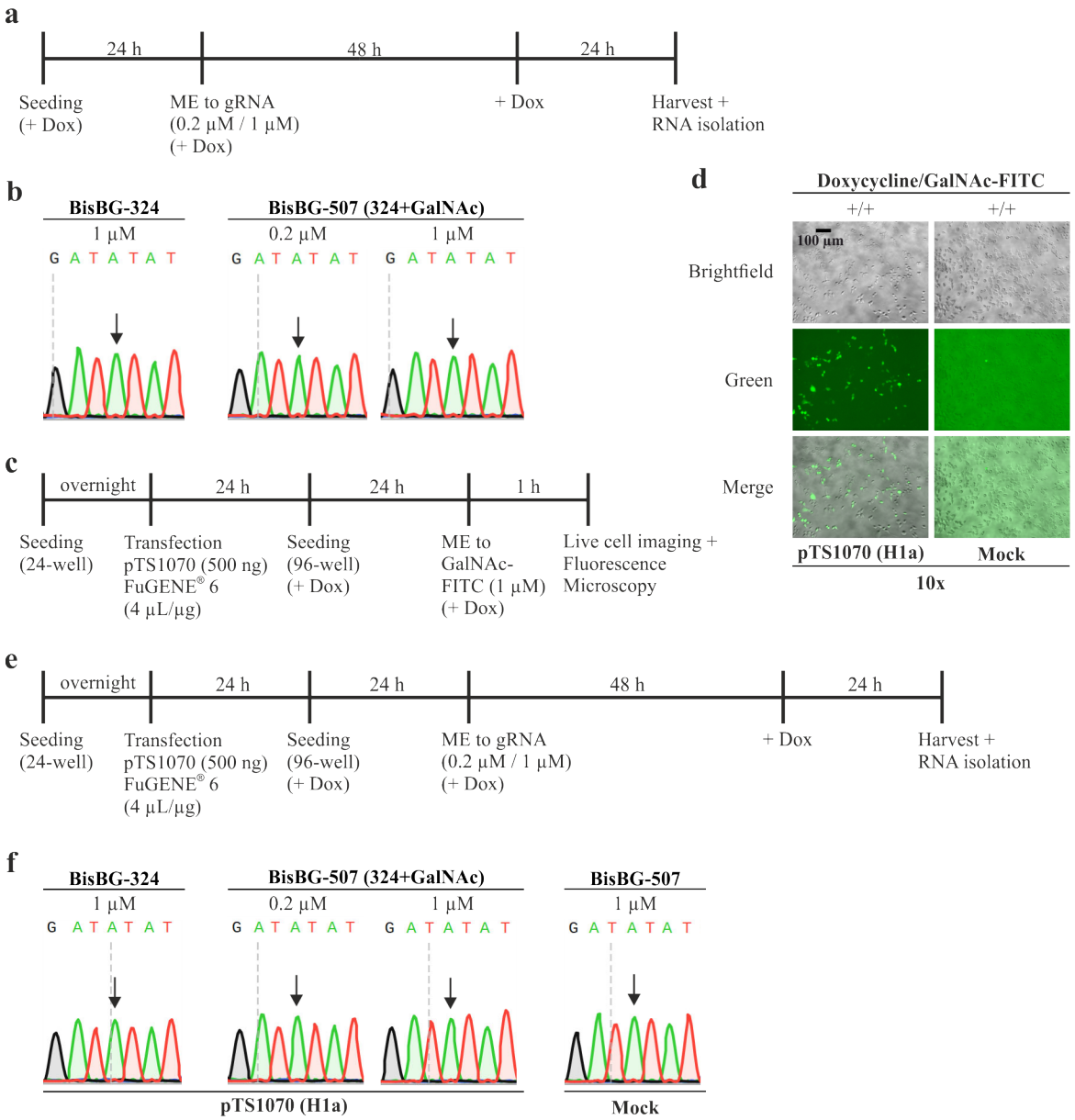


Figure 21: Receptor mediated endocytosis of SNAP®-ADAR gRNAs 324 and 507. *(a) General workflow of the receptor mediated endocytosis of gRNAs 324 and 507 into SA1Q and H1a expressing FlpIn™ T-REx™ 293 cells. 1.25•10^4 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well cell culture plates and incubated for 24 h under doxycycline induction (10 ng/mL). After 24 h, the media were exchanged and the gRNAs were added with final concentrations of 0.2 µM and 1 µM and the cells were incubated for 72h in total. After 48 h, additional doxycycline was added and after further 24 h, the cells were harvested and the total RNA was isolated. (b) Sanger sequencing traces of the STAT1 Y701 target to determine the editing yield. (c) and (e) General workflow of the transfection of pTS1070 (pcDNA5-H1a) into SA1Q expressing FlpIn™ T-REx™ 293 cells. 4 x 3·10^5 cells/well were seeded into 24-well cell culture plates and incubated overnight. The adherent cells were transfected for 24 h with pTS1070 (500 ng) using FuGENE® 6 (4 µL/µg) in OptiMEM™ and as negative control, no vector was transfected (mock). After transfection, 1.25•10^4 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well cell culture plates and incubated for 24 h under doxycycline induction (10 ng/mL). (c) After 24 h, the media were exchanged and GalNAc-FITC was added with a final concentration of 1 µM and the cells were incubated for 1 h and analyzed via fluorescence microscopy. (e) After 24 h, the media were exchanged and the gRNAs were added with final concentrations of 0.2 µM and 1 µM and the cells were incubated for 72h in total. After 48 h, additional doxycycline was added and after further 24 h, the cells were harvested and the total RNA was isolated. (d) Fluorescence imaging of the internalization of GalNAc-FITC into the H1a transfected SA1Q expressing FlpIn™ T-REx™ 293 cells. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree.*

For an overview of all samples and controls, see supplementary information Figure S15. (f) Sanger sequencing traces of the STAT1 Y701 target to determine the editing yield.

For the investigation of the ASGPR mediated internalization of gRNA 324 and 507, the cells were seeded for 24 h under doxycycline induction (10 ng/mL) into coated 96-well cell culture plates. The media were exchanged and the gRNAs were added with final concentrations of 0.2 μM and 1 μM. The cells were incubated for 72 h in total with an additional doxycycline addition after 48 h (see Figure 21a). The incubated cells were harvested and the total RNA was isolated and further processed as described before for the transfection-based experiments. The results of the Sanger sequencings are shown in Figure 21b, and as can be seen, no guanosine signal was observed at the target position (black arrow), and therefore no A-to-I editing took place within the use of any gRNA.

To exclude any issues regarding the integration of the bidirectional GOI of the generated SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell line, the receptor subunit H1a was additionally transfected transiently into the exclusively SA1Q expressing FlpIn™ T-REx™ 293 cells prior to the addition of gRNA. For the transient transfection of the pcDNA™5/FRT and receptor subunit H1a containing vector (pTS1070), the cells were seeded into 24-well cell culture plated and incubated overnight. The adherent cells were transfected using FuGENE® 6 according to the manufacturer's protocol for 24 h and after transfection, the cells were seeded under doxycycline induction (10 ng/mL) into coated 96-well cell culture plates and incubated for 24 h. The media were exchanged and the gRNAs were added with final concentrations of 0.2 μM and 1 μM. The cells were incubated for 72 h in total, with an additional doxycycline addition after 48 h (see Figure 21e), followed by RNA isolation and further sample processing. Unfortunately, and as before, no RNA editing was observed via Sanger sequencing (Figure 21f) within all samples.

To further evaluate the transfection efficiency, the transfected cells were additionally seeded under doxycycline induction (10 ng/mL) into coated 96-well cell culture plates and incubated for 24 h, followed by the internalization of GalNAc-FITC (**15**) for 1 h similar to the previous described settings (see Figure 21c). The results of the fluorescence microscopy are shown in Figure 21d, and a successful transfection as well as internalization of GalNAc-FITC was observed within about 40 % of the cells. This is even less than the degree of integration of the bidirectional SA1Q and H1a construct, but within both experiments, at least a minor amount of RNA editing was expected on the basis of a successful internalization of gRNA 507. This is not in accordance with the promising results of the intracellular distributed GalNAc-FITC, but

as described before, oligonucleotides are molecules with different physico-chemical properties, and only minor amounts of siRNAs and ASOs are reported to undergo an endosomal release.

Therefore, an endosomal entrapment of the used gRNAs could be a possible explanation for the missing intracellular availability and the unsuccessful RNA editing. However, all samples of the transfection-based and endocytosis-based experiments were treated similarly, but any preparative issues during RNA isolation and/or further sample processing can not be excluded completely. To further investigate the hypothesis of an endosomal entrapment, both NH_2-terminal gRNAs (324 and 507) were fluorescently labeled with ATTO 594, a hydrophilic and rhodamine inspired fluorescent dye and the transfection as well as receptor mediated internalization of the labeled gRNAs were analyzed via fluorescence microscopy.

3.3.1.3. Transfection and receptor mediated endocytosis of ATTO 594 labeled gRNAs 324 and 507

To investigate the transfection as well as the receptor mediated internalization of labeled gRNAs into SA1Q and H1a expressing FlpIn™ T-REx™ 293 cells, it was preliminary necessary to conjugate the desired gRNA to the dye-of-interest. Therefore, each NH_2-terminal gRNA was incubated overnight at RT with the commercially available and activated ATTO 594 NHS ester in gRNA labeling buffer (pH 8.3) and purified by NaOAc precipitation in EtOH. A detailed procedure is described in section 6.6.10, and the DOL (degree of labeling) was determined as recommended by the manufacturer (see equation (1)). For gRNA 324 and 507, a DOL of 114 % and 108 % were obtained, respectively, and the gRNAs were used without further purification to compare the transfection with the receptor mediated internalization of labeled gRNAs.

For the transfection of the two ATTO 594 labeled gRNAs 324 and 507, the cells were seeded into 96-well imaging plates and incubated for 24 h without doxycycline induction. The adherent cells were transfected for 24 h with 5 pmol of each gRNA using Lipofectamine™ RNAiMAX. In comparison, and for the receptor mediated endocytosis of the labeled gRNAs, the cells were also seeded into 96-well imaging plates but incubated for 24 h under doxycycline induction (10 ng/mL). The media were exchanged and the gRNAs were added with a final concentration of 1 µM. The cells were incubated for 24 h under doxycycline induction (10 ng/mL) and the cells within all conditions (transfection and receptor mediated endocytosis) were fixated using *p*-formaldehyde and analyzed via fluorescence microscopy. The results of the two different internalizations are shown in Figure 22c and d and different localizations were observed for the different internalization pathways. Besides scattered and bright signals, which

are most probably related to residual liposomes, both transiently transfected gRNAs showed a predominant localization within the nucleus (Figure 22c). This was additionally confirmed via merging the signals of the ATTO 594 dye (red) and the nuclei staining (blue). While a nuclear localization was observed for both gRNAs, any impact of the additionally conjugated triantennary GalNAc regarding the intracellular localization can be excluded. This is in accordance with the obtained editing yields, whereby only minor differences were observed for the editing yield of gRNA 324 compared to 507. The ubiquitous and therefore cytoplasmic and nuclear localization of SA1Q (see section 3.2.2.1) is also concordant to the observed localization of gRNAs, which can be assumed to be necessary to induce RNA editing.

In contrast, different localizations of the ATTO 594 signals were observed for the receptor mediated endocytosis of the used gRNAs (Figure 22d). As expected, and in accordance with the results of Clemens Lochmann, no internalization was observed for gRNA 324, which is lacking any 3'-modification such as cholesterol or the triantennary GalNAc. Interestingly, bright and concentrated cytosolic ATTO 594 signals were observed for gRNA 507, which is containing an additional 3'-conjugated triantennary GalNAc. However, and as for the internalization of GalNAc-FITC (**15**), no homogenous and cytoplasmic distribution was observed, and the accumulated signals are most probably endosomal-like structures. This is in accordance with the hypothesis of endosomal entrapped gRNAs, which could be responsible for the unsuccessful editing experiments utilizing the receptor mediated endocytosis. Additionally, no nuclear localization was observed for the internalized gRNA 507, whereby a detection of weak fluorescence signals next to high-intense signals is very challenging. Based on the obtained data, any assumption or discussion about the nuclear localization of the internalized gRNAs by endocytosis is not meaningful. Furthermore, a DOL of >100 % was obtained for both gRNAs, indicating a residual amount of unconjugated dye. According to that, any relation between the detected signal and the residual amount of dye is also possible. However, due to a missing signal for the use of gRNA 324, any relation of the observed signal to residual amounts of dye can be excluded and the authenticity of the detected ATTO labeled gRNAs is highly plausible.

In combination with the hypothesis of endosomal entrapment of gRNAs, the lysosomal stability of gRNAs is another aspect that could prevent a receptor mediated internalization to induce RNA editing. Especially a very low fraction (< 0.1 %) of internalized PS ASOs is reported to be intracellularly available within the cytosol or the nucleus, and a similar degree of endosomal release is assumed for other oligonucleotides with comparable lengths and

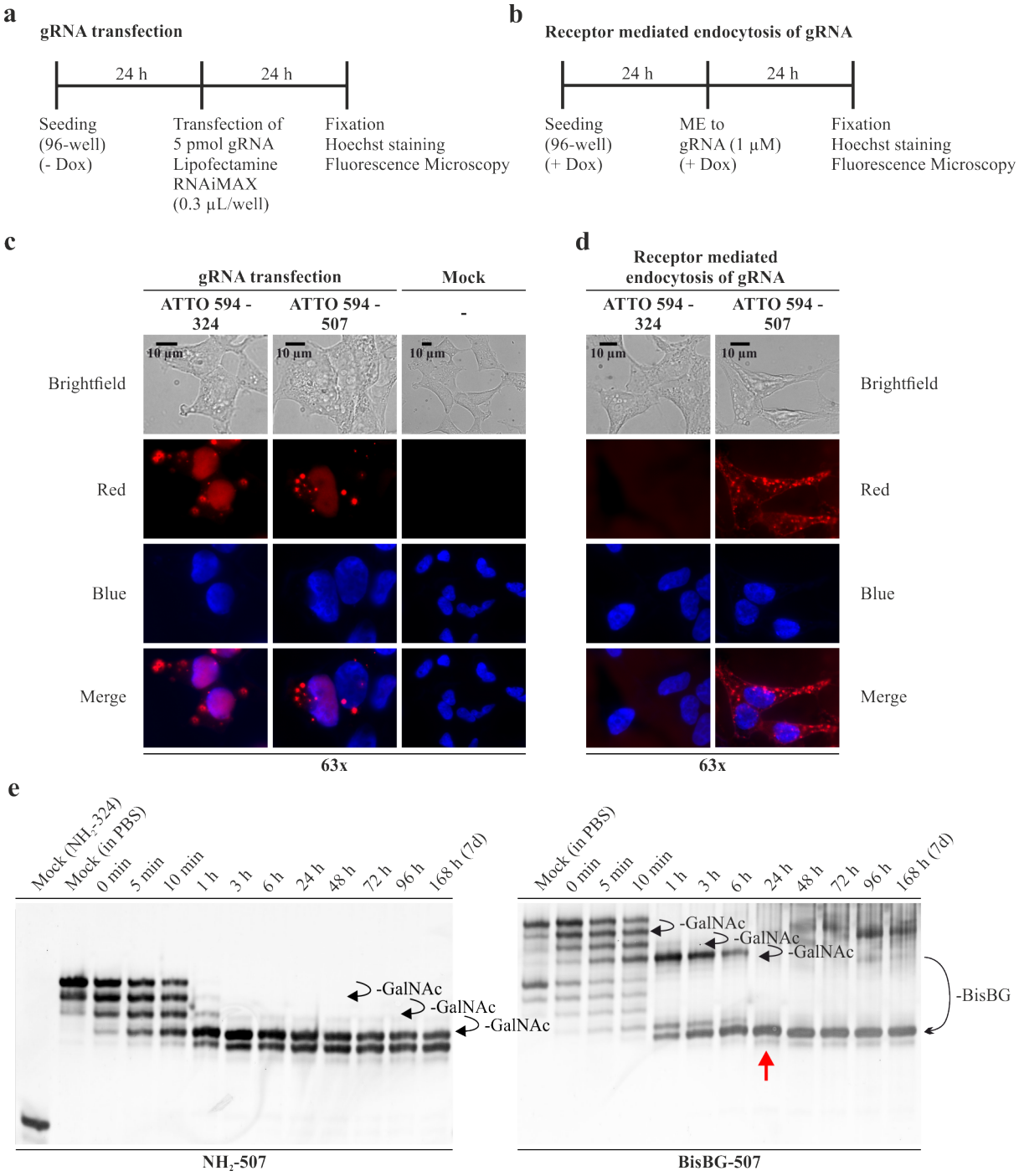


Figure 22: Transfection and receptor mediated uptake of ATTO 594 labeled BisBG gRNAs. *(a) General workflow of the transient transfection of gRNAs. 1.5•10^4 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well imaging plates and incubated for 24 h without doxycycline induction. The adherent cells were transfected for 24 h with gRNA (5 pmol) using Lipofectamine™ RNAiMAX (0.3 µL/well) in OptiMEM™ and as negative control, no gRNA was transfected (mock). The cells were fixated with p-formaldehyde, the nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed by fluorescence microscopy. (b) General workflow of the receptor mediated endocytosis of gRNAs. 1.5•10^4 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well imaging plates and incubated for 24 h under doxycycline induction (10 ng/mL). After 24 h, the media were exchanged, the gRNAs were added with final concentrations of 1 µM and the cells were incubated for 24h. The cells were fixated with p-formaldehyde, the nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed by fluorescence microscopy. ME = Media exchange (c) and (d) Fluorescence imaging of the transient transfection (c) and receptor mediated uptake (d) of ATTO 594 labeled gRNAs using a 63x magnification. Within a single light channel, similar exposure times and intensities are applied for the different conditions and*

within a single experiment, the contrasts of all ATTO 594 signals (red) are adjusted to a similar degree. (e) Stability assay of the NH_2-terminal and BisBG conjugated gRNA 507 in rat liver tritosomes. 180 pmol of gRNA (with a final concentration of 15 µM) were incubated in rat liver tritosomes (54 mU, 0.5 U/mL, SEKISUI XenoTech, Kansas City (MO), USA) at 37 °C and a sample of 15 pmol was taken at each time point. The tritosomes of each sample were inactivated with proteinase K (60 µg) and the samples were separated using a denaturing TBE-7 M Urea-PAGE (20 %) and visualized using fluorescence imaging as described in section 6.6.9. The stability assay is not part of this doctoral thesis and was performed by Ngadhnjim Latifi as previously described[280–282].

chemical modifications, such as SNAP®-ADAR gRNAs[184,200,214,215]. Therefore, and to ensure any long-term or delayed release-based effects, a high lysosomal stability of the utilized gRNAs is very important. Within 6-8 hours after internalization using a productive or non-productive pathway, PS ASOs are transported from early endosomes to late endosomes and/or lysosomes by Annexin A2 (ANXA2)[283]. ANXA2 is a Ca^{2+} dependent membrane binding protein and is proposed to facilitate ASO trafficking and/or release during endosomal maturing[284]. To investigate the lysosomal stability, the used gRNA 507 was tested regarding their stability against rat liver tritosomes (Triton WR 1339 treated rat liver lysosomes) and the results are shown in Figure 22e. The data of the stability assay of gRNA 507 were kindly provided by Ngadhnjim Latifi and are in general not part of this thesis.

As can be seen, a comparable pattern is observed for both, the NH_2-terminal and the BisBG conjugated gRNAs 507 within the first hour of incubation. The symmetric pattern is most probably related to the subsequent metabolization of single *N*-acetyl galactosamines (-GalNAc), whereby only minor changes are expected regarding the molecular masses and volumes in combination with an unchanged charge. This is in accordance with the results of Wang *et al.* where a rapid metabolism (within minutes) of the *N*-acetyl galactosamines by β-*N*-acetylglucosaminidases (lysosomal hydrolases) is reported.[285] In contrast, a further metabolite was observed exclusively within the stability assay of the BisBG conjugated gRNA, and after 24 h (red arrow) a single metabolite is observable, which is located at a similar migrational shift as the final metabolite of the NH_2-terminal gRNA. The additional metabolite is expected as the BisBG conjugated intermediate and a subsequent cleavage of the additionally conjugated BisBG modification (-BisBG) is most reasonable. Unfortunately, both final metabolites of the NH_2-terminal and the BisBG conjugated gRNAs 507 were expected at a comparable migrational shift as the NH_2-terminal gRNA 324. However, any further metabolization of the residual triantennary linker by DNAse II[285] or the cleavage of a terminal phosphodiester by RNAses or phosphatases can only be assumed. Especially the cleavage of a terminal phosphodiester would also change the overall charge of the metabolite, which would impact the migrational shift in an opposite manner. Therefore, a reason for the different migrational shifts of the NH_2-terminal gRNA 324 and the final metabolites of gRNA 507 remains unknown,

but the described metabolization of gRNA 507 (-GalNAc and -BisBG) is likely. The less tritosomal and therefore lysosomal stability of the BisBG conjugation of oligonucleotides could also explain the unsuccessful editing experiment utilizing a receptor mediated endocytosis, especially when transported form the early endosome to the late endosome and/or the lysosome. As seen within the transfection-based experiments, NH_2-terminal gRNAs are not able to hybridize with the SNAP®-ADAR fusion proteins, which is mandatory to induce RNA editing. Therefore, it is most probably not possible to evaluate any long-term or delayed release-based effects of a receptor mediated endocytosis utilizing gRNAs with the described design and stability.

To summarize the results of the receptor mediated uptake of SNAP®-ADAR gRNAs into ASGPR and SA1Q expressing FlpIn™ T-REx™ 293 cells, the fundamental functionality of the receptor mediated internalization of gRNAs was confirmed. The generation of a novel ASGPR H1a and SA1 E406Q expressing cell line was successful, and the integration of the bidirectional construct was confirmed using different characterizations, such as the receptor mediated internalization of a fluorescently active GalNAc-derivative or the immunofluorescent staining of the SNAP®-ADAR fusion protein as well as the membrane bound receptor subunit H1. The functionality of the generated cell line and the chosen gRNAs, to perform A-to-I editing, was additionally confirmed via transfection-based experiments and any impact of the 3'-conjugated triantennary GalNAc to the editing performance can be excluded. Despite the fact that no RNA editing was observed utilizing the receptor mediated endocytosis, plausible and consistent arguments were elaborated, which likely are related to the unsuccessful editing experiments. In contrast to the transfection of fluorescently labeled gRNAs, no nuclear localization was observable utilizing the receptor mediated endocytosis. However, the general internalization of exclusively 3'-GalNAc conjugated gRNA was successful and endosomal-like structures were observed. In general, the method of fluorescence microscopy is limited due to the requirement of relatively high concentrations of the labeled oligonucleotides, and as semi quantitative method, any visualization of finely distributed signals next to punctate accumulations remains challenging[283]. It is therefore not possible to propose any quantification or (sub-) cellular localization utilizing the receptor mediated endocytosis of fluorescently labeled gRNAs. A further aspect, which likely is related to the failed editing experiment using the receptor mediated uptake is the lysosomal stability of the BisBG modification. Before endosomal release, the metabolization of the conjugated BisBG during endosomal maturing, is abolishing the mandatory hybridization of the SNAP®-ADAR fusion proteins with the mRNA targeting

gRNA. The missing functionality of the metabolized gRNAs was also shown by transfecting NH_2-terminal gRNAs into SA1Q or SA1Q and H1a expressing cells, whereby no RNA editing was observed as well. Therefore, the use of gRNAs with an increased lysosomal stability is assumed to be highly beneficial regarding the endosomal release and the intracellular availability, and therefore the capability to perform RNA editing.

3.3.2. RNA Editing of ASGPR expressing HeLa cells targeting GAPDH L157L

To investigate the generated ASGPR H1a expressing HeLa cell line concerning their capability to perform A-to-I editing, the ORF of GAPDH (L157L) was chosen as endogenous target and gRNAs TMR189 and TMR236 were used. GAPDH is an endogenous housekeeping gene, which is involved in the glycolysis pathway and therefore the metabolization of glucose, as well as non-metabolic processes, such as transcription activation, initiation of apoptosis or the vesicle shuttling from the ER to the Golgi[286]. As described before, RESTORE v1 gRNAs are R/G motif containing and chemically stabilized gRNAs, which are able to recruit endogenous ADAR in a very promising way with high editing yields up to 77 % in a wide range of immortalized or even primary *human* cell lines and under interferon α induction (see Introduction, section 1.3.2.6). The corresponding RESOTRE v1 gRNAs were developed and further improved by Tobias Merkle during his doctoral thesis, resulting in RESTORE v2 gRNAs[287]. The improved gRNA design is lacking the R/G motif, which was thought to be mandatory for the dsRBD binding, but high editing yields between 15 % and 74 % were reported within several immortalized cell lines. Compared to the transfection of RESTORE v1 gRNAs, the interferon α dependency was negligible, and additionally, editing yields up to 88 % were obtained within different primary *human* cell lines. Especially the high editing yield of > 70 % within HeLa cells provided the fundamental situation to compare the ASGPR mediated uptake of 3'-GalNAc conjugated RESTORE v2 gRNAs (TMR236) to unconjugated ones (TMR189) within a H1a expressing HeLa cell line. As for the SNAP®-ADAR gRNAs, and regarding monetary aspects, the commercially available and 3'-conjugated GalNAc gRNA TMR236 was used, whereby gRNA TMR236 is similar in length and sequence to gRNA TMR189, with an additional 3'-conjugated triantennary *N*-acetyl galactosamine.

3.3.2.1. Transfection of RESTORE v2 gRNAs TMR189 and TMR236

For the transfection of the RESTORE v2 gRNAs TMR189 and TMR236 into the unsorted H1a expressing HeLa cells, the cells were seeded into 24-well cell culture plates 24 h before transfection with and without doxycycline induction (200 ng/mL). The seeded cells were

transfected for 24 h with 25 pmol of gRNA using Lipofectamine™ RNAiMAX according to the manufacturer's protocol. The transfected cells were harvested and the total RNA was isolated using a Monarch® RNA Cleanup Kit (10 µg), followed by a reverse transcription and amplification of the desired target using a One Step RT-PCR Kit according to the manufacturer's protocols. The purified cDNA amplicons were analyzed by Sanger sequencing and the editing yield was determined. Detailed information about procedures and primers are described within methods and materials, section 6.5.9, and the results of the described experiments are shown in Figure 23.

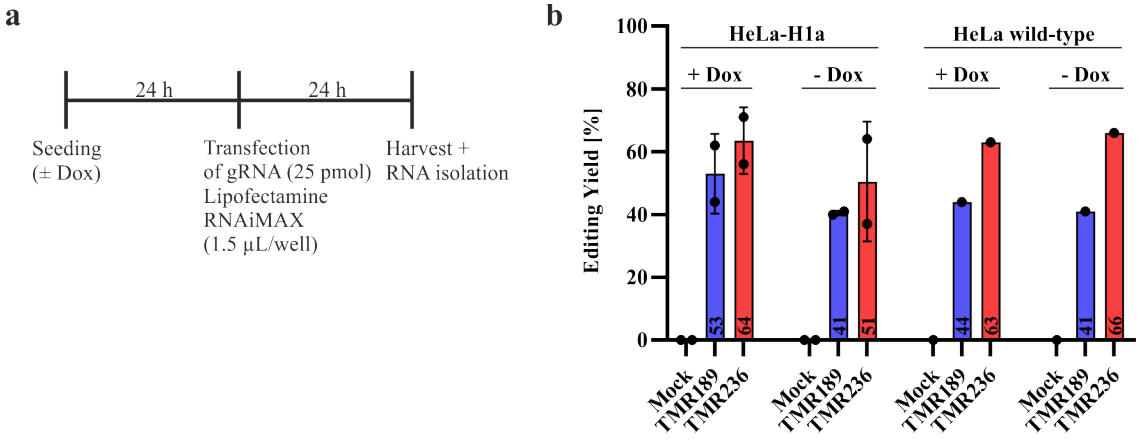


Figure 23: Transfection of RESTORE v2 gRNAs into HeLa cells. *(a) General workflow of the experiment. Within a 24-well scale, 100.000 cells/well were induced with and without doxycycline (200 ng/mL) for 24 h. The cells were transfected with 25 pmol gRNA using Lipofectamine™ RNAiMAX (1.5 µL/well) in OptiMEM™ for 24 h with and without doxycycline induction (200 ng/mL). As negative control, no gRNA was transfected (mock). The transfected cells were harvested and the total RNA was isolated. The desired target RNA was further processed using reverse transcription and Sanger sequencing to evaluate the editing yield. (b) Mean editing yields in [%] of the GAPDH L157L codon using H1a expressing HeLa cells compared to HeLa wild-type cells. Each experimental condition using H1a expressing HeLa cells was performed in duplicates (n=2) and each experimental condition using HeLa wild-type cells was performed in singlets (n=1). Standard deviations are indicated as error bars and measurement data are indicated as dots.*

In general, the functionality of the used gRNAs TMR189 and TMR236 and the generated H1a expressing cell line was confirmed and high editing yields were obtained. For the induced H1a expressing HeLa cells editing yields between 53 % and 64 % were observed, whereby the uninduced cells showed slightly decreased editing yields between 41 % and 51 %. Interestingly, comparable editing yields were observed for use of wild-type HeLa cells and neither the addition of doxycycline, nor the expression of the receptor subunit H1a seems to impact the overall editing performance. Additionally, the GalNAc conjugated gRNA TMR236 showed a higher mean editing yield than the unconjugated ones in all cases. Therefore, and in contrast to the SNAP®-ADAR gRNAs, an impact of the additional GalNAc modification can not be completely excluded. For unknown reasons, it was not possible to reproduce the previously described editing yield of > 70 % within the use of HeLa cells and the use of TMR189. But in

summary, the basic functionality of the used gRNAs, and of the generated cell line, to perform A-to-I editing, was confirmed and both were used for the further investigation of the receptor mediated uptake of RESTORE v2 gRNAs into H1a expressing HeLa cells.

3.3.2.2. Receptor mediated endocytosis of RESTORE v2 gRNAs TMR189 and TMR236

Based on the results of the transfection, the next step was to investigate the receptor mediated endocytosis of the used gRNAs. The passive or gymnotic, and therefore undirected uptake of the RESTORE v2 gRNA TMR189 was previously investigated by Laura Pfeiffer during her master's thesis and editing yields up to 35 % were reported utilizing a gRNA concentration of 1 µM over 10 days[288]. Therefore, an additional GalNAc modification was thought to be highly beneficial to direct the internalization into a highly selective and productive pathway, the ASGPR mediated endocytosis, and with this, to decrease to time of incubation, which is necessary to obtain sufficient editing yields. Additionally, and in relation to the hypothesis of endosomal entrapped gRNAs, which arose through the application of SNAP®-ADAR gRNAs, the addition of Chloroquine was also evaluated within the investigation of the receptor mediated internalization. Compared to other small molecule-based endosomal disruptors, such as Siramesine or Loperamide, Chloroquine (Figure 24b) was recently reported by Du Rietz *et al.* to enhance the target knockdown efficiency of cholesterol conjugated siRNAs within tumor cell spheroids of two cancer cell lines in the most promising manner[289].

For the investigation of the receptor mediated internalization of gRNA TMR189 and TMR236, the unsorted H1a expressing HeLa cells were seeded into 24-well cell culture plates for 24 h under doxycycline induction (200 ng/mL). The media were exchanged and the gRNAs were added with final concentrations of 0.2 µM and 1 µM and with or without the addition of 60 µM Chloroquine. The cells were incubated for 24 h, 48 h or 72 h and after incubation, the cells were harvested and the total RNA was isolated and further processed as described before for the transfection-based experiments. An overview of the Sanger sequencing results is shown in Figure 24c, and promising editing yields were obtained.

First of all, the reported results of Laura Pfeiffer for the gymnotic uptake of TMR189 into wild-type HeLa cells, decreased editing yields were obtained for the internalization of both gRNAs into H1a expressing HeLa cells over 72 h (Figure 24c, -Chloroquine). This is in accordance to the previously described results of the gRNA transfection and the decreased editing yields of both used gRNAs. Furthermore, a highly beneficial effect of Chloroquine was observed within all samples and mean editing yields up to 32 % and 56 % were observed for gRNA concentrations of 0.2 µM and 1 µM, respectively (Figure 24c, +Chloroquine).

Additionally, a similar progression was obtained within all samples using Chloroquine, whereby an increasing editing yield was observed from 24 h to 48 h, followed by a decrease from 48 h to 72 h. This is most probably related to the cellular toxicity of Chloroquine, which was also reflected by an increasing degree of dead cells over time (data not shown). However, and in accordance with the recently reported application as endosomal disruptor[289], the highly increased editing yield when using Chloroquine, is supporting the hypothesis of an endosomal entrapment of gRNAs and therefore, the highly hindered cytosolic and nuclear availability applying a receptor mediated endocytosis.

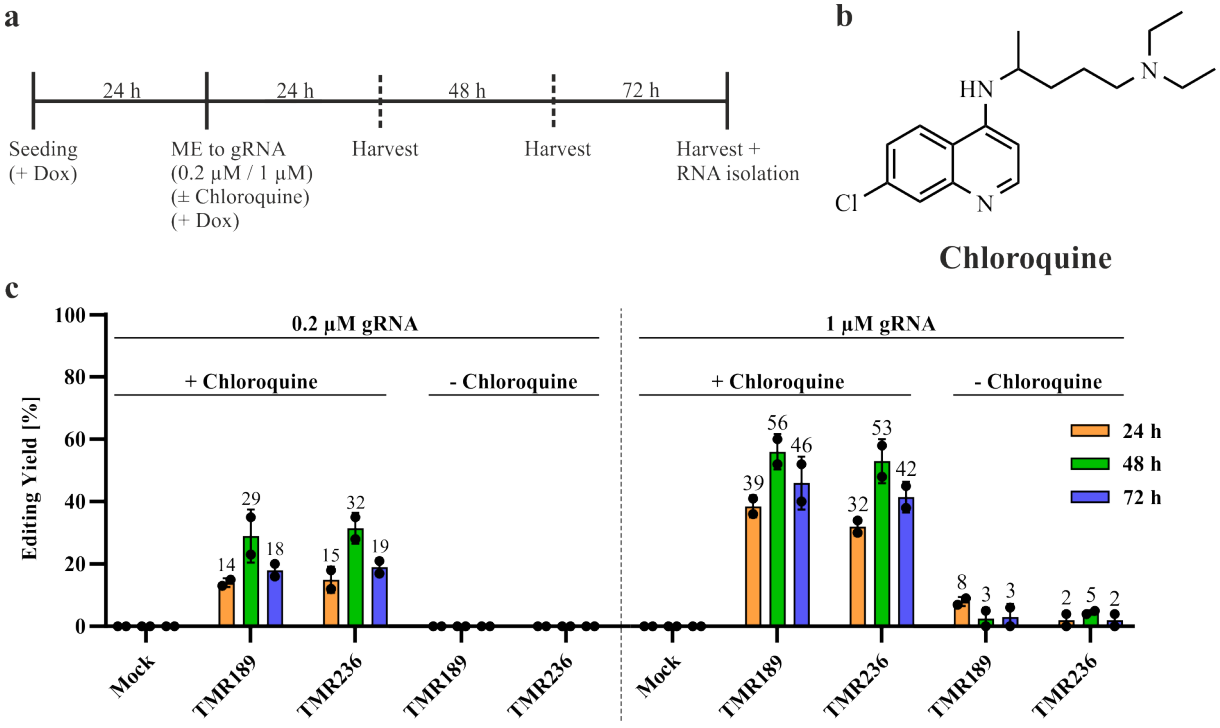


Figure 24: Receptor mediated uptake of RESTORE v2 into unsorted HeLa cells. *(a) General workflow of the receptor mediated endocytosis of gRNAs TMR189 and TMR236 into unsorted and H1a expressing HeLa cells. 5•10[4] cells/well were seeded into 24-well cell culture plates and incubated for 24 h under doxycycline induction (200 ng/mL). After 24 h, the media were exchanged and the gRNAs were added with final concentrations of 0.2 µM and 1 µM with and without 60 µM Chloroquine and the cells were incubated for 24 h, 48 h or 72 h. The cells were harvested at each time point and the total RNA was isolated. As negative control, no gRNA was added (mock). (b) Chemical structure of Chloroquine. (c) Mean editing yields in [%] of the GAPDH L157L codon. Each experimental condition was performed in duplicates (n=2). Standard deviations are indicated as error bars and measurement data are indicated as dots.*

Most interestingly, and independent of the gRNA concentration, highly consistent editing yields were obtained for the use of both gRNAs, and compared to the unmodified gRNA TMR189, no beneficial effect of the receptor mediated endocytosis can be addressed to the additional GalNAc modification of gRNA TMR236. This is in contrast to the current state-of-knowledge regarding the receptor mediated internalization of GalNAc modified ASOs and siRNAs into ASGPR expressing cell lines or within *in vivo* models as well as clinical

trials[144,146,251,290,291]. However, the capability of ASGPR expressing cells to internalize PS ASOs was previously indicated by the reported results of Tanowitz *et al.*[252] as well as Liang *et al.*[193] and especially the use of high ASO concentrations (≥ 1 µM) lead to an adequate internalization of unmodified PS ASOs[252]. With an increasing PS content, the beneficial effect of an additional GalNAc modification became less relevant and a significant role of the ASGPR to the internalization and the activity of unmodified PS ASOs was concluded within *in vitro* and *in vivo* studies. While a high PS content of the two used gRNAs TMR189 and TMR236 was demonstrated to be highly beneficial regarding serum stability and editing efficiency[287,288], the high internalization efficiency and the consistent editing yields of both gRNAs is most likely also related to the fully PS modified design.

Another aspect, which could impact any effect of the additional GalNAc modification, might be the expression level the ASGPR receptor itself. For this purpose, the effect of the additional GalNAc modification was further investigated, using both described gRNAs in combination with HeLa cells showing different expression levels of the receptor subunit H1a (see cell sorting, section 6.8.2.3). The general expression of the receptor was also investigated in the absence of doxycycline and for all conditions, the endosomal disruptive effect of Chloroquine was evaluated as well. Therefore, and as described before, the different cell lines were seeded with or without doxycycline induction (200 ng/mL) and after 24 h, the media were exchanged, and based on the previously described results, the cells were further incubated in the absence or presence of doxycycline (200 ng/mL) and/or Chloroquine (60 µM) for 48 h and with a final gRNA concentration of 0.2 µM. The total RNA was isolated, further processed as described before and the results are shown in Figure 25.

In general, different results were observed for the different conditions (Figure 25b, ±Chloroquine, ±Dox). First of all, only minor editing yields were obtained for the absence of Chloroquine, and by this no definite dependency of the receptor expression or the expression level was observable (Figure 25b, -Chloroquine, +Dox vs. -Dox). In contrast, higher editing yields were observed under Chloroquine induction, which is in accordance with the expectations and the previously described results. While only minor differences were observed between the editing yields of the unsorted cells and the sorted cells of the FITC dim fraction (low FITC signal) under doxycycline induction, both cell lines showed higher mean editing yields than the uninduced cells (Figure 25b, +Chloroquine, +Dox vs. -Dox). This effect was more pronounced when comparing the editing yields of the induced and sorted cells of the FITC pos fraction (high FITC signal) with the uninduced cells. For this, mean editing yields of 37 %

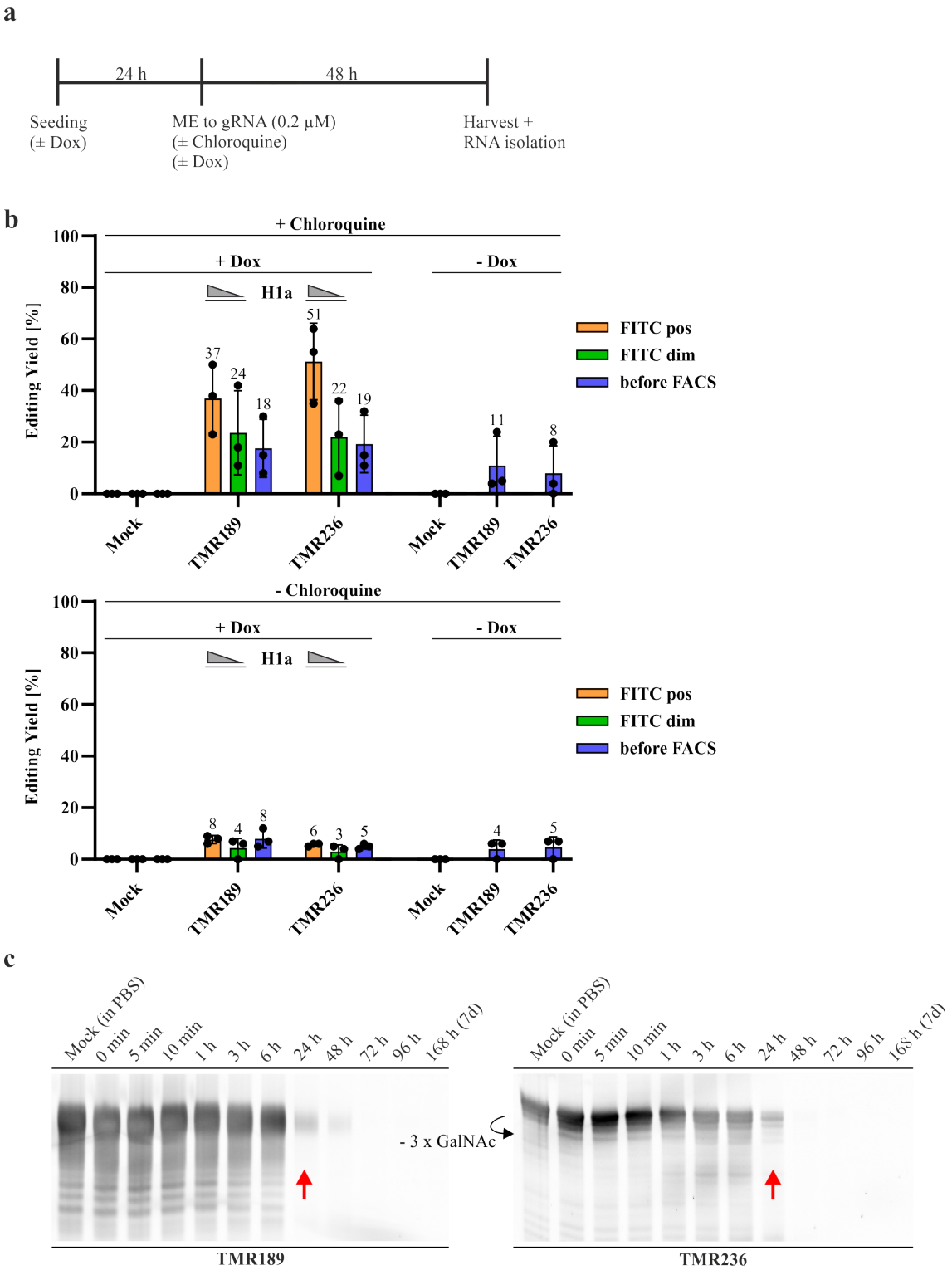


Figure 25: Receptor mediated uptake of RESTORE v2 into sorted HeLa cells. *(a) General workflow of the receptor mediated endocytosis of gRNAs TMR189 and TMR236 into sorted and H1a expressing HeLa cells. 5•10^4 cells/well (+Chloroquine) or 1•10^4 cells/well (-Chloroquine) were seeded into 24-well cell culture plates and incubated for 24 h with and without doxycycline induction (200 ng/mL). After 24 h, the media were exchanged, the gRNAs were added with a final concentration of 0.2 µM, with and without 60 µM Chloroquine and with or without doxycycline (200 ng/ml) and the cells were incubated for 48 h. The cells were harvested and the total RNA was isolated. As negative control, no gRNA was added (mock). (b) Mean editing yields in [%] of the GAPDH L157L codon. Each experimental condition was performed in duplicates (n=3). Standard deviations are indicated as error bars and measurement data are indicated as dots. (c) Stability assay of the RESTORE v2 gRNAs TMR189*

(left) and TMR236 (right) in rat liver tritosomes. 180 pmol of gRNA (with a final concentration of 15 µM) were incubated in rat liver tritosomes (54 mU, 0.5 U/mL, SEKISUI XenoTech, Kansas City (MO), USA) at 37 °C and a sample of 15 pmol was taken at each time point. The samples were separated using a denaturing TBE-7 M Urea-PAGE (15 %) and visualized using fluorescence imaging as described in section 6.6.9. The stability assay is not part of this doctoral thesis and was performed by Laura Pfeiffer as previously described[280–282].

and 51 % were observed for the two gRNAs TMR189 and TMR236, respectively, and a dependency of the degree of internalization and the general receptor expression as well as the expression level can be confirmed. Therefore, the increased editing efficiency, which correlates with a higher degree of internalized gRNAs, can be directly addressed to the receptor expression, and thus to the receptor mediated endocytosis (Figure 25b, +Chloroquine, +Dox vs. -Dox).

An increased editing yield was observed for gRNA TMR236 compared to TMR189 using the sorted cells of the FITC pos fraction (high FITC signal), and by this, the beneficial effect of the additional GalNAc modification is exclusively indicated within the use of the cell line with a high ASGPR H1a expression level and in combination with Chloroquine. However, and because of the calculated standard deviations and varying measurement data, the effect of the GalNAc modification can not be completely confirmed and a further verification is necessary. A correlation between the receptor expression level and the splice modulating effect of *SMN2* targeting splice switching ASOs was also observed by Scharner *et al.* within H1a transduced U87 glioblastoma cell lines, and a 5-fold increased dose-response curve was reported for the use of GalNAc conjugated ASOs compared to unconjugated ones[269].

Unfortunately, any beneficial effect of the GalNAc modification was not observed for the use of TMR189 and TMR236 in the absence of Chloroquine and the high H1a expressing cell line. This is most probably related to the hypothesis of the endosomal entrapped gRNAs, which is further confirmed by the overall and highly differing editing yields between the presence and absence of Chloroquine (Figure 25b, +Chloroquine vs. -Chloroquine). This is also in accordance with the literature concerning the endosomal release of oligonucleotides and the very low fraction of < 0.1 % that becomes available within the cytosol or the nucleus during endosomal maturing. RNA editing is also assumed to be less sensitive than the RNAse H1 or RISC mediated degradation of the target mRNAs. This might also be reflected by the necessity of higher gRNAs concentrations (0.2 µM) and the further addition of Chloroquine, which are both needed to obtain sufficient editing yields. Compared to other sequencing or detection techniques such as next generation sequencing, luciferase-based assays or quantitative PCRs (qPCRs), Sanger sequencing is less sensitive including noticeable and internal deviations, which is also limiting the exact evaluation of minor editing yields.

Another aspect, which is of particular importance and which is in combination with the endosomal release or any long-term or delayed release-based effects, is the lysosomal stability of gRNAs. The importance of the lysosomal stability was previously shown within the results of the receptor mediated endocytosis of SNAP®-ADAR gRNAs and hence, no RNA editing was observed, due to the likely metabolization of gRNAs within the lysosome and the cleavage of the mandatory BisBG modification. Because of this, the lysosomal stability of the used gRNAs TMR189 and TMR236 was also tested regarding their stability against rat liver tritosomes (Triton WR 1339 treated rat liver lysosomes) and the results are shown in Figure 25c. The data of the stability assay of the two gRNAs TMR189 and TMR236 were kindly provided by Laura Pfeiffer and are in general not part of this thesis.

In comparison to the SNAP®-ADAR gRNAs, a similar pattern was observed for the tritosomal metabolization of the GalNAc modified RESTORE v2 gRNA TMR236, and the characteristic metabolic pattern of the three *N*-acetyl galactosamines was indicated. Most importantly, both RESTORE v2 gRNAs showed an almost complete metabolization, and therefore degradation after 24 h (Figure 25c, red arrows). Next to the endosomal release, the reduced tritosomal and lysosomal stability of the two gRNAs provides a further and major bottleneck, which is most likely related to the minor editing yields utilizing a receptor mediated endocytosis. Especially the combination of both aspects, the limited endosomal release and the less lysosomal stability, is highly inhibiting the availability of a sufficient amount of gRNA within the cytosol or the nucleus, which is necessary to induce of RNA editing in a detectable degree. Therefore, it is unlikely that any long-term or delayed release-based effects of a receptor mediated endocytosis utilizing gRNAs with the described design and stability can be evaluated in a meaningful way. In conclusion, and as shown within the described results of the receptor mediated endocytosis and the addition of Chloroquine, a sufficient amount of gRNA is productively internalized into the cells, but not cytosolic and/or nuclear available, due to the limited endosomal release and the reduced lysosomal stability.

3.3.2.3. Transfection and receptor mediated endocytosis of ATTO 594 labeled gRNAs TMR189 and TMR236

To further investigate the endosomal entrapment and the effect of Chloroquine to enhance the endosomal release of the GalNAc modified and unmodified oligonucleotides, fluorescently labeled gRNAs were used within transfection-based as well as endocytosis-based experiments and analyzed using fluorescence microscopy. This is in accordance with the application of fluorescently labeled SNAP®-ADAR gRNAs, and the corresponding RESTORE v2 gRNAs

were conjugated to ATTO 594 in the same vein and as described before (see section 6.6.10). Therefore, for gRNA TMR189 and TMR236, DOLs of 103 % and 95 % were obtained, respectively, and the gRNAs were used without further purification.

For the transfection of the two ATTO 594 labeled gRNAs, the cells were seeded into 96-well imaging plates and incubated for 24 h without doxycycline induction, followed by the transfection of the adherent cells for 24 h with 5 pmol of each gRNA using Lipofectamine™ RNAiMAX (Figure 26a). In comparison, and for the receptor mediated endocytosis of the labeled gRNAs, the cells were also seeded into 96-well imaging plates but incubated for 24 h in the absence or presence of doxycycline (200 ng/mL). The media were exchanged and the gRNAs were added with a final concentration of 1 µM and the cells were incubated for 24 h in the absence or presence of doxycycline (200 ng/mL). To investigate the effect of Chloroquine, additional cells were also incubated for 24 h with a final gRNA concentration of 1 µM and in the absence or presence of doxycycline (200 ng/mL) and/or Chloroquine (60 µM) (Figure 26b). The cells within all conditions (transfection and receptor mediated endocytosis) were fixated using *p*-formaldehyde and analyzed via fluorescence microscopy. The results of two different internalizations are shown in Figure 26c and d and different localizations were observed for the different internalization pathways. Besides scattered and bright signals, which are most probably related to residual liposomes, both transiently transfected gRNAs showed nuclear localized accumulations (Figure 26c, white arrows). In general, PS oligonucleotides are known to shuttle between the nucleus and the cytoplasm in a saturable and carrier mediated manner, and especially for the use of transfection reagents, PS oligonucleotides are known to accumulate in the nucleus as punctate structures[216,292]. These accumulated structures, so-called PS bodies, are biochemically not completely characterized, but were first described by Lorenz *et al.* and are thought to buffer the availability of PS oligonucleotides in the nucleoplasm[292]. In contrast to the transfection-based internalization, different localizations of the ATTO 594 signals were observed for the receptor mediated endocytosis of the used gRNAs (Figure 26d), and no PS bodies was observed within the nucleus. In accordance with the results of the RNA editing experiments and the described effect of the high PS content, both gRNAs showed intense and most likely exclusively endosomal-like signals. Additionally, intense signals were observed for both gRNAs under doxycycline induction, and therefore under ASGPR expression, which is also in correlation to the previously described results. In contrast to the generated FlpIn™ T-REx™ 293 cell line, similar and endosomal-like signals were observed for both, the ATTO 594 labeled gRNAs and GalNAc-FITC (**15**) using H1a expressing HeLa cells. Within FlpIn™ T-

REx™ 293 cells, a ubiquitously distributed FITC signal was observed for the internalization of GalNAc-FITC (**15**), whereby endosomal-like signals were obtained for the use of gRNA 507 (see section 3.3.1.2). However, the transient transfection of the receptor subunit H1a into A549, Huh7, U2OS and U87 cell lines and the subsequent receptor mediated internalization of GalNAc-FITC (**15**), showed endosomal-like structures within all samples (supplementary information, Figure S19). Therefore, the ubiquitously distributed FITC signal, which was observed for the internalization of GalNAc-FITC (**15**), is potentially related to an exceptional endosomal behavior of the FlpIn™ T-REx™ 293 cells compared to the observed and endosomal-like structures of all other used cell lines.

Additionally, different observations were made for the receptor mediated uptake of the ATTO 594 labeled gRNAs within the investigation of the effect of Chloroquine to enhance the endosomal release (Figure 26e). As indicated, the high editing yields of the transfected gRNAs and the comparatively low editing yields of the receptor mediated uptake (without the use of Chloroquine), might be reflected by the presence of PS bodies, and therefore the nuclear availability of the internalized gRNAs. As observed within the transfection-based experiments, accumulated ATTO 594 signals within the nucleus became exclusively observable by the receptor mediated uptake, when additional Chloroquine was added (Figure 26e, white arrows). This is in correlation to the observed editing yields, which were also highly increased when adding Chloroquine as endosomal disruptor. Therefore, both the increased editing yields and the occurrence of PS bodies using Chloroquine as additive during the receptor mediated uptake are further confirming the hypothesis of the endosomal entrapment and the related reduced intracellular and nuclear availability of gRNAs, which is necessary for high editing yields.

However, a general correlation between the antisense activity and the nuclear localization is reported for PS ASOs, which can not be characterized in detail utilizing fluorescence microscopy as semi quantitative and less sensitive technique[283]. A nuclear activity of RESTORE v2 gRNAs was also suggested by T. Merkle during their development, which was additionally reflected by the independence of interferon α as well as the remaining editing activity after ADAR1p150 or ADAR2 knockdown[287]. Therefore, the recruitment of the predominantly nuclear localized ADAR1p110 is plausible. However, a direct correlation between the nuclear activity of the used gRNAs and their accumulation in nuclear PS bodies, which are detected via fluorescence microscopy, can only be assumed.

Further to the occurring and nuclear localized PS bodies using Chloroquine, and as described before, similar fluorescence signals were observed for the GalNAc modified and

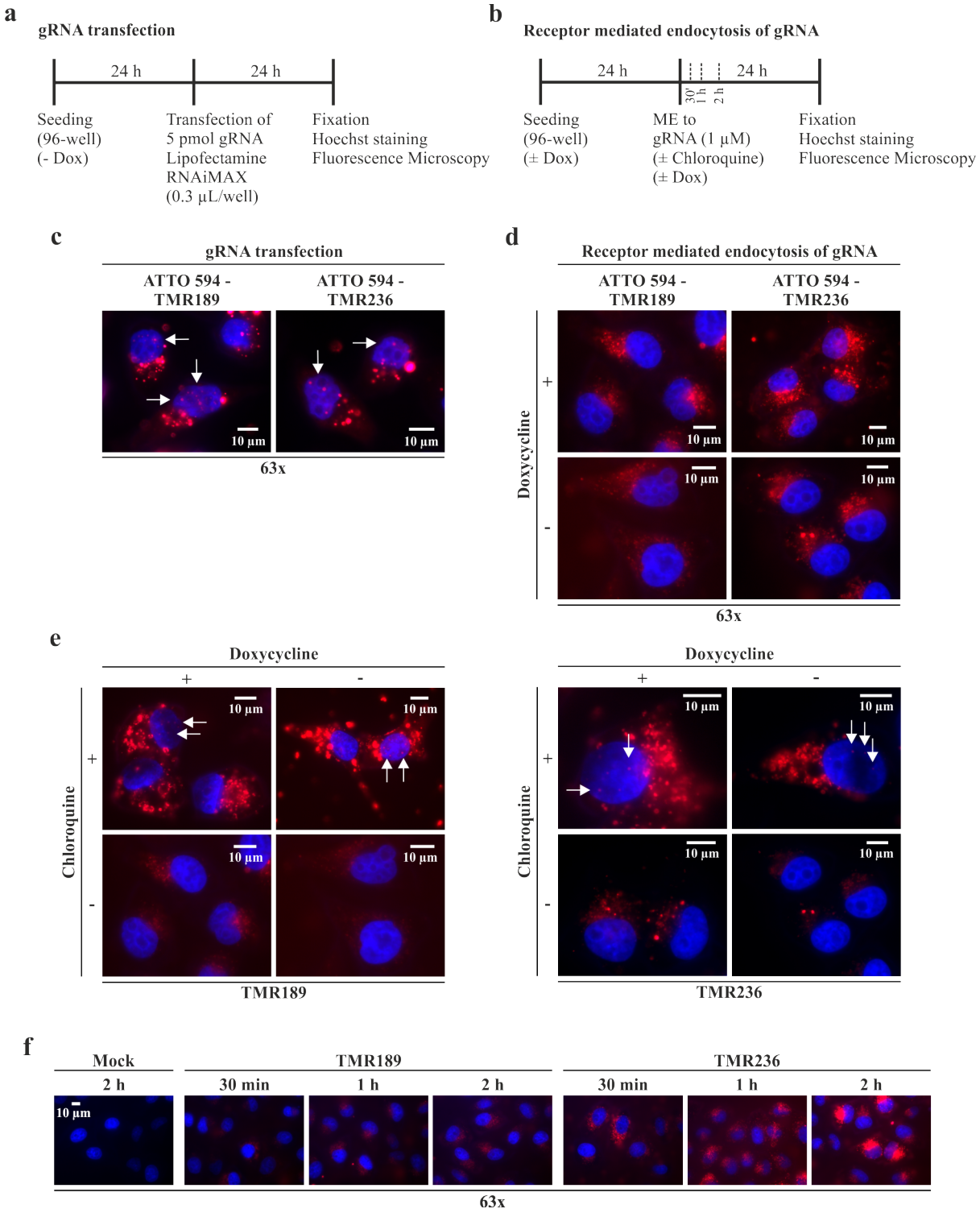


Figure 26: Transfection and receptor mediated endocytosis of ATTO 594 labeled RESTORE v2 gRNAs into HeLa cells. *(a) General workflow of the transient transfection of gRNAs. 5•10^3 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well imaging plates and incubated for 24 h without doxycycline induction. The adherent cells were transfected for 24 h with gRNA (5 pmol) using Lipofectamine^{TM} RNAiMAX (0.3 µL/well) in OptiMEM^{TM} and as negative control, no gRNA was transfected (mock). The cells were fixated with p-formaldehyde, the nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed by fluorescence microscopy. (b) General workflow of the receptor mediated endocytosis of gRNAs. 5•10^3 cells/well were seeded into poly-D-lysine HBr$_{aq}$ coated 96-well imaging plates and incubated for 24 h with and without doxycycline induction (200 ng/mL). After 24 h, the media were exchanged, the gRNAs were added with final concentrations of 1 µM and the cells were incubated in absence or presence of 60 µM Chloroquine and or doxycycline induction (200 ng/mL) for 24h. The*

cells were fixated with p-formaldehyde, the nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed by fluorescence microscopy. (c) Merged fluorescence imaging of the red and blue channel of the transient transfection of gRNAs. All single channels are shown in supplementary information Figure S16. (d) Merged fluorescence imaging of the red and blue channel of the receptor mediated uptake of gRNAs without Chloroquine after 24 h. All single channels are shown in supplementary information Figure S16 (e) Merged fluorescence imaging of the red and blue channel of the receptor mediated uptake of gRNAs in absence or presence of Chloroquine after 24 h. All single channels are shown in supplementary information Figure S17. (f) Merged fluorescence imaging of the red and blue channel of the receptor mediated uptake of gRNAs without Chloroquine from 0-2h. All single channels are shown in supplementary information Figure S18. Within a single light channel, similar exposure times and intensities are applied for the different conditions and within a single section, the contrasts of all ATTO 594 signals (red) are adjusted to a similar degree. ME = Media exchange.

unmodified gRNAs utilizing the receptor mediated endocytosis and an incubation time of 24 h. To investigate a possible saturation-based effect of the productive endocytosis, shorter incubations times (0-2 h) were analyzed as well using live cell imaging, and in fact, clear differences were observed between the two different gRNAs (Figure 26f). For the use of the GalNAc modified gRNA TMR236, high fluorescence signals were observed within several cells, while only minor fluorescence signals were observed for the use of the unmodified gRNA TMR189. The heterogeneous integration of gRNA TMR236 is in accordance with the use of the unsorted and therefore heterogeneously H1a expressing HeLa cell line, and at least for short incubation times, a beneficial effect of the additional GalNAc modification became particularly obvious. Therefore, a kind of plateau effect or saturation of the productive and receptor mediated endocytosis might be assumed, but a further investigation and quantification is necessary.

In total, and within the transfection-based editing experiments, the general functionality of the GalNAc modified RESTORE v2 gRNA was confirmed, and any impact of the ASGPR H1a expression to the overall editing performance was excluded. While only low editing yields were observed for the application of the receptor mediated endocytosis, highly increased editing yields were observed due to the addition of Chloroquine as endosomal disruptor. This is in accordance with the hypothesis of endosomal entrapped gRNAs and the general state-of-knowledge about the endosomal release of oligonucleotides and the very low fraction of < 0.1 % that becomes available within the cytoplasm or the nucleoplasm during endosomal maturing. The aspect of endosomal entrapped gRNAs was further confirmed by the application of ATTO 594 labeled gRNAs and the occurring nuclear availability of gRNAs after Chloroquine treatment (PS bodies). The limitation of the endosomal release is even more important in combination with the less lysosomal stability of the used gRNA designs, and the combination of both aspects is highly inhibiting the cytoplasmic or nucleoplasmic availability. Furthermore, a high impact of the internalization capability of RESTORE v2 gRNAs was addressable to the

ASGPR expression level, and especially for the use of a high receptor expressing HeLa cell line, a beneficial effect of the additional GalNAc modification was indicated within the editing results. While a conspicuous and beneficial effect of the GalNAc modification was also observable via fluorescence microscopy applying shorter incubation times, any effect of the mentioned modification was not reflected within the editing results or the fluorescence imaging of HeLa cells with a heterogeneous or lower receptor expression level utilizing longer incubation times. For this, a further verification via RNA editing utilizing shorter incubation times is necessary. This is in accordance with the still uncharacterized, but not neglectable gymnotic uptake or receptor mediated endocytosis of PS oligonucleotides as competing pathways for the cellular internalization. This was also observed by the internalization of gRNA TMR189 into induced and uninduced HeLa cells. Unfortunately, it is not possible to determine any quantitative or stoichiometric relations due to the limitations of fluorescence microscopy as a semi quantitative method. High signal intensities, and therefore high concentrations, are required for a sufficient visualization and especially co-localizations of high intense signals next to low concentrated signals remains challenging[283]. However, and as shown within the described results, it is possible to conclude that a sufficient amount of gRNA is productively internalized into the cells via ASGPR mediated endocytosis, but not available within the cytosol or the nucleus, due to the limited endosomal release and the less lysosomal stability.

4. Summary and Conclusion

In the past decades, antisense oligonucleotide (ASO)- or short-interfering RNA (siRNA)-based drug systems pioneered promising state-of-the-art RNA targeted drugs to treat several diseases via gene regulation as well as transcriptional or translational regulations, such as splice site alterations or transcript degradations[144–146]. However, the targeted delivery of these drugs was very challenging. With the discovery of the asialoglycoprotein receptor (ASGPR), and the subsequent development of synthetic ligands, a promising strategy arose to deliver ASOs and siRNAs into hepatocytes to target several liver diseases[222,253,254]. This advantage was also thought to be highly beneficial regarding the targeted internalization of either SNAP®-ADAR or RESTORE v2 gRNAs into different ASGPR expressing cells, to avoid the necessity of any cell toxic transfection reagents. Especially since site-directed RNA editing arose as a highly promising post-transcriptional modification to introduce wide ranging consequences in RNA function by A-to-I base substitutions and thus, as a highly advantageous possibility to recode incorrect genetic information on the RNA level.

With the idea to enable and investigate the receptor mediated delivery of gRNAs into ASGPR expressing cells, the first-off synthesis of the reported GalNAc ligand was successful and the correct anomic configuration of the galactosamines as well as the trimeric conjugation was confirmed via ^{1}H-NMR spectroscopy as well as high resolution mass spectrometry. The synthesized GalNAc was successfully functionalized with fluorescein isothiocyanate as well as a maleimide and a glutaric acid conjugate, and a novel wet chemical synthetic route was successfully established to provide GalNAc modified gRNAs from commercially available and disulfide terminal oligonucleotides. The identity of the synthesized and GalNAc modified oligonucleotides was additionally confirmed via mass spectrometry.

Besides the GalNAc modification of gRNAs and to investigate the receptor mediated delivery of gRNAs into ASGPR expressing cells, it was further necessary generate two different receptor expressing cell lines: One for the application of the SNAP®-ADAR system and one for the recruitment of endogenous ADAR using RESTORE v2 gRNAs. A novel FlpIn™ T-REx™ 293 cell line was successfully generated, which expresses the SNAP®-ADAR1 E406Q as well as the receptor subunit H1a in a bidirectional manner. Further, a HeLa cell line was created, which exclusively expresses the receptor subunit H1a. Especially FACS sorting of the generated HeLa cell line provided a homogeneously expressing cell line with a high expression level of the receptor. The integration of the GOIs into all cell lines was confirmed utilizing different characterizations, such as the receptor mediated internalization of the fluorescently active GalNAc-FITC derivative, immunofluorescence as well as Western Blot analysis.

The functionality of both, the generated cell lines as well as the two commercially available, and GalNAc modified gRNAs regarding their editing performance was additionally confirmed via transfection-based experiments and the obtained results were consistent with the previously obtained results of the cell line characterizations. Additionally, no negative impact of the additional GalNAc modification or the H1a expression to the editing performance was observable within the transfection-based experiments of both systems. This provided a promising initial situation for the investigation of the receptor mediated internalization. However, different results were obtained for the application of the receptor mediated uptake of the two different systems.

First of all, no RNA editing was observed utilizing the receptor mediated endocytosis in combination with the SNAP®-ADAR system. While a general endocytosis was exclusively observed for the GalNAc modified and ATTO labeled gRNA by the detection of endosomal-like structures, no RNA editing was obtained using the BisBG modified equivalent. Furthermore, a nuclear localization of the ATTO labeled gRNA was exclusively observable within the transfection of gRNA, and the terminal, but absolutely essential amino modification of the SNAP®-ADAR gRNA was shown to be less stable under lysosomal conditions (< 24 h). Therefore, a less endosomal release (< 0.1 %[184,200,214,215]) in combination with a less lysosomal stability of the mandatory BisBG modification, were elaborated as most plausible and consistent arguments for the failed editing performance.

In contrast, different results were observed for the application of the RESTORE v2 gRNAs in H1a expressing HeLa cells. While only minor editing yields were observed for the application of the receptor mediated endocytosis, a highly beneficial effect of Chloroquine was observable and 2- to 5-fold increased editing yields were obtained for both used gRNAs. This was additionally reflected by the appearance of nuclear localized PS bodies after Chloroquine treatment and the endocytosis of ATTO labeled gRNAs. Furthermore, an impact of the internalization capability of both, GalNAc modified and unmodified gRNAs, was addressable to the ASGPR expression level and a beneficial effect of the additional GalNAc modification was exclusively indicated within the editing results using a HeLa cell line with a high receptor expression level, or fluorescence imaging utilizing short incubation times (0-2 h). In addition, a less lysosomal stability (< 24 h) of both used gRNAs was observed as well, and especially for the use of the unconjugated, but total PS modified RESTORE v2 gRNA, a highly competing internalization was observed, which is evidencing a still uncharacterized, but not neglectable gymnotic uptake or receptor mediated endocytosis of PS oligonucleotides. This was

additionally reflected by the highly differing internalization capabilities of the two SNAP®-ADAR gRNAs. These gRNAs do contain less PS linkages, and by this, only the GalNAc modified variant of the ATTO labeled gRNA was internalized via the receptor mediated endocytosis.

In a general description, comparable results were obtained for both, the SNAP®-ADAR and the RESTORE system. The less lysosomal stability of all used gRNAs in combination with the beneficial effect of Chloroquine as endosomal disruptor lead to a hypothesis of endosomal entrapped gRNAs, which would be in accordance with the general state-of-knowledge about the endosomal release of oligonucleotides (< 0.1 %) during endosomal maturing[184,200,214,215]. For both systems, it is possible to conclude that a sufficient amount of gRNA is productively internalized into the cells via ASGPR mediated endocytosis, but not available within the cytosol or the nucleus, due to the limited endosomal release and the less lysosomal stability. Therefore, the ASGPR mediated endocytosis of gRNAs to induce RNA editing was at least partially successful, but does not allow for the evaluation of any long-term or delayed release-based effects of a receptor mediated endocytosis utilizing gRNAs with the described designs and stabilities. Furthermore, it was not possible to determine any quantitative or stoichiometric relations of the cellular localizations due to the limitations of fluorescence microscopy as semi quantitative method, or to evaluate the differences of minor editing yields using Sanger sequencing as insensitive detection method. Therefore, further research is necessary to fathom the relationships of the receptor mediated endocytosis of gRNAs, their endosomal release as well as their nuclear concentrations, which are necessary for a sufficient editing activity.

5. Outlook

First of all, a major challenge of the established GalNAc conjugation to gRNAs, is the very low yield, which may be attributed to the necessity of two Urea-PAGE purifications. Therefore, a purification with a higher oligonucleotide recovery, such as preparative ion exchange HPLC, could be beneficial regarding the overall yield of the gRNA synthesis. Furthermore, lowering the excess of the used GalNAc-maleimide derivative could also be highly beneficial regarding the choice of purifications, such as size exclusion chromatography, which showed a less retention of GalNAc derivatives.

Another major challenge is the improvement of the lysosomal stability of the used gRNAs. Increasing the gRNA stabilities could be the key regarding their survivability during endosomal maturing, the related endosomal release and the evaluation of long-term or delayed release-based effects of a receptor mediated endocytosis. In addition, further gRNA designs with a varying modification pattern, such as less PS linkages, could also enable an increased opportunity of the GalNAc modification and the suppression of competing internalization pathways, such as gymnotic uptakes or receptor mediated internalizations of PS oligonucleotides.

Furthermore, quantification-based assays, such as qPCR or HPLC analysis using hybridization-based assays with fluorescently labeled peptide nucleic acids, could also be a relevant approach regarding the general understanding of cellular internalizations. Especially an organelle specific quantification of transfected versus endocytosed gRNAs would increase the general knowledge of gRNA localizations and by this, concentration dependent editing activities. This could also be an important aspect regarding any future therapeutic applications and the evaluation of dose-response relationships or toxicities, and the related bottlenecks during administration, distribution and tissue accumulation.

6. Methods and Materials

6.1. General approach

All chemical reactions were conducted under nitrogen atmosphere if not stated otherwise. Schlenk flasks used were heated in vacuo and cooled under nitrogen atmosphere. This procedure was repeated three times. All biological experiments were conducted using nanopure water, or if necessary, nuclease free water. All cell culture experiments were conducted under sterile conditions using sterile consumables and a laminar flow cabinet. All editing yields were calculated from the peaks heights of the adenosines divided by the sum of the guanosine and adenosine at the target site, which were obtained from Sanger Sequencing. This thesis was written with Microsoft Word 2016, Sanger sequencing data were evaluated with SnapGENE v4.2.11 and analyzed (mean, SD) and plotted with GraphPad Prism v8.4.3. All Schemes and figures were prepared with CorelDraw 2017 and ChemDraw Professional v19.1.0.8.

6.2. Consumables

6.2.1. Antibodies

Table 6: Antibodies used for immunofluorescence and western blotting.

Antibody	Product No.	Company
Goat α-*mouse* Alexa Fluor™ 488	A11001	Thermo Fisher Scientific Inc.
Goat α-*mouse* Alexa Fluor™ 647	A21235	Thermo Fisher Scientific Inc.
Goat α-*mouse* HRP	115-035-003	Jackson ImmunoResearch Laboratories, Inc.
Goat α-*rabbit* HRP	111-035-003	Jackson ImmunoResearch Laboratories, Inc.
Mouse α-ASGPR1 (8D7)	sc-52623	Santa Cruz Biotechnology, Inc.
Rabbit α-GAPDH	5174S	Cell Signaling Technology, Inc.
Rabbit α-SNAP	P9310S	New England Biolabs GmbH

6.2.2. Solvents and Chemicals

Table 7: Solvents used for chemical synthesis.

Solvent	Product No.	Company
1,4-Dioxane	296309-250mL	Sigma-Aldrich (Merck KGaA)
Acetone (pa grade)	71031003	Chemical supply of the University
$CDCl_3$	71010007	Chemical supply of the University
$CHCl_3$ (anhydrous)	364325000	Acros Organics
Cyclohexane (pa grade)	71031012	Chemical supply of the University

Table 7: Continued.

Solvent	Product No.	Company
Dichloromethane (pa grade)	71010017	Chemical supply of the University
Diethyl ether (pa grade)	71031014	Chemical supply of the University
DMSO	A994.1	Carl Roth GmbH + Co. KG
DMSO-d_6	71010019	Chemical supply of the University
Ethanol (HPLC grade)	71031020	Chemical supply of the University
Ethyl acetate (pa grade)	71031022	Chemical supply of the University
Methanol (HPLC grade)	71031031	Chemical supply of the University
N,N-Dimethyl formamide (anhydrous)	227056-100mL	Sigma-Aldrich (Merck KGaA)
Propan-2-ol (HPLC grade)	71031038	Chemical supply of the University
Toluene (HPLC grade)	71031047	Chemical supply of the University

Table 8: Chemicals used for synthesis.

Chemical	Product No.	Company
1-Ethyl-3-(3-dimethyl aminopropyl) carbodiimide	A10807	Alfa Aesar
1-Hydroxybenzotriazole (HOBt)	54802-100G-F	Sigma-Aldrich (Merck KGaA)
4-Dimethylaminopyridine (DMAP)	29224-10G	Honeywell *Fluka*™
4-Methoxybenzyl mercaptan	B22542	Alfa Aesar
5-Hexen-1-ol	A15766.06	Alfa Aesar
Acetic acid	33209-2,5L	Sigma-Aldrich (Merck KGaA)
Acrylonitrile	110213-5ML	Sigma-Aldrich (Merck KGaA)
Benzyl chloroformate	152941000	Acros Organics
Citric acid	6490.1	Carl Roth GmbH + Co. KG
Dodecyl sulfate sodium salt	CN30.1	Carl Roth GmbH + Co. KG
Ellman's reagent	22582	Thermo Fisher Scientific Inc.
Ethylenediaminetetraacetic acid (EDTA)	8040.2	Carl Roth GmbH + Co. KG
Fluorescein 5(6)-isothiocyanate	L09315	Alfa Aesar
Formic acid (FA)	33015-500mL	Sigma-Aldrich (Merck KGaA)
Glutaric anhydride	A11152	Alfa Aesar
Glycine	3908.1	Carl Roth GmbH + Co. KG

Table 8: Continued.

Chemical	Product No.	Company
H_2O_2, 30 % $(w/v)_{aq.}$	31642-1L-M	Sigma-Aldrich (Merck KGaA)
H_2SO_4 (96 %)	71802018	Chemical supply of the University
HCl_{aq} (37 %)	30721-2,5L	Honeywell *Fluka*™
Hexafluorophosphate benzotriazole tetramethyl uronium (HBTU)	AB 128869	abcr GmbH
$KMNO_4$ (pa grade)	71512110	Chemical supply of the University
KOH (pa grade)	71802003	Chemical supply of the University
LiOH	L9650-100g	Sigma-Aldrich (Merck KGaA)
Luminol	4203.1	Carl Roth GmbH + Co. KG
Methyl amine (33 wt% in ethanol)	396731000	Acros Organics
$MgSO_4$ mono hydrate	71010211	Chemical supply of the University
Mono-boc-1,3- propane amine	H50304.06	Alfa Aesar
N,N′-Diisopropylcarbodiimide	D1254407-5G	Sigma-Aldrich (Merck KGaA)
N,N-Diisopropylethylamine	AB182190	abcr GmbH
Na_2CO_3 (pa grade)	71010040	Chemical supply of the University
NaCl (pa grade)	71010043	Chemical supply of the University
$NaHCO_3$ (pa grade)	71010229	Chemical supply of the University
$NaIO_4$	13798	Alfa Aesar
$NH_4(HCOO)$	14517	Alfa Aesar
N-Hydroxysuccinimide	130672-25G	Sigma-Aldrich (Merck KGaA)
N-Succinimidyl 4-Maleimidobutyrate	S0399	TCI Deutschland GmbH
p-Anisaldehyde	A15793.14	Alfa Aesar
p-Coumaric acid	9906.1	Carl Roth GmbH + Co. KG
Pd/C (10 %, dry)	A12012	Alfa Aesar
Pentafluorophenol (PfpOH)	A15574	Alfa Aesar
p-Formaldehyde	2137.1011	Th. Geyer
$RuCl_3$	206229-1G	Sigma-Aldrich (Merck KGaA)
Tetrahydrofurane (THF)	1.08107.0500	Merck KGaA
TRI Reagent®	T9424-200ML	Sigma-Aldrich (Merck KGaA)
Triethyl amine	760.1000	Th. Geyer GmbH & Co. KG
Trifluoroacetic acid (TFA)	56508-500mL	Sigma-Aldrich (Merck KGaA)

Table 8: Continued.

Chemical	Product No.	Company
Trimethylsilyl trifluoromethanesulfonate (TMSOTf)	91741-50mL	Sigma-Aldrich (Merck KGaA)
Tris-(hydroxymethyl)-aminomethane (tris) (PUFFERAN® ≥99,9 %)	4855.5	Carl Roth GmbH + Co. KG

6.2.3. Media, buffers, solutions and additives

Table 9: Commercially available media, buffers, solutions and additives.

Component	Product No.	Company
1 kb Plus DNA Ladder (2Log Ladder)	N3200L	New England Biolabs GmbH
96-well cell imaging plates	0030741030	Eppendorf AG
Agarose NEEO Ultra-Quality	2267.4	Carl Roth GmbH + Co. KG
ATTO 594 NHS ester	AD 594-31	ATTO-TEC GmbH
APS	A3678-25G	Sigma-Aldrich (Merck KGaA)
Blasticidin S hydrochloride	CP14.4	Carl Roth GmbH + Co. KG
Bovine Serum Albumin	A2153-10G	Sigma-Aldrich (Merck KGaA)
Bromphenol blue	A512.1	Carl Roth GmbH + Co. KG
Chloroquine diphosphate salt, 98%	J64459.14	VWR International
Clarity™ Western ECL Substrate	170-5060	Bio-Rad Laboratories, Inc.
Collagen-I	A1048301	Thermo Fisher Scientific Inc.
Cover glass (⌀ 12 mm)	631-1577P	VWR International
Doxycycline	A2951,0005	AppliChem GmbH
FuGENE® 6	E2692	Promega GmbH
Geniticin™ (G418)	10131035	Thermo Fisher Scientific Inc.
Gibco™ DMEM, high glucose	41965062	Thermo Fisher Scientific Inc.
Gibco™ Fetal Bovine Serum, qualified, Brazil	10270106	Thermo Fisher Scientific Inc.
Gibco™ Zeocin™ Selection Reagent	R25001	Thermo Fisher Scientific Inc.
Hygromycin B	CP12.1	Carl Roth GmbH + Co. KG
LB Media	X968.2	Carl Roth GmbH + Co. KG
Lipofectamine™ 2000	11668019	Thermo Fisher Scientific Inc.

Table 9: Continued.

Component	Product No.	Company
Lipofectamine™ RNAiMAX	13778150	Thermo Fisher Scientific Inc.
Non-fat dry milk	A0830,0500	AppliChem GmbH
Novex™ WedgeWell™ 8 to 16%, Tris-Glycine, 1.0 mm, Mini Protein Gels	XP08165BOX	Thermo Fisher Scientific Inc.
OptiMEM™	11058021	Thermo Fisher Scientific Inc.
PageRuler™ Plus Prestained Protein Ladder (10 to 250 kDa)	26619	Thermo Fisher Scientific Inc.
Poly-D-Lysine HBr	P6407-5MG	Sigma-Aldrich (Merck KGaA)
RIPA Lysis and Extraction Buffer	89901	Thermo Fisher Scientific Inc.
ROTI® GelStain	3865.1	Carl Roth GmbH + Co. KG
ROTIPHORESE® Sequencing gel buffer concentrate	3050.1	Carl Roth GmbH + Co. KG
ROTIPHORESE® Sequencing gel concentrate	3043.1	Carl Roth GmbH + Co. KG
ROTIPHORESE® Sequencing gel diluent (50 % (w/v, 8.3 M)	3047.1	Carl Roth GmbH + Co. KG
ROTIPHORESE®10x SDS-PAGE, 5 l	3060.2	Carl Roth GmbH + Co. KG
Sep-Pak® Plus C18 cartridges	WAT020515	Waters GmbH
SYBR™ Gold Nucleic Acid Gel Stain (10,000X Concentrate in DMSO)	S11494	Thermo Fisher Scientific Inc.
TEMED	2367.3	Carl Roth GmbH + Co. KG
Triton X 100	3051.3	Carl Roth GmbH + Co. KG
Trypsin/EDTA solution	T4049-500ML	Sigma-Aldrich (Merck KGaA)
TWEEN® 20	M147-1L	VWR International
Xylene cyanol	A513.1	Carl Roth GmbH + Co. KG
Zeba™ Spin Desalting Columns, 40K MWCO, 0.5 mL	87766	Thermo Fisher Scientific Inc.

Table 10: Prepared buffers with corresponding components and concentrations*. If not stated different, nanopure water is used as solvent for all buffers.*

Buffer	Component	Final concentration
0.1 M PBS (pH 7.4)	8.0 g NaCl	137 mM
	0.2 g KCl	2.73 mM
	1.42 g $NaH_2PO_4 \cdot 2\ H_2O$	9.24 mM
	0.27 g KH_2PO_4	1.8 mM

Table 10: Continued.

Buffer	Component	Final concentration
0.5 M Na-EDTA (pH 8.0)	73.06 g EDTA Nanopure water up to 500 mL $NaOH_s$	0.5 M
10 % APS	APS (5 g)	10 % (w/v)
6 x Lämmli buffer	1.2 mL 0.5 M Tris-HCl_{aq} (pH 6.8) 0.93 g DTT 1.2 g SDS 4.7 mL Glycerol (86 %) 6 mg Bromphenol blue	60 mM 0.6 mM 12 % 47 % 0.06 %
gRNA labeling buffer (pH 8.3)	20 parts 0.1 M PBS (pH 7.4) 1 part 0.2 M $NaHCO_3$ (pH 9.0)	130.5 mM NaCl 2.6 mM KCl 8.8 mM NaH_2PO_4 1.7 mM KH_2PO_4 9.5 mM $NaHCO_3$
Homemade ECL solution (100 mL)	10 mL 1 M Tris-HCl_{aq} (pH 8.5) 500 µL 0.25 M Luminol in DMSO 220 µL 0.09 M p-Coumaric acid in DMSO	100 mM 1.25 mM 100 mM + 1 µL/mL H_2O_2 (30 % w/v) right before use
LB	25 g LB-Media Nanopure water up to 1 L	
NP40 lysis buffer	5 mL 10 % $NP40_{aq}$ 1.5 mL 5 M $NaCl_{aq}$ 5 mL 0.5 M $Tris_{aq}$ (pH 8.0) Nanopure water up to 50 mL	1 % 150 mM 50 mM
SB	0.4 g $NaOH_s$ 2.25 g $H_3BO_{3\ s}$ Nanopure water up to 1 L	10 mM 36 mM
SDS-PAGE running buffer	100 mL ROTIPHORESE®10x SDS-PAGE 900 mL Nanopure water	192 mM Glycine 25 mM Tris 0.1 % (w/v)
TAE (pH 8.3)	4.84 g Tris 1.142 mL acetic acid 2 mL 0.5 M EDTA (pH 8.0) Nanopure water up to 1 L	40 mM 20 mM 1 mM
TBE (pH 8.3)	10.8 g Tris 5.5 g $H_3BO_{3\ s}$ 4 mL 0.5 M Na-EDTA (pH 8.0) Nanopure water up to 1 L	8.9 mM 8.9 mM 0.2 mM
TE (pH 8.3)	0.8 mL 0.5 M Tris-HCl_{aq} 0.08 mL 0.5 M Na-EDTA (pH 8.0) Nanopure water up to 40 mL HCl_{aq}	10 mM 1 mM

Table 10: Continued.

Buffer	Component	Final concentration
TBS	Tris-HCl_{aq}	50 mM
	NaCl	150 mM
TBST	Tris-HCl_{aq}	50 mM
	NaCl	150 mM
	TWEEN® 20	0.1 %
Transfer buffer	14.26 g Glycine	190 mM
	3.03 g Tris	25 mM
	200 mL Methanol	
	Nanopure water up to 1 L	

6.2.4. Commercially available kits

Table 11: Commercially available kits.

Kits / Equipment	Product No.	Company
Monarch® RNA Cleanup Kit (10 µg)	T2030L	New England Biolabs GmbH
NucleoSpin® Gel and PCR Clean-up	740609.250	MACHEREY-NAGEL GmbH & Co. KG
NucleoSpin® Plasmid	740588.250	MACHEREY-NAGEL GmbH & Co. KG
One Step RT-PCR Kit	BR0400103	biotechrabbit GmbH
Pierce™ BCA Protein Assay Kit	23225	Thermo Fisher Scientific Inc.

6.2.5. Enzymes and restriction enzymes

Table 12: Commercially available enzymes and restriction enzymes.

Enzyme	Product No.	Company
AgeI-HF	R3552S	New England Biolabs GmbH
AvrII	R0174S	New England Biolabs GmbH
BamHI-HF	R3136S	New England Biolabs GmbH
ClaI	R0197S	New England Biolabs GmbH
DNase I (RNase Free)	M0303L	New England Biolabs GmbH
dNTPs	N0447L	New England Biolabs GmbH
HindIII-HF	R3104S	New England Biolabs GmbH
KpnI-HF	R3142S	New England Biolabs GmbH
M-MuLV Reverse Transcriptase	M0253L	New England Biolabs GmbH
Murine RNase Inhibitor	M0314L	New England Biolabs GmbH
NgoMIV	R0564S	New England Biolabs GmbH

Table 12: Continued.

Enzyme	Product No.	Company
NotI-HF	R3189S	New England Biolabs GmbH
PacI	R0547S	New England Biolabs GmbH
Phusion High-Fidelity DNA Polymerase	M0530 L	New England Biolabs GmbH
SalI-HF	R3138S	New England Biolabs GmbH
SpeI-HF	R3133S	New England Biolabs GmbH
T4 DNA ligase	M0202L	New England Biolabs GmbH
Taq DNA Polymerase	M0267L	New England Biolabs GmbH
XbaI	R0145S	New England Biolabs GmbH

6.2.6. Primers

Table 13: Primers used for molecular cloning, PCR amplification and sequencing.

No.	Name	Sequence
144	BGH backward	CTAGAAGGCACAGTCGAGGC
213	CMV_fw_Primer	CGCAAATGGGCGGTAGGCGTG
265	ADAR1_E/Qfw	GTGGAGAACGGACAAGGCACA ATCCCTG
280	TOPO SA1-fwd	CACCATGGACAAAGACTGCG
418	bw_fullADAR2_pTS57_BstX-I	CCAAACAGATGGCTGGCAACTA GAAGGCAC
532	psilencer_XbaI_bw	CGGTCTAGAAGCGGAAGAGCGC CCAATACGCAAACC
565	Snap-Tag end _fw	CGGGCTCGCCGTGAAAGAG
647	pTS57_CMV_fw	CATGAAGAATCTGCTTAGGGTTA GG
690	pcDNA3_postTATA_fw	GAACCCACTGCTTACTGGCTTAT CG
927	GAPDH_fw	CTCAAGATCATCAGCAATGCCTC CTGC
928	GAPDH_bw	GAGCACAGGGTACTTTATTGATG GTACATGACAAGG
1065	AH_10aaLink_Fw	CCTGCAGGCGGAGGCGCGCCAG G
1159	GAPDH_ORF_seq_bw	GCTGTTGAAGTCAGAGGAGACC
1787	ASGPR1_fw_KpnI_BamHI	GCTTGGTACCGAGCTCGGATCCA CCATGACCAAGGAGTATCAAGA CCTTCAG
1788	ASGPR1_bw_XbaI	GCCCTCTAGATTAAAGGAGAGG TGGCTCCTGGC

Table 13: Continued.

No.	Name	Sequence
1789	ASGPR2_fw_BamHI	GCTCGGATCCACCATGGCCAAG GACTTTCAAGATATCC
1790	ASGPR2_bw_XbaI	GCCCTCTAGATCAGGCCACCTC GCCGGTG
2016	AH_pcDNA5duo3_Seq1_Rev	ACTCTCTTACCCGTCATTGGC
2422	AH_AvrII_SV40-PolyA_Rev	ACACCTAGGATCTCCAGAGGA TCATAATCAGCCATACC
2423	AH_SV40-PolyA_SalI_Rev	ACGCCAAGGTCGACTTAACCC
2473	F1_Ori_bw	AGGGAAGAAAGCGAAAGGAG
2517	AH_NotI_Kozac_SNAPf	ACAGCGGCCGCCACCATGGACA AAGACTGCGAAATGAAGC
2528	CMV_3RTfw	CGTGGATAGCGGTTTGACTC
2686	AvrII_NeoR_bw	GCTCCCTAGGCGCTCAGAAGAA CTCGTCAAGAAGGC
2687	P2A_AgeI_NeoR_fw	GCTCACCGGTGACGTGGAGGAG AACCCCGGCCCCATGATTGAACA AGATGGATTGCAC
2726	ADAR1-Stop-PacI_Bw	ACTATTAATTAATCATACTGGGC AGAGATAAAAGTTCTTTTCC
2770	HindIII_TRE3GS_fw	GCTCAAGCTTTGCTTATGTAAAC CAGGG
2771	TRE3GS_SacII/NotI_bw	GGTGGCGGCCGCGGTACCTTTAC GAGGGTAGGAAGTGG
2772	ADAR1E406Q_Stop_ClaI/PacI/SpeI_bw	GAGCACTAGTTAATTAATCGATT TATACTGGGCAGAGATAAAAGTT CTTTTC
2773	ADAR2E403Q_Stop_ClaI/PacI/SpeI_bw	GAGCACTAGTTAATTAATCGATTT AGGGCGTGAGTGAGAACTG
2782	KpnI_Kozak_H1a_fw	GCTCGGTACCGCCACCATGACCA AGGAGTATCAAGACCT
2783	H1a_NotI_bw	GCTCGCGGCCGCTTAAAGGAGAG GTGGCTCCTGG
3245	H1a_ClaI_bw	GCTCATCGATTTAAAGGAGAGGT GGCTCCTG
3246	AvrII_tet2O_CMV_bw	GCTCCCTAGGCCCCAGAGTAAAG CTATTCGG
3247	AvrII_Kozak_H1a_fw	GCTCCCTAGGGCCACCATGACCA AGGAGTATCA
3248	NotI_tet2O_EF1a_bw	GCTCGCGGCCGCCCGATGGCGTC TCCAG
3499	NotI_Kozak_H1a_fw	GCTCGCGGCCGCCACCATGACC AAG
3532	STAT1_fw	GCTTCATCAGCAAGGAGCGAGA GCG

Table 13: Continued.

No.	Name	Sequence
3533	STAT1_bw	CTTCAGACACAGAAATCAACTC AGTC
3534	STAT1_sequencing	GGCTGCTGAGAATATTCCTGAG AATC

6.2.7. gRNAs

Table 14: Overview of used gRNAs. *The following chemical modification are used: * = PS linkage,* N *= RNA,* N_f *= 2'-F, N = 2'-OMe,* **N** *= DNA, {*N*} = LNA,* N *= Target site missmatch. All used gRNAs were designed by colleagues: gRNAs 257 and 258 were designed by Paul Vogel, gRNAs 324, 507, 471, 472 and 473 were by Ngadhnjim Latifi and gRNAs TMR53, TMR189 and TMR239 were designed by Tobias Merkle. Nt's = nucleotides. All gRNAs were provided form BioSpring Gesellschaft für Biotechnologie mbH (Frankfurt am Main, Germany) or Kaneka Eurogentec S.A. (Seraing, Belgium).*

No. of gRNA	Target	Sequences and base modifications (5' to 3')	No. of Nt's	Terminal modifications
257	GAPDH 3'UTR	*G*A*ACAAGGGGU***C**N**C***A**CAUGGCA*A*C*U*G*	25	5': C_6-NH_2 3': C_6-Disulfide
258	GAPDH L249L	*C*C*GAGGUUUUU***C**N**C***A**GACGGCA*G*G*U*C*	25	5': C_6-NH_2 3': C_6-Disulfide
324	STAT1 Y701C	*A*G*U*{G}*U*{C}*UUGAU***A**N**C**A**UCCAGUU*C*{C}*U*{T}*	25	5': C_6-NH_2 3': none
507	STAT1 Y701C	*A*G*U*{G}*U*{C}*UUGAU***A**N**C**A**UCCAGUU*C*{C}*U*{T}*	25	5': C_6-NH_2 3': GalNAc
471	STAT1 Y701C	{G}**U**{C}**U***U***G***A***U****A*****N*****C*****A****U***C***C***A***G***U***U***C**{C}**U**{T}	22	5': C_6-NH_2 3': C_6-Disulfide
472	STAT1 Y701C	{T}**U**{G}**A***U****A*****N*****C*****A****U***C***C***A***G**{T}**U**{C}	16	5': C_6-Disulfide 3': C_6-NH_2
473	STAT1 Y701C	{G}**U**{C}**U***U***G***A***U****A*****N*****C*****A****U***C**{C}**A**{G}	16	5': C_6-Disulfide 3': C_6-NH_2
TMR 53	GAPDH L157L	*CCAACUGCUUNGCACCCCUGGCCAAGGUCAUCCAUGACAA*	40	5': none 3': none
TMR 189	GAPDH L157L	*U*U*G**U_f*C_f*A*U_f*G*G*A*U_f*G*A*C_f*C_f*U_f*U_f*G*G*C_f*C_f*A*G*G*G*GU$_f$G**CN**AAGC$_f$*A*G*U_f*U_f*G*G*U_f*G*G*U_f*G*C_f*A*G*G*A*G*G*C_f*A*U_f*U_f**G*C*U*	59	5': C_6-NH_2 3': none
TMR 236	GAPDH L157L	*U*U*G**U_f*C_f*A*U_f*G*G*A*U_f*G*A*C_f*C_f*U_f*U_f*G*G*C_f*C_f*A*G*G*G*GU$_f$G**CN**AAGC$_f$*A*G*U_f*U_f*G*G*U_f*G*G*U_f*G*C_f*A*G*G*A*G*G*C_f*A*U_f*U_f**G*C*U*	59	5': C_6-NH_2 3': GalNAc

6.2.8. Cell lines

All cell lines (A549 (European Collection of Authenticated Cell Cultures ECACC 86012804), FlpIn™ T-REx™ 293 (Thermo Fisher Scientific Inc., Waltham (MA), USA, cat. no. R78007), HEK 293T (DSMZ Braunschweig, Germany, cat. no. ACC-635), HeLa (ATCC, Manassas (VA), USA, cat. no. ATCC CCL-2), HepG2 (DSMZ, Braunschweig, Germany, cat. no. ACC180), Huh7 (CLS GmbH, Heidelberg, Germany, cat. no. 300156), U2OS FlpIn™ T-REx™ (kind donation from Elmar Schiebel), U87MG (ATCC, Manassas (VA), USA, cat. no. ATCC HTB-14)) were cultured in Dulbecco's Modified Eagle Medium (Thermo Fisher Scientific Inc., Waltham (MA), USA, cat. no. 41965062) supplemented with 10 % fetal bovine serum (Thermo Fisher Scientific Inc., Waltham (MA), USA, cat. no. 10270106) under standard conditions (37 °C and 5 % CO_2 in a water saturated steam atmosphere) and subcultured as described below. *Human* primary hepatocytes (HUCPI, Lot: HUM4190, Lonza Group Ltd, Basel, Switzerland) were cultured and plated as recommended by the manufacturer in fresh prepared maintenance medium or plating medium, respectively and cultured under standard conditions (37 °C and 5 % CO_2 in a water saturated steam atmosphere).

6.2.9. Plasmids

Table 15: Overview of generated plasmids.

pTS No.	Insert				Restriction Enzymes	Vector			
	GOI	NCBI No. of GOI	Origin of GOI	Primer set (Cloning)		Backbone	Promotor	Resistances	Primer set (Sequencing)
689	ASGPR H1a	NM_001671.4	HepG2	1787/1788	KpnI-HF/ XbaI	pcDNA™3.1(+)*	CMV	NeoR/Amp	213/144
690	ASGPR H2b	NM_001201352.1	HepG2	1789/1790	BamHI-HF/ XbaI	pcDNA™3.1(+)	CMV	NeoR/Amp	690/144
691	ASGPR H2c	NM_080913.3	HepG2	1789/1790	BamHI-HF/ XbaI	pcDNA™3.1(+)	CMV	NeoR/Amp	213/144
1032	PiggyBac right (3') inv. Rep. NeoR	-	pTS1355	2422/2423	SalI/ AvrII	XLone-GFP-BSD**	EF1α-core	NeoR/Amp	532/2686
		-	pTS656	2686/2687	AgeI (NgoMIV)/ AvrII				
1037	SNAP®-ADAR1 E406Q TRE3G	-	pTS814	2517/2772	NotI/ SpeI	Xlone-NeoR from pTS1032	TRE3G	NeoR/Amp	1065/2016/ 2347/2770
		-	pTS1032	2770/2771	HindIII/ NotI				
1040	SNAP®-ADAR2 E488Q TRE3G	-	pTS312	2517/2773	NotI/ SpeI	XLone-NeoR from pTS1032	TRE3G	NeoR/Amp	565/2016/ 2347/2770
		-	pTS1032	2770/2771	HindIII/ NotI				
1070	ASGPR H1a	NM_001671.4	HepG2	2782/2783	KpnI-HF/ NotI	pcDNA™5/ FRT***	CMV	HygR/Amp	418/2528
1251	ASGPR H1a	NM_001671.4	HepG2	3247/3245	AvrII/ ClaI	pcDNA™5/ FRT bi-duo****	CMV	HygR/Amp	213/265/280/647/ 1065/2473/3246
	SNAP®-ADAR1 E406Q CMV-EF1α-	-	pTS814	2517/2726	NotI/ PacI		EF1α-core		
		-	pTS1084	3246/3248	AvrII/ NotI				
1340	ASGPR H1a	NM_001671.4	HepG2	3245/3499	NotI/ ClaI	XLone-NeoR from pTS1040	TRE3G	NeoR/Amp	2016/2770

*Thermo Fisher Scientific Inc., **Addgene, ***Thermo Fisher Scientific Inc., **** Stroppel, A. S.; Latifi, N.; Hanswillemenke, A.; Tasakis, R. N.; Papavasiliou, F. N.; Stafforst, T. Harnessing Self-Labeling Enzymes for Selective and Concurrent A-to-I and C-to-U RNA Base Editing. Nucleic Acids Res. 2021. https://doi.org/10.1093/NAR/GKAB541.

6.3. Analytics and Equipment

6.3.1. Nuclear magnetic resonance spectroscopy

All NMR spectra were recorded on a BRUKER Avance III HD 300 and a BRUKER Avance III HDX 400 with a frequency of 300.13 MHz or 400.13 MHz for ^{1}H-NMR and a frequency of 75.48 MHz or 100.62 MHz for broadband decoupled ^{13}C-NMR spectra at room temperature. Spectra were calibrated to the solvent signal, chemical shifts δ are reported in parts per million (ppm) and coupling constants *J* in Hertz (Hz). Signals were assigned by literature or correlation spectroscopy (^{1}H,^{1}H-COSY, ^{13}C,^{1}H-HSQC, ^{13}C,^{1}H-HMBC).

6.3.2. LCMS analysis

All LCMS analytics using positive electron spray ionization (ESI) were recorded on a SHIMADZU LC/MS system, equipped with a CBM-20A system controller, two LC-20AD solvent delivery pumps, a SIL-20AXR autosampler, an SPD-20A UV detector and a LCMS-2020 single quadrupole mass spectrometer. Analytes were separated using a Kinetex C18 column (100 x 2.1 mm, 2.6 µm, 100 Å, Phenomenex, Torrance (CA), USA) and a linear binary gradient elution from 5 % to 95 % buffer B_{LCMS} (80 % ACN + 0.1 % FA in H_2O) in buffer A_{LCMS} (H_2O + 0.1 % FA) in 12.75 min. A flow rate of 0.2 mL/min was applied and absorptions were detected at 218 nm and 280 nm. Chromatograms were evaluated using LabSolutions V5.97 SP1.

6.3.3. HPLC analysis

All HPLC analytics were recorded on a SHIMADZU Nexera X3 UHPLC system, equipped with a SCL-40 system controller, a mobile phase monitor, one DGU-405 and one DGU-403 degassing unit, two LC-40D X3 solvent delivery pumps, a SIL-40C X3 autosampler, a CTO-40C column oven, an SPD-M40 PDA detector and a RF-20Axs spectrofluorometric detector. Analytes were separated using a ReproSil-Pur Basic C18 column (125 x 4 mm, 5 µm, 100 Å, Dr. A. Maisch HPLC GmbH, Ammerbuch-Entringen, Germany) and a linear binary gradient elution from 5 %-95 % buffer B_{HPLC} (90 % ACN + 0.1 % FA in H_2O) in buffer A_{HPLC} (H_2O + 0.1 % FA) in 24 min or 5 % to 35 % buffer B_{HPLC} (90 % ACN + 0.1 % FA in H_2O) in buffer A_{HPLC} (H_2O + 0.1 % FA) in 19 min. A flow rate of 1 mL/min was applied and absorptions were detected from 200 nm to 320 nm. Chromatograms were evaluated using LabSolutions V5.97 SP1.

6.3.4. Preparative HPLC separation

All separations using preparative HPLC were performed on a SHIMADZU LC-20AT system, equipped with a SCL-10A VP system controller, two LC-20AT solvent delivery pumps and a SPD-20AV UV detector. Analytes were separated using a VP NUCLEODUR C18ec column (250 x 10 mm, 5 µm, 110 Å, MACHERY-NAGEL GmbH & Co. KG, Düren, Germany), a linear binary gradient elution using varying ratios of buffer A_{HPLC} (H_2O + 0.1 % TFA or FA) and buffer B_{HPLC} (90 % ACN in H_2O + 0.1 % TFA or FA), a flow rate of 3 ml/min and absorptions at 218 nm and 254 nm or 218 nm and 280 nm. Detailed buffer compositions for each separation are mentioned in the corresponding procedures. Chromatograms were evaluated using LabSolutions V5.97 SP1.

6.3.5. HRMS analysis

All HRMS analytics using an electron spray ionization (ESI) were recorded on a BRUKER Daltonics maXis 4G system, equipped with a TOF (time-of-flight) mass spectrometer. Analytes were separated using an UltiMate3000 (Thermo Fisher Scientific Inc., Waltham (MA), USA) equipped with a column oven, autosampler and DAD (diode array detector).

6.3.6. LCMS analysis of oligonucleotides

All MS analytics of oligonucleotides were performed in cooperation with BioSpring GmbH (Frankfurt, Germany) using electron spray ionization and were recorded on a BRUKER amaZon SL system, equipped with an ion trap mass spectrometer (ESI-ITMS). Analytes were separated using liquid chromatography and spectra were detected with a DAD (diode array detector). All spectra were evaluated using BRUKER Compass DataAnalysis 4.4.

6.3.7. UV spectroscopy

All UV spectra were recorded with a Cary 300 Scan UV/Visible spectrophotometer (Agilent Technologies, Inc., Santa Clara (CA), USA). Concentrations were calculated using the law of Lambert-Beer and the substance characteristic molar attenuation coefficient ε_λ.

6.3.8. Thin layer and column chromatography

Thin layer chromatography was performed on TLC Silica Gel 60 F_{254} aluminium sheets (Merck KGaA, Darmstadt, Germany). Column chromatography was performed using silica (60 M, 0.04-0.063 mm, MACHERY-NAGEL GmbH & Co. KG, Düren, Germany). The following staining agents were used:

$KMnO_4$	Aqueous solution of 0.5 % (w/V) $KMnO_4$

Molybdenum Blue	Solution of 5 g $(NH_4)_2MoO_4$, 0.1 g $Ce(SO_4)_2$, 10.5 mL H_2SO_4 (conc.) and 89.5 mL H_2O
Sulfuric acid/Anisaldehyde	Solution of 0.5 mL *p*-Anisaldehyde, 1 mL H_2SO_4 (conc.) and 50 mL glacial acetic acid

6.3.9. Microscopy

All microscopy images were performed using a ZEISS AXIO Observer.Z1 equipped with a Colibri.2 light source under 5x, 10x, 40x or 63x magnification. The excitation and emission wavelengths λ are listed below (Table 16). Pictures were processed using ImageJ 1.49m.

Table 16: Excitation and emission wavelengths λ.

Channel	**green**		**blue**		**red**	
	λ [nm]	**band pass filter [nm]**	**λ [nm]**	**Band pass filter [nm]**	**λ [nm]**	**band pass filter [nm]**
Excitation	488	460-488	353	350-390	587	567-602
Emission	509	500-557	465	402-448	610	615-4095

6.3.10. Western Blot imaging

Chemiluminescence of HRP conjugated western blot membranes was detected using a Fusion S2 Vilber Lourmat imaging system equipped with a CCD camera (Peqlab, VWR Life Science) or an Odyssey Fc Imaging System (LI-COR® Biosciences, Lincoln (NE), USA). Pictures were processed using ImageJ 1.49m.

6.3.11. Nucleic acid and protein quantification

All nucleic acid quantifications were determined using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham (MA), USA) or a Spark 10M Luminescence Multi Mode Microplate Reader (Tecan Group AG, Männedorf, Switzerland). For isolated RNA and PCR amplicons, concentrations were calculated using device internal parameters. Concentrations of oligonucleotides were calculated using the law of Lambert-Beer, the substance characteristic molar attenuation coefficient ε_{260} and the absorbance at a wavelength λ of 260 nm. For each calculation, the used solvent was applied as reference.

Protein concentrations were determined using a Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific Inc., Waltham (MA), USA) according to the manufacturer's protocol with bovine serum albumin (BSA) as standard in concentrations of 0 to 1500 mg/mL. Measurements were carried out using a Spark 10M Luminescence Multi Mode Microplate Reader (Tecan

Group AG, Männedorf, Switzerland) and the absorbance at a wavelength λ of 562 nm was determined. For each calculation, the used cell lysis buffer was applied as reference.

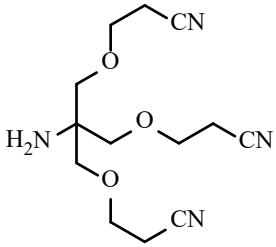
6.3.12. Fluorescent imaging of Urea-PAGE

Separated oligonucleotides by Urea-PAGE were stained using SYBR™ Gold Nucleic Acid Gel Stain (10,000 X Concentrate in DMSO, Thermo Fisher Scientific Inc., Waltham (MA), USA) for 20 min in TBE buffer according to the manufacturer's protocol. The stained oligonucleotides were visualized using a FLA-5100 Fluorescent Image Analyzer (FUJI PHOTO FILM Co., Ltd., Tokyo, Japan) and fluorescence was detected with an excitation wavelength λ_{ex} of 473 nm using a long pass blue (LPB, Y510) filter.

6.4. Synthesis of triantennary GalNAc

6.4.1. Synthesis of compound 2[293–295]

Tris (100.00 mmol) was added to a solution of 1,4-Dioxane (20 mL) and KOH_{aq} (15 mmol, dissolved in 1.26 mL H_2O) to give a colorless suspension. Acylonitrile (350.00 mmol) was added dropwise over 1.5 h, whereby the suspension slowly dissolved over 4 h. The solution turned slight yellow and was stirred overnight at RT. The turned dark orange reaction was neutralized with 2.5 M HCl_{aq} and the formed suspension was filtered. The filtrate was reduced in vacuo and precipitated in DCM (100 mL). The precipitate was filtered and the organic layer was washed with brine (4 x 50 mL), dried over $MgSO_4$ and reduced in vacuo to give the crude product (22.89 g) as dark yellow oil. The crude product was purified by chromatography (silica, MeOH in DCM, 0-5 %) to give the product (14.22 g, 50.7 mmol, 51 %) as yellow oil.



$C_{13}H_{20}N_4O_3$
280.15

2

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 1.67 (s, 2H, N*H*$_2$), 2.57 (t, *J* = 6 Hz, 6H, C*H*$_2$CN), 3.40 (s, 6H, NC*H*$_2$OCH$_2$), 3.64 (t, *J* = 6 Hz, 6H, OC*H*$_2$CH$_2$)

^{13}C-NMR {^{1}H} (100 MHz, CDCl3): δ (ppm) = 118.3 (*C*N), 72.7 (*C*H$_2$ tris), 65.9 (O*C*H$_2$CH$_2$CN), 56.3 (*C*$_q$ tris), 19.0 (*C*H$_2$CN)

MS (ESI): *m/z* = 281.4 $[M+H]^+$, 303.2 $[M+Na]^+$

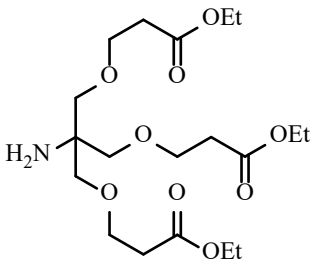
t_R (HPLC) = 2.27 min (5-95 % Buffer B_{HPLC})

R_f ($KMnO_4$) = 0.53 (MeOH in DCM, 12 % + 0.1 % Et_3N)

The spectra are in accordance with the literature.

6.4.2. Synthesis of compound 3[293–295]

Compound **2** (23.74 mmol) was diluted in 100 mL ethanol, and 43.07 mL conc. H_2SO_4 (96 %, 807.16 mmol) was added dropwise at 0 °C. The solution was stirred for 7 h under reflux and overnight at RT. The formed suspension was quenched with water, neutralized with 1 M $NaOH_{aq}$ and concentrated in vacuo. The residue was dissolved in water and extracted with DCM (3 x 100 mL). The combined organic layer was dried over $MgSO_4$ and reduced in vacuo to give the crude product as dark yellow oil. The crude product was purified by chromatography (silica, MeOH in toluene, 10 % + 1 % Et_3N) to give the product (4.74 g, 11.25 mmol, 47 %) as yellow oil.



$C_{19}H_{35}NO_9$
421.23

3

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 1.22 (t, *J* = 7 Hz, 9H, CH_2CH_3, 2.34 (s, 2H, NH_2), 2.50 (t, *J* = 6.3 Hz, 6H, CH_2C=O), 3.31 (s, 6 H, CCH_2O), 3.65 (t, *J* = 6.3 Hz, 6H, OCH_2CH_2), 4.10 (q, *J* = 7 Hz, 6H, OCH_2CH_3

^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$): δ (ppm) = 14.3 (CH_2CH_3), 35.1 (CH_2C=O), 56.4 (C_q tris), 60.5 (CH_2CH_3), 66.9 (OCH_2CH_2), 72.4 (C_qCH_2O), 171.7 (*C*=O)

MS (ESI): m/z = 422.5 $[M+H]^+$

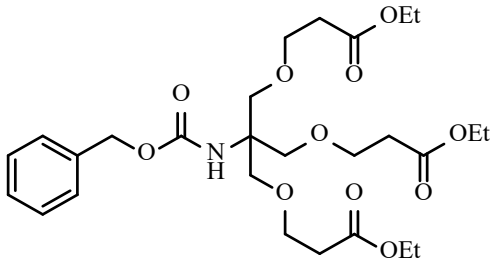
t_R (HPLC) = 7.73 min (5-95 % Buffer B_{HPLC})

R_f (KMnO4) = 0.88 (MeOH in toluene, 50 % + 0.1 % Et3N)

The spectra are in accordance with the literature.

6.4.3. Synthesis of compound 4[296]

Compound **3** (11.25 mmol) was diluted in 30 mL 1,4-Dioxane and Na_2CO_3 (12.38 mmol, dissolved in 9.5 mL H_2O) was added dropwise. Benzyl chloroformate (12.38 mmol) was added dropwise at 0 °C and the suspension was stirred overnight at RT. The suspension was reduced in vacuo and the aqueous residue was diluted in EtOAc (50 mL) and extracted with EtOAc (3 x 50 mL). The combined organic layer was washed with brine, dried over $MgSO_4$ and reduced in vacuo to the give the crude product as slight yellow oil (6.45 g). The crude product was purified by chromatography three times (silica, EtOAc in cyclohexane, 25-33 % + 0.5 % Et_3N, MeOH in DCM, 10 % + 0.5 % Et_3N, MeOH in toluene 0-10 %) to give the product (5.21 g, 9.38 mmol, 83 %) as a colorless oil.



$C_{27}H_{41}NO_{11}$
555.27

4

^{1}H-NMR (300 MHz, $CDCl_3$): δ (ppm) = 1.23 (t, *J* = 7 Hz, 9H, CH_2CH_3), 2.51 (t, *J* = 6.3 Hz, 6H, CH_2C=O), 3.67 (t + s, 12H, OCH_2CH_2 + CCH_2O) 4.11 (q, *J* = 7 Hz, 6H, OCH_2CH_3), 5.02 (s, 2H, $ArCH_2O$), 5.24 (s, 1H, N*H*), 7.27-7.38 (m, 5H, Ar-*H*)

^{13}C-NMR {^{1}H} (75 MHz, $CDCl_3$): δ (ppm) = 14.3 (CH_2CH_3), 35.1 (CH_2C=O), 58.8 (C_q tris), 60.5 (CH_2CH_3), 66.2 ($ArCH_2O$), 66.9 (OCH_2CH_2), 69.5 (C_qCH_2O), 128.0, 128.1, 128.5, 136.8 ($C_{Ar}H$), 155.2 (O*C*=ONH), 171.6 (*C*=O)

MS (ESI): *m/z* = 556.2 $[M+H]^+$, 578.2 $[M+Na]^+$

t_R (HPLC) = 19.01 min (5-95 % Buffer B_{HPLC})

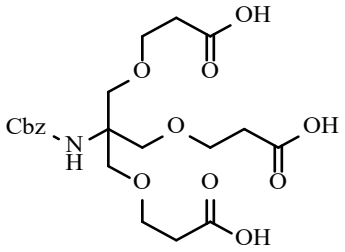
R_f (Molybdenum Blue) = 0.20 (EtOAc in cyclohexane, 20 % + 0.1 % Et_3N)

R_f (Molybdenum Blue) = 0.74 (MeOH in DCM, 9 % + 0.1 % Et_3N)

The spectra are in accordance with the literature.

6.4.4. Synthesis of compound 5[296]

Compound **4** (1.76 mmol) was diluted in 5 mL MeOH, LiOH (18.48 mmol, dissolved in 4 mL THF:H_2O 1:1) was added dropwise at 0 °C and the solution was stirred for 5 h at RT. The reaction was reduced in vacuo, diluted with 20 mL H_2O, acidified with 2.5 M HCl_{aq} to pH 1 and extracted with EtOAc (3 x 50 mL). The combined organic layer was washed with brine, dried over MgSO4 and reduced in vacuo to the give the crude product as colorless oil (0.79 g, 1.68 mmol, 95 %). The crude product was used without further purification.



$C_{21}H_{29}NO_{11}$
471.17

5

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 2.56 (t, *J* = 6.3 Hz, 6H, C*H*$_2$C=O), 3.67 (t + s, 12H, OC*H*$_2$CH$_2$ + CC*H*$_2$O), 5.04 (s, 2H, ArC*H*$_2$O), 5.30 (br, 1H, N*H*), 7.28-7.37 (m, 5H, Ar-*H*), 9.38 (br, COO*H*)

MS (ESI): *m/z* = 472.00 $[M+H]^+$, 493.90 $[M+Na]^+$, 427.95 $[M+H-CO_2]^+$

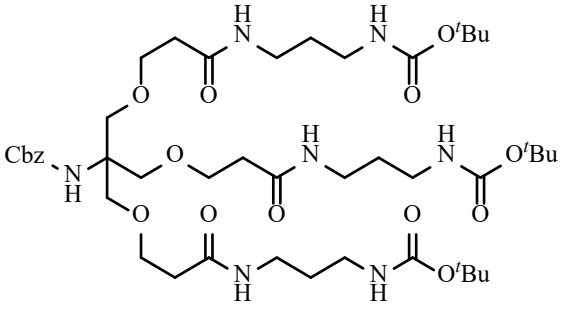
t_R (HPLC) = 10.25 min (5-95 % Buffer B_{HPLC})

R_f (Molybdenum Blue) = 0.33 (MeOH in DCM, 9 %)

The spectra are in accordance with the literature.

6.4.5. Synthesis of Compound 6[253]

Compound **5** (7.10 mmol) and mono-boc 1,3-propane diamine (28.4 mmol) were diluted in 100 mL abs. DMF. HBTU (22.0 mmol) and DIPEA (42.6 mmol) was added and the reaction was stirred overnight at RT (TLC). The reaction was poured into ice water and extracted with DCM (3 x 100 mL). The combined organic layer was washed with sat. $NaHCO_{3\ aq}$ and brine, dried over $MgSO_4$ and concentrated in vacuo to give the crude product as dark yellow oil (14.15 g). The crude product was purified by chromatography (silica, MeOH in DCM, 0-10 %) to give the product as colorless solid/foam (6.02 g, 6.41 mmol, 90 %).



$C_{45}H_{77}N_7O_{14}$
939.55

6

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.36 (s, 27H, C(C*H*$_3$)$_3$), 1.48 (q, 6H, *J* = 7.2 Hz), 2.27 (t, *J* = 6.2 Hz, 6H), 2.91 (m, 6H), 3.01 (m, 6H), 3.48 (br, 6H), 3.54 (t, *J* = 6.3 Hz, 6H), 4.97 (s, 2H), 6.51 (br, 1H, OC=ON*H*C$_q$), 6.74 (t, *J* = 5.5 Hz, 3H, CH$_2$N*H*COO), 7.28-7.38 (m, 5H), 7.0 (t, *J* = 5.5 Hz, 3H, CH$_2$C=ON*H*CH$_2$)

MS (ESI): *m/z* = 940.50 $[M+H]^+$, 962.50 $[M+Na]^+$, 840.50 $[M+2H-CO_2{}^t Bu]^{2+}$

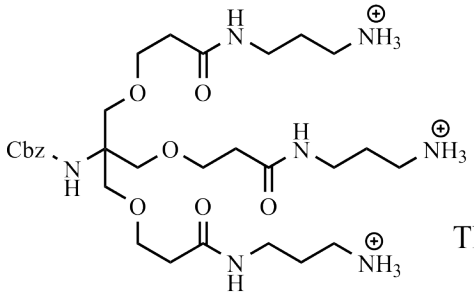
t_R (HPLC) = 16.75 min (5-95 % Buffer B_{HPLC})

R_f ($KMnO_4$) = 0.34 (MeOH in DCM, 10 %)

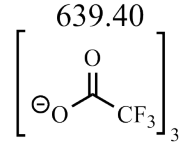
The spectra are in accordance with the literature.

6.4.6. Synthesis of Compound 7[253]

Compound **6** (0.32 mmol) was diluted in 3 mL abs. $CHCl_3$, TFA (9.58 mmol) was added dropwise at 0 °C and the reaction was stirred for 1 h at RT. The reaction was diluted in 25 mL MeOH and concentrated under reduced pressure. The residue was co-evaporated 4 times with 25 mL MeOH, concentrated in vacuo and dried using a high vacuum pump to give the crude product as TFA salt as colorless oil (0.33 g, 0.33 mmol, 103 %). The crude product was used without further purification.



$C_{30}H_{53}N_7O_8$
639.40



TFA salt: $C_{36}H_{56}N_7O_{14}F_9$
981.37

7

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.67 (quint, *J* = 7.4 Hz, 6H, CH$_2$C*H*$_2$CH$_2$), 2.30 (t, *J* = 6.4 Hz, 6H, C*H*$_2$C=ONH), 2.77 (m, 6H, CH$_2$C*H*$_2$NH$_3^+$), 3.11 (m, 6H, C=ONHC*H*$_2$), 3.48 (br, 6H, C$_q$C*H*$_2$O), 3.56 (t, *J* = 6.4 Hz, 6H, OC*H*$_2$CH$_2$), 4.98 (s, 2H, ArC*H*$_2$), 6.53 (br, 1H,

OC=ON*H*C_q), 7.28-7.38 (m, 5H, Ar*H*), 7.88 (br, 9H, $CH_2NH_3^+$), 8.07 (t, *J* = 5.7 Hz, 3H, CH_2C=ON*H*CH_2)

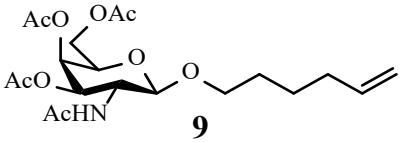
^{13}C-NMR {^{1}H} (100 MHz, DMSO-d_6): δ (ppm) = 27.5 ($CH_2CH_2CH_2$), 35.6 (C=ONHCH_2), 36.0 (CH_2C=ONH), 36.8 ($CH_2CH_2NH_3^+$), 58.9 (C_q tris), 64.9 (ArCH_2), 67.3 (OCH_2CH_2), 68.3 (C_qCH_2O), 116.7 (q, ^1J(C,F) = 296 Hz, CF_3COO^-), 127.6, 127.8, 128.4, 137.2, (C_{Ar}-H), 154.6 (O*C*=ONH), 158.6 (q, ^2J(C,F) = 33 Hz, CF_3COO^-), 170.7 (*C*ONH)

MS (ESI): *m/z* = 640.50 $[M+H]^+$, 320.85 $[M+2H]^{2+}$, 214.20 $[M+3H]^{3+}$

The spectra are in accordance with the literature.

6.4.7. Synthesis of Compound 9[254,297]

Commercially available β-D-Galactosamine penta acetate (**8**) (25.75 mmol) was diluted in 5 mL abs. $CHCl_3$, equipped with freshly activated molecular sieve (3 Å). TMSOTf (28.34 mmol) was added dropwise and the solution was stirred for 70 h at RT. 5-Hexen-1-ol (77.25 mmol) was added dropwise and stirring at RT was continued for 4 h. The reaction was neutralized with Et_3N, filtered and concentrated in vacuo to give the crude product as slight yellow oil. The crude product was purified by chromatography (silica, Acetone in DCM, 10-20 %) to give the product as colorless foam (6.52 g, 15.19 mmol, 59 %).



9

$C_{20}H_{31}NO_9$
429.20

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.32-1.39 (m, 2H, CH_2), 1.44-1.51 (m, 2H, CH_2), 1.76 (s, 3H, C=OCH_3), 1.89 (s, 3H, C=OCH_3), 1.99-2.03 (m, 5H, C=OCH_3, CH_2), 2.10 (s, 3H, C=OCH_3), 3.42 (dt, *J* = 6.6 Hz, 9.9 Hz, 1H, OCH_2CH_2), 3.71 (dt, *J* = 6.6 Hz, 9.9 Hz, 1H, OCH_2CH_2), 3.87 (dt, *J* = 8.9 Hz, 11.9 Hz, 1H, H-2) 4.00-4.05 (m, 3H, H-5, H-6, H-6'), 4.48 (d, *J* = 8.4 Hz, 1H, H-1), 4.93-5.02 (m, 3H,CH=CH_2, H-3), 5.21 (d, *J* = 3.4 Hz, 1H, H-4), 5.73-5.83 (m, 1H,C*H*=CH_2), 7.81 (d, *J* = 9.2 Hz, 1H, N*H*)

^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 1.41 (m, 2H, CH_2CH_2CH), 1.57 (m, 2H, OCH_2CH_2), 1.92 (s, 3H, NHC=OCH_3), 1.97 (s, 3H, OC=OCH_3), 2.02-2.03 (m, 5H, OC=OCH_3, CH_2CH=CH_2), 2.11 (s, 3H, OC=OCH_3), 3.46 (m, 1H, *H*-5), 3.83-3.94 (m, 3H, *H*-2, *H*-6, *H*-6'), 4.12 (m, 2H, OCH_2), 4.68 (d, *J* = 8 Hz, 1H, *H*-1), 4.90-4.99 (m, 2H, CH=CH_2), 5.27-5.34 (m, 2H, *H*-3, *H*-4), 5.68-5.81 (m, 2H, N*H*, C*H*=CH_2)

^{13}C-NMR {^{1}H} (100 MHz, $CDCl_3$): δ (ppm) = 20.8 (C=O*C*H$_3$), 23.5 (NHC=O*C*H$_3$), 25.2 (*C*H$_2$CH$_2$CH), 28.9 (OCH$_2$*C*H$_2$), 33.4 (CH$_2$*C*H$_2$CH), 51.8 (C$_{H\text{-}2}$), 61.6 (O*C*H$_2$), 66.9 (*C*$_{H\text{-}4}$), 69.8 (*C*$_{H\text{-}5}$), 70.0 (*C*$_{H\text{-}3}$), 70.6 (*C*$_{H\text{-}6}$), 101.0 (*C*$_{H\text{-}1}$), 114.7 (CH=*C*H$_2$), 138.6 (*C*H=CH$_2$), 170.4, 170.4, 170.6 (*C*=OCH$_3$)

MS (ESI): *m/z* = 430.5 $[M+H]^+$, 452.1 $[M+Na]^+$, 330.0 [3,4,6-*O*-Acetyl-*N*-Acetyl-Galactosamine$^{\cdot}$]

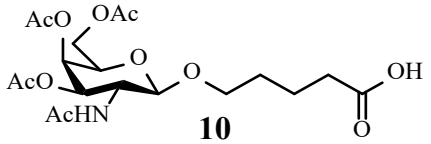
t_R (HPLC) = 12.56 min (5-95 % Buffer B$_{HPLC}$)

R_f (Sulfuric acid/Anisaldehyde) = 0.66 (Acetone in DCM, 25 %)

The spectra are in accordance with the literature.

6.4.8. Synthesis of Compound 10[254]

Compound **9** (15.20 mmol) was dissolved in 105 mL of a mixture of DCM/ACN/H_2O (2:2:3) and the reaction was cooled to < 10 °C. Solid $NaIO_4$ (60.80 mmol) was added portion wise and the reaction was stirred at < 10 °C for 15 min. $RuCl_3$ (cat., 0.26 mmol) was added to the cold reaction and the temperature was maintained below 35 °C during the addition. The reaction was stirred for 1 h at RT, additional solid $NaIO_4$ (15.20 mmol) was added and stirring was continued for 1.5 h at RT. The complete reaction was diluted with water (100 mL) and DCM (100 mL) and the pH was adjusted to 7.5 by adding solid $NaHCO_3$. The dark organic layer was removed, the aqueous layer was washed with DCM (3 x 50 mL) and the organic extracts were discarded. The aqueous layer was acidified to pH 3 by the addition of solid citric acid, and extracted with DCM (10 x 100 ml). The combined organic layer was washed with brine, dried over $MgSO_4$ and reduced in vacuo to give the crude product as orange foam (4.71 g). The crude product was purified by chromatography (silica, MeOH in DCM, 5 % + 1 % AcOH) to give the product as colorless oil/foam (4.24 g, 9.48 mmol, 62 %).



10

$C_{19}H_{29}NO_{11}$
447.17

^{1}H-NMR (400 MHz, DMSO-d$_6$): δ (ppm) = 1.48-1.50 (m, 4H, OCH$_2$C*H*$_2$C*H*$_2$), 1.77 (s, 3H, C=OC*H*$_3$), 1.89 (s, 3H, C=OC*H*$_3$), 1.99 (s, 3H, C=OC*H*$_3$), 2.10 (s, 3H, C=OC*H*$_3$), 2.19 (t, *J* = 7.2 Hz, 2H, C*H*$_2$COOH), 3.39-3.44 (m, 1H, OC*H*$_2$CH$_2$), 3.68-3.73 (m, 1H, OC*H*$_2$CH$_2$), 3.87 (dt, *J* = 8.9 Hz, 11.2 Hz, 1H, *H*-2), 4.01-4.05 (m, 3H, *H*-5, *H*-6, *H*-6'), 4.48 (d, *J* = 8.5 Hz, 1H,

H-1), 4.96 (dd, *J* = 3.4 Hz, 11.2 Hz, 1H, *H*-3), 5.21 (d, *J* = 3.3 Hz, 1H, *H*-4), 7.80 (d, *J* = 9.2 Hz, 1H, N*H*), 11.97 (br, 1H, COO*H*)

^{13}C-NMR {^{1}H} (100 MHz, DMSO-d_6): δ (ppm) = 20.5, 20.5, 20.5, 21.0, 22.7, 28.4, 33.3, 49.4, 61.5, 66.7, 68.5, 69.9, 70.5, 101.0, 169.3, 169.7, 170.0, 170.0. 174.5

MS (ESI): *m/z* = 448.1 $[M+H]^+$, 470.1 $[M+Na]^+$, 330.0 [3,4,6-*O*-Acetyl-*N*-Acetyl-Galactosamine$^+$]

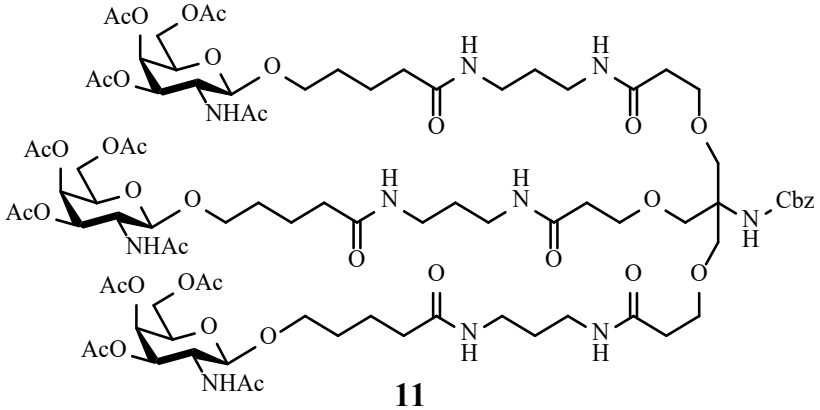
t_R (HPLC) = 7.75 min (5-95 % Buffer B_{HPLC})

R_f (Sulfuric acid/Anisaldehyde) = 0.47 (MeOH in DCM, 9 % + 1% AcOH)

The spectra are in accordance with the literature.

6.4.9. Synthesis of Compound 11 in a small scale approach[253,254]

Compound **10** (0.58 mmol) was dissolved in 2.5 mL abs. DMF. HBTU (0.65 mmol), HOBt (0.65 mmol) and DIPEA (1.86 mmol) was added and the reaction was stirred for 15 min at RT. Compound **11** (0.166 mmol, dissolved in 0.5 mL abs. DMF) was added and the reaction was stirred overnight at RT. The reaction was reduced in vacuo, the residue was dissolved in water (20 mL) and DCM (30 mL) and the aqueous phase was extracted with DCM (3 x 30 mL). The combined organic layer was washed with sat. $NaHCO_{3\ aq}$, water and brine, dried over $MgSO_4$ and reduced in vacuo to give the crude product as yellow oil (401.1 mg). The crude product was purified by chromatography (silica, MeOH in DCM 0-20%) to give the product as colorless oil (226.6 mg, 0.117 mmol, 71 %).



$C_{87}H_{134}N_{10}O_{38}$
1926.89

11

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.44-1.53 (m, 18H, $NHCH_2C\mathit{H}_2CH_2NH$, $OCH_2C\mathit{H}_2C\mathit{H}_2CH_2C{=}ONH$), 1.77 (s, 9H, $C{=}OC\mathit{H}_3$), 1.88 (s, 9H, $C{=}OC\mathit{H}_3$), 1.99 (s, 9H, $C{=}OC\mathit{H}_3$), 2.04 (t, *J* = 7.0 Hz, 6H, $OCH_2CH_2CH_2C\mathit{H}_2C{=}ONH$), 2.10 (s, 9H, $C{=}OC\mathit{H}_3$), 2.28 (t, *J* = 6.4 Hz, 6H, $OCH_2C\mathit{H}_2C{=}ONH$), 3.00-3.06 (m, 12H, $C{=}ONHC\mathit{H}_2CH_2C\mathit{H}_2NHC{=}O$), 3.38-

3.43 (m, *J* = 6.3 Hz, 9.9 Hz, 3H, $C_{anomer}OCH_2CH_2$), 3.48 (br, 6H, C_qCH_2O), 3.54 (t, *J* = 6.3 Hz, 6H, $OCH_2CH_2C{=}ONH$), 3.67-3.72 (m, 3H, $C_{anomer}OCH_2CH_2$), 3.87 (dt, *J* = 8.9 Hz, 11.1 Hz, 3H, *H*-2), 3.99-4.04 (m, 9H, *H*-5, *H*-6, *H*-6'), 4.49 (d, *J* = 8.5 Hz, 3H, *H*-1), 4.95-4.99 (m, 5H, *H*-3, $ArCH_2$), 5.21 (d, *J* = 3.4 Hz, 3H, *H*-4), 6.53 (br, 1H, $OC{=}ONHC_q$), 7.27-7.37 (m, 5H, Ar*H*), 7.77 (t, *J* = 5.6 Hz, 3H, $OCH_2CH_2C{=}ONH$), 7.84-7.87 (m, 6H, $NHC{=}OCH_3$, $CH_{2,GalNAc}C{=}ONHCH_2$)

^{13}C-NMR {^{1}H} (100 MHz, DMSO-d_6): δ (ppm) = 20.4, 20.5, 20.5, 21.8, 22.8, 28.6, 29.3, 35.0, 36.1, 36.3, 36.4, 49.4, 58.8, 61.4, 66.7, 67.3, 68.3, 68.7, 69.8, 70.5, 101.0, 127.5, 127.7, 128.3, 137.2, 154.6, 169.4, 169.6, 169.9, 170.0, 170.1, 172.0

MS (ESI): *m/z* = 1927.5 $[M+H]^+$, 1928.9 $[M+2H]^{2+}$, 1929.9 $[M+3H]^{3+}$, 1949.6 $[M+Na]^+$, 1950.1 $[M+H+Na]^{2+}$, 1950.9 $[M+Na+2H]^{3+}$, 964.8 $[M+2H]^{2+}$, 643.6 $[M+3H]^{3+}$, 330.1 [3,4,6-*O*-Acetyl-*N*-Acetyl-Galactosamine$^+$]

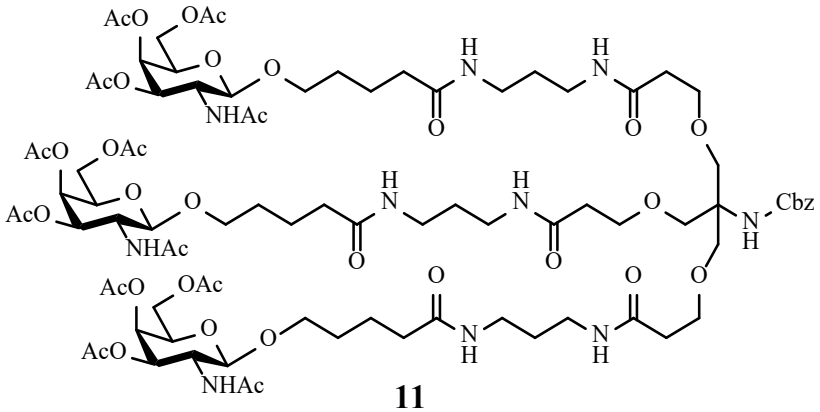
t_R (HPLC) = 12.52 min (5-95 % Buffer B_{HPLC})

R_f (Sulfuric acid/Anisaldehyde) = 0.11 (MeOH in DCM, 9 %)

The spectra are in accordance with the literature.

6.4.10. Synthesis of Compound 11 in a large scale approach[253,254]

Compound **10** (9.48 mmol) was dissolved in 40 mL abs. DMF, HBTU (10.57 mmol), HOBt (10.57 mmol) and DIPEA (30.35 mmol) was added and the reaction was stirred for 15 min at RT. Compound **11** (2.71 mmol) was added in abs. DMF (~ 10 mL) and the reaction was stirred overnight at RT. The reaction was reduced in vacuo, the residue was dissolved in water (30 mL) and DCM (50 mL) and the aqueous phase was extracted with DCM (3 x 50 mL). The combined organic layer was washed with sat. $NaHCO_{3\ aq}$, water and brine, dried over $MgSO_4$ and reduced in vacuo to give the crude product as yellow oil (10.27 g). The crude product was purified by chromatography (silica, MeOH in DCM, 0-20%) to give the product as colorless foam (0.45 g, 0.234 mmol, 9 %)



$C_{87}H_{134}N_{10}O_{38}$
1926.89

11

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.44-1.53 (m, 18H, NHCH$_2$C*H*$_2$CH$_2$NH, OCH$_2$C*H*$_2$C*H*$_2$CH$_2$C=ONH), 1.77 (s, 9H, C=OC*H*$_3$), 1.89 (s, 9H, C=OC*H*$_3$), 1.99 (s, 9H, C=OC*H*$_3$), 2.04 (t, *J* = 7.0 Hz, 6H, OCH$_2$CH$_2$CH$_2$C*H*$_2$C=ONH), 2.10 (s, 9H, C=OC*H*$_3$), 2.27 (t, *J* = 6.4 Hz, 6H, OCH$_2$C*H*$_2$C=ONH), 2.99-3.05 (m, 12H, C=ONHC*H*$_2$CH$_2$C*H*$_2$NHC=O), 3.37-3.43 (m, *J* = 6.3 Hz, 9.9 Hz, 3H, C_{anomer}OC*H*$_2$CH$_2$), 3.48 (br, 6H, C_qC*H*$_2$O), 3.54 (t, *J* = 6.3 Hz, 6H, OC*H*$_2$CH$_2$C=ONH), 3.67-3.72 (m, 3H, C_{anomer}OC*H*$_2$CH$_2$), 3.87 (dt, *J* = 8.9 Hz, 11.1 Hz, 3H, *H*-2), 4.00-4.04 (m, 9H, *H*-5, *H*-6, *H*-6'), 4.49 (d, *J* = 8.5 Hz, 3H, *H*-1), 4.95-4.98 (m, 5H, *H*-3, ArC*H*$_2$), 5.21 (d, *J* = 3.4 Hz, 3H, *H*-4), 6.53 (br, 1H, OC=ON*H*C_q), 7.27-7.38 (m, 5H, Ar*H*), 7.74 (t, *J* = 5.5 Hz, 3H, OCH$_2$CH$_2$C=ON*H*), 7.82-7.85 (m, 6H, N*H*C=OCH$_3$, $CH_{2,GalNAc}$C=ON*H*CH$_2$)

MS (ESI): *m/z* = 1927.8 $[M+H]^+$, 1928.8 $[M+2H]^{2+}$, 965.0 $[M+2H]^{2+}$, 643.6 $[M+3H]^{3+}$

R_f (Sulfuric acid/Anisaldehyde) = 0.11 (MeOH in DCM, 9 %)

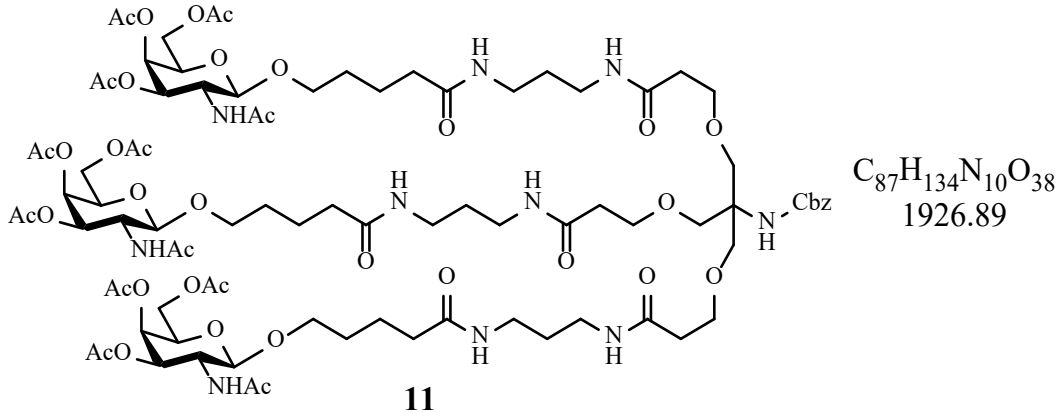
^{1}H-NMR of the unsaturated conjugate (side product):

^{1}H-NMR (400 MHz, DMSO-d_6, found/expected protons): δ (ppm) = 1.17-1.18 (br, 6/0H, probably N*H*$_2$), 1.44-1.53 (m, 15/18H, NHCH$_2$C*H*$_2$CH$_2$NH, OCH$_2$C*H*$_2$C*H*$_2$CH$_2$C=ONH), 1.77 (s, 6/9H, C=OC*H*$_3$), 1.89 (s, 6/9H, C=OC*H*$_3$), 1.99 (s, 6/9H, C=OC*H*$_3$), 2.04 (t, *J* = 7.0 Hz, 4/6H, OCH$_2$CH$_2$CH$_2$C*H*$_2$C=ONH), 2.10 (s, 6/9H, C=OC*H*$_3$), 2.27 (t, *J* = 6.4 Hz, 6/6H, OCH$_2$C*H*$_2$C=ONH), 2.99-3.06 (m, 12/12H, C=ONHC*H*$_2$CH$_2$C*H*$_2$NHC=O), 3.37-3.43 (m, 3/3H, C_{anomer}OC*H*$_2$CH$_2$), 3.48 (br, 6/6H, C_qC*H*$_2$O), 3.54 (t, *J* = 6.3 Hz, 6/6H, OC*H*$_2$CH$_2$C=ONH), 3.67-3.72 (m, 2/3H, C_{anomer}OC*H*$_2$CH$_2$), 3.87 (dt, *J* = 8.9 Hz, 11.1 Hz, 2/3H, *H*-2), 4.00-4.04 (m, 6/9H, *H*-5, *H*-6, *H*-6'), 4.48 (d, *J* = 8.5 Hz, 2/3H, *H*-1), 4.95-4.98 (m, 4/5H, *H*-3, ArC*H*$_2$), 5.21 (d, *J* = 3.4 Hz, 2/3H, *H*-4), 6.53 (br, 1/1H, OC=ON*H*C_q), 7.20-7.38 (m, 7/5H,

Ar*H*), 7.73 (t, *J* = 5.6 Hz, 3/3H, $OCH_2CH_2C{=}ONH$), 7.77-7.84 (m, 7/6H, $NHC{=}OCH_3$, $CH_{2,GalNAc}C{=}ONHCH_2$)

6.4.11. Synthesis of Compound 11 using active ester 12[298]

Compound **12** (0.366 mmol) was dissolved in 2 mL abs. DMF. Compound **7** (0.105 mmol) was dissolved in 2 mL abs. DMF and the solution was added. DIPEA (1.89 mmol) was added and the reaction was stirred overnight at RT. DMF and DIPEA were removed in vacuo to give the crude product as slight yellow oil (432.6 mg). The crude product was purified by chromatography (silica, MeOH in DCM, 5-18 %) to give the product as colorless foam (203.4 mg, 0.105 mmol, 99 %).



11

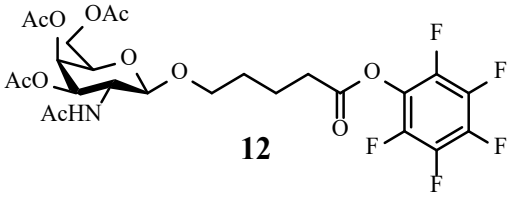
^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.44-1.53 (m, 18H, $NHCH_2C\mathit{H}_2CH_2NH$, $OCH_2C\mathit{H}_2C\mathit{H}_2CH_2C{=}ONH$), 1.77 (s, 9H, $C{=}OC\mathit{H}_3$), 1.89 (s, 9H, $C{=}OC\mathit{H}_3$), 1.99 (s, 9H, $C{=}OC\mathit{H}_3$), 2.04 (m, 6H, $OCH_2CH_2CH_2C\mathit{H}_2C{=}ONH$), 2.10 (s, 9H, $C{=}OC\mathit{H}_3$), 2.27 (t, *J* = 6.4 Hz, 6H, $OCH_2C\mathit{H}_2C{=}ONH$), 3.00-3.06 (m, 12H, $C{=}ONHC\mathit{H}_2CH_2C\mathit{H}_2NHC{=}O$), 3.39-3.43 (m, 3H, $C_{anomer}OC\mathit{H}_2CH_2$), 3.48 (br, 6H, $C_qC\mathit{H}_2O$), 3.54 (t, *J* = 6.3 Hz, 6H, $OC\mathit{H}_2CH_2C{=}ONH$), 3.67-3.72 (m, 3H, $C_{anomer}OC\mathit{H}_2CH_2$), 3.87 (dt, *J* = 8.9 Hz, 11.1 Hz, 3H, *H*-2), 4.00-4.04 (m, 9H, *H*-5, *H*-6, *H*-6'), 4.48 (d, *J* = 8.5 Hz, 3H, *H*-1), 4.95-4.98 (m, 5H, *H*-3, $ArC\mathit{H}_2$), 5.21 (d, *J* = 3.4 Hz, 3H, *H*-4), 6.53 (br, 1H, $OC{=}ONHC_q$), 7.28-7.37 (m, 5H, Ar*H*), 7.72 (t, *J* = 5.6 Hz, 3H, $OCH_2CH_2C{=}ONH$), 7.81-7.84 (m, 6H, $NHC{=}OCH_3$, $CH_{2,GalNAc}C{=}ONHCH_2$)

R_f (Sulfuric acid/Anisaldehyde) = 0.1 (MeOH in DCM, 9 %)

6.4.12. Synthesis of Compound 12[298]

Compound **10** (0.63 mmol) and Pentafluorophenol (0.82 mmol) were dissolved in 5 mL abs. $CHCl_3$, EDCI (0.82 mmol) was added and the reaction was stirred for 1 h at 0 °C and overnight at RT. The reaction was diluted with DCM (20 mL) and extracted with water (3 x 20 ml). The combined aqueous layer was back extracted with DCM (2 x 20 mL) and the combined organic

layer was dried over $MgSO_4$ and reduced in vacuo to give the crude product as yellow oil (334.3 mg). The crude product was purified by chromatography (silica, Acetone in DCM, 0-20 %) to give the product as colorless oil (249.9 mg, 0.41 mmol, 65 %).



12

$C_{25}H_{28}F_5NO_{11}$
613.16

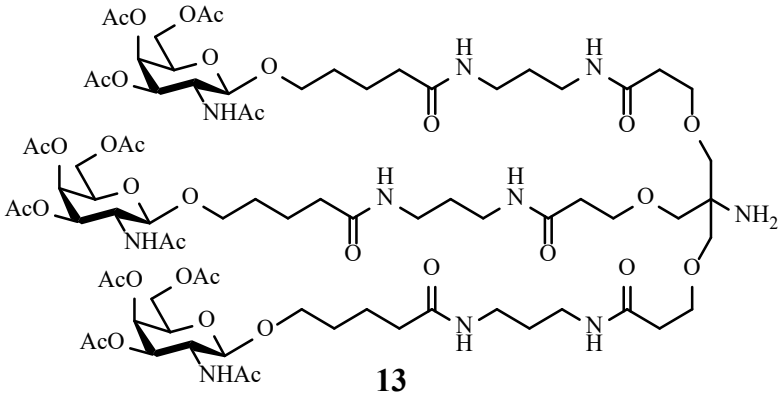
^{1}H-NMR (300 MHz, DMSO-d_6): δ (ppm) = 1.56-1.69 (m, 4H, $OCH_2CH_2CH_2$), 1.76 (s, 3H, $C{=}OCH_3$), 1.89 (s, 3H, $C{=}OCH_3$), 2.00 (s, 3H, $C{=}OCH_3$), 2.10 (s, 3H, $C{=}OCH_3$), 2.79 (t, *J* = 7.2 Hz, 2H, $CH_2COOPfp$), 3.44-3.51 (m, 1H, OCH_2CH_2), 3.73-3.80 (m, 1H, OCH_2CH_2), 3.89 (dt, *J* = 8.9 Hz, 11.2 Hz, 1H, *H*-2), 4.03 (m, 3H, *H*-5, *H*-6, *H*-6'), 4.50 (d, *J* = 8.5 Hz, 1H, *H*-1), 4.96 (dd, *J* = 3.4 Hz, 11.2 Hz, 1H, *H*-3), 5.22 (d, *J* = 3.5 Hz, 1H, *H*-4), 7.83 (d, *J* = 9.3 Hz, 1H, N*H*)

MS (ESI): *m/z* = 636.10 $[M+Na]^+$, 614.1 $[M+H]^+$, 329.9 [3,4,6-*O*-Acetyl-*N*-Acetyl-Galactosamine$^+$]

R_f (Sulfuric acid/Anisaldehyde) = 0.56 (Acetone in DCM, 16.6 %)

6.4.13. Synthesis of Compound 13[253,254]

Compound **11** (0.104 mmol) was dissolved in 2 mL abs. MeOH and Pd/C (20.0 mg, 10 wt%, dry) was added. The flask was flushed with hydrogen and the reaction mixture was hydrogenated (balloon pressure) for 24 h. The reaction mixture was filtered through celite and a 0.2 µm filter, washed with methanol and reduced in vacuo. TFA (0.16 mmol) was added and the solution was stirred for 10 min. The solvent was removed in vacuo, the residue was co-evaporated twice with methanol and dried using a high vacuum pump to yield the crude product as colorless oil (163.8 mg, 83%, TFA salt). The crude product was used without further purification.



$C_{79}H_{128}N_{10}O_{36}$
1792.85

13

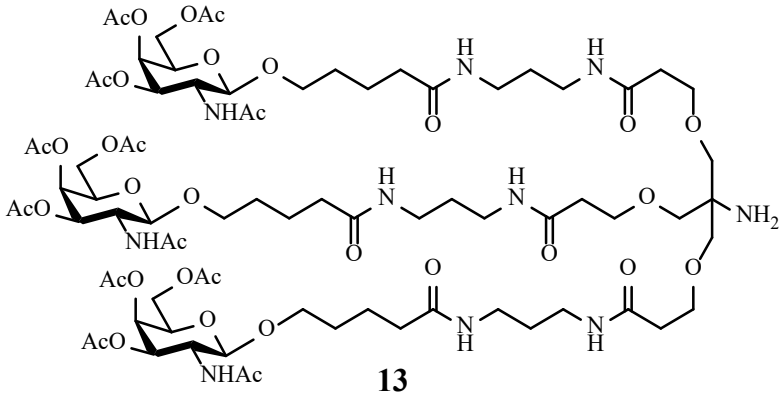
^{1}H-NMR (400 MHz, DMSO-d$_6$): δ (ppm) = 1.44-1.54 (m, 34/18H, NHCH$_2$C*H*$_2$CH$_2$NH, OCH$_2$C*H*$_2$C*H*$_2$CH$_2$C=ONH), 1.77 (s, 9H, C=OC*H*$_3$), 1.89 (s, 9H, C=OC*H*$_3$), 1.99 (s, 9H, C=OC*H*$_3$), 2.04 (m, *J* = 7.0 Hz, 11/6H, OCH$_2$CH$_2$CH$_2$C*H*$_2$C=ONH), 2.10 (s, 9H, C=OC*H*$_3$), 2.28 (t, *J* = 6.4 Hz, 12/6H, OCH$_2$C*H*$_2$C=ONH), 3.01-3.11 (m, 33/12H, C=ONHC*H*$_2$CH$_2$C*H*$_2$NHC=O), 3.33-3.42 (m, 6/3H, C$_{anomer}$OC*H*$_2$CH$_2$), 3.46 (br, 9/6H, C$_q$C*H*$_2$O), 3.63 (t, *J* = 6.2 Hz, 13/6H, OC*H*$_2$CH$_2$C=ONH), 3.67-3.73 (m, 7/3H, C$_{anomer}$OC*H*$_2$CH$_2$), 3.81-3.91 (m, *J* = 8.9 Hz, 11.1 Hz, 4/3H, *H*-2), 4.00-4.04 (m, 9H, *H*-5, *H*-6, *H*-6'), 4.49 (d, *J* = 8.5 Hz, 3H, *H*-1), 4.95-4.99 (dd, *J* = 3.4 Hz, 11.3 Hz, 3H, *H*-3), 5.21 (d, *J* = 3.4 Hz, 3H, *H*-4), 7.85 (d, *J* = 5.5 Hz, 3H, N*H*C=OCH$_3$), 7.92-7.95 (m, 6H, C=ON*H*CH$_2$CH$_2$CH$_2$N*H*C=O)

MS (ESI): *m/z* = 1795.0 $[M+2H]^{2+}$, 897.9 $[M+2H]^{2+}$, 598.9 $[M+3H]^{3+}$, 330.0 [3,4,6-*O*-Acetyl-*N*-Acetyl-Galactosamine$^+$]

The spectra are in accordance with the literature.

6.4.14. Synthesis of Compound 13[253,254]

Compound **11** (15.57 µmol) was dissolved in 3 mL abs. MeOH and Pd/C (3 mg, 10 wt%, dry) was added. HCO_2NH_4 (155.69 µmol) was added and the reaction was stirred overnight at RT. The reaction mixture was filtered through celite and a 0.2 µm sterile filter, washed with methanol and reduced in vacuo to give the crude product as colorless solid/foam (29.6 mg, 16.51 µmol, 105 %). The crude product was used without further purification.

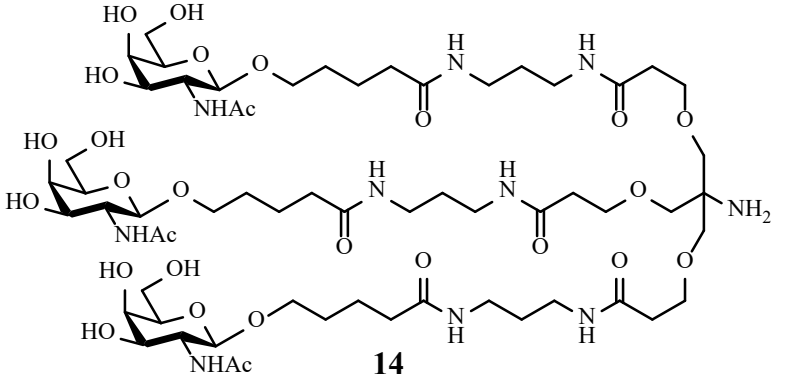


13

MS (ESI): *m/z* = 1793.9 $[M+H]^+$, 897.8 $[M+2H]^{2+}$, 598.8 $[M+3H]^{3+}$, 330.1 [3,4,6-*O*-Acetyl-*N*-Acetyl-Galactosamine$^+$]

6.4.15. Synthesis of Compound 14[253,254,279]

In a sealed glass vessel, crude compound **13** (0.013 mmol) was dissolved in 2 mL $MeNH_2$ (33wt% in EtOH) and the reaction was stirred overnight at RT. The reaction mixture was reduced in vacuo to give the crude product as slight yellow oil. The crude product was diluted in a small amount of MeOH and the product was precipitated with cold Et_2O, filtered using a frit, washed with cold Et_2O and eluted with MeOH. The procedure was repeated with the filtrate. The combined product was reduced in vacuo to give the product as slight yellow oil (16 mg, 0.011 mmol, 85 %). After reducing in vacuo, the crude product can also be purified directly by preparative HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 5-30 % + 0.1 % TFA in 40 min) to give the product as colorless TFA salt.



14

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.40-1.45 (m, 6H, $NHCH_2CH_2CH_2NH$), 1.46-1.53 (m, 12H, $OCH_2CH_2CH_2CH_2C{=}ONH$), 1.80 (s, 9H, $NHC{=}OCH_3$), 2.04 (t, *J* = 7.0 Hz, 6H, $OCH_2CH_2CH_2CH_2C{=}ONH$), 2.32 (t, *J* = 6.1 Hz, 6H, $OCH_2CH_2C{=}ONH$), 3.04 (q, *J* = 6.1 Hz, 12H, $C{=}ONHCH_2CH_2CH_2NHC{=}O$), 3.27-3.70 (m, 45H [3.28 (6H, C_qCH_2O), 3.34 (3H, *H*-5), 3.42 (3H, *H*-3), 3.51 (6H, *H*-6, *H*-6'), 3.62 (6H, $C_{anomer}OCH_2$), 3.62 (6H, $OCH_2CH_2C{=}ONH$),

3.69 (3H, *H*-2), 3.70 (3H, *H*-4)]), 4.21 (d, *J* = 8.4 Hz, 3H, *H*-1), 4.48 (br, 3H, CH_2OH), 4.57 (br, 6H, CHO*H*), 7.64 (d, *J* = 9 Hz, 3H, N*H*C=OCH_3), 7.78 (t, *J* = 5.6 Hz, 3H, $OCH_2CH_2CH_2CH_2C$=ON*H*), 7.91 (t, *J* = 5.5 Hz, 3H, OCH_2CH_2C=ON*H*)

^{13}C-NMR {^{1}H} (100 MHz, DMSO-d_6): δ (ppm) = 21.98 ($OCH_2CH_2CH_2CH_2C$=ONH), 23.05 (NHC=OCH_3), 28.63 ($NHCH_2CH_2CH_2NH$), 29.31 ($OCH_2CH_2CH_2CH_2C$=ONH), 35.09 ($OCH_2CH_2CH_2CH_2C$=ONH), 35.85 (OCH_2CH_2C=ONH), 36.26, 36.37 ($NHCH_2CH_2CH_2NH$), 52.05 (*C*H-2), 60.46 (*C*H-6/H-6'), 67.50 (OCH_2CH_2C=ONH), 67.57 (*C*H-4), 67.89 ($C_{anomer}OCH_2$), 69.78 (*C*H-3), 75.26 (C_qCH_2O), 101.40 (*C*H-1), 169.53 (NH*C*=OCH_3), 170.06 ($OCH_2CH_2CH_2CH_2C$=ONH), 170.10 (OCH_2CH_2C=ONH)

HRMS (ESI): m/z_{calc} = 719.37534 $[M+H+Na]^{2+}$

m/z_{found} = 719.37572 $[M+H+Na]^{2+}$

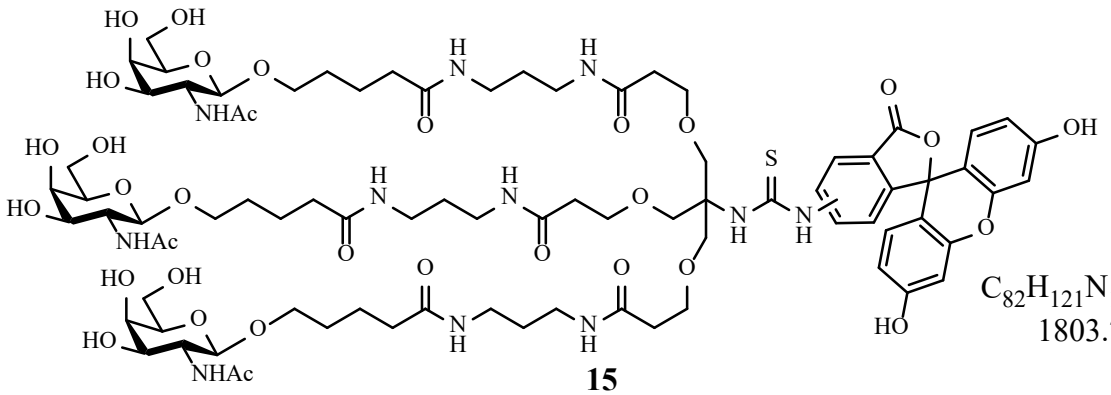
MS (ESI): *m/z* = 1416.3 $[M+H]^+$, 709.0 $[M+2H]^{2+}$, 472.9 $[M+3H]^{3+}$, 203.9 [*N*-Acetyl-Galactosamine$^+$]

t_R (HPLC) = 5.64 min (5-35 % Buffer B_{HPLC})

The spectra are in accordance with the literature.

6.4.16. Synthesis of Compound 15

Compound **14** (0.014 mmol) was dissolved in 0.5 mL abs. DMF, fluorescein 5(6)-isothiocyanate (0.014 mmol, dissolved in 0.5 mL abs. DMF) was added and the reaction was stirred overnight at RT. After 20 h, Et_3N (0.140 mmol) was added and the reaction was stirred for further 5 h at RT. The reaction was reduced in vacuo to give the crude product as yellow oil. The crude product was purified by preparative HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 15-65 % + 0.1 % TFA in 40 min) to give the product as yellow oil (6.60 mg, 0.00366 mmol, 26 %). The product was dissolved in 1 mL DMSO-d_6 and stored as DMSO stock at -20°C.



$C_{82}H_{121}N_{11}O_{32}S$
1803.79

15

^{1}H-NMR (400 MHz, DMSO-d_6): δ (ppm) = 1.35-1.52 (m, 18H, NHCH$_2$C*H*$_2$CH$_2$NH, OCH$_2$C*H*$_2$C*H*$_2$CH$_2$C=ONH), 1.78 (s, 6/9H, NHC=OC*H*$_3$), 2.03 (t, *J* = 7.2 Hz, 6H, OCH$_2$CH$_2$CH$_2$C*H*$_2$C=ONH), 2.32 (t, *J* = 6.1 Hz, 6H, OCH$_2$C*H*$_2$C=ONH), 3.04 (m, 12H, C=ONHC*H*$_2$CH$_2$C*H*$_2$NHC=O), 3.27-3.76 (m, 35/36H [3.28 (6H, C$_q$C*H*$_2$O), 3.34 (3H, *H*-5), 3.42 (3H, *H*-3), 3.51 (6H, *H*-6, *H*-6'), 3.62 (6H, C$_{anomer}$OC*H*$_2$), 3.62 (6H, OC*H*$_2$CH$_2$C=ONH), 3.69 (3H, *H*-2), 3.70 (3H, *H*-4)]), 4.21 (d, *J* = 8.4 Hz, 7H, *H*-1, overlaying with 4.10 (200H, probably water, 4.48 (br, 3H, CH$_2$O*H*), 4.57 (br, 6H, CHO*H*) and Ar-O*H*), 6.53-6.70 (m, 6/9H, Ar-*H*) 7.64 (d, *J* = 9 Hz, 2/3H, N*H*C=OCH$_3$), 7.72-7.75 (m, 3H, OCH$_2$CH$_2$CH$_2$CH$_2$C=ON*H*), 7.86-7.89 (m, 3H, OCH$_2$CH$_2$C=ON*H*), 8.30 (br s, 1H, NHC=SN*H*), 10.09 (br s, 1H, N*H*C=SNH)

HRMS (ESI): m/z_{calc} = 921.87574 $[M+Ca]^{2+}$

m/z_{found} = 921.87573 $[M+Ca]^{2+}$

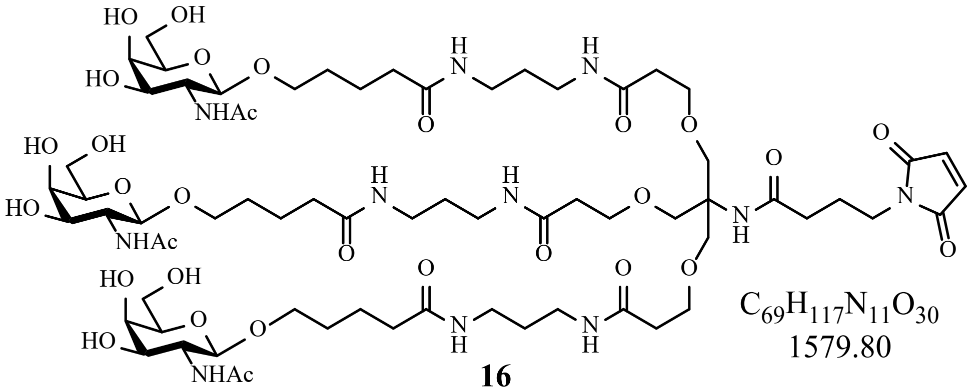
MS (ESI): *m/z* = 903.3 $[M+2H]^{2+}$, 602.5 $[M+3H]^{3+}$, 204.0 [N-Acetyl-Galactosamine$^+$]

UV: λ_{max} = 494 nm

$\varepsilon_{494\ nm}$ (FITC) = 89215.18 $M^{-1}\cdot cm^{-1}$ [assumption: $\varepsilon_{494\ nm}$ (FITC) = $\varepsilon_{494\ nm}$ (GalNAc-FITC)]

6.4.17. Synthesis of Compound 16

Compound **14** (8.27 µmol) was dissolved in 0.5 mL abs. DMF, *N*-Succinimidyl 4-Maleimidobutyrate (62.45 µmol, dissolved in 0.5 mL abs. DMF) and Et$_3$N (248.1 µmol) were added and the reaction was stirred for 48 h at RT. Additional Et$_3$N (144.2 µmol) was added and the reaction was stirred for further 24 h at RT. The reaction was reduced in vacuo and lyophilized from ACN in H$_2$O (50 %), to give the crude product as slight orange powder. The crude product was purified preparative HPLC (C18, buffer B$_{HPLC}$ in buffer A$_{HPLC}$, 5-40 % + 0.1 % TFA in 40 min). Product containing fractions were extracted with EtOAc (3 x 3 mL) and the aqueous phase was lyophilized to give the product as colorless powder (3.3 mg, 2.09 µmol, 25 %).



16

$C_{69}H_{117}N_{11}O_{30}$
1579.80

HRMS (ESI): m/z_{calc} = 809.87914 $[M+Ca]^{2+}$

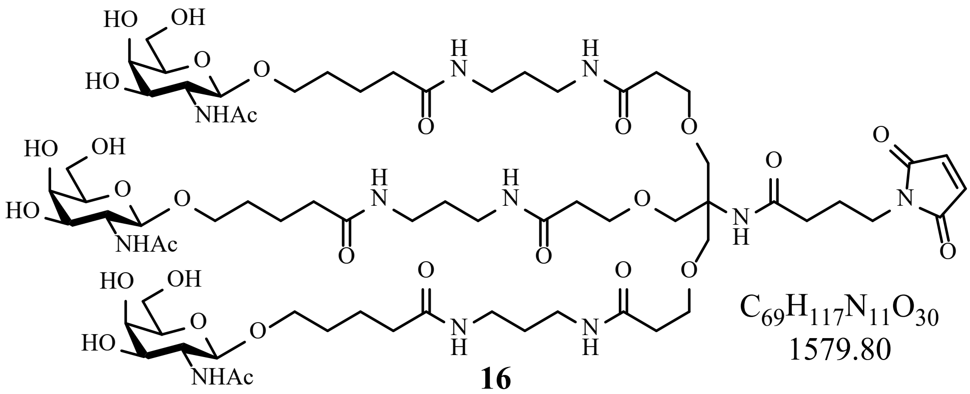
m/z_{found} = 809.87909 $[M+Ca]^{2+}$

MS (ESI): m/z = 1581.6 $[M+H]^{+}$, 791.5 $[M+2H]^{2+}$, 528.0 $[M+3H]^{3+}$, 204.1 [*N*-Acetyl-Galactosamine$^{\cdot}$]

t_R (HPLC) = 9.45 min (5 – 35 % Buffer B_{HPLC})

6.4.18. Synthesis of Compound 16[279]

N-Succinimidyl 4-Maleimidobutyrat (10.06 µmol) was dissolved in 300 µL abs. DMF and the solution was added to compound **14** (8.56 µmol). Et_3N (50.3 µmol) was added and the reaction was shaken for 19 h at 35 °C and 300 rpm. The reaction was lyophilized from DMF and the crude product was purified by preparative HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 5-30 % + 0.1 % TFA in 40 min) to give the product as a white solid (3.6 mg, 2.28 µmol, 27 %).



16

$C_{69}H_{117}N_{11}O_{30}$
1579.80

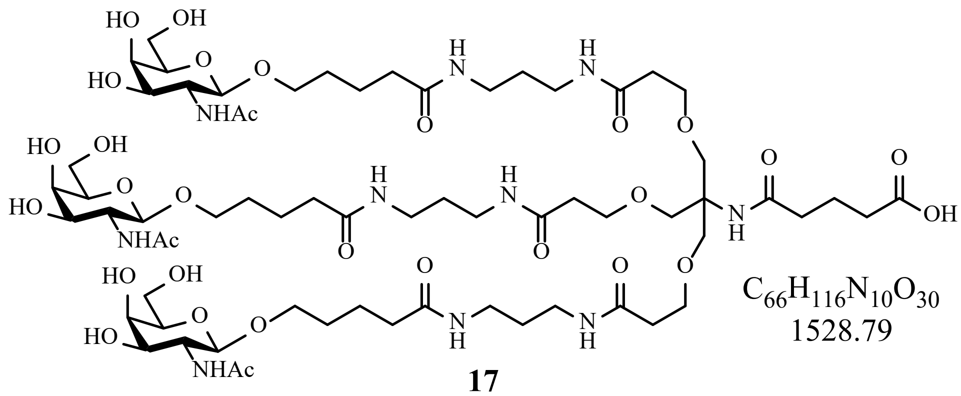
HRMS (ESI): m/z_{calc} = 812.88761 $[M+2Na]^{2+}$

m/z_{found} = 812.88763 $[M+2Na]^{2+}$

t_R (HPLC) = 9.45 min (5-35 % Buffer B_{HPLC})

6.4.19. Synthesis of compound 17[250,299]

Compound **14** (3.32 µmol) was dissolved in 0.5 mL abs. DMF, DIPEA (6.64 µmol) and DMAP (1.66 µmol) were added. Glutaric anhydride (4.98 µmol, dissolved in 50 µL abs. DMF) was added in 10 µL potions over 1 h. The reaction was stirred for 5 h at 50 °C and overnight at RT. The reaction was lyophilized from ACN in H_2O (50 %) and purified by preparative HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 5-35 % + 0.1 % TFA in 40 min) to give the product as colorless solid (2.4 mg, 1.56 µmol, 20 %).



17

HRMS (ESI): m/z_{calc} = 787.38216 $[M+2Na]^{2+}$

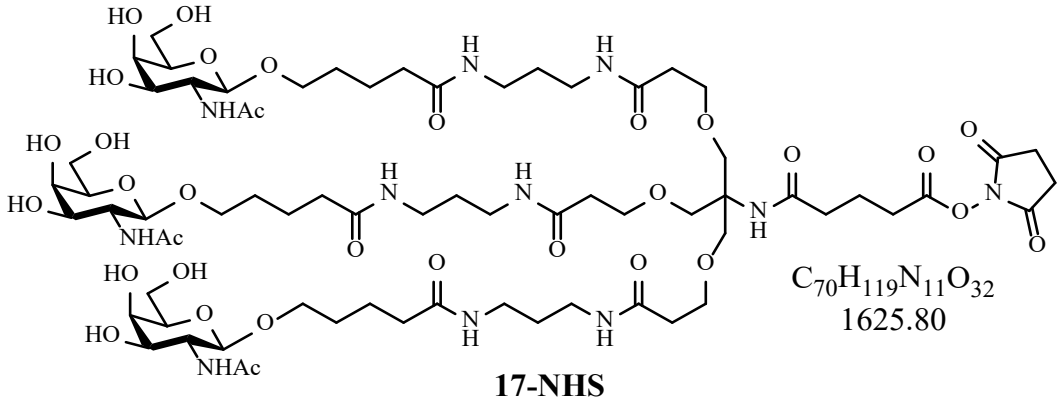
m/z_{found} = 787.38184 $[M+2Na]^{2+}$

MS (ESI): m/z = 765.8 $[M+2H]^{2+}$, 510.7 $[M+3H]^{3+}$, 204.2 [*N*-Acetyl Galactosamine$^{\cdot}$]

t_R (HPLC) = 8.17 min (5-35 % Buffer B_{HPLC})

6.4.20. Synthesis of compound 17-NHS

Compound **17** (240 nmol / 4 µL, 60 mM in DMSO), DIC (2160 nmol / 8 µL, 270 mM in DMSO), NHS (3680 nmol / 8 µL, 460 mM in DMSO) and DIPEA (2240 nmol / 8 µL, 280 mM in DMSO / sonicated for 15 min before use) were added to a reaction cup and the mixture was shaken overnight at 45 °C and 900 rpm. The reaction progress was monitored by analytical HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 5-35 % + 0.1 % FA in 19 min). The reaction was lyophilized from DMSO and purified by preparative HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 5-35 % + 0.1 % FA in 40 min) to give the product as colorless solid. No yield was determined.



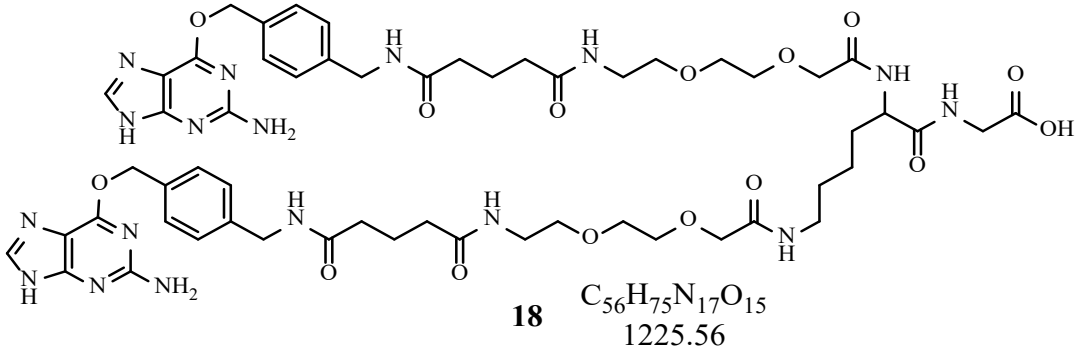
17-NHS

MS (ESI): *m/z* = 814.3 $[M+2H]^{2+}$, 543,3 $[M+3H]^{3+}$, 204.1 [*N*-Acetyl Galactosamine$^-$]

t_R (HPLC) = 9.63 min (5-35 % Buffer B_{HPLC})

6.4.21. Synthesis of Compound 18[119]

The preparation of compound **18** was conducted according to the reported solid phase synthesis by Stroppel *et al.*.



18

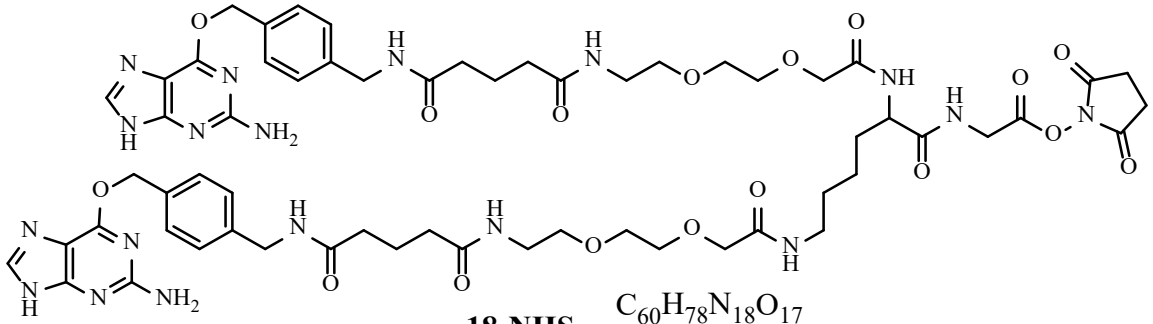
HRMS (ESI): m/z_{calc} = 613.78870 $[M+2H]^{2+}$

m/z_{found} = 613.78891 $[M+2H]^{2+}$

t_R (HPLC) = 6.91 min (5-95 % Buffer B_{HPLC})

6.4.22. Synthesis of Compound 18-NHS

BisBG-COOH (**18**) (240 nmol / 4 µL, 60 mM in DMSO), DIC (1080 nmol / 4 µL, 270 mM in DMSO), NHS (1840 nmol / 4 µL, 460 mM in DMSO) and DIPEA (1120 nmol / 4 µL, 280 mM in DMSO / sonicated for 15 min before use) were added to a reaction tube and the mixture was shaken for 4 h at 45 °C and 900 rpm. The reaction progress was monitored by analytical HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 5-95 % + 0.1 % FA in 24 min). With a progress of > 50 %, the reaction was lyophilized from DMSO and used without further purification or analysis for gRNA conjugations.

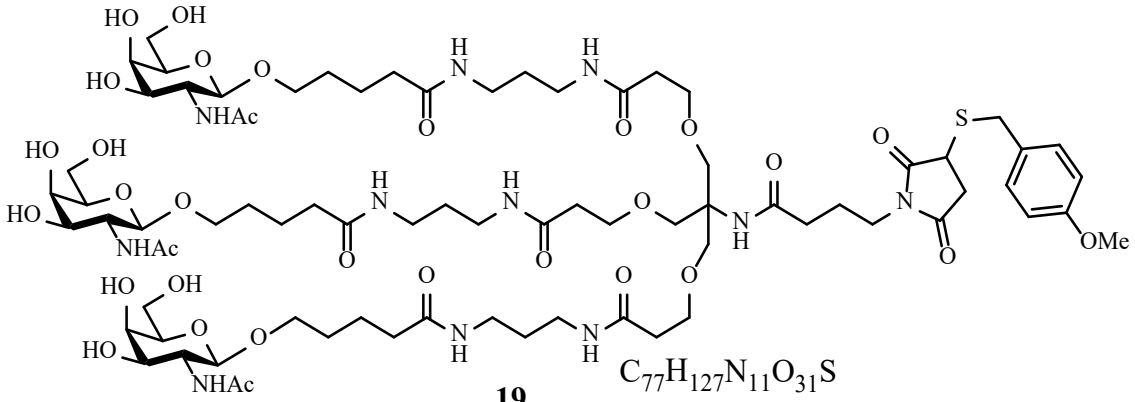


18-NHS $C_{60}H_{78}N_{18}O_{17}$ 1322.58

t_R (HPLC) = 7.32 min (5-95 % Buffer B_{HPLC})

6.4.23. Synthesis of Compound 19[300,301]

Compound **16** (0.16 µmol) was dissolved in 413 µL PBS (pH 7.4), 4-Methoxybenzyl mercaptan (0.16 µmol) was added and the reaction was shaken for 2 h at RT and 450 rpm. The complete reaction mixture was lyophilized from ACN in H_2O (50 %) to give the crude product as white powder. The crude product was purified by preparative HPLC (C18, buffer B_{HPLC} in buffer A_{HPLC}, 10-65 % + 0.1 % TFA in 40 min) to give the product as white powder. No Yield was determined.



19 $C_{77}H_{127}N_{11}O_{31}S$ 1733.84

MS (ESI): *m/z* = 1757.8 $[M+Na]^+$, 868.3 $[M+2H]^{2+}$, 579.3 $[M+3H]^{3+}$, 204.3 [*N*-Acetyl-Galactosamine$^+$]

6.5. Molecular biology methods

6.5.1. General procedures

6.5.2. DNAse digest with DNAse I

DNAse digests were performed according to the following approach (Table 17).

Table 17: General approach for DNAse digests using DNAse I

Component	Amount
Isolated RNA	1 µg
10x DNAse Buffer	2.5 µL
DNAse I (2 U/µL)	1 µL
V_{total} (RNAse free H_2O)	25 µL
⇨ **Incubation for 30 min at 37 °C**	
EDTA (25 mM)	2.5 µL
⇨ **Incubation for 10 min at 75 °C**	
⇨ **Store at 0 °C / -20 °C**	

6.5.3. Amplification of template DNA using Phusion DNA polymerase

The amplification of cDNA (reverse transcribed RNA) or recombinant DNA, using Phusion DNA polymerase, was performed according to the following procedure (Table 18). Amplification was performed according to the thermocycler program described below (Table 19). PCR amplicon was purified using gel electrophoresis (see section 6.5.5) or stored at -20 °C.

Table 18: General procedure for template DNA amplifications using Ph-DNA polymerase.

Component	Amount
DNA template (cDNA / recombinant DNA)	5 µL / 0.5 L
5x Ph-DNA Buffer	10 µL
dNTPs	1.25 µL
Forward primer (10 µM)	2.5 µL
Backward primer (10µM)	2.5 µL
Ph-DNA polymerase	1 µL
Nanopure water	27.75 µL / 32.25 µL

Table 19: Thermocycler program for PCRs using Phuison DNA polymerase. *Elongation time was adjusted to the template length and annealing temperature was adjusted to the used primer set.*

Cycle Step	Temperature [° C]	Time	Cycles
Initial Denaturation	98	30 sec	1
Denaturation	98	10 sec	36
Annealing	55-65*	30 sec	36
Elongation	72	30 sec / kb	36
Final extension	72	10 min	1
Storage	8	∞	1

*Annealing temperature is highly dependent on the melting temperature (T_M) of the used primer set

6.5.4. Amplification of template DNA using *Taq* DNA polymerase

The amplification of template DNA, using *Taq* DNA polymerase, was performed according to the following procedure (Table 20). Amplification was performed according to the thermocycler program described below (Table 21). PCR amplicon was purified using gel electrophoresis (see section 6.5.5) or stored at -20 °C.

Table 20: General procedure for template DNA amplifications using Taq DNA polymerase.

Component	Amount
DNA template	1 µL
10x ThermoPol Buffer	5 µL
dNTPs	1.25 µL
Forward primer (10 µM)	2.5 µL
Backward primer (10 µM)	2.5 µL
Taq DNA polymerase	0.5 µL
Nanopure water	32.25 µL

Table 21: Thermocycler program for PCRs using Taq DNA polymerase. *Elongation time was adjusted to the template length and annealing temperature was adjusted to the used primer set.*

Cycle Step	Temperature [° C]	Time	Cycles
Initial Denaturation	95	5 min	1
Denaturation	95	15 sec	32
Annealing	55-65*	30 sec	32
Elongation	68	60 sec / kb	32
Final extension	68	10 min	1
Storage	8	∞	1

*Annealing temperature is highly dependent on the melting temperature (T_M) of the used primer set. For the nested PCR amplification of GAPDH using P927/P1159 an annealing temperature of 56 °C was used, and for the nested PCR amplification of STAT1 using P3533/P3534 an annealing temperature of 57 °C was used.

6.5.5. Agarose gel electrophoresis and gel extraction

PCR amplicons were separated using agarose gel electrophoresis. Dependent on the amplicon size, different buffers and concentrations were applied for the separation. For an overview, see Table 22. For loading, each sample was diluted with 6x Gel loading Dye and 1 kb Plus DNA ladder was used as reference. DNA was visualized using ROTI®GelStain in-gel staining (5µL for 100 mL gel) and UV illumination at a wavelength λ = 365 nm. Separated PCR amplicons were excised and extracted using NucleoSpin® Gel and PCR Clean-Up kit (MACHERY-NAGEL GmbH & Co. KG, Düren, Germany) according to the manufacturer's protocol. Drying time was prolonged to 5 min. Amplicons were eluted in 25 µL NE elution buffer and stored at -20 °C.

Table 22: Overview over the conditions used for agarose gel electrophoresis. *For PCR and restriction digest cleanup, gels were casted in Mini-Sub Cell GT UV-Transparent Gel Trays (7x7 cm with 8 wells or 7x10 cm with 2x8 wells). For colony-PCR, gels were casted in Mini-Sub Cell GT UV-Transparent Gel Trays (7x7 cm with 15 wells) (Bio-Rad Laboratories GmbH, Feldkirchen, Germany).*

Amplicon size	Conditions
< 500 bp	1.4 % (w/v) SB, 200 V, 15 min
500-2000 bp	1.4 % (w/v) TAE, 120 V, 30 min
> 2000 bp	1.0 % (w/v) TAE, 120 V, 30 min

6.5.6. Restriction digest and ligation for subcloning

PCR amplicons or plasmids were double digested according to the manufacturer's protocol (New England Biolabs GmbH, Frankfurt am Main, Germany) using the desired restriction enzymes. All restriction digests were performed in CutSmart buffer. Only CutSmart buffer

compatible restriction enzymes were used and a final restriction enzyme activity of 10 U/µg was applied. An exemplary restriction digest is shown in Table 23. Restriction digests were incubated for 1 h at 37 °C and purified using agarose gel electrophoresis and gel extraction (see section 6.5.5). If no separation of fragments is necessary (digested overhangs < 70 bp), restriction digests were purified directly using NucleoSpin® Gel and PCR Clean-Up kit (MACHERY-NAGEL GmbH & Co. KG, Düren, Germany).

Table 23: Exemplary restriction digest.

Component	Amount
PCR amplicon or Plasmid	1 µg
10x CutSmart Buffer	5 µL
Restriction enzyme I (20.000 U/mL)	1 µL
Restriction enzyme II (10.000 U/mL)	2 µL
V_{total} (Nanopure water)	50 µL

For insert and vector ligation, a molar ratio of insert and vector of 3:1 was applied. 200 ng vector and the corresponding amount of insert were added to 2 µL T4 DNA ligase buffer and 2 µL T4 DNA ligase. The final volume adjusted to 20 µL using Nanopure water and the ligation was incubated for 4 h at RT. The ligation was stored at -20 °C or used directly for heat-shock transformations into $CaCl_2$ competent *E.Coli* XL1-blue (see section 6.5.7).

6.5.7. Heat-shock transformation into $CaCl_2$ competent *E.Coli* XL1-blue, overnight culture and plasmid isolation

For heat-shock transformations of $CaCl_2$ competent *E.Coli* XL1-blue, 100 µL of cells were thawed on ice. 100 µL TE buffer and 0.2 µL isolated plasmid or 5 µL ligation mixture were added and the cells were incubated for 30 min on ice, 2 min at 42 °C and 5 min on RT, subsequently. 800 µL antibiotic free LB medium was added and the transformed cells were incubated for 30 min at 37 °C. The cells were centrifuged for 2 min at 5000 rpm and 900 µL of the supernatant was removed. The cells were resuspended in the residual supernatant, plated on selective LB-Agar plates and incubated overnight at 37 °C. For overnight cultures, 6 mL of selective LB medium was inoculated with a single colony and incubated overnight at 37 °C and 180 rpm. Overnight cultures were analyzed by Colony-PCR and the plasmids of positive cultures were isolated using NucleoSpin® Plasmid kit (MACHERY-NAGEL GmbH & Co. KG,

Düren, Germany) according to the manufacturer's protocol. Plasmids were eluted in 50 µL AE elution buffer, analyzed using Sanger sequencing and stored at -20 °C.

6.5.8. Colony-PCR using *Taq* DNA polymerase

Overnight cultures were screened for plasmid or ligated vector containing cultures using colony-PCR. Colony-PCR was performed according to the following procedure (Table 24). Amplification was performed using the thermocycler program described before for amplifications of template DNA using *Taq* DNA polymerase. Colony-PCR amplicon was analyzed using gel electrophoresis (see section 6.5.5). Plasmids of positive cultures (plasmid or ligated vector containing) were isolated and analyzed as described before (see section 6.5.7).

Table 24: General procedure for colony-PCRs using Taq DNA polymerase.

Component	Amount
Overnight culture	1 µL
10x ThermoPol Buffer	5 µL
dNTPs	0.63 µL
Forward primer (10 µM)	1.25 µL
Backward primer (10 µM)	1.25 µL
Taq DNA polymerase	0.25 µL
Nanopure water	15.62 µL

6.5.9. RNA isolation using Monarch® RNA Cleanup Kit (10 µg)

RNA isolations of cell culture experiments were performed using Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany). For cell culture experiments using a 96-well scale, the cells of each well were lysed using 50 µL RLT Lysis Buffer (QIAGEN GmbH – Germany, Düsseldorf, Germany). 100 µL RNA Cleanup Binding Buffer was added to each well, both similar wells were pooled and transferred in a new tube, containing 300 µL ethanol (99.8 %). For cell culture experiments using a 24-well scale, the cells of each well were lysed using 100 µL RLT Lysis Buffer. 200 µL RNA Cleanup Binding Buffer was added to each well and the sample was transferred in a new tube, containing 300 µL ethanol (99.8 %). The samples were mixed by pipetting, transferred into the intended spin columns and processed according to the manufacturer's protocol. The sample was washed twice using 500 µL RNA Cleanup Wash Buffer, dried for 5 min at 13.000 rpm, eluted in 15 µL nuclease free water and stored at -80 °C.

6.5.10. DNAse digest and RT-PCR using One Step RT-PCR Kit

Table 25: General procedure for combined reverse transcription of isolated RNA and amplification of cDNA targeting GAPDH.

Component	Amount
DNAse digest (250 ng)	6.25 µL
P928 (GAPDH_bw) (10 µM)	1 µL
TMR53 (GAPDH_ORF1_sense_OMe) (10 µM)	1 µL
Nuclease free water	2 µL
⇨ **Incubation for 2 min at 95 °C**	
One Step Mix (2x)	12.5 µL
RT-RI Blend (20x)	1.25 µL
P927 (GAPDH_fw) (10µM)	1 µL

Table 26: General procedure for combined reverse transcription of isolated RNA and amplification of cDNA targeting STAT1. *For sample preparation, 500 ng isolated RNA was diluted to a volume of 8.25 µL using nuclease free water.*

Component	Amount
Isolated RNA (500 ng)	8.25 µL
P3533 (STAT1_bw) (10 µM)	1 µL
P3535 (STAT1_sense_oligo_NH180) (10 µM)	1 µL
⇨ **Incubation for 2 min at 95 °C**	
One Step Mix (2x)	12.5 µL
RT-RI Blend (20x)	1.25 µL
P3532 (STAT1_fw) (10µM)	1 µL

Preliminary to the reverse transcription of the isolated RNA and amplification of synthesized cDNA targeting GAPDH, residual DNA was digested according to the general procedure (see Table 25). For samples targeting STAT1, no DNAse digest was performed. For both, cDNA was synthesized and amplified using One Step RT-PCR Kit (biotechrabbit GmbH, Henningsdorf, Germany) according to the manufacturer's protocol. Samples were prepared according to the following general procedure (For GAPDH see Table 25, for STAT1 see Table 26). Combined reverse transcription and amplification was performed according to the thermocycler program described below (Table 27). Amplicons were separated and purified using gel electrophoresis and gel extraction as described before (see section 6.5.5). Sequencing was performed as described below (see section 6.5.12).

Table 27: Thermocycler program of the reverse transcription and amplification of cDNA. *Elongation time was adjusted to the template length and annealing temperature was adjusted to the used primer set.*

Cycle Step	Temperature [° C]	Time	Cycles
Reverse transcription	50	30 min	1
Enzyme inactivation	95	10 min	1
Initial Denaturation	95	15 sec	1
Denaturation	95	15 sec	35
Annealing	58*	30 sec	35
Elongation	72	60 sec / kb	35
Final extension	72	10 min	1
Storage	4	∞	1

*Annealing temperature is highly dependent on the melting temperature (T_M) of the used primer set. The same annealing temperature of 58 °C was used for the amplification of GAPDH using P927/P928, as well as STAT1 using P3532/P3233.

6.5.11. Nested-PCR

If an additional purification of particular samples was required, samples were further amplified using *Taq* DNA polymerase (Nested-PCR). Samples were prepared according to the general procedure for the amplification using *Taq* DNA polymerase as describes before (see section 6.5.4). As DNA template, 1 µL purified RT-PCR was used and either the forward or reverse primer was replaced by the sequencing primer, dependent on the sequencing primer's orientation. Nested-PCR amplicons were separated and purified using gel electrophoresis and gel extraction as described before (see section 6.5.5). Sequencing was performed as described below and the same sequencing primer was used for sequencing (see section 6.5.12).

6.5.12. Sequencing of plasmid DNA and PCR amplicons

Isolated plasmids and purified PCR amplicons were analyzed using Sanger sequencing by Eurofins Genomics Germany GmbH (Ebersberg, Germany) or Microsynth AG (Balgach, Switzerland). Primers used for sequencing are listed in Table 28.

For sequencing by Eurofins, 1500 ng plasmid DNA or 120 ng PCR amplicon was mixed with 1.5 µL Primer (10 µM) and the total volume was adjusted to 15 µL using Nanopure water.

For sequencing by Microsynth, samples were prepared according to the manufacturer's requirements. 18 ng per 100 bp of purified PCR amplicon was adjusted to a volume of 9 µL using Nanopure water and 6 µL primer (10 µM) was added.

Table 28: Used sequencing primers of PCR amplicons for evaluation of editing yields. *Plasmids were sequenced using several different primers to ensure sequencing of the complete insert and restrictions sites. The detailed primer sets, used for plasmid sequencing, are listed in Table 15.*

Primer No.	Primer description	Target
P1159	GAPDH_ORF_seq_bw	GAPDH ORF1
P3534	STAT1_sequencing	STAT1Y701C

6.6. gRNA synthesis and purification

The conjugation of BisBG-COOH (**18**) to NH_2-functionalized oligonucleotides was performed as previously described by Stroppel *et al.*[119].

6.6.1. Conjugation of BisBG-COONHS to NH_2-functionalized oligonucleotides

Lyophilized and activated BisBG-COONHS (**18-NHS**) was dissolved in DMSO/DIPEA (12 µL, 60:1, sonicated for 15 min before use), NH_2 functionalized oligonucleotides (8.33 µL, 6 µg/µL) were added and the reaction was shaken for 2 h at 37 °C and 900 rpm. The conjugated oligonucleotide was separated from unconjugated oligonucleotide and excess BisBG-COOH (**18**) using denaturing TBE-7 M Urea-PAGE (20 %) and visualized using fluorescence as described in section 6.3.12.

6.6.2. Purification of conjugated oligonucleotides using preparative Urea-PAGE

Conjugated oligonucleotides were purified using denaturing TBE-7 M Urea-PAGE (20 %, 37 x 31 cm x 0.8 mm), followed by NaOAc precipitation or Sep-Pak® Plus C18 cartridges (Waters Corporation, Milford (MA), USA). Polyacrylamide gels were cast according to the compositions described below (see Table 29). For sample preparation, 0.1 volumes of RNA loading dye (90 % ROTIPHORESE® Sequencing gel diluent, 10 % ROTIPHORESE® Sequencing gel buffer concentrate, Bromphenol blue, Xylene cyanol) was added to each sample. As reference, unconjugated NH_2 functionalized oligonucleotide (2 µL, 6 µg/µL) was diluted with 4 µL DMSO and 0.1 volumes of gRNA loading dye. All samples were incubated for 3 min at 70 °C and chilled on ice before loading. Electrophoresis was performed for 5-6 h maintaining the electrical power between 45-50 W, the voltage between 1200-1500 V and the electric current between 60-90 mA. The separated oligonucleotides were visualized using TLC Silica Gel 60 F_{254} aluminium sheets (Merck KGaA, Darmstadt, Germany) and UV light illumination at 254 nm. Conjugated oligonucleotides were excised and extracted using the crush-soak method into 600 µL nuclease free water overnight at 4 °C and 1100 rpm. To remove urea and buffer ingredients, the extracted oligonucleotides were further purified by NaOAc

precipitation or C18 column purification using Sep-Pak® Plus C18 cartridges. For NaOAc precipitation, the extracted oligonucleotides were split in 2 x 300 µL and precipitated with 0.1 volumes 3 M $NaOAc_{aq}$ and 3 volumes EtOH (99.8 %) for at least 24 h at -20 °C. The precipitated oligonucleotides were centrifuged for 90 min at -4 °C and 14.000 rpm, the supernatant was removed and the pellet was washed with $EtOH_{aq}$ (70 % v/v, precooled to -20 °C). Centrifugation was repeated, the supernatant was removed, the pellet was dried for 3 min at 70 °C and the oligonucleotide was dissolved in 15 µL nuclease free water. The concentration was determined as described before (see section 6.3.11). For C18 column purification, Sep-Pak® Plus C18 cartridges was used according to the manufacturer's protocol. Oligonucleotide containing fractions were combined, lyophilized from elution buffer, dissolved in 15 µL nuclease free water and the concentration was determined as described bevor (see section 6.3.11).

HRMS (-ESI): m/z_{calc} (BisBG-471) = 9181.3 $[M-H]^-$, 9203.3 $[M+Na-2H]^-$, 9219.2 $[M+K-2H]^-$, 9261.2 $[M+Br]^-$, 4590.1 $[M-2H]^{2-}$

m/z_{found} (BisBG-471) = 9187.5, 9258.6, 4593.8

***Table 29: Polyacrylamide gels (15 % and 20 %) used for oligonucleotide purification**. Polymerization was carried out overnight at RT.*

Component	15 %	20%
Nanopure water	20.28 mL	13 mL
ROTIPHORESE® Sequencing gel buffer concentrate	13 mL	13 mL
ROTIPHORESE® Sequencing gel diluent	18.72 mL	-
ROTIPHORESE® Sequencing gel concentrate	78 mL	104 mL
APS_{aq} (10 % w/v)	650 µL	650 µL
TEMED	65 µL	65µL

6.6.3. Quantification of DTT using Ellman's reagent

The quantification of DTT was performed according to the manufacturer's protocol using Ellman's reagent (Thermo Fisher Scientific Inc., Waltham (MA), USA). In a 96-well plate, 25 µL of the desired sample was diluted with phosphate buffer (50 µL, 0.1 M, pH 7.0) and Ellman's reagent (25 µL, 25 mM in phosphate buffer (0.1 M, pH 7.0)) and incubated at RT for 5 min. The absorbance was determined using a Spark 10M Luminescence Multi Mode

Microplate Reader (Tecan Group AG, Männedorf, Switzerland) at a wavelength λ of 412 nm. The concentration of DTT was calculated using the law of Lambert-Beer and the molecular attenuation coefficient ε_{412} = 14,150 $M^{-1} \bullet cm^{-1}$ of 2-nitro-5-thiobenzoic acid. As blank value, the absorbance of Ellman's reagent (25 µL, 25 mM in phosphate buffer (0.1 M, pH 7.0)), diluted in 75 µL phosphate buffer (0.1 M, pH 7.0), was determined and subtracted from all measurements.

6.6.4. Deprotection of Thiol functionalized gRNAs 257 and 258

For the deprotection of thiol functionalized oligonucleotides 257 and 258, 35 pmol of oligonucleotide was diluted in phosphate buffer (0.1 M, pH 8.4) to a final volume of 25 µL. DTT (1000 nmol, 10 µL, 100 mM in phosphate buffer (0.1 M, pH 8.4)) was added and the reaction was shaken at RT for 1 h and 300 rpm. To remove the excess of DTT, the oligonucleotide was precipitated with 0.1 volumes 3 M $NaOAc_{aq}$ and 3 volumes EtOH (99.8 %) for at least 24 h at -20 °C. The precipitated oligonucleotides were centrifuged for 90 min at -4 °C and 14.000 rpm, the supernatant was removed and the pellet was washed 3 times with $EtOH_{aq}$ (70 % v/v, precooled to -20 °C). Centrifugation was repeated during each washing step. After washing, the supernatant was removed, the pellet was dried for 3 min at 70 °C and the oligonucleotide was dissolved in 11 µL nuclease free water. The concentration was determined as described before (see section 6.3.11) and the deprotected oligonucleotide was used directly for the subsequent conjugation to GalNAc-Maleimide (**16**). The amount of DTT was quantified for each obtained supernatant as described within section 6.6.3.

6.6.5. Conjugation of GalNAc-Maleimide to thiol functionalized gRNAs 257 and 258

For the conjugation of GalNAc-Maleimide **(16)** to the deprotected and thiol functionalized oligonucleotides, GalNAc-Maleimide (350 pmol, 290 µM in DMSO) was added to each gRNA and the reactions were shaken overnight at 4 °C and 300 rpm. The reactions were separated using denaturing TBE-7 M Urea-PAGE (15 %, 37 x 31 cm x 0.8 mm) and visualized using fluorescence imaging (see section in section 6.3.12).

6.6.6. Deprotection of Thiol functionalized gRNAs 471-473

For the deprotection of thiol functionalized oligonucleotides 471-473, 2 nmol of oligonucleotide was lyophilized from water and dissolved in freshly degassed phosphate buffer (27.5 µL, 0.1 M, pH 8.4). Only resuspension is recommended, do not vortex. DTT (250 nmol, 2.5 µL, 100 mM in degassed phosphate buffer (0.1 M, pH 8.4)) was added and the reaction was shaken at RT for 1 h and 300 rpm. The excess DTT was removed using Zeba™ Spin Desalting Columns (40 kDa MWCO, 0.5 mL, Thermo Fisher Scientific Inc., Waltham (MA), USA)

according to the manufacturer's protocol for buffer exchange. Zeba™ Spin Desalting Columns were equilibrated with freshly degassed phosphate buffer (4 x 300 µL, 0.1 M, pH 7.0), and after elution, the column was additionally washed with freshly degassed phosphate buffer (70 µL, 0.1 M, pH 7.0). The oligonucleotide concentration of both, elution and wash fractions were determined as described before (see section 6.3.11). The deprotected oligonucleotide was used directly for the subsequent conjugation to GalNAc-Maleimide (**16**).

6.6.7. Conjugation of GalNAc-Maleimide to thiol functionalized gRNAs 471-473

For the conjugation of GalNAc-Maleimide (**16**) to the deprotected, thiol functionalized oligonucleotides, GalNAc-Maleimide (200 nmol, 60 mM in DMSO) was added to each fraction (elution and wash) and the reactions were shaken for 3 h at RT and 300 rpm (If the wash fraction is only containing a minor amount of oligonucleotide, the amount of GalNAc-Maleimide can be reduced, or the wash fraction can be disposed). The reactions of elution and wash fractions were combined, lyophilized from water/DMSO and purified using denaturing TBE-7 M Urea-PAGE (20 %, 37 x 31 cm x 0.8 mm), followed by Sep-Pak® Plus C18 cartridges (Waters Corporation, Milford (MA), USA) as described before (see section 6.6.2).

HRMS (ESI): m/z_{calc} (BisBG-471-GalNAc) = 10626.8 $[M-H]^-$, 10648.8 $[M+Na-2H]^-$, 5312.9 $[M-2H]^{2-}$

m/z_{found} (BisBG-471-GalNAc) = 10635.2, 10617.9, 5326.0

6.6.8. Conjugation of GalNAc-COONHS to NH_2-functionalized oligonucleotides

The respective oligonucleotide (30 pmol) was lyophilized from water and dissolved in 15 µL gRNA labeling buffer (pH 8.3; 130.5 mM NaCl, 2.6 mM KCl, 8.8 mM NaH_2PO_4, 1.7 mM KH_2PO_4, 9.5 mM $NaHCO_3$; PBS (pH 7.4) : 0.2 M $NaHCO_3$ (pH 9.0) 20:1). GalNAc-COONHS (**17-NHS**) (60 mM in DMSO, assumed) was added and the reaction was shaken overnight at RT and 600 rpm. The conjugated oligonucleotide was separated using denaturing TBE-7 M Urea-PAGE (15 %) and visualized using fluorescence imaging as described in section 6.3.12.

6.6.9. Analytical Urea-PAGE of conjugated oligonucleotides

The modification of GalNAc derivatives to oligonucleotides was also evaluated using denaturing TBE-7 M Urea-PAGE (15 % or 20 %, 37 x 31 cm x 0.8 mm) in an analytical scale. Urea-PAGE's were cast as described before (see section 6.6.2) and 15 pmol of oligonucleotide in 11 µL nuclease free water were used. 4 µL DMSO and 2 µL RNA loading dye (90 % ROTIPHORESE® Sequencing gel diluent, 10 % ROTIPHORESE® Sequencing gel buffer

concentrate, Bromphenol blue, Xylene cyanol) were added to each sample. All samples were incubated for 3 min at 70 °C and chilled on ice before loading. Electrophoresis was performed for 5-6 h maintaining the electrical power between 45-50 W, the voltage between 1200-1500 V and the electric current between 60-90 mA. The gel was washed with TBE and the separated oligonucleotides stained using SYBR™ Gold Nucleic Acid Gel Stain (10,000 X Concentrate in DMSO, Thermo Fisher Scientific Inc., Waltham (MA), USA) for 20 min in TBE buffer according to the manufacturer's protocol. The gel was washed with TBE and the stained oligonucleotides were visualized using a FLA-5100 Fluorescent Image Analyzer (FUJI PHOTO FILM Co., Ltd., Tokyo, Japan) and fluorescence was detected with an excitation wavelength λ_{ex} of 473 nm.

6.6.10. Labeling of oligonucleotides with ATTO 594-NHS ester

Labeling of oligonucleotides using ATTO 594-NHS esters (ATTO-TEC GmbH, Siegen, Germany) was performed considering the manufacturer's protocol. Oligonucleotides (1 nmol) were lyophilized from water and dissolved in 20 µL gRNA labeling buffer (pH 8.3; 130.5 mM NaCl, 2.6 mM KCl, 8.8 mM NaH_2PO_4, 1.7 mM KH_2PO_4, 9.5 mM $NaHCO_3$; PBS (pH 7.4) : 0.2 M $NaHCO_3$ (pH 9.0) 20:1). ATTO 594-NHS ester (15 nmol, 3.6 mM in DMSO) was added and the reaction was shaken overnight at RT and 500 rpm. Labeled oligonucleotides were purified using NaOAc precipitation as described bevor (see section 6.6.2). The concentration was determined as described bevor (see section 6.3.11) and the degree of labeling (DOL) was determined according to the manufacturer's protocol and as described in equation (1).

$$\mathrm{DOL} = \frac{\mathrm{c\,(dye)}}{\mathrm{c\,(gRNA)}} = \frac{{}^{\mathrm{A}_{603}}/_{\varepsilon_{603}}}{{}^{\mathrm{A}_{260}}/_{\varepsilon_{260}}} = \frac{\mathrm{A}_{603} \cdot \varepsilon_{260}}{(\mathrm{A}_{260} - \mathrm{A}_{603} \cdot \mathrm{CF}_{260}) \cdot \varepsilon_{603}} \qquad (1)$$

with A_λ = absorbance, ε_λ = molar attenuation coefficient and CF_λ = correction factor

6.7. RNA isolation and cDNA synthesis for molecular cloning of ASGPR

6.7.1. RNA isolation of HepG2 cells using TRI Reagent®

For mRNA isolation of HepG2 cells using TRI Reagent®, 200.000 cells were thawed, 500 µL Trizol and 100 µl RNAse free Chloroform were added, subsequently. The sample was vortexed for 30 sec and incubated at RT for 10 min. The sample was vortexed for 10 sec and centrifuged for 20 min at 4 °C and 14.000 rpm. The upper aqueous phase was transferred to a fresh tube, 350 µl ice cold RNAse free isopropanol and 1.5 µL linear acrylamide were added, subsequently. The tube was inverted three times and the mRNA was precipitated overnight at

-20 °C. The precipitation was centrifuged for 1 h at 4 °C and 14.000 rpm and the supernatant was removed. 500 µL $EtOH_{aq}$ (75 % v/v) was added to the pellet and the sample was inverted and centrifuged for 5 min at RT and 14.000 rpm. This step was repeated, the supernatant was removed and the pellet was dried for 4 min at 70 °C. The pellet was dissolved in 30 µL RNAse free H2O, concentration and purity were determined and the sample was stored at -80 °C.

6.7.2. DNAse digestion and reverse transcription

Table 30: General procedure for cDNA syntheses of the ASGPR subunits using reverse transcription. *For reverse transcriptions, only reverse primers are used.*

Component	Amount	
	ASGPR1	**ASGPR2**
DNAse digest	14.75 µL	14.75 µL
P1788 (ASGPR1_bw_XbaI) (10 µM)	1 µL	-
P1790 (ASGPR2_bw_XbaI) (10 µM)	-	1µL
dNTPs	1 µL	1 µL
⇨ **Incubation for 3 min at 70 °C**		
10x MulV Buffer	2 µL	2 µL
RNAse Inhibitor	0.25 µL	0.25 µL
Reverse transcriptase	1 µL	1 µL

Preliminary to the reverse transcription of the isolated RNA, residual DNA was digested according to the general procedure (see Figure 11). cDNA synthesis of the two different ASGPR subunits, using reverse transcription, was performed according to the following procedure (Table 30). Reverse transcription was performed according to the thermocycler program described below (Table 31). cDNA was stored at -20 °C.

Table 31: Thermocycler program for the reverse transcription.

Cycle Step	Temperature [° C]	Time [min]	Cycles
Reverse transcription	42	120	1
Enzyme inactivation	90	10	1
Storage	8	∞	1

6.7.3. Amplification of cDNA and cloning of ASGPR subunits

The cDNA of the two different ASGPR subunits was amplified with PCR using Phusion DNA polymerase according to the general procedure (see section 6.5.3). The following primer sets

were used for the different subunits (Table 32). PCR amplicon was purified using gel electrophoreses (1.4 % (w/v) Agarose in TAE buffer) according to the general procedure (see section 6.5.5) and stored at -20 °C. Restriction digests, using KpnI-HF/XbaI for ASGPR1 and BamHI-HF/XbaI for ASGPR2, ligation into pcDNA™3.1(+), transformation into *E.Coli* XL1-blue and plasmid isolation was performed according to the general procedures described before (see sections 6.5.7). Isolated plasmid DNA was stored at -20 °C and sequenced to reveal specific isoforms for the different subunits.

Table 32: Primer set used for amplification of ASGPR cDNA.

Subunit	Primer No.	Description
ASGPR1	P1787	ASGPR1_fw_KpnI_BamHI
	P1788	ASGPR1_bw_XbaI
ASGPR2	P1789	ASGPR2_fw_BamHI
	P1790	ASGPR2_bw_XbaI

6.8. Cell culture techniques for human cell lines

6.8.1. General procedures for cell culture

6.8.1.1. Cell freezing

For freezing, culture media was removed and the cells were washed with PBS (5 mL for 25 cm^2 or 2 x 5 mL for 75 cm^2 cell culture flasks). PBS was removed and the cells were detached using trypsin/EDTA solution (600 µL for 25 cm^2 or 1 mL for 75 cm^2 cell culture flasks). For detaching, different incubation times are necessary for different cell lines. The detached cells were resuspended in fresh culture medium (4.4 mL for 25 cm^2 or 9 mL for 75 cm^2 cell culture flasks). The resuspended cells were centrifuged for 5 min at 300 *g*, the media was exchanged to DMEM supplemented with 10 % FBS and 10 % DMSO and the concentration was adjusted to $3 \cdot 10^6$ cells/mL. The desired amount of aliquotes (1 mL) was distributed into cryo vials and frozen to -80 °C using a Mr. Frosty™ (Thermo Fisher Scientific Inc., Waltham (MA), USA) freezing container. After 24 h, the frozen cell aliquots were transferred to a cryogenic tank using liquid nitrogen as cryogenic fluid for long time storage.

6.8.1.2. Cell storage

For long time storage, all cell lines were stored under cryogenic conditions using liquid nitrogen as cryogenic fluid. Cells were stored with a concentration of $3 \cdot 10^6$ cells/mL in DMEM supplemented with 10 % FBS and 10 % DMSO.

6.8.1.3. Cell thawing

All cell lines were thawed using Nanopure water at RT. Each vial was placed in fresh nanopure water and was thawed over 5 min at RT. The thawed cell suspension was transferred in a sterile 1.5 mL tube and centrifuged for 5 min at 300 *g*. The supernatant was removed and the cells were resuspended in fresh culture medium (DMEM supplemented with 10 % FBS). The cell suspension was transferred into a fresh cell culture flask, additional culture medium was added up to the recommended volume and the cells were incubated at 37 °C and 5 % CO_2. Cells were cultivated in 25 cm^2 or 75 cm^2 cell culture flasks, the recommended total volumes are 5 mL or 10 mL, respectively. After 24 h, the medium was exchange to fresh culture medium, or if necessary, selective medium containing the required selection markers.

6.8.1.4. Cell cultivation and subcultivation

For cell cultivation, cells were generally cultured in antibiotic free culture medium (DMEM supplemented with 10 % FBS) or in selective medium (DMEM supplemented with 10 % FBS and selective markers). Stable transfected Flp-In™ T-REx™ 293 cell lines were cultured in selective cell culture medium DMEM supplemented with 10 % FBS, 15 µg/mL Blasticidin S hydrochloride and 100 µg/mL Hygromycin B as selective markers. Stable integrated Hela or HepG2 cell lines, using the PiggyBac transposon system, were cultured in selective medium containing DMEM supplemented with 10 % FBS and 1000 µg/mL or 1250 µg/mL G418 as selective markers, respectively.

For subcultivation, culture medium was removed and the adherent cells were washed using PBS (5 mL for 25 cm^2 or 2 x 5 mL for 75 cm^2 cell culture flasks). PBS was removed and the cells were detached using trypsin/EDTA solution (600 µL for 25 cm^2 or 1 mL for 75 cm^2 cell culture flasks) and incubated at 37 °C and 5 % CO_2. For detaching, different incubation times are necessary for different cell lines. The detached cells were resuspended in fresh culture medium (4.4 mL for 25 cm^2 or 9 mL for 75 cm^2 cell culture flasks) and subcultured by adding an appropriate aliquot of the cell suspension to the cell culture flask. Fresh culture medium was added up to the recommended volume as described before.

For further cell culture experiments, the concentration of the cell suspension was determined using a Neubauer-improved Hemocytometer (HBG Henneberg-Sander GmbH, Giessen-Lützellinden, Germany) according to the manufacturer's protocol.

6.8.2. Generation of stable expressing cell lines

6.8.2.1. Generation of FlpIn™ T-REx™ 293

For the generation of stable expressing FlpIn™ T-REx™ 293 cells, parental FlpIn™ T-REx™ 293 host cells, containing the FRT integration site, were cultivated in cell culture medium (DMEM) supplemented with 10 % FBS, 100 µg/mL Zeocin and 15 µg/mL Blasticidin S hydrochloride. $4 \cdot 10^6$ cells were seeded into a 10 cm cell culture dish and the cells were incubated for 24 h at 37 °C and 5 % CO_2. Before transfection, the cell culture medium was exchanged to 10 mL antibiotic free DMEM supplemented with 10 % FBS and the cells were transfected with 1 µg pcDNA5™/FRT vector, carrying the gene of interest and 9 µg pOG44 (Thermo Fisher Scientific Inc., Waltham (MA), USA) using 30 µL Lipofectamine™ 2000 (Thermo Fisher Scientific Inc., Waltham (MA), USA) according to the manufacturer's protocol. Each, plasmid combination (pcDNA5 + pOG44) and Lipofectamine™ 2000 were diluted in OptiMEM™ (Thermo Fisher Scientific Inc., Waltham (MA), USA) to a total volume of 1.5 mL and the separate solutions were incubated for 5 min at RT. Both solutions were combined carefully and incubated for further 20 min at RT. The transfection mixture was added dropwise to the cell culture dish and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, the medium was exchanged to selective culture medium (DMEM supplemented with 10 % FBS, 15 µg/mL Blasticidin S hydrochloride and 100 µg/mL Hygromycin B as selective markers). The medium was exchanged every 3 days and the cells were cultured for 14 days until colonies were formed. The colonies were carefully washed with PBS (10 mL), detached using trypsin/EDTA solution (1 mL), resuspended in selective culture medium and transferred to a desired cell cultures flask.

6.8.2.2. Generation of HepG2 and Hela cells using the PiggyBac transposon system

For the generation of stable expressing HepG2 cells using the PiggyBac transposon system, $1 \cdot 10^5$ cells were seeded per well into a 24-well plate in 500 µL antibiotic free DMEM supplemented with 10 % FBS and the cells were incubated for 24 h at 37 °C and 5 % CO_2. Before transfection, the culture medium was exchanged to 400 µL fresh cell culture medium (antibiotic free DMEM supplemented with 10 % FBS) and the cells were transfected with 0.75 µg carrier plasmid (XLone or PB), containing the gene of interest and 0.25 µg transposase containing plasmid using 3 µL FuGENE® 6 (Promega GmbH, Walldorf, Germany) according to the manufacturer's protocol. Each, plasmid combination (carrier and transposase containing plasmids) and FuGENE® 6 were diluted in OptiMEM™ (Thermo Fisher Scientific Inc., Waltham (MA), USA) to a total volume of 50 µL and the separate solutions were incubated for

5 min at RT. Both solutions were combined carefully and incubated for further 20 min at RT. The transfection mixture was added dropwise to the cell containing well and the cells were incubated for 48 h at 37 °C and 5 % CO_2. After 48 h, the cells were washed with PBS (500 µL), detached with trypsin/EDTA solution (60 µL), resuspended in fresh selective cell culture medium (DMEM supplemented with 10 % FBS and 1250 µg/mL G418 as selective markers) and transferred to a 6-well plate. The medium was exchanged every 3 days and, if necessary, the cells were subcultivated, for 14 days. The cells were washed with PBS (2.5 mL), detached using trypsin/EDTA solution (300 µL), resuspended in selective cell culture medium and transferred to a desired cell cultures flask.

For the generation of stable expressing Hela cells using the PiggyBac transposon system, $1.5 \cdot 10^5$ cells were seeded per well into a 24-well plate in 500 µL antibiotic free DMEM supplemented with 10 % FBS and the cells were incubated for 24 h at 37 °C and 5 % CO_2. Before transfection, the culture medium was exchanged to 400 µL fresh cell culture medium (antibiotic free DMEM supplemented with 10 % FBS) and the cells were transfected with 0.75 µg carrier plasmid (pTS1340), containing the gene of interest and 0.25 µg transposase containing plasmid (pTS687) using 3 µL FuGENE® 6 (Promega GmbH, Walldorf, Germany) according to the manufacturer's protocol. Each, plasmid combination (carrier and transposase containing plasmids) and FuGENE® 6 were diluted in OptiMEM™ (Thermo Fisher Scientific Inc., Waltham (MA), USA) to a total volume of 50 µL and the separate solutions were incubated for 5 min at RT. Both solutions were combined carefully and incubated for further 20 min at RT. The transfection mixture was added dropwise to the cell containing well and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, the cells were washed with PBS (500 µL), detached with trypsin/EDTA solution (60 µL), resuspended in fresh antibiotic free cell culture medium (DMEM supplemented with 10 % FBS), transferred to a 6 cm cell culture dish and incubated for 24 h at 37 °C and 5 % CO_2. After 24 h the medium was exchanged to selective cell culture medium (DMEM supplemented with 10 % FBS and 1000 µg/mL Geniticin™ (G418) as selective marker). The medium was exchanged every 3 days and the cells were cultured for 8 days until colonies were formed. The colonies were washed with PBS (2 x 5 mL), detached using trypsin/EDTA solution (500 µL), resuspended in selective cell culture medium and transferred to a desired cell cultures flask.

6.8.2.3. FACS sorting of HeLa cells stably expressing ASGPR1

For FACS sorting of the stable integrated Hela cells, $4 \cdot 10^6$ cells were seeded into a 10 cm cell culture dish in 10 mL antibiotic free DMEM, supplemented with 10 % FBS and 200 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. As reference, and for fluorophore compensation, $4 \cdot 10^6$ cells were additionally seeded into a 10 cm cell culture dish in 10 mL antibiotic free DMEM, supplemented with 10 % FBS, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, the culture medium of the induced cells was exchanged to 10 mL fresh antibiotic free DMEM supplemented with 10 % FBS, 200 ng/mL doxycycline and 1 µM GalNAc-FITC (**15**, 3.66 mM in DMSO) and the cells were further incubated for 1.5 h at 37 °C and 5 % CO_2. The culture medium of the uninduced cells was exchanged to 10 mL antibiotic free DMEM supplemented with 10 % FBS and the cells were further incubated for 1.5 h at 37 °C and 5 % CO_2. Each condition of the seeded cells was washed with PBS (2 x 5 mL) and analyzed via fluorescence microscopy. After imaging, PBS was removed and the cells were detached using trypsin/EDTA solution (1 mL) at 37 °C and 5 % CO_2. The detached cells were resuspended in 10 mL fresh antibiotic free DMEM, supplemented with 10 % FBS, centrifuged for 5 min at 300 *g* and resuspended in 10 mL PBS. The cells were centrifuged again for 5 min at 300 *g* and resuspended in 1 mL PBE (PBS, supplemented with 0.5 % FBS and 5 mM EDTA) and stored on ice in coated tubes (PBE, 4 °C, overnight). The cell suspension was filtered into coated tubes (PBE, 4 °C, overnight) using a 40 µm filter and sorted according to their FITC signal intensity using a MA900 Multi-Application Cell Sorter (Sony Biotechnology Inc., San Jose (CA), USA) into coated tubes (antibiotic free DMEM, supplemented with 10 % FBS, 4 °C, overnight) containing 500 µL antibiotic free DMEM, supplemented with 10 % FBS. For fluorescence detection, a wavelength of λ_{ex} = 488 nm was used for excitation, and a FL1 (525/50) optical filter set was used for detection. The sorted cells were centrifuged for 5 min at 300 *g,* resuspended 2.5 mL DMEM, supplemented with 10 % FBS and 1000 µg/mL Geniticin™ (G418), plated into a 6-well cell culture plate and incubated at 37 °C and 5 % CO_2. After 48 h, the media was exchanged to fresh DMEM, supplemented with 10 % FBS and 1000 µg/mL Geniticin™ (G418). The confluent cells were washed with 2.5 mL PBS, detached using trypsin/EDTA solution (300 µL) at 37 °C and 5 % CO_2, resuspended 4.7 mL DMEM, supplemented with 10 % FBS and 1000 µg/mL Geniticin™ (G418) and transferred into a 25 cm^2 cell culture flask. The sorted cells were subcultured for further cell culture experiments and long time storage under cryogenic conditions as described within the general procedures.

6.8.3. Characterizations of stable expressing cell lines

6.8.3.1. Coating

All imaging experiments and passive uptake experiments with FlpIn™ T-REx™ 293 cell lines were performed using poly-D-lysin coated well plates. For immunofluorescence experiments using a 24-well scale, each well was equipped with a cover glass (⌀ 12 mm, VWR International, Leuven, Belgium). For live cell imaging experiments using a 96-well scale, imaging was performed using 96-well cell imaging plates (Eppendorf AG, Hamburg, Germany). For coating, each 24-well was coated using 500 µL and each 96-well was coated using 100 µL poly-*D*-Lysine HBr_{aq} (0.1 mg/mL, 70,000-150,000 mol wt, Merck KGaA, Darmstadt, Germany). The wells were incubated for 30 min at RT and washed twice with PBS applying the same volume as for coating. The coated wells were dried for 5 min at RT and sterilized under UV light illumination overnight or at least for 30 min. The coating solution was reused 5-7 times.

6.8.3.2. Transfection of the ASGPR subunits to induce receptor mediated uptake of GalNAc-FITC

For the transfection of the ASGPR subunits into HEK 293T cells, 30.000 cells/well were seeded into poly-*D*-lysine HBr_{aq} coated 96-well imaging plates in 100 µL antibiotic free culture medium (DMEM supplemented with 10 % FBS). The cells were incubated for 24 h and prior to transfection, the media was exchanged to 90 µL fresh antibiotic free culture medium (DMEM supplemented with 10 % FBS). For each transfection, a total amount of 100 ng plasmid (H1a (100 ng), H2b (100 ng) or H1a + H2b (50 ng + 50 ng)) was diluted with OptiMEM™ to final volume of 10 µL. As negative control, no vector was used for the transfection. In addition 0.4 µL Lipofectamine™ 2000 (4 µL/µg) were diluted with OptiMEM™ to final volume of 10 µL for each transfection, and the separate dilutions were incubated for 5 min RT. The corresponding dilutions were combined and incubated for further 20 min at RT. The transfection mixture was added to the cells and the cells were incubated for 24 h. The media was exchanged to fresh antibiotic free culture medium (DMEM supplemented with 10 % FBS) containing 1 µM GalNAc-FITC (**15**) and the cells were incubated for further 24 h. After 24 h, the media was removed and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100, Thermo Fisher Scientific Inc., Waltham (MA), USA) in 40 µL PBS. The cells were washed with PBS (2 x 50 µL) and maintained in 40 µL phenol red free DMEM supplemented with 10 % FBS for live cell imaging and fluorescence microscopy.

6.8.3.3. Receptor mediated uptake of GalNAc-FITC into H1a expressing cell lines

For the receptor mediated uptake of GalNAc-FITC (**15**) into stable integrated ASGPR H1a FlpIn™ T-Rex cells, 15.000 cells/well were seeded into poly-*D*-lysine HBr_{aq} coated 96-well imaging plates in 100 µL antibiotic free culture medium (DMEM supplemented with 10 % FBS) under doxycycline induction (10 ng/mL). The cells were incubated for 24 h and the media was exchanged to 100 µL fresh antibiotic free culture medium (DMEM supplemented with 10 % FBS) containing 1 µM GalNAc-FITC (**15**) as well as doxycycline (10 ng/mL) and the cells were incubated for further 24 h. After 24 h, the media was removed and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100, Thermo Fisher Scientific Inc., Waltham (MA), USA) in 40 µL PBS. The cells were washed with PBS (2 x 50 µL) and maintained in 40 µL phenol red free DMEM supplemented with 10 % FBS for live cell imaging and fluorescence microscopy.

For the receptor mediated uptake of GalNAc-FITC (**15**) into stable integrated ASGPR H1a HeLa cells, 10.000 cells/well were seeded into poly-*D*-lysine HBr_{aq} coated 96-well imaging plates in 100 µL antibiotic free culture medium (DMEM supplemented with 10 % FBS) under doxycycline induction (200 ng/mL). The cells were incubated for 24 h and the media was exchanged to 100 µL fresh antibiotic free culture medium (DMEM supplemented with 10 % FBS) containing 1 µM GalNAc-FITC (**15**) as well as doxycycline (200 ng/mL) and the cells were incubated for further 24 h. After 24 h, the media was removed and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100, Thermo Fisher Scientific Inc., Waltham (MA), USA) in 40 µL PBS. The cells were washed with PBS (2 x 50 µL) and maintained in 40 µL phenol red free DMEM supplemented with 10 % FBS for live cell imaging and fluorescence microscopy.

6.8.3.4. BG-FITC staining of SNAP®-ADAR expressing cells

BG-FITC staining of SNAP®-ADAR expressing cells was performed considering the general procedure for immunofluorescence as described below (see section 6.8.3.5). Cells were seeded on coated cover glasses and incubated under the desired conditions at 37 °C and 5 % CO_2. For BG-FITC staining, the cell culture medium was removed to a remaining volume of 197 µL, 3 µL imaging solution, containing 0.04 µL acetylated BG-FITC (10 mM in DMSO), 0.96 µL DMEM supplemented with 10 % FBS, and 2 µL NucBlue™ Live ReadyProbes™ Reagent, was added to each well and the cells were incubated for 30 min at RT. For fixation, *p*-formaldehyde (21.6 µL, 37 % w/w) was added to each well and the cells were incubated for 10 min at RT.

The cells were washed with PBS (3 x 500 µL) and for permeabilization, the cells were incubated for 15 min at RT with 0.1 % Triton X 100 in 200 µL PBS. The cells were washed with PBS (3 x 500 µL) and were either mounted on microscope slides using Dako Fluorescent mounting medium for microscopy or the cells were blocked overnight at 4 °C in 500 µL PBS, supplemented with 10 % FBS and further immunofluorescence staining was applied as described below (see section 6.8.3.5).

6.8.3.5. Immunofluorescence staining of *p*-formaldehyde fixated cells

Immunofluorescence experiments were performed on poly-*D*-Lysine HBr_{aq} coated cover classes in a 24-well scale (see section 6.8.3.1). Cells were seeded on coated cover glasses and incubated under the desired conditions at 37 °C and 5 % CO_2. Cells were fixated by removing the cell culture medium to a remaining volume of 200 µL. For fixation, *p*-formaldehyde (21.6 µL, 37 % w/w) was added to each well and the cells were incubated for 10 min at RT. The cells were washed with PBS (3 x 500 µL) and blocked overnight at 4 °C in 500 µL PBS, supplemented with 10 % FBS. The cells were incubated for 2 h at RT with or without the primary antibody *mouse* α-ASGPR1 (1:1000, 8D7, Santa Cruz Biotechnology, Inc., Heidelberg, Germany) in 200 µL PBS supplemented with 5 % FBS and the cells were washed with PBS (3 x 500 µL). The cells were incubated for 1 h at RT with or without secondary antibodies *goat* α-*mouse* Alexa Fluor 488 (A11001) or *goat* α-*mouse* Alexa Fluor 647 (A21235) (1:1000, Thermo Fisher Scientific Inc., Waltham (MA), USA) in 250 µL PBS supplemented with 10 % FBS and the cells were washed with PBS (2 x 500 µL). Nuclei were stained for 30 min at RT using NucBlue™ Live ReadyProbes™ Reagent (1:100, Thermo Fisher Scientific Inc., Waltham (MA), USA) in 200 µL PBS. The cells were washed with PBS (3 x 500 µL) and mounted on microscope slides (76 x 26 mm, VWR International, Leuven, Belgium) using Dako Fluorescent mounting medium (Agilent Technologies, Inc., Santa Clara (CA), USA). The mounted cells were dried for 15 min at RT and overnight at 4 °C before microscopy.

6.8.4. Preparation of whole cell lysates, protein separation and western blotting

For the analysis of doxycycline inducible expression of proteins using western blot analysis, different generated cell lines were seeded in antibiotic free cell culture medium (DMEM supplemented with 10 % FBS) with and without doxycycline induction. The cells were incubated for 24 h at 37 °C and 5 % CO_2. For detailed information about cell seeding and doxycycline induction, see Table 33.

Table 33: Information about sample preparation for western blotting. *To ensure ASGPR protein stability, FlpIn™ T-REx™ 293 and Hela cells were lysed according to Scharner et al.[269] using RIPA cell lysis buffer. HepG2, FlpIn™ T-REx™ and Hela cells were cultured in DMEM supplemented with 10 % FBS. Primary Hepatocytes were seeded in fresh plating medium into Collagen I coated plates (35 µg/mL Collagen I in 0.1 % acetic acid) and cultured in maintenance medium (medium was exchanged 4 h after seeding) according to the manufacturer's protocol (Lonza Walkersville Inc., Walkersville (MD), USA). The listed volume of lysis buffer was used for each well.*

Cell line	Well scale	c (doxy-cycline) [ng/mL]	V (cell culture medium) [mL]	Number of cells	Lysis Buffer	V (Lysis buffer) [µL]
HepG2	2 x 24	0-1000	0.5	100.000	NP40	100
	2 x 24	-	0.5	100.000	NP40	100
FlpIn™ T-REx™	2 x 24	10	0.5	300.000	RIPA	50
	2 x 24	-	0.5	300.000	RIPA	50
Hela	2 x 6	200	2.5	750.000	RIPA	250
	2 x 6	-	2.5	750.000	RIPA	250
Primary Hepatocytes	1 x 24	-	0.5	450.000	RIPA	50

6.8.4.1. Cell lysis

For cell lysis, as indicated in Table 33, cells were lysed in different cell lysis buffers. HepG2 cells were washed with 500 µL cold PBS and detached using 60 µL trypsin/EDTA. The detached cells were resuspended in 440 µL antibiotic free cell culture medium, transferred in a fresh 1.5 ml tube and centrifuged for 5 min at 210 *g* and 4 °C. The pellet was washed twice with 500 µL cold PBS and lysed for 30 min on ice in 100 µL NP40 lysis buffer, supplemented with a cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (1 tablet/10 mL, Roche, Basel, Switzerland). The lyse was vortexed in 10 min intervals. The lysate was transferred in a fresh 1.5 mL tube and centrifuged for 10 min at 13.000 rpm and 4 °C. The clear supernatant was transferred in a fresh 1.5mL tube and the whole cell lysate was stored at -80 °C. FlpIn™ T-REx™ 293 and Hela cells were carefully washed twice with 500 µL cold PBS and the cells were lysed for 30 min on ice and overnight at -80 °C in the desired amount (see Table 33) RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific Inc., Waltham (MA), USA), supplemented with a cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (1 tablet/10 mL, Roche, Basel, Switzerland). The lysate was transferred in a fresh 1.5 mL tube and centrifuged for 15 min at 18.000 rpm and 4 °C. The clear supernatant was transferred in a fresh 1.5mL tube and the whole cell lysate was stored at -80 °C.

The total amount of protein of the whole cell lysates was determined using a Pierce™ BCA Pierce™ BCA Protein Assay Kit as described before (see section 6.3.11)

6.8.4.2. SDS-PAGE for protein separation

Protein separation was performed using Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Novex™ WedgeWell™ 8-16 %, tris-glycine, 1.0 mm, Mini Protein Gels (15-well, Thermo Fisher Scientific Inc., Waltham (MA), USA). Whole cell lysates were denatured using 6 x Lämmli buffer (60 mM Tris-HCl_{aq}, pH 6.8, 0.6 M DTT, 12 % SDS, 47 % glycerol, 0.06 % bromophenol blue), incubated for 5 min at 95 °C, chilled on ice and applied to the gel. As reference, PageRuler™ Plus Prestained Protein Ladder (10 to 250 kDa, Thermo Fisher Scientific Inc., Waltham (MA), USA) was used. Protein separation was performed in SDS-PAGE running buffer (192 mM Glycine, 25 mM Tris, 0.1 % (w/v) SDS) for 10 min at 90 V and 90 min at 150 V, subsequently. SDS-PAGE was carried out using a ROTIPHORESE® PROclamp MINI electrophoresis unit (Carl Roth GmbH + Co. KG, Karlsruhe, Germany).

6.8.4.3. Western blotting

For western blotting, the separated proteins using SDS-PAGE, were transferred onto PVDF membranes using a Mini Trans-Blot® Cell (Bio-Rad Laboratories GmbH, Feldkirchen, Germany). Before membrane-sandwich assembly, PVDF membranes (Thermo Fisher Scientific Inc., Waltham (MA), USA) were activated in methanol and transfer buffer (25 mM Tris-HCl_{aq}, 190 mM glycine, 20 % methanol), and proteins were transferred for 18 h at 4 °C and 30 V. After membrane sandwich disassembly, the membranes were washed with TBST (50 mM Tris-HCl_{aq}, 150 mM NaCl, 0.1 % TWEEN® 20) and blocked with 5 % non-fat dry milk in TBST for 1 h at RT. Primary and secondary antibodies were diluted in 5 % non-fat dry milk or 5 % BSA in TBST. Detailed information about antibody dilutions, blocking reagent and incubation conditions are listed in Table 34. After each incubation (primary and secondary antibodies) membranes were washed with TBST (3 x 5 mL). For visualization, each membrane was treated with Clarity™ Western ECL Substrates (Bio-Rad Laboratories GmbH, Feldkirchen, Germany) according to the manufacturer's protocol or with homemade ECl solution (100 mM Tris-HCl_{aq} (pH 8.5), 0.12 mM Luminol, 0.2 mM *p*-Coumaric acid in Nanopure water + 1 µL/mL H_2O_2 (30 % w/v) right before use). Chemiluminescence was detected as described above, using a Fusion S2 Vilber Lourmat imaging system equipped with a CCD

camera (Peqlab, VWR Life Science) or an Odyssey Fc Imaging System (LI-COR® Biosciences, Lincoln (NE), USA) (see section 6.3.10). Pictures were processed using ImageJ 1.49m.

Table 34: Information about antibody treatment used for western blotting.

Antibody	Host	Blocking reagent	Dilution	Incubation condition
α-SNAP	*rabbit*	5 % non-fat dry milk	1:1000	1 h / RT + overnight / 4 °C
α-GAPDH	*mouse*	5 % non-fat dry milk	1:1000	overnight / 4 °C
α-*mouse* HRP	*goat*	5 % non-fat dry milk	1:5000	2 h / RT
α-*rabbit* HRP	*goat*	5 % non-fat dry milk	1:5000	2 h / RT
α-SNAP	*rabbit*	5 % BSA	1:1000	1 h / RT + overnight / 4 °C
α-GAPDH	*mouse*	5 % BSA	1:1000	overnight / 4 °C
α-ASGPR1	*mouse*	5 % BSA	1:500	1 h / RT + 72 h / 4 °C

6.8.5. Transfections and receptor mediated uptakes of gRNAs

6.8.5.1. Transfection of SNAP®-ADAR gRNAs 324 and 507

For the transfection of gRNA 324 and 507 into SA1Q or SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell lines, 3 x 300.000 cells/well were seeded into 24-well cell culture plates in 500 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and 10 ng/mL doxycycline and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After induction, the cells were washed with PBS (500 µL) and detached using trypsin/EDTA solution (60 µL) at 37 °C and 5 % CO_2. The detached cells were resuspended in 440 µL fresh antibiotic free DMEM, supplemented with 10 % FBS, and the cell suspension of each cell line was combined. The cell concentration was determined using a Neubauer-improved Hemocytometer (HBG Henneberg-Sander GmbH, Giessen-Lützellinden, Germany) as described before and 800.000 cells of each cell line were centrifuged for 5 min at 300 x *g*. The supernatant was removed and the cells were resuspended in 1 mL antibiotic free culture medium (DMEM), supplemented with 10 % FBS and 10 ng/mL doxycycline to a final concentration of 80.000 cells/100 µL. For each transfection, 2 pmol of each gRNA were diluted with OptiMEM™ to final volume of 25 µL. As negative control, unmodified and NH_2-terminal gRNAs were used and each condition was performed in duplicates (2 x 96-well, see below). In addition, 0.75 µL Lipofectamine™ 2000 were diluted with OptiMEM™ to a final volume of 25 µL for each

transfection, and the separate dilutions were incubated for 5 min RT. The corresponding dilutions were combined and incubated for further 20 min at RT. Each transfection mixture (50 µL) was added in advance to a 96-well cell culture plate, followed by the addition of 80.000 cells/well in 100 µL antibiotic free culture medium (DMEM), supplemented with 10 % FBS and 10 ng/mL doxycycline and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After transfection, the media was removed and the cells were harvested using a Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany) and further processed as described before (see section 6.5.9).

6.8.5.2. Receptor mediated endocytosis of ASGPR H1a and SA1Q expressing FlpIn™ T-REx™ 293 cells

For the receptor mediated endocytosis of gRNAs into ASGPR H1a and SA1Q expressing FlpIn™ T-REx™ 293 cells, $1.25 \cdot 10^4$ cells/well were seeded into coated 96-well plates in 100 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS and 10 ng/mL doxycycline and incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, 70 µL of the medium was removed and 30 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS, 10 ng/mL doxycycline and 0.4 or 2 µM gRNA was added and the cells were incubated for 48 h at 37 °C and 5 % CO_2. The final gRNA concentration is accordingly 0.2 or 1 µM. After 48h, 0.6 µL doxycycline (1 µg/mL) was spiked in into each well and the cells were incubated for further 24 h at 37 °C and 5 % CO_2. The media were removed and the cells were harvested using a Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany) and further processed as described before (see section 6.5.9).

6.8.5.3. Transfection of ASGPR H1a into SA1Q expressing FlpIn™ T-REx™ 293 cells and receptor mediated endocytosis

For the transfection of the receptor subunit H1a into SA1Q expressing FlpIn™ T-REx™ 293 cells, 4 x 300.000 cells/well were seeded into a 24-well cell culture plate in 500 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and the cells were incubated overnight at 37 °C and 5 % CO_2. Prior to transfection, the medium of each well was exchanged to 500 µL fresh antibiotic free culture medium (DMEM) supplemented with 10 % FBS and for each transfection, a total amount of 500 ng plasmid (pTS1070) was diluted with OptiMEM™ to final volume of 25 µL. Each transfection was performed in duplicates, and as negative control, no vector was used for the transfection (mock control). In addition, 2 µL FuGENE® 6 (4 µL/µg) were diluted with OptiMEM™ to final volume of 25 µL for each

transfection, and the separate dilutions were incubated for 5 min RT. The corresponding dilutions were combined and incubated for further 20 min at RT. The transfection mixture was added dropwise onto the cells and the cells were incubated for 24 h. The transfected cells were detached using trypsin/EDTA solution (60 µL) at 37 °C and 5 % CO_2 and the detached cells were resuspended in 440 µL fresh antibiotic free DMEM, supplemented with 10 % FBS and 10 ng/mL doxycycline. The cell concentration was determined using a Neubauer-improved Hemocytometer (HBG Henneberg-Sander GmbH, Giessen-Lützellinden, Germany) as described before. $1.25 \cdot 10^4$ cells/well were seeded into coated 96-well plates in 100 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS and 10 ng/mL doxycycline and incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, 70 µL of the medium was removed and 30 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS, 10 ng/mL doxycycline and 0.4 or 2 µM gRNA was added and the cells were incubated for 48 h at 37 °C and 5 % CO_2. The final gRNA concentration is accordingly 0.2 or 1 µM. After 48h, 0.6 µL doxycycline (1 µg/mL) was spiked in into each well and the cells were incubated for further 24 h at 37 °C and 5 % CO_2. The media were removed and the cells were harvested using a Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany) and further processed as described before (see section 6.5.9).

To determine the transfection efficiency, $1.25 \cdot 10^4$ cells/well were seeded into a coated 96-well plates in 100 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS and 10 ng/mL doxycycline and incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, 70 µL of the medium was removed and 30 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS, 10 ng/mL doxycycline and 2 µM GalNAc-FITC (DMSO, 3.66 mM) was added and the cells were incubated for 1 h at 37 °C and 5 % CO_2. The final GalNAc-FITC (**15**) concentration is accordingly 1 µM. The cells were washed with PBS (2 x 100 µL) and maintained in 40 µL phenol red free DMEM supplemented with 10 % FBS for live cell imaging and fluorescence microscopy.

6.8.5.4. Transfection of ATTO 594 labeled SNAP®-ADAR gRNAs 324 and 507

For the transfection of ATTO 594 (ATTO-TEC GmbH, Siegen, Germany) labeled gRNA 324 and 507 into SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell lines, 15.000 cells/well were seeded into coated 96-well cell imaging plates (Eppendorf AG, Hamburg, Germany) in 100 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and the cells were incubated for 24 h at 37 °C and 5 % CO_2. Prior to transfection, 50 µL of the medium was

removed and 50 µL fresh antibiotic free culture medium, (DMEM) supplemented with 10 % FBS, was added. For each transfection, 5 pmol of each gRNA were diluted with OptiMEM™ to a final volume of 10 µL. As negative control, no gRNA was used. In addition, 0.3 µL Lipofectamine™ RNAiMAX (0.3 µL/well, Thermo Fisher Scientific Inc., Waltham (MA), USA) were diluted with OptiMEM™ to a final volume of 10 µL for each transfection, and the separate dilutions were incubated for 5 min at RT. The corresponding dilutions were combined and incubated for further 20 min at RT. Each transfection mixture (20 µL) was added to the cells and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After transfection, 80 µL of the medium was removed, 4.32 µL *p*-formaldehyde (37 %) was added and the cells were fixated for 10 min at RT. The cells were washed carefully with PBS (3 x 100 µL) and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100) in 40 µL PBS. The cells were washed carefully with PBS (2 x 100 µL) and the cells were maintained in 40 µL phenol red free DMEM, supplemented with 10 % FBS, for fluorescence microscopy.

6.8.5.5. Receptor mediated endocytosis of ATTO 594 labeled SNAP®-ADAR gRNAs 324 and 507

For the receptor mediated endocytosis of ATTO 594 labeled gRNA 324 and 507 into SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell lines, 15.000 cells/well were seeded into coated 96-well cell imaging plates (Eppendorf AG, Hamburg, Germany) in 100 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and 10 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, 70 µL of the medium was removed and 30 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % FBS, 10 ng/mL doxycycline and 2 µM gRNA, was added and the cells were incubated for 24 h at 37 °C and 5 % CO_2. The final gRNA concentration is accordingly 1 µM. After 24 h, 20 µL of the medium was removed, 4.32 µL *p*-formaldehyde (37 %) was added and the cells were fixated for 10 min at RT. The cells were washed with PBS (3 x 100 µL) and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100) in 40 µL PBS. The cells were washed carefully with PBS (2 x 100 µL) and the cells were maintained in 40 µL phenol red free DMEM, supplemented with 10 % FBS, for fluorescence microscopy.

6.8.5.6. Transfection of RESTORE gRNAs TMR189 and TMR236

For the transfection of gRNA TMR189 and TMR236 into H1a expressing HeLa cells, 100.000 cells/well were seeded into 24-well cell culture plates in 500 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and 200 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. Prior to transfection, the media was exchanged to 450 µL fresh antibiotic free culture medium (DMEM) supplemented with 10 % FBS and for each transfection, 25 pmol of each gRNA were diluted with OptiMEM™ to a final volume of 50 µL. As negative control, no gRNA was transfected. In addition, 1.5 µL Lipofectamine™ RNAiMAX were diluted with OptiMEM™ to a final volume of 50 µL for each transfection, and the separate dilutions were incubated for 5 min RT. The corresponding dilutions were combined and incubated for further 20 min at RT. Each transfection mixture (100 µL) was added dropwise to the cells and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After transfection, the media was removed and the cells were harvested using a Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany) and further processed as described before (see section 6.5.9).

6.8.5.7. Receptor mediated endocytosis of unsorted ASGPR H1a expressing HeLa cells

For passive uptake experiments with H1a expressing HeLa cells and chloroquine as additive, 50.000 cells/well were seeded into 24-well plates in 500 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % FBS and with or without 200 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, the cell culture medium was exchanged to 300 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % FBS, and with or without 0.2 µM or 1 µM gRNA, 200 ng/mL doxycycline and 60 µM chloroquine (1 mM in nuclease free water), and the cells were incubated for 24 / 48 / 72 h at 37 °C and 5 % CO_2. After incubation, the cells were harvested using a Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany) and further processed as described before (see section 6.5.9).

6.8.5.8. Receptor mediated endocytosis of sorted ASGPR H1a expressing HeLa cells

For passive uptake experiments with H1a expressing HeLa cells and chloroquine as additive, 50.000 cells/well were seeded into 24-well plates in 500 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % FBS and with or without 200 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. For passive uptake experiments with H1a expressing HeLa cells and without chloroquine as additive, 10.000 cells/well were seeded into

24-well plates in 500 µL antibiotic free cell culture medium (DMEM), supplemented with 3 % FBS and with or without 200 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, the cell culture medium was exchanged to 300 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % or 3 % FBS, respectively, and with or without 0.2 µM gRNA, 200 ng/mL doxycycline and 60 µM chloroquine (1 mM in nuclease free water), and the cells were incubated for 48 h at 37 °C and 5 % CO_2. After incubation, the cells were harvested using a Monarch® RNA Cleanup Kit 10 µg (New England Biolabs GmbH, Frankfurt am Main, Germany) and further processed as described before (see section 6.5.9).

6.8.5.9. Transfection of ATTO 594 labeled RESTORE v2 gRNAs TMR189 and TMR236

For the transfection of ATTO 594 (ATTO-TEC GmbH, Siegen, Germany) labeled gRNA TMR189 and TMR236 into unsorted H1a expressing HeLa cells, 5.000 cells/well were seeded into coated 96-well cell imaging plates (Eppendorf AG, Hamburg, Germany) in 100 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and the cells were incubated for 24 h at 37 °C and 5 % CO_2. Prior to transfection, 50 µL of the medium was removed and 50 µL fresh antibiotic free culture medium, (DMEM) supplemented with 10 % FBS, was added. For each transfection, 5 pmol of each gRNA were diluted with OptiMEM™ to a final volume of 10 µL. As negative control, no gRNA was used. In addition, 0.3 µL Lipofectamine™ RNAiMAX (0.3 µL/well, Thermo Fisher Scientific Inc., Waltham (MA), USA) were diluted with OptiMEM™ to a final volume of 10 µL for each transfection, and the separate dilutions were incubated for 5 min at RT. The corresponding dilutions were combined and incubated for further 20 min at RT. Each transfection mixture (20 µL) was added to the cells and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After transfection, 80 µL of the medium was removed, 4.32 µL *p*-formaldehyde (37 %) was added and the cells were fixated for 10 min at RT. The cells were washed with PBS (3 x 100 µL) and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100) in 40 µL PBS. The cells were washed carefully with PBS (2 x 100 µL) and the cells were maintained in 40 µL phenol red free DMEM, supplemented with 10 % FBS, for fluorescence microscopy.

6.8.5.10. Receptor mediated endocytosis of ATTO 594 labeled RESTORE v2 gRNAs TMR189 and TMR236

For the transfection of ATTO 594 (ATTO-TEC GmbH, Siegen, Germany) labeled gRNA TMR189 and TMR236 into unsorted H1a expressing HeLa cells, 5.000 cells/well were seeded into coated 96-well cell imaging plates (Eppendorf AG, Hamburg, Germany) in 100 µL (each)

antibiotic free culture medium (DMEM), supplemented with 10 % FBS and in the absence or presence of 200 ng/mL doxycycline, and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, 70 µL of the medium was removed and 30 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % FBS, with or without 200 ng/mL doxycycline and 2 µM gRNA, was added and the cells were incubated for 24 h at 37 °C and 5 % CO_2. For the addition of chloroquine, 70 µL of the medium was removed and 30 µL antibiotic free cell culture medium (DMEM), supplemented with 10 % FBS, with or without 200 ng/mL doxycycline, and with or without 120 µM chloroquine (1 mM in nuclease free water) and 2 µM gRNA, was added and the cells were incubated for 24 h at 37 °C and 5 % CO_2. The final gRNA concentration is accordingly 1 µM and the final chloroquine concentration 60 µM. After 24 h, 20 µL of the medium was removed, 4.32 µL *p*-formaldehyde (37 %) was added and the cells were fixated for 10 min at RT. The cells were washed with PBS (3 x 100 µL) and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100) in 40 µL PBS. The cells were washed carefully with PBS (2 x 100 µL) and the cells were maintained in 40 µL phenol red free DMEM, supplemented with 10 % FBS, for fluorescence microscopy.

The samples for 0-2 h were further prepared without fixation after the desired time of incubation. The cells were washed with PBS (3 x 100 µL) and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100) in 40 µL PBS. The cells were washed carefully with PBS (2 x 100 µL) and the cells were maintained in 40 µL phenol red free DMEM, supplemented with 10 % FBS, for fluorescence microscopy.

6.8.5.11. Transfection of ASGPR H1a into A549, Huh7, U2OS and U87 cells and receptor mediated endocytosis of GalNAc-FITC

For the transfection of the receptor subunit H1a into A549, Huh7, U2OS and U87 cells, for each cell line 10 x 75.000 cells/well were seeded into a 24-well cell culture plate, in 500 µL (each) antibiotic free culture medium (DMEM), supplemented with 10 % FBS and the cells were incubated overnight at 37 °C and 5 % CO_2. Prior to transfection, the medium of each well was exchanged to 500 µL fresh antibiotic free culture medium (DMEM) supplemented with 10 % FBS and for each transfection, 500 ng plasmid (pTS1070) were diluted with OptiMEM™ to final volume of 25 µL. Each transfection was performed in duplicates, and as negative control, no vector was used for the transfection (mock). In addition, 2 µL FuGENE® 6 (4 µL/µg) were diluted with OptiMEM™ to final volume of 25 µL for each transfection, and the separate

dilutions were incubated for 5 min RT. The corresponding dilutions were combined and incubated for further 20 min at RT. The transfection mixture was added dropwise onto the cells and the cells were incubated for 24 h. The transfected cells were detached using trypsin/EDTA solution (60 µL) at 37 °C and 5 % CO_2 and the detached cells were resuspended in 440 µL fresh antibiotic free DMEM, supplemented with 10 % FBS. The cell concentration was determined using a Neubauer-improved Hemocytometer (HBG Henneberg-Sander GmbH, Giessen-Lützellinden, Germany) as described before. $5 \cdot 10^3$ cells/well were seeded into coated 96-well plates in 100 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS and incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, The media was removed and 60 µL antibiotic free cell culture medium (DMEM) supplemented with 10 % FBS, 1 µM GalNAc-FITC (**15**) was added and the cells were incubated for 24 h at 37 °C and 5 % CO_2. After 24 h, the cells were washed with PBS (3 x 100 µL) and the nuclei were stained for 15 min at 37 °C and 5 % CO_2 using NucBlue™ Live ReadyProbes™ Reagent (1:100) in 40 µL PBS. The cells were washed carefully with PBS (2 x 100 µL) and the cells were maintained in 40 µL phenol red free DMEM, supplemented with 10 % FBS, for fluorescence microscopy.

7. References

(1) Bass, B. L.; Weintraub, H. A Developmentally Regulated Activity That Unwinds RNA Duplexes. *Cell* **1987**, *48* (4), 607–613. https://doi.org/10.1016/0092-8674(87)90239-X.

(2) Bass, B. L.; Weintraub, H. An Unwinding Activity That Covalently Modifies Its Double-Stranded RNA Substrate. *Cell* **1988**, *55* (6), 1089–1098. https://doi.org/10.1016/0092-8674(88)90253-X.

(3) Wagner, R. W.; Smith, J. E.; Cooperman, B. S.; Nishikura, K. A Double-Stranded RNA Unwinding Activity Introduces Structural Alterations by Means of Adenosine to Inosine Conversions in Mammalian Cells and Xenopus Eggs. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (8), 2647–2651. https://doi.org/10.1073/pnas.86.8.2647.

(4) Bass, B. L. RNA Editing by Adenosine Deaminases That Act on RNA. *Annual Review of Biochemistry*. NIH Public Access 2002, pp 817–846. https://doi.org/10.1146/annurev.biochem.71.110601.135501.

(5) Nishikura, K. Functions and Regulation of RNA Editing by ADAR Deaminases. *Annual Review of Biochemistry*. NIH Public Access July 7, 2010, pp 321–349. https://doi.org/10.1146/annurev-biochem-060208-105251.

(6) Colina, C.; Palavicini, J. P.; Srikumar, D.; Holmgren, M.; Rosenthal, J. J. C. Regulation of Na+/K+ ATPase Transport Velocity by RNA Editing. *PLoS Biol.* **2010**, *8* (11), e1000540. https://doi.org/10.1371/journal.pbio.1000540.

(7) Yeo, J.; Goodman, R. A.; Schirle, N. T.; David, S. S.; Beal, P. A. RNA Editing Changes the Lesion Specificity for the DNA Repair Enzyme NEIL1. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (48), 20715–20719. https://doi.org/10.1073/pnas.1009231107.

(8) Rueter, S. M.; Dawson, T. R.; Emeson, R. B. Regulation of Alternative Splicing by RNA Editing. *Nature* **1999**, *399* (6731), 75–80. https://doi.org/10.1038/19992.

(9) Wang, Q.; Hui, H.; Guo, Z.; Zhang, W.; Hu, Y.; He, T.; Tai, Y.; Peng, P.; Wang, L. ADAR1 Regulates ARHGAP26 Gene Expression through RNA Editing by Disrupting MiR-30b-3p and MiR-573 Binding. *RNA* **2013**, *19* (11), 1525–1536. https://doi.org/10.1261/rna.041533.113.

(10) Nishikura, K. A-to-I Editing of Coding and Non-Coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **2016**, *17* (2), 83–96. https://doi.org/10.1038/nrm.2015.4.

(11) Piazzi, M.; Bavelloni, A.; Gallo, A.; Blalock, W. L. AKT-Dependent Phosphorylation of ADAR1p110 and ADAR2 Represents a New and Important Link between Cell Signaling and RNA Editing. *DNA Cell Biol.* **2020**, *39* (3), 343–348. https://doi.org/10.1089/dna.2020.5351.

(12) Smith, H. C.; Bennett, R. P.; Kizilyer, A.; McDougall, W. M.; Prohaska, K. M. Functions and Regulation of the APOBEC Family of Proteins. *Seminars in Cell and Developmental Biology*. Elsevier Ltd 2012, pp 258–268. https://doi.org/10.1016/j.semcdb.2011.10.004.

(13) Auxilien, S.; Crain, P. F.; Trewyn, R. W.; Grosjean, H. Mechanism, Specificity and General Properties of the Yeast Enzyme Catalysing the Formation of Inosine 34 in the Anticodon of Transfer RNA. *J. Mol. Biol.* **1996**, *262* (4), 437–458. https://doi.org/10.1006/jmbi.1996.0527.

(14) Crick, F. H. C. Codon-Anticodon Pairing: The Wobble Hypothesis. *J. Mol. Biol.* **1966**, *19* (2), 548–555. https://doi.org/10.1016/S0022-2836(66)80022-0.

(15) Kim, U.; Wang, Y.; Sanford, T.; Zeng, Y.; Nishikura, K. Molecular Cloning of CDNA for Double-Stranded RNA Adenosine Deaminase, a Candidate Enzyme for Nuclear RNA Editing. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (24), 11457–11461. https://doi.org/10.1073/pnas.91.24.11457.

(16) Melcher, T.; Maas, S.; Herb, A.; Sprengel, R.; Seeburg, P. H.; Higuchi, M. A Mammalian RNA Editing Enzyme. *Nature* **1996**, *379* (6564), 460–464. https://doi.org/10.1038/379460a0.

(17) Chen, C. X.; Cho, D. S. C.; Wang, Q.; Lai, F.; Carter, K. C.; Nishikura, K. A Third Member of the RNA-Specific Adenosine Deaminase Gene Family, ADAR3, Contains Both Single- and Double-Stranded RNA Binding Domains. *RNA* **2000**, *6* (5), 755–767. https://doi.org/10.1017/S1355838200000170.

(18) Melcher, T.; Maas, S.; Herb, A.; Sprengel, R.; Higuchi, M.; Seeburg, P. H. RED2, a Brain-Specific Member of the RNA-Specific Adenosine Deaminase Family. *J. Biol. Chem.* **1996**, *271* (50), 31795–31798. https://doi.org/10.1074/jbc.271.50.31795.

(19) Lykke-Andersen, S.; Piñol-Roma, S.; Kjems, J. Alternative Splicing of the ADAR1 Transcript in a Region That Functions Either as a 5′-UTR or an ORF. *RNA* **2007**, *13* (10), 1732–1744. https://doi.org/10.1261/rna.567807.

(20) Herbert, A.; Alfken, J.; Kim, Y. G.; Mian, I. S.; Nishikura, K.; Rich, A. A Z-DNA Binding Domain Present in the Human Editing Enzyme, Double-Stranded RNA Adenosine Deaminase. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (16), 8421–8426. https://doi.org/10.1073/pnas.94.16.8421.

(21) Desterro, J. M. P.; Keegan, L. P.; Lafarga, M.; Berciano, M. T.; O'Connell, M.; Carmo-Fonseca, M. Dynamic Association of RNA-Editing Enzymes with the Nucleolus. *J. Cell Sci.* **2003**, *116* (9), 1805–1818. https://doi.org/10.1242/jcs.00371.

(22) Barraud, P.; Banerjee, S.; Mohamed, W. I.; Jantsch, M. F.; Allain, F. H. T. A Bimodular Nuclear Localization Signal Assembled via an Extended Double-Stranded RNA-Binding Domain Acts as an RNA-Sensing Signal for Transportin 1. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (18), E1852–E1861. https://doi.org/10.1073/pnas.1323698111.

(23) Fritz, J.; Strehblow, ‡ Alexander; Taschner, A.; Schopoff, S.; Pasierbek, P.; Jantsch, M. F. RNA-Regulated Interaction of Transportin-1 and Exportin-5 with the Double-Stranded RNA-Binding Domain Regulates Nucleocytoplasmic Shuttling of ADAR1 †. *Mol. Cell. Biol.* **2009**, *29* (6), 1487–1497. https://doi.org/10.1128/MCB.01519-08.

(24) Poulsen, H.; Nilsson, J.; Damgaard, C. K.; Egebjerg, J.; Kjems, J. CRM1 Mediates the Export of ADAR1 through a Nuclear Export Signal within the Z-DNA Binding Domain. *Mol. Cell. Biol.* **2001**, *21* (22), 7862–7871. https://doi.org/10.1128/MCB.21.22.7862-7871.2001.

(25) Maas, S.; Gommans, W. M. Identification of a Selective Nuclear Import Signal in Adenosine Deaminases Acting on RNA. *Nucleic Acids Res.* **2009**, *37* (17), 5822–5829. https://doi.org/10.1093/nar/gkp599.

(26) Fagerberg, L.; Hallstrom, B. M.; Oksvold, P.; Kampf, C.; Djureinovic, D.; Odeberg, J.; Habuka, M.; Tahmasebpoor, S.; Danielsson, A.; Edlund, K.; Asplund, A.; Sjostedt, E.; Lundberg, E.; Szigyarto, C. A. K.; Skogs, M.; Ottosson Takanen, J.; Berling, H.; Tegel, H.; Mulder, J.; Nilsson, P.; Schwenk, J. M.; Lindskog, C.; Danielsson, F.; Mardinoglu, A.; Sivertsson, A.; Von Feilitzen, K.; Forsberg, M.; Zwahlen, M.; Olsson, I.; Navani, S.; Huss, M.; Nielsen, J.; Ponten, F.; Uhlen, M. Analysis of the Human Tissue-Specific Expression by Genome-Wide Integration of Transcriptomics and Antibody-Based Proteomics. *Mol. Cell. Proteomics* **2014**, *13* (2), 397–406. https://doi.org/10.1074/mcp.M113.035600.

(27) Matthews, M. M.; Thomas, J. M.; Zheng, Y.; Tran, K.; Phelps, K. J.; Scott, A. I.; Havel, J.; Fisher, A. J.; Beal, P. A. Structures of Human ADAR2 Bound to DsRNA Reveal Base-Flipping Mechanism and Basis for Site Selectivity. *Nat. Struct. Mol. Biol.* **2016**, *23* (5), 426–433. https://doi.org/10.1038/nsmb.3203.

(28) Macbeth, M. R.; Schubert, H. L.; VanDemark, A. F.; Lingam, A. T.; Hill, C. P.; Bass, B. L. Structural Biology: Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science (80-.).* **2005**, *309* (5740), 1534–1539. https://doi.org/10.1126/science.1113150.

(29) Lai, F.; Drakas, R.; Nishikura, K. Mutagenic Analysis of Double-Stranded RNA Adenosine Deaminase, a Candidate Enzyme for RNA Editing of Glutamate-Gated Ion Channel Transcripts. *J. Biol. Chem.* **1995**, *270* (29), 17098–17105. https://doi.org/10.1074/jbc.270.29.17098.

(30) Cho, D. S. C.; Yang, W.; Lee, J. T.; Shiekhattar, R.; Murray, J. M.; Nishikura, K. Requirement of Dimerization for RNA Editing Activity of Adenosine Deaminases Acting on RNA. *J. Biol. Chem.* **2003**, *278* (19), 17093–17102. https://doi.org/10.1074/jbc.M213127200.

(31) Wang, Y.; Beal, P. A. Probing RNA Recognition by Human ADAR2 Using a High-Throughput Mutagenesis Method. *Nucleic Acids Res.* **2016**, *44* (20), 9872–9880. https://doi.org/10.1093/nar/gkw799.

(32) Thuy-Boun, A. S.; Thomas, J. M.; Grajo, H. L.; Palumbo, C. M.; Park, S.; Nguyen, L. T.; Fisher, A. J.; Beal, P. A. Asymmetric Dimerization of Adenosine Deaminase Acting on RNA Facilitates Substrate Recognition. *Nucleic Acids Res.* **2020**, *48* (14), 7958–7972. https://doi.org/10.1093/nar/gkaa532.

(33) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 1–9. https://doi.org/10.1038/ncomms1324.

(34) Wong, S. K.; Sato, S.; Lazinski, D. W. Substrate Recognition by ADAR1 and ADAR2. *RNA* **2001**, *7* (6), 846–858. https://doi.org/10.1017/S135583820101007X.

(35) Kuttan, A.; Bass, B. L. Mechanistic Insights into Editing-Site Specificity of ADARs. *Proc. Natl. Acad. Sci.* **2012**, *109* (48), E3295–E3304. https://doi.org/10.1073/pnas.1212548109.

(36) Ota, H.; Sakurai, M.; Gupta, R.; Valente, L.; Wulff, B. E.; Ariyoshi, K.; Iizasa, H.; Davuluri, R. V.; Nishikura, K. ADAR1 Forms a Complex with Dicer to Promote MicroRNA Processing and RNA-Induced Gene Silencing. *Cell* **2013**, *153* (3), 575–589. https://doi.org/10.1016/j.cell.2013.03.024.

(37) Gallo, A.; Keegan, L. P.; Ring, G. M.; O'Connell, M. A. An ADAR That Edits Transcripts Encoding Ion Channel Subunits Functions as a Dimer. *EMBO J.* **2003**, *22* (13), 3421–3430. https://doi.org/10.1093/emboj/cdg327.

(38) Jaikaran, D. C. J.; Collins, C. H.; MacMillan, A. M. Adenosine to Inosine Editing by ADAR2 Requires Formation of a Ternary Complex on the GluR-B R/G Site. *J. Biol. Chem.* **2002**, *277* (40), 37624–37629. https://doi.org/10.1074/jbc.M204126200.

(39) Chilibeck, K. A.; Wu, T.; Liang, C.; Schellenberg, M. J.; Gesner, E. M.; Lynch, J. M.; MacMillan, A. M. FRET Analysis of in Vivo Dimerization by RNA-Editing Enzymes. *J. Biol. Chem.* **2006**, *281* (24), 16530–16535. https://doi.org/10.1074/jbc.M511831200.

(40) Valente, L.; Nishikura, K. RNA Binding-Independent Dimerization of Adenosine Deaminases Acting on RNA and Dominant Negative Effects of Nonfunctional Subunits on Dimer Functions. *J. Biol. Chem.* **2007**, *282* (22), 16054–16061. https://doi.org/10.1074/jbc.M611392200.

(41) Poulsen, H.; Jorgensen, R.; Heding, A.; Nielsen, F. C.; Bonven, B.; Egebjerg, J. Dimerization of ADAR2 Is Mediated by the Double-Stranded RNA Binding Domain. *RNA* **2006**, *12* (7), 1350–1360. https://doi.org/10.1261/rna.2314406.

(42) Macbeth, M. R.; Bass, B. L. Large-Scale Overexpression and Purification of ADARs from Saccharomyces Cerevisiae for Biophysical and Biochemical Studies. In *Methods in Enzymology*; Academic Press Inc., 2007; Vol. 424, pp 319–331. https://doi.org/10.1016/S0076-6879(07)24015-7.

(43) Lazzaretti, D.; Bandholz-Cajamarca, L.; Emmerich, C.; Schaaf, K.; Basquin, C.; Irion, U.; Bono, F. The Crystal Structure of Staufen1 in Complex with a Physiological RNA Sheds Light on Substrate Selectivity. *Life Sci. Alliance* **2018**, *1* (5). https://doi.org/10.26508/lsa.201800187.

(44) Fu, Q.; Yuan, Y. A. Structural Insights into RISC Assembly Facilitated by DsRNA-Binding Domains of Human RNA Helicase A (DHX9). *Nucleic Acids Res.* **2013**, *41* (5),

3457–3470. https://doi.org/10.1093/nar/gkt042.

(45) Gan, J.; Tropea, J. E.; Austin, B. P.; Court, D. L.; Waugh, D. S.; Ji, X. Structural Insight into the Mechanism of Double-Stranded RNA Processing by Ribonuclease III. *Cell* **2006**, *124* (2), 355–366. https://doi.org/10.1016/j.cell.2005.11.034.

(46) Stefl, R.; Oberstrass, F. C.; Hood, J. L.; Jourdan, M.; Zimmermann, M.; Skrisovska, L.; Maris, C.; Peng, L.; Hofr, C.; Emeson, R. B.; Allain, F. H. T. The Solution Structure of the ADAR2 DsRBM-RNA Complex Reveals a Sequence-Specific Readout of the Minor Groove. *Cell* **2010**, *143* (2), 225–237. https://doi.org/10.1016/j.cell.2010.09.026.

(47) Oakes, E.; Anderson, A.; Cohen-Gadol, A.; Hundley, H. A. Adenosine Deaminase That Acts on RNA 3 (Adar3) Binding to Glutamate Receptor Subunit B Pre-MRNA Inhibits RNA Editing in Glioblastoma. *J. Biol. Chem.* **2017**, *292* (10), 4326–4335. https://doi.org/10.1074/jbc.M117.779868.

(48) Schneider, M. F.; Wettengel, J.; Hoffmann, P. C.; Stafforst, T. Optimal GuideRNAs for Re-Directing Deaminase Activity of HADAR1 and HADAR2 in Trans. *Nucleic Acids Res.* **2014**, *42* (10), 1–9. https://doi.org/10.1093/nar/gku272.

(49) Shevchenko, G.; Morris, K. V. All I's on the RADAR: Role of ADAR in Gene Regulation. *FEBS Letters*. Wiley Blackwell September 1, 2018, pp 2860–2873. https://doi.org/10.1002/1873-3468.13093.

(50) Lomeli, H.; Mosbacher, J.; Melcher, T.; Höger, T.; Geiger, J. R. P.; Kuner, T.; Monyer, H.; Higuchi, M.; Bach, A.; Seeburg, P. H. Control of Kinetic Properties of AMPA Receptor Channels by Nuclear RNA Editing. *Science (80-.).* **1994**, *266* (5191), 1709–1713. https://doi.org/10.1126/science.7992055.

(51) Higuchi, M.; Single, F. N.; Köhler, M.; Sommer, B.; Sprengel, R.; Seeburg, P. H. RNA Editing of AMPA Receptor Subunit GluR-B: A Base-Paired Intron-Exon Structure Determines Position and Efficiency. *Cell* **1993**, *75* (7), 1361–1370. https://doi.org/10.1016/0092-8674(93)90622-W.

(52) Sommer, B.; Köhler, M.; Sprengel, R.; Seeburg, P. H. RNA Editing in Brain Controls a Determinant of Ion Flow in Glutamate-Gated Channels. *Cell* **1991**, *67* (1), 11–19. https://doi.org/10.1016/0092-8674(91)90568-J.

(53) Kawahara, Y.; Ito, K.; Sun, H.; Aizawa, H.; Kanazawa, I.; Kwak, S. RNA Editing and

Death of Motor Neurons: There Is a Glutamate-Receptor Defect in Patients with Amyotrophic Lateral Sclerosis. *Nature* **2004**, *427* (6977), 801. https://doi.org/10.1038/427801a.

(54) Whitney, N. P.; Peng, H.; Erdmann, N. B.; Tian, C.; Monaghan, D. T.; Zheng, J. C. Calcium-permeable AMPA Receptors Containing Q/R-unedited GluR2 Direct Human Neural Progenitor Cell Differentiation to Neurons. *FASEB J.* **2008**, *22* (8), 2888–2900. https://doi.org/10.1096/fj.07-104661.

(55) Hideyama, T.; Yamashita, T.; Suzuki, T.; Tsuji, S.; Higuchi, M.; Seeburg, P. H.; Takahashi, R.; Misawa, H.; Kwak, S. Induced Loss of ADAR2 Engenders Slow Death of Motor Neurons from Q/R Site-Unedited GluR2. *J. Neurosci.* **2010**, *30* (36), 11917–11925. https://doi.org/10.1523/JNEUROSCI.2021-10.2010.

(56) Burns, C. M.; Chu, H.; Rueter, S. M.; Hutchinson, L. K.; Canton, H.; Sanders-Bush, E.; Emeson, R. B. Regulation of Serotonin-2C Receptor G-Protein Coupling by RNA Editing. *Nature* **1997**, *387* (6630), 303–308. https://doi.org/10.1038/387303a0.

(57) Bhalla, T.; Rosenthal, J. J. C.; Holmgren, M.; Reenan, R. Control of Human Potassium Channel Inactivation by Editing of a Small MRNA Hairpin. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 950–956. https://doi.org/10.1038/nsmb825.

(58) Daniel, C.; Wahlstedt, H.; Ohlson, J.; Björk, P.; Öhman, M. Adenosine-to-Inosine RNA Editing Affects Trafficking of the γ-Aminobutyric Acid Type A (GABAA) Receptor. *J. Biol. Chem.* **2011**, *286* (3), 2031–2040. https://doi.org/10.1074/jbc.M110.130096.

(59) Ohlson, J.; Pedersen, J. S.; Haussler, D.; Öhman, M. Editing Modifies the GABAA Receptor Subunit A3. *RNA* **2007**, *13* (5), 698–703. https://doi.org/10.1261/rna.349107.

(60) Niswender, C. M.; Sanders-Bush, E.; Emeson, R. B. Identification and Characterization of RNA Editing Events within the 5-HT(2C) Receptor. In *Annals of the New York Academy of Sciences*; New York Academy of Sciences, 1998; Vol. 861, pp 38–48. https://doi.org/10.1111/j.1749-6632.1998.tb10171.x.

(61) Tan, M. H.; Li, Q.; Shanmugam, R.; Nishikura, K.; Li, J. B. Dynamic Landscape and Regulation of RNA Editing in Mammals. *Nature* **2017**, *550* (7675), 249–254. https://doi.org/10.1038/nature24041.

(62) Cordaux, R.; Batzer, M. A. The Impact of Retrotransposons on Human Genome

Evolution. *Nature Reviews Genetics*. NIH Public Access October 2009, pp 691–703. https://doi.org/10.1038/nrg2640.

(63) Tajaddod, M.; Jantsch, M. F.; Licht, K. The Dynamic Epitranscriptome: A to I Editing Modulates Genetic Information. *Chromosoma*. Springer Science and Business Media Deutschland GmbH March 1, 2016, pp 51–63. https://doi.org/10.1007/s00412-015-0526-9.

(64) Liddicoat, B. J.; Piskol, R.; Chalk, A. M.; Ramaswami, G.; Higuchi, M.; Hartner, J. C.; Li, J. B.; Seeburg, P. H.; Walkley, C. R. RNA Editing by ADAR1 Prevents MDA5 Sensing of Endogenous DsRNA as Nonself. *Science (80-.).* **2015**, *349* (6252), 1115–1120. https://doi.org/10.1126/science.aac7049.

(65) Mannion, N. M.; Greenwood, S. M.; Young, R.; Cox, S.; Brindle, J.; Read, D.; Nellåker, C.; Vesely, C.; Ponting, C. P.; McLaughlin, P. J.; Jantsch, M. F.; Dorin, J.; Adams, I. R.; Scadden, A. D. J.; Öhman, M.; Keegan, L. P.; O'Connell, M. A. The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA. *Cell Rep.* **2014**, *9* (4), 1482–1494. https://doi.org/10.1016/j.celrep.2014.10.041.

(66) Ahmad, S.; Mu, X.; Yang, F.; Greenwald, E.; Park, J. W.; Jacob, E.; Zhang, C. Z.; Hur, S. Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation. *Cell* **2018**, *172* (4), 797-810.e13. https://doi.org/10.1016/j.cell.2017.12.016.

(67) Chung, H.; Calis, J. J. A.; Wu, X.; Sun, T.; Yu, Y.; Sarbanes, S. L.; Dao Thi, V. L.; Shilvock, A. R.; Hoffmann, H. H.; Rosenberg, B. R.; Rice, C. M. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. *Cell* **2018**, *172* (4), 811-824.e14. https://doi.org/10.1016/j.cell.2017.12.038.

(68) Wang, Q.; Miyakoda, M.; Yang, W.; Khillan, J.; Stachura, D. L.; Weiss, M. J.; Nishikura, K. Stress-Induced Apoptosis Associated with Null Mutation of ADAR1 RNA Editing Deaminase Gene. *J. Biol. Chem.* **2004**, *279* (6), 4952–4961. https://doi.org/10.1074/jbc.M310162200.

(69) Hartner, J. C.; Schmittwolf, C.; Kispert, A.; Müller, A. M.; Higuchi, M.; Seeburg, P. H. Liver Disintegration in the Mouse Embryo Caused by Deficiency in the RNA-Editing Enzyme ADAR1. *J. Biol. Chem.* **2004**, *279* (6), 4894–4902.

https://doi.org/10.1074/jbc.M311347200.

(70) Pestal, K.; Funk, C. C.; Snyder, J. M.; Price, N. D.; Treuting, P. M.; Stetson, D. B. Isoforms of RNA-Editing Enzyme ADAR1 Independently Control Nucleic Acid Sensor MDA5-Driven Autoimmunity and Multi-Organ Development. *Immunity* **2015**, *43* (5), 933–944. https://doi.org/10.1016/j.immuni.2015.11.001.

(71) Rice, G. I.; Kasher, P. R.; Forte, G. M. A.; Lovell, S. C.; Crow, Y. J. Mutations in ADAR1 Cause Aicardi-Goutières Syndrome Associated with a Type i Interferon Signature. *Nat. Genet.* **2012**, *44* (11), 1243–1248. https://doi.org/10.1038/ng.2414.

(72) Lev-Maor, G.; Sorek, R.; Levanon, E. Y.; Paz, N.; Eisenberg, E.; Ast, G. RNA-Editing-Mediated Exon Evolution. *Genome Biol.* **2007**, *8* (2), R29. https://doi.org/10.1186/gb-2007-8-2-r29.

(73) Athanasiadis, A.; Rich, A.; Maas, S. Widespread A-to-I RNA Editing of Alu-Containing MRNAs in the Human Transcriptome. *PLoS Biol.* **2004**, *2* (12), e391. https://doi.org/10.1371/journal.pbio.0020391.

(74) Feng, Y.; Sansam, C. L.; Singh, M.; Emeson, R. B. Altered RNA Editing in Mice Lacking ADAR2 Autoregulation. *Mol. Cell. Biol.* **2006**, *26* (2), 480–488. https://doi.org/10.1128/mcb.26.2.480-488.2006.

(75) Scadden, A. D. J.; Smith, C. W. J. Specific Cleavage of Hyper-Edited DsRNAs. *EMBO J.* **2001**, *20* (15), 4243–4252. https://doi.org/10.1093/emboj/20.15.4243.

(76) Scadden, A. D. J. The RISC Subunit Tudor-SN Binds to Hyper-Edited Double-Stranded RNA and Promotes Its Cleavage. *Nat. Struct. Mol. Biol.* **2005**, *12* (6), 489–496. https://doi.org/10.1038/nsmb936.

(77) Ha, M.; Kim, V. N. Regulation of MicroRNA Biogenesis. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group July 16, 2014, pp 509–524. https://doi.org/10.1038/nrm3838.

(78) Mendell, J. T.; Olson, E. N. MicroRNAs in Stress Signaling and Human Disease. *Cell*. Cell March 16, 2012, pp 1172–1187. https://doi.org/10.1016/j.cell.2012.02.005.

(79) Croce, C. M.; Calin, G. A. MiRNAs, Cancer, and Stem Cell Division. *Cell*. Cell July 15, 2005, pp 6–7. https://doi.org/10.1016/j.cell.2005.06.036.

(80) Kawahara, Y.; Zinshteyn, B.; Sethupathy, P.; Iizasa, H.; Hatzigeorgiou, A. G.; Nishikura, K. Redirection of Silencing Targets by Adenosine-to-Inosine Editing of MiRNAs. *Science (80-.).* **2007**, *315* (5815), 1137–1140. https://doi.org/10.1126/science.1138050.

(81) Gandy, S. Z.; Linnstaedt, S. D.; Muralidhar, S.; Cashman, K. A.; Rosenthal, L. J.; Casey, J. L. RNA Editing of the Human Herpesvirus 8 Kaposin Transcript Eliminates Its Transforming Activity and Is Induced during Lytic Replication. *J. Virol.* **2007**, *81* (24), 13544–13551. https://doi.org/10.1128/jvi.01521-07.

(82) Yang, W.; Chendrimada, T. P.; Wang, Q.; Higuchi, M.; Seeburg, P. H.; Shiekhattar, R.; Nishikura, K. Modulation of MicroRNA Processing and Expression through RNA Editing by ADAR Deaminases. *Nat. Struct. Mol. Biol.* **2006**, *13* (1), 13–21. https://doi.org/10.1038/nsmb1041.

(83) Kawahara, Y.; Megraw, M.; Kreider, E.; Iizasa, H.; Valente, L.; Hatzigeorgiou, A. G.; Nishikura, K. Frequency and Fate of MicroRNA Editing in Human Brain. *Nucleic Acids Res.* **2008**, *36* (16), 5270–5280. https://doi.org/10.1093/nar/gkn479.

(84) Iizasa, H.; Wulff, B. E.; Alla, N. R.; Maragkakis, M.; Megraw, M.; Hatzigeorgiou, A.; Iwakiri, D.; Takada, K.; Wiedmer, A.; Showe, L.; Lieberman, P.; Nishikura, K. Editing of Epstein-Barr Virus-Encoded BART6 MicroRNAs Controls Their Dicer Targeting and Consequently Affects Viral Latency. *J. Biol. Chem.* **2010**, *285* (43), 33358–33370. https://doi.org/10.1074/jbc.M110.138362.

(85) Bass, B. L. Double-Stranded RNA as a Template for Gene Silencing. *Cell*. Cell Press April 28, 2000, pp 235–238. https://doi.org/10.1016/S0092-8674(02)71133-1.

(86) Nishikura, K.; Sakurai, M.; Ariyoshi, K.; Ota, H. Antagonistic and Stimulative Roles of ADAR1 in RNA Silencing: An Editor's Point-of-View. *RNA Biology*. Taylor and Francis Inc. 2013, pp 1240–1247. https://doi.org/10.4161/rna.25947.

(87) Ma, E.; MacRae, I. J.; Kirsch, J. F.; Doudna, J. A. Autoinhibition of Human Dicer by Its Internal Helicase Domain. *J. Mol. Biol.* **2008**, *380* (1), 237–243. https://doi.org/10.1016/j.jmb.2008.05.005.

(88) Ji, X.; Dadon, D. B.; Abraham, B. J.; Lee, T. I.; Jaenisch, R.; Bradner, J. E.; Young, R. A. Chromatin Proteomic Profiling Reveals Novel Proteins Associated with Histone-Marked Genomic Regions. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (12), 3841–3846.

https://doi.org/10.1073/pnas.1502971112.

(89) Soldi, M.; Bonaldi, T. The Proteomic Investigation of Chromatin Functional Domains Reveals Novel Synergisms among Distinct Heterochromatin Components. *Mol. Cell. Proteomics* **2013**, *12* (3), 764–780. https://doi.org/10.1074/mcp.M112.024307.

(90) Cao, Q.; Wang, X.; Zhao, M.; Yang, R.; Malik, R.; Qiao, Y.; Poliakov, A.; Yocum, A. K.; Li, Y.; Chen, W.; Cao, X.; Jiang, X.; Dahiya, A.; Harris, C.; Feng, F. Y.; Kalantry, S.; Qin, Z. S.; Dhanasekaran, S. M.; Chinnaiyan, A. M. The Central Role of EED in the Orchestration of Polycomb Group Complexes. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4127.

(91) Lehmann, L.; Ferrari, R.; Vashisht, A. A.; Wohlschlegel, J. A.; Kurdistani, S. K.; Careys, M. Polycomb Repressive Complex 1 (PRC1) Disassembles RNA Polymerase II Preinitiation Complexes. *J. Biol. Chem.* **2012**, *287* (43), 35784–35794. https://doi.org/10.1074/jbc.M112.397430.

(92) Zhou, J.; Wang, Q.; Chen, L. L.; Carmichael, G. G. On the Mechanism of Induction of Heterochromatin by the RNA-Binding Protein Vigilin. *RNA* **2008**, *14* (9), 1773–1781. https://doi.org/10.1261/rna.1036308.

(93) Wang, Q.; Zhang, Z.; Blackwell, K.; Carmichael, G. G. Vigilins Bind to Promiscuously A-to-I-Edited RNAs and Are Involved in the Formation of Heterochromatin. *Curr. Biol.* **2005**, *15* (4), 384–391. https://doi.org/10.1016/j.cub.2005.01.046.

(94) Kondo, Y.; Issa, J. P. J. Enrichment for Histone H3 Lysine 9 Methylation at Alu Repeats in Human Cells. *J. Biol. Chem.* **2003**, *278* (30), 27658–27662. https://doi.org/10.1074/jbc.M304072200.

(95) Sakurai, M.; Shiromoto, Y.; Ota, H.; Song, C.; Kossenkov, A. V.; Wickramasinghe, J.; Showe, L. C.; Skordalakes, E.; Tang, H. Y.; Speicher, D. W.; Nishikura, K. ADAR1 Controls Apoptosis of Stressed Cells by Inhibiting Staufen1-Mediated MRNA Decay. *Nat. Struct. Mol. Biol.* **2017**, *24* (6), 534–543. https://doi.org/10.1038/nsmb.3403.

(96) Anantharaman, A.; Tripathi, V.; Khan, A.; Yoon, J. H.; Singh, D. K.; Gholamalamdari, O.; Guang, S.; Ohlson, J.; Wahlstedt, H.; Öhman, M.; Jantsch, M. F.; Conrad, N. K.; Ma, J.; Gorospe, M.; Prasanth, S. G.; Prasanth, K. V. ADAR2 Regulates RNA Stability by Modifying Access of Decay-Promoting RNA-Binding Proteins. *Nucleic Acids Res.*

2017, *45* (7), 4189–4201. https://doi.org/10.1093/nar/gkw1304.

(97) Rees, H. A.; Liu, D. R. Base Editing: Precision Chemistry on the Genome and Transcriptome of Living Cells. *Nat. Rev. Genet.* **2018**, *19* (12), 770–788. https://doi.org/10.1038/s41576-018-0059-1.

(98) Zhang, F. Development of CRISPR-Cas Systems for Genome Editing and Beyond. *Q. Rev. Biophys.* **2019**, *52*. https://doi.org/10.1017/s0033583519000052.

(99) Van Der Oost, J.; Westra, E. R.; Jackson, R. N.; Wiedenheft, B. Unravelling the Structural and Mechanistic Basis of CRISPR-Cas Systems. *Nature Reviews Microbiology*. Nature Publishing Group 2014, pp 479–492. https://doi.org/10.1038/nrmicro3279.

(100) Makarova, K. S.; Wolf, Y. I.; Alkhnbashi, O. S.; Costa, F.; Shah, S. A.; Saunders, S. J.; Barrangou, R.; Brouns, S. J. J.; Charpentier, E.; Haft, D. H.; Horvath, P.; Moineau, S.; Mojica, F. J. M.; Terns, R. M.; Terns, M. P.; White, M. F.; Yakunin, A. F.; Garrett, R. A.; Van Der Oost, J.; Backofen, R.; Koonin, E. V. An Updated Evolutionary Classification of CRISPR-Cas Systems. *Nat. Rev. Microbiol.* **2015**, *13* (11), 722–736. https://doi.org/10.1038/nrmicro3569.

(101) Ceccaldi, R.; Rondinelli, B.; D'Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in Cell Biology*. Elsevier Ltd January 1, 2016, pp 52–64. https://doi.org/10.1016/j.tcb.2015.07.009.

(102) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **2012**, *337* (6096), 816–821. https://doi.org/10.1126/science.1225829.

(103) Nishimasu, H.; Ran, F. A.; Hsu, P. D.; Konermann, S.; Shehata, S. I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **2014**, *156* (5), 935–949. https://doi.org/10.1016/j.cell.2014.02.001.

(104) Paquet, D.; Kwart, D.; Chen, A.; Sproul, A.; Jacob, S.; Teo, S.; Olsen, K. M.; Gregg, A.; Noggle, S.; Tessier-Lavigne, M. Efficient Introduction of Specific Homozygous and Heterozygous Mutations Using CRISPR/Cas9. *Nature* **2016**, *533* (7601), 125–129. https://doi.org/10.1038/nature17664.

(105) Lin, S.; Staahl, B. T.; Alla, R. K.; Doudna, J. A. Enhanced Homology-Directed Human Genome Engineering by Controlled Timing of CRISPR/Cas9 Delivery. *Elife* **2014**, *3*, e04766. https://doi.org/10.7554/eLife.04766.

(106) Komor, A. C.; Kim, Y. B.; Packer, M. S.; Zuris, J. A.; Liu, D. R. Programmable Editing of a Target Base in Genomic DNA without Double-Stranded DNA Cleavage. *Nature* **2016**, *533* (7603), 420–424. https://doi.org/10.1038/nature17946.

(107) Gaudelli, N. M.; Komor, A. C.; Rees, H. A.; Packer, M. S.; Badran, A. H.; Bryson, D. I.; Liu, D. R. Programmable Base Editing of T to G C in Genomic DNA without DNA Cleavage. *Nature* **2017**, *551* (7681), 464–471. https://doi.org/10.1038/nature24644.

(108) Losey, H. C.; Ruthenburg, A. J.; Verdine, G. L. Crystal Structure of Staphylococcus Aureus TRNA Adenosine Deaminase TadA in Complex with RNA. *Nat. Struct. Mol. Biol.* **2006**, *13* (2), 153–159. https://doi.org/10.1038/nsmb1047.

(109) Kim, D.; Lim, K.; Kim, S. T.; Yoon, S. H.; Kim, K.; Ryu, S. M.; Kim, J. S. Genome-Wide Target Specificities of CRISPR RNA-Guided Programmable Deaminases. *Nat. Biotechnol.* **2017**, *35* (5), 475–480. https://doi.org/10.1038/nbt.3852.

(110) Woolf, T. M.; Chase, J. M.; Stinchcomb, D. T. Toward the Therapeutic Editing of Mutated RNA Sequences. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (18), 8298–8302. https://doi.org/10.1073/pnas.92.18.8298.

(111) Fry, L. E.; Peddle, C. F.; Barnard, A. R.; McClements, M. E.; Maclaren, R. E. RNA Editing as a Therapeutic Approach for Retinal Gene Therapy Requiring Long Coding Sequences. *Int. J. Mol. Sci.* **2020**, *21* (3). https://doi.org/10.3390/ijms21030777.

(112) Stafforst, T.; Schneider, M. F. An RNA-Deaminase Conjugate Selectively Repairs Point Mutations. *Angew. Chemie - Int. Ed.* **2012**, *51* (44), 11166–11169. https://doi.org/10.1002/anie.201206489.

(113) Vogel, P.; Stafforst, T. Critical Review on Engineering Deaminases for Site-Directed RNA Editing. *Current Opinion in Biotechnology*. Elsevier Ltd February 1, 2019, pp 74–80. https://doi.org/10.1016/j.copbio.2018.08.006.

(114) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A General Method for the Covalent Labeling of Fusion Proteins with Small Molecules in Vivo. *Nat. Biotechnol.* **2003**, *21* (1), 86–89. https://doi.org/10.1038/nbt765.

(115) Gronemeyer, T.; Chidley, C.; Juillerat, A.; Heinis, C.; Johnsson, K. Directed Evolution of O6-Alkylguanine-DNA Alkyltransferase for Applications in Protein Labeling. *Protein Eng. Des. Sel.* **2006**, *19* (7), 309–316. https://doi.org/10.1093/protein/gzl014.

(116) Vogel, P.; Schneider, M. F.; Wettengel, J.; Stafforst, T. Improving Site-Directed RNA Editing in Vitro and in Cell Culture by Chemical Modification of the GuideRNA. *Angew. Chemie - Int. Ed.* **2014**, *53* (24), 6267–6271. https://doi.org/10.1002/anie.201402634.

(117) Hanswillemenke, A.; Kuzdere, T.; Vogel, P.; Jékely, G.; Stafforst, T. Site-Directed RNA Editing in Vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein. *J. Am. Chem. Soc.* **2015**, *137* (50), 15875–15881. https://doi.org/10.1021/jacs.5b10216.

(118) Vogel, P.; Moschref, M.; Li, Q.; Merkle, T.; Selvasaravanan, K. D.; Li, J. B.; Stafforst, T. Efficient and Precise Editing of Endogenous Transcripts with SNAP-Tagged ADARs. *Nat. Methods* **2018**, *15* (July). https://doi.org/10.1038/s41592-018-0017-z.

(119) Stroppel, A. S.; Latifi, N.; Hanswillemenke, A.; Tasakis, R. N.; Papavasiliou, F. N.; Stafforst, T. Harnessing Self-Labeling Enzymes for Selective and Concurrent A-to-I and C-to-U RNA Base Editing. *Nucleic Acids Res.* **2021,** 49 (16), e95, 1-12. https://doi.org/10.1093/NAR/GKAB541.

(120) Vogel, P.; Hanswillemenke, A.; Stafforst, T. Switching Protein Localization by Site-Directed RNA Editing under Control of Light. *ACS Synth. Biol.* **2017**, *6* (9), 1642–1649. https://doi.org/10.1021/acssynbio.7b00113.

(121) Montiel-Gonzalez, M. F.; Vallecillo-Viejo, I.; Yudowski, G. A.; Rosenthal, J. J. C. Correction of Mutations within the Cystic Fibrosis Transmembrane Conductance Regulator by Site-Directed RNA Editing. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (45), 18285–18290. https://doi.org/10.1073/pnas.1306243110.

(122) Chattopadhyay, S.; Garcia-Mena, J.; DeVito, J.; Wolska, K.; Das, A. Bipartite Function of a Small RNA Hairpin in Transcription Antitermination in Bacteriophage λ. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (9), 4061–4065. https://doi.org/10.1073/pnas.92.9.4061.

(123) Montiel-Gonźalez, M. F.; Vallecillo-Viejo, I. C.; Rosenthal, J. J. C. An Efficient System for Selectively Altering Genetic Information within MRNAs. *Nucleic Acids Res.* **2016**, *44* (21). https://doi.org/10.1093/nar/gkw738.

(124) Sinnamon, J. R.; Kim, S. Y.; Corson, G. M.; Song, Z.; Nakai, H.; Adelman, J. P.; Mandel, G. Site-Directed RNA Repair of Endogenous Mecp2 RNA in Neurons. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (44), E9395–E9402. https://doi.org/10.1073/pnas.1715320114.

(125) Azad, M. T. A.; Bhakta, S.; Tsukahara, T. Site-Directed RNA Editing by Adenosine Deaminase Acting on RNA for Correction of the Genetic Code in Gene Therapy. *Gene Ther.* **2017**, *24* (12), 779–786. https://doi.org/10.1038/gt.2017.90.

(126) Shmakov, S.; Smargon, A.; Scott, D.; Cox, D.; Pyzocha, N.; Yan, W.; Abudayyeh, O. O.; Gootenberg, J. S.; Makarova, K. S.; Wolf, Y. I.; Severinov, K.; Zhang, F.; Koonin, E. V. Diversity and Evolution of Class 2 CRISPR-Cas Systems. *Nat. Rev. Microbiol.* **2017**, *15* (3), 169–182. https://doi.org/10.1038/nrmicro.2016.184.

(127) Smargon, A. A.; Cox, D. B. T.; Pyzocha, N. K.; Zheng, K.; Slaymaker, I. M.; Gootenberg, J. S.; Abudayyeh, O. A.; Essletzbichler, P.; Shmakov, S.; Makarova, K. S.; Koonin, E. V.; Zhang, F. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol. Cell* **2017**, *65* (4), 618-630.e7. https://doi.org/10.1016/j.molcel.2016.12.023.

(128) Slaymaker, I. M.; Mesa, P.; Kellner, M. J.; Kannan, S.; Brignole, E.; Koob, J.; Feliciano, P. R.; Stella, S.; Abudayyeh, O. O.; Gootenberg, J. S.; Strecker, J.; Montoya, G.; Zhang, F. High-Resolution Structure of Cas13b and Biochemical Characterization of RNA Targeting and Cleavage. *Cell Rep.* **2019**, *26* (13), 3741-3751.e5. https://doi.org/10.1016/j.celrep.2019.02.094.

(129) Liu, L.; Li, X.; Ma, J.; Li, Z.; You, L.; Wang, J.; Wang, M.; Zhang, X.; Wang, Y. The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell* **2017**, *170* (4), 714-726.e10. https://doi.org/10.1016/j.cell.2017.06.050.

(130) Liu, L.; Li, X.; Wang, J.; Wang, M.; Chen, P.; Yin, M.; Li, J.; Sheng, G.; Wang, Y. Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. *Cell* **2017**, *168* (1–2), 121-134.e12. https://doi.org/10.1016/j.cell.2016.12.031.

(131) Tambe, A.; East-Seletsky, A.; Knott, G. J.; Doudna, J. A.; O'Connell, M. R. RNA Binding and HEPN-Nuclease Activation Are Decoupled in CRISPR-Cas13a. *Cell Rep.* **2018**, *24* (4), 1025–1036. https://doi.org/10.1016/j.celrep.2018.06.105.

(132) Abudayyeh, O. O.; Gootenberg, J. S.; Essletzbichler, P.; Han, S.; Joung, J.; Belanto, J.

J.; Verdine, V.; Cox, D. B. T.; Kellner, M. J.; Regev, A.; Lander, E. S.; Voytas, D. F.; Ting, A. Y.; Zhang, F. RNA Targeting with CRISPR-Cas13. *Nature* **2017**, *550* (7675), 280–284. https://doi.org/10.1038/nature24049.

(133) Konermann, S.; Lotfy, P.; Brideau, N. J.; Oki, J.; Shokhirev, M. N.; Hsu, P. D. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* **2018**, *173* (3), 665-676.e14. https://doi.org/10.1016/j.cell.2018.02.033.

(134) Abudayyeh, O. O.; Gootenberg, J. S.; Konermann, S.; Joung, J.; Slaymaker, I. M.; Cox, D. B. T.; Shmakov, S.; Makarova, K. S.; Semenova, E.; Minakhin, L.; Severinov, K.; Regev, A.; Lander, E. S.; Koonin, E. V.; Zhang, F. C2c2 Is a Single-Component Programmable RNA-Guided RNA-Targeting CRISPR Effector. *Science (80-.).* **2016**, *353* (6299). https://doi.org/10.1126/science.aaf5573.

(135) East-Seletsky, A.; O'Connell, M. R.; Knight, S. C.; Burstein, D.; Cate, J. H. D.; Tjian, R.; Doudna, J. A. Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection. *Nature* **2016**, *538* (7624), 270–273. https://doi.org/10.1038/nature19802.

(136) Cox, D. B. T.; Gootenberg, J. S.; Abudayyeh, O. O.; Franklin, B.; Kellner, M. J.; Joung, J.; Zhang, F. RNA Editing with CRISPR-Cas13. *Science (80-.).* **2017**, *358* (6366), 1019–1027. https://doi.org/10.1126/science.aaq0180.

(137) Rauch, S.; He, E.; Srienc, M.; Zhou, H.; Zhang, Z.; Dickinson, B. C. Programmable RNA-Guided RNA Effector Proteins Built from Human Parts. *Cell* **2019**, *178* (1), 122-134.e12. https://doi.org/10.1016/j.cell.2019.05.049.

(138) Wettengel, J.; Reautschnig, P.; Geisler, S.; Kahle, P. J.; Stafforst, T. Harnessing Human ADAR2 for RNA Repair – Recoding a PINK1 Mutation Rescues Mitophagy. *Nucleic Acids Res.* **2016**, *45* (5), gkw911. https://doi.org/10.1093/nar/gkw911.

(139) Valente, E. M.; Abou-Sleiman, P. M.; Caputo, V.; Muqit, M. M. K.; Harvey, K.; Gispert, S.; Ali, Z.; Del Turco, D.; Bentivoglio, A. R.; Healy, D. G.; Albanese, A.; Nussbaum, R.; González-Maldonado, R.; Deller, T.; Salvi, S.; Cortelli, P.; Gilks, W. P.; Latchman, D. S.; Harvey, R. J.; Dallapiccola, B.; Auburger, G.; Wood, N. W. Hereditary Early-Onset Parkinson's Disease Caused by Mutations in PINK1. *Science (80-.).* **2004**, *304* (5674), 1158–1160. https://doi.org/10.1126/science.1096284.

(140) Heep, M.; Mach, P.; Reautschnig, P.; Wettengel, J.; Stafforst, T. Applying Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing—G/C Substitution Stabilizes GuideRNAs against Editing. *Genes (Basel).* **2017**, *8* (1), 34. https://doi.org/10.3390/genes8010034.

(141) Merkle, T.; Merz, S.; Reautschnig, P.; Blaha, A.; Li, Q.; Vogel, P.; Wettengel, J.; Li, J. B.; Stafforst, T. Precise RNA Editing by Recruiting Endogenous ADARs with Antisense Oligonucleotides. *Nat. Biotechnol.* **2019**, *37* (2), 133–138. https://doi.org/10.1038/s41587-019-0013-6.

(142) Qu, L.; Yi, Z.; Zhu, S.; Wang, C.; Cao, Z.; Zhou, Z.; Yuan, P.; Yu, Y.; Tian, F.; Liu, Z.; Bao, Y.; Zhao, Y.; Wei, W. Programmable RNA Editing by Recruiting Endogenous ADAR Using Engineered RNAs. *Nat. Biotechnol.* **2019**, *37* (9), 1059–1069. https://doi.org/10.1038/s41587-019-0178-z.

(143) Lomas, D. A.; Mahadeva, R. A1-Antitrypsin Polymerization and the Serpinopathies: Pathobiology and Prospects for Therapy. *J. Clin. Invest.* **2002**, *110* (11), 1585–1590. https://doi.org/10.1172/jci16782.

(144) Crooke, S. T.; Liang, X.-H.; Baker, B. F.; Crooke, R. M. Antisense Technology: A Review. *J. Biol. Chem.* **2021**, *296*, 100416. https://doi.org/10.1016/j.jbc.2021.100416.

(145) Khan, P.; Siddiqui, J. A.; Lakshmanan, I.; Ganti, A. K.; Salgia, R.; Jain, M.; Batra, S. K.; Nasser, M. W. RNA-Based Therapies: A Cog in the Wheel of Lung Cancer Defense. *Molecular Cancer*. BioMed Central Ltd December 1, 2021, pp 1–24. https://doi.org/10.1186/s12943-021-01338-2.

(146) Hu, B.; Zhong, L.; Weng, Y.; Peng, L.; Huang, Y.; Zhao, Y.; Liang, X. J. Therapeutic SiRNA: State of the Art. *Signal Transduct. Target. Ther.* **2020**, *5* (1). https://doi.org/10.1038/s41392-020-0207-x.

(147) Liang, X.; Li, D.; Leng, S.; Zhu, X. RNA-Based Pharmacotherapy for Tumors: From Bench to Clinic and Back. *Biomedicine and Pharmacotherapy*. Elsevier Masson SAS May 1, 2020, p 109997. https://doi.org/10.1016/j.biopha.2020.109997.

(148) Bandi, N.; Vassella, E. MiR-34a and MiR-15a/16 Are Co-Regulated in Non-Small Cell Lung Cancer and Control Cell Cycle Progression in a Synergistic and Rb-Dependent Manner. *Mol. Cancer* **2011**, *10* (1), 55. https://doi.org/10.1186/1476-4598-10-55.

(149) Ortiz-Quintero, B. Extracellular MicroRNAs as Intercellular Mediators and Noninvasive Biomarkers of Cancer. *Cancers (Basel).* **2020**, *12* (11), 3455. https://doi.org/10.3390/cancers12113455.

(150) Svoronos, A. A.; Engelman, D. M.; Slack, F. J. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. *Cancer Research*. American Association for Cancer Research Inc. July 1, 2016, pp 3666–3670. https://doi.org/10.1158/0008-5472.CAN-16-0359.

(151) Cao, W.; Stricker, E.; Hotz-Wagenblatt, A.; Heit-Mondrzyk, A.; Pougialis, G.; Hugo, A.; Kuźmak, J.; Materniak-Kornas, M.; Löchelt, M. Functional Analyses of Bovine Foamy Virus-Encoded MiRNAs Reveal the Importance of a Defined MiRNA for Virus Replication and Host–Virus Interaction. *Viruses* **2020**, *12* (11), 1250. https://doi.org/10.3390/v12111250.

(152) Bochnakian, A.; Zhen, A.; Zisoulis, D. G.; Idica, A.; KewalRamani, V. N.; Neel, N.; Daugaard, I.; Hamdorf, M.; Kitchen, S.; Lee, K.; Pedersen, I. M. Interferon-Inducible MicroRNA MiR-128 Modulates HIV-1 Replication by Targeting TNPO3 MRNA. *J. Virol.* **2019**, *93* (20). https://doi.org/10.1128/jvi.00364-19.

(153) Eniafe, J.; Jiang, S. MicroRNA-99 Family in Cancer and Immunity. *Wiley Interdisciplinary Reviews: RNA*. Blackwell Publishing Ltd May 1, 2020, p e1635. https://doi.org/10.1002/wrna.1635.

(154) Bueno, M. J.; Malumbres, M. MicroRNAs and the Cell Cycle. *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Elsevier May 1, 2011, pp 592–601. https://doi.org/10.1016/j.bbadis.2011.02.002.

(155) Heyn, G. S.; Corrêa, L. H.; Magalhães, K. G. The Impact of Adipose Tissue–Derived MiRNAs in Metabolic Syndrome, Obesity, and Cancer. *Frontiers in Endocrinology*. Frontiers Media S.A. October 6, 2020, p 801. https://doi.org/10.3389/fendo.2020.563816.

(156) Setten, R. L.; Rossi, J. J.; Han, S. ping. The Current State and Future Directions of RNAi-Based Therapeutics. *Nature Reviews Drug Discovery*. Nature Publishing Group June 1, 2019, pp 421–446. https://doi.org/10.1038/s41573-019-0017-4.

(157) Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Rådmark, O.;

Kim, S.; Kim, V. N. The Nuclear RNase III Drosha Initiates MicroRNA Processing. *Nature* **2003**, *425* (6956), 415–419. https://doi.org/10.1038/nature01957.

(158) Kim, Y. K.; Kim, B.; Kim, V. N. Re-Evaluation of the Roles of DROSHA, Exportin 5, and DICER in MicroRNA Biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (13), E1881–E1889. https://doi.org/10.1073/pnas.1602532113.

(159) Liu, J.; Carmell, M. A.; Rivas, F. V.; Marsden, C. G.; Thomson, J. M.; Song, J. J.; Hammond, S. M.; Joshua-Tor, L.; Hannon, G. J. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science (80-.).* **2004**, *305* (5689), 1437–1441. https://doi.org/10.1126/science.1102513.

(160) Daugaard, I.; Hansen, T. B. Biogenesis and Function of Ago-Associated RNAs. *Trends in Genetics*. Elsevier Ltd March 1, 2017, pp 208–219. https://doi.org/10.1016/j.tig.2017.01.003.

(161) Schürmann, N.; Trabuco, L. G.; Bender, C.; Russell, R. B.; Grimm, D. Molecular Dissection of Human Argonaute Proteins by DNA Shuffling. *Nat. Struct. Mol. Biol.* **2013**, *20* (7), 818–826. https://doi.org/10.1038/nsmb.2607.

(162) Olina, A. V.; Kulbachinskiy, A. V.; Aravin, A. A.; Esyunina, D. M. Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry (Moscow)*. Pleiades Publishing May 1, 2018, pp 483–497. https://doi.org/10.1134/S0006297918050024.

(163) Wilson, R. C.; Doudna, J. A. Molecular Mechanisms of RNA Interference. *Annu. Rev. Biophys.* **2013**, *42* (1), 217–239. https://doi.org/10.1146/annurev-biophys-083012-130404.

(164) Farooqi, A. A.; Fayyaz, S.; Poltronieri, P.; Calin, G.; Mallardo, M. Epigenetic Deregulation in Cancer: Enzyme Players and Non-Coding RNAs. *Semin. Cancer Biol.* **2020**. https://doi.org/10.1016/j.semcancer.2020.07.013.

(165) Gerthoffer, W. Epigenetic Targets for Oligonucleotide Therapies of Pulmonary Arterial Hypertension. *Int. J. Mol. Sci.* **2020**, *21* (23), 9222. https://doi.org/10.3390/ijms21239222.

(166) Liang, G.; Weisenberger, D. J. DNA Methylation Aberrancies as a Guide for Surveillance and Treatment of Human Cancers. *Epigenetics*. Taylor and Francis Inc.

June 3, 2017, pp 416–432. https://doi.org/10.1080/15592294.2017.1311434.

(167) Arif, K. M. T.; Elliott, E. K.; Haupt, L. M.; Griffiths, L. R. Regulatory Mechanisms of Epigenetic MiRNA Relationships in Human Cancer and Potential as Therapeutic Targets. *Cancers (Basel).* **2020**, *12* (10), 2922. https://doi.org/10.3390/cancers12102922.

(168) Vickers, T. A.; Koo, S.; Bennett, C. F.; Crooke, S. T.; Dean, N. M.; Baker, B. F. Efficient Reduction of Target RNAs by Small Interfering RNA and RNase H-Dependent Antisense Agents. A Comparative Analysis. *J. Biol. Chem.* **2003**, *278* (9), 7108–7118. https://doi.org/10.1074/jbc.M210326200.

(169) Jackson, A. L.; Linsley, P. S. Recognizing and Avoiding SiRNA Off-Target Effects for Target Identification and Therapeutic Application. *Nature Reviews Drug Discovery*. Nature Publishing Group January 2010, pp 57–67. https://doi.org/10.1038/nrd3010.

(170) Crooke, S. T.; Witztum, J. L.; Bennett, C. F.; Baker, B. F. RNA-Targeted Therapeutics. *Cell Metabolism*. Cell Press April 3, 2018, pp 714–739. https://doi.org/10.1016/j.cmet.2018.03.004.

(171) Liang, X. H.; Sun, H.; Shen, W.; Wang, S.; Yao, J.; Migawa, M. T.; Bui, H. H.; Damle, S. S.; Riney, S.; Graham, M. J.; Crooke, R. M.; Crooke, S. T. Antisense Oligonucleotides Targeting Translation Inhibitory Elements in 5′ UTRs Can Selectively Increase Protein Levels. *Nucleic Acids Res.* **2017**, *45* (16), 9528–9546. https://doi.org/10.1093/nar/gkx632.

(172) Liang, X. H.; Shen, W.; Sun, H.; Migawa, M. T.; Vickers, T. A.; Crooke, S. T. Translation Efficiency of MRNAs Is Increased by Antisense Oligonucleotides Targeting Upstream Open Reading Frames. *Nat. Biotechnol.* **2016**, *34* (8), 875–880. https://doi.org/10.1038/nbt.3589.

(173) Hua, Y.; Vickers, T. A.; Okunola, H. L.; Bennett, C. F.; Krainer, A. R. Antisense Masking of an HnRNP A1/A2 Intronic Splicing Silencer Corrects SMN2 Splicing in Transgenic Mice. *Am. J. Hum. Genet.* **2008**, *82* (4), 834–848. https://doi.org/10.1016/j.ajhg.2008.01.014.

(174) Lima, W., Wu, H., and Crooke, S. T. The RNase H Mechanism. In *Antisense Drug Technology - Principles, Strategies, and Applications*; Stanley T. Crooke, Ed.; CRC

Press, 2007; pp 47–74.

(175) Wu, H.; Lima, W. F.; Crooke, S. T. Molecular Cloning and Expression of CDNA for Human RNase H. *Antisense Nucleic Acid Drug Dev.* **1998**, *8* (1), 53–61. https://doi.org/10.1089/oli.1.1998.8.53.

(176) Sazani, P., Graziewicz, M. A., and Kole, R. Splice Switching Oligonucleotides as Potential Therapeutics. In *Antisense Drug Technology: Principles, Strategies, and Applications*; Stanley T. Crooke, Ed.; CRC Press, 2007; pp 89–114.

(177) Hodges, D.; Crooke, S. T. Inhibition of Splicing of Wild-Type and Mutated Luciferase-Adenovirus Pre-MRNAs by Antisense Oligonucleotides. *Mol. Pharmacol.* **1995**, *48* (5).

(178) Pollak, A. J.; Hickman, J. H.; Liang, X. H.; Crooke, S. T. Gapmer Antisense Oligonucleotides Targeting 5S Ribosomal RNA Can Reduce Mature 5S Ribosomal RNA by Two Mechanisms. *Nucleic Acid Ther.* **2020**, *30* (5), 312–324. https://doi.org/10.1089/nat.2020.0864.

(179) Shen, W.; Liang, X.; Sun, H.; De Hoyos, C. L.; Crooke, S. T. Depletion of NEAT1 LncRNA Attenuates Nucleolar Stress by Releasing Sequestered P54nrb and PSF to Facilitate C-Myc Translation. *PLoS One* **2017**, *12* (3), e0173494. https://doi.org/10.1371/journal.pone.0173494.

(180) Liang, X. H.; Vickers, T. A.; Guo, S.; Crooke, S. T. Efficient and Specific Knockdown of Small Non-Coding RNAs in Mammalian Cells and in Mice. *Nucleic Acids Res.* **2011**, *39* (3). https://doi.org/10.1093/nar/gkq1121.

(181) Crooke, S. T., Vickers, T. A., Lima, W. F., and Wu, H.-J. Mechanisms of Antisense Drug Action - an Introduction. In *Antisense Drug Technology: Principles, Strategies, and Applications*; Crooke, S. T., Ed.; CRC Press, 2007; pp 3–46.

(182) Swinnen, B.; Robberecht, W.; Van Den Bosch, L. RNA Toxicity in Non-coding Repeat Expansion Disorders. *EMBO J.* **2020**, *39* (1). https://doi.org/10.15252/embj.2018101112.

(183) Bennett, C. F. Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu. Rev. Med.* **2019**, *70* (1), 307–321. https://doi.org/10.1146/annurev-med-041217-010829.

(184) Crooke, S. T.; Wang, S.; Vickers, T. A.; Shen, W.; Liang, X. Cellular Uptake and

Trafficking of Antisense Oligonucleotides. *Nat. Biotechnol.* **2017**, *35* (3), 230–237. https://doi.org/10.1038/nbt.3779.

(185) Khvorova, A.; Watts, J. K. The Chemical Evolution of Oligonucleotide Therapies of Clinical Utility. *Nature Biotechnology*. Nature Publishing Group March 1, 2017, pp 238–248. https://doi.org/10.1038/nbt.3765.

(186) Swayze, E. E., and Bhat, B. The Medicinal Chemistry of Oligonucleotides. In *Antisense Drug Technology: Principles, Strategies, and Applications*; Stanley T. Crooke, Ed.; CRC Press, 2007; pp 143–182.

(187) Janssen, H. L. A.; Reesink, H. W.; Lawitz, E. J.; Zeuzem, S.; Rodriguez-Torres, M.; Patel, K.; van der Meer, A. J.; Patick, A. K.; Chen, A.; Zhou, Y.; Persson, R.; King, B. D.; Kauppinen, S.; Levin, A. A.; Hodges, M. R. Treatment of HCV Infection by Targeting MicroRNA. *N. Engl. J. Med.* **2013**, *368* (18), 1685–1694. https://doi.org/10.1056/NEJMoa1209026.

(188) Hagedorn, P. H.; Persson, R.; Funder, E. D.; Albæk, N.; Diemer, S. L.; Hansen, D. J.; Møller, M. R.; Papargyri, N.; Christiansen, H.; Hansen, B. R.; Hansen, H. F.; Jensen, M. A.; Koch, T. Locked Nucleic Acid: Modality, Diversity, and Drug Discovery. *Drug Discov. Today* **2018**, *23* (1), 101–114. https://doi.org/10.1016/j.drudis.2017.09.018.

(189) Eckstein, F. Phosphorothioate Oligodeoxynucleotides: What Is Their Origin and What Is Unique about Them? *Antisense and Nucleic Acid Drug Development*. Mary Ann Liebert Inc. January 30, 2000, pp 117–121. https://doi.org/10.1089/oli.1.2000.10.117.

(190) Oka, N.; Wada, T.; Saigo, K. An Oxazaphospholidine Approach for the Stereocontrolled Synthesis of Oligonucleoside Phosphorothioates. *J. Am. Chem. Soc.* **2003**, *125* (27), 8307–8317. https://doi.org/10.1021/ja034502z.

(191) Wan, W. B.; Migawa, M. T.; Vasquez, G.; Murray, H. M.; Nichols, J. G.; Gaus, H.; Berdeja, A.; Lee, S.; Hart, C. E.; Lima, W. F.; Swayze, E. E.; Seth, P. P. Synthesis, Biophysical Properties and Biological Activity of Second Generation Antisense Oligonucleotides Containing Chiral Phosphorothioate Linkages. *Nucleic Acids Res.* **2014**, *42* (22), 13456–13468. https://doi.org/10.1093/nar/gku1115.

(192) Liang, X. H.; Shen, W.; Sun, H.; Kinberger, G. A.; Prakash, T. P.; Nichols, J. G.; Crooke, S. T. Hsp90 Protein Interacts with Phosphorothioate Oligonucleotides Containing

Hydrophobic 2′-Modifications and Enhances Antisense Activity. *Nucleic Acids Res.* **2016**, *44* (8), 3892–3907. https://doi.org/10.1093/nar/gkw144.

(193) Liang, X. H.; Sun, H.; Shen, W.; Crooke, S. T. Identification and Characterization of Intracellular Proteins That Bind Oligonucleotides with Phosphorothioate Linkages. *Nucleic Acids Res.* **2015**, *43* (5), 2927–2945. https://doi.org/10.1093/nar/gkv143.

(194) Steinke, C. A.; Reeves, K. K.; Powell, J. W.; Lee, S. A.; Chen, Y. Z.; Wyrzykiewicz, T.; Griffey, R. H.; Mohan, V. Vibrational Analysis of Phosphorothioate Dna: Ii. the Pos Group in the Model Compound Dimethyl Phosphorothioate [(Ch3O)2(POS)]. *J. Biomol. Struct. Dyn.* **1997**, *14* (4), 509–516. https://doi.org/10.1080/07391102.1997.10508149.

(195) Baraniak, J.; Frey, P. A. Effect of Ion Pairing on Bond Order and Charge Localization in Alkyl Phosphorothioates. *J. Am. Chem. Soc.* **1988**, *110* (12), 4059–4060. https://doi.org/10.1021/ja00220a067.

(196) Crooke, S. T.; Vickers, T. A.; Liang, X. H. Phosphorothioate Modified Oligonucleotide-Protein Interactions. *Nucleic Acids Research*. Oxford University Press June 4, 2021, pp 5235–5253. https://doi.org/10.1093/NAR/GKAA299.

(197) Gaus, H. J.; Gupta, R.; Chappell, A. E.; Østergaard, M. E.; Swayze, E. E.; Seth, P. P. Characterization of the Interactions of Chemically-Modified Therapeutic Nucleic Acids with Plasma Proteins Using a Fluorescence Polarization Assay. *Nucleic Acids Res.* **2019**, *47* (3), 1110–1122. https://doi.org/10.1093/nar/gky1260.

(198) Bennett, C. F. Pharmacological Properties of 2'-O-Methoxyethyl- Modified Oligonucleotides. In *Antisense Drug Technology*; Crooke, S. T., Ed.; CRC Press, 2007; pp 273–304. https://doi.org/10.1201/9780849387951.

(199) Witztum, J. L.; Gaudet, D.; Freedman, S. D.; Alexander, V. J.; Digenio, A.; Williams, K. R.; Yang, Q.; Hughes, S. G.; Geary, R. S.; Arca, M.; Stroes, E. S. G.; Bergeron, J.; Soran, H.; Civeira, F.; Hemphill, L.; Tsimikas, S.; Blom, D. J.; O'Dea, L.; Bruckert, E. Volanesorsen and Triglyceride Levels in Familial Chylomicronemia Syndrome. *N. Engl. J. Med.* **2019**, *381* (6), 531–542. https://doi.org/10.1056/NEJMoa1715944.

(200) Crooke, S. T.; Liang, X. hai; Crooke, R. M.; Baker, B. F.; Geary, R. S. Antisense Drug Discovery and Development Technology Considered in a Pharmacological Context. *Biochemical Pharmacology*. Elsevier Inc. August 13, 2020, p 114196.

https://doi.org/10.1016/j.bcp.2020.114196.

(201) Crooke, S. T.; Baker, B. F.; Crooke, R. M.; Liang, X. Antisense Technology: An Overview and Prospectus. *Nat. Rev. Drug Discov.* **2021**. https://doi.org/10.1038/s41573-021-00162-z.

(202) Gaus, H. J.; Gupta, R.; Chappell, A. E.; Østergaard, M. E.; Swayze, E. E.; Seth, P. P. Characterization of the Interactions of Chemically-Modified Therapeutic Nucleic Acids with Plasma Proteins Using a Fluorescence Polarization Assay. *Nucleic Acids Res.* **2019**, *47* (3), 1110–1122. https://doi.org/10.1093/nar/gky1260.

(203) Crooke, S. T.; Baker, B. F.; Pham, N. C.; Hughes, S. G.; Kwoh, T. J.; Cai, D.; Tsimikas, S.; Geary, R. S.; Bhanot, S. The Effects of 2'-o-Methoxyethyl Oligonucleotides on Renal Function in Humans. *Nucleic Acid Ther.* **2018**, *28* (1), 10–22. https://doi.org/10.1089/nat.2017.0693.

(204) Crooke, S. T.; Baker, B. F.; Kwoh, T. J.; Cheng, W.; Schulz, D. J.; Xia, S.; Salgado, N.; Bui, H. H.; Hart, C. E.; Burel, S. A.; Younis, H. S.; Geary, R. S.; Henry, S. P.; Bhanot, S. Integrated Safety Assessment of 2′-O-Methoxyethyl Chimeric Antisense Oligonucleotides in NonHuman Primates and Healthy Human Volunteers. *Mol. Ther.* **2016**, *24* (10), 1771–1782. https://doi.org/10.1038/mt.2016.136.

(205) Tillman, L. G.; Geary, R. S.; Hardee, G. E. Oral Delivery of Antisense Oligonucleotides in Man. *J. Pharm. Sci.* **2008**, *97* (1), 225–236. https://doi.org/10.1002/jps.21084.

(206) Grillone, L. R.; Lanz, R. FOMIVIRSEN. *Drugs of Today* **2001**, *37* (4), 245–255.

(207) Gale, J. D.; Jensen, J.; Berman, G.; Freimuth, W.; Li, G.; Pleil, A.; Kutty, M.; Rosenthal, A.; Boswell, C. B.; Noah, V. . E. M.; Young, L. A Placebo-Controlled Study of PF-06473871 (Anti-Connective Tissue Growth Factor Antisense Oligonucleotide) in Reducing Hypertrophic Skin Scarring. *Plast. Reconstr. Surg. - Glob. Open* **2018**, *6* (9), e1861. https://doi.org/10.1097/GOX.0000000000001861.

(208) Bennett, C. F. Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu. Rev. Med.* **2019**, *70* (1), 307–321. https://doi.org/10.1146/annurev-med-041217-010829.

(209) Finkel, R. S.; Mercuri, E.; Darras, B. T.; Connolly, A. M.; Kuntz, N. L.; Kirschner, J.; Chiriboga, C. A.; Saito, K.; Servais, L.; Tizzano, E.; Topaloglu, H.; Tulinius, M.; Montes, J.; Glanzman, A. M.; Bishop, K.; Zhong, Z. J.; Gheuens, S.; Bennett, C. F.;

Schneider, E.; Farwell, W.; De Vivo, D. C. Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy. *N. Engl. J. Med.* **2017**, *377* (18), 1723–1732. https://doi.org/10.1056/nejmoa1702752.

(210) Bennett, C. F.; Krainer, A. R.; Cleveland, D. W. Antisense Oligonucleotide Therapies for Neurodegenerative Diseases. *Annu. Rev. Neurosci.* **2019**, *42* (1), 385–406. https://doi.org/10.1146/annurev-neuro-070918-050501.

(211) MINER, P. B.; WEDEL, M. K.; XIA, S.; BAKER, B. F. Safety and Efficacy of Two Dose Formulations of Alicaforsen Enema Compared with Mesalazine Enema for Treatment of Mild to Moderate Left-Sided Ulcerative Colitis: A Randomized, Double-Blind, Active-Controlled Trial. *Aliment. Pharmacol. Ther.* **2006**, *23* (10), 1403–1413. https://doi.org/10.1111/j.1365-2036.2006.02837.x.

(212) Greuter, T.; Rogler, G. Alicaforsen in the Treatment of Pouchitis. *Immunotherapy* **2017**, *9* (14), 1143–1152. https://doi.org/10.2217/imt-2017-0085.

(213) Jairath, V.; Khanna, R.; Feagan, B. G. Alicaforsen for the Treatment of Inflammatory Bowel Disease. *Expert Opin. Investig. Drugs* **2017**, *26* (8), 991–997. https://doi.org/10.1080/13543784.2017.1349753.

(214) Liang, X. H.; Sun, H.; Hsu, C. W.; Nichols, J. G.; Vickers, T. A.; de Hoyos, C. L.; Crooke, S. T. Golgi-Endosome Transport Mediated by M6PR Facilitates Release of Antisense Oligonucleotides from Endosomes. *Nucleic Acids Res.* **2021**, *48* (3), 1372–1391. https://doi.org/10.1093/NAR/GKZ1171.

(215) Liang, X. hai; Sun, H.; Nichols, J. G.; Allen, N.; Wang, S.; Vickers, T. A.; Shen, W.; Hsu, C. W.; Crooke, S. T. COPII Vesicles Can Affect the Activity of Antisense Oligonucleotides by Facilitating the Release of Oligonucleotides from Endocytic Pathways. *Nucleic Acids Res.* **2018**, *46* (19), 10225–10245. https://doi.org/10.1093/nar/gky841.

(216) Lorenz, P.; Misteli, T.; Baker, B. F.; Bennett, C. F.; Spector, D. L. Nucleocytoplasmic Shuttling: A Novel in Vivo Property of Antisense Phosphorothioate Oligodeoxynucleotides. *Nucleic Acids Res.* **2000**, *28* (2), 582–592. https://doi.org/10.1093/nar/28.2.582.

(217) Liang, X. H.; Shen, W.; Sun, H.; Prakash, T. P.; Crooke, S. T. TCP1 Complex Proteins

Interact with Phosphorothioate Oligonucleotides and Can Co-Localize in Oligonucleotide-Induced Nuclear Bodies in Mammalian Cells. *Nucleic Acids Res.* **2014**, *42* (12), 7819–7832. https://doi.org/10.1093/nar/gku484.

(218) Shen, W.; Liang, X. H.; Crooke, S. T. Phosphorothioate Oligonucleotides Can Displace NEAT1 RNA and Form Nuclear Paraspeckle-like Structures. *Nucleic Acids Res.* **2014**, *42* (13), 8648–8662. https://doi.org/10.1093/nar/gku579.

(219) Morell, A. G.; Irvine, R. A.; Sternlieb, I.; Scheinberg, I. H.; Ashwell, G. Physical and Chemical Studies on Ceruloplasmin. V. Metabolic Studies on Sialic Acid-Free Ceruloplasmin in Vivo. *J. Biol. Chem.* **1968**, *243* (1), 155–159. https://doi.org/10.1016/S0021-9258(18)99337-3.

(220) Weigel, P. H.; Oka, J. A. The Large Intracellular Pool of Asialoglycoprotein Receptors Functions during the Endocytosis of Asialoglycoproteins by Isolated Rat Hepatocytes. *J. Biol. Chem.* **1983**, *258* (8), 5095–5102. https://doi.org/10.1016/s0021-9258(18)32543-2.

(221) Monroe, R. S.; Huber, B. E. The Major Form of the Murine Asialoglycoprotein Receptor: CDNA Sequence and Expression in Liver, Testis and Epididymis. *Gene* **1994**, *148* (2), 237–244. https://doi.org/10.1016/0378-1119(94)90694-7.

(222) Hubbard, A. L.; Wilson, G.; Ashwell, G.; Stukenbrok, H. An Electron Microscope Autoradiographic Study of the Carbohydrate Recognition Systems in Rat Liver. I. Distribution of 125I-Ligands among the Liver Cell Types. *J. Cell Biol.* **1979**, *83* (1), 47–64. https://doi.org/10.1083/jcb.83.1.47.

(223) Geffen, I.; Spiess, M. Asialoglycoprotein Receptor. *Int. Rev. Cytol.* **1993**, *137*, 181–219. https://doi.org/10.1016/S0074-7696(08)62605-4.

(224) Bider, M. D.; Spiess, M. Ligand-Induced Endocytosis of the Asialoglycoprotein Receptor: Evidence for Heterogeneity in Subunit Oligomerization. *FEBS Lett.* **1998**, *434* (1–2), 37–41. https://doi.org/10.1016/S0014-5793(98)00947-8.

(225) Spiess, M. The Asialoglycoprotein Receptor: A Model for Endocytic Transport Receptors. *Biochemistry* **1990**, *29* (43), 10009–10018. https://doi.org/10.1021/bi00495a001.

(226) Schwartz, A. L.; Fridovich, S. E.; Lodish, H. F. Kinetics of Internalization and Recycling

of the Asialoglycoprotein Receptor in a Hepatoma Cell Line. *J. Biol. Chem.* **1982**, *257* (8), 4230–4237. https://doi.org/10.1016/s0021-9258(18)34710-0.

(227) Weigel, P. H.; Yik, J. H. N. Glycans as Endocytosis Signals: The Cases of the Asialoglycoprotein and Hyaluronan/Chondroitin Sulfate Receptors. *Biochimica et Biophysica Acta - General Subjects*. Biochim Biophys Acta September 19, 2002, pp 341–363. https://doi.org/10.1016/S0304-4165(02)00318-5.

(228) Paietta, E.; Stockert, R. J.; Racevskis, J. Alternatively Spliced Variants of the Human Hepatic Asialoglycoprotein Receptor, H2, Differ in Cellular Trafficking and Regulation of Phosphorylation. *J. Biol. Chem.* **1992**, *267* (16), 11078–11084. https://doi.org/10.1016/s0021-9258(19)49877-3.

(229) Bischoff, J.; Lodish, H. F. Two Asialoglycoprotein Receptor Polypeptides in Human Hepatoma Cells. *J. Biol. Chem.* **1987**, *262* (24), 11825–11832. https://doi.org/10.1016/s0021-9258(18)60888-9.

(230) Braiterman, L. T.; Chance, S. C.; Porter, W. R.; Lee, Y. C.; Townsend, R. R.; Hubbard, A. L. The Major Subunit of the Rat Asialoglycoprotein Receptor Can Function Alone as a Receptor. *J. Biol. Chem.* **1989**, *264* (3), 1682–1688. https://doi.org/10.1016/s0021-9258(18)94240-7.

(231) Saxena, A.; Yik, J. H. N.; Weigel, P. H. H2, the Minor Subunit of the Human Asialoglycoprotein Receptor, Trafficks Intracellularly and Forms Homo-Oligomers, but Does Not Bind Asialo-Orosomucoid. *J. Biol. Chem.* **2002**, *277* (38), 35297–35304. https://doi.org/10.1074/jbc.M205653200.

(232) Ruiz, N. I.; Drickamer, K. Differential Ligand Binding by Two Subunits of the Rat Liver Asialoglycoprotein Receptor. *Glycobiology* **1996**, *6* (5), 551–559. https://doi.org/10.1093/glycob/6.5.551.

(233) Henis, Y. I.; Katzir, Z.; Shia, M. A.; Lodish, H. F. Oligomeric Structure of the Human Asialoglycoprotein Receptor: Nature and Stoichiometry of Mutual Complexes Containing H1 and H2 Polypeptides Assessed by Fluorescence Photobleaching Recovery. *J. Cell Biol.* **1990**, *111* (4), 1409–1418. https://doi.org/10.1083/jcb.111.4.1409.

(234) Hardy, M. R.; Townsend, R. R.; Parkhurst, S. M.; Lee, Y. C. Different Modes of Ligand

Binding to the Hepatic Galactose/N-Acetylgalactosamine Lectin on the Surface of Rabbit Hepatocytes. *Biochemistry* **1985**, *24* (1), 22–28. https://doi.org/10.1021/bi00322a004.

(235) Bider, M. D.; Cescato, R.; Jenö, P.; Spiess, M. High-Affinity Ligand Binding to Subunit HI of the Asialoglycoprotein Receptor in the Absence of Subunit H2. *Eur. J. Biochem.* **1995**, *230* (1), 207–212. https://doi.org/10.1111/j.1432-1033.1995.0207i.x.

(236) Grewal, P. K. The Ashwell-Morell Receptor. In *Methods in Enzymology*; Academic Press Inc., 2010; Vol. 479, pp 223–241. https://doi.org/10.1016/S0076-6879(10)79013-3.

(237) Spiess, M.; Schwartz, a L.; Lodish, H. F. Sequence of Human Asialoglycoprotein Receptor CDNA. An Internal Signal Sequence for Membrane Insertion. *J. Biol. Chem.* **1985**, *260* (4), 1979–1982.

(238) Spiess, M.; Lodish, H. F. Sequence of a Second Human Asialoglycoprotein Receptor: Conservation of Two Receptor Genes during Evolution. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (19), 6465–6469. https://doi.org/10.1073/pnas.82.19.6465.

(239) Stefanescu, R.; Born, R.; Moise, A.; Ernst, B.; Przybylski, M. Epitope Structure of the Carbohydrate Recognition Domain of Asialoglycoprotein Receptor to a Monoclonal Antibody Revealed by High-Resolution Proteolytic Excision Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (1), 148–157. https://doi.org/10.1007/s13361-010-0010-y.

(240) Chiacchia, K. B.; Drickamers, K. Direct Evidence for the Transmembrane Orientation of the Hepatic Glycoprotein Receptors. *J. Biol. Chem.* **1984**, *259* (24), 15440–15446. https://doi.org/10.1016/s0021-9258(17)42568-3.

(241) Liu, J.; Hu, B.; Yang, Y.; Ma, Z.; Yu, Y.; Liu, S.; Wang, B.; Zhao, X.; Lu, M.; Yang, D. A New Splice Variant of the Major Subunit of Human Asialoglycoprotein Receptor Encodes a Secreted Form in Hepatocytes. *PLoS One* **2010**, *5* (9). https://doi.org/10.1371/journal.pone.0012934.

(242) Yik, J. H. N.; Saxena, A.; Weigel, P. H. The Minor Subunit Splice Variants, H2b and H2c, of the Human Asialoglycoprotein Receptor Are Present with the Major Subunit H1 in Different Hetero-Oligomeric Receptor Complexes. *J. Biol. Chem.* **2002**, *277* (25),

23076–23083. https://doi.org/10.1074/jbc.M202748200.

(243) Spiess, M.; Lodish, H. F. Sequence of a Second Human Asialoglycoprotein Receptor: Conservation of Two Receptor Genes during Evolution. *Proc. Natl. Acad. Sci.* **1985**, *82* (19), 6465–6469. https://doi.org/10.1073/pnas.82.19.6465.

(244) Paietta, E.; Stockert, R. J.; Racevskis, J. Differences in the Abundance of Variably Spliced Transcripts for the Second Asialoglycoprotein Receptor Polypeptide, H2, in Normal and Transformed Human Liver. *Hepatology* **1992**, *15* (3), 395–402. https://doi.org/10.1002/hep.1840150307.

(245) Sanhueza, C. A.; Baksh, M. M.; Thuma, B.; Roy, M. D.; Dutta, S.; Pr??ville, C.; Chrunyk, B. A.; Beaumont, K.; Dullea, R.; Ammirati, M.; Liu, S.; Gebhard, D.; Finley, J. E.; Salatto, C. T.; King-Ahmad, A.; Stock, I.; Atkinson, K.; Reidich, B.; Lin, W.; Kumar, R.; Tu, M.; Menhaji-Klotz, E.; Price, D. A.; Liras, S.; Finn, M. G.; Mascitti, V. Efficient Liver Targeting by Polyvalent Display of a Compact Ligand for the Asialoglycoprotein Receptor. *J. Am. Chem. Soc.* **2017**, *139* (9), 3528–3536. https://doi.org/10.1021/jacs.6b12964.

(246) Merwin, J. R.; Noell, G. S.; Thomas, W. L.; Chiou, H. C.; DeRome, M. E.; McKee, T. D.; Spitalny, G. L.; Findeis, M. A. Targeted Delivery of DNA Using YEE(GalNAcAH)3, a Synthetic Glycopeptide Ligand for the Asialoglycoprotein Receptor. *Bioconjug. Chem.* **1994**, *5* (6), 612–620. https://doi.org/10.1021/bc00030a017.

(247) Huang, X.; Leroux, J. C.; Castagner, B. Well-Defined Multivalent Ligands for Hepatocytes Targeting via Asialoglycoprotein Receptor. *Bioconjug. Chem.* **2017**, *28* (2), 283–295. https://doi.org/10.1021/acs.bioconjchem.6b00651.

(248) D'Souza, A. A.; Devarajan, P. V. Asialoglycoprotein Receptor Mediated Hepatocyte Targeting - Strategies and Applications. *J. Control. Release* **2015**, *203*, 126–139. https://doi.org/10.1016/j.jconrel.2015.02.022.

(249) Iobst, S. T.; Drickamer, K. Selective Sugar Binding to the Carbohydrate Recognition Domains of the Rat Hepatic and Macrophage Asialoglycoprotein Receptors. *J. Biol. Chem.* **1996**, *271* (12), 6686–6693. https://doi.org/10.1074/jbc.271.12.6686.

(250) Prakash, T. P.; Yu, J.; Migawa, M. T.; Kinberger, G. A.; Wan, W. B.; Østergaard, M. E.; Carty, R. L.; Vasquez, G.; Low, A.; Chappell, A.; Schmidt, K.; Aghajan, M.; Crosby, J.;

Murray, H. M.; Booten, S. L.; Hsiao, J.; Soriano, A.; MacHemer, T.; Cauntay, P.; Burel, S. A.; Murray, S. F.; Gaus, H.; Graham, M. J.; Swayze, E. E.; Seth, P. P. Comprehensive Structure-Activity Relationship of Triantennary N-Acetylgalactosamine Conjugated Antisense Oligonucleotides for Targeted Delivery to Hepatocytes. *J. Med. Chem.* **2016**, *59* (6), 2718–2733. https://doi.org/10.1021/acs.jmedchem.5b01948.

(251) Debacker, A. J.; Voutila, J.; Catley, M.; Blakey, D.; Habib, N. Delivery of Oligonucleotides to the Liver with GalNAc: From Research to Registered Therapeutic Drug. *Mol. Ther.* **2020**, *28* (8), 1759–1771. https://doi.org/10.1016/j.ymthe.2020.06.015.

(252) Tanowitz, M.; Hettrick, L.; Revenko, A.; Kinberger, G. A.; Prakash, T. P.; Seth, P. P. Asialoglycoprotein Receptor 1 Mediates Productive Uptake of N-Acetylgalactosamine-Conjugated and Unconjugated Phosphorothioate Antisense Oligonucleotides into Liver Hepatocytes. *Nucleic Acids Res.* **2017**, *45* (21), 12388–12400. https://doi.org/10.1093/nar/gkx960.

(253) Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. Targeted Delivery of Antisense Oligonucleotides to Hepatocytes Using Triantennary N -Acetyl Galactosamine Improves Potency 10-Fold in Mice. *Nucleic Acids Res.* **2014**, *42* (13), 8796–8807. https://doi.org/10.1093/nar/gku531.

(254) Nair, J. K.; Willoughby, J. L. S.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Kandasamy, P.; Kelin, A. V.; Milstein, S.; Taneja, N.; Oshea, J.; Shaikh, S.; Zhang, L.; Van Der Sluis, R. J.; Jung, M. E.; Akinc, A.; Hutabarat, R.; Kuchimanchi, S.; Fitzgerald, K.; Zimmermann, T.; Van Berkel, T. J. C.; Maier, M. A.; Rajeev, K. G.; Manoharan, M. Multivalent N -Acetylgalactosamine-Conjugated SiRNA Localizes in Hepatocytes and Elicits Robust RNAi-Mediated Gene Silencing. *J. Am. Chem. Soc.* **2014**, *136* (49), 16958–16961. https://doi.org/10.1021/ja505986a.

(255) Wang, S.; Allen, N.; Vickers, T. A.; Revenko, A. S.; Sun, H.; Liang, X. H.; Crooke, S. T. Cellular Uptake Mediated by Epidermal Growth Factor Receptor Facilitates the Intracellular Activity of Phosphorothioate-Modified Antisense Oligonucleotides. *Nucleic Acids Res.* **2018**, *46* (7), 3579–3594. https://doi.org/10.1093/nar/gky145.

(256) Ämmälä, C.; Drury, W. J.; Knerr, L.; Ahlstedt, I.; Stillemark-Billton, P.; Wennberg-

Huldt, C.; Andersson, E. M.; Valeur, E.; Jansson-Löfmark, R.; Janzén, D.; Sundström, L.; Meuller, J.; Claesson, J.; Andersson, P.; Johansson, C.; Lee, R. G.; Prakash, T. P.; Seth, P. P.; Monia, B. P.; Andersson, S. Targeted Delivery of Antisense Oligonucleotides to Pancreatic β-Cells. *Sci. Adv.* **2018**, *4* (10), 3386–3403. https://doi.org/10.1126/sciadv.aat3386.

(257) Viney, N. J.; van Capelleveen, J. C.; Geary, R. S.; Xia, S.; Tami, J. A.; Yu, R. Z.; Marcovina, S. M.; Hughes, S. G.; Graham, M. J.; Crooke, R. M.; Crooke, S. T.; Witztum, J. L.; Stroes, E. S.; Tsimikas, S. Antisense Oligonucleotides Targeting Apolipoprotein(a) in People with Raised Lipoprotein(a): Two Randomised, Double-Blind, Placebo-Controlled, Dose-Ranging Trials. *Lancet* **2016**, *388* (10057), 2239–2253. https://doi.org/10.1016/S0140-6736(16)31009-1.

(258) Setten, R. L.; Rossi, J. J.; Han, S. ping. The Current State and Future Directions of RNAi-Based Therapeutics. *Nature Reviews Drug Discovery*. Nature Publishing Group June 1, 2019, pp 421–446. https://doi.org/10.1038/s41573-019-0017-4.

(259) Tang, X.; Zhang, S.; Fu, R.; Zhang, L.; Huang, K.; Peng, H.; Dai, L.; Chen, Q. Therapeutic Prospects of MRNA-Based Gene Therapy for Glioblastoma. *Frontiers in Oncology*. Frontiers Media S.A. November 8, 2019, p 1208. https://doi.org/10.3389/fonc.2019.01208.

(260) Fougerolles, A. R.; Maraganore, J. M. Discovery and Development of RNAi Therapeutics. In *Antisense Drug Technology: Principles, Strategies, and Applications*; Crooke, S. T., Ed.; CRC Press, 2007; pp 465–486.

(261) Janas, M. M.; Zlatev, I.; Liu, J.; Jiang, Y.; Barros, S. A.; Sutherland, J. E.; Davis, W. P.; Liu, J.; Brown, C. R.; Liu, X.; Schlegel, M. K.; Blair, L.; Zhang, X.; Das, B.; Tran, C.; Aluri, K.; Li, J.; Agarwal, S.; Indrakanti, R.; Charisse, K.; Nair, J.; Matsuda, S.; Rajeev, K. G.; Zimmermann, T.; Sepp-Lorenzino, L.; Xu, Y.; Akinc, A.; Fitzgerald, K.; Vaishnaw, A. K.; Smith, P. F.; Manoharan, M.; Jadhav, V.; Wu, J. T.; Maier, M. A. Safety Evaluation of 2-Deoxy-2-Fluoro Nucleotides in GalNAc-SiRNA Conjugates. *Nucleic Acids Res.* **2019**, *47* (7), 3306–3320. https://doi.org/10.1093/nar/gkz140.

(262) Severgnini, M.; Sherman, J.; Sehgal, A.; Jayaprakash, N. K.; Aubin, J.; Wang, G.; Zhang, L.; Peng, C. G.; Yucius, K.; Butler, J.; Fitzgerald, K. A Rapid Two-Step Method for Isolation of Functional Primary Mouse Hepatocytes: Cell Characterization and

Asialoglycoprotein Receptor Based Assay Development. *Cytotechnology* **2012**, *64* (2), 187–195. https://doi.org/10.1007/s10616-011-9407-0.

(263) Wooddell, C.; Zhu, R.; Hamilton, H.; Chu, Q.; Sternard, H.; Schumacher, J.; Schluep, T.; Seefeld, M.; Li, Z.; Given, B. Development of Subcutaneously Administered RNAi Therapeutic ARO-HBV for Chronic Hepatitis B Virus Infection. *J. Hepatol.* **2018**, *68*, S18–S19. https://doi.org/10.1016/s0168-8278(18)30255-1.

(264) Wooddell, C. I.; Blomenkamp, K.; Peterson, R. M.; Subbotin, V. M.; Schwabe, C.; Hamilton, J.; Chu, Q.; Christianson, D. R.; Hegge, J. O.; Kolbe, J.; Hamilton, H. L.; Branca-Afrazi, M. F.; Given, B. D.; Lewis, D. L.; Gane, E.; Kanner, S. B.; Teckman, J. H. Development of an RNAi Therapeutic for Alpha-1-Antitrypsin Liver Disease. *JCI Insight* **2020**, *5* (12). https://doi.org/10.1172/jci.insight.135348.

(265) Sardh, E.; Harper, P.; Balwani, M.; Stein, P.; Rees, D.; Bissell, D. M.; Desnick, R.; Parker, C.; Phillips, J.; Bonkovsky, H. L.; Vassiliou, D.; Penz, C.; Chan-Daniels, A.; He, Q.; Querbes, W.; Fitzgerald, K.; Kim, J. B.; Garg, P.; Vaishnaw, A.; Simon, A. R.; Anderson, K. E. Phase 1 Trial of an RNA Interference Therapy for Acute Intermittent Porphyria. *N. Engl. J. Med.* **2019**, *380* (6), 549–558. https://doi.org/10.1056/NEJMoa1807838.

(266) Balwani, M.; Sardh, E.; Ventura, P.; Peiró, P. A.; Rees, D. C.; Stölzel, U.; Bissell, D. M.; Bonkovsky, H. L.; Windyga, J.; Anderson, K. E.; Parker, C.; Silver, S. M.; Keel, S. B.; Wang, J.-D.; Stein, P. E.; Harper, P.; Vassiliou, D.; Wang, B.; Phillips, J.; Ivanova, A.; Langendonk, J. G.; Kauppinen, R.; Minder, E.; Horie, Y.; Penz, C.; Chen, J.; Liu, S.; Ko, J. J.; Sweetser, M. T.; Garg, P.; Vaishnaw, A.; Kim, J. B.; Simon, A. R.; Gouya, L. Phase 3 Trial of RNAi Therapeutic Givosiran for Acute Intermittent Porphyria. *N. Engl. J. Med.* **2020**, *382* (24), 2289–2301. https://doi.org/10.1056/nejmoa1913147.

(267) Regulus Therapeutics Inc. RG-125(AZD4076), a microRNA Therapeutic Targeting microRNA-103/107 Being Developed for the Treatment of NASH in Patients with Type 2 Diabetes/Pre-Diabetes, Enters Phase I Clinical Development http://ir.regulusrx.com/news-releases/news-release-details/rg-125azd4076-microrna-therapeutic-targeting-microrna-103107 (accessed Apr 14, 2021).

(268) Hu, B.; Yang, Y.; Liu, J.; Ma, Z.; Huang, H.; Liu, S.; Yu, Y.; Hao, Y.; Wang, B.; Lu, M.; Yang, D. Establishment of a Functional Cell Line Expressing Both Subunits of H1a

and H2c of Human Hepatocyte Surface Molecule ASGPR. *J. Huazhong Univ. Sci. Technol. - Med. Sci.* **2010**, *30* (5), 556–561. https://doi.org/10.1007/s11596-010-0542-1.

(269) Scharner, J.; Qi, S.; Rigo, F.; Bennett, C. F.; Krainer, A. R. Delivery of GalNAc-Conjugated Splice-Switching ASOs to Non-Hepatic Cells through Ectopic Expression of Asialoglycoprotein Receptor. *Mol. Ther. - Nucleic Acids* **2019**, *16* (June), 313–325. https://doi.org/10.1016/j.omtn.2019.02.024.

(270) Schirrmann, T.; Pecher, G. Tumor-Specific Targeting of a Cell Line with Natural Killer Cell Activity by Asialoglycoprotein Receptor Gene Transfer. *Cancer Immunol. Immunother.* **2001**, *50* (10), 549–556. https://doi.org/10.1007/s00262-001-0236-4.

(271) Dowdy, S. F. Overcoming Cellular Barriers for RNA Therapeutics. *Nat. Biotechnol.* **2017**, *35* (3), 222–229. https://doi.org/10.1038/nbt.3802.

(272) Kürti, L.; Czakó, B. *Strategic Applications of Named Reactions in Organic Synthesis*; Elsevier Academic Press, 2005.

(273) Bruckner, R. *Organic Mechanisms: Reactions, Stereochemistry and Synthesis*; Springer-Verlag Berlin Heidelberg, 2010. https://doi.org/10.1007/978-3-642-03651-4.

(274) Østergaard, M. E.; Yu, J.; Kinberger, G. A.; Wan, W. B.; Migawa, M. T.; Vasquez, G.; Schmidt, K.; Gaus, H. J.; Murray, H. M.; Low, A.; Swayze, E. E.; Prakash, T. P.; Seth, P. P. Efficient Synthesis and Biological Evaluation of 5′-GalNAc Conjugated Antisense Oligonucleotides. *Bioconjug. Chem.* **2015**, *26* (8), 1451–1455. https://doi.org/10.1021/acs.bioconjchem.5b00265.

(275) Lochmann, C. Engineering of Nuclease-Resistant GuideRNAs for Site-Directed A-to-I RNA Editing, Tuebingen University, 2019.

(276) Miklossy, G.; Hilliard, T. S.; Turkson, J. Therapeutic Modulators of STAT Signalling for Human Diseases. *Nature Reviews Drug Discovery*. NIH Public Access August 2013, pp 611–629. https://doi.org/10.1038/nrd4088.

(277) Shia, M. a; Lodish, H. F. The Two Subunits of the Human Asialoglycoprotein Receptor Have Different Fates When Expressed Alone in Fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (4), 1158–1162. https://doi.org/10.1073/pnas.86.4.1158.

(278) Randolph, L. N.; Bao, X.; Zhou, C.; Lian, X. An All-in-One, Tet-On 3G Inducible

PiggyBac System for Human Pluripotent Stem Cells and Derivatives. *Sci. Rep.* **2017**, *7* (1), 1–8. https://doi.org/10.1038/s41598-017-01684-6.

(279) Stahl, Y. Application of SNAP-ADAR GuideRNAs and Synthesis of Their GalNAc-Modified Derivatives, Tuebingen University, 2020.

(280) Berk, C.; Civenni, G.; Wang, Y.; Steuer, C.; Catapano, C. V.; Hall, J. Pharmacodynamic and Pharmacokinetic Properties of Full Phosphorothioate Small Interfering RNAs for Gene Silencing In Vivo. *https://home.liebertpub.com/nat* **2021**, *31* (3), 237–244. https://doi.org/10.1089/NAT.2020.0852.

(281) Parmar, R.; Willoughby, J. L. S.; Liu, J.; Foster, D. J.; Brigham, B.; Theile, C. S.; Charisse, K.; Akinc, A.; Guidry, E.; Pei, Y.; Strapps, W.; Cancilla, M.; Stanton, M. G.; Rajeev, K. G.; Sepp-Lorenzino, L.; Manoharan, M.; Meyers, R.; Maier, M. A.; Jadhav, V. 5′-(E)-Vinylphosphonate: A Stable Phosphate Mimic Can Improve the RNAi Activity of SiRNA-GalNAc Conjugates. *ChemBioChem* **2016**, *17* (11), 985–989. https://doi.org/10.1002/cbic.201600130.

(282) Nair, J. K.; Attarwala, H.; Sehgal, A.; Wang, Q.; Aluri, K.; Zhang, X.; Gao, M.; Liu, J.; Indrakanti, R.; Schofield, S.; Kretschmer, P.; Brown, C. R.; Gupta, S.; Willoughby, J. L. S.; Boshar, J. A.; Jadhav, V.; Charisse, K.; Zimmermann, T.; Fitzgerald, K.; Manoharan, M.; Rajeev, K. G.; Akinc, A.; Hutabarat, R.; Maier, M. A. Impact of Enhanced Metabolic Stability on Pharmacokinetics and Pharmacodynamics of GalNAc-SiRNA Conjugates. *Nucleic Acids Res.* **2017**, *45* (19), 10969–10977. https://doi.org/10.1093/nar/gkx818.

(283) Crooke, S. T.; Wang, S.; Vickers, T. A.; Shen, W.; Liang, X. H. Cellular Uptake and Trafficking of Antisense Oligonucleotides. *Nature Biotechnology*. Nature Publishing Group March 1, 2017, pp 230–237. https://doi.org/10.1038/nbt.3779.

(284) Wang, S.; Sun, H.; Tanowitz, M.; Liang, X.; Crooke, S. T. Annexin A2 Facilitates Endocytic Trafficking of Antisense Oligonucleotides. *Nucleic Acids Res.* **2016**, *44* (15), 7314. https://doi.org/10.1093/NAR/GKW595.

(285) Wang, Y.; Yu, R. Z.; Henry, S.; Geary, R. S. Pharmacokinetics and Clinical Pharmacology Considerations of GalNAc3-Conjugated Antisense Oligonucleotides. *Expert Opin. Drug Metab. Toxicol.* **2019**, *15* (6), 475–485. https://doi.org/10.1080/17425255.2019.1621838.

(286) Tarze, A.; Deniaud, A.; Bras, M. Le; Maillier, E.; Molle, D.; Larochette, N.; Zamzami, N.; Jan, G.; Kroemer, G.; Brenner, C. GAPDH, a Novel Regulator of the pro-Apoptotic Mitochondrial Membrane Permeabilization. *Oncogene 2007 2618* **2006**, *26* (18), 2606–2620. https://doi.org/10.1038/sj.onc.1210074.

(287) Merkle, T. Engineering Antisense Oligonucleotides for Site-Directed RNA Editing with Endogenous ADAR, Tuebingen University, 2020.

(288) Pfeiffer, L. Optimizing Antisense Oligonucleotidemediated RNA Editing in an Eye Disease Model, Tuebingen University, 2020.

(289) Du Rietz, H.; Hedlund, H.; Wilhelmson, S.; Nordenfelt, P.; Wittrup, A. Imaging Small Molecule-Induced Endosomal Escape of SiRNA. *Nat. Commun.* **2020**, *11* (1). https://doi.org/10.1038/s41467-020-15300-1.

(290) Crooke, S. T. Molecular Mechanisms of Antisense Oligonucleotides. *Nucleic Acid Ther.* **2017**, *27* (2), 70–77. https://doi.org/10.1089/nat.2016.0656.

(291) K, D.; C, B.; E, Q.; H, P.; A, G.; A, V.; R, B. Antisense Oligonucleotides: An Emerging Area in Drug Discovery and Development. *J. Clin. Med.* **2020**, *9* (6), 2004. https://doi.org/10.3390/JCM9062004.

(292) Lorenz, P.; Baker, B. F.; Bennett, C. F.; Spector, D. L. Phosphorothioate Antisense Oligonucleotides Induce the Formation of Nuclear Bodies. *Mol. Biol. Cell* **1998**, *9* (May), 1007–1023.

(293) Sliedregt, L. A. J. M.; Rensen, P. C. N.; Rump, E. T.; Van Santbrink, P. J.; Bijsterbosch, M. K.; Valentijn, A. R. P. M.; Van Der Marel, G. A.; Van Boom, J. H.; Van Berkel, T. J. C.; Biessen, E. A. L. Design and Synthesis of Novel Amphiphilic Dendritic Galactosides for Selective Targeting of Liposomes to the Hepatic Asialoglycoprotein Receptor. *J. Med. Chem.* **1999**, *42* (4), 609–618. https://doi.org/10.1021/jm981078h.

(294) Miller, G. J.; Gardiner, J. M. Adaptable Synthesis of C-Glycosidic Multivalent Carbohydrates and Succinamide-Linked Derivatization. *Org. Lett.* **2010**, *12* (22), 5262–5265. https://doi.org/10.1021/ol102310x.

(295) Calderon, M.; Martinelli, M.; Froimowicz, P.; Leiva, A.; Gargallo, L.; Radi??, D.; Strum??a, M. C. Synthesis and Characterization of Dendronized Polymers. *Macromol. Symp.* **2007**, *258*, 53–62. https://doi.org/10.1002/masy.200751207.

(296) Kinberger, G. A.; Cai, W.; Goodman, M. Collagen Mimetic Dendrimers. *J. Am. Chem. Soc.* **2002**, *124* (51), 15162–15163. https://doi.org/10.1021/ja021203l.

(297) Pérez, M.; Muñoz, F. J.; Muñoz, E.; Fernández, M.; Sinisterra, J. V.; Hernáiz, M. J. Synthesis of Novel Glycoconjugates and Evaluation as Inhibitors against β-Glucosidase from Almond. *J. Mol. Catal. B Enzym.* **2008**, *52–53* (1–4), 153–157. https://doi.org/10.1016/j.molcatb.2007.11.021.

(298) Diez-Castellnou, M.; Mancin, F.; Scrimin, P. Efficient Phosphodiester Cleaving Nanozymes Resulting from Multivalency and Local Medium Polarity Control. *J. Am. Chem. Soc.* **2014**, *136* (4), 1158–1161. https://doi.org/10.1021/ja411969e.

(299) Odenwald, C. Revolutionizing Site-Directed RNA Editing with Bivalent BG-GuideRNAs, Tuebingen University, 2019.

(300) Stolz, R. M.; Northrop, B. H. Experimental and Theoretical Studies of Selective Thiol-Ene and Thiol-Yne Click Reactions Involving N -Substituted Maleimides. *J. Org. Chem.* **2013**, *78* (16), 8105–8116. https://doi.org/10.1021/jo4014436.

(301) Zhang, Y.; Zhou, X.; Xie, Y.; Greenberg, M. M.; Xi, Z.; Zhou, C. Thiol Specific and Tracelessly Removable Bioconjugation via Michael Addition to 5-Methylene Pyrrolones. *J. Am. Chem. Soc.* **2017**, *139* (17), 6146–6151. https://doi.org/10.1021/jacs.7b00670.

8. Appendix

8.1. Supplementary Information

8.1.1. Supplementary Figures

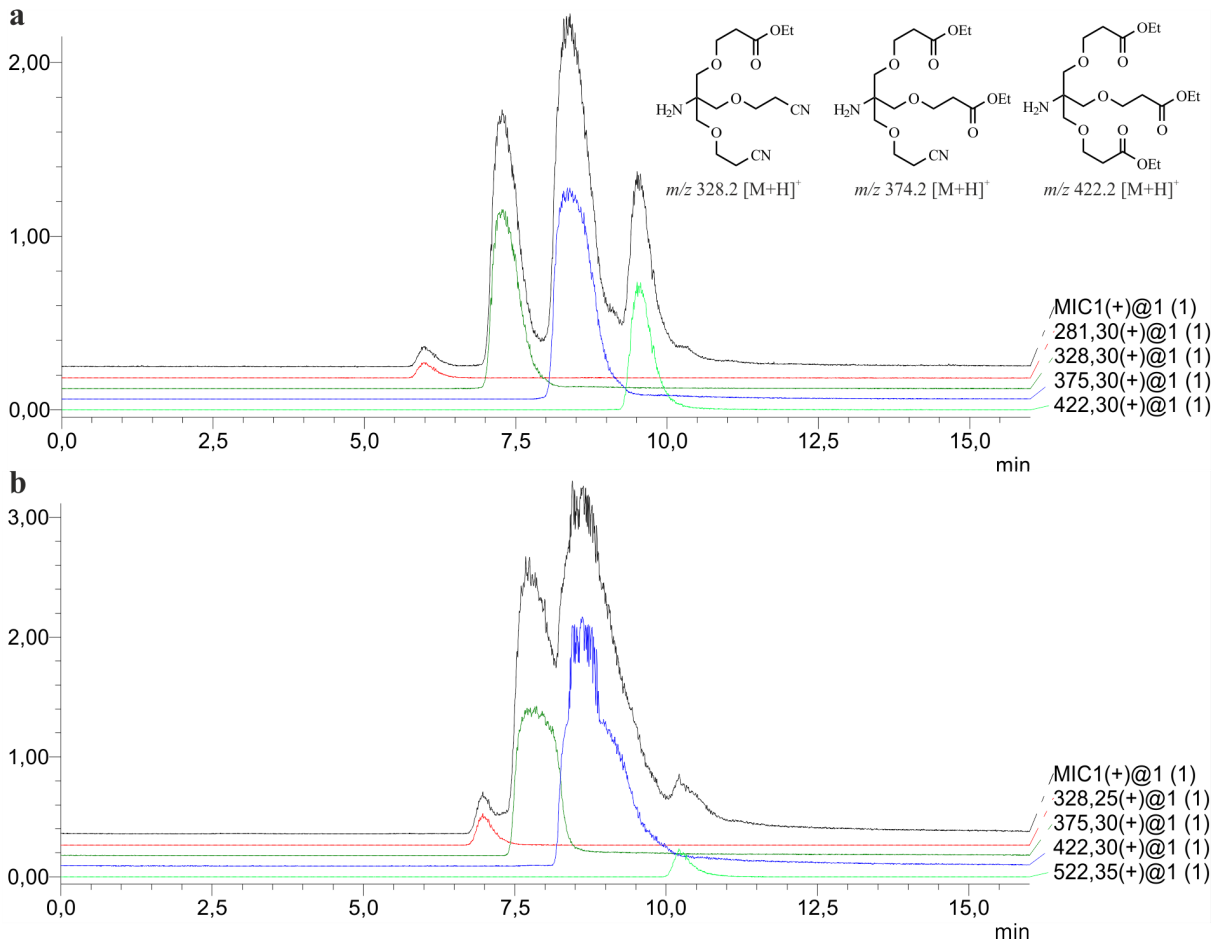


Figure S1: Mixed ion count (MIC) and mass scans of the preparation of compound 3. *(a) Pinner reaction using aq. conc. HCl (37 %) in EtOH. Single (m/z 328.3 $[M+H]^+$), double (m/z 375.3 $[M+H]^+$) and triple (m/z 422.3 $[M+H]^+$) substituted products were observed. (b) Pinner reaction using conc. H_2SO_4 (96 %) in EtOH. Mainly double (m/z 375.3 $[M+H]^+$) and triple (m/z 422.3 $[M+H]^+$) substituted products were observed. LCMS of both samples was performed as described within the general procedures.*

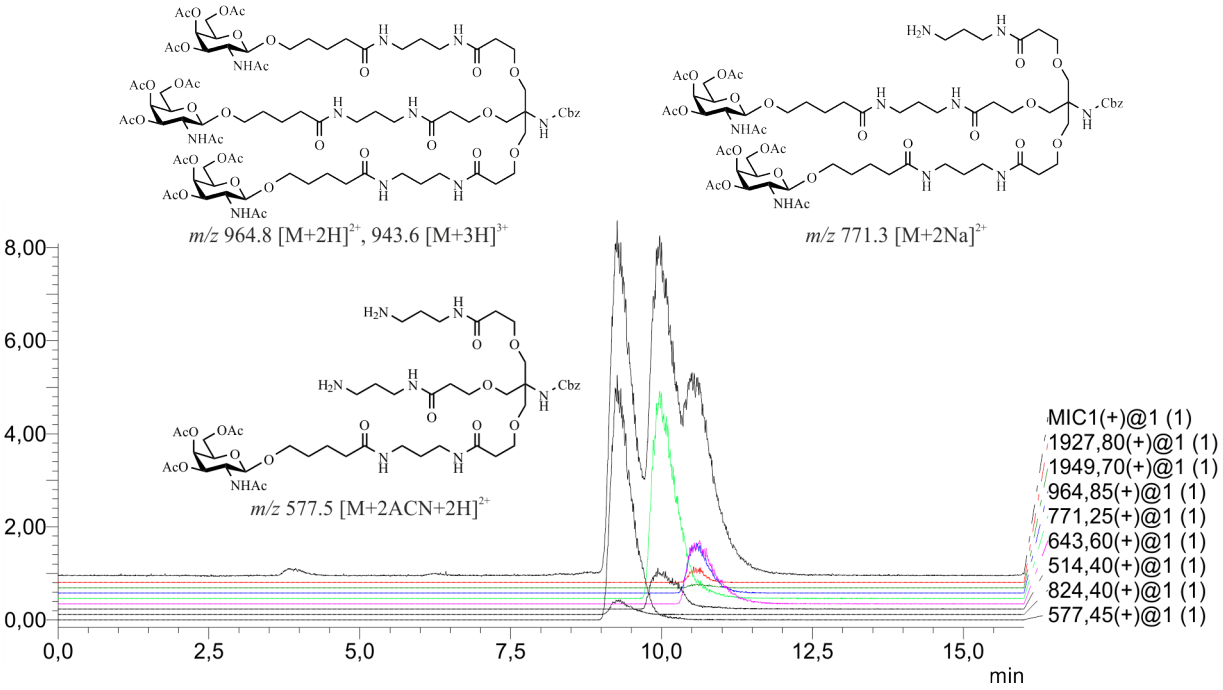


Figure S2: Mixed ion count (MIC) and mass scans of the preparation of compound 11. in a large-scale reaction. *Mainly single (m/z 577.5 [M+2ACN+2H]$^{2+}$) and double (m/z 771.3 [M+2H]$^{2+}$) conjugated products as well as a minor amount of triple (m/z 964.8 [M+2H]$^{2+}$, 943.6 [M+3H]$^{3+}$) conjugated products were observed. LCMS was performed as described within the general procedures.*

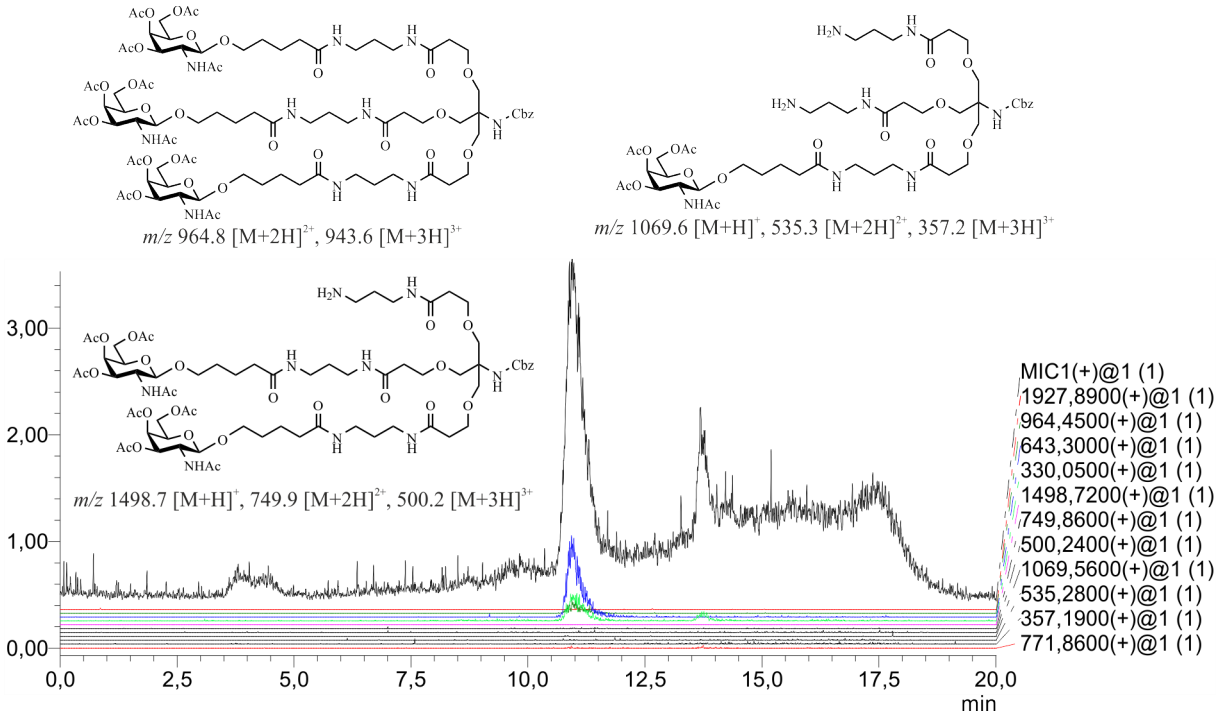


Figure S3: Mixed ion count (MIC) and mass scans of the preparation of compound 11 using compound 12. *Only triple conjugated products (m/z 964.8 [M+2H]$^{2+}$, 943.6 [M+3H]$^{3+}$) was observed. LCMS was performed as described within the general procedures.*

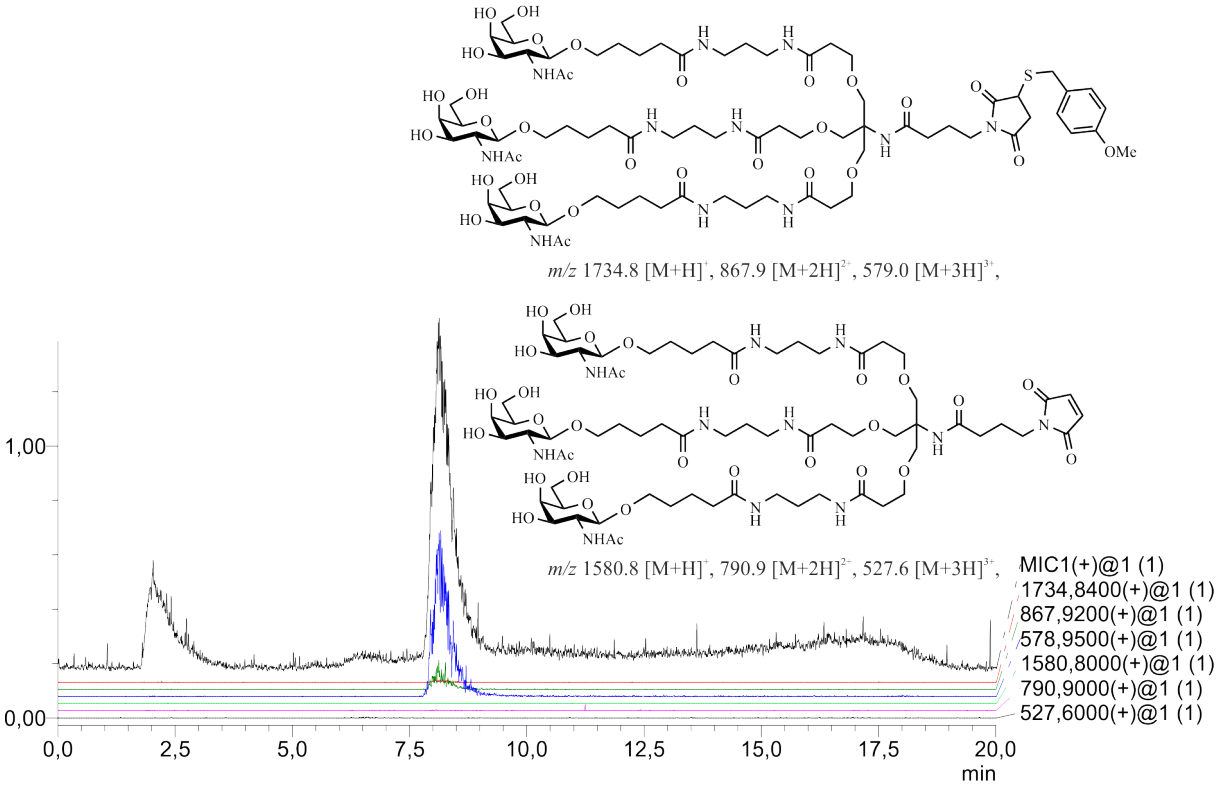


Figure S4: Mixed ion count (MIC) and mass scans of the crude product of compound 19. *Only masses of the product (m/z 1734.8 [M+H]* $^{+}$*, 867.9 [M+2H]*$^{2+}$*, 579.0 [M+3H]*$^{3+}$*) but no masses of the educt (**16**) (m/z 1580.8 [M+H]* $^{+}$*, 790.9 [M+2H]*$^{2+}$*, 527.6 [M+3H]*$^{3+}$*) were observed. LCMS was performed as described within the general procedures.*

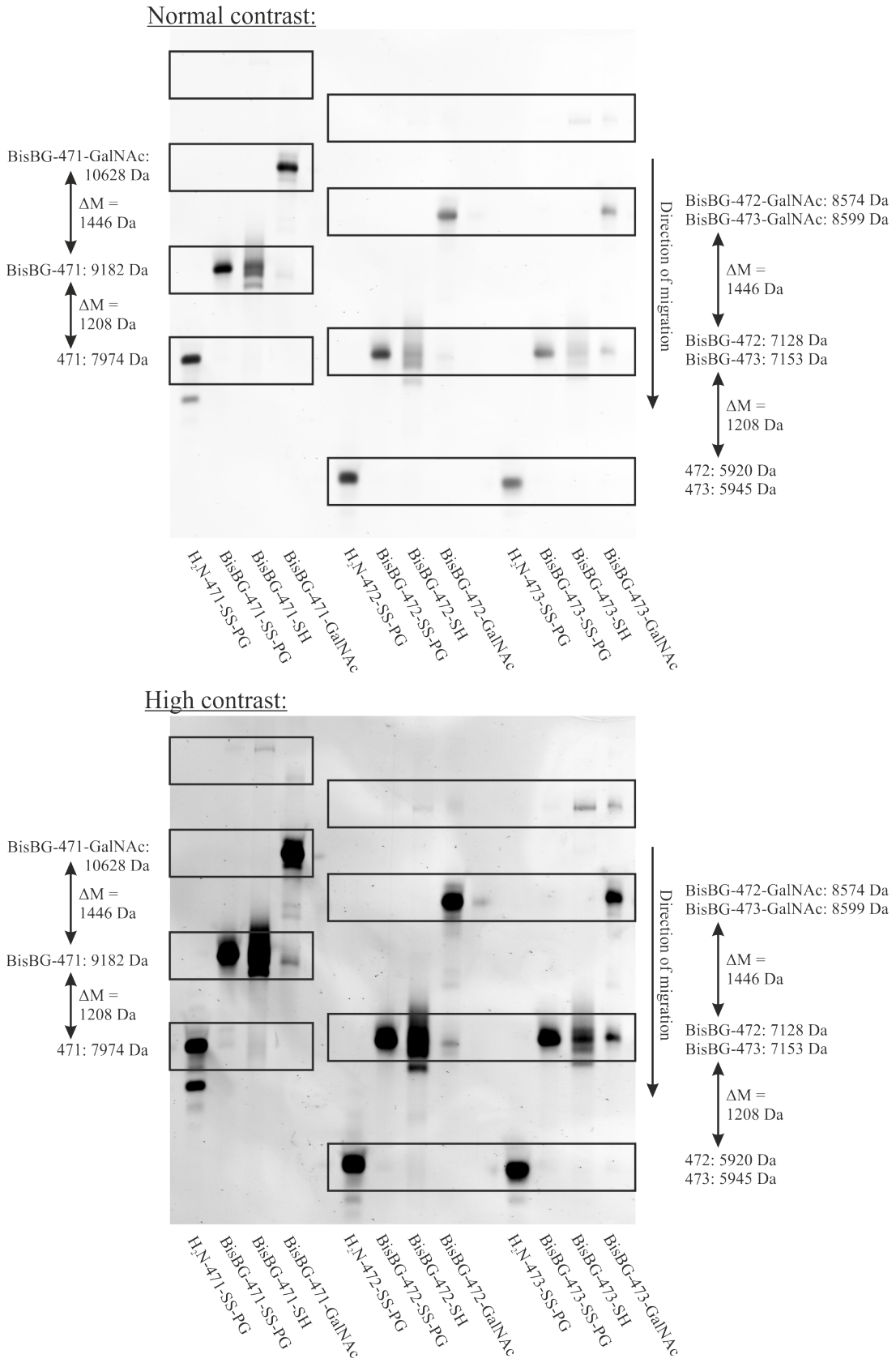


Figure S5: Uniform and high contrast adjustments of the fluorescence imaging of GalNAc conjugated gRNA 471-473. *For a detailed description, see Figure 11.*

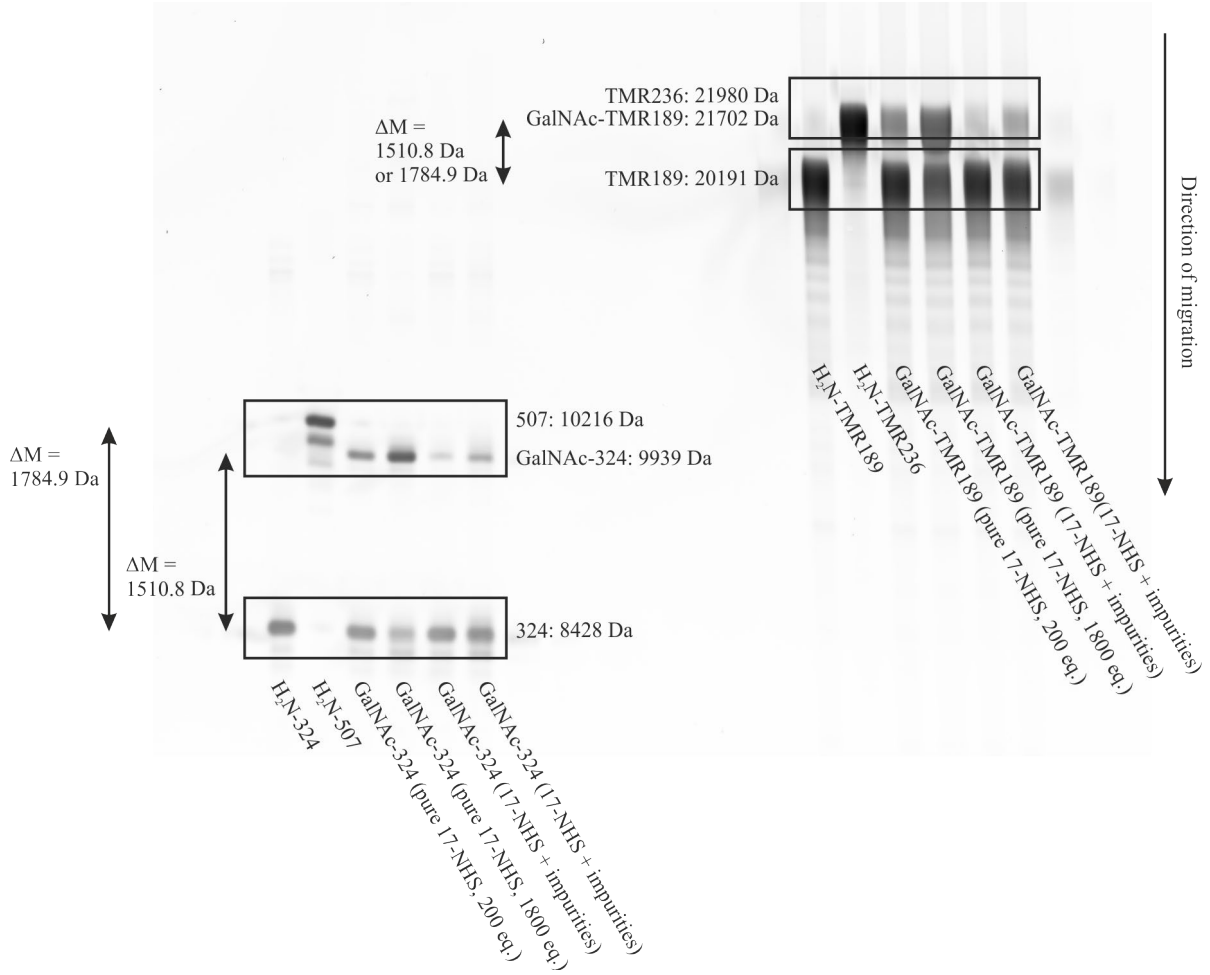


Figure S6: Uniform and high contrast adjustments of the fluorescence imaging of GalNAc conjugated gRNAs 324 and TMR189.. *For a detailed description, see Figure 12. An impure fraction of preparative HPLC purification was used for the conjugation in a similar approach (30 pmol gRNA) and a less product formation was observed within all samples.*

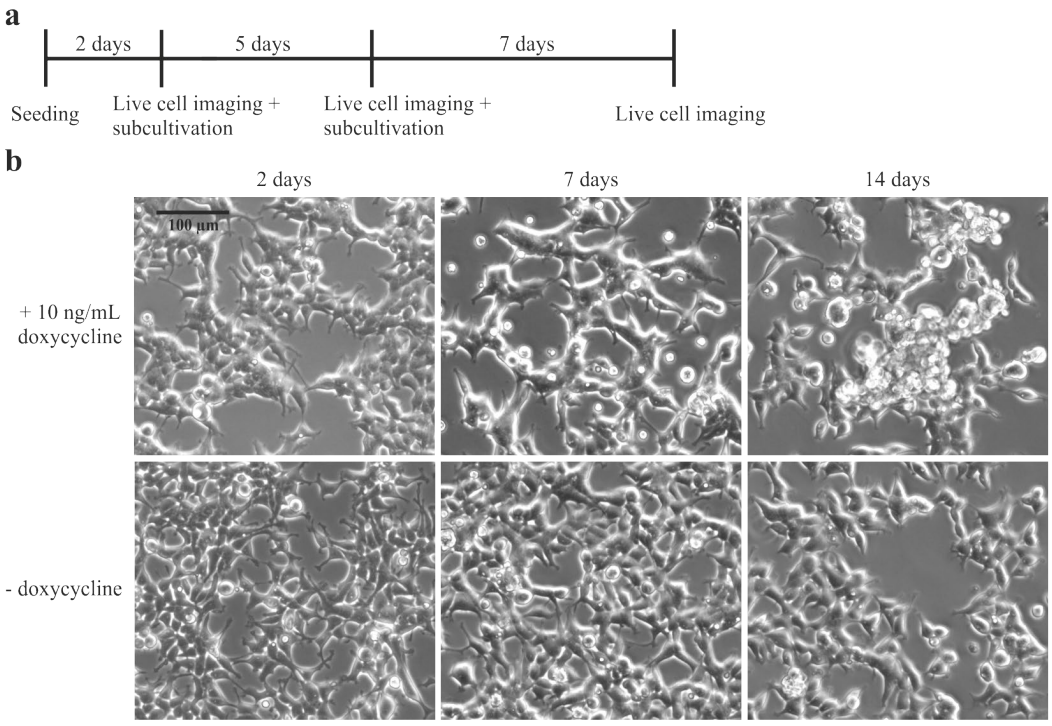


Figure S7: Live cell imaging of stable integrated ASGPR H1a FlpIn™ T-REx™ 293 cells. *(a) General workflow of the experiment. 5•10^5 cells/well were seeded into a 6-well plate and incubated in DMEM supplemented with 10 % FBS and with or without 10 ng/mL doxycycline for 2 to 14 days. The cells were subcultured after 2 and 7 days post imaging. (b) Imaging of the two different conditions after 2, 7 and 14 days utilizing a 10x magnification. The ASGPR H1a expressing FlpIn™ T-REx™ 293 cell line was generated according to the manufacturer's protocol. ASGPR H1a was cloned into a pcDNA™5/FRT expression vector via standard molecular cloning techniques utilizing KpnI and NotI restriction sites, and the isolated expression vector (pTS1070) was transfected into a FlpIn™ T-REx™ 293 host cell line according to the manufacturer's protocol using Lipofectamine™2000 for transfection and Hygromycin B for selection. Within all images, similar exposure times and intensities are applied. Due to different cell densities and for visualization, the contrast is not adjusted to a similar degree.*

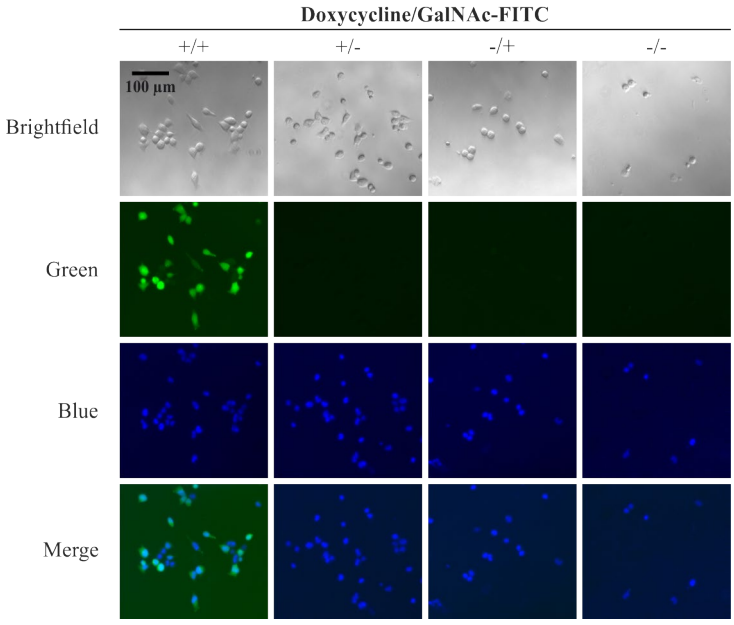


Figure S8: Internalization of GalNAc-FITC into a stably SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell line. *1.5•10^5 cells/well were seeded into poly-D-lysine HBr_{aq} coated 96-well imaging plates and incubated for 24 h with and without doxycycline induction (10 ng/mL). The adherent cells were further incubated for 24 h with 1 µM GalNAc-FITC (**15**). The nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed*

*using fluorescence microscopy. Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC (**15**) using a 10x magnification. An overview of all samples using the 63x magnification is shown in Figure 15. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree. ME = Media exchange*

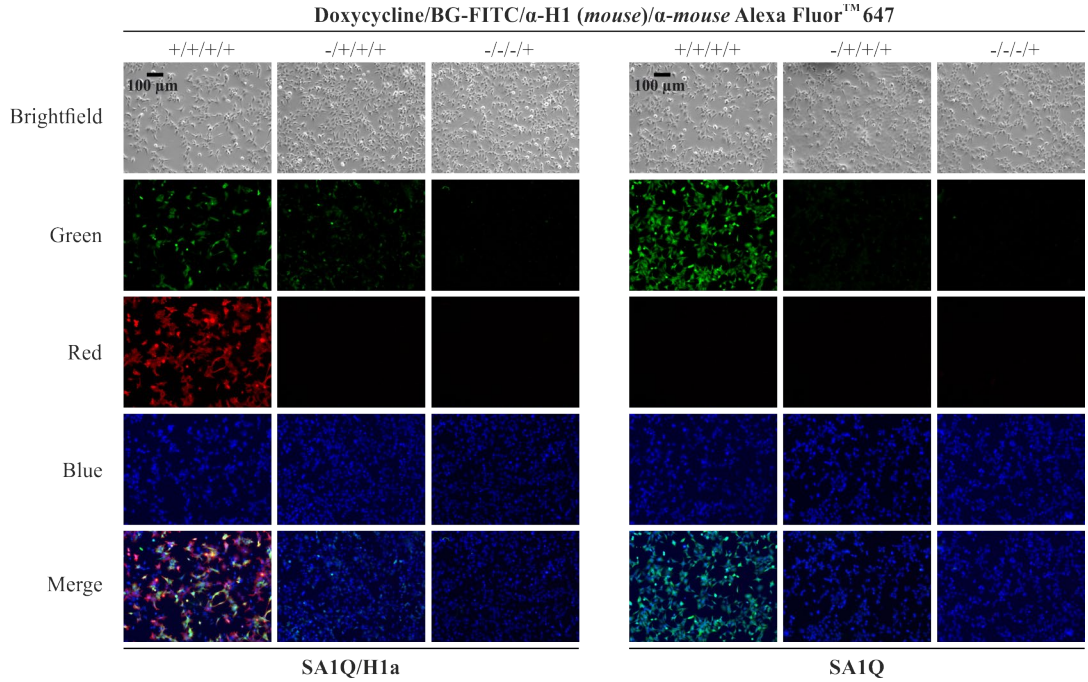


Figure S9: Immunofluorescence of a stably SA1Q and H1a expressing FlpIn™ T-REx™ 293 cell line. *1.5•10^5 cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h with and without doxycycline induction (10 ng/mL). The expressed SA1Q of the adherent cells was stained with monoacetylated BG-FITC and the nuclei were stained with NucBlue™ Live ReadyProbes. The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 647 conjugated secondary antibody (Goat α-Mouse Alexa FluorTM 647). The cells were mounted analyzed by fluorescence microscopy. Fluorescence imaging of the immunofluorescence staining of SA1Q and H1a expressing FlpIn™ T-REx™ 293 cells using a 10x magnification (left) and fluorescence imaging of the immunofluorescence staining of SA1Q expressing FlpIn™ T-REx™ 293 cells using a 10x magnification (right). An overview of all samples using the 63x magnification is shown in Figure 16. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC (green) and Alexa Fluor™ 647 (red) signals are adjusted to a similar degree.*

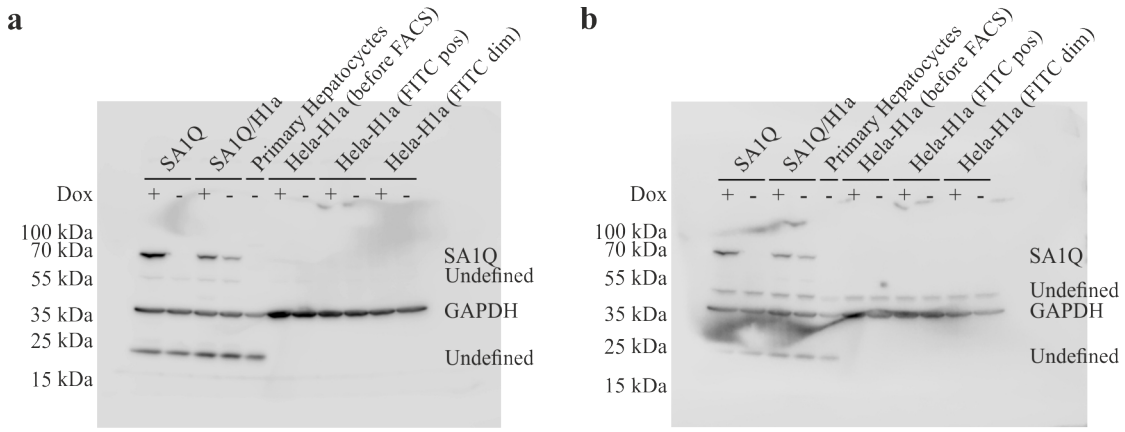


Figure S10: Western Blot analysis of the SA1Q and H1a expressing FlpIn™ T-REx™ 293 compared to the SA1Q expressing FlpIn™ T-REx™ 293 as well as primary hepatocytes and ASGPR expressing HeLa cells. *FlpIn™ T-REx™ 293 were induced with a doxycycline concentration of 10 ng/mL and HeLa cells were induced with a concentration of 200 ng/mL). Primary Hepatocytes were used as a reference for the ASGPR expression. Dox = Doxycycline, HRP = horseradish peroxidase. (a) Western Blot imaging after staining and detection of SNAP®-ADAR and GAPDH using rabbit α-SNAP (1:1000) and rabbit α-GAPDH (1:1000) as well as goat α-rabbit HRP (1:5000). (b) Western Blot imaging after staining and detection of ASGPR H1 using mouse α-ASGPR1 (1:500) as well as goat α-mouse HRP (1:5000).*

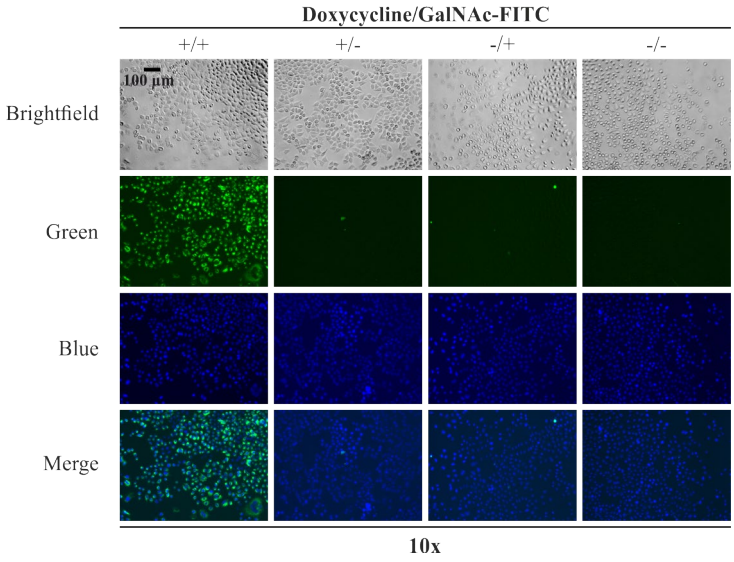


Figure S11: Internalization of GalNAc-FITC into a stably H1a expressing HeLa cell line *$1 \cdot 10^4$ cells/well were seeded into poly-D-lysine HBr_{aq} coated 96-well imaging plates and incubated for 24 h with and without doxycycline induction (200 ng/mL). The adherent cells were further incubated for 24 h with 1 µM GalNAc-FITC (**15**). The nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed using fluorescence microscopy. Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC (**15**) using a 10x magnification. An overview of all samples using the 63x magnification is shown in Figure 17. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree.*

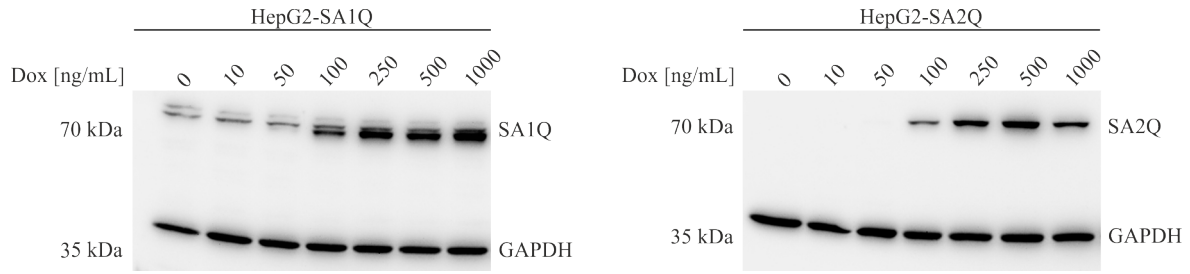


Figure S12: Western Blot analysis of SA1Q and SA2Q expressing HepG2 cells. *1•10^5 cells/well were seeded into a 24-well plate and incubated for 24 h with and without doxycycline induction (0-1000 ng/mL). The cells were lysed using NP40 lysis buffer (supplemented with cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail) and the total amount of protein was determined using a Pierce™ BCA Protein Assay Kit with BSA as reference (0-1.5 mg/mL). 15 µg whole cell lysates were used for western blotting. Subsequently, SNAP®-ADAR and GAPDH were stained and detected using rabbit α-SNAP (1:1000) and rabbit α-GAPDH (1:1000), respectively. For detection goat α-rabbit HRP (1:5000) was used and SNAP®-ADAR as well as GAPDH were detected simultaneously. Dox = Doxycycline, HRP = horseradish peroxidase. The SA1Q and SA2Q expressing HepG2 cell lines were generated according to the previously described protocol. SA1Q and SA2Q were cloned into a NeoR containing XLone PiggyBac expression vector (pTS1032) via standard molecular cloning techniques utilizing NotI and SpeI restriction sites, and the isolated expression vectors (SA1Q: pTS1037, SA2Q: pTS1040; 750 ng) were co-transfected with the transposase containing vector (pTS687; 250 ng) into a wild-type HepG2 cell line using FuGENE® 6 (3 µl/µg) according to the manufacturer's protocol for 48 h. After transfection, the cells were transferred into a 6-well cell culture plate and selected for 14 days using Geniticin™ (G418) (1250 ng/mL).*

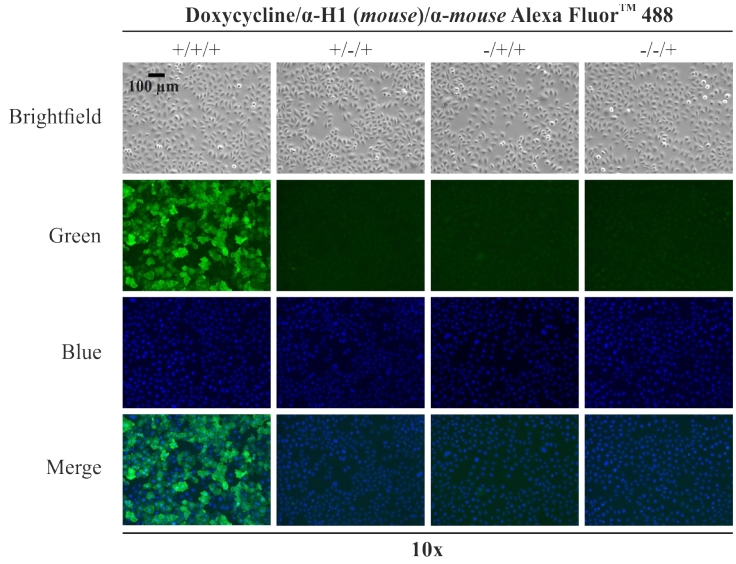


Figure S13: Immunofluorescence of a stably H1a expressing HeLa cell line. *1•10^5 cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀12 mm) and incubated for 24 h with and without doxycycline induction (200 ng/mL). The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed by fluorescence microscopy using a 10x magnification. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC (green) signals are adjusted to a similar degree.*

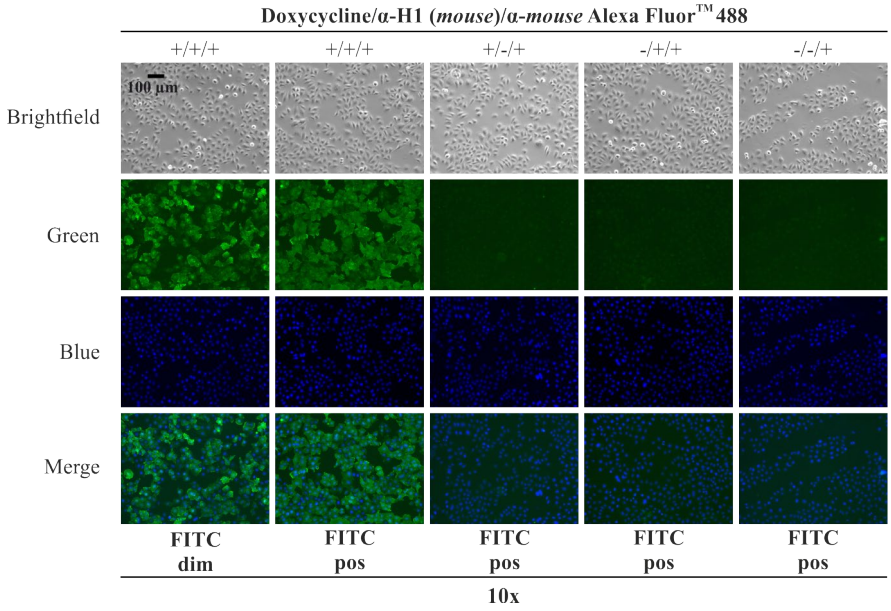


Figure S14: Immunofluorescence of the sorted and stably H1a expressing HeLa cell line. *$1 \cdot 10^5$ cells/well were seeded onto poly-D-lysine HBr_{aq} coated cover glasses (⌀ 12 mm) and incubated for 24 h with and without doxycycline induction (200 ng/mL). The cells were fixated using p-formaldehyde and membrane bound ASGPR H1a was stained using an antibody against the receptor subunit H1 (Mouse α-ASGPR1) and an Alexa Fluor™ 488 conjugated secondary antibody (Goat α-Mouse Alexa Fluor™ 488). The nuclei were stained with NucBlue™ Live ReadyProbes and the mounted cells were analyzed by fluorescence microscopy using a 10x magnification. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC (green) signals are adjusted to a similar degree.*

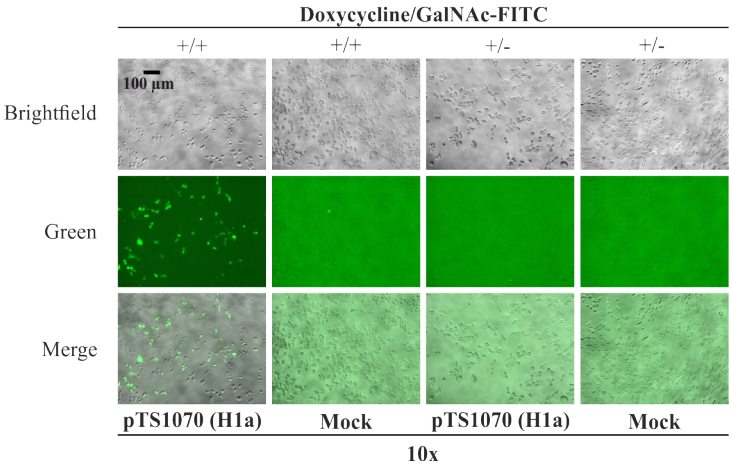


Figure S15: Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC into pcDNA5-H1a transfected SA1Q expressing FlpIn™ T-REx™ 293 cells. *$3 \cdot 10^5$ cells/well were seeded into 24-well cell culture plates and incubated overnight. The adherent cells were transfected for 24 h with pTS1070 (500 ng) using FuGENE® 6 (4 µL/µg) in OptiMEM™ and as negative control, no vector was transfected (mock). After transfection, $1.25 \cdot 10^4$ cells/well were seeded into poly-D-lysine HBr_{aq} coated 96-well cell culture plates and incubated for 24 h under doxycycline induction (10 ng/mL). After 24 h, the media were exchanged and GalNAc-FITC (**15**) was added with a final concentration of 1 µM and the cells were incubated for 1 h and analyzed via fluorescence microscopy. Within a single light channel and magnification, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree.*

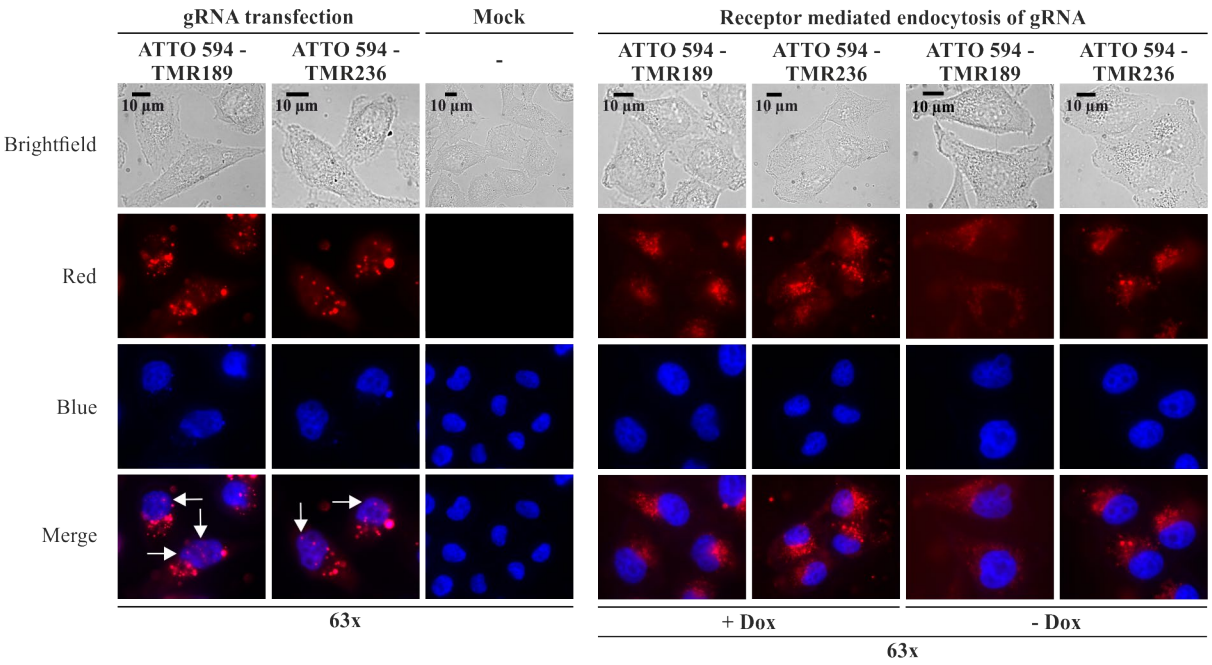


Figure S16: Fluorescence imaging of the transfection and receptor mediated endocytosis of ATTO 594 labeled RESTORE v2 gRNAs into unsorted and H1a expressing HeLa cells. *For a detailed description, see Figure 26c and d. Mock = negative control (no gRNA). Within a single light channel, similar exposure times and intensities are applied for the different conditions and within a single experiment, the contrasts of all ATTO 594 signals (red) are adjusted to a similar degree.*

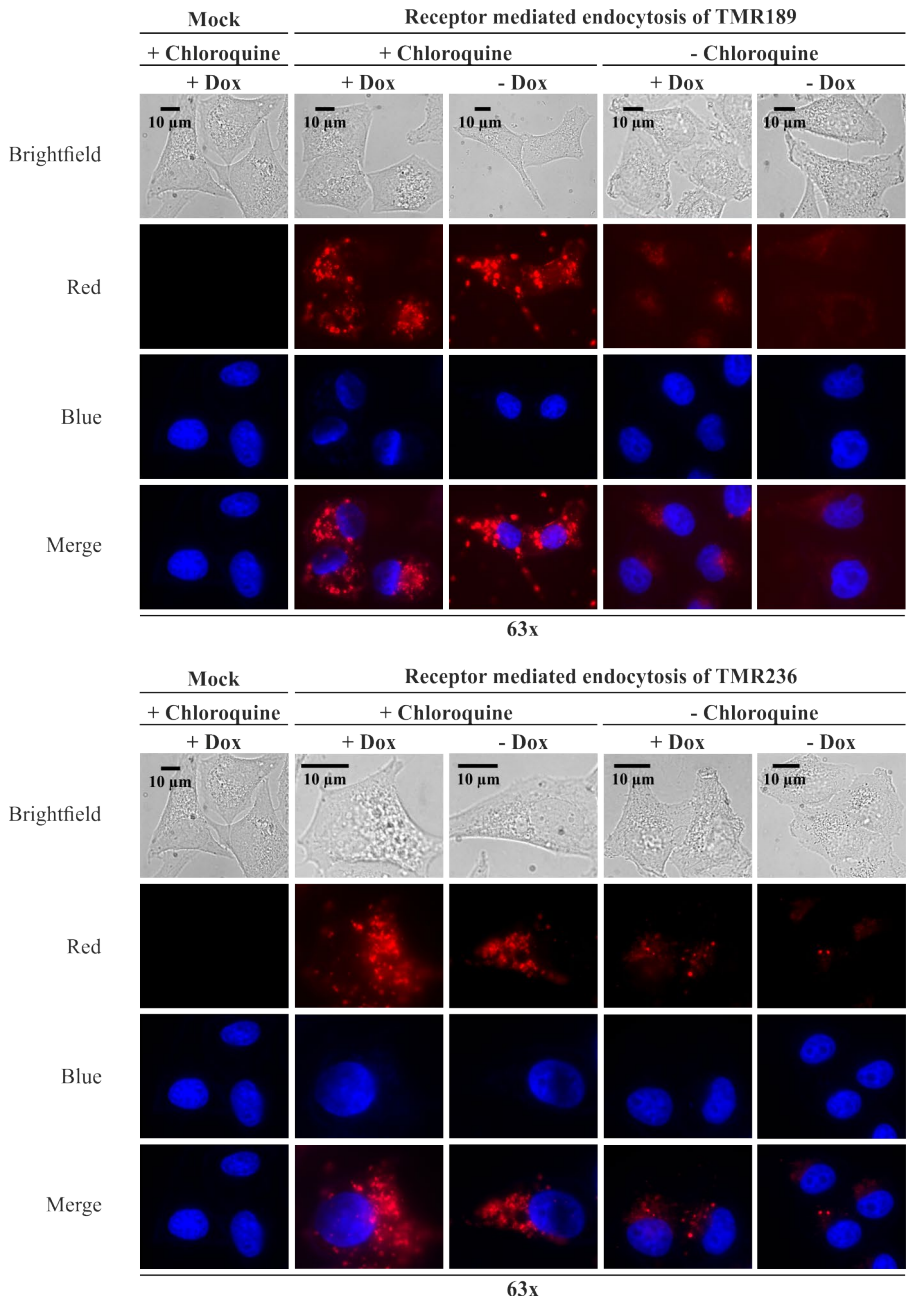


Figure S17: Fluorescence imaging of the receptor mediated endocytosis of ATTO 594 labeled RESTORE v2 gRNAs into unsorted and H1a expressing HeLa cells in the absence or presence of chloroquine. *For a detailed description, see Figure 26e. Mock = negative control (no gRNA). Within a single light channel, similar exposure times and intensities are applied for the different conditions and within a single experiment, the contrasts of all ATTO 594 signals (red) are adjusted to a similar degree.*

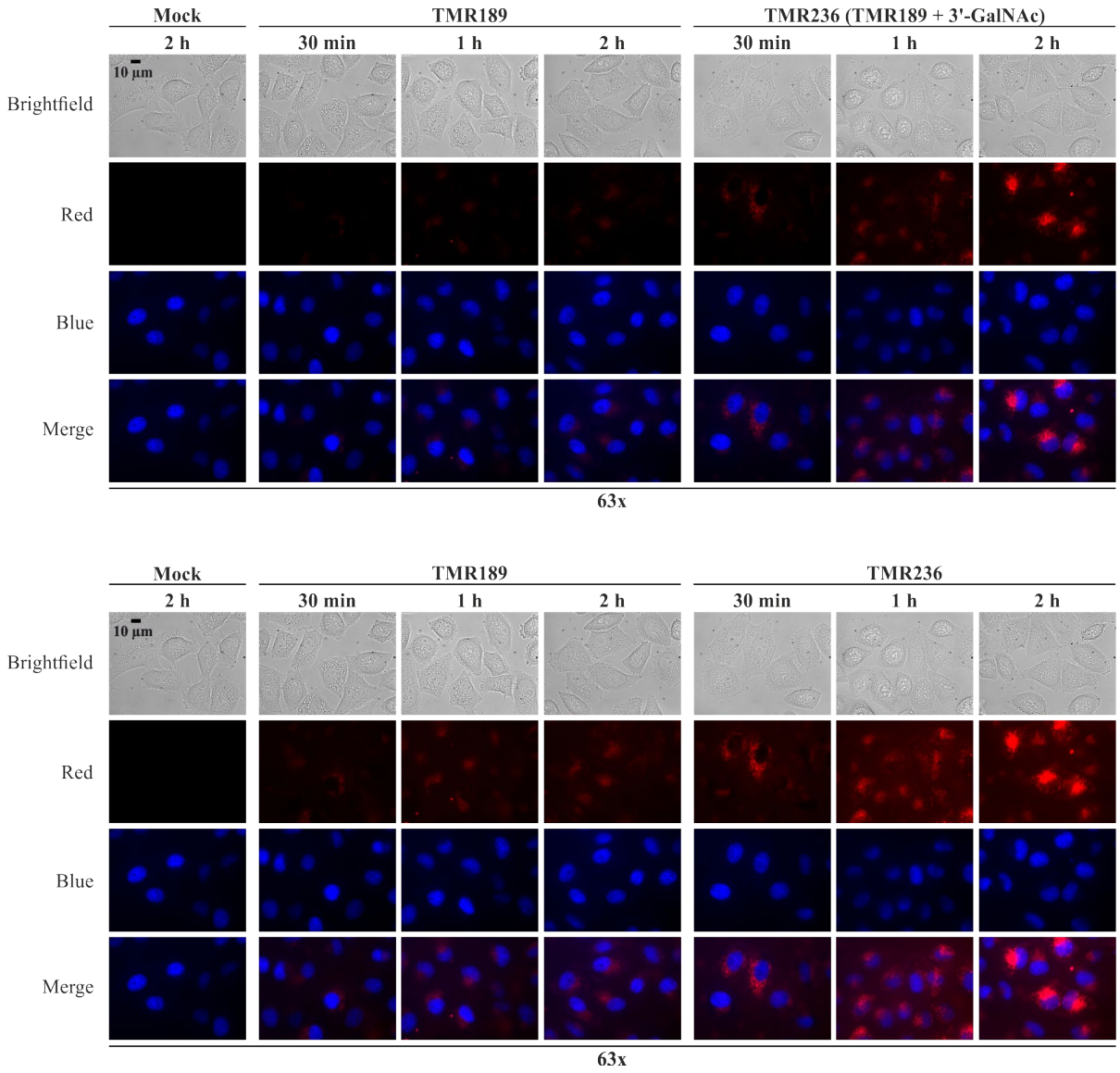


Figure S18: Fluorescence imaging of the receptor mediated endocytosis of ATTO 594 labeled RESTORE v2 gRNAs into unsorted and H1a expressing HeLa cells for 0-2h. *For a detailed description, see Figure 26f. For reasons of illustration, high contrast adjustments are shown in Figure 26f (lower part). Lower contrast adjustments are illustrated in the upper part. Mock = negative control (no gRNA). Within a single light channel, similar exposure times and intensities are applied for the different conditions and within a single experiments, the contrasts of all ATTO 594 signals (red) are adjusted to a similar degree.*

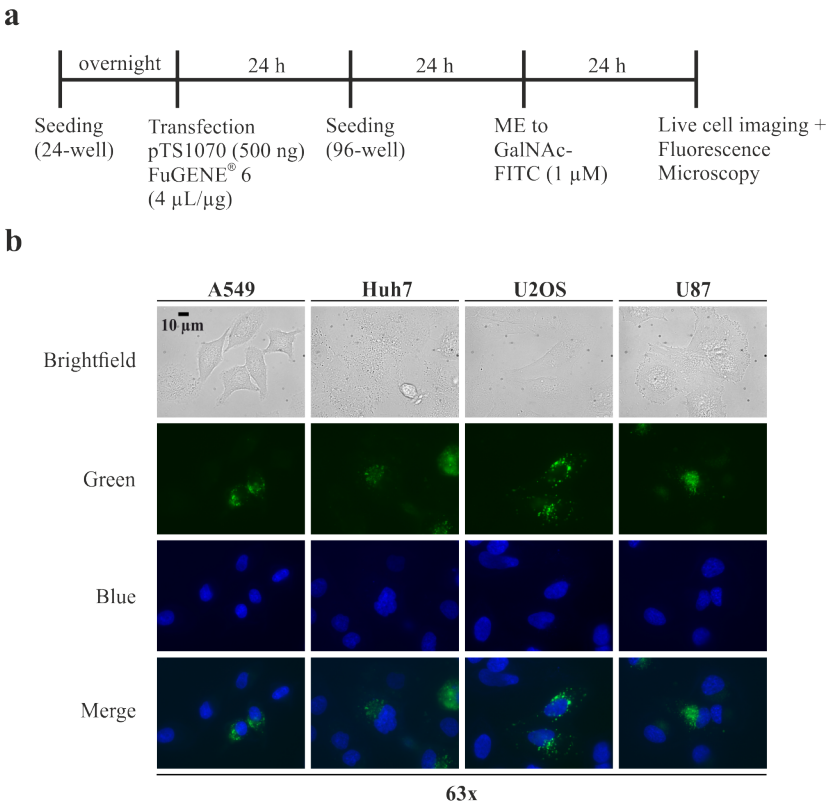
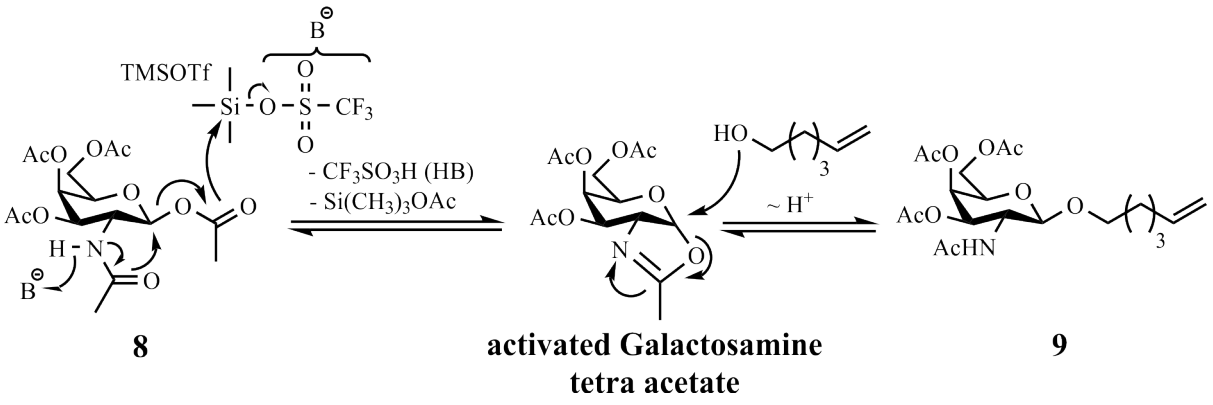
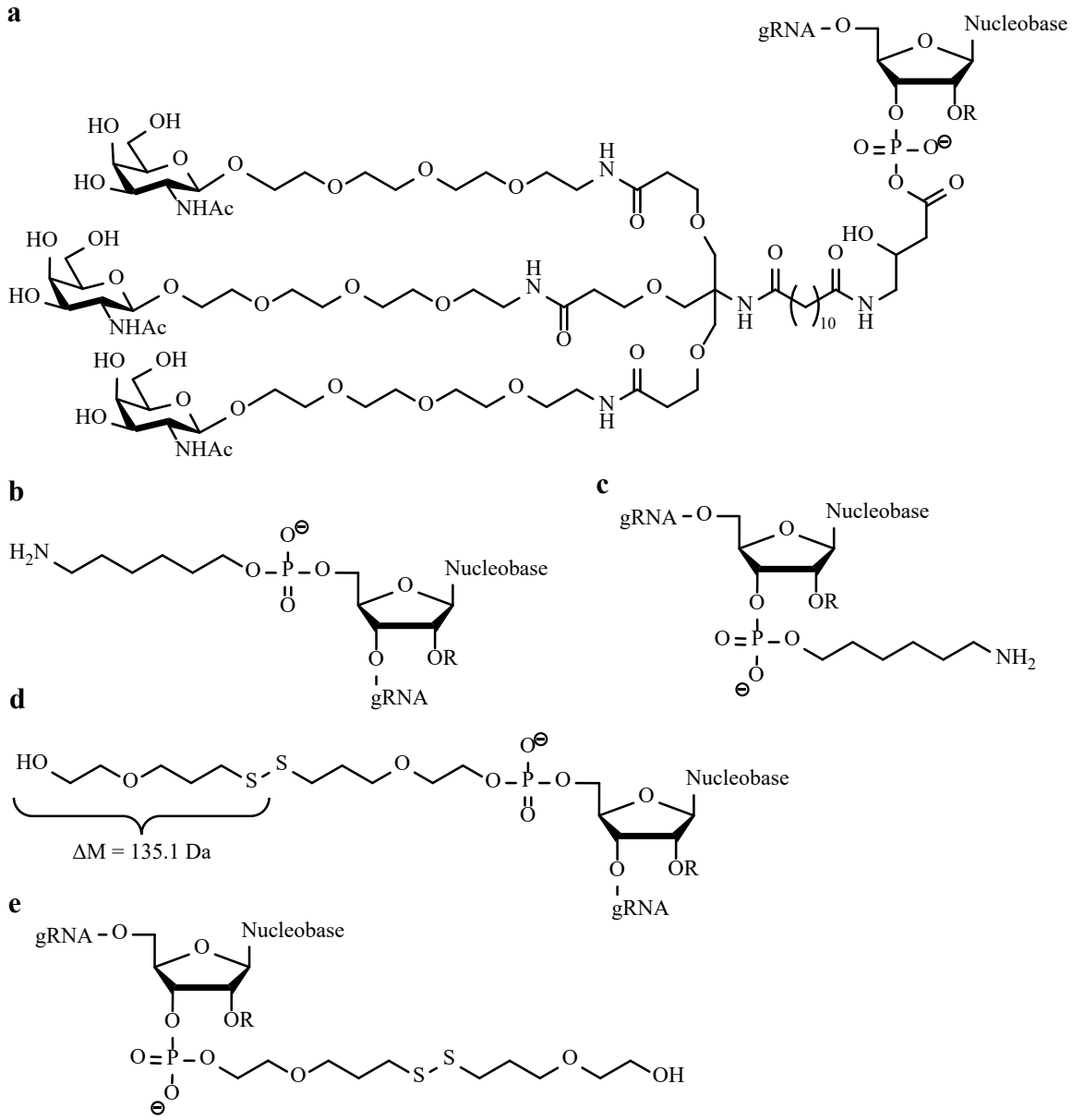


Figure S19: Fluorescence imaging of the receptor mediated uptake of GalNAc-FITC into H1a transfected A549, Huh7, U2OS and U87 cells. *$7.5 \cdot 10^4$ cells/well were seeded into 24-well cell culture plates and incubated overnight. The adherent cells were transfected for 24 h with pTS1070 (500 ng) using FuGENE® 6 (4 µL/µg) in OptiMEM™. After transfection, $5 \bullet 10^3$ cells/well were seeded into poly-D-lysine HBr_{aq} coated 96-well cell culture plates and incubated for 24 h. The media were exchanged and GalNAc-FITC (**15**) was added with a final concentration of 1 µM and the cells were incubated for 24 h. The nuclei were stained with NucBlue™ Live ReadyProbes and the cells were analyzed by fluorescence microscopy using a 63x magnification. Within a single light channel, similar exposure times and intensities are applied for the different conditions and the contrasts of all FITC signals (green) are adjusted to a similar degree. ME = Media exchange.*

8.1.2. Supplementary schemes



Scheme S1: Reaction mechanism of the TMSOTf induced activation of β-D-Galactosamine penta acetate and β-selective substitution of 5-Hexen-1-ol. *The trimethylsilyl, a lewis acceptor, is activating the carbonyl group of the anomeric O-acetyl, which serves as a leaving group, due to an intramolecular electron transfer and the deprotonation of the amide proton. The result is a bicyclic intermediate, whereby the oxazoline derivative is occupying the α-anomeric position of the galactosamine. The subsequent nucleophile substitution (S_N2) of 5-Hexen-1-ol to the bicyclic intermediate is preferring the β-configurated product formation (**9**), while the N-acetylation is restored.*



Scheme S2: Different 5'- and 3'-terminal modification of gRNAs provided by the manufacturer. *(a) 3'-terminal tris-based triantennary N-acetyl galactosamine modification. (b) 5'-terminal amino modification including a C_6 linker. (c) 3'-terminal amino modification including a C_6 linker. (d) 5'-terminal disulfide modification including a C_6 linker. (e) 3'-terminal disulfide modification including a C_6 linker. R = H or desired chemical modification.*

8.1.3. Supplementary tables

Table S1: Calculation of the amount of DTT used for the deprotection of gRNA 257 within the corresponding supernatant using the law of Lambert-Beer and the molecular attenuation coefficient ε_{412} = 14,150 $M^{-1} \cdot cm^{-1}$ of 2-nitro-5-thiobenzoic acid. *The samples were prepared as described in section 6.6.3. The absorbance of Ellman's reagent (25 µL, 25 mM in phosphate buffer (0.1 M, pH 7.0)), diluted in 75 µL phosphate buffer (0.1 M, pH 7.0), was determined and subtracted from all measurements and the measurements are stated in arbitrary units (a.u.). The path length was calculated from the volume (100 µL) and a cylindrical shape was assumed. An initial amount of 1 µmol DTT was used. Abs. = Absorbance*

	Difference = Abs.-blank [a.u.]	c (DTT, measurement) [µM]	n (DTT, sample) [µmol]	n (DTT, supernatant) [µmol]
Supernatant (Precipitation)	2.1781	0.5497	0.0550	1.1610
Supernatant (Wash step I)	0.0131	0.0033	0.0003	0.0066
Supernatant (Wash step II)	-0.0114	-0.0029	-0.0003	-0.0058
Supernatant (Wash step III)	-0.0007	-0.0002	0.0000	-0.0004

Table S2: Calculation of the amount of DTT used for the deprotection of gRNA 258 within the corresponding supernatant using the law of Lambert-Beer and the molecular attenuation coefficient ε_{412} = 14,150 $M^{-1} \cdot cm^{-1}$ of 2-nitro-5-thiobenzoic acid. *The samples were prepared as described in section 6.6.3. The absorbance of Ellman's reagent (25 µL, 25 mM in phosphate buffer (0.1 M, pH 7.0)), diluted in 75 µL phosphate buffer (0.1 M, pH 7.0), was determined and subtracted from all measurements and the measurements are stated in arbitrary units (a.u.). The path length d = 0.28 cm was calculated from the volume (100 µL) and a cylindrical shape was assumed. An initial amount of 1 µmol DTT was used. Abs. = Absorbance.*

	Difference = Abs.-blank [a.u.]	c (DTT, measurement) [µM]	n (DTT, sample) [µmol]	n (DTT, supernatant) [µmol]
Supernatant (Precipitation)	2.2361	0.5644	0.0564	1.1920
Supernatant (Wash step I)	0.0118	0.0030	0.0003	0.0063
Supernatant (Wash step II)	-0.0068	-0.0017	-0.0002	-0.0036
Supernatant (Wash step III)	-0.0078	-0.0020	-0.0002	-0.0042

Table S3: Calculation of the residual amount of DTT within the eluent of Zeba™ Spin Desalting Columns 7 kDa MWCO using the law of Lambert-Beer and the molecular attenuation coefficient ε_{412} = 14,150 $M^{-1} \cdot cm^{-1}$ of 2-nitro-5-thiobenzoic acid. *5 µL DTT (100 mM in phosphate buffer (0.1 M, pH 8.4)) were diluted with 45 µL phosphate buffer (0.1 M, pH 7.0) and purified with the described desalting columns according to the manufacturer's protocol. The columns were eluted twice with 70 µL phosphate buffer (0.1 M, pH 7.0) and the amount of DTT was determined within each eluent. The samples were prepared as described in section 6.6.3. The absorbance of Ellman's reagent (25 µL, 25 mM in phosphate buffer (0.1 M, pH 7.0)), diluted in 75 µL phosphate buffer (0.1 M, pH 7.0), was determined and subtracted from all measurements and the measurements are stated in arbitrary units (a.u.). As reference, the initial amount of DTT was determined according to the same protocol, whereby a 1:10 dilution was necessary due to an exceeding signal intensity. The path length d = 0.28 cm was calculated from the volume (100 µL) and a cylindrical shape was assumed. Abs. = Absorbance. This data were previously reported by Yannis Stahl during his bachelor's thesis.*

	Difference = Abs.-blank [a.u.]	n (DTT, sample) [nmol]	Percentage (DTT) [%]
DTT (1:10)	0.9834	496.42	99.23
Eluent I	-0.0095	-0.48	-0.10
Eluent II	0.0929	4.69	0.94

8.1.4. NMR spectra

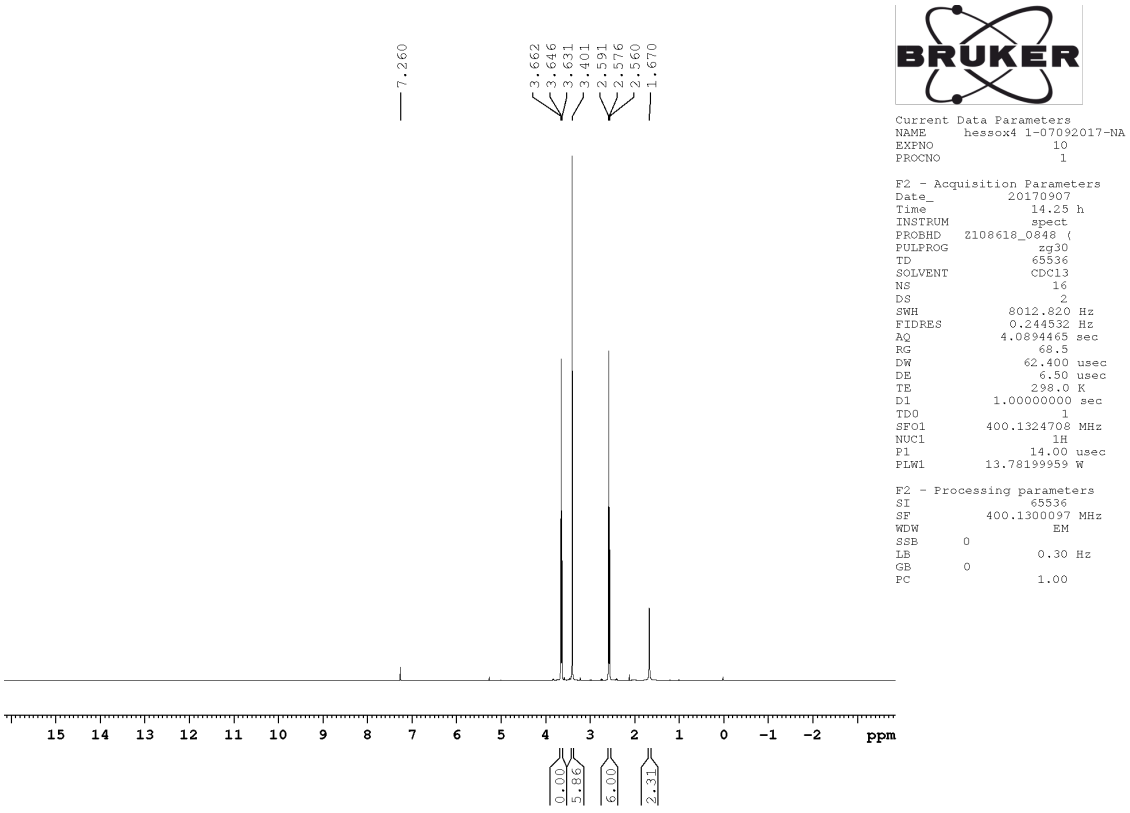


Figure S20: ^{1}H-NMR spectrum of compound 2 in $CDCl_3$.

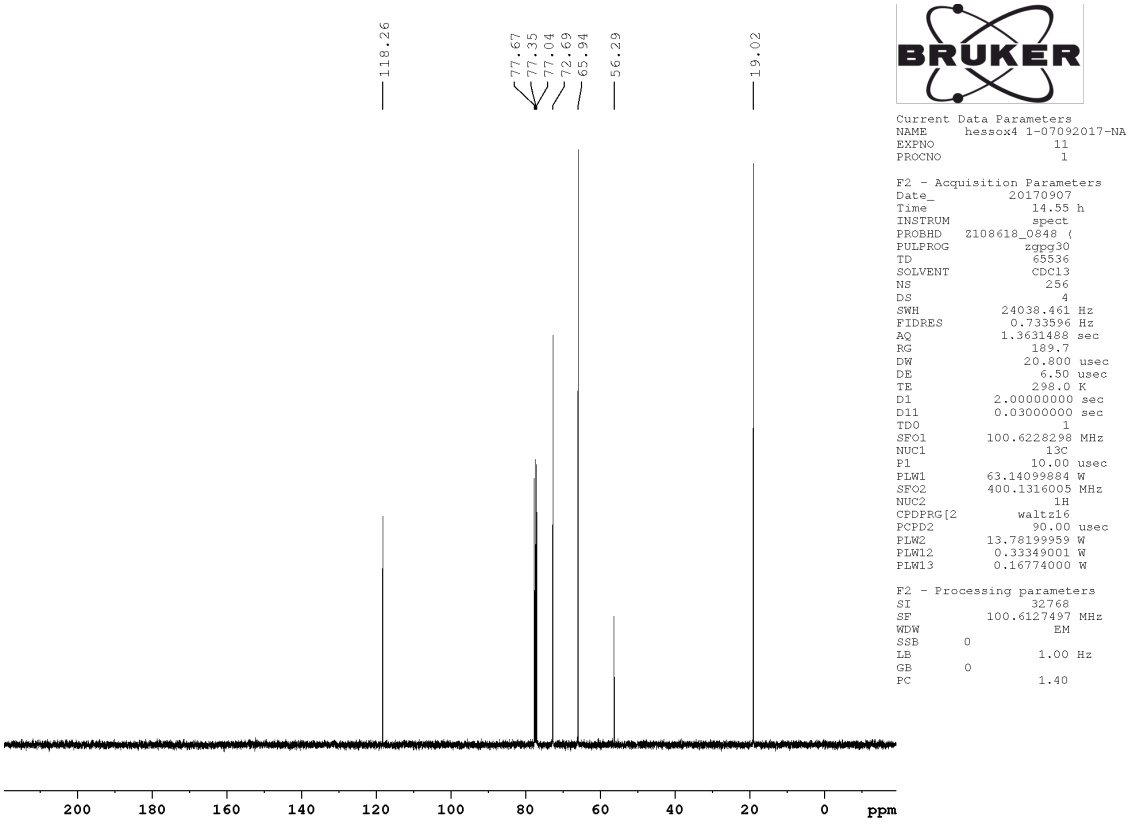


Figure S21:[13]C-NMR spectrum of compound 2 in $CDCl_3$.

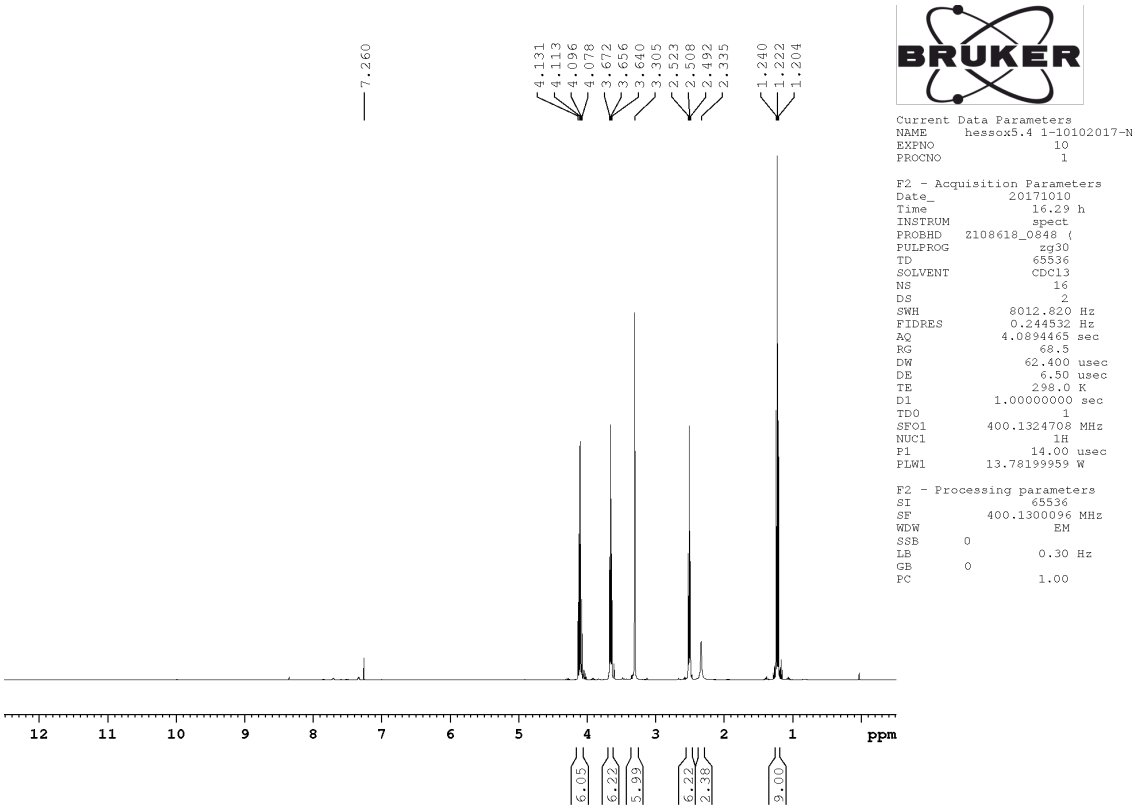


Figure S22: [1]H-NMR spectrum of compound 3 in $CDCl_3$.

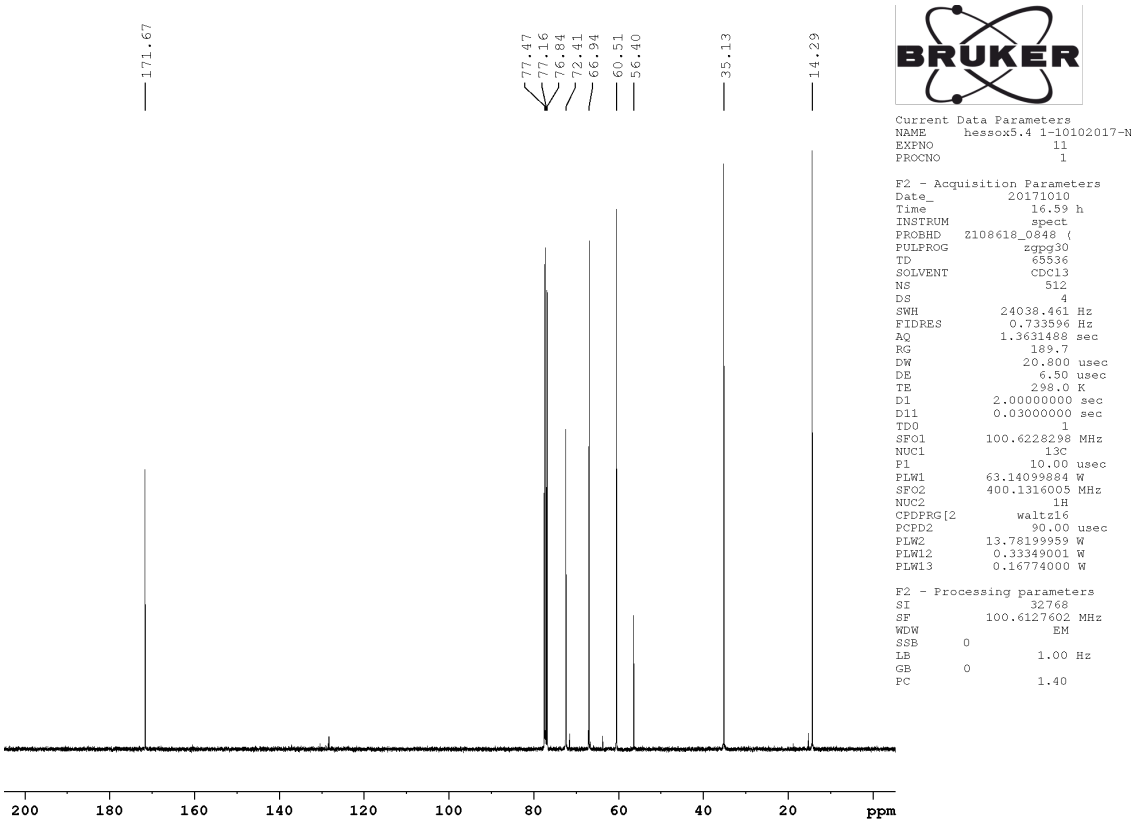


Figure S23: ^{13}C-NMR spectrum of compound 3 in $CDCl_3$.

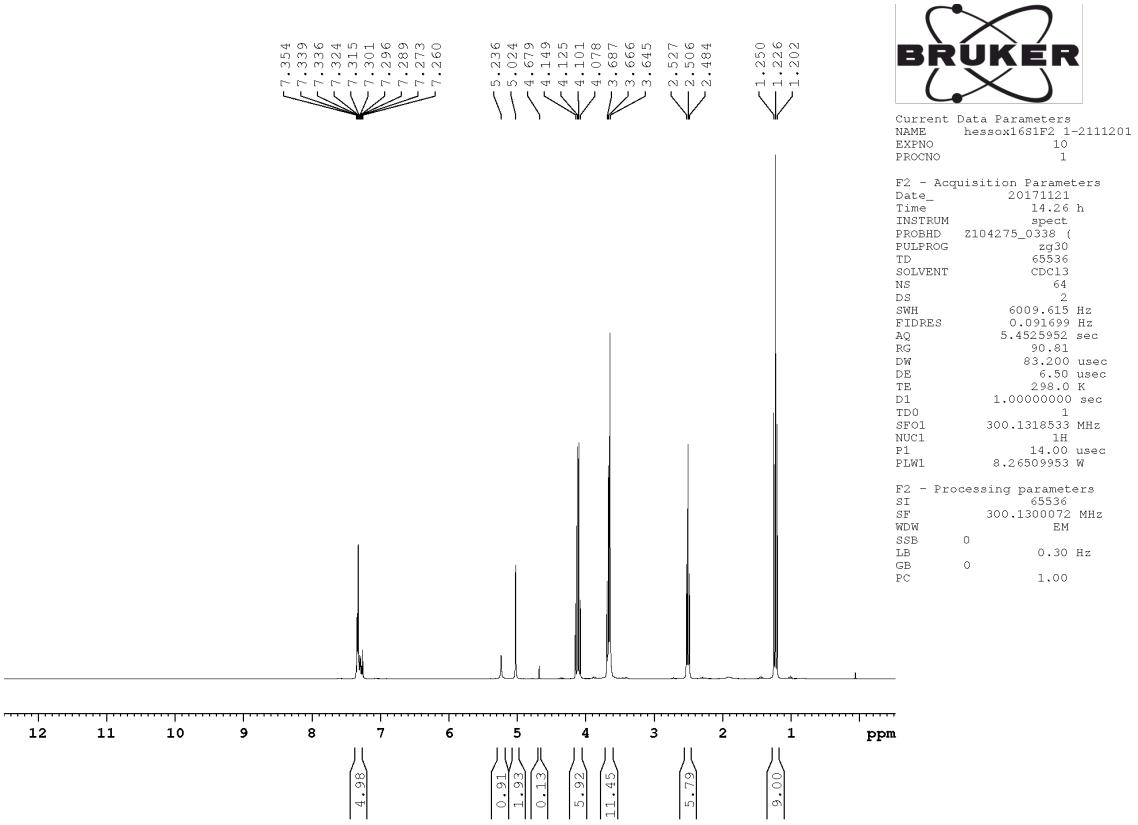


Figure S24: ^{1}H-NMR spectrum of compound 4 in $CDCl_3$.

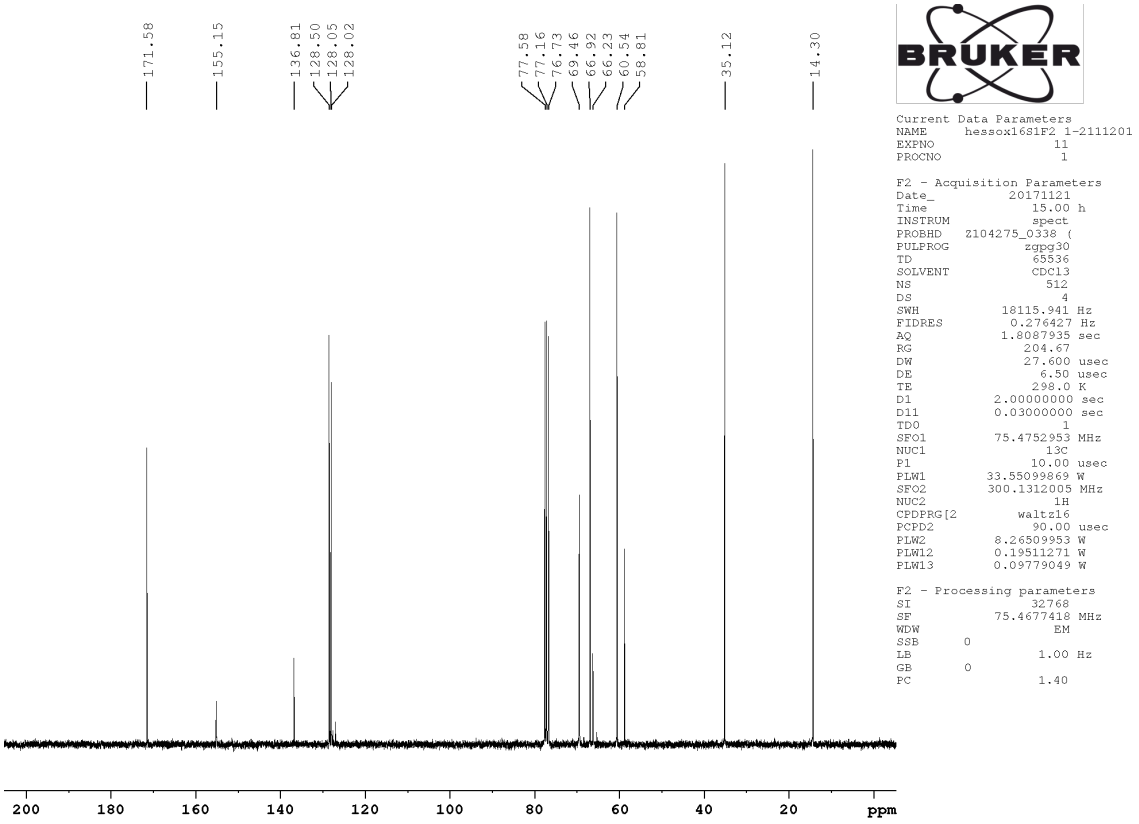


Figure S25: ^{13}C-NMR spectrum of compound 4 in $CDCl_3$.

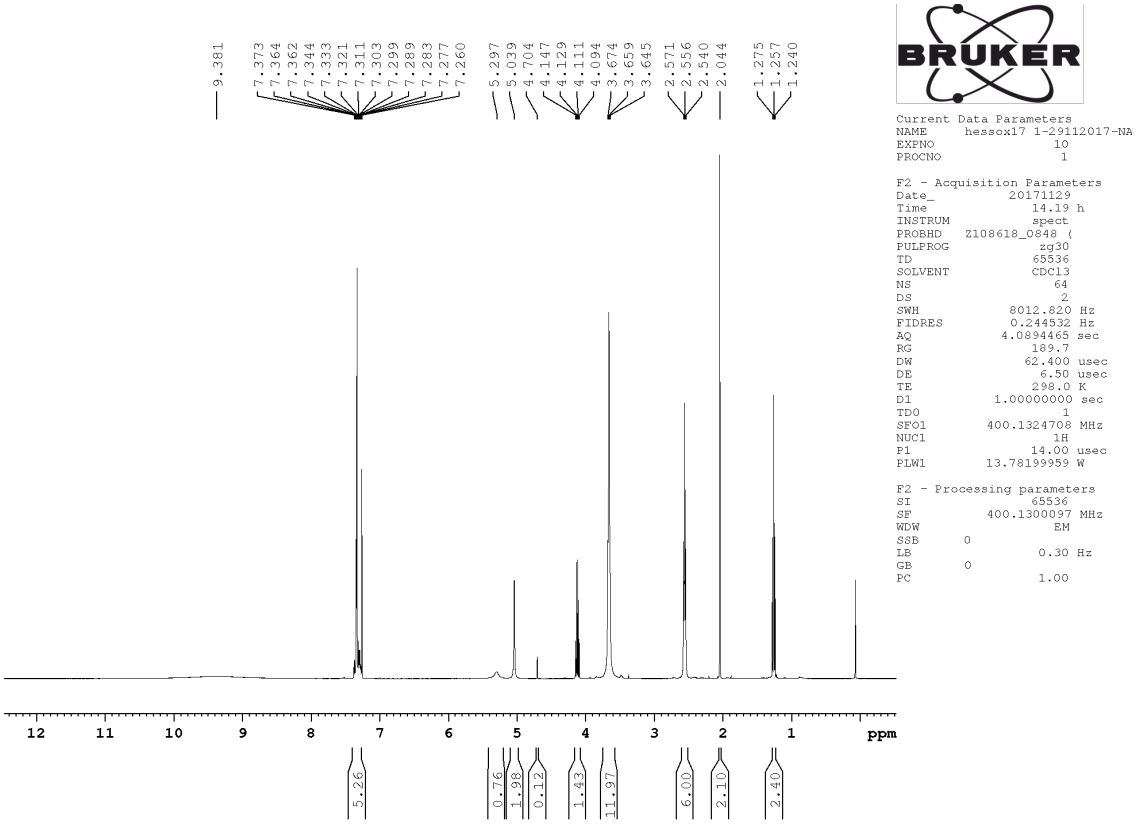


Figure S26: ^{1}H-NMR spectra of compound 5 in $CDCl_3$.

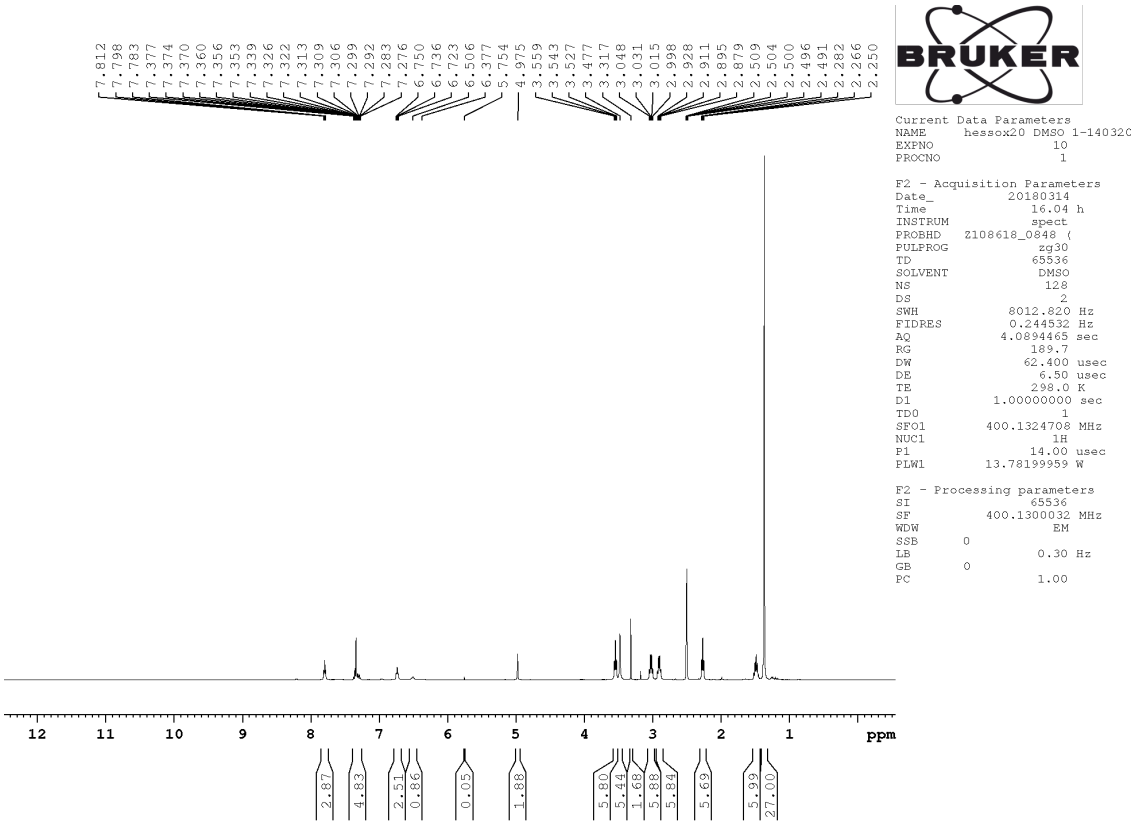


Figure S27: ^{1}H-NMR spectra of compound 6 in DMSO-d$_6$.

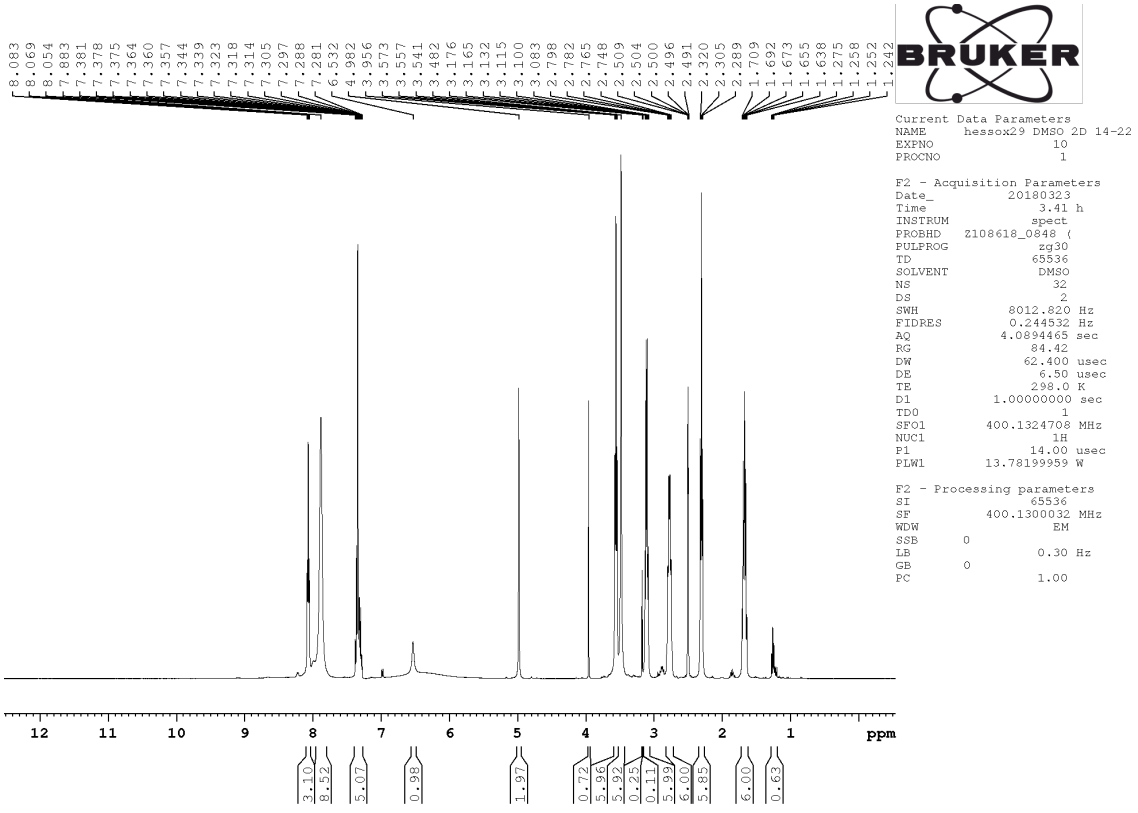


Figure S28: ^{1}H-NMR spectrum of compound 7 in DMSO-d$_6$.

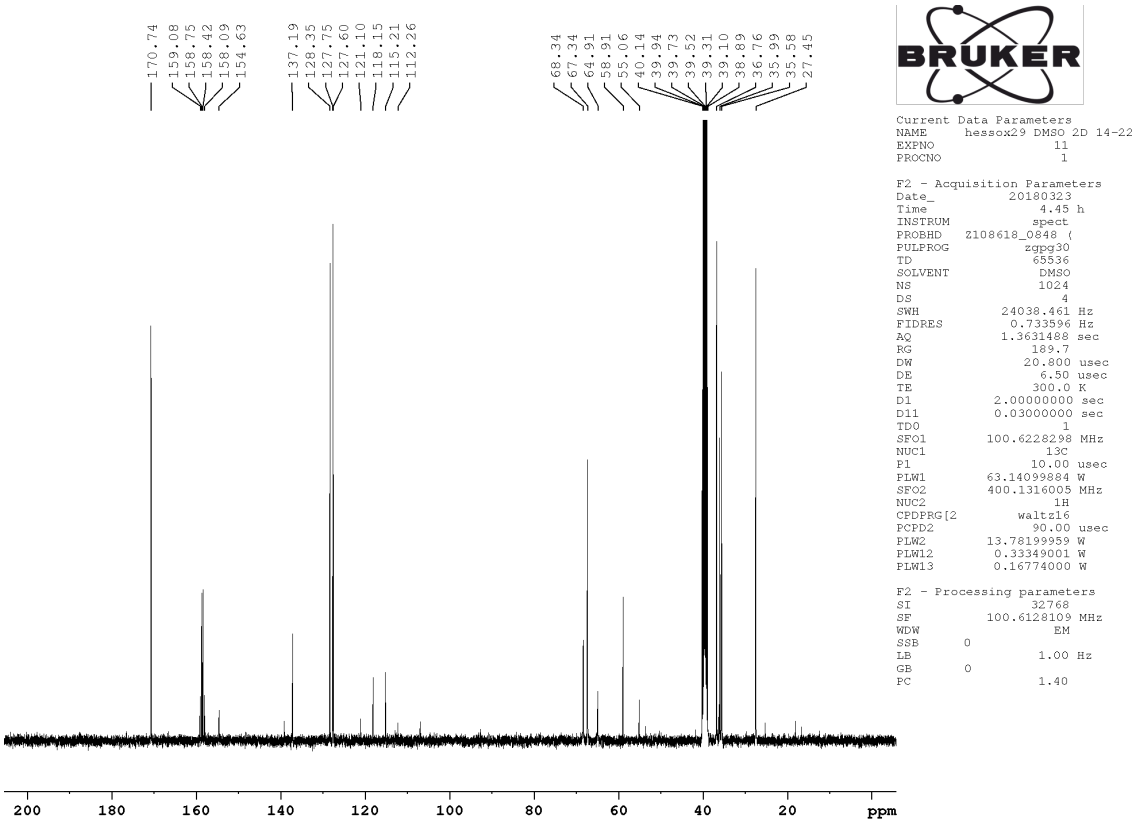


Figure S29: ^{13}C-NMR spectrum of compound 7 in DMSO-d$_6$.

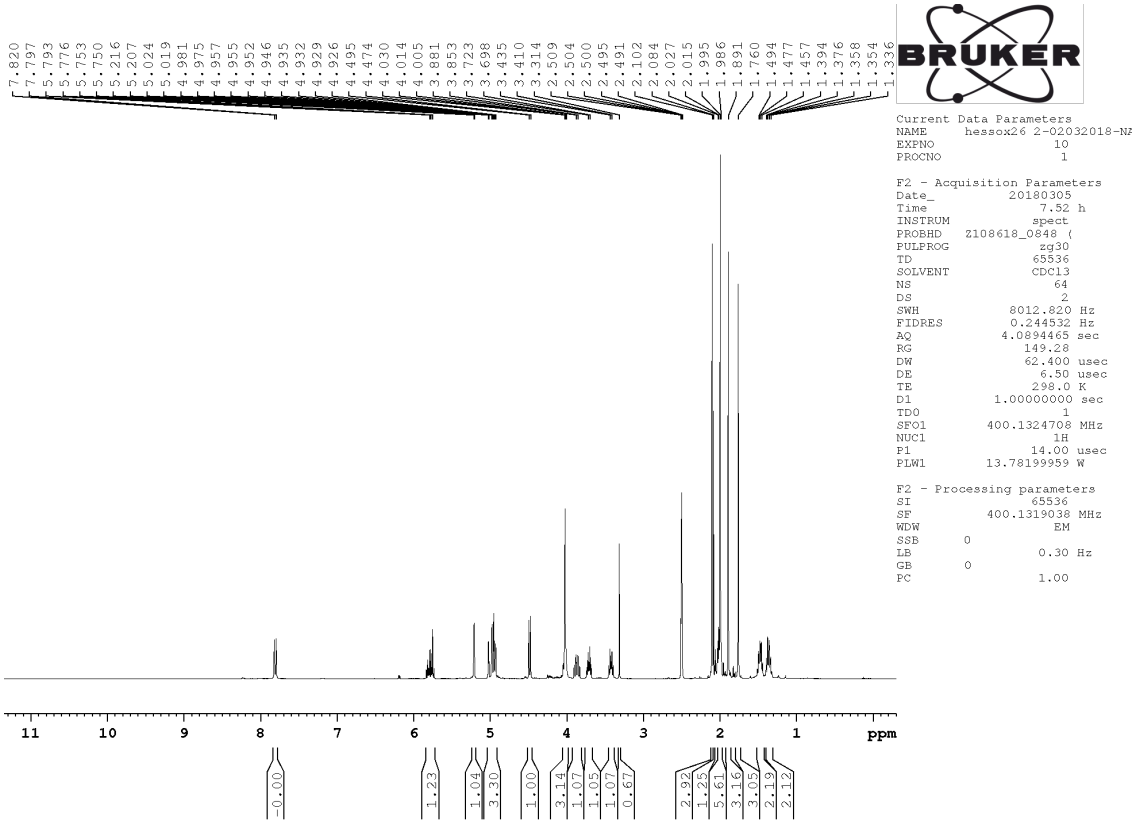


Figure S30: ^{1}H-NMR spectra of compound 9 in DMSO-d$_6$.

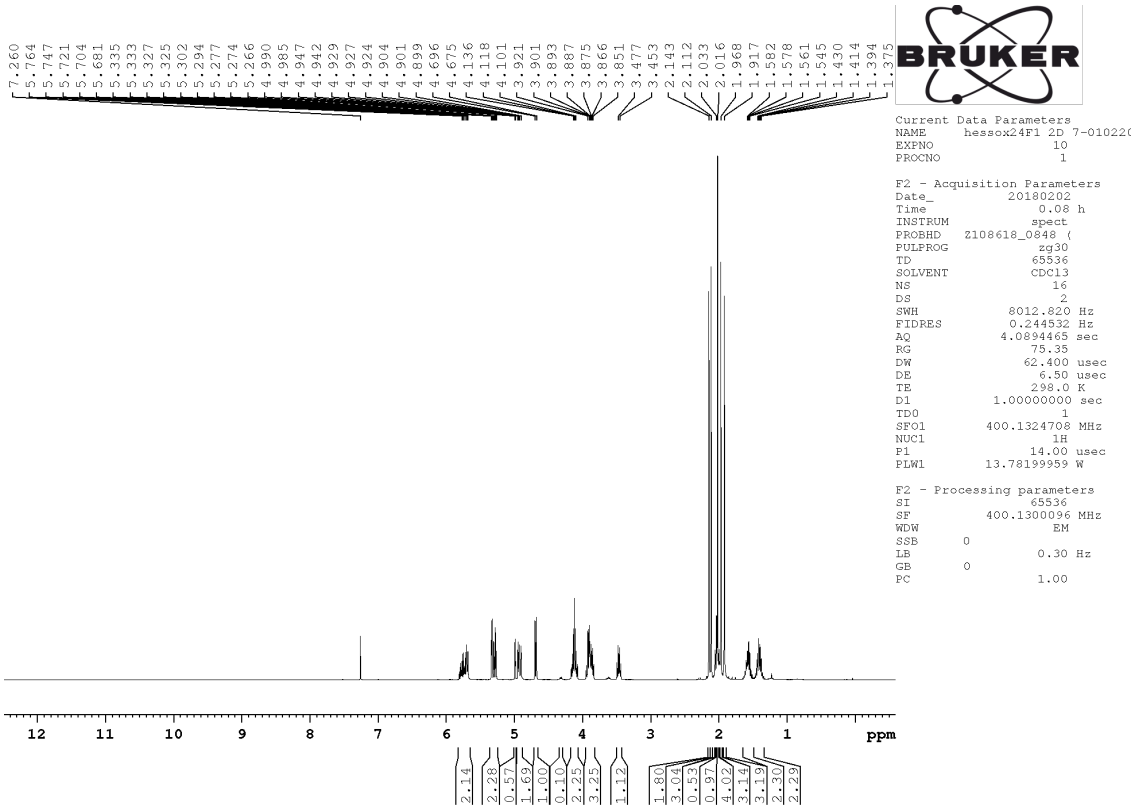


Figure S31: ^{1}H.NMR spectrum of compound 9 in $CDCl_3$.

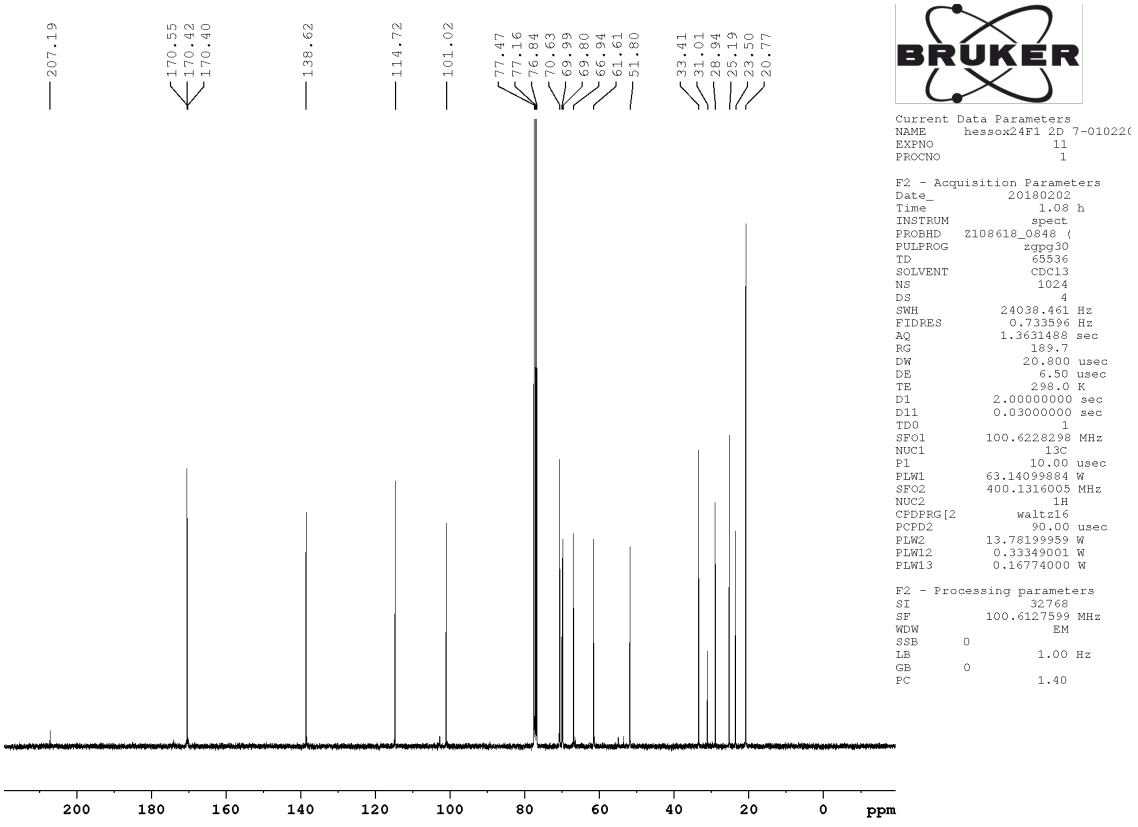


Figure S32: ^{13}C-NMR spectrum of compound 9 in $CDCl_3$.

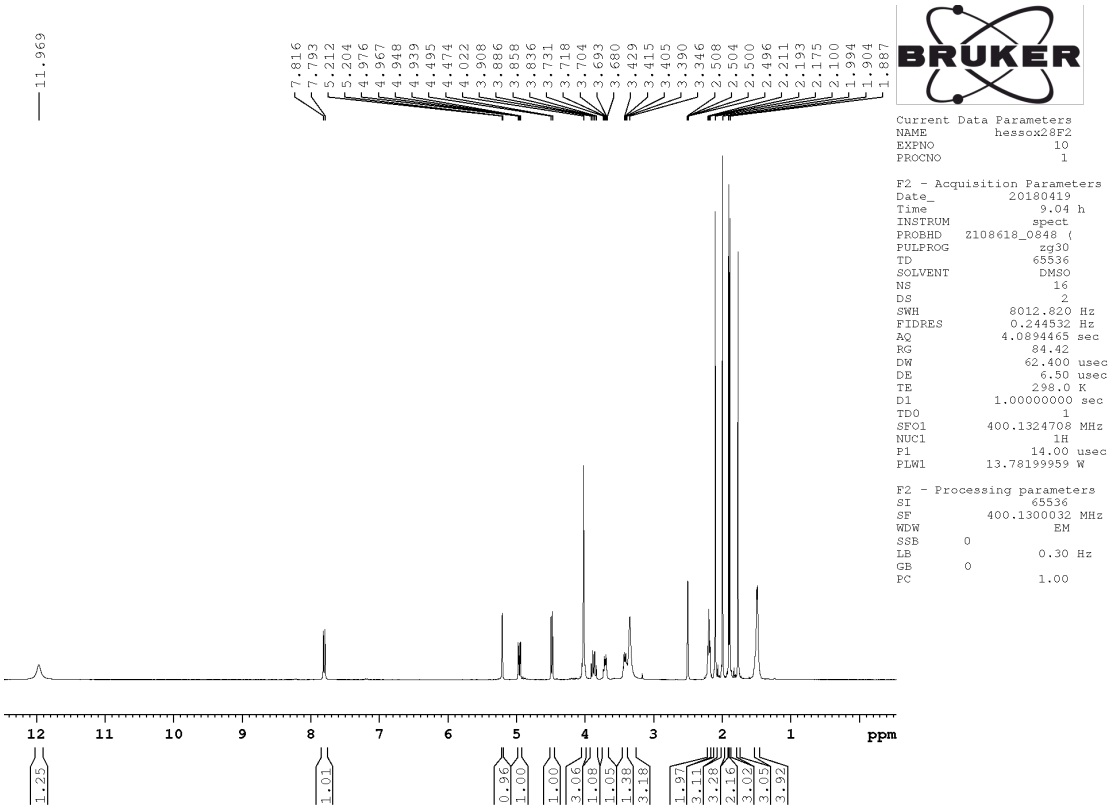


Figure S33: ^{1}H-NMR spectrum of compound 10 in DMSO-d$_6$.

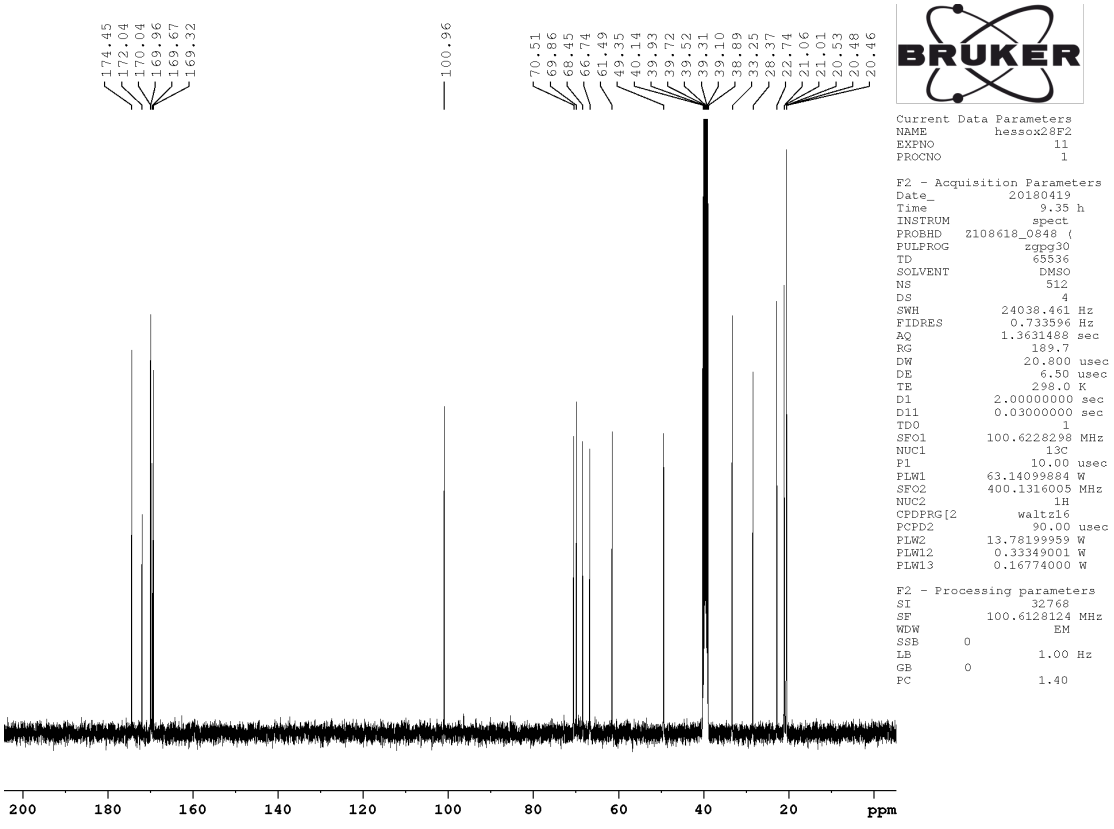


Figure S34: ^{13}C-NMR spectrum of compound 10 in DMSO-d$_6$.

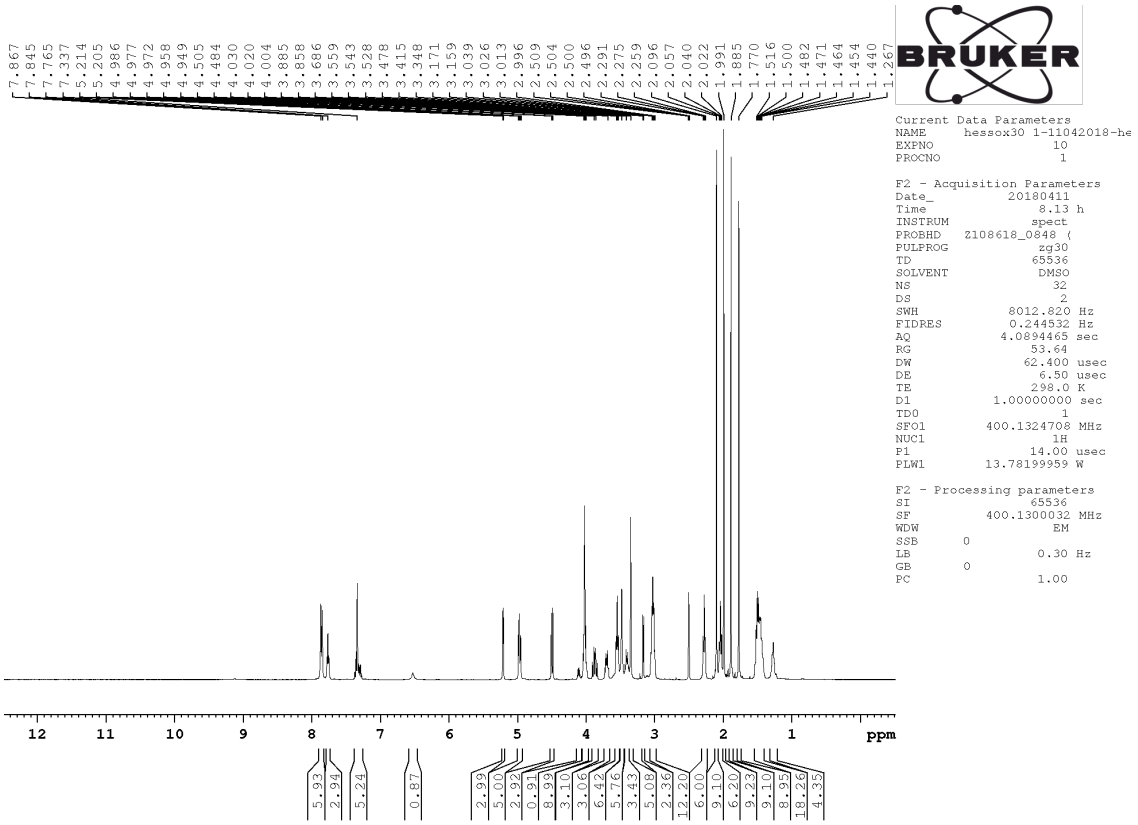


Figure S35: ^{1}H-NMR spectrum of compound 11 in DMSO-d$_6$ obtained from a small scale reaction.

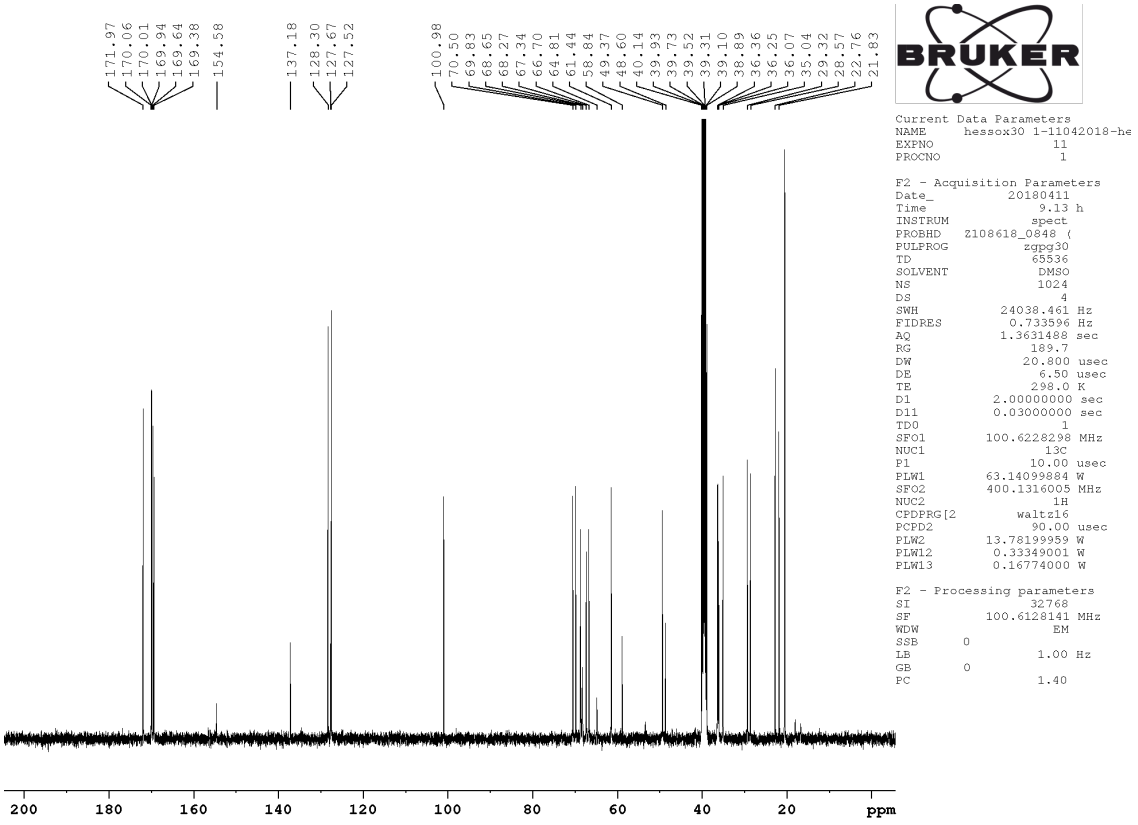


Figure S36: ^{13}C-NMR spectrum of compound 11 in DMSO-d$_6$ obtained from a small scale reaction.

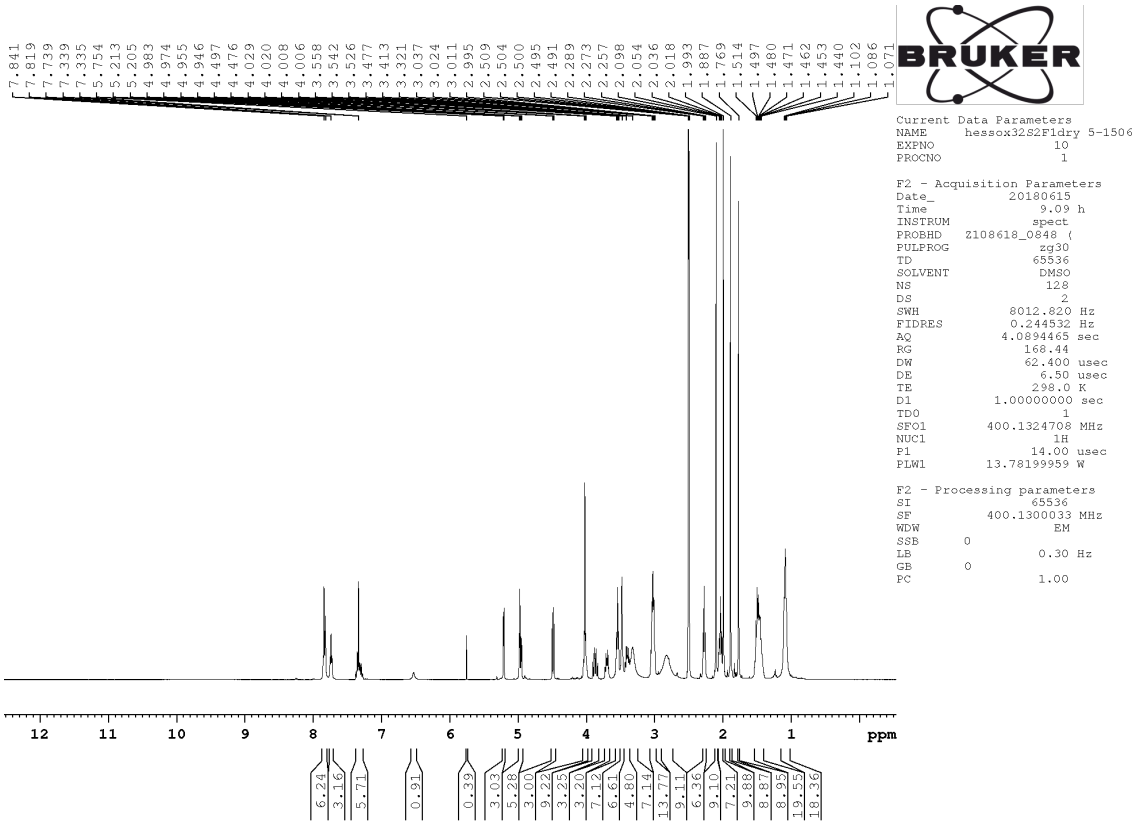


Figure S37: ^{1}H-NMR spectrum of compound 11 in DMSO-d_6 obtained from a large scale reaction.

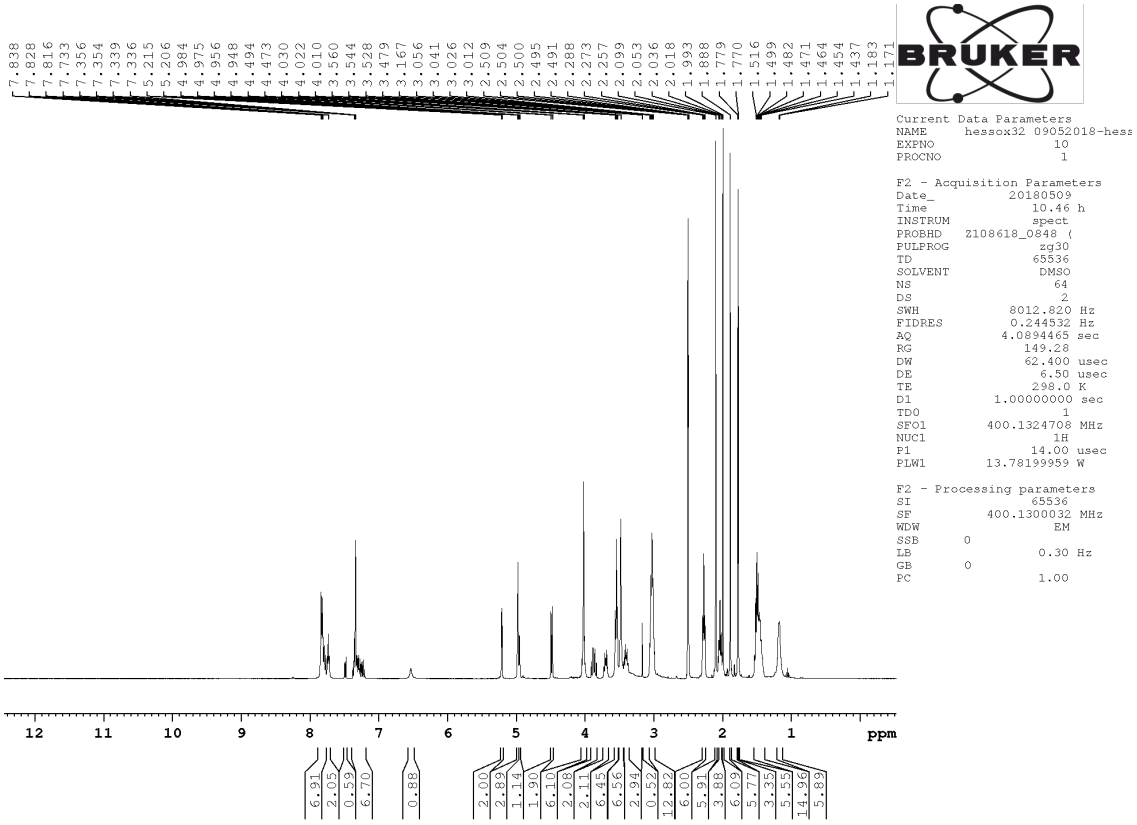


Figure S38: ^{1}H-NMR spectrum of the incomplete conjugated compound 11 in DMSO-d_6 obtained from the large scale reaction.

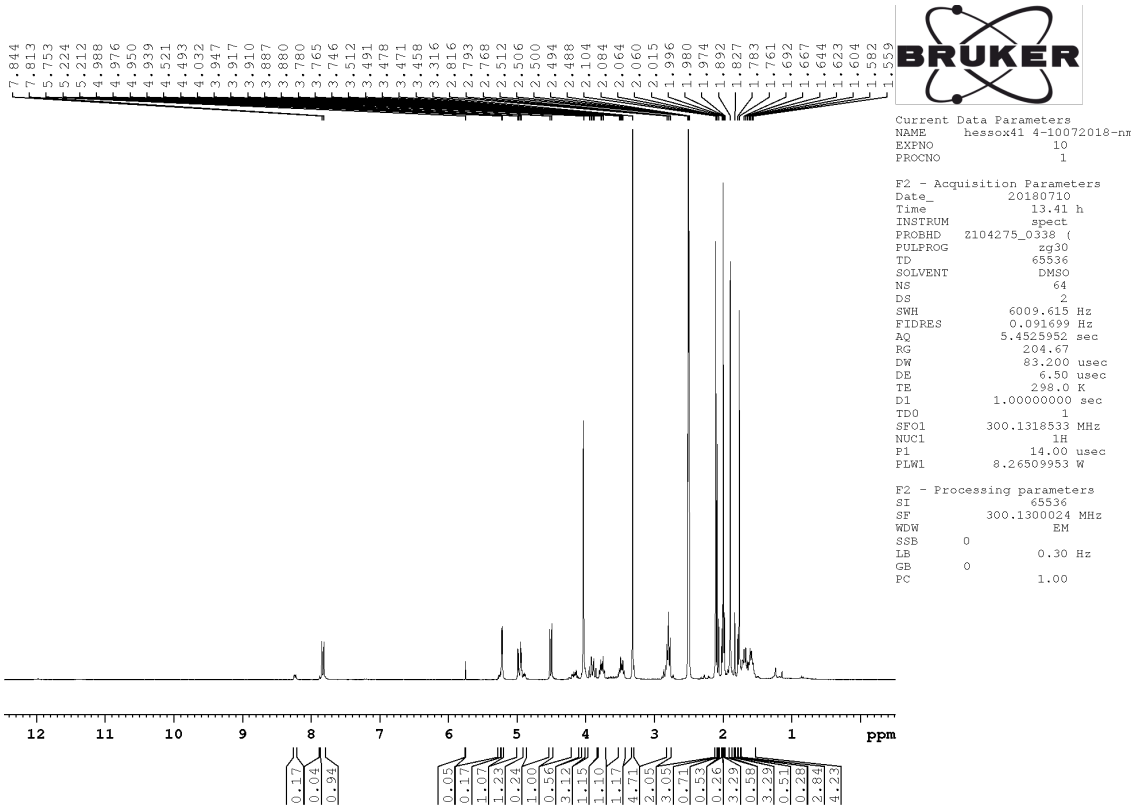


Figure S39: ^{1}H-NMR spectrum of compound 12 in DMSO-d$_6$.

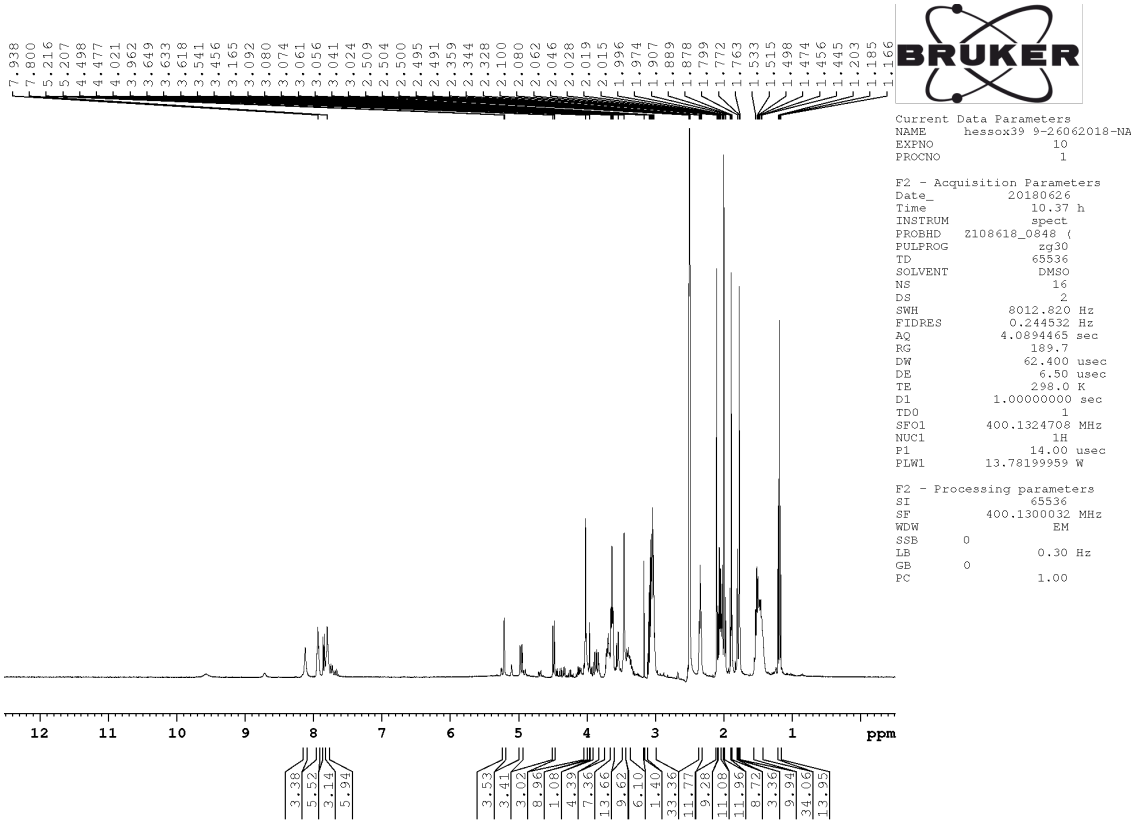


Figure S40: ^{1}H-NMR spectrum of compound 13 in DMSO-d$_6$.

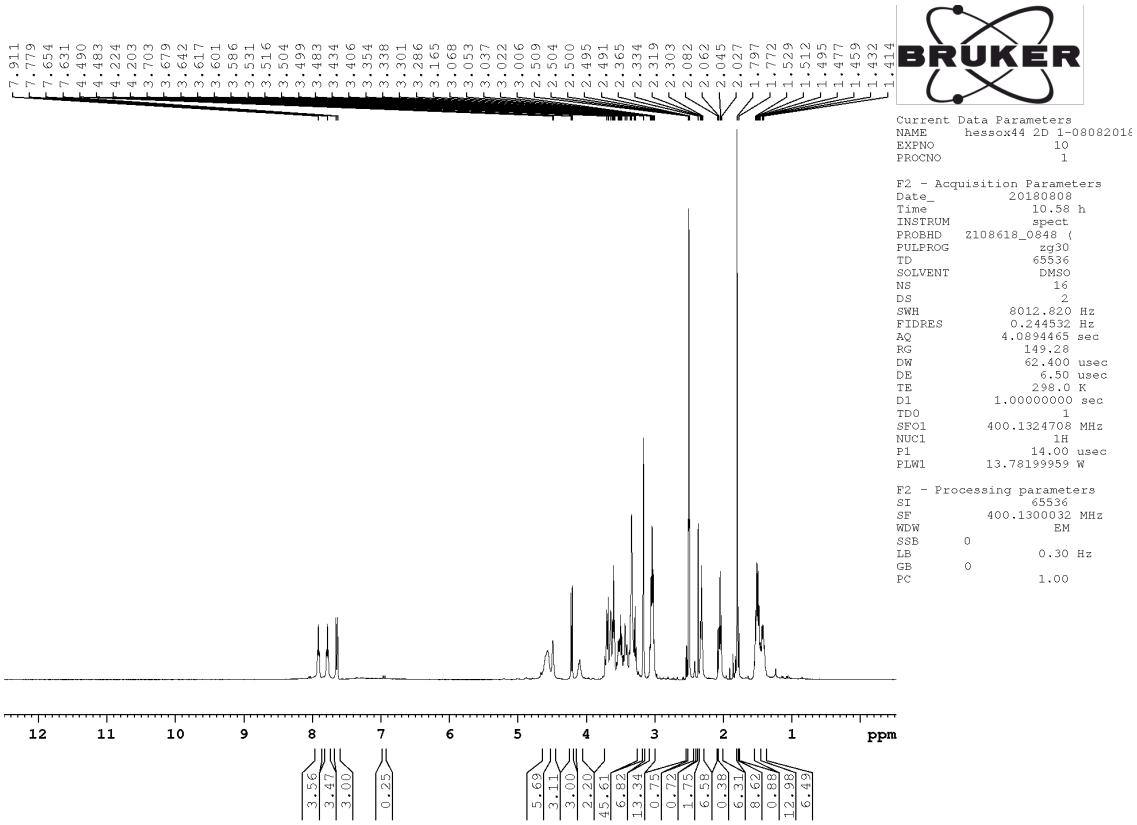


Figure S41: ^{1}H-NMR spectrum of compound 14 in DMSO-d$_6$.

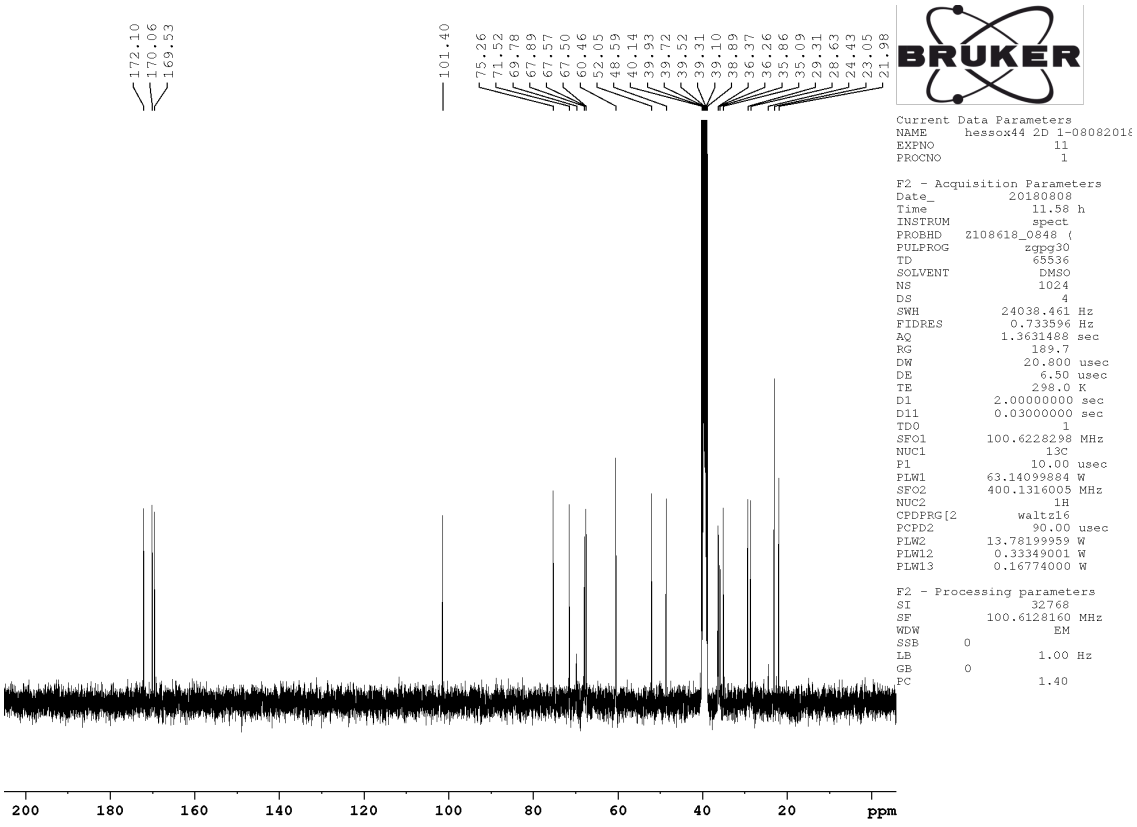


Figure S42: ^{13}C-NMR spectrum of compound 14 in DMSO-d$_6$.

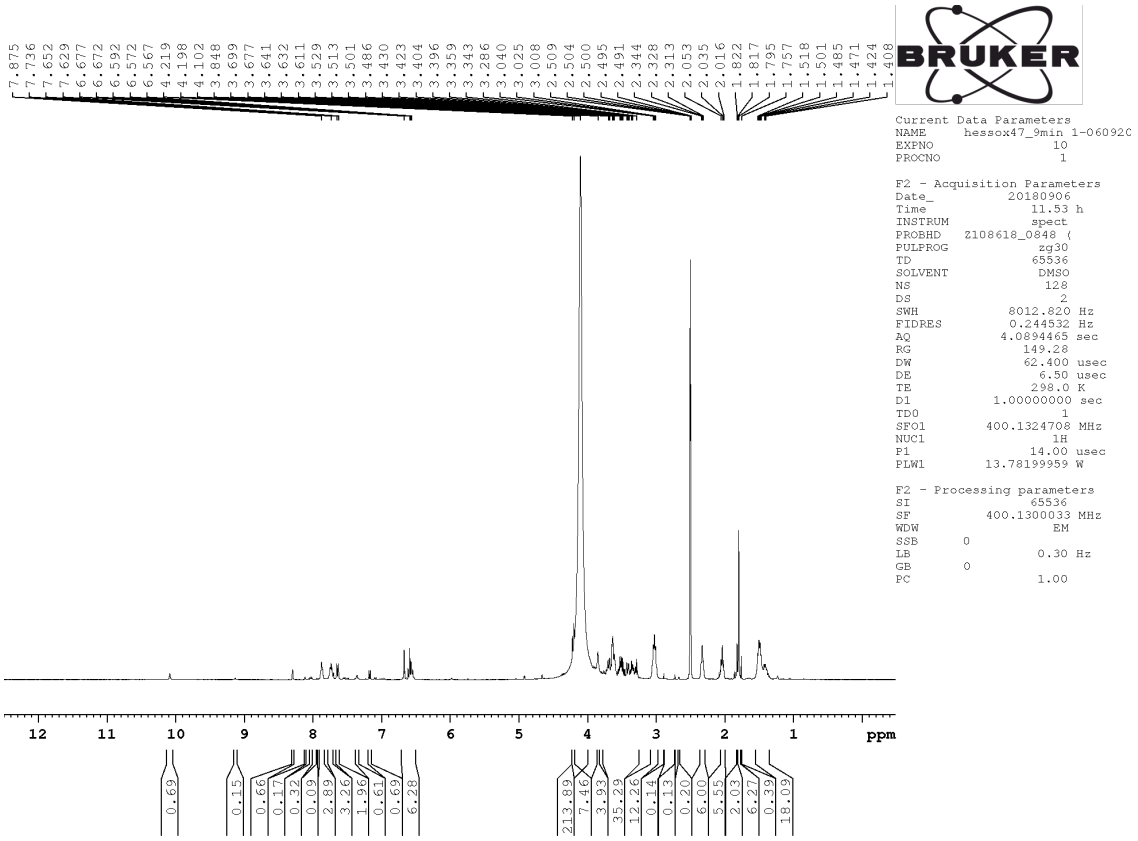


Figure S43: ^{1}H-NMR spectrum of compound 15 in DMSO-d_6.

8.1.5. High resolution mass spectra

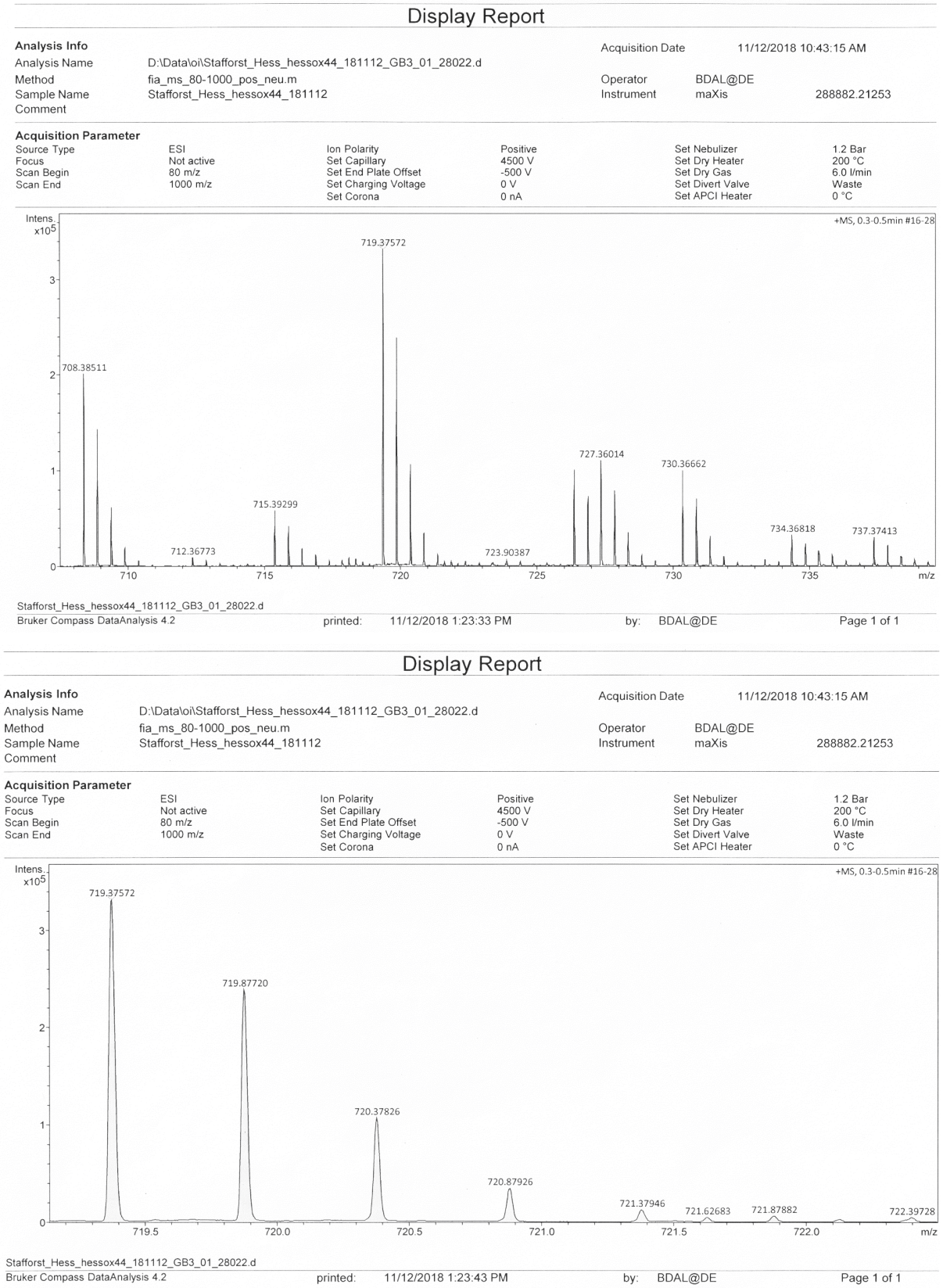


Figure S44: High resolution mass spectra of compound 14.

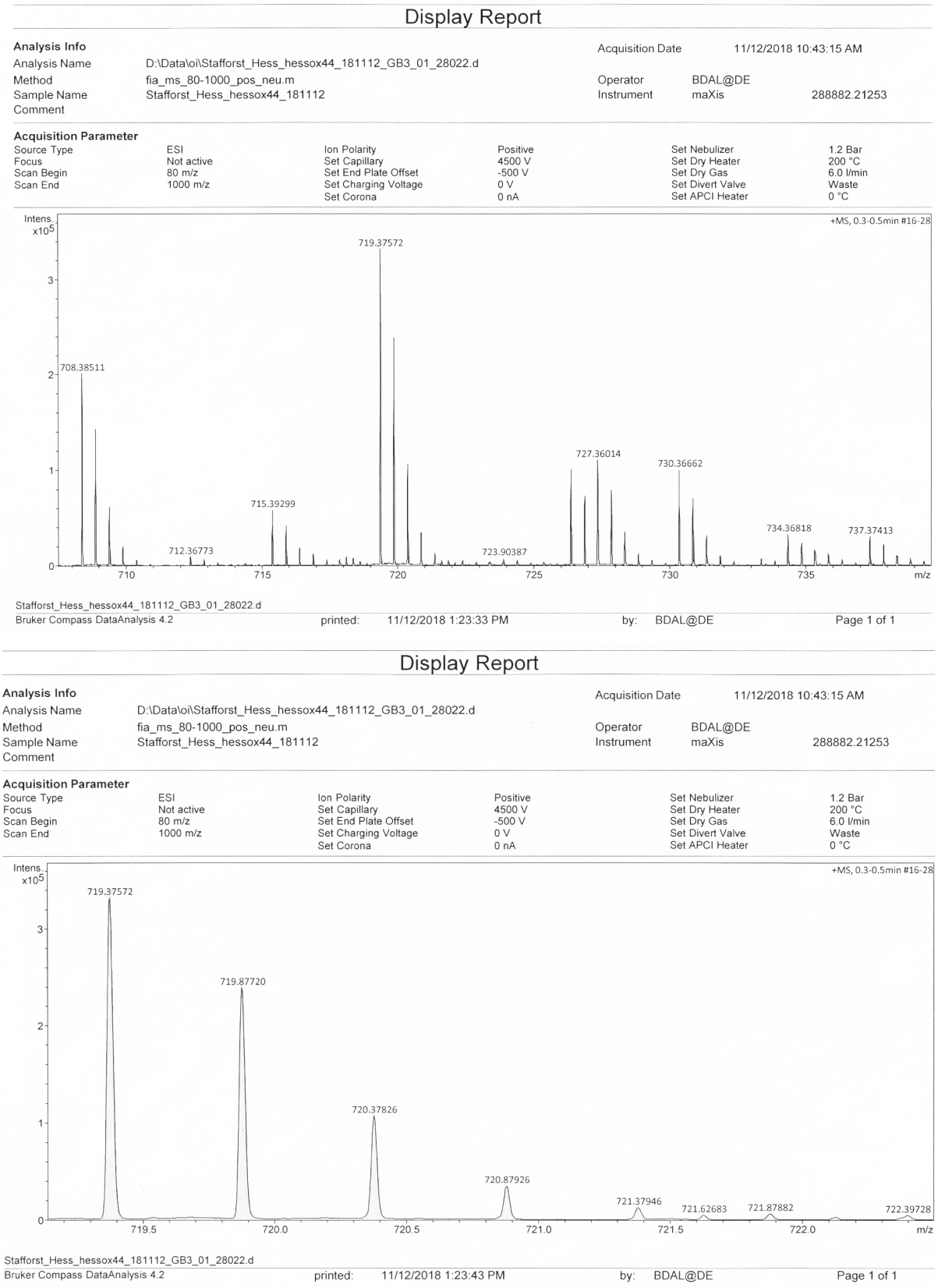


Figure S45: High resolution mass spectra of compound 15.

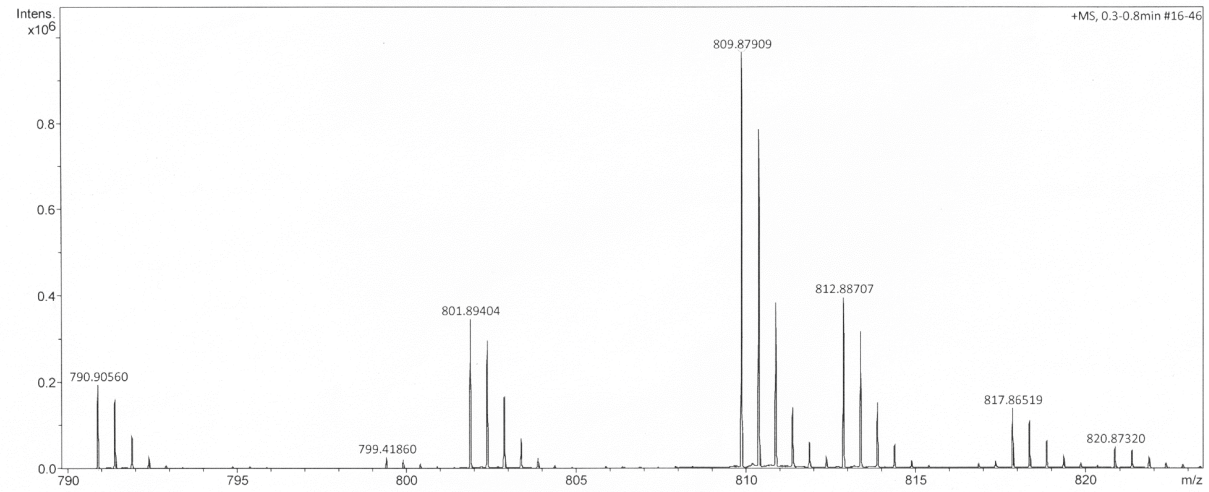
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_hessox49_181112_GB4_01_28023.d	Acquisition Date	11/12/2018 10:48:09 AM
Method	fia_ms_80-1000_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_hessox49_181112	Instrument	maXis 288882.21253
Comment			

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	80 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1000 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_hessox49_181112_GB4_01_28023.d

Bruker Compass DataAnalysis 4.2 printed: 11/12/2018 1:43:03 PM by: BDAL@DE Page 1 of 1

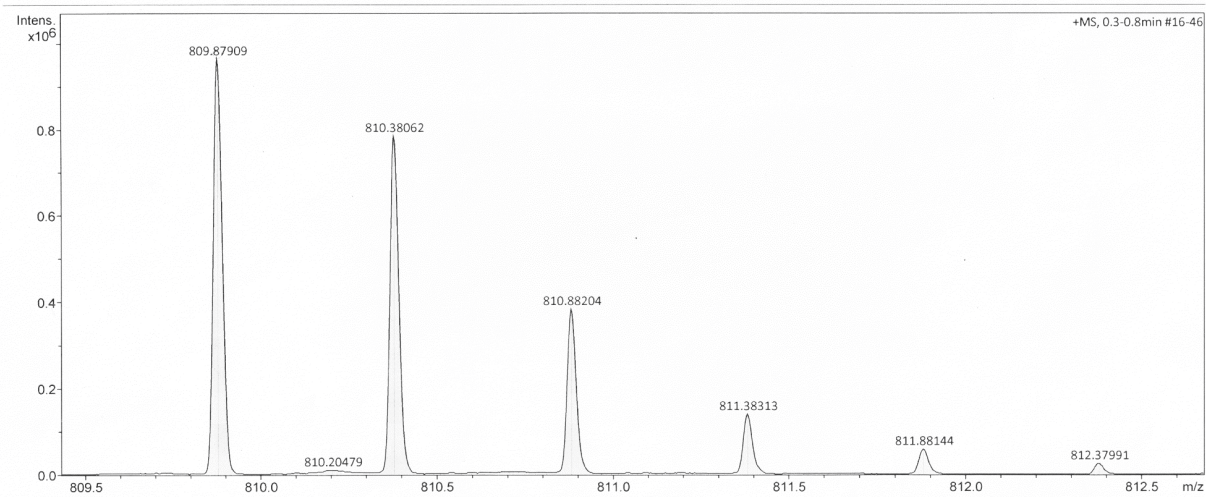
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_hessox49_181112_GB4_01_28023.d	Acquisition Date	11/12/2018 10:48:09 AM
Method	fia_ms_80-1000_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_hessox49_181112	Instrument	maXis 288882.21253
Comment			

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	80 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1000 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_hessox49_181112_GB4_01_28023.d

Bruker Compass DataAnalysis 4.2 printed: 11/12/2018 1:43:13 PM by: BDAL@DE Page 1 of 1

Figure S46: High resolution mass spectra of compound 16.

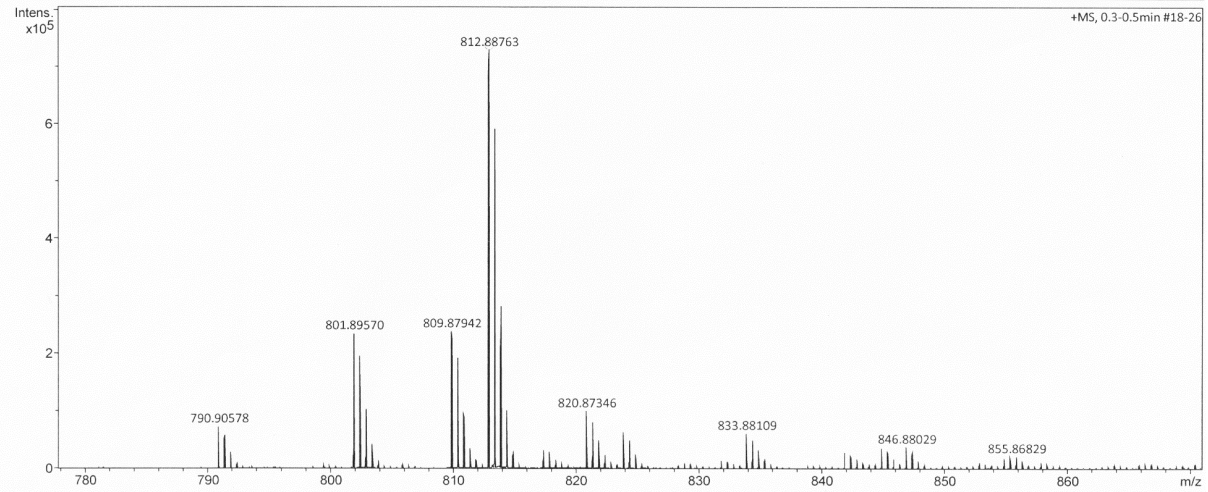
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_YS3_GC3_01_42741.d	Acquisition Date	8/6/2020 12:38:55 PM
Method	fia_ms_100-1700_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_YS3	Instrument	maXis 288882.21253
Comment	YS3		

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1700 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_YS3_GC3_01_42741.d

Bruker Compass DataAnalysis 4.2 printed: 8/7/2020 4:24:06 PM by: BDAL@DE Page 1 of 1

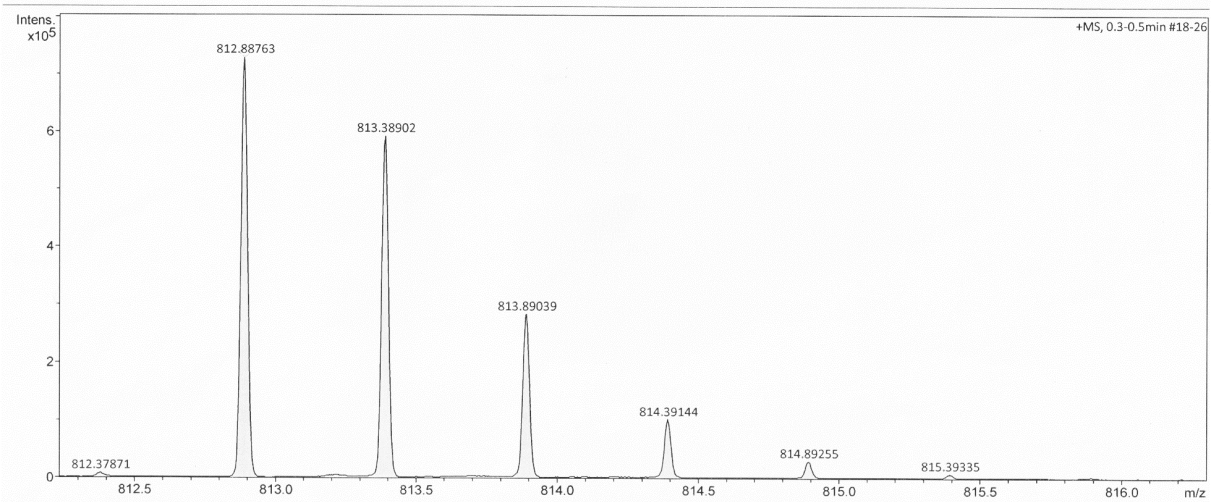
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_YS3_GC3_01_42741.d	Acquisition Date	8/6/2020 12:38:55 PM
Method	fia_ms_100-1700_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_YS3	Instrument	maXis 288882.21253
Comment	YS3		

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1700 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_YS3_GC3_01_42741.d

Bruker Compass DataAnalysis 4.2 printed: 8/7/2020 4:24:17 PM by: BDAL@DE Page 1 of 1

Figure S47: High resolution mass spectra of compound 16.

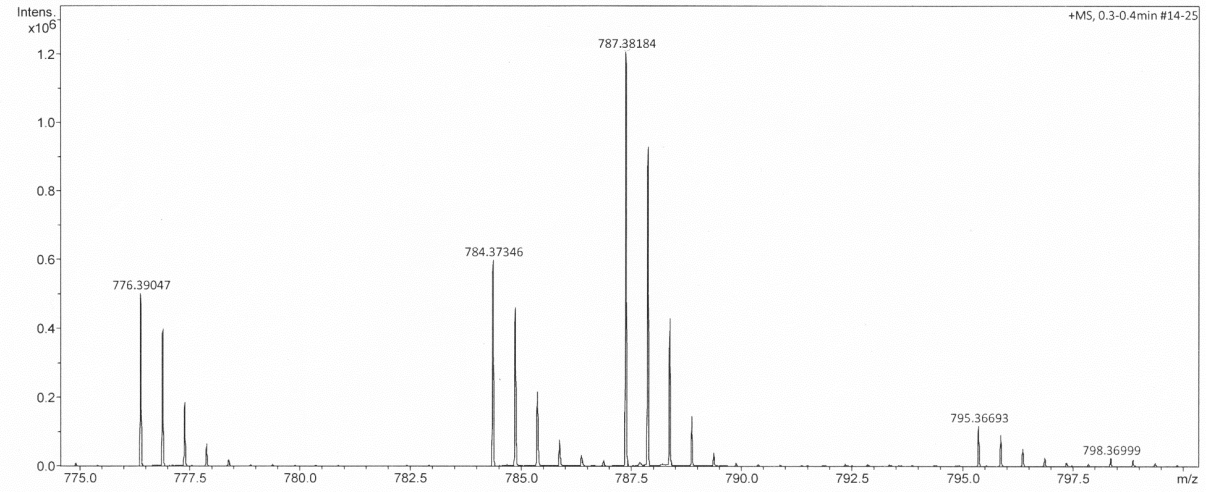
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_hessox77_GE3_01_40952.d	Acquisition Date	5/26/2020 5:44:56 PM
Method	fia_ms_100-1700_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_hessox77	Instrument	maXis 288882.21253
Comment			

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1700 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_hessox77_GE3_01_40952.d

Bruker Compass DataAnalysis 4.2 printed: 5/26/2020 5:53:50 PM by: BDAL@DE Page 1 of 1

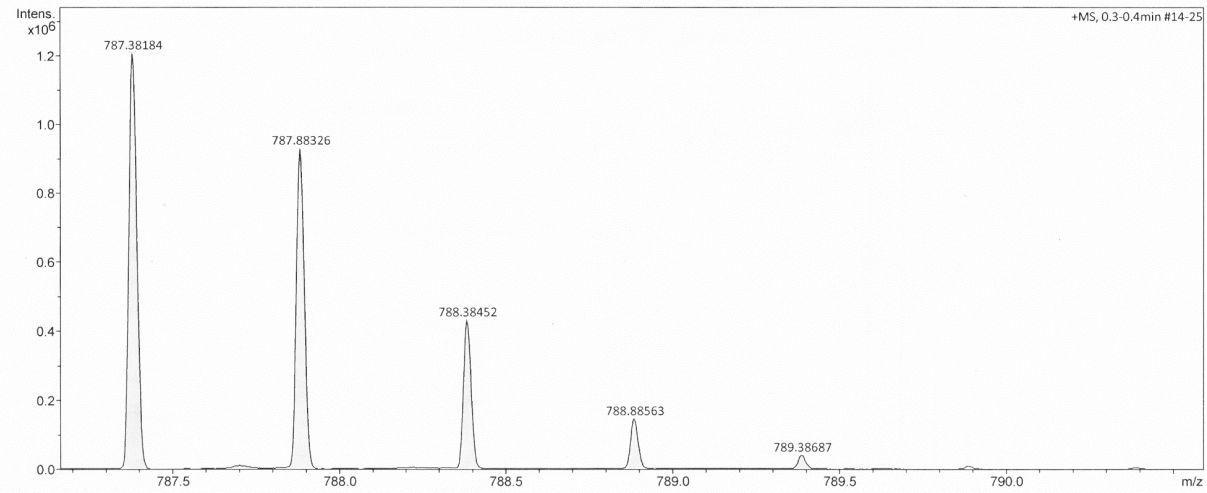
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_hessox77_GE3_01_40952.d	Acquisition Date	5/26/2020 5:44:56 PM
Method	fia_ms_100-1700_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_hessox77	Instrument	maXis 288882.21253
Comment			

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1700 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_hessox77_GE3_01_40952.d

Bruker Compass DataAnalysis 4.2 printed: 5/26/2020 5:54:01 PM by: BDAL@DE Page 1 of 1

Figure S48: High resolution mass spectra of compound 17.

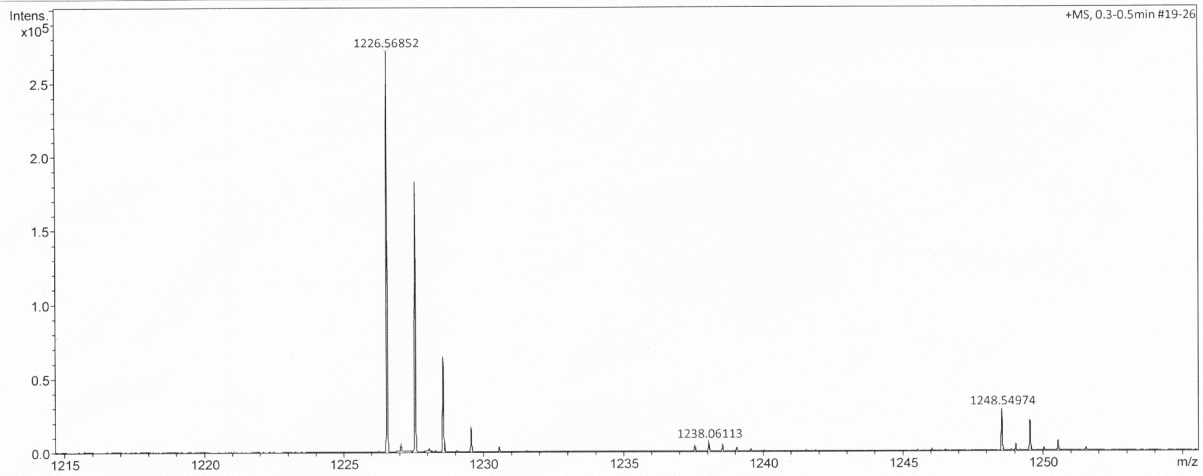
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_BisBG-Lin-YS1_2_GC2_01_42739.d	Acquisition Date	8/6/2020 12:18:48 PM
Method	fia_ms_100-1700_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_BisBG-Lin-YS1_2	Instrument	maXis 288882.21253
Comment	Stafforst Oliver Hess BisBG-Linker-YS1.2		

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1700 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_BisBG-Lin-YS1_2_GC2_01_42739.d
Bruker Compass DataAnalysis 4.2 printed: 8/6/2020 4:03:21 PM by: BDAL@DE Page 1 of 1

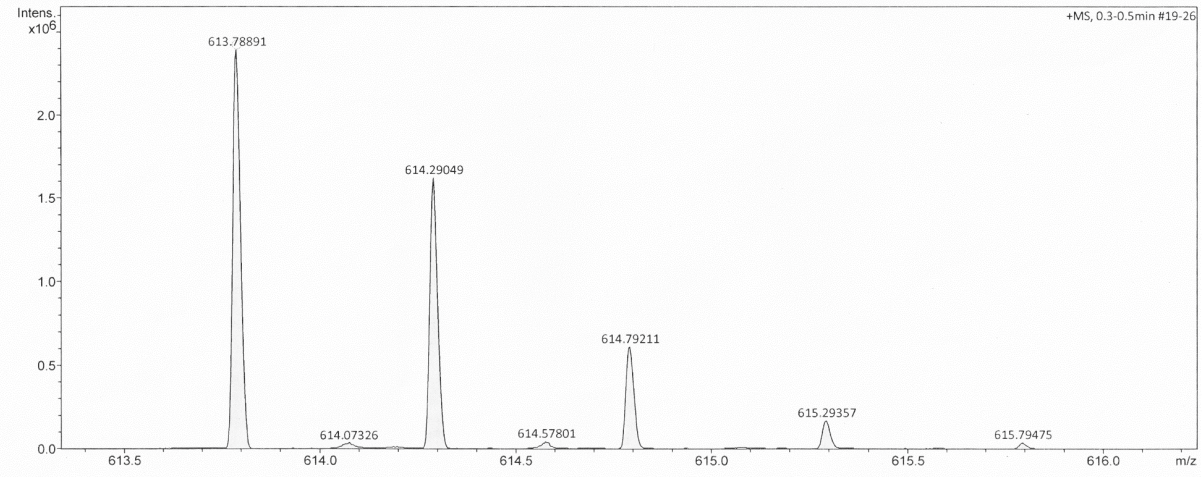
Display Report

Analysis Info

Analysis Name	D:\Data\oi\Stafforst_Hess_BisBG-Lin-YS1_2_GC2_01_42739.d	Acquisition Date	8/6/2020 12:18:48 PM
Method	fia_ms_100-1700_pos_neu.m	Operator	BDAL@DE
Sample Name	Stafforst_Hess_BisBG-Lin-YS1_2	Instrument	maXis 288882.21253
Comment	Stafforst Oliver Hess BisBG-Linker-YS1.2		

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.2 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	6.0 l/min
Scan End	1700 m/z	Set Charging Voltage	0 V	Set Divert Valve	Waste
		Set Corona	0 nA	Set APCI Heater	0 °C



Stafforst_Hess_BisBG-Lin-YS1_2_GC2_01_42739.d
Bruker Compass DataAnalysis 4.2 printed: 8/6/2020 4:04:03 PM by: BDAL@DE Page 1 of 1

Figure S49: High resolution mass spectra of compound 18 (BisBG)

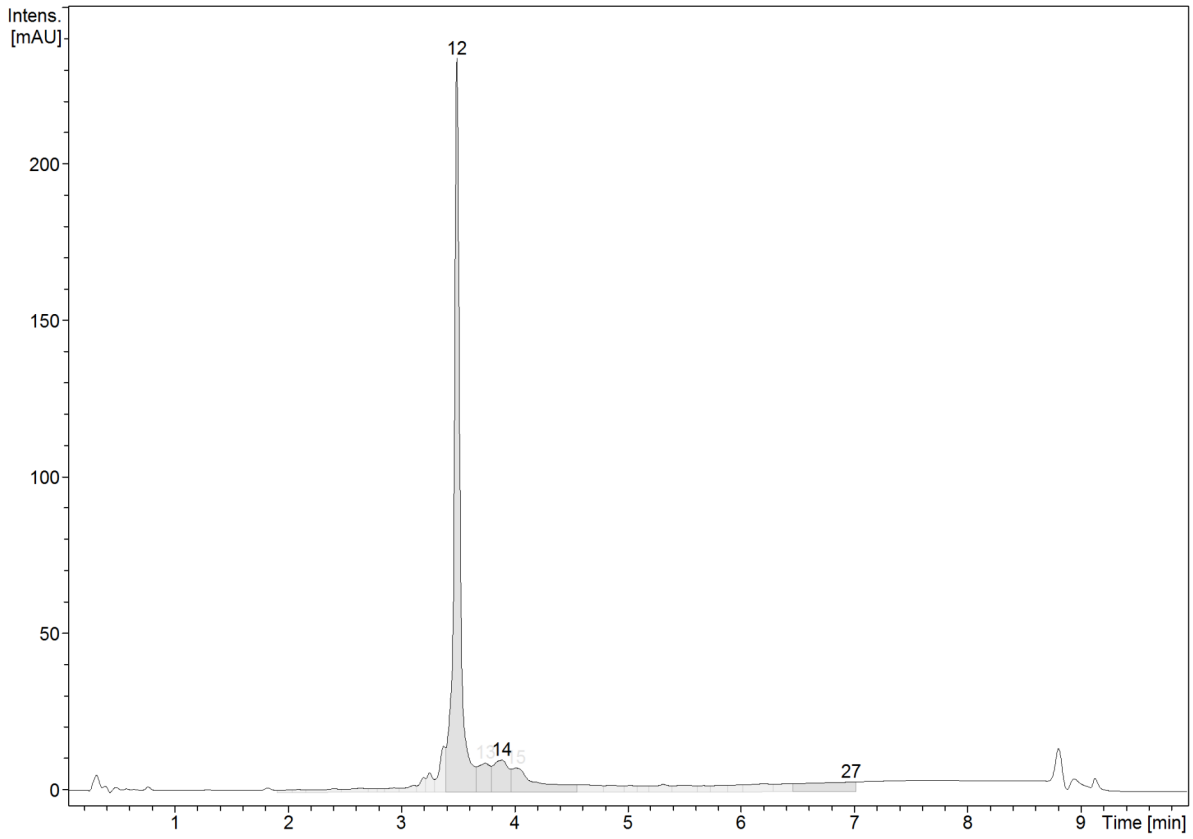
8.1.6. Mass spectrometry reports

Compound Spectrum List Report

BIOSPRING The Oligo Company

Analysis Info

Analysis Name	D:\Data\LCMS X\NON-GMP\Prod\2021\05_Mai\282214_Mod_1_GD5_01_11608.d	Acquisition Date	5/21/2021 12:16:04 PM
Method	uplc_41_ms_3000_13000_11608.m	Operator	lab
Sample Name	282214_Mod_1	Instrument	amaZon SL
Comment			



RT [min]	#	Area	Area Frac. %	S/N
3.5	12	875.1542	51.53	274209.2
3.7	13	70.9020	4.17	10845.1
3.9	14	97.5514	5.74	11972.5
4.0	15	128.3800	7.56	8990.6
7.0	27	94.5648	5.57	3650.4

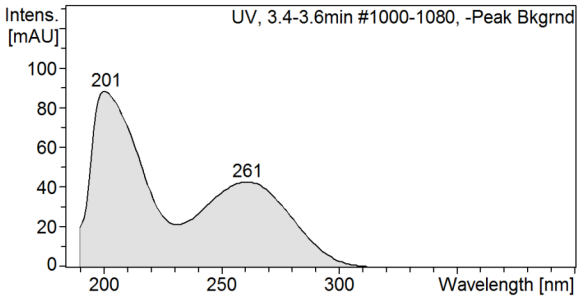
Bruker Compass DataAnalysis 4.4 printed: 5/21/2021 12:26:25 PM by: lab Page 1 of 6

Figure S50: Mass spectrometry report of gRNA BisBG-471 obtained from BioSpring GmbH. *Only pages 1 and 2 are shown within the supplementary information to confirm the purity as well as identity of the desired compound. The data of the other fractions are not shown.*

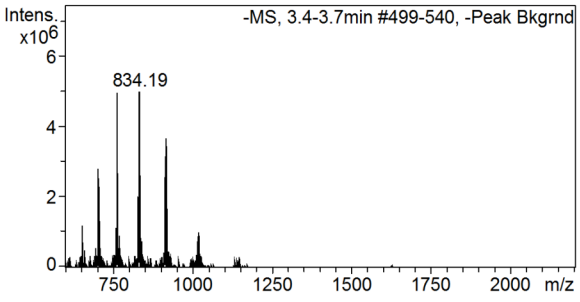
Compound Spectrum List Report

BIOSPRING The Oligo Company

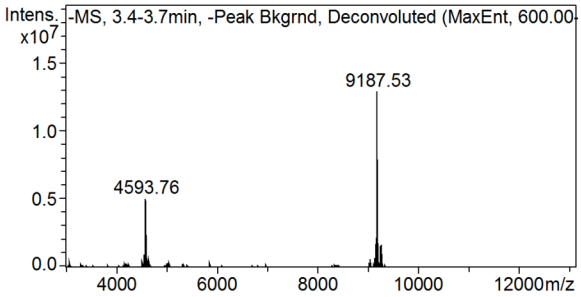
Cmpd 12, 3.5 min



#	Wavelength	Intensity
0	201	88.3
1	261	43.0



#	m/z	I	S/N
1	655.24	1190053	16.6
2	704.64	1528641	21.3
3	705.80	2804905	39.2
4	762.00	1125179	15.7
5	763.33	2497901	34.9
6	764.62	4954203	69.2
7	832.76	1967272	27.5
8	834.19	4992287	69.7
9	917.75	3650146	51.0
10	1019.81	1011315	14.1



#	m/z	I	S/N
1	4570.00	523051	27.5
2	4577.37	942435	49.5
3	4585.74	1041958	54.8
4	4593.76	5077562	266.8
5	9052.79	628789	33.0
6	9138.87	675671	35.5
7	9155.44	1537260	80.8
8	9171.63	5946621	312.5
9	9187.53	12908669	678.3
10	9258.58	1580428	83.0

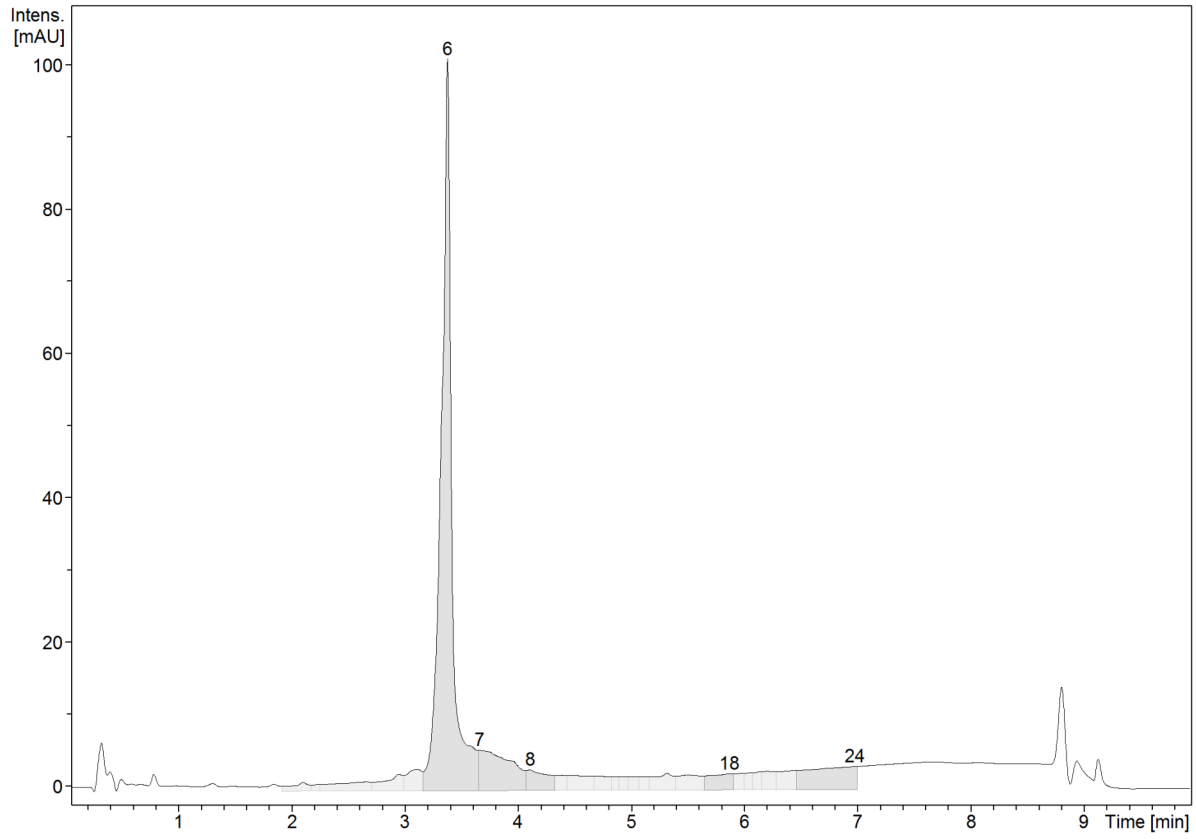
Figure S50: Continued.

Compound Spectrum List Report

BIOSPRING The Oligo Company

Analysis Info

Acquisition Date	5/21/2021 12:28:42 PM		
Analysis Name	D:\Data\LCMS X\NON-GMP\Prod\2021\05_Mai\282214_Mod_2_GE5_01_11609.d		
Method	uplc_41_ms_3000_13000_11609.m	Operator	lab
Sample Name	282214_Mod_2	Instrument	amaZon SL
Comment			



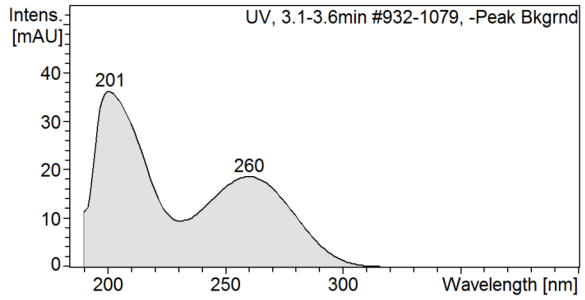
RT [min]	#	Area	Area Frac. %	S/N
3.4	6	673.1158	52.18	118597.8
3.7	7	111.2826	8.63	6479.0
4.1	8	37.3875	2.90	3325.1
5.9	18	32.6568	2.53	2602.4
7.0	24	94.3430	7.31	3751.4

Figure S51: Mass spectrometry report of gRNA BisBG-471-GalNAc obtained from BioSpring GmbH. *Only pages 1 and 2 are shown within the supplementary information to confirm the purity as well as identity of the desired compound. The data of the other fractions are not shown.*

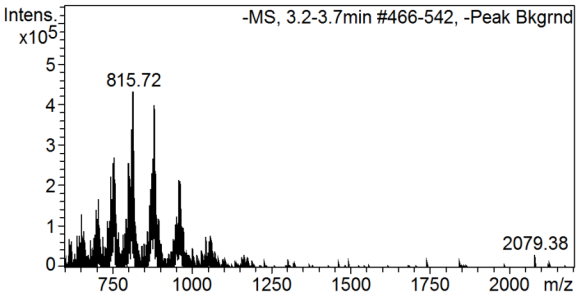
Compound Spectrum List Report

BIOSPRING The Oligo Company

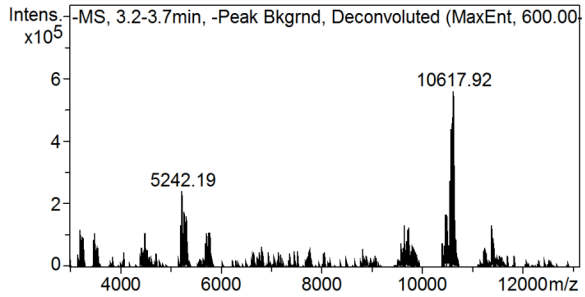
Cmpd 6, 3.4 min



#	Wavelength	Intensity
0	201	36.2
1	260	18.7



#	m/z	I	S/N
1	753.92	257487	9.7
2	757.54	269602	10.2
3	812.08	267360	10.1
4	813.27	306572	11.6
5	814.72	338346	12.8
6	815.72	432312	16.3
7	816.98	326057	12.3
8	881.07	267385	10.1
9	882.64	293191	11.1
10	883.77	399891	15.1



#	m/z	I	S/N
1	5242.19	242985	35.5
2	5284.84	175988	25.7
3	5325.97	161426	23.6
4	10483.00	166885	24.4
5	10554.57	181411	26.5
6	10569.90	276097	40.4
7	10585.50	437879	64.0
8	10603.53	455908	66.7
9	10617.92	560412	81.9
10	10635.15	554612	81.1

Bruker Compass DataAnalysis 4.4 printed: 5/21/2021 12:39:03 PM by: lab Page 2 of 6

Figure S51: Continued.

8.1.7. Mass spectra

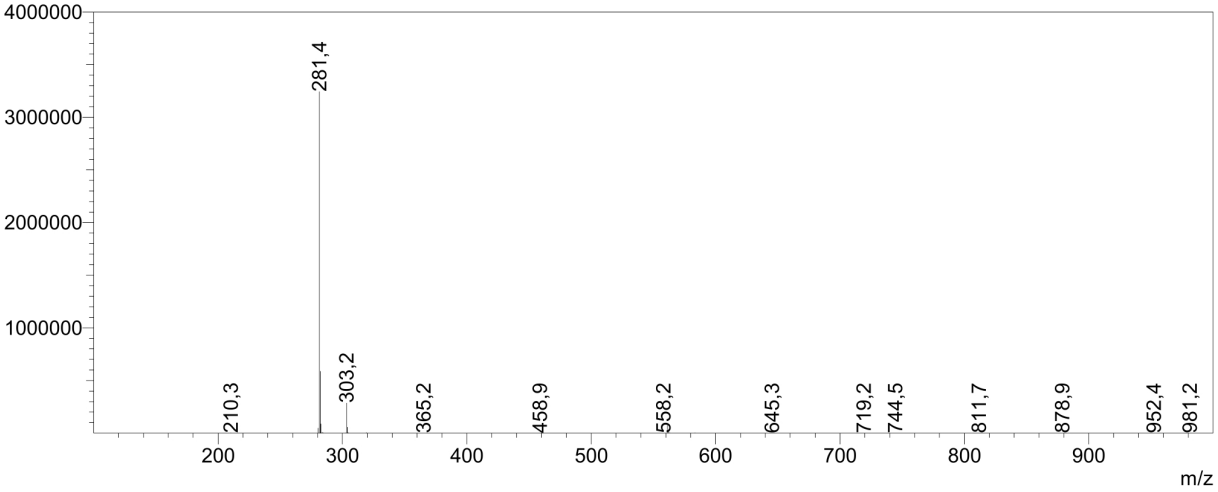


Figure S52: Mass spectrum of compound 2 obtained from LCMS.

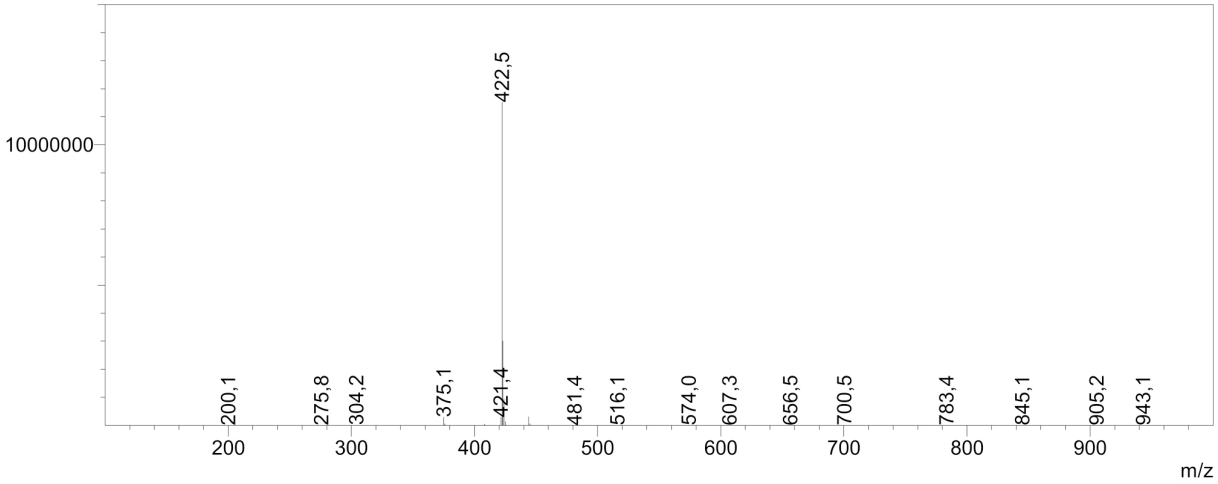


Figure S53: Mass spectrum of compound 3 obtained from LCMS.

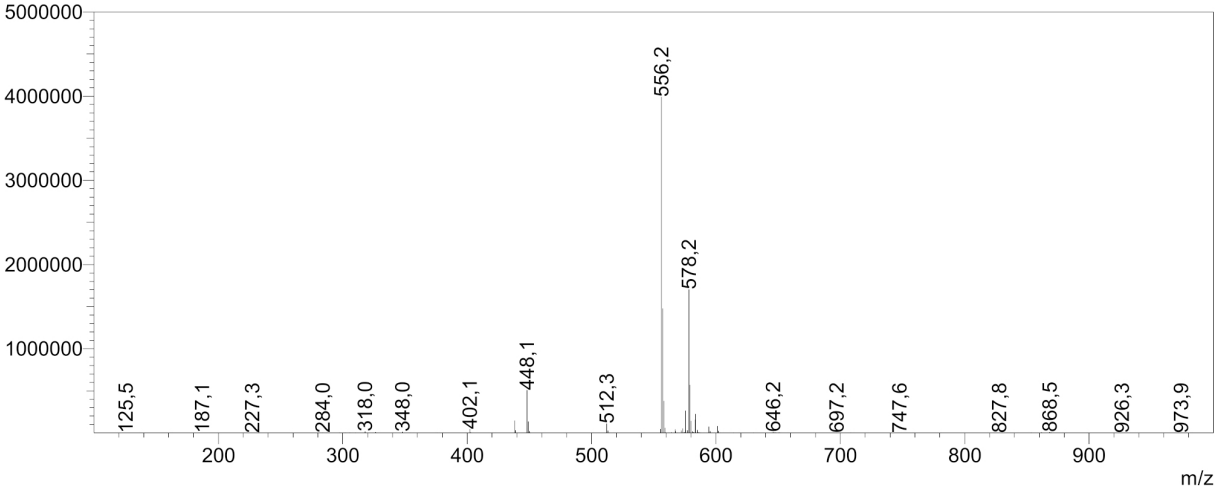


Figure S54: Mass spectrum of compound 4 obtained from LCMS.

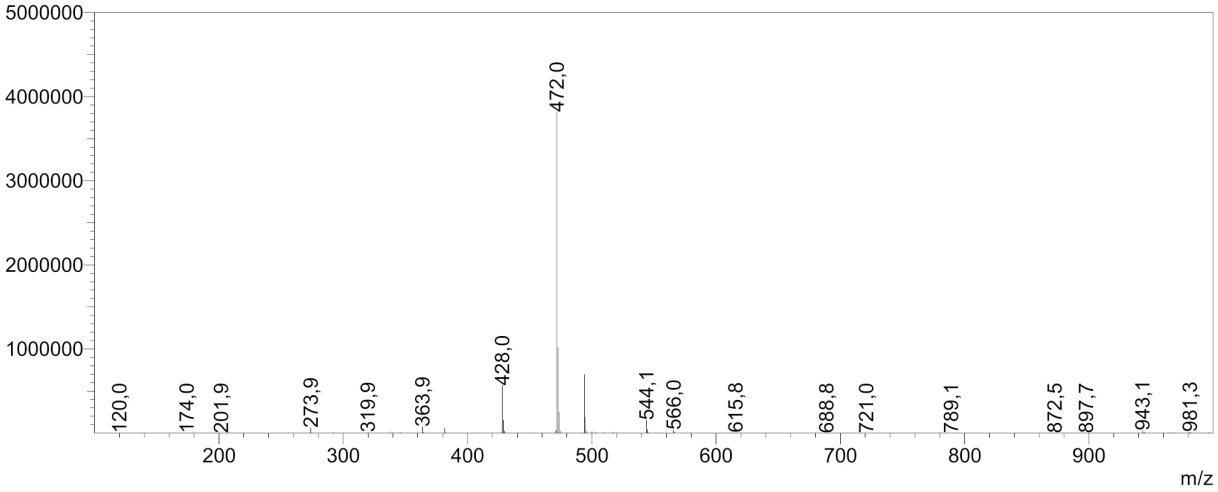


Figure S55: Mass spectrum of compound 5 obtained from LCMS.

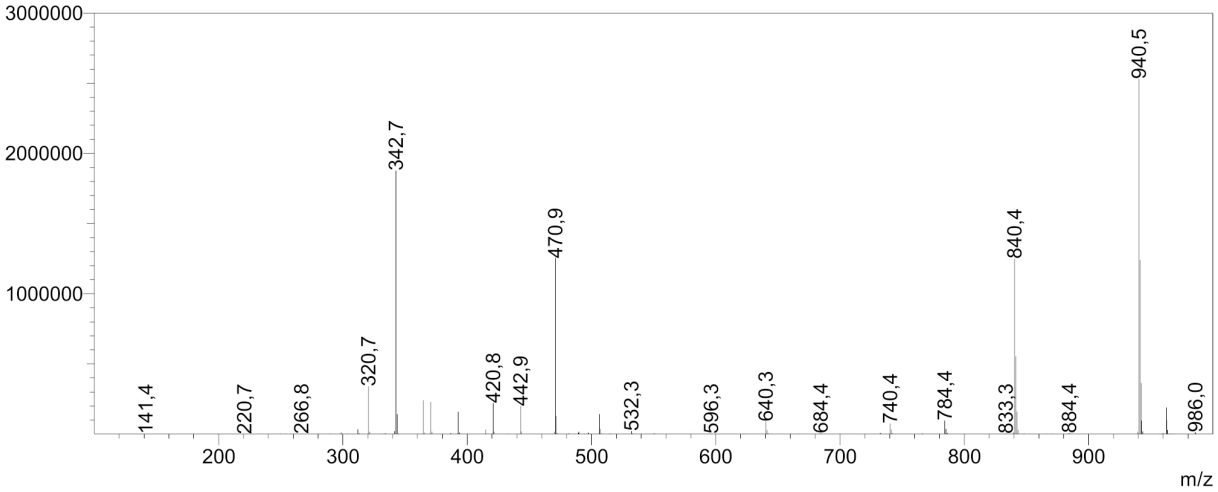


Figure S56: Mass spectrum of compound 6 obtained from LCMS.

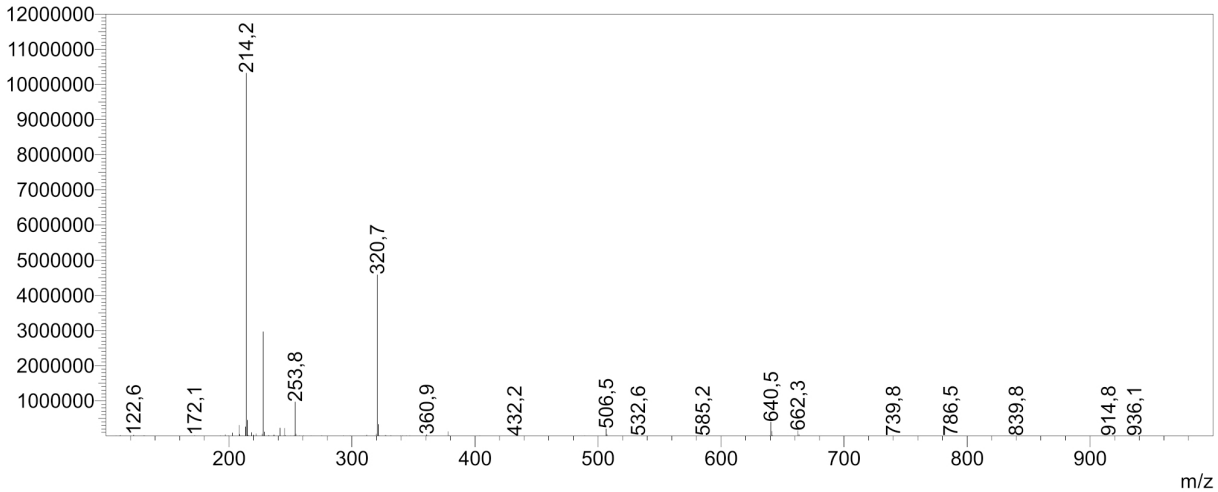


Figure S57: Mass spectrum of compound 7 obtained from LCMS.

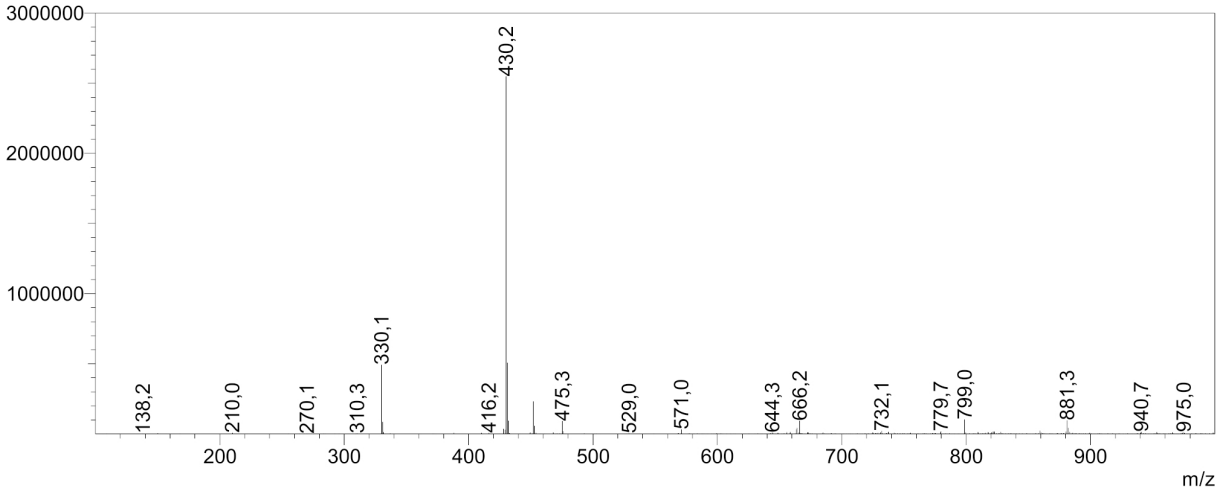


Figure S58: Mass spectrum of compound 9 obtained from LCMS.

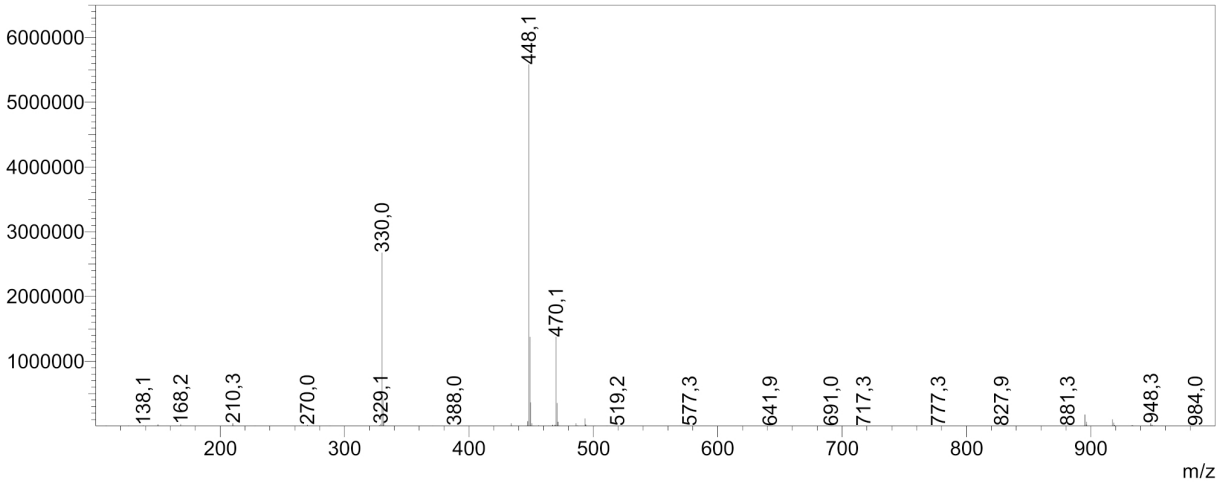


Figure S59: Mass spectrum of compound 10 obtained from LCMS.

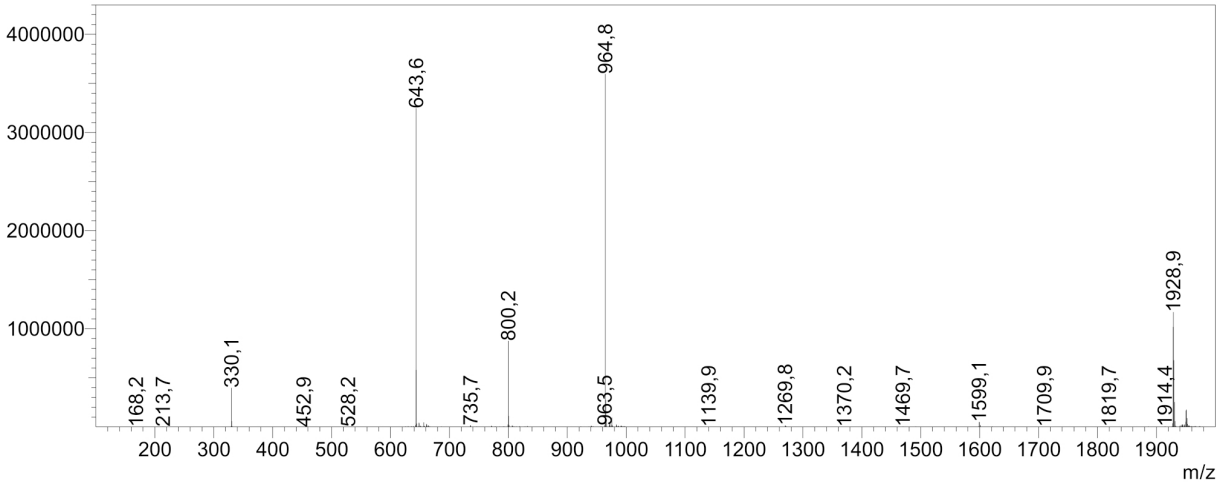


Figure S60: Mass spectrum of full conjugated compound 11 (small scale) obtained from LCMS.

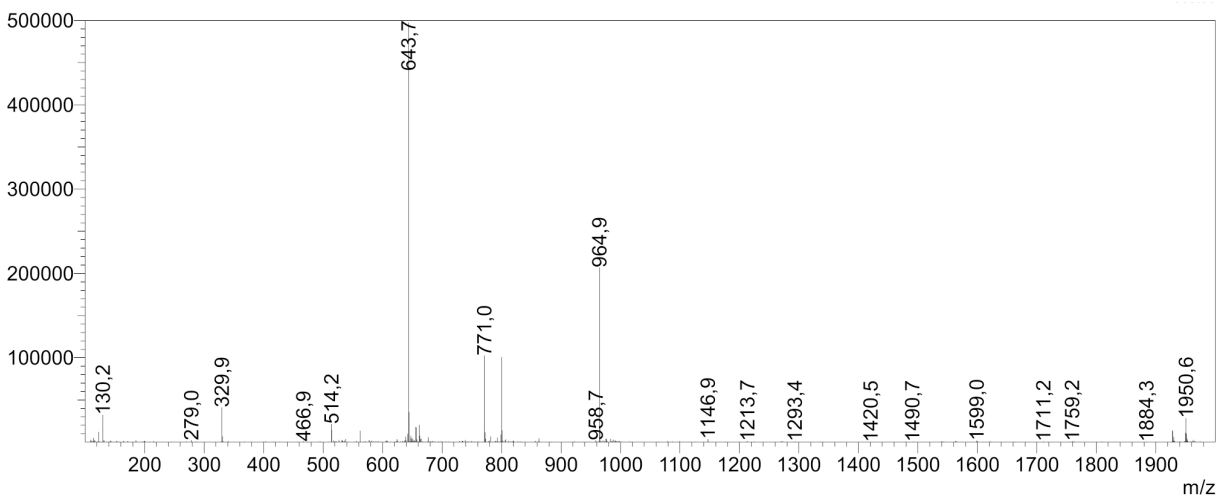


Figure S61: Mass spectrum of full conjugated compound 11 (large scale) obtained from LCMS.

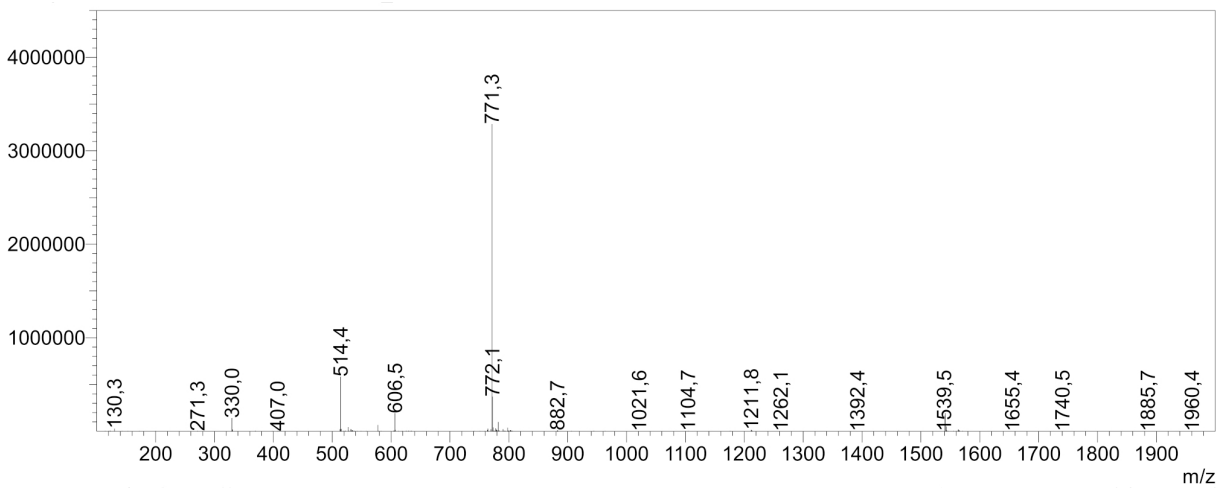


Figure S62: Mass spectrum of double conjugated compound 11 (large scale) obtained from LCMS (see supporting figure S2).

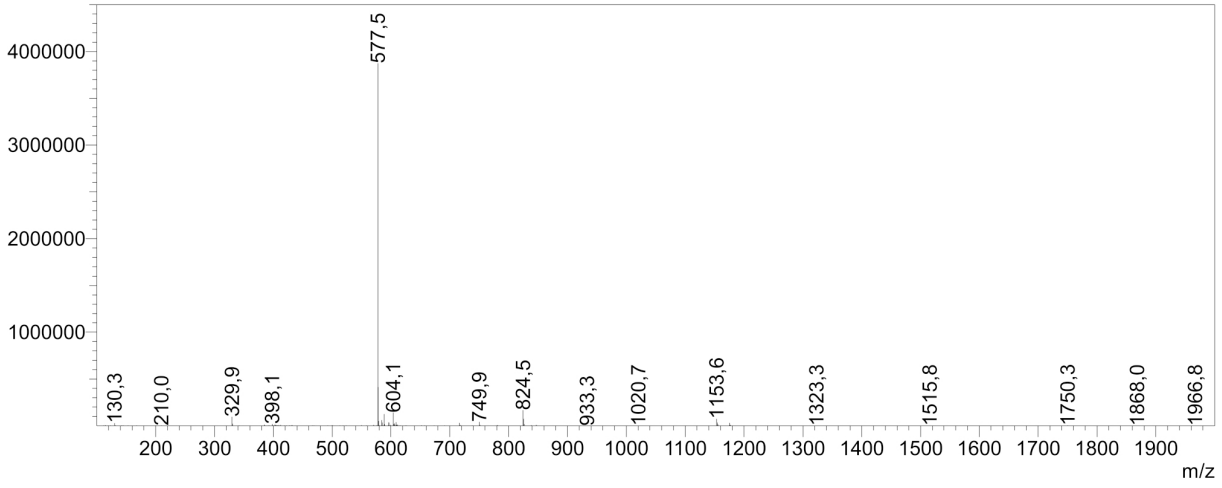


Figure S63: Mass spectrum of single conjugated compound 11 (large scale) obtained from LCMS (see supporting figure S2).

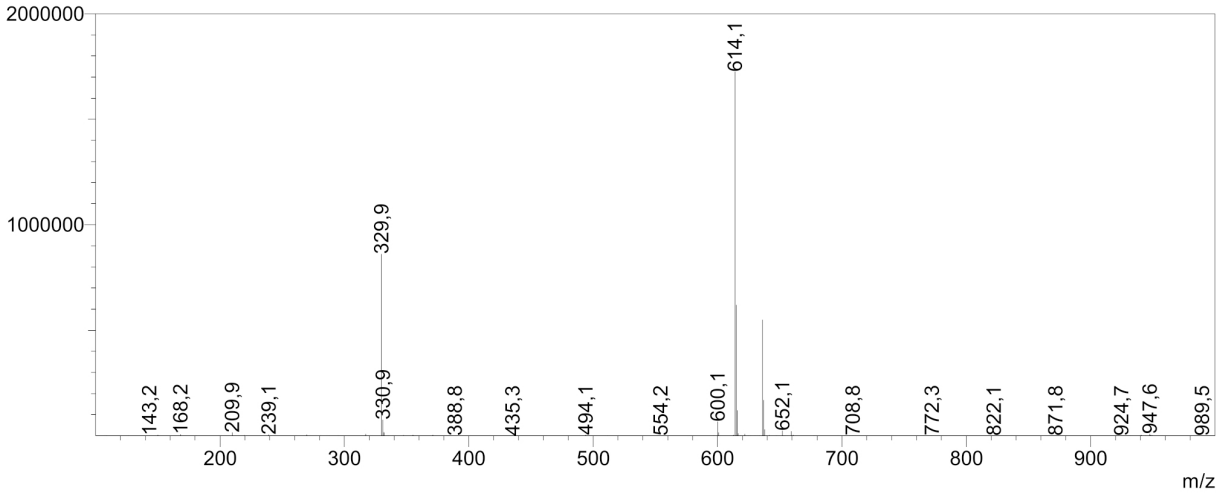


Figure S64: Mass spectrum of compound 12 obtained from LCMS.

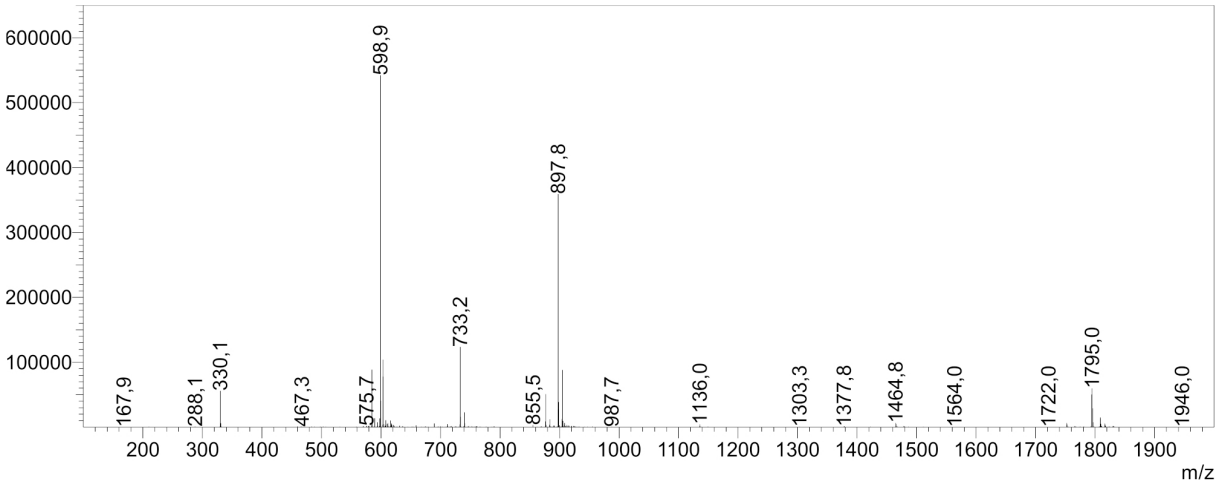


Figure S65: Mass spectrum of compound 13 generated by hydrogenation using molecular nitrogen. Spectrum is obtained from LCMS.

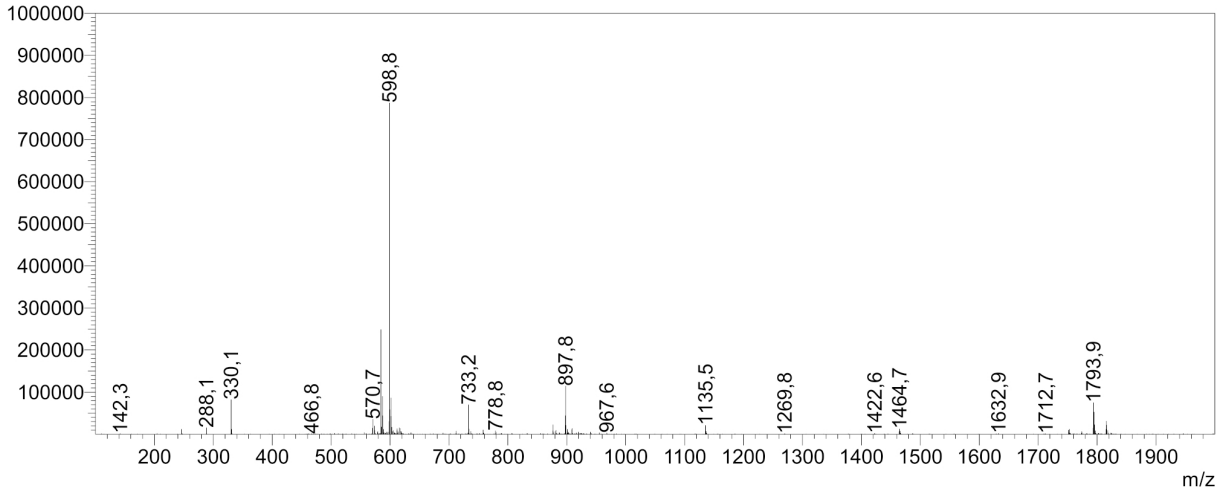


Figure S66: Mass spectrum of compound 13 generated by hydrogenation using ammonium formiate. Spectrum is obtained from LCMS.

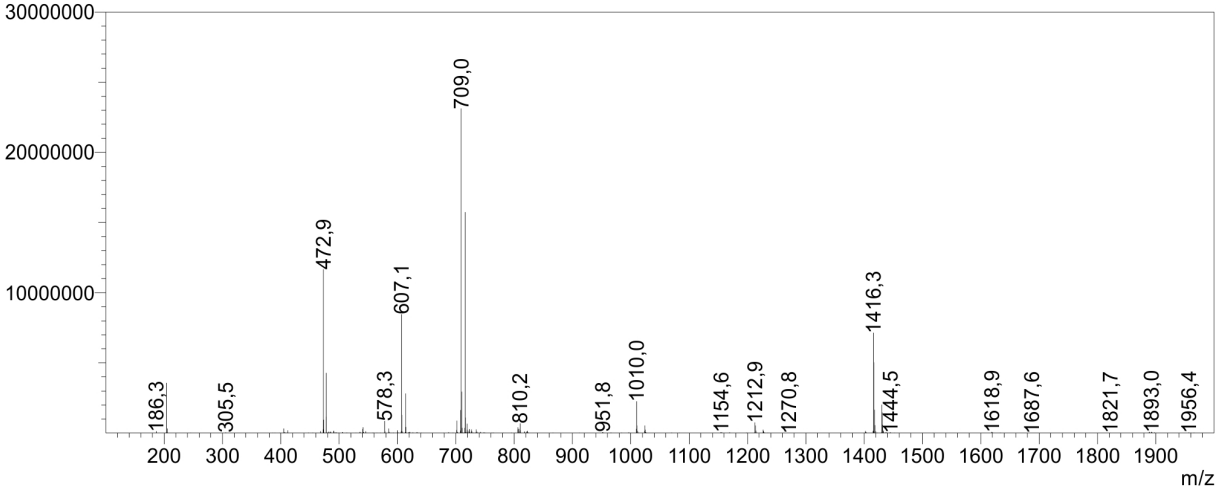


Figure S67: Mass spectrum of compound 14 obtained from LCMS.

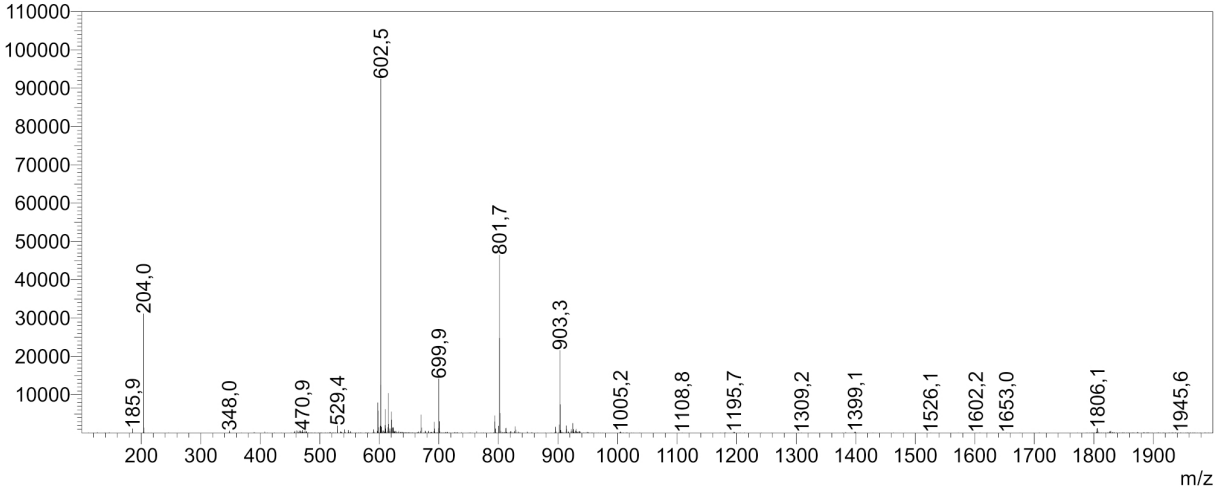


Figure S68: Mass spectrum of compound 15 obtained from LCMS.

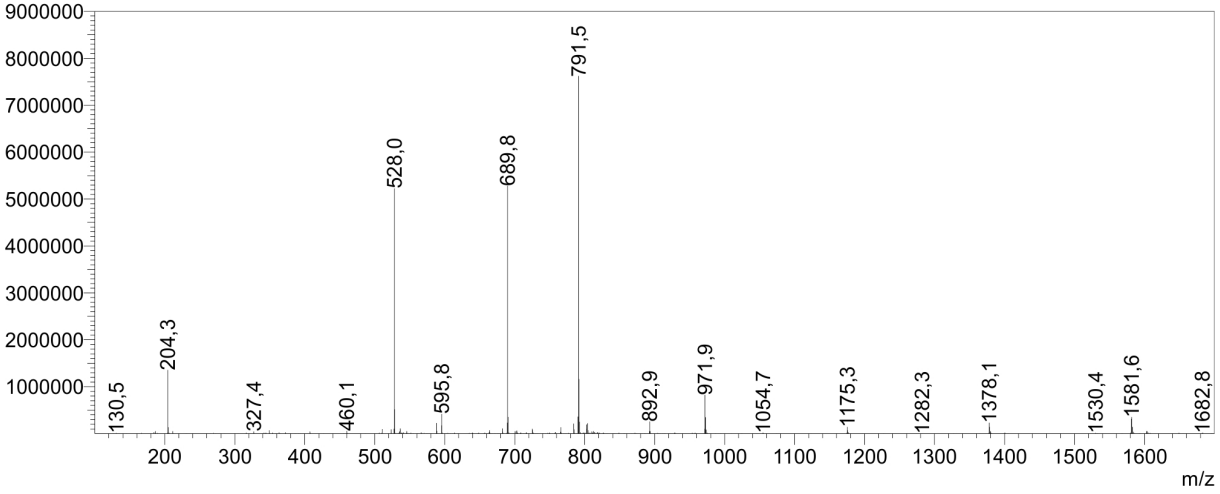


Figure S69: Mass spectrum of compound 16 obtained from LCMS.

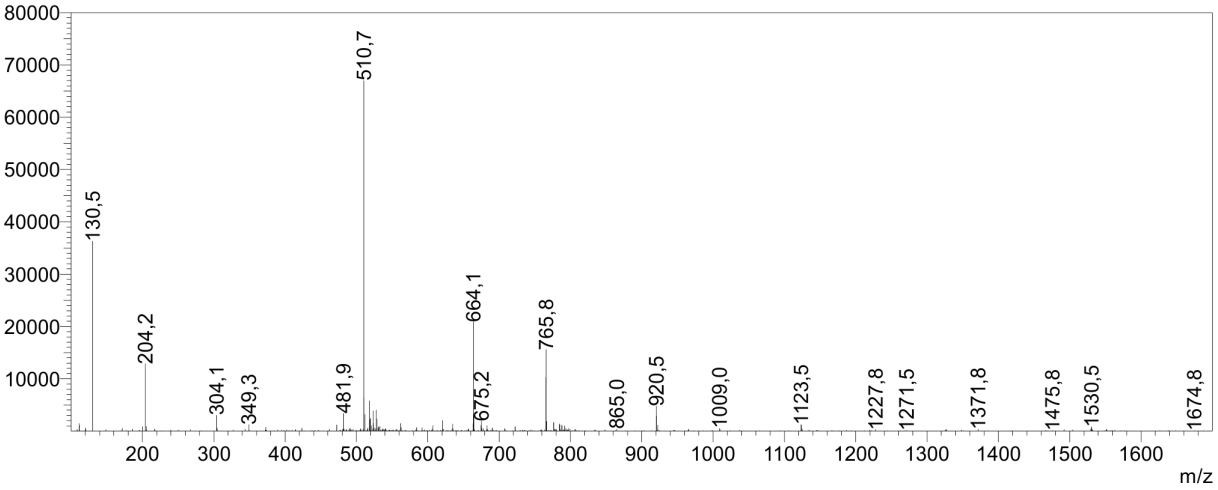


Figure S70: Mass spectrum of compound 17 obtained from LCMS.

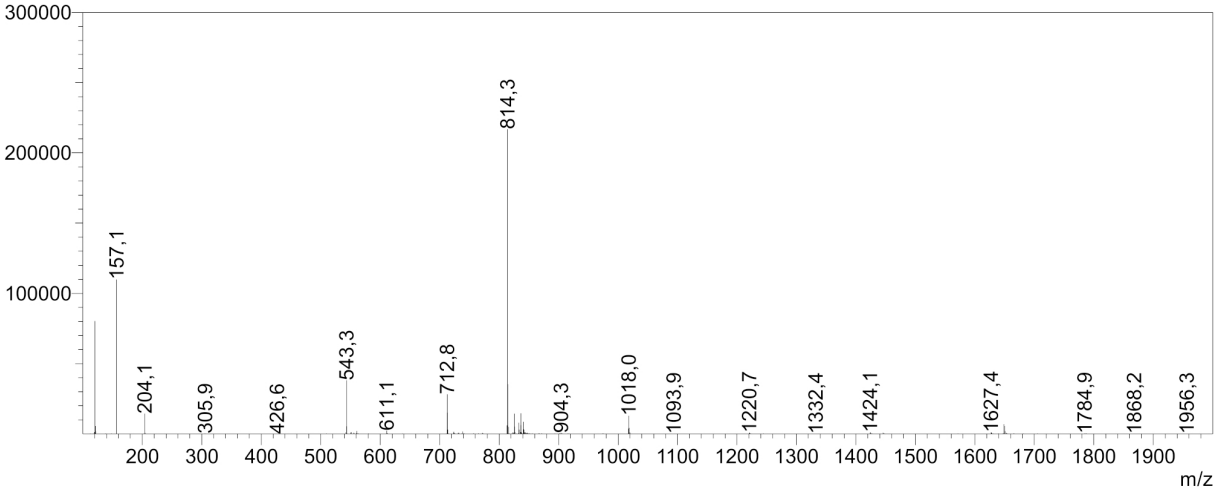


Figure S71: Mass spectrum of compound 17-NHS obtained from LCMS.

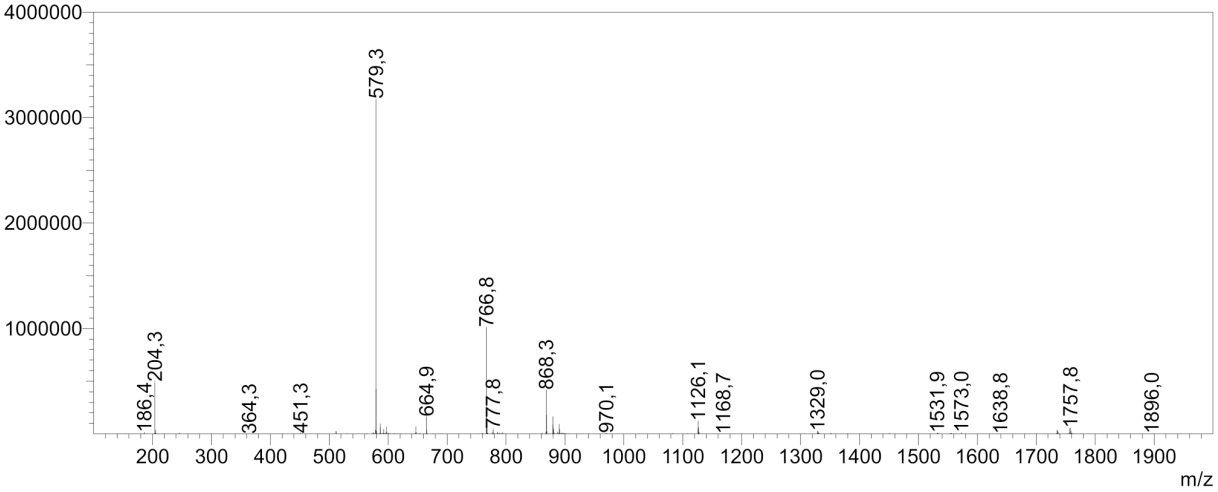


Figure S72: Mass spectrum of compound 19 obtained from LCMS.

8.1.8. HPLC chromatograms

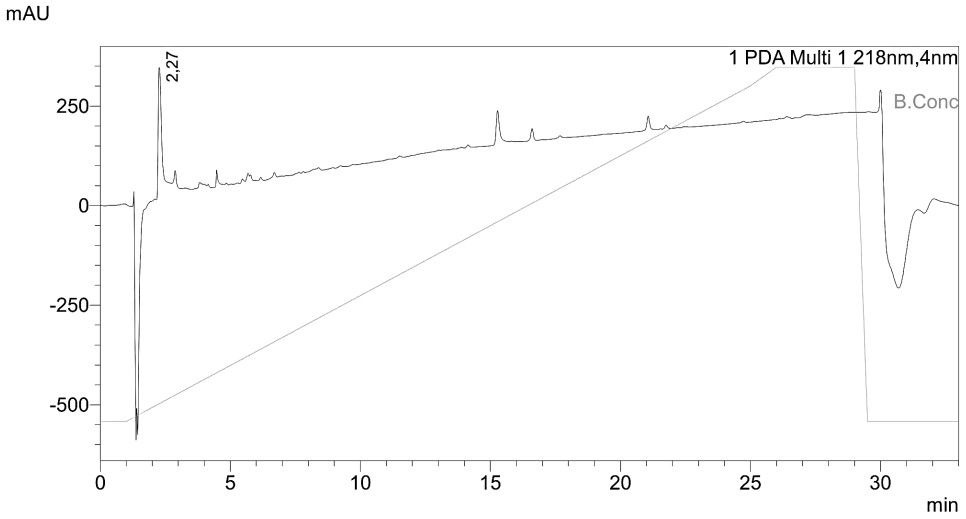


Figure S73: HPLC chromatogram of compound 2 at a detection wavelength of λ = 218 nm.

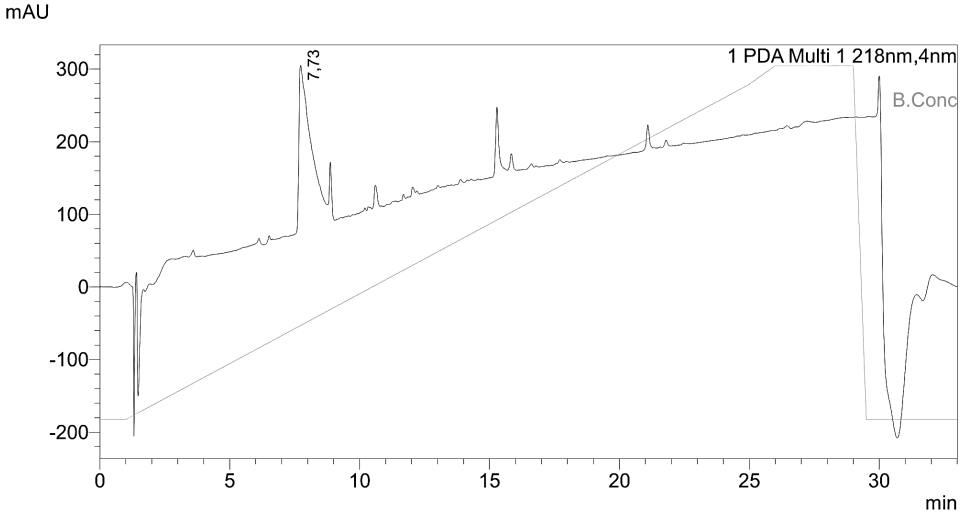


Figure S74: HPLC chromatogram of compound 3 at a detection wavelength of λ = 218 nm.

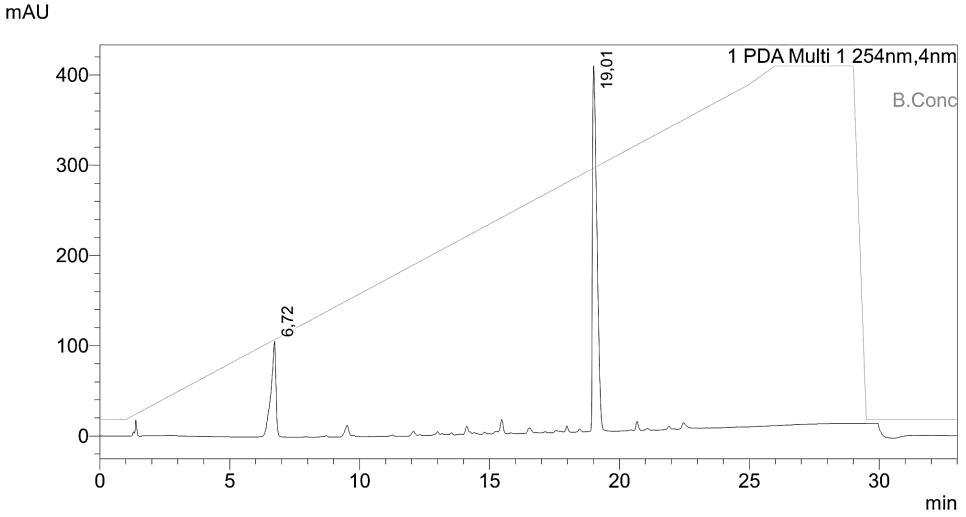


Figure S75: HPLC chromatogram of compound 4 at a detection wavelength of λ = 254 nm.

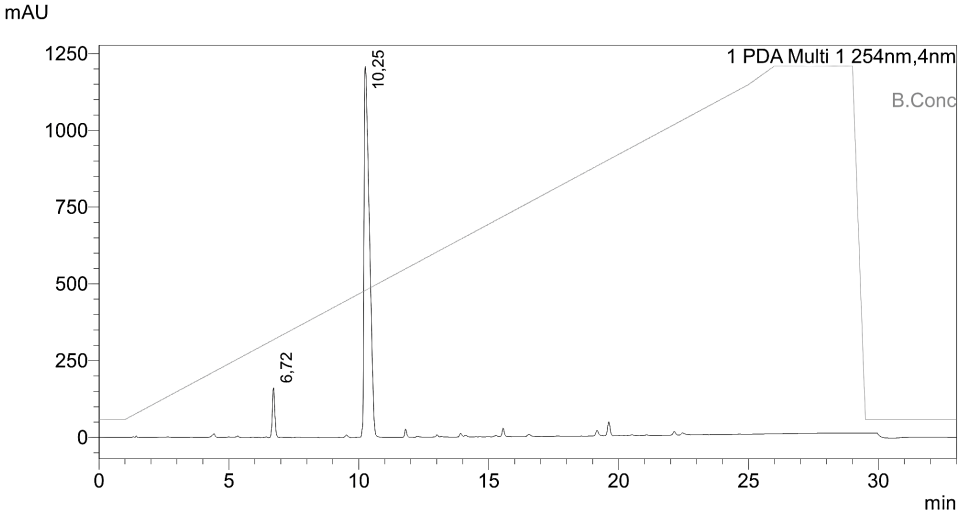


Figure S76: HPLC chromatogram of compound 5 at a detection wavelength of λ = 254 nm.

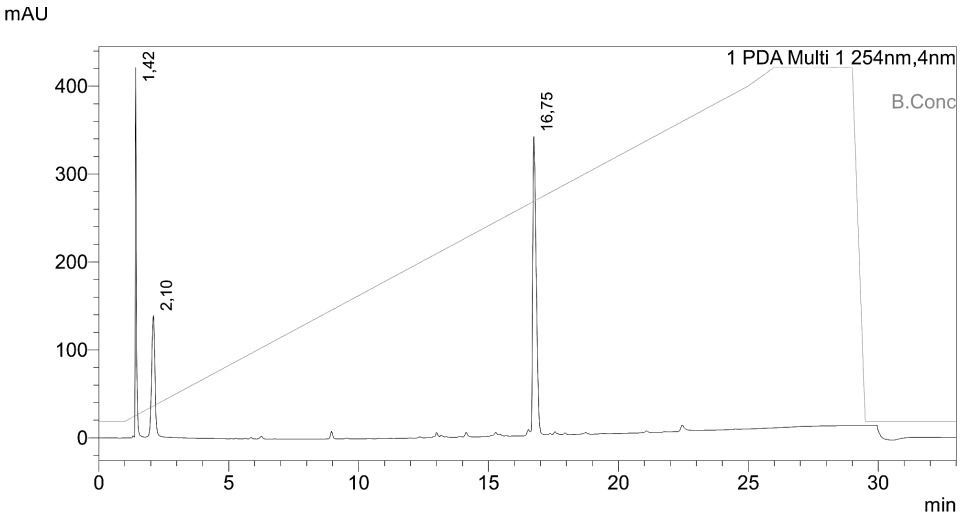


Figure S77: HPLC chromatogram of compound 6 at a detection wavelength of λ = 254 nm.

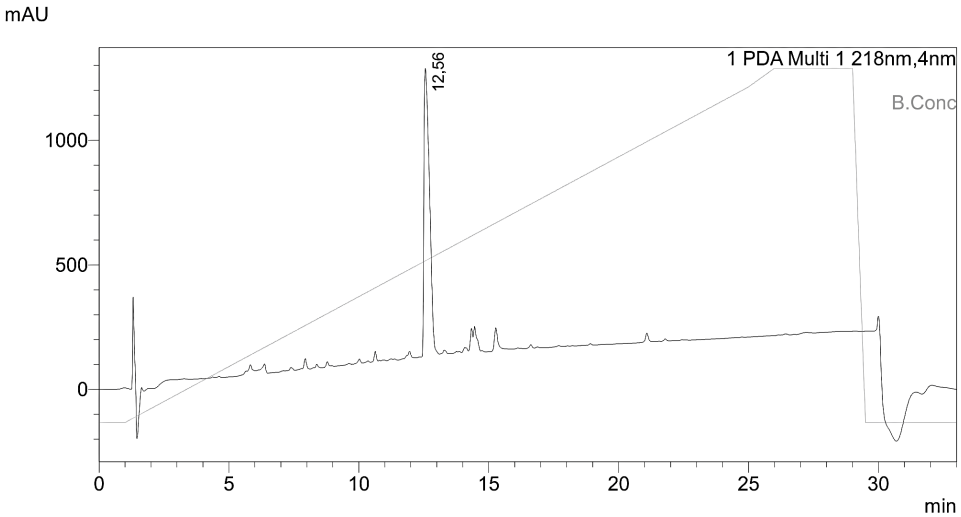


Figure S78: HPLC chromatogram of compound 9 at a detection wavelength of λ = 218 nm.

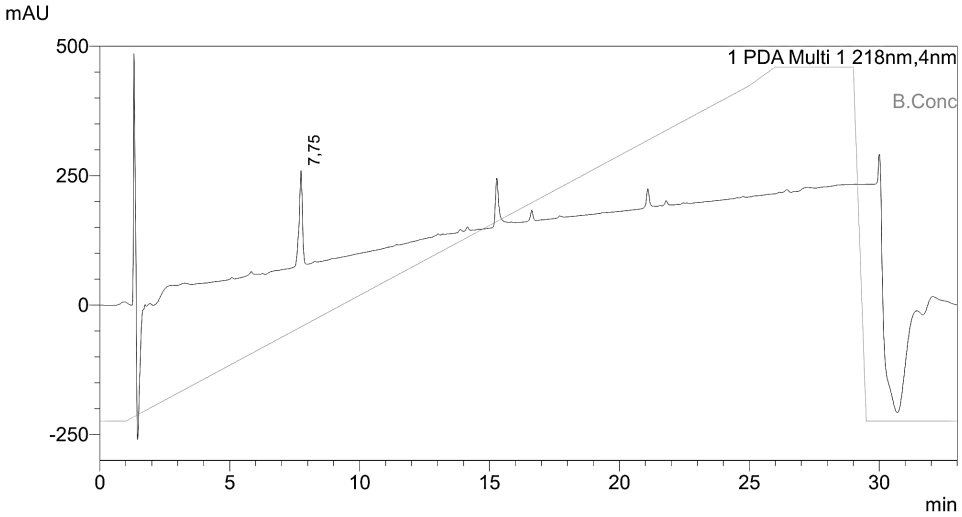


Figure S79: HPLC chromatogram of compound 10 at a detection wavelength of λ = 218 nm.

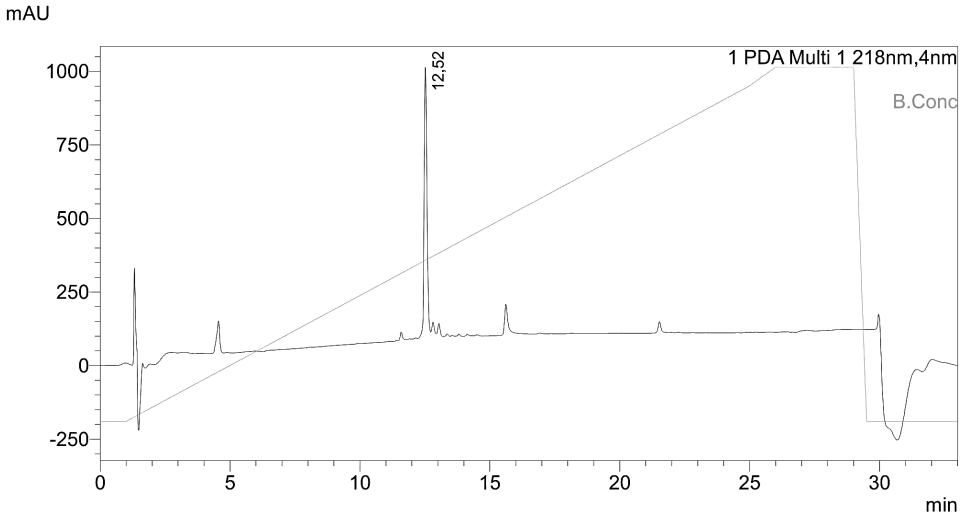


Figure S80: HPLC chromatogram of compound 12 at a detection wavelength of λ = 218 nm.

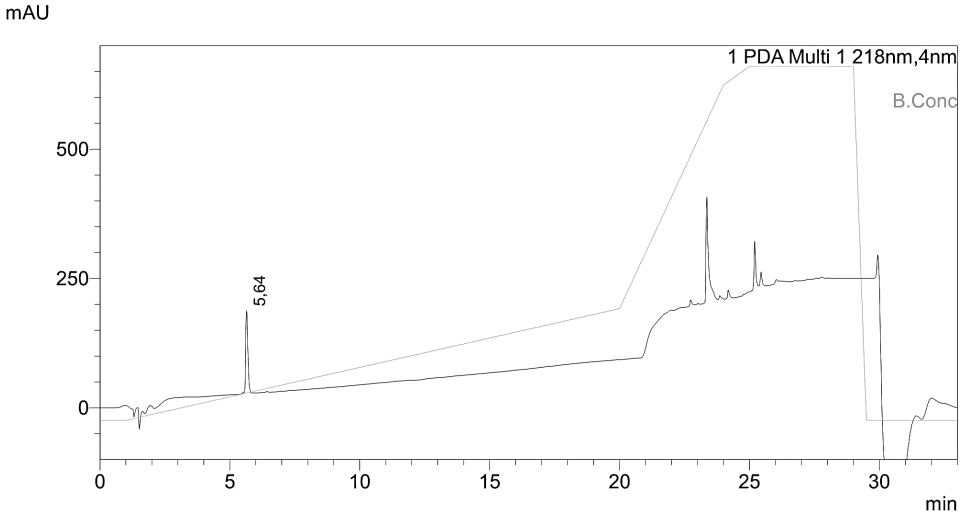


Figure S81: HPLC chromatogram of compound 14 at a detection wavelength of λ = 218 nm.

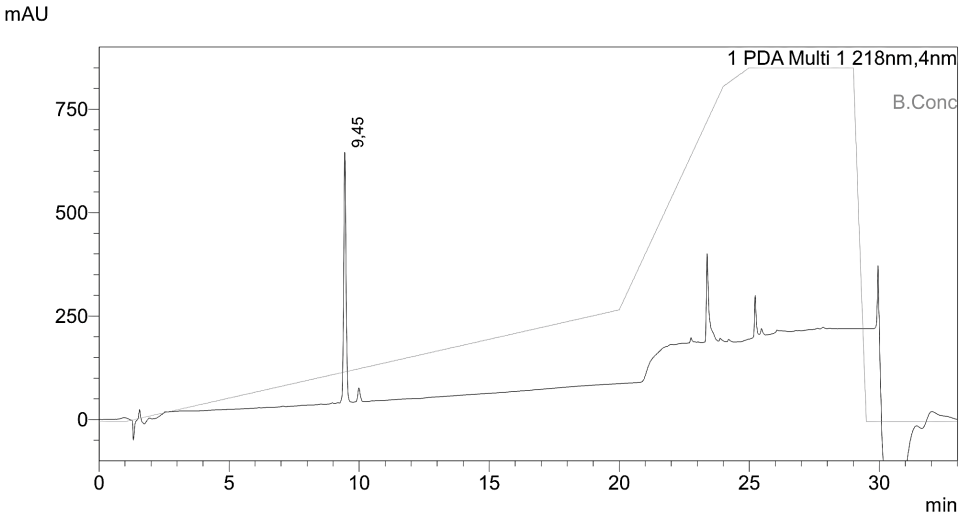


Figure S82: HPLC chromatogram of compound 16 at a detection wavelength of λ = 218 nm.

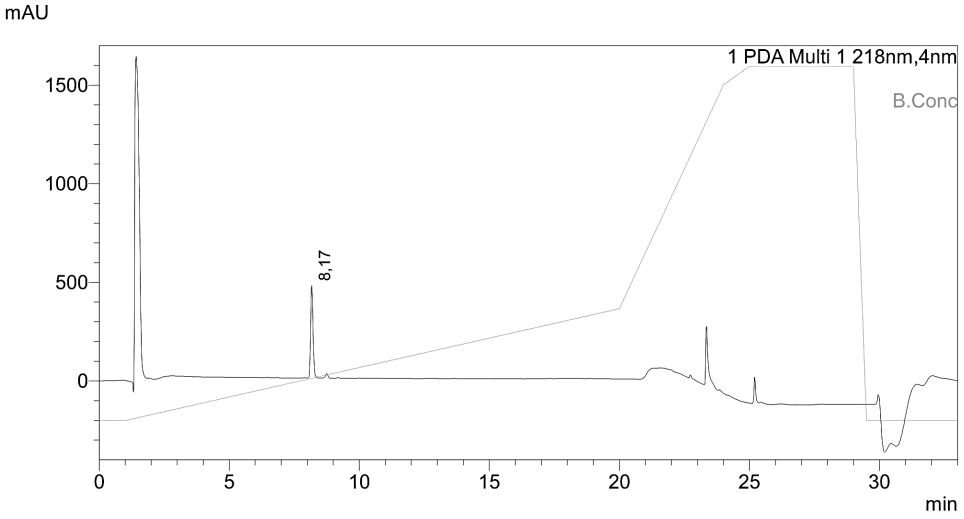


Figure S83: HPLC chromatogram of compound 17 at a detection wavelength of λ = 218 nm.

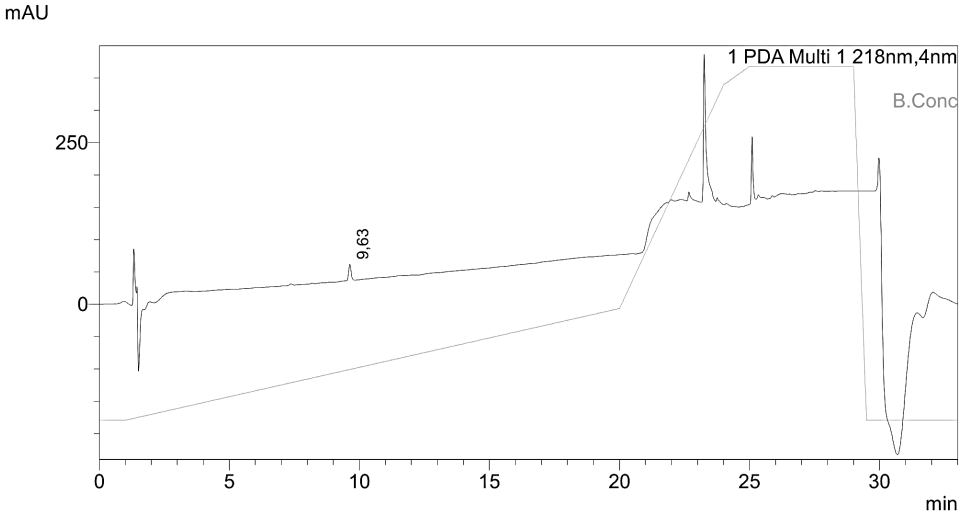


Figure S84: HPLC chromatogram of compound 17-NHS at a detection wavelength of λ = 218 nm.

8.1.9. UV/VIS spectrometry

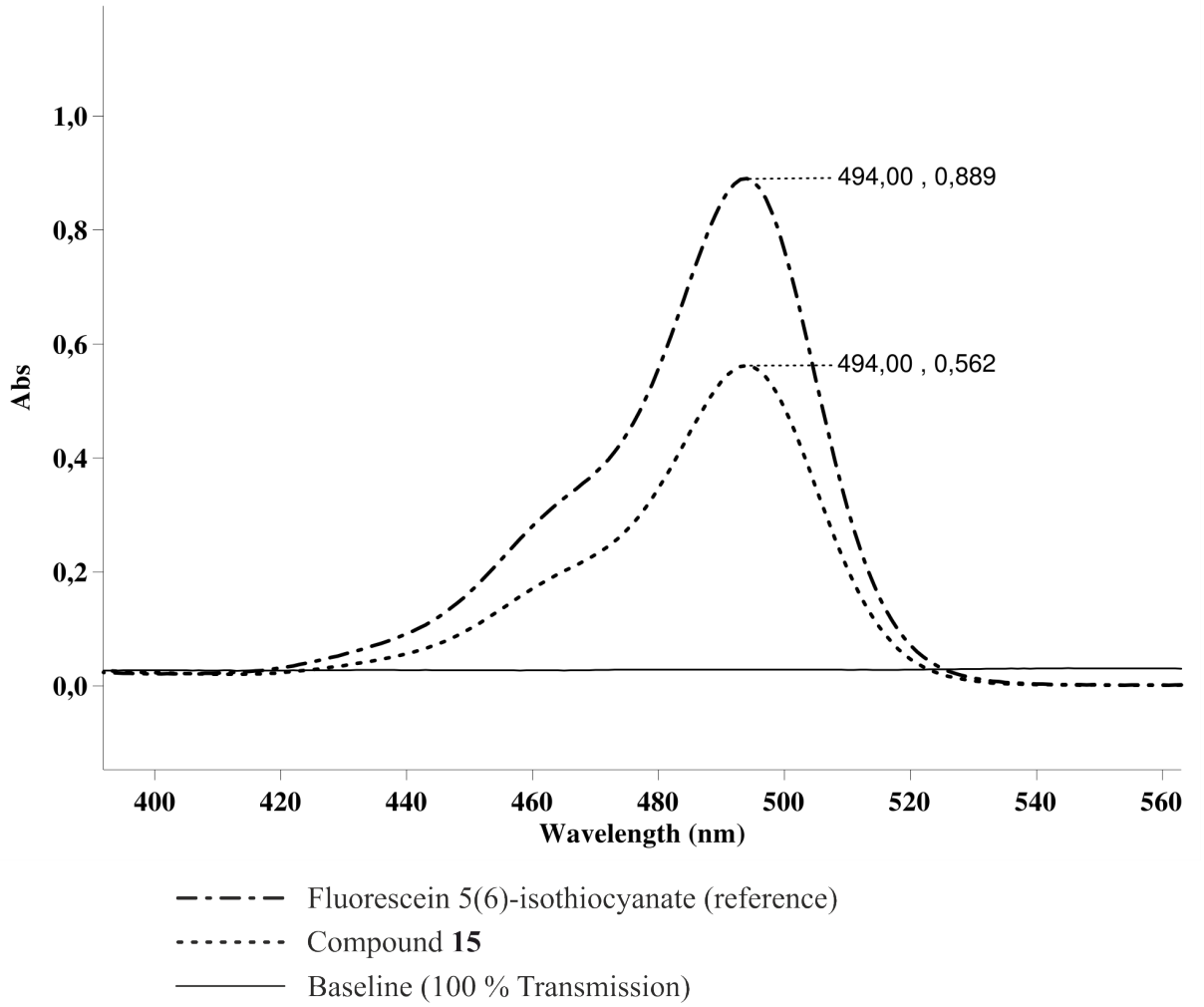


Figure S85: UV spectrum of compound 15. *Compound **15** was preliminary dissolved in 1 mL DMSO-d_6 for NMR analysis and the concentration was calculated using the mass of the purified product ($c_{calculated}$ (**15**) = 5.82 mM). 3.44 µL of the stock solution (5.82 mM) were diluted with PBS (pH 7.4) to a final concentration of 10 µM and a final volume of 2 mL. As reference, 3.44 µL of a stock solution of fluorescein 5(6)-isothiocyanate (5.8 mM in DMSO) were diluted with PBS (pH 7.4) to a final concentration of 10 µM and a final volume of 2 mL. The UV/VIS spectra were recorded with a Cary 300 Scan UV/Visible spectrophotometer (see methods and materials) and the molar attenuation coefficient ε_λ of fluorescein 5(6)-isothiocyanate was calculated from the absorption at λ_{max} = 494 nm and the given concentration applying the law of Lambert-Beer. The final concentration of compound **15** was calculated using the molar attenuation coefficient ε_λ of fluorescein 5(6)-isothiocyanate and the obtained absorption at λ_{max} = 494 nm.*

Law of Lambert-Beer:

$$E_\lambda = \varepsilon_\lambda \cdot c \cdot d \qquad (2)$$

with E_λ = absorbance, ε_λ = molar attenuation coefficient, c = concentration, d = path length

Molar attenuation coefficient ε_λ (FITC):

c_{stock} (FITC in DMSO) $= 5.8$ mM

$c_{measurement}$ (FITC) $= 10$ µM (in PBS)

$\lambda_{max} = 494$ nm

$d = 1$ cm

E_{494} (FITC) $= 0.889$

$$E_\lambda = \varepsilon_\lambda \cdot c \cdot d \Leftrightarrow \varepsilon_\lambda = \frac{E_\lambda}{c \cdot d} \quad (2)$$

$$\varepsilon_{494}\,(\text{FITC}) = \frac{E_{494}}{c \cdot d} = \frac{0.889}{10\ \mu\text{M} \cdot \text{cm}} = 89215.15\ \frac{1}{\text{M} \cdot \text{cm}}$$

Determination of the concentration and yield of compound **15**:

$$\varepsilon_{494}\,(\text{FITC}) = \varepsilon_{494}\,(\mathbf{15}) = 89215.15\ \frac{1}{\text{M} \cdot \text{cm}}$$

$\lambda_{max} = 494$ nm

$d = 1$ cm

$E_{494}\,(\mathbf{15}) = 0.562$

$$E_\lambda = \varepsilon_\lambda \cdot c \cdot d \Leftrightarrow c = \frac{E_\lambda}{\varepsilon_\lambda \cdot d} \quad (2)$$

$$c\,(\mathbf{15}) = \frac{E_{494}\,(\mathbf{15})}{\varepsilon_{494}\,(\text{FITC}) \cdot d} = \frac{0.562}{89215.15\ \frac{1}{\text{M} \cdot \text{cm}} \cdot 1\ \text{cm}} = 3.66\ \text{mM}$$

$$n\,(\mathbf{15}) = c \cdot V = 3.66\ \text{mM} \cdot 1\ \text{mL (DMSO)} = 3.66\ \mu\text{mol}$$

$$\text{Yield} = \frac{n\,(\mathbf{15})}{n\,(\mathbf{14})} \cdot 100\ \% = \frac{3.66\ \mu\text{mol}}{14\ \mu\text{mol}} \cdot 100\ \% = 26\ \%$$

8.1.10. Vector sequences

8.1.10.1. Vector sequence of pTS689

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241     GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301     TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
```

```
361     CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421     ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481     ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541     ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601     TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661     ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
841     CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
901     GTTTAAACTT AAGCTTGGTA CCGAGCTCGG ATCCACCATG ACCAAGGAGT ATCAAGACCT
961     TCAGCATCTG GACAATGAGG AGAGTGACCA CCATCAGCTC AGAAAAGGGC CACCTCCTCC
1021    CCAGCCCCTC CTGCAGCGTC TCTGCTCCGG ACCTCGCCTC CTCCTGCTCT CCCTGGGCCT
1081    CAGCCTCCTG CTGCTTGTGG TTGTCTGTGT GATCGGATCC CAAAACTCCC AGCTGCAGGA
1141    GGAGCTGCGG GGCCTGAGAG AGACGTTCAG CAACTTCACA GCGAGCACGG AGGCCCAGGT
1201    CAAGGGCTTG AGCACCCAGG GAGGCAATGT GGGAAGAAAG ATGAAGTCGC TAGAGTCCCA
1261    GCTGGAGAAA CAGCAGAAGG ACCTGAGTGA AGATCACTCC AGCCTGCTGC TCCACGTGAA
1321    GCAGTTCGTG TCTGACCTGC GGAGCCTGAG CTGTCAGATG GCGGCGCTCC AGGGCAATGG
1381    CTCAGAAAGG ACCTGCTGCC CGGTCAACTG GGTGGAGCAC GAGCGCAGCT GCTACTGGTT
1441    CTCTCGCTCC GGGAAGGCCT GGGCTGACGC CGACAACTAC TGCCGGCTGG AGGACGCGCA
1501    CCTGGTGGTG GTCACGTCCT GGGAGGAGCA GAAATTTGTC CAGCACCACA TAGGCCCTGT
1561    GAACACCTGG ATGGGCCTCC ACGACCAAAA CGGGCCCTGG AAGTGGGTGG ACGGGACGGA
1621    CTACGAGACG GGCTTCAAGA ACTGGAGGCC GGAGCAGCCG GACGACTGGT ACGGCCACGG
1681    GCTCGGAGGA GGCGAGGACT GTGCCCACTT CACCGACGAC GGCCGCTGGA ACGACGACGT
1741    CTGCCAGAGG CCCTACCGCT GGGTCTGCGA GACAGAGCTG GACAAGGCCA GCCAGGAGCC
1801    ACCTCTCCTT TAATCTAGAG GGCCCTTCGA ACAAAAACTC ATCTCAGAAG AGGATCTGAA
1861    TATGCATACC GGTCATCATC ACCATCACCA TTGAGTTTAA ACCCGCTGAT CAGCCTCGAC
1921    TGTGCCTTCT AGTTGCCAGC CATCTGTTGT TTGCCCCTCC CCCGTGCCTT CCTTGACCCT
1981    GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT CGCATTGTCT
2041    GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG
2101    GGAAGACAAT AGCAGGCATG CTGGGGATGC GGTGGGCTCT ATGGCTTCTG AGGCGGAAAG
2161    AACCAGCTGG GGCTCTAGGG GGTATCCCCA CGCGCCCTGT AGCGGCGCAT TAAGCGCGGC
2221    GGGTGTGGTG GTTACGCGCA GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCCGCTCC
2281    TTTCGCTTTC TTCCCTTCCT TTCTCGCCAC GTTCGCCGGC TTTCCCCGTC AAGCTCTAAA
2341    TCGGGGGCTC CCTTTAGGGT TCCGATTTAG TGCTTTACGG CACCTCGACC CCAAAAAACT
2401    TGATTAGGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTCGCCCTTT
2461    GACGTTGGAG TCCACGTTCT TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA
2521    CCCTATCTCG GTCTATTCTT TTGATTTATA AGGGATTTTG CCGATTTCGG CCTATTGGTT
2581    AAAAAATGAG CTGATTTAAC AAAAATTTAA CGCGAATTAA TTCTGTGGAA TGTGTGTCAG
2641    TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA GTATGCAAAG CATGCATCTC
2701    AATTAGTCAG CAACCAGGTG TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG AAGTATGCAA
2761    AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC TAACTCCGCC CATCCCGCCC
2821    CTAACTCCGC CCAGTTCCGC CCATTCTCCG CCCCATGGCT GACTAATTTT TTTTATTTAT
2881    GCAGAGGCCG AGGCCGCCTC TGCCTCTGAG CTATTCCAGA AGTAGTGAGG AGGCTTTTTT
2941    GGAGGCCTAG GCTTTTGCAA AAAGCTCCCG GGAGCTTGTA TATCCATTTT CGGATCTGAT
3001    CAAGAGACAG GATGAGGATC GTTTCGCATG ATTGAACAAG ATGGATTGCA CGCAGGTTCT
3061    CCGGCCGCTT GGGTGGAGAG GCTATTCGGC TATGACTGGG CACAACAGAC AATCGGCTGC
3121    TCTGATGCCG CCGTGTTCCG GCTGTCAGCG CAGGGGCGCC CGGTTCTTTT TGTCAAGACC
3181    GACCTGTCCG GTGCCCTGAA TGAACTGCAG GACGAGGCAG CGCGGCTATC GTGGCTGGCC
3241    ACGACGGGCG TTCCTTGCGC AGCTGTGCTC GACGTTGTCA CTGAAGCGGG AAGGGACTGG
3301    CTGCTATTGG GCGAAGTGCC GGGGCAGGAT CTCCTGTCAT CTCACCTTGC TCCTGCCGAG
3361    AAAGTATCCA TCATGGCTGA TGCAATGCGG CGGCTGCATA CGCTTGATCC GGCTACCTGC
3421    CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCGAGCAC GTACTCGGAT GGAAGCCGGT
3481    CTTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGGC TCGCGCCAGC CGAACTGTTC
3541    GCCAGGCTCA AGGCGCGCAT GCCCGACGGC GAGGATCTCG TCGTGACCCA TGGCGATGCC
3601    TGCTTGCCGA ATATCATGGT GGAAAATGGC CGCTTTTCTG GATTCATCGA CTGTGGCCGG
3661    CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA CCCGTGATAT TGCTGAAGAG
3721    CTTGGCGGCG AATGGGCTGA CCGCTTCCTC GTGCTTTACG GTATCGCCGC TCCCGATTCG
3781    CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT GAGCGGGACT CTGGGGTTCG
3841    AAATGACCGA CCAAGCGACG CCCAACCTGC CATCACGAGA TTTCGATTCC ACCGCCGCCT
3901    TCTATGAAAG GTTGGGCTTC GGAATCGTTT TCCGGGACGC CGGCTGGATG ATCCTCCAGC
```

```
3961    GCGGGGATCT CATGCTGGAG TTCTTCGCCC ACCCCAACTT GTTTATTGCA GCTTATAATG
4021    GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT
4081    CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGTATA CCGTCGACCT
4141    CTAGCTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT TGTTATCCGC
4201    TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG GGTGCCTAAT
4261    GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG TCGGGAAACC
4321    TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT TTGCGTATTG
4381    GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTCGTTCGG CTGCGGCGAG
4441    CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC AGAATCAGGG GATAACGCAG
4501    GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG GCCGCGTTGC
4561    TGGCGTTTTT CCATAGGCTC CGCCCCCCTG ACGAGCATCA CAAAAATCGA CGCTCAAGTC
4621    AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCCCCCT GGAAGCTCCC
4681    TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC TTTCTCCCTT
4741    CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG GTGTAGGTCG
4801    TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCGACCGC TGCGCCTTAT
4861    CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA CTGGCAGCAG
4921    CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG TTCTTGAAGT
4981    GGTGGCCTAA CTACGGCTAC ACTAGAAGAA CAGTATTTGG TATCTGCGCT CTGCTGAAGC
5041    CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC ACCGCTGGTA
5101    GCGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC
5161    CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT
5221    TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT
5281    TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA
5341    GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG
5401    TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC
5461    CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG
5521    CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC
5581    GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA
5641    CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC
5701    GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC
5761    CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC
5821    TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT
5881    CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA
5941    TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT
6001    CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA
6061    CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA
6121    AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC
6181    TCATACTCTT CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG
6241    GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC
6301    GAAAAGTGCC ACCTGACGTC
```

8.1.10.2. Vector sequence of pTS690

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241     GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301     TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
361     CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421     ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481     ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541     ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601     TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661     ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
841     CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
901     GTTTAAACTT AAGCTTGGTA CCGAGCTCGG ATCCACCATG GCCAAGGACT TTCAAGATAT
961     CCAGCAGCTG AGCTCGGAGG AAAATGACCA TCCTTTCCAT CAAGGTGAGG GGCCAGGCAC
```

```
1021    TCGCAGGCTG AATCCCAGGA GAGGAAATCC ATTTTTGAAA GGGCCACCTC CTGCCCAGCC
1081    CCTGGCACAG CGTCTCTGCT CCATGGTCTG CTTCAGTCTG CTTGCCCTGA GCTTCAACAT
1141    CCTGCTGCTG GTGGTCATCT GTGTGACTGG GTCCCAAAGT GCACAGCTGC AAGCCGAGCT
1201    GCGGAGCCTG AAGGAAGCTT TCAGCAACTT CTCCTCGAGC ACCCTGACGG AGGTCCAGGC
1261    AATCAGCACC CACGGAGGCA GCGTGGGTGA CAAGATCACA TCCCTAGGAG CCAAGCTGGA
1321    GAAACAGCAG CAGGACCTGA AAGCAGATCA CGATGCCCTG CTCTTCCATC TGAAGCACTT
1381    CCCCGTGGAC CTGCGCTTCG TGGCCTGCCA GATGGAGCTC CTCCACAGCA ACGGCTCCCA
1441    AAGGACCTGC TGCCCCGTCA ACTGGGTGGA GCACCAAGGC AGCTGCTACT GGTTCTCTCA
1501    CTCCGGGAAG GCCTGGGCTG AGGCGGAGAA GTACTGCCAG CTGGAGAACG CACACCTGGT
1561    GGTCATCAAC TCCTGGGAGG AGCAGAAATT CATTGTACAA CACACGAACC CCTTCAATAC
1621    CTGGATAGGT CTCACGGACA GTGATGGCTC TTGGAAATGG GTGGATGGCA CAGACTATAG
1681    GCACAACTAC AAGAACTGGG CTGTCACTCA GCCAGATAAT TGGCACGGGC ACGAGCTGGG
1741    TGGAAGTGAA GACTGTGTTG AAGTCCAGCC GGATGGCCGC TGGAACGATG ACTTCTGCCT
1801    GCAGGTGTAC CGCTGGGTGT GTGAGAAAAG GCGGAATGCC ACCGGCGAGG TGGCCTGATC
1861    TAGAGGGCCC TTCGAACAAA AACTCATCTC AGAAGAGGAT CTGAATATGC ATACCGGTCA
1921    TCATCACCAT CACCATTGAG TTTAAACCCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
1981    CCAGCCATCT GTTGTTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC
2041    CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATTC
2101    TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG GATTGGGAAG ACAATAGCAG
2161    GCATGCTGGG GATGCGGTGG GCTCTATGGC TTCTGAGGCG GAAAGAACCA GCTGGGGCTC
2221    TAGGGGGTAT CCCCACGCGC CCTGTAGCGG CGCATTAAGC GCGGCGGGTG TGGTGGTTAC
2281    GCGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC GCTCCTTTCG CTTTCTTCCC
2341    TTCCTTTCTC GCCACGTTCG CCGGCTTTCC CCGTCAAGCT CTAAATCGGG GGCTCCCTTT
2401    AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAAA AAACTTGATT AGGGTGATGG
2461    TTCACGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTCGC CCTTTGACGT TGGAGTCCAC
2521    GTTCTTTAAT AGTGGACTCT TGTTCCAAAC TGGAACAACA CTCAACCCTA TCTCGGTCTA
2581    TTCTTTTGAT TTATAAGGGA TTTTGCCGAT TTCGGCCTAT TGGTTAAAAA ATGAGCTGAT
2641    TTAACAAAAA TTTAACGCGA ATTAATTCTG TGGAATGTGT GTCAGTTAGG GTGTGGAAAG
2701    TCCCCAGGCT CCCCAGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA GTCAGCAACC
2761    AGGTGTGGAA AGTCCCCAGG CTCCCCAGCA GGCAGAAGTA TGCAAAGCAT GCATCTCAAT
2821    TAGTCAGCAA CCATAGTCCC GCCCCTAACT CCGCCCATCC CGCCCCTAAC TCCGCCCAGT
2881    TCCGCCCATT CTCCGCCCCA TGGCTGACTA ATTTTTTTTA TTTATGCAGA GGCCGAGGCC
2941    GCCTCTGCCT CTGAGCTATT CCAGAAGTAG TGAGGAGGCT TTTTTGGAGG CCTAGGCTTT
3001    TGCAAAAAGC TCCCGGGAGC TTGTATATCC ATTTTCGGAT CTGATCAAGA GACAGGATGA
3061    GGATCGTTTC GCATGATTGA ACAAGATGGA TTGCACGCAG GTTCTCCGGC CGCTTGGGTG
3121    GAGAGGCTAT TCGGCTATGA CTGGGCACAA CAGACAATCG GCTGCTCTGA TGCCGCCGTG
3181    TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT CTTTTTGTCA AGACCGACCT GTCCGGTGCC
3241    CTGAATGAAC TGCAGGACGA GGCAGCGCGG CTATCGTGGC TGGCCACGAC GGGCGTTCCT
3301    TGCGCAGCTG TGCTCGACGT TGTCACTGAA GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA
3361    GTGCCGGGGC AGGATCTCCT GTCATCTCAC CTTGCTCCTG CCGAGAAAGT ATCCATCATG
3421    GCTGATGCAA TGCGGCGGCT GCATACGCTT GATCCGGCTA CCTGCCCATT CGACCACCAA
3481    GCGAAACATC GCATCGAGCG AGCACGTACT CGGATGGAAG CCGGTCTTGT CGATCAGGAT
3541    GATCTGGACG AAGAGCATCA GGGGCTCGCG CCAGCCGAAC TGTTCGCCAG GCTCAAGGCG
3601    CGCATGCCCG ACGGCGAGGA TCTCGTCGTG ACCCATGGCG ATGCCTGCTT GCCGAATATC
3661    ATGGTGGAAA ATGGCCGCTT TTCTGGATTC ATCGACTGTG GCCGGCTGGG TGTGGCGGAC
3721    CGCTATCAGG ACATAGCGTT GGCTACCCGT GATATTGCTG AAGAGCTTGG CGGCGAATGG
3781    GCTGACCGCT TCCTCGTGCT TTACGGTATC GCCGCTCCCG ATTCGCAGCG CATCGCCTTC
3841    TATCGCCTTC TTGACGAGTT CTTCTGAGCG GGACTCTGGG GTTCGAAATG ACCGACCAAG
3901    CGACGCCCAA CCTGCCATCA CGAGATTTCG ATTCCACCGC CGCCTTCTAT GAAAGGTTGG
3961    GCTTCGGAAT CGTTTTCCGG GACGCCGGCT GGATGATCCT CCAGCGCGGG GATCTCATGC
4021    TGGAGTTCTT CGCCCACCCC AACTTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAAGCA
4081    ATAGCATCAC AAATTTCACA AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTTGT
4141    CCAAACTCAT CAATGTATCT TATCATGTCT GTATACCGTC GACCTCTAGC TAGAGCTTGG
4201    CGTAATCATG GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA
4261    ACATACGAGC CGGAAGCATA AAGTGTAAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA
4321    CATTAATTGC GTTGCGCTCA CTGCCCGCTT TCCAGTCGGG AAACCTGTCG TGCCAGCTGC
4381    ATTAATGAAT CGGCCAACGC GCGGGGAGAG GCGGTTTGCG TATTGGGCGC TCTTCCGCTT
4441    CCTCGCTCAC TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT
4501    CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG
4561    CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTTCCATA
```

```
4621    GGCTCCGCCC CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC
4681    CGACAGGACT ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG
4741    TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC
4801    TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC TCCAAGCTGG
4861    GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC
4921    TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA
4981    TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG
5041    GCTACACTAG AAGAACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA
5101    AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT TTTTTTGTTT
5161    GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG ATCTTTTCTA
5221    CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC ATGAGATTAT
5281    CAAAAAGGAT CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTAAA TCAATCTAAA
5341    GTATATATGA GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT
5401    CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG TAGATAACTA
5461    CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA GACCCACGCT
5521    CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG
5581    GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA
5641    GTAGTTCGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT
5701    CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA
5761    CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA
5821    GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT AATTCTCTTA
5881    CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT
5941    GAGAATAGTG TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG GATAATACCG
6001    CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC
6061    TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT GCACCCAACT
6121    GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA GGAAGGCAAA
6181    ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT
6241    TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT
6301    GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA GTGCCACCTG
6361    ACGTC
```

8.1.10.3. Vector sequence of pTS691

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241     GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301     TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
361     CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421     ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481     ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541     ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601     TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661     ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
841     CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
901     GTTTAAACTT AAGCTTGGTA CCGAGCTCGG ATCCACCATG GCCAAGGACT TTCAAGATAT
961     CCAGCAGCTG AGCTCGGAGG AAAATGACCA TCCTTTCCAT CAAGGGCCAC CTCCTGCCCA
1021    GCCCCTGGCA CAGCGTCTCT GCTCCATGGT CTGCTTCAGT CTGCTTGCCC TGAGCTTCAA
1081    CATCCTGCTG CTGGTGGTCA TCTGTGTGAC TGGGTCCCAA AGTGCACAGC TGCAAGCCGA
1141    GCTGCGGAGC CTGAAGGAAG CTTTCAGCAA CTTCTCCTCG AGCACCCTGA CGGAGGTCCA
1201    GGCAATCAGC ACCCACGGAG GCAGCGTGGG TGACAAGATC ACATCCCTAG GAGCCAAGCT
1261    GGAGAAACAG CAGCAGGACC TGAAAGCAGA TCACGATGCC CTGCTCTTCC ATCTGAAGCA
1321    CTTCCCCGTG GACCTGCGCT TCGTGGCCTG CCAGATGGAG CTCCTCCACA GCAACGGCTC
1381    CCAAAGGACC TGCTGCCCCG TCAACTGGGT GGAGCACCAA GGCAGCTGCT ACTGGTTCTC
1441    TCACTCCGGG AAGGCCTGGG CTGAGGCGGA GAAGTACTGC CAGCTGGAGA ACGCACACCT
1501    GGTGGTCATC AACTCCTGGG AGGAGCAGAA ATTCATTGTA CAACACACGA ACCCCTTCAA
1561    TACCTGGATA GGTCTCACGG ACAGTGATGG CTCTTGGAAA TGGGTGGATG GCACAGACTA
```

```
1621    TAGGCACAAC TACAAGAACT GGGCTGTCAC TCAGCCAGAT AATTGGCACG GGCACGAGCT
1681    GGGTGGAAGT GAAGACTGTG TTGAAGTCCA GCCGGATGGC CGCTGGAACG ATGACTTCTG
1741    CCTGCAGGTG TACCGCTGGG TGTGTGAGAA AAGGCGGAAT GCCACCGGCG AGGTGGCCTG
1801    ATCTAGAGGG CCCTTCGAAC AAAAACTCAT CTCAGAAGAG GATCTGAATA TGCATACCGG
1861    TCATCATCAC CATCACCATT GAGTTTAAAC CCGCTGATCA GCCTCGACTG TGCCTTCTAG
1921    TTGCCAGCCA TCTGTTGTTT GCCCCTCCCC CGTGCCTTCC TTGACCCTGG AAGGTGCCAC
1981    TCCCACTGTC CTTTCCTAAT AAAATGAGGA AATTGCATCG CATTGTCTGA GTAGGTGTCA
2041    TTCTATTCTG GGGGGTGGGG TGGGGCAGGA CAGCAAGGGG GAGGATTGGG AAGACAATAG
2101    CAGGCATGCT GGGGATGCGG TGGGCTCTAT GGCTTCTGAG GCGGAAAGAA CCAGCTGGGG
2161    CTCTAGGGGG TATCCCCACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG GTGTGGTGGT
2221    TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT TCGCTTTCTT
2281    CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC
2341    TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTAGGGTGA
2401    TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA CGTTGGAGTC
2461    CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC CTATCTCGGT
2521    CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGCC TATTGGTTAA AAAATGAGCT
2581    GATTTAACAA AAATTTAACG CGAATTAATT CTGTGGAATG TGTGTCAGTT AGGGTGTGGA
2641    AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA
2701    ACCAGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA GTATGCAAAG CATGCATCTC
2761    AATTAGTCAG CAACCATAGT CCCGCCCCTA ACTCCGCCCA TCCCGCCCCT AACTCCGCCC
2821    AGTTCCGCCC ATTCTCCGCC CCATGGCTGA CTAATTTTTT TTATTTATGC AGAGGCCGAG
2881    GCCGCCTCTG CCTCTGAGCT ATTCCAGAAG TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC
2941    TTTTGCAAAA AGCTCCCGGG AGCTTGTATA TCCATTTTCG GATCTGATCA AGAGACAGGA
3001    TGAGGATCGT TTCGCATGAT TGAACAAGAT GGATTGCACG CAGGTTCTCC GGCCGCTTGG
3061    GTGGAGAGGC TATTCGGCTA TGACTGGGCA CAACAGACAA TCGGCTGCTC TGATGCCGCC
3121    GTGTTCCGGC TGTCAGCGCA GGGGCGCCCG GTTCTTTTTG TCAAGACCGA CCTGTCCGGT
3181    GCCCTGAATG AACTGCAGGA CGAGGCAGCG CGGCTATCGT GGCTGGCCAC GACGGGCGTT
3241    CCTTGCGCAG CTGTGCTCGA CGTTGTCACT GAAGCGGGAA GGGACTGGCT GCTATTGGGC
3301    GAAGTGCCGG GGCAGGATCT CCTGTCATCT CACCTTGCTC CTGCCGAGAA AGTATCCATC
3361    ATGGCTGATG CAATGCGGCG GCTGCATACG CTTGATCCGG CTACCTGCCC ATTCGACCAC
3421    CAAGCGAAAC ATCGCATCGA GCGAGCACGT ACTCGGATGG AAGCCGGTCT TGTCGATCAG
3481    GATGATCTGG ACGAAGAGCA TCAGGGGCTC GCGCCAGCCG AACTGTTCGC CAGGCTCAAG
3541    GCGCGCATGC CCGACGGCGA GGATCTCGTC GTGACCCATG GCGATGCCTG CTTGCCGAAT
3601    ATCATGGTGG AAAATGGCCG CTTTTCTGGA TTCATCGACT GTGGCCGGCT GGGTGTGGCG
3661    GACCGCTATC AGGACATAGC GTTGGCTACC CGTGATATTG CTGAAGAGCT TGGCGGCGAA
3721    TGGGCTGACC GCTTCCTCGT GCTTTACGGT ATCGCCGCTC CCGATTCGCA GCGCATCGCC
3781    TTCTATCGCC TTCTTGACGA GTTCTTCTGA GCGGGACTCT GGGGTTCGAA ATGACCGACC
3841    AAGCGACGCC CAACCTGCCA TCACGAGATT TCGATTCCAC CGCCGCCTTC TATGAAAGGT
3901    TGGGCTTCGG AATCGTTTTC CGGGACGCCG GCTGGATGAT CCTCCAGCGC GGGGATCTCA
3961    TGCTGGAGTT CTTCGCCCAC CCCAACTTGT TTATTGCAGC TTATAATGGT TACAAATAAA
4021    GCAATAGCAT CACAAATTTC ACAAATAAAG CATTTTTTTC ACTGCATTCT AGTTGTGGTT
4081    TGTCCAAACT CATCAATGTA TCTTATCATG TCTGTATACC GTCGACCTCT AGCTAGAGCT
4141    TGGCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC ACAATTCCAC
4201    ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA GTGAGCTAAC
4261    TCACATTAAT TGCGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG TCGTGCCAGC
4321    TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG CGCTCTTCCG
4381    CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG GTATCAGCTC
4441    ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA AAGAACATGT
4501    GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC
4561    ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA
4621    ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC
4681    CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG
4741    CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC
4801    TGGGCTGTGT GCACGAACCC CCCGTTCAGC CCGACCGCTG CGCCTTATCC GGTAACTATC
4861    GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA
4921    GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT
4981    ACGGCTACAC TAGAAGAACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA GTTACCTTCG
5041    GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC GGTTTTTTTG
5101    TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA AGAAGATCCT TTGATCTTTT
5161    CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCACGTTA AGGGATTTTG GTCATGAGAT
```

```
5221     TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT
5281     AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA
5341     TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA
5401     CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG CGAGACCCAC
5461     GCTCACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA
5521     GTGGTCCTGC AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG
5581     TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG
5641     TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGCGAG
5701     TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG
5761     TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC
5821     TTACTGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT
5881     TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAATA CGGGATAATA
5941     CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA
6001     AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA
6061     ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC
6121     AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC
6181     TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG
6241     AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC
6301     CTGACGTC
```

8.1.10.4. Vector sequence of pTS1032

```
1        ATCACCTCGA GTTTACTCCC TATCAGTGAT AGAGAACGTA TGAAGAGTTT ACTCCCTATC
61       AGTGATAGAG AACGTATGCA GACTTTACTC CCTATCAGTG ATAGAGAACG TATAAGGAGT
121      TTACTCCCTA TCAGTGATAG AGAACGTATG ACCAGTTTAC TCCCTATCAG TGATAGAGAA
181      CGTATCTACA GTTTACTCCC TATCAGTGAT AGAGAACGTA TATCCAGTTT ACTCCCTATC
241      AGTGATAGAG AACGTATAAG CTTTGCTTAT GTAAACCAGG GCGCCTATAA AAGAGTGCTG
301      ATTTTTTGAG TAAACTTCAA TTCCACAACA CTTTTGTCTT ATACCAACTT TCCGTACCAC
361      TTCCTACCCT CGTAAAGGTA CCATGGTGAG CAAGGGCGAG GAGCTGTTCA CCGGGGTGGT
421      GCCCATCCTG GTCGAGCTGG ACGGCGACGT AAACGGCCAC AAGTTCAGCG TGTCCGGCGA
481      GGGCGAGGGC GATGCCACCT ACGGCAAGCT GACCCTGAAG TTCATCTGCA CCACCGGCAA
541      GCTGCCCGTG CCCTGGCCCA CCCTCGTGAC CACCCTGACC TACGGCGTGC AGTGCTTCAG
601      CCGCTACCCC GACCACATGA AGCAGCACGA CTTCTTCAAG TCCGCCATGC CCGAAGGCTA
661      CGTCCAGGAG CGCACCATCT TCTTCAAGGA CGACGGCAAC TACAAGACCC GCGCCGAGGT
721      GAAGTTCGAG GGCGACACCC TGGTGAACCG CATCGAGCTG AAGGGCATCG ACTTCAAGGA
781      GGACGGCAAC ATCCTGGGGC ACAAGCTGGA GTACAACTAC AACAGCCACA ACGTCTATAT
841      CATGGCCGAC AAGCAGAAGA ACGGCATCAA GGTGAACTTC AAGATCCGCC ACAACATCGA
901      GGACGGCAGC GTGCAGCTCG CCGACCACTA CCAGCAGAAC ACCCCCATCG GCGACGGCCC
961      CGTGCTGCTG CCCGACAACC ACTACCTGAG CACCCAGTCC GCCCTGAGCA AAGACCCCAA
1021     CGAGAAGCGC GATCACATGG TCCTGCTGGA GTTCGTGACC GCCGCCGGGA TCACTCTCGG
1081     CATGGACGAG CTGTACAAGG AATTCTAAAC TAGTAGACCA CCTCCCCTGC GAGCTAAGCT
1141     GGACAGCCAA TGACGGGTAA GAGAGTGACA TTTTTCACTA ACCTAAGACA GGAGGGCCGT
1201     CAGAGCTACT GCCTAATCCA AAGACGGGTA AAAGTGATAA AAATGTATCA CTCCAACCTA
1261     AGACAGGCGC AGCTTCCGAG GGATTTGAGA TCCAGACATG ATAAGATACA TTGATGAGTT
1321     TGGACAAACC AAAACTAGAA TGCAGTGAAA AAAATGCCTT ATTTGTGAAA TTTGTGATGC
1381     TATTGCCTTA TTTGTAACCA TTATAAGCTG CAATAAACAA GTTTGATATC TATAACAAGA
1441     AAATATATAT ATAATAAGTT ATCACGTAAG TAGAACATGA AATAACAATA TAATTATCGT
1501     ATGAGTTAAA TCTTAAAAGT CACGTAAAAG ATAATCATGC GTCATTTTGA CTCACGCGGT
1561     CGTTATAGTT CAAAATCAGT GACACTTACC GCATTGACAA GCACGCCTCA CGGGAGCTCC
1621     AAGCGGCGAC TGAGATGTCC TAAATGCACA GCGACGGATT CGCGCTATTT AGAAAGAGAG
1681     AGCAATATTT CAAGAATGCA TGCGTCAATT TTACGCAGAC TATCTTTCTA GGGTTAAGAA
1741     TTCACTGGCC GTCGTTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA
1801     TCGCCTTGCA GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA
1861     TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG CGCCTGATGC GGTATTTTCT
1921     CCTTACGCAT CTGTGCGGTA TTTCACACCG CATATGGTGC ACTCTCAGTA CAATCTGCTC
1981     TGATGCCGCA TAGTTAAGCC AGCCCCGACA CCCGCCAACA CCCGCTGACG CGCCCTGACG
2041     GGCTTGTCTG CTCCCGGCAT CCGCTTACAG ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT
2101     GTGTCAGAGG TTTTCACCGT CATCACCGAA ACGCGCGAGA CGAAAGGGCC TCGTGATACG
2161     CCTATTTTTA TAGGTTAATG TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT
2221     TCGGGGAAAT GTGCGCGGAA CCCCTATTTG TTTATTTTTC TAAATACATT CAAATATGTA
```

```
2281     TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT
2341     GAGTATTCAA CATTTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT
2401     TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGTGCACG
2461     AGTGGGTTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTCGCCCCGA
2521     AGAACGTTTT CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG
2581     TATTGACGCC GGGCAAGAGC AACTCGGTCG CCGCATACAC TATTCTCAGA ATGACTTGGT
2641     TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG
2701     CAGTGCTGCC ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG
2761     AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA
2821     TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
2881     TGTAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA CTCTAGCTTC
2941     CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC
3001     GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG
3061     CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC
3121     GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC
3181     ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATTT
3241     AAAACTTCAT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC
3301     CAAAATCCCT TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
3361     AGGATCTTCT TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC
3421     ACCGCTACCA GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT
3481     AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG
3541     CCACCACTTC AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC
3601     AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT
3661     ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA
3721     GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT
3781     TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG
3841     CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA
3901     CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA
3961     CGCCAGCAAC GCGGCCTTTT TACGGTTCCT GGCCTTTTGC TGGCCTTTTG CTCACATGTT
4021     CTTTCCTGCG TTATCCCCTG ATTCTGTGGA TAACCGTATT ACCGCCTTTG AGTGAGCTGA
4081     TACCGCTCGC CGCAGCCGAA CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA
4141     GCGCCCAATA CGCAAACCGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGGCA
4201     CGACAGGTTT CCCGACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTTAGCT
4261     CACTCATTAG GCACCCCAGG CTTTACACTT TATGCTTCCG GCTCGTATGT TGTGTGGAAT
4321     TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGGTCGA
4381     CTTAACCCTA GAAAGATAAT CATATTGTGA CGTACGTTAA AGATAATCAT GCGTAAAATT
4441     GACGCATGTG TTTTATCGGT CTGTATATCG AGGTTTATTT ATTAATTTGA ATAGATATTA
4501     AGTTTTATTA TATTTACACT TACATACTAA TAATAAATTC AACAAACAAT TTATTTATGT
4561     TTATTTATTT ATTAAAAAAA AACAAAAACT CAAAATTTCT TCTATAAAGT AACAAAACTT
4621     TTAGCAGTGA AAAAAATGCT TTATTTGTGA AATTTGTGAT GCTATTGCTT TATTTGTAAC
4681     CATTATAAGC TGCAATAAAC AAGTTAACAA CAACAATTGC ATTCATTTTA TGTTTCAGGT
4741     TCAGGGGGAG GTGTGGGAGG TTTTTTAAAG CAAGTAAAAC CTCTACAAAT GTGGTATGGC
4801     TGATTATGAT CCTCTGGAGA TCCTAGGCGC TCAGAAGAAC TCGTCAAGAA GGCGATAGAA
4861     GGCGATGCGC TGCGAATCGG GAGCGGCGAT ACCGTAAAGC ACGAGGAAGC GGTCAGCCCA
4921     TTCGCCGCCA AGCTCTTCAG CAATATCACG GGTAGCCAAC GCTATGTCCT GATAGCGGTC
4981     CGCCACACCC AGCCGGCCAC AGTCGATGAA TCCAGAAAAG CGGCCATTTT CCACCATGAT
5041     ATTCGGCAAG CAGGCATCGC CATGGGTCAC GACGAGATCC TCGCCGTCGG GCATGCGCGC
5101     CTTGAGCCTG GCGAACAGTT CGGCTGGCGC GAGCCCCTGA TGCTCTTCGT CCAGATCATC
5161     CTGATCGACA AGACCGGCTT CCATCCGAGT ACGTGCTCGC TCGATGCGAT GTTTCGCTTG
5221     GTGGTCGAAT GGGCAGGTAG CCGGATCAAG CGTATGCAGC CGCCGCATTG CATCAGCCAT
5281     GATGGATACT TTCTCGGCAG GAGCAAGGTG AGATGACAGG AGATCCTGCC CCGGCACTTC
5341     GCCCAATAGC AGCCAGTCCC TTCCCGCTTC AGTGACAACG TCGAGCACAG CTGCGCAAGG
5401     AACGCCCGTC GTGGCCAGCC ACGATAGCCG CGCTGCCTCG TCCTGCAGTT CATTCAGGGC
5461     ACCGGACAGG TCGGTCTTGA CAAAAAGAAC CGGGCGCCCC TGCGCTGACA GCCGGAACAC
5521     GGCGGCATCA GAGCAGCCGA TTGTCTGTTG TGCCCAGTCA TAGCCGAATA GCCTCTCCAC
5581     CCAAGCGGCC GGAGAACCTG CGTGCAATCC ATCTTGTTCA ATCATGGGGC CGGGGTTCTC
5641     CTCCACGTCA CCGGCCTGCT TCAGCAGGCT GAAGTTGGTG GCGCCGCTGC CCCCGGGGAG
5701     CATGTCAAGG TCAAAATCGT CAAGAGCGTC AGCAGGCAGC ATATCAAGGT CAAAGTCGTC
5761     AAGGGCATCG GCTGGGAGCA TGTCTAAGTC AAAATCGTCA AGGGCGTCGG TCGGCCCGCC
5821     GCTTTCGCAC TTTAGCTGTT TCTCCAGGCC ACATATGATT AGTTCCAGGC CGAAAAGGAA
```

```
5881    GGCAGGTTCG GCTCCCTGCC GGTCGAACAG CTCAATTGCT TGTTTCAGAA GTGGGGGCAT
5941    AGAATCGGTG GTAGGTGTCT CTCTTTCCTC TTTTGCTACT TGATGCTCCT GTTCCTCCAA
6001    TACGCAGCCC AGTGTAAAGT GGCCCACGGC GGACAGAGCG TACAGTGCGT TCTCCAGGGA
6061    GAAGCCTTGC TGACACAGGA ACGCGAGCTG ATTTTCCAGG GTTTCGTACT GTTTCTCTGT
6121    TGGGCGGGTG CCGAGATGCA CTTTAGCCCC GTCGCGATGT GAGAGGAGAG CACAGCGGTA
6181    TGACTTGGCG TTGTTCCGCA GAAAGTCTTG CCATGACTCG CCTTCCAGGG GGCAGGAGTG
6241    GGTATGATGC CTGTCCAGCA TCTCGATTGG CAGGGCATCG AGCAGGGCCC GCTTGTTCTT
6301    CACGTGCCAG TACAGGGTAG GCTGCTCAAC TCCCAGCTTT TGAGCGAGTT TCCTTGTCGT
6361    CAGGCCTTCG ATACCGACTC CATTGAGTAA TTCCAGAGCA GAGTTTATGA CTTTGCTCTT
6421    GTCCAGTCTA GACATCTTAT CGTCATCGTC TTTGTAATCC ATGGTGGCGG ATCCCGCGTC
6481    ACGACACCTG TGTTCTGGCG GCAAACCCGT TGCGAAAAAG AACGTTCACG GCGACTACTG
6541    CACTTATATA CGGTTCTCCC CCACCCTCGG GAAAAAGGCG GAGCCAGTAC ACGACATCAC
6601    TTTCCCAGTT TACCCCGCGC CACCTTCTCT AGGCACCGGT TCAATTGCCG ACCCCTCCCC
6661    CCAACTTCTC GGGGACTGTG GGCGATGTGC GCTCTGCCCA CTGACGGGCA CCGGAGCCAC
6721    TCGAGTGGAA TT
```

8.1.10.5. Vector sequence of pTS1037

```
1       ATCACCTCGA GTTTACTCCC TATCAGTGAT AGAGAACGTA TGAAGAGTTT ACTCCCTATC
61      AGTGATAGAG AACGTATGCA GACTTTACTC CCTATCAGTG ATAGAGAACG TATAAGGAGT
121     TTACTCCCTA TCAGTGATAG AGAACGTATG ACCAGTTTAC TCCCTATCAG TGATAGAGAA
181     CGTATCTACA GTTTACTCCC TATCAGTGAT AGAGAACGTA TATCCAGTTT ACTCCCTATC
241     AGTGATAGAG AACGTATAAG CTTTGCTTAT GTAAACCAGG GCGCCTATAA AAGAGTGCTG
301     ATTTTTTGAG TAAACTTCAA TTCCACAACA CTTTTGTCTT ATACCAACTT TCCGTACCAC
361     TTCCTACCCT CGTAAAGGTA CCGCGGCCGC CACCATGGAC AAAGACTGCG AAATGAAGCG
421     CACCACCCTG GATAGCCCTC TGGGCAAGCT GGAACTGTCT GGGTGCGAAC AGGGCCTGCA
481     CCGTATCATC TTCCTGGGCA AAGGAACATC TGCCGCCGAC GCCGTGGAAG TGCCTGCCCC
541     AGCCGCCGTG CTGGGCGGAC CAGAGCCACT GATGCAGGCC ACCGCCTGGC TCAACGCCTA
601     CTTTCACCAG CCTGAGGCCA TCGAGGAGTT CCCTGTGCCA GCCCTGCACC ACCCAGTGTT
661     CCAGCAGGAG AGCTTTACCC GCCAGGTGCT GTGGAAACTG CTGAAAGTGG TGAAGTTCGG
721     AGAGGTCATC AGCTACAGCC ACCTGGCCGC CCTGGCCGGC AATCCCGCCG CCACCGCCGC
781     CGTGAAAACC GCCCTGAGCG GAAATCCCGT GCCCATTCTG ATCCCCTGCC ACCGGGTGGT
841     GCAGGGCGAC CTGGACGTGG GGGGCTACGA GGGCGGGCTC GCCGTGAAAG AGTGGCTGCT
901     GGCCCACGAG GGCCACAGAC TGGGCAAGCC TGGGCTGGGT CCTGCAGGCG GAGGCGCGCC
961     AGGGTCTGGC GGCGGCAGTA AGGCAGAACG CATGGGTTTC ACAGAGGTAA CCCCAGTGAC
1021    AGGGGCCAGT CTCAGAAGAA CTATGCTCCT CCTCTCAAGG TCCCCAGAAG CACAGCCAAA
1081    GACACTCCCT CTCACTGGCA GCACCTTCCA TGACCAGATA GCCATGCTGA GCCACCGGTG
1141    CTTCAACACT CTGACTAACA GCTTCCAGCC CTCCTTGCTC GGCCGCAAGA TTCTGGCCGC
1201    CATCATTATG AAAAAAGACT CTGAGGACAT GGGTGTCGTC GTCAGCTTGG GAACAGGGAA
1261    TCGCTGTGTA AAAGGAGATT CTCTCAGCCT AAAAGGAGAA ACTGTCAATG ACTGCCATGC
1321    AGAAATAATC TCCCGGAGAG GCTTCATCAG GTTTCTCTAC AGTGAGTTAA TGAAATACAA
1381    CTCCCAGACT GCGAAGGATA GTATATTTGA ACCTGCTAAG GGAGGAGAAA AGCTCCAAAT
1441    AAAAAAGACT GTGTCATTCC ATCTGTATAT CAGCACTGCT CCGTGTGGAG ATGGCGCCCT
1501    CTTTGACAAG TCCTGCAGCG ACCGTGCTAT GGAAAGCACA GAATCCCGCC ACTACCCTGT
1561    CTTCGAGAAT CCCAAACAAG GAAAGCTCCG CACCAAGGTG GAGAACGGAC AAGGCACAAT
1621    CCCTGTGGAA TCCAGTGACA TTGTGCCTAC GTGGGATGGC ATTCGGCTCG GGGAGAGACT
1681    CCGTACCATG TCCTGTAGTG ACAAAATCCT ACGCTGGAAC GTGCTGGGCC TGCAAGGGGC
1741    ACTGTTGACC CACTTCCTGC AGCCCATTTA TCTCAAATCT GTCACATTGG GTTACCTTTT
1801    CAGCCAAGGG CATCTGACCC GTGCTATTTG CTGTCGTGTG ACAAGAGATG GGAGTGCATT
1861    TGAGGATGGA CTACGACATC CCTTTATTGT CAACCACCCC AAGGTTGGCA GAGTCAGCAT
1921    ATATGATTCC AAAAGGCAAT CCGGGAAGAC TAAGGAGACA AGCGTCAACT GGTGTCTGGC
1981    TGATGGCTAT GACCTGGAGA TCCTGGACGG TACCAGAGGC ACTGTGGATG GGCCACGGAA
2041    TGAATTGTCC CGGGTCTCCA AAAAGAACAT TTTTCTTCTA TTTAAGAAGC TCTGCTCCTT
2101    CCGTTACCGC AGGGATCTAC TGAGACTCTC CTATGGTGAG GCCAAGAAAG CTGCCCGTGA
2161    CTACGAGACG GCCAAGAACT ACTTCAAAAA AGGCCTGAAG GATATGGGCT ATGGGAACTG
2221    GATTAGCAAA CCCCAGGAGG AAAAGAACTT TTATCTCTGC CCAGTATAAA TCGATTAATT
2281    AACTAGTAGA CCACCTCCCC TGCGAGCTAA GCTGGACAGC CAATGACGGG TAAGAGAGTG
2341    ACATTTTTCA CTAACCTAAG ACAGGAGGGC CGTCAGAGCT ACTGCCTAAT CCAAAGACGG
2401    GTAAAAGTGA TAAAAATGTA TCACTCCAAC CTAAGACAGG CGCAGCTTCC GAGGGATTTG
2461    AGATCCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCAAAACTA GAATGCAGTG
```

```
2521     AAAAAAATGC CTTATTTGTG AAATTTGTGA TGCTATTGCC TTATTTGTAA CCATTATAAG
2581     CTGCAATAAA CAAGTTTGAT ATCTATAACA AGAAAATATA TATATAATAA GTTATCACGT
2641     AAGTAGAACA TGAAATAACA ATATAATTAT CGTATGAGTT AAATCTTAAA AGTCACGTAA
2701     AAGATAATCA TGCGTCATTT TGACTCACGC GGTCGTTATA GTTCAAAATC AGTGACACTT
2761     ACCGCATTGA CAAGCACGCC TCACGGGAGC TCCAAGCGGC GACTGAGATG TCCTAAATGC
2821     ACAGCGACGG ATTCGCGCTA TTTAGAAAGA GAGAGCAATA TTTCAAGAAT GCATGCGTCA
2881     ATTTTACGCA GACTATCTTT CTAGGGTTAA GAATTCACTG GCCGTCGTTT TACAACGTCG
2941     TGACTGGGAA AACCCTGGCG TTACCCAACT TAATCGCCTT GCAGCACATC CCCCTTTCGC
3001     CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT TGCGCAGCCT
3061     GAATGGCGAA TGGCGCCTGA TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTCACA
3121     CCGCATATGG TGCACTCTCA GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGCCCCG
3181     ACACCCGCCA ACACCCGCTG ACGCGCCCTG ACGGGCTTGT CTGCTCCCGG CATCCGCTTA
3241     CAGACAAGCT GTGACCGTCT CCGGGAGCTG CATGTGTCAG AGGTTTTCAC CGTCATCACC
3301     GAAACGCGCG AGACGAAAGG GCCTCGTGAT ACGCCTATTT TTATAGGTTA ATGTCATGAT
3361     AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT
3421     TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA
3481     AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT
3541     TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA
3601     AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA
3661     CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
3721     TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG
3781     TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA
3841     TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA
3901     CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT
3961     GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC
4021     CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
4081     ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA
4141     GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC
4201     TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA
4261     TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA
4321     ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA
4381     CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT
4441     CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT
4501     CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT
4561     GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC
4621     GGATCAAGAG CTACCAACTC TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC
4681     AAATACTGTT CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC
4741     GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC
4801     GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG
4861     AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
4921     CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
4981     TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC
5041     CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG
5101     ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT
5161     CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT
5221     GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA
5281     GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC
5341     CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAAGCGGG
5401     CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA
5461     CTTTATGCTT CCGGCTCGTA TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG
5521     AAACAGCTAT GACCATGATT ACGCCAAGGT CGACTTAACC CTAGAAAGAT AATCATATTG
5581     TGACGTACGT TAAAGATAAT CATGCGTAAA ATTGACGCAT GTGTTTTATC GGTCTGTATA
5641     TCGAGGTTTA TTTATTAATT TGAATAGATA TTAAGTTTTA TTATATTTAC ACTTACATAC
5701     TAATAATAAA TTCAACAAAC AATTTATTTA TGTTTATTTA TTTATTAAAA AAAAACAAAA
5761     ACTCAAAATT TCTTCTATAA AGTAACAAAA CTTTTAGCAG TGAAAAAAAT GCTTTATTTG
5821     TGAAATTTGT GATGCTATTG CTTTATTTGT AACCATTATA AGCTGCAATA AACAAGTTAA
5881     CAACAACAAT TGCATTCATT TTATGTTTCA GGTTCAGGGG GAGGTGTGGG AGGTTTTTTA
5941     AAGCAAGTAA AACCTCTACA AATGTGGTAT GGCTGATTAT GATCCTCTGG AGATCCTAGG
6001     CGCTCAGAAG AACTCGTCAA GAAGGCGATA GAAGGCGATG CGCTGCGAAT CGGGAGCGGC
6061     GATACCGTAA AGCACGAGGA AGCGGTCAGC CCATTCGCCG CCAAGCTCTT CAGCAATATC
```

```
6121    ACGGGTAGCC AACGCTATGT CCTGATAGCG GTCCGCCACA CCCAGCCGGC CACAGTCGAT
6181    GAATCCAGAA AAGCGGCCAT TTTCCACCAT GATATTCGGC AAGCAGGCAT CGCCATGGGT
6241    CACGACGAGA TCCTCGCCGT CGGGCATGCG CGCCTTGAGC CTGGCGAACA GTTCGGCTGG
6301    CGCGAGCCCC TGATGCTCTT CGTCCAGATC ATCCTGATCG ACAAGACCGG CTTCCATCCG
6361    AGTACGTGCT CGCTCGATGC GATGTTTCGC TTGGTGGTCG AATGGGCAGG TAGCCGGATC
6421    AAGCGTATGC AGCCGCCGCA TTGCATCAGC CATGATGGAT ACTTTCTCGG CAGGAGCAAG
6481    GTGAGATGAC AGGAGATCCT GCCCCGGCAC TTCGCCCAAT AGCAGCCAGT CCCTTCCCGC
6541    TTCAGTGACA ACGTCGAGCA CAGCTGCGCA AGGAACGCCC GTCGTGGCCA GCCACGATAG
6601    CCGCGCTGCC TCGTCCTGCA GTTCATTCAG GGCACCGGAC AGGTCGGTCT TGACAAAAAG
6661    AACCGGGCGC CCCTGCGCTG ACAGCCGGAA CACGGCGGCA TCAGAGCAGC CGATTGTCTG
6721    TTGTGCCCAG TCATAGCCGA ATAGCCTCTC CACCCAAGCG GCCGGAGAAC CTGCGTGCAA
6781    TCCATCTTGT TCAATCATGG GGCCGGGGTT CTCCTCCACG TCACCGGCCT GCTTCAGCAG
6841    GCTGAAGTTG GTGGCGCCGC TGCCCCCGGG GAGCATGTCA AGGTCAAAAT CGTCAAGAGC
6901    GTCAGCAGGC AGCATATCAA GGTCAAAGTC GTCAAGGGCA TCGGCTGGGA GCATGTCTAA
6961    GTCAAAATCG TCAAGGGCGT CGGTCGGCCC GCCGCTTTCG CACTTTAGCT GTTTCTCCAG
7021    GCCACATATG ATTAGTTCCA GGCCGAAAAG GAAGGCAGGT TCGGCTCCCT GCCGGTCGAA
7081    CAGCTCAATT GCTTGTTTCA GAAGTGGGGG CATAGAATCG GTGGTAGGTG TCTCTCTTTC
7141    CTCTTTTGCT ACTTGATGCT CCTGTTCCTC CAATACGCAG CCCAGTGTAA AGTGGCCCAC
7201    GGCGGACAGA GCGTACAGTG CGTTCTCCAG GGAGAAGCCT TGCTGACACA GGAACGCGAG
7261    CTGATTTTCC AGGGTTTCGT ACTGTTTCTC TGTTGGGCGG GTGCCGAGAT GCACTTTAGC
7321    CCCGTCGCGA TGTGAGAGGA GAGCACAGCG GTATGACTTG GCGTTGTTCC GCAGAAAGTC
7381    TTGCCATGAC TCGCCTTCCA GGGGGCAGGA GTGGGTATGA TGCCTGTCCA GCATCTCGAT
7441    TGGCAGGGCA TCGAGCAGGG CCCGCTTGTT CTTCACGTGC CAGTACAGGG TAGGCTGCTC
7501    AACTCCCAGC TTTTGAGCGA GTTTCCTTGT CGTCAGGCCT TCGATACCGA CTCCATTGAG
7561    TAATTCCAGA GCAGAGTTTA TGACTTTGCT CTTGTCCAGT CTAGACATCT TATCGTCATC
7621    GTCTTTGTAA TCCATGGTGG CGGATCCCGC GTCACGACAC CTGTGTTCTG GCGGCAAACC
7681    CGTTGCGAAA AAGAACGTTC ACGGCGACTA CTGCACTTAT ATACGGTTCT CCCCCACCCT
7741    CGGGAAAAAG GCGGAGCCAG TACACGACAT CACTTTCCCA GTTTACCCCG CGCCACCTTC
7801    TCTAGGCACC GGTTCAATTG CCGACCCCTC CCCCCAACTT CTCGGGGACT GTGGGCGATG
7861    TGCGCTCTGC CCACTGACGG GCACCGGAGC CACTCGAGTG GAATT
```

8.1.10.6. Vector sequence of pTS1040

```
1       ATCACCTCGA GTTTACTCCC TATCAGTGAT AGAGAACGTA TGAAGAGTTT ACTCCCTATC
61      AGTGATAGAG AACGTATGCA GACTTTACTC CCTATCAGTG ATAGAGAACG TATAAGGAGT
121     TTACTCCCTA TCAGTGATAG AGAACGTATG ACCAGTTTAC TCCCTATCAG TGATAGAGAA
181     CGTATCTACA GTTTACTCCC TATCAGTGAT AGAGAACGTA TATCCAGTTT ACTCCCTATC
241     AGTGATAGAG AACGTATAAG CTTTGCTTAT GTAAACCAGG GCGCCTATAA AAGAGTGCTG
301     ATTTTTTGAG TAAACTTCAA TTCCACAACA CTTTTGTCTT ATACCAACTT TCCGTACCAC
361     TTCCTACCCT CGTAAAGGTA CCGCGGCCGC CACCATGGAC AAAGACTGCG AAATGAAGCG
421     CACCACCCTG GATAGCCCTC TGGGCAAGCT GGAACTGTCT GGGTGCGAAC AGGGCCTGCA
481     CCGTATCATC TTCCTGGGCA AAGGAACATC TGCCGCCGAC GCCGTGGAAG TGCCTGCCCC
541     AGCCGCCGTG CTGGGCGGAC CAGAGCCACT GATGCAGGCC ACCGCCTGGC TCAACGCCTA
601     CTTTCACCAG CCTGAGGCCA TCGAGGAGTT CCCTGTGCCA GCCCTGCACC ACCCAGTGTT
661     CCAGCAGGAG AGCTTTACCC GCCAGGTGCT GTGGAAACTG CTGAAAGTGG TGAAGTTCGG
721     AGAGGTCATC AGCTACAGCC ACCTGGCCGC CCTGGCCGGC AATCCCGCCG CCACCGCCGC
781     CGTGAAAACC GCCCTGAGCG GAAATCCCGT GCCCATTCTG ATCCCCTGCC ACCGGGTGGT
841     GCAGGGCGAC CTGGACGTGG GGGGCTACGA GGGCGGGCTC GCCGTGAAAG AGTGGCTGCT
901     GGCCCACGAG GGCCACAGAC TGGGCAAGCC TGGGCTGGGT CCTGCAGGCG GAGGCGCGCC
961     AGGGTCTGGC GGCGGCAGTA AGAAGCTTGC CAAGGCCCGG GCTGCGCAGT CTGCCCTGGC
1021    CGCCATTTTT AACTTGCACT TGGATCAGAC GCCATCTCGC CAGCCTATTC CCAGTGAGGG
1081    TCTTCAGCTG CATTTACCGC AGGTTTTAGC TGACGCTGTC TCACGCCTGG TCCTGGGTAA
1141    GTTTGGTGAC CTGACCGACA ACTTCTCCTC CCCTCACGCT CGCAGAAAAG TGCTGGCTGG
1201    AGTCGTCATG ACAACAGGCA CAGATGTTAA AGATGCCAAG GTGATAAGTG TTTCTACAGG
1261    AACAAAATGT ATTAATGGTG AATACATGAG TGATCGTGGC CTTGCATTAA ATGACTGCCA
1321    TGCAGAAATA ATATCTCGGA GATCCTTGCT CAGATTTCTT TATACACAAC TTGAGCTTTA
1381    CTTAAATAAC AAAGATGATC AAAAAAGATC CATCTTTCAG AAATCAGAGC GAGGGGGGTT
1441    TAGGCTGAAG GAGAATGTCC AGTTTCATCT GTACATCAGC ACCTCTCCCT GTGGAGATGC
1501    CAGAATCTTC TCACCACATG AGCCAATCCT GGAAGAACCA GCAGATAGAC ACCCAAATCG
1561    TAAAGCAAGA GGACAGCTAC GGACCAAAAT AGAGTCTGGT CAGGGGACGA TTCCAGTGCG
```

```
1621    CTCCAATGCG AGCATCCAAA CGTGGGACGG GGTGCTGCAA GGGGAGCGGC TGCTCACCAT
1681    GTCCTGCAGT GACAAGATTG CACGCTGGAA CGTGGTGGGC ATCCAGGGAT CCCTGCTCAG
1741    CATTTTCGTG GAGCCCATTT ACTTCTCGAG CATCATCCTG GGCAGCCTTT ACCACGGGGA
1801    CCACCTTTCC AGGGCCATGT ACCAGCGGAT CTCCAACATA GAGGACCTGC CACCTCTCTA
1861    CACCCTCAAC AAGCCTTTGC TCAGTGGCAT CAGCAATGCA GAAGCACGGC AGCCAGGGAA
1921    GGCCCCCAAC TTCAGTGTCA ACTGGACGGT AGGCGACTCC GCTATTGAGG TCATCAACGC
1981    CACGACTGGG AAGGATGAGC TGGGCCGCGC GTCCCGCCTG TGTAAGCACG CGTTGTACTG
2041    TCGCTGGATG CGTGTGCACG GCAAGGTTCC CTCCCACTTA CTACGCTCCA AGATTACCAA
2101    ACCCAACGTG TACCATGAGT CCAAGCTGGC GGCAAAGGAG TACCAGGCCG CCAAGGCGCG
2161    TCTGTTCACA GCCTTCATCA AGGCGGGGCT GGGGGCCTGG GTGGAGAAGC CCACCGAGCA
2221    GGACCAGTTC TCACTCACGC CCTAAATCGA TTAATTAACT AGTAGACCAC CTCCCCTGCG
2281    AGCTAAGCTG GACAGCCAAT GACGGGTAAG AGAGTGACAT TTTTCACTAA CCTAAGACAG
2341    GAGGGCCGTC AGAGCTACTG CCTAATCCAA AGACGGGTAA AAGTGATAAA AATGTATCAC
2401    TCCAACCTAA GACAGGCGCA GCTTCCGAGG GATTTGAGAT CCAGACATGA TAAGATACAT
2461    TGATGAGTTT GGACAAACCA AAACTAGAAT GCAGTGAAAA AAATGCCTTA TTTGTGAAAT
2521    TTGTGATGCT ATTGCCTTAT TTGTAACCAT TATAAGCTGC AATAAACAAG TTTGATATCT
2581    ATAACAAGAA AATATATATA TAATAAGTTA TCACGTAAGT AGAACATGAA ATAACAATAT
2641    AATTATCGTA TGAGTTAAAT CTTAAAAGTC ACGTAAAAGA TAATCATGCG TCATTTTGAC
2701    TCACGCGGTC GTTATAGTTC AAAATCAGTG ACACTTACCG CATTGACAAG CACGCCTCAC
2761    GGGAGCTCCA AGCGGCGACT GAGATGTCCT AAATGCACAG CGACGGATTC GCGCTATTTA
2821    GAAAGAGAGA GCAATATTTC AAGAATGCAT GCGTCAATTT TACGCAGACT ATCTTTCTAG
2881    GGTTAAGAAT TCACTGGCCG TCGTTTTACA ACGTCGTGAC TGGGAAAACC CTGGCGTTAC
2941    CCAACTTAAT CGCCTTGCAG CACATCCCCC TTTCGCCAGC TGGCGTAATA GCGAAGAGGC
3001    CCGCACCGAT CGCCCTTCCC AACAGTTGCG CAGCCTGAAT GGCGAATGGC GCCTGATGCG
3061    GTATTTTCTC CTTACGCATC TGTGCGGTAT TTCACACCGC ATATGGTGCA CTCTCAGTAC
3121    AATCTGCTCT GATGCCGCAT AGTTAAGCCA GCCCCGACAC CCGCCAACAC CCGCTGACGC
3181    GCCCTGACGG GCTTGTCTGC TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG
3241    GAGCTGCATG TGTCAGAGGT TTTCACCGTC ATCACCGAAA CGCGCGAGAC GAAAGGGCCT
3301    CGTGATACGC CTATTTTTAT AGGTTAATGT CATGATAATA ATGGTTTCTT AGACGTCAGG
3361    TGGCACTTTT CGGGGAAATG TGCGCGGAAC CCCTATTTGT TTATTTTTCT AAATACATTC
3421    AAATATGTAT CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT ATTGAAAAAG
3481    GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT CCCTTTTTTG CGGCATTTTG
3541    CCTTCCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG AAGATCAGTT
3601    GGGTGCACGA GTGGGTTACA TCGAACTGGA TCTCAACAGC GGTAAGATCC TTGAGAGTTT
3661    TCGCCCCGAA GAACGTTTTC CAATGATGAG CACTTTTAAA GTTCTGCTAT GTGGCGCGGT
3721    ATTATCCCGT ATTGACGCCG GGCAAGAGCA ACTCGGTCGC CGCATACACT ATTCTCAGAA
3781    TGACTTGGTT GAGTACTCAC CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG
3841    AGAATTATGC AGTGCTGCCA TAACCATGAG TGATAACACT GCGGCCAACT TACTTCTGAC
3901    AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGCAC AACATGGGGG ATCATGTAAC
3961    TCGCCTTGAT CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG AGCGTGACAC
4021    CACGATGCCT GTAGCAATGG CAACAACGTT GCGCAAACTA TTAACTGGCG AACTACTTAC
4081    TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG GATAAAGTTG CAGGACCACT
4141    TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG CCGGTGAGCG
4201    TGGGTCTCGC GGTATCATTG CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT
4261    TATCTACACG ACGGGGAGTC AGGCAACTAT GGATGAACGA AATAGACAGA TCGCTGAGAT
4321    AGGTGCCTCA CTGATTAAGC ATTGGTAACT GTCAGACCAA GTTTACTCAT ATATACTTTA
4381    GATTGATTTA AAACTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC TTTTTGATAA
4441    TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCCGTAGA
4501    AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC
4561    AAAAAAACCA CCGCTACCAG CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT
4621    TCCGAAGGTA ACTGGCTTCA GCAGAGCGCA GATACCAAAT ACTGTTCTTC TAGTGTAGCC
4681    GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG CTCTGCTAAT
4741    CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT TGGACTCAAG
4801    ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT GCACACAGCC
4861    CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC TATGAGAAAG
4921    CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC
4981    AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG
5041    GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC TCGTCAGGGG GGCGGAGCCT
5101    ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCCTG GCCTTTTGCT GGCCTTTTGC
5161    TCACATGTTC TTTCCTGCGT TATCCCCTGA TTCTGTGGAT AACCGTATTA CCGCCTTTGA
```

```
5221     GTGAGCTGAT ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG TGAGCGAGGA
5281     AGCGGAAGAG CGCCCAATAC GCAAACCGCC TCTCCCCGCG CGTTGGCCGA TTCATTAATG
5341     CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGGCAGT GAGCGCAACG CAATTAATGT
5401     GAGTTAGCTC ACTCATTAGG CACCCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATGTT
5461     GTGTGGAATT GTGAGCGGAT AACAATTTCA CACAGGAAAC AGCTATGACC ATGATTACGC
5521     CAAGGTCGAC TTAACCCTAG AAAGATAATC ATATTGTGAC GTACGTTAAA GATAATCATG
5581     CGTAAAATTG ACGCATGTGT TTTATCGGTC TGTATATCGA GGTTTATTTA TTAATTTGAA
5641     TAGATATTAA GTTTTATTAT ATTTACACTT ACATACTAAT AATAAATTCA ACAAACAATT
5701     TATTTATGTT TATTTATTTA TTAAAAAAAA ACAAAAACTC AAAATTTCTT CTATAAAGTA
5761     ACAAAACTTT TAGCAGTGAA AAAAATGCTT TATTTGTGAA ATTTGTGATG CTATTGCTTT
5821     ATTTGTAACC ATTATAAGCT GCAATAAACA AGTTAACAAC AACAATTGCA TTCATTTTAT
5881     GTTTCAGGTT CAGGGGGAGG TGTGGGAGGT TTTTTAAAGC AAGTAAAACC TCTACAAATG
5941     TGGTATGGCT GATTATGATC CTCTGGAGAT CCTAGGCGCT CAGAAGAACT CGTCAAGAAG
6001     GCGATAGAAG GCGATGCGCT GCGAATCGGG AGCGGCGATA CCGTAAAGCA CGAGGAAGCG
6061     GTCAGCCCAT TCGCCGCCAA GCTCTTCAGC AATATCACGG GTAGCCAACG CTATGTCCTG
6121     ATAGCGGTCC GCCACACCCA GCCGGCCACA GTCGATGAAT CCAGAAAAGC GGCCATTTTC
6181     CACCATGATA TTCGGCAAGC AGGCATCGCC ATGGGTCACG ACGAGATCCT CGCCGTCGGG
6241     CATGCGCGCC TTGAGCCTGG CGAACAGTTC GGCTGGCGCG AGCCCCTGAT GCTCTTCGTC
6301     CAGATCATCC TGATCGACAA GACCGGCTTC CATCCGAGTA CGTGCTCGCT CGATGCGATG
6361     TTTCGCTTGG TGGTCGAATG GGCAGGTAGC CGGATCAAGC GTATGCAGCC GCCGCATTGC
6421     ATCAGCCATG ATGGATACTT TCTCGGCAGG AGCAAGGTGA GATGACAGGA GATCCTGCCC
6481     CGGCACTTCG CCCAATAGCA GCCAGTCCCT TCCCGCTTCA GTGACAACGT CGAGCACAGC
6541     TGCGCAAGGA ACGCCCGTCG TGGCCAGCCA CGATAGCCGC GCTGCCTCGT CCTGCAGTTC
6601     ATTCAGGGCA CCGGACAGGT CGGTCTTGAC AAAAAGAACC GGGCGCCCCT GCGCTGACAG
6661     CCGGAACACG GCGGCATCAG AGCAGCCGAT TGTCTGTTGT GCCCAGTCAT AGCCGAATAG
6721     CCTCTCCACC CAAGCGGCCG GAGAACCTGC GTGCAATCCA TCTTGTTCAA TCATGGGGCC
6781     GGGGTTCTCC TCCACGTCAC CGGCCTGCTT CAGCAGGCTG AAGTTGGTGG CGCCGCTGCC
6841     CCCGGGGAGC ATGTCAAGGT CAAAATCGTC AAGAGCGTCA GCAGGCAGCA TATCAAGGTC
6901     AAAGTCGTCA AGGGCATCGG CTGGGAGCAT GTCTAAGTCA AAATCGTCAA GGGCGTCGGT
6961     CGGCCCGCCG CTTTCGCACT TTAGCTGTTT CTCCAGGCCA CATATGATTA GTTCCAGGCC
7021     GAAAAGGAAG GCAGGTTCGG CTCCCTGCCG GTCGAACAGC TCAATTGCTT GTTTCAGAAG
7081     TGGGGGCATA GAATCGGTGG TAGGTGTCTC TCTTTCCTCT TTTGCTACTT GATGCTCCTG
7141     TTCCTCCAAT ACGCAGCCCA GTGTAAAGTG GCCCACGGCG GACAGAGCGT ACAGTGCGTT
7201     CTCCAGGGAG AAGCCTTGCT GACACAGGAA CGCGAGCTGA TTTTCCAGGG TTTCGTACTG
7261     TTTCTCTGTT GGGCGGGTGC CGAGATGCAC TTTAGCCCCG TCGCGATGTG AGAGGAGAGC
7321     ACAGCGGTAT GACTTGGCGT TGTTCCGCAG AAAGTCTTGC CATGACTCGC CTTCCAGGGG
7381     GCAGGAGTGG GTATGATGCC TGTCCAGCAT CTCGATTGGC AGGGCATCGA GCAGGGCCCG
7441     CTTGTTCTTC ACGTGCCAGT ACAGGGTAGG CTGCTCAACT CCCAGCTTTT GAGCGAGTTT
7501     CCTTGTCGTC AGGCCTTCGA TACCGACTCC ATTGAGTAAT TCCAGAGCAG AGTTTATGAC
7561     TTTGCTCTTG TCCAGTCTAG ACATCTTATC GTCATCGTCT TTGTAATCCA TGGTGGCGGA
7621     TCCCGCGTCA CGACACCTGT GTTCTGGCGG CAAACCCGTT GCGAAAAAGA ACGTTCACGG
7681     CGACTACTGC ACTTATATAC GGTTCTCCCC CACCCTCGGG AAAAAGGCGG AGCCAGTACA
7741     CGACATCACT TTCCCAGTTT ACCCCGCGCC ACCTTCTCTA GGCACCGGTT CAATTGCCGA
7801     CCCCTCCCCC CAACTTCTCG GGGACTGTGG GCGATGTGCG CTCTGCCCAC TGACGGGCAC
7861     CGGAGCCACT CGAGTGGAAT T
```

8.1.10.7. Vector sequence of pTS1070

```
1        GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61       CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121      CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181      TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
241      GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
301      TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
361      CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
421      ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
481      ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
541      ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
601      TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
661      ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
```

```
721     AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
781     GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CCCTATCAGT GATAGAGATC
841     TCCCTATCAG TGATAGAGAT CGTCGACGAG CTCGTTTAGT GAACCGTCAG ATCGCCTGGA
901     GACGCCATCC ACGCTGTTTT GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGGA
961     CTCTAGCGTT TAAACTTAAG CTTGGTACCG CCACCATGAC CAAGGAGTAT CAAGACCTTC
1021    AGCATCTGGA CAATGAGGAG AGTGACCACC ATCAGCTCAG AAAAGGGCCA CCTCCTCCCC
1081    AGCCCCTCCT GCAGCGTCTC TGCTCCGGAC CTCGCCTCCT CCTGCTCTCC CTGGGCCTCA
1141    GCCTCCTGCT GCTTGTGGTT GTCTGTGTGA TCGGATCCCA AAACTCCCAG CTGCAGGAGG
1201    AGCTGCGGGG CCTGAGAGAG ACGTTCAGCA ACTTCACAGC GAGCACGGAG GCCCAGGTCA
1261    AGGGCTTGAG CACCCAGGGA GGCAATGTGG GAAGAAAGAT GAAGTCGCTA GAGTCCCAGC
1321    TGGAGAAACA GCAGAAGGAC CTGAGTGAAG ATCACTCCAG CCTGCTGCTC CACGTGAAGC
1381    AGTTCGTGTC TGACCTGCGG AGCCTGAGCT GTCAGATGGC GGCGCTCCAG GGCAATGGCT
1441    CAGAAAGGAC CTGCTGCCCG GTCAACTGGG TGGAGCACGA GCGCAGCTGC TACTGGTTCT
1501    CTCGCTCCGG GAAGGCCTGG GCTGACGCCG ACAACTACTG CCGGCTGGAG GACGCGCACC
1561    TGGTGGTGGT CACGTCCTGG GAGGAGCAGA AATTTGTCCA GCACCACATA GGCCCTGTGA
1621    ACACCTGGAT GGGCCTCCAC GACCAAAACG GGCCCTGGAA GTGGGTGGAC GGGACGGACT
1681    ACGAGACGGG CTTCAAGAAC TGGAGGCCGG AGCAGCCGGA CGACTGGTAC GGCCACGGGC
1741    TCGGAGGAGG CGAGGACTGT GCCCACTTCA CCGACGACGG CCGCTGGAAC GACGACGTCT
1801    GCCAGAGGCC CTACCGCTGG GTCTGCGAGA CAGAGCTGGA CAAGGCCAGC CAGGAGCCAC
1861    CTCTCCTTTA AGCGGCCGCT CGAGTCTAGA GGGCCCGTTT AAACCCGCTG ATCAGCCTCG
1921    ACTGTGCCTT CTAGTTGCCA GCCATCTGTT GTTTGCCCCT CCCCCGTGCC TTCCTTGACC
1981    CTGGAAGGTG CCACTCCCAC TGTCCTTTCC TAATAAAATG AGGAAATTGC ATCGCATTGT
2041    CTGAGTAGGT GTCATTCTAT TCTGGGGGGT GGGGTGGGGC AGGACAGCAA GGGGGAGGAT
2101    TGGGAAGACA ATAGCAGGCA TGCTGGGGAT GCGGTGGGCT CTATGGCTTC TGAGGCGGAA
2161    AGAACCAGCT GGGGCTCTAG GGGGTATCCC CACGCGCCCT GTAGCGGCGC ATTAAGCGCG
2221    GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT
2281    CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG GCTTTCCCCG TCAAGCTCTA
2341    AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCAAAAAA
2401    CTTGATTAGG GTGATGGTTC ACGTACCTAG AAGTTCCTAT TCCGAAGTTC CTATTCTCTA
2461    GAAAGTATAG GAACTTCCTT GGCCAAAAAG CCTGAACTCA CCGCGACGTC TGTCGAGAAG
2521    TTTCTGATCG AAAAGTTCGA CAGCGTCTCC GACCTGATGC AGCTCTCGGA GGGCGAAGAA
2581    TCTCGTGCTT TCAGCTTCGA TGTAGGAGGG CGTGGATATG TCCTGCGGGT AAATAGCTGC
2641    GCCGATGGTT TCTACAAAGA TCGTTATGTT TATCGGCACT TTGCATCGGC CGCGCTCCCG
2701    ATTCCGGAAG TGCTTGACAT TGGGGAATTC AGCGAGAGCC TGACCTATTG CATCTCCCGC
2761    CGTGCACAGG GTGTCACGTT GCAAGACCTG CCTGAAACCG AACTGCCCGC TGTTCTGCAG
2821    CCGGTCGCGG AGGCCATGGA TGCGATCGCT GCGGCCGATC TTAGCCAGAC GAGCGGGTTC
2881    GGCCCATTCG GACCGCAAGG AATCGGTCAA TACACTACAT GGCGTGATTT CATATGCGCG
2941    ATTGCTGATC CCCATGTGTA TCACTGGCAA ACTGTGATGG ACGACACCGT CAGTGCGTCC
3001    GTCGCGCAGG CTCTCGATGA GCTGATGCTT TGGGCCGAGG ACTGCCCCGA AGTCCGGCAC
3061    CTCGTGCACG CGGATTTCGG CTCCAACAAT GTCCTGACGG ACAATGGCCG CATAACAGCG
3121    GTCATTGACT GGAGCGAGGC GATGTTCGGG GATTCCCAAT ACGAGGTCGC CAACATCTTC
3181    TTCTGGAGGC CGTGGTTGGC TTGTATGGAG CAGCAGACGC GCTACTTCGA GCGGAGGCAT
3241    CCGGAGCTTG CAGGATCGCC GCGGCTCCGG GCGTATATGC TCCGCATTGG TCTTGACCAA
3301    CTCTATCAGA GCTTGGTTGA CGGCAATTTC GATGATGCAG CTTGGGCGCA GGGTCGATGC
3361    GACGCAATCG TCCGATCCGG AGCCGGGACT GTCGGGCGTA CACAAATCGC CCGCAGAAGC
3421    GCGGCCGTCT GGACCGATGG CTGTGTAGAA GTACTCGCCG ATAGTGGAAA CCGACGCCCC
3481    AGCACTCGTC CGAGGGCAAA GGAATAGCAC GTACTACGAG ATTTCGATTC CACCGCCGCC
3541    TTCTATGAAA GGTTGGGCTT CGGAATCGTT TTCCGGGACG CCGGCTGGAT GATCCTCCAG
3601    CGCGGGGATC TCATGCTGGA GTTCTTCGCC CACCCCAACT TGTTTATTGC AGCTTATAAT
3661    GGTTACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTTTT TTCACTGCAT
3721    TCTAGTTGTG GTTTGTCCAA ACTCATCAAT GTATCTTATC ATGTCTGTAT ACCGTCGACC
3781    TCTAGCTAGA GCTTGGCGTA ATCATGGTCA TAGCTGTTTC CTGTGTGAAA TTGTTATCCG
3841    CTCACAATTC CACACAACAT ACGAGCCGGA AGCATAAAGT GTAAAGCCTG GGGTGCCTAA
3901    TGAGTGAGCT AACTCACATT AATTGCGTTG CGCTCACTGC CCGCTTTCCA GTCGGGAAAC
3961    CTGTCGTGCC AGCTGCATTA ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT
4021    GGGCGCTCTT CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA
4081    GCGGTATCAG CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA
4141    GGAAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
4201    CTGGCGTTTT TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT
4261    CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC
```

```
4321    CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
4381    TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
4441    GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA
4501    TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
4561    GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
4621    TGGTGGCCTA ACTACGGCTA CACTAGAAGA ACAGTATTTG GTATCTGCGC TCTGCTGAAG
4681    CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT
4741    AGCGGTGGTT TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA
4801    GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG
4861    ATTTTGGTCA TGAGATTATC AAAAAGGATC TTCACCTAGA TCCTTTTAAA TTAAAAATGA
4921    AGTTTTAAAT CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA CCAATGCTTA
4981    ATCAGTGAGG CACCTATCTC AGCGATCTGT CTATTTCGTT CATCCATAGT TGCCTGACTC
5041    CCCGTCGTGT AGATAACTAC GATACGGGAG GGCTTACCAT CTGGCCCCAG TGCTGCAATG
5101    ATACCGCGAG ACCCACGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA GCCAGCCGGA
5161    AGGGCCGAGC GCAGAAGTGG TCCTGCAACT TTATCCGCCT CCATCCAGTC TATTAATTGT
5221    TGCCGGGAAG CTAGAGTAAG TAGTTCGCCA GTTAATAGTT TGCGCAACGT TGTTGCCATT
5281    GCTACAGGCA TCGTGGTGTC ACGCTCGTCG TTTGGTATGG CTTCATTCAG CTCCGGTTCC
5341    CAACGATCAA GGCGAGTTAC ATGATCCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC
5401    GGTCCTCCGA TCGTTGTCAG AAGTAAGTTG GCCGCAGTGT TATCACTCAT GGTTATGGCA
5461    GCACTGCATA ATTCTCTTAC TGTCATGCCA TCCGTAAGAT GCTTTTCTGT GACTGGTGAG
5521    TACTCAACCA AGTCATTCTG AGAATAGTGT ATGCGGCGAC CGAGTTGCTC TTGCCCGGCG
5581    TCAATACGGG ATAATACCGC GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA
5641    CGTTCTTCGG GGCGAAAACT CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA
5701    CCCACTCGTG CACCCAACTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT TTCTGGGTGA
5761    GCAAAAACAG GAAGGCAAAA TGCCGCAAAA AAGGGAATAA GGGCGACACG GAAATGTTGA
5821    ATACTCATAC TCTTCCTTTT TCAATATTAT TGAAGCATTT ATCAGGGTTA TTGTCTCATG
5881    AGCGGATACA TATTTGAATG TATTTAGAAA AATAAACAAA TAGGGGTTCC GCGCACATTT
5941    CCCCGAAAAG TGCCACCTGA CGTC
```

8.1.10.8. Vector sequence of pTS1251

```
1       GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
61      CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
121     CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
181     TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTCCATAGA
241     GCCCACCGCA TCCCCAGCAT GCCTGCTATT GTCTTCCCAA TCCTCCCCCT TGCTGTCCTG
301     CCCCACCCCA CCCCCCAGAA TAGAATGACA CCTACTCAGA CAATGCGATG CAATTTCCTC
361     ATTTTATTAG GAAAGGACAG TGGGAGTGGC ACCTTCCAGG GTCAAGGAAG GCACGGGGGA
421     GGGGCAAACA ACAGATGGCT GGCAACTAGA AGGCACAGTC GAGGCTGATC AGCGGGTTTA
481     AACATCGATT TAAAGGAGAG GTGGCTCCTG GCTGGCCTTG TCCAGCTCTG TCTCGCAGAC
541     CCAGCGGTAG GGCCTCTGGC AGACGTCGTC GTTCCAGCGG CCGTCGTCGG TGAAGTGGGC
601     ACAGTCCTCG CCTCCTCCGA GCCCGTGGCC GTACCAGTCG TCCGGCTGCT CCGGCCTCCA
661     GTTCTTGAAG CCCGTCTCGT AGTCCGTCCC GTCCACCCAC TTCCAGGGCC CGTTTTGGTC
721     GTGGAGGCCC ATCCAGGTGT TCACAGGGCC TATGTGGTGC TGGACAAATT TCTGCTCCTC
781     CCAGGACGTG ACCACCACCA GGTGCGCGTC CTCCAGCCGG CAGTAGTTGT CGGCGTCAGC
841     CCAGGCCTTC CCGGAGCGAG AGAACCAGTA GCAGCTGCGC TCGTGCTCCA CCCAGTTGAC
901     CGGGCAGCAG GTCCTTTCTG AGCCATTGCC CTGGAGCGCC GCCATCTGAC AGCTCAGGCT
961     CCGCAGGTCA GACACGAACT GCTTCACGTG GAGCAGCAGG CTGGAGTGAT CTTCACTCAG
1021    GTCCTTCTGC TGTTTCTCCA GCTGGGACTC TAGCGACTTC ATCTTTCTTC CCACATTGCC
1081    TCCCTGGGTG CTCAAGCCCT TGACCTGGGC CTCCGTGCTC GCTGTGAAGT TGCTGAACGT
1141    CTCTCTCAGG CCCCGCAGCT CCTCCTGCAG CTGGGAGTTT TGGGATCCGA TCACACAGAC
1201    AACCACAAGC AGCAGGAGGC TGAGGCCCAG GGAGAGCAGG AGGAGGCGAG GTCCGGAGCA
1261    GAGACGCTGC AGGAGGGGCT GGGGAGGAGG TGGCCCTTTT CTGAGCTGAT GGTGGTCACT
1321    CTCCTCATTG TCCAGATGCT GAAGGTCTTG ATACTCCTTG GTCATGGTGG CCCTAGGCCC
1381    CAGAGTAAAG CTATTCGGTA ATTCGTCACC CAAGAGATCA ATCGGTCTCT CTCTATCACT
1441    GATAGGGAGA TCTCTATCAC TGATAGGGAG AGCTCTGCTT ATATAGACCT CCCACCGTAC
1501    ACGCCTACCG CCCATTTGCG TCAATGGGGC GGAGTTGTTA CGACATTTTG GAAAGTCCCG
1561    TTGATTTTGG TGCCAAAACA AACTCCCATT GACGTCAATG GGGTGGAGAC TTGGAAATCC
1621    CCGTGAGTCA AACCGCTATC CACGCCCATT GATGTACTGC CAAAACCGCA TCACCATGGA
1681    CGTGTCGAGG TGATAATTCC ACTCGAGTGG CTCCGGTGCC CGTCAGTGGG CAGAGCGCAC
```

```
1741    ATCGCCCACA GTCCCCGAGA AGTTGGGGGG AGGGGTCGGC AATTGAACCG GTGCCTAGAG
1801    AAGGTGGCGC GGGGTAAACT GGGAAAGTGA TGTCGTGTAC TGGCTCCGCC TTTTTCCCGA
1861    GGGTGGGGGA GAACCGTATA TAAGTGCAGT AGTCGCCGTG AACGTTCTTT TTCGCAACGG
1921    GTTTGCCGCC AGAACACAGG TCCCTATCAG TGATAGAGAT CTCCCTATCA GTGATAGAGA
1981    TCGTCGACGA GCTCGTTTAG TGAACCGTCA GATCGCCTGG AGACGCCATC GGGCGGCCGC
2041    CACCATGGAC AAAGACTGCG AAATGAAGCG CACCACCCTG GATAGCCCTC TGGGCAAGCT
2101    GGAACTGTCT GGGTGCGAAC AGGGCCTGCA CCGTATCATC TTCCTGGGCA AAGGAACATC
2161    TGCCGCCGAC GCCGTGGAAG TGCCTGCCCC AGCCGCCGTG CTGGGCGGAC CAGAGCCACT
2221    GATGCAGGCC ACCGCCTGGC TCAACGCCTA CTTTCACCAG CCTGAGGCCA TCGAGGAGTT
2281    CCCTGTGCCA GCCCTGCACC ACCCAGTGTT CCAGCAGGAG AGCTTTACCC GCCAGGTGCT
2341    GTGGAAACTG CTGAAAGTGG TGAAGTTCGG AGAGGTCATC AGCTACAGCC ACCTGGCCGC
2401    CCTGGCCGGC AATCCCGCCG CCACCGCCGC CGTGAAAACC GCCCTGAGCG GAAATCCCGT
2461    GCCCATTCTG ATCCCCTGCC ACCGGGTGGT GCAGGGCGAC CTGGACGTGG GGGGCTACGA
2521    GGGCGGGCTC GCCGTGAAAG AGTGGCTGCT GGCCCACGAG GGCCACAGAC TGGGCAAGCC
2581    TGGGCTGGGT CCTGCAGGCG GAGGCGCGCC AGGGTCTGGC GGCGGCAGTA AGGCAGAACG
2641    CATGGGTTTC ACAGAGGTAA CCCCAGTGAC AGGGGCCAGT CTCAGAAGAA CTATGCTCCT
2701    CCTCTCAAGG TCCCCAGAAG CACAGCCAAA GACACTCCCT CTCACTGGCA GCACCTTCCA
2761    TGACCAGATA GCCATGCTGA GCCACCGGTG CTTCAACACT CTGACTAACA GCTTCCAGCC
2821    CTCCTTGCTC GGCCGCAAGA TTCTGGCCGC CATCATTATG AAAAAAGACT CTGAGGACAT
2881    GGGTGTCGTC GTCAGCTTGG GAACAGGGAA TCGCTGTGTA AAAGGAGATT CTCTCAGCCT
2941    AAAAGGAGAA ACTGTCAATG ACTGCCATGC AGAAATAATC TCCCGGAGAG GCTTCATCAG
3001    GTTTCTCTAC AGTGAGTTAA TGAAATACAA CTCCCAGACT GCGAAGGATA GTATATTTGA
3061    ACCTGCTAAG GGAGGAGAAA AGCTCCAAAT AAAAAAGACT GTGTCATTCC ATCTGTATAT
3121    CAGCACTGCT CCGTGTGGAG ATGGCGCCCT CTTTGACAAG TCCTGCAGCG ACCGTGCTAT
3181    GGAAAGCACA GAATCCCGCC ACTACCCTGT CTTCGAGAAT CCCAAACAAG GAAAGCTCCG
3241    CACCAAGGTG GAGAACGGAC AAGGCACAAT CCCTGTGGAA TCCAGTGACA TTGTGCCTAC
3301    GTGGGATGGC ATTCGGCTCG GGGAGAGACT CCGTACCATG TCCTGTAGTG ACAAAATCCT
3361    ACGCTGGAAC GTGCTGGGCC TGCAAGGGGC ACTGTTGACC CACTTCCTGC AGCCCATTTA
3421    TCTCAAATCT GTCACATTGG GTTACCTTTT CAGCCAAGGG CATCTGACCC GTGCTATTTG
3481    CTGTCGTGTG ACAAGAGATG GGAGTGCATT TGAGGATGGA CTACGACATC CCTTTATTGT
3541    CAACCACCCC AAGGTTGGCA GAGTCAGCAT ATATGATTCC AAAAGGCAAT CCGGGAAGAC
3601    TAAGGAGACA AGCGTCAACT GGTGTCTGGC TGATGGCTAT GACCTGGAGA TCCTGGACGG
3661    TACCAGAGGC ACTGTGGATG GGCCACGGAA TGAATTGTCC CGGGTCTCCA AAAAGAACAT
3721    TTTTCTTCTA TTTAAGAAGC TCTGCTCCTT CCGTTACCGC AGGGATCTAC TGAGACTCTC
3781    CTATGGTGAG GCCAAGAAAG CTGCCCGTGA CTACGAGACG GCCAAGAACT ACTTCAAAAA
3841    AGGCCTGAAG GATATGGGCT ATGGGAACTG GATTAGCAAA CCCCAGGAGG AAAAGAACTT
3901    TTATCTCTGC CCAGTATGAT TAATTAAGTT TAAACCCGCT GATCAGCCTC GACTGTGCCT
3961    TCTAGTTGCC AGCCATCTGT TGTTTGCCCC TCCCCCGTGC CTTCCTTGAC CCTGGAAGGT
4021    GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG CATCGCATTG TCTGAGTAGG
4081    TGTCATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA AGGGGGAGGA TTGGGAAGAC
4141    AATAGCAGGC ATGCTGGGGA TGCGGTGGGC TCTATGGCTT CTGAGGCGGA AAGAACCAGC
4201    TGGGGCTCTA GGGGGTATCC CCACGCGCCC TGTAGCGGCG CATTAAGCGC GGCGGGTGTG
4261    GTGGTTACGC GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT
4321    TTCTTCCCTT CCTTTCTCGC CACGTTCGCC GGCTTTCCCC GTCAAGCTCT AAATCGGGGG
4381    CTCCCTTTAG GGTTCCGATT TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG
4441    GGTGATGGTT CACGTACCTA GAAGTTCCTA TTCCGAAGTT CCTATTCTCT AGAAAGTATA
4501    GGAACTTCCT TGGCCAAAAA GCCTGAACTC ACCGCGACGT CTGTCGAGAA GTTTCTGATC
4561    GAAAAGTTCG ACAGCGTCTC CGACCTGATG CAGCTCTCGG AGGGCGAAGA ATCTCGTGCT
4621    TTCAGCTTCG ATGTAGGAGG GCGTGGATAT GTCCTGCGGG TAAATAGCTG CGCCGATGGT
4681    TTCTACAAAG ATCGTTATGT TTATCGGCAC TTTGCATCGG CCGCGCTCCC GATTCCGGAA
4741    GTGCTTGACA TTGGGGAATT CAGCGAGAGC CTGACCTATT GCATCTCCCG CCGTGCACAG
4801    GGTGTCACGT TGCAAGACCT GCCTGAAACC GAACTGCCCG CTGTTCTGCA GCCGGTCGCG
4861    GAGGCCATGG ATGCGATCGC TGCGGCCGAT CTTAGCCAGA CGAGCGGGTT CGGCCCATTC
4921    GGACCGCAAG GAATCGGTCA ATACACTACA TGGCGTGATT TCATATGCGC GATTGCTGAT
4981    CCCCATGTGT ATCACTGGCA AACTGTGATG GACGACACCG TCAGTGCGTC CGTCGCGCAG
5041    GCTCTCGATG AGCTGATGCT TTGGGCCGAG GACTGCCCCG AAGTCCGGCA CCTCGTGCAC
5101    GCGGATTTCG GCTCCAACAA TGTCCTGACG GACAATGGCC GCATAACAGC GGTCATTGAC
5161    TGGAGCGAGG CGATGTTCGG GGATTCCCAA TACGAGGTCG CCAACATCTT CTTCTGGAGG
5221    CCGTGGTTGG CTTGTATGGA GCAGCAGACG CGCTACTTCG AGCGGAGGCA TCCGGAGCTT
5281    GCAGGATCGC CGCGGCTCCG GGCGTATATG CTCCGCATTG GTCTTGACCA ACTCTATCAG
```

```
5341     AGCTTGGTTG ACGGCAATTT CGATGATGCA GCTTGGGCGC AGGGTCGATG CGACGCAATC
5401     GTCCGATCCG GAGCCGGGAC TGTCGGGCGT ACACAAATCG CCCGCAGAAG CGCGGCCGTC
5461     TGGACCGATG GCTGTGTAGA AGTACTCGCC GATAGTGGAA ACCGACGCCC CAGCACTCGT
5521     CCGAGGGCAA AGGAATAGCA CGTACTACGA GATTTCGATT CCACCGCCGC CTTCTATGAA
5581     AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA GCGCGGGGAT
5641     CTCATGCTGG AGTTCTTCGC CCACCCCAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA
5701     TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
5761     GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGTA TACCGTCGAC CTCTAGCTAG
5821     AGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT
5881     CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC
5941     TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC AGTCGGGAAA CCTGTCGTGC
6001     CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT
6061     TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA
6121     GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC
6181     ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT
6241     TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG
6301     CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC
6361     TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC
6421     GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC
6481     AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC
6541     TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT
6601     AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT
6661     AACTACGGCT ACACTAGAAG AACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC
6721     TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT
6781     TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG
6841     ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC
6901     ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTAAA
6961     TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG
7021     GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG
7081     TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA
7141     GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG
7201     CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA
7261     GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC
7321     ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA
7381     AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG
7441     ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT
7501     AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGACTGGTGA GTACTCAACC
7561     AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG
7621     GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG
7681     GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT
7741     GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA
7801     GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA
7861     CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC
7921     ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA
7981     GTGCCACCTG ACGTC
```

8.1.10.9. Vector sequence of pTS1340

```
1        ATCACCTCGA GTTTACTCCC TATCAGTGAT AGAGAACGTA TGAAGAGTTT ACTCCCTATC
61       AGTGATAGAG AACGTATGCA GACTTTACTC CCTATCAGTG ATAGAGAACG TATAAGGAGT
121      TTACTCCCTA TCAGTGATAG AGAACGTATG ACCAGTTTAC TCCCTATCAG TGATAGAGAA
181      CGTATCTACA GTTTACTCCC TATCAGTGAT AGAGAACGTA TATCCAGTTT ACTCCCTATC
241      AGTGATAGAG AACGTATAAG CTTTGCTTAT GTAAACCAGG GCGCCTATAA AAGAGTGCTG
301      ATTTTTTGAG TAAACTTCAA TTCCACAACA CTTTTGTCTT ATACCAACTT TCCGTACCAC
361      TTCCTACCCT CGTAAAGGTA CCGCGGCCGC CACCATGACC AAGGAGTATC AAGACCTTCA
421      GCATCTGGAC AATGAGGAGA GTGACCACCA TCAGCTCAGA AAAGGGCCAC CTCCTCCCCA
481      GCCCCTCCTG CAGCGTCTCT GCTCCGGACC TCGCCTCCTC CTGCTCTCCC TGGGCCTCAG
541      CCTCCTGCTG CTTGTGGTTG TCTGTGTGAT CGGATCCCAA AACTCCCAGC TGCAGGAGGA
601      GCTGCGGGGC CTGAGAGAGA CGTTCAGCAA CTTCACAGCG AGCACGGAGG CCCAGGTCAA
661      GGGCTTGAGC ACCCAGGGAG GCAATGTGGG AAGAAAGATG AAGTCGCTAG AGTCCCAGCT
```

```
721     GGAGAAACAG CAGAAGGACC TGAGTGAAGA TCACTCCAGC CTGCTGCTCC ACGTGAAGCA
781     GTTCGTGTCT GACCTGCGGA GCCTGAGCTG TCAGATGGCG GCGCTCCAGG GCAATGGCTC
841     AGAAAGGACC TGCTGCCCGG TCAACTGGGT GGAGCACGAG CGCAGCTGCT ACTGGTTCTC
901     TCGCTCCGGG AAGGCCTGGG CTGACGCCGA CAACTACTGC CGGCTGGAGG ACGCGCACCT
961     GGTGGTGGTC ACGTCCTGGG AGGAGCAGAA ATTTGTCCAG CACCACATAG GCCCTGTGAA
1021    CACCTGGATG GGCCTCCACG ACCAAAACGG GCCCTGGAAG TGGGTGGACG GGACGGACTA
1081    CGAGACGGGC TTCAAGAACT GGAGGCCGGA GCAGCCGGAC GACTGGTACG GCCACGGGCT
1141    CGGAGGAGGC GAGGACTGTG CCCACTTCAC CGACGACGGC CGCTGGAACG ACGACGTCTG
1201    CCAGAGGCCC TACCGCTGGG TCTGCGAGAC AGAGCTGGAC AAGGCCAGCC AGGAGCCACC
1261    TCTCCTTTAA ATCGATTAAT TAACTAGTAG ACCACCTCCC CTGCGAGCTA AGCTGGACAG
1321    CCAATGACGG GTAAGAGAGT GACATTTTTC ACTAACCTAA GACAGGAGGG CCGTCAGAGC
1381    TACTGCCTAA TCCAAAGACG GGTAAAAGTG ATAAAAATGT ATCACTCCAA CCTAAGACAG
1441    GCGCAGCTTC CGAGGGATTT GAGATCCAGA CATGATAAGA TACATTGATG AGTTTGGACA
1501    AACCAAAACT AGAATGCAGT GAAAAAAATG CCTTATTTGT GAAATTTGTG ATGCTATTGC
1561    CTTATTTGTA ACCATTATAA GCTGCAATAA ACAAGTTTGA TATCTATAAC AAGAAAATAT
1621    ATATATAATA AGTTATCACG TAAGTAGAAC ATGAAATAAC AATATAATTA TCGTATGAGT
1681    TAAATCTTAA AAGTCACGTA AAAGATAATC ATGCGTCATT TTGACTCACG CGGTCGTTAT
1741    AGTTCAAAAT CAGTGACACT TACCGCATTG ACAAGCACGC CTCACGGGAG CTCCAAGCGG
1801    CGACTGAGAT GTCCTAAATG CACAGCGACG GATTCGCGCT ATTTAGAAAG AGAGAGCAAT
1861    ATTTCAAGAA TGCATGCGTC AATTTTACGC AGACTATCTT TCTAGGGTTA AGAATTCACT
1921    GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT
1981    TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
2041    TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCCTG ATGCGGTATT TTCTCCTTAC
2101    GCATCTGTGC GGTATTTCAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC
2161    CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG
2221    TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA
2281    GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGACGAAAG GGCCTCGTGA TACGCCTATT
2341    TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG TCAGGTGGCA CTTTTCGGGG
2401    AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT
2461    CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
2521    TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC
2581    TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
2641    TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG
2701    TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA
2761    CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT CAGAATGACT TGGTTGAGTA
2821    CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT TATGCAGTGC
2881    TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT CTGACAACGA TCGGAGGACC
2941    GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG
3001    GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA TGCCTGTAGC
3061    AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGGCA
3121    ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTGC GCTCGGCCCT
3181    TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT
3241    CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG
3301    GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
3361    TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT
3421    TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT
3481    CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC
3541    TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT
3601    ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACTGG
3661    CTTCAGCAGA GCGCAGATAC CAAATACTGT TCTTCTAGTG TAGCCGTAGT TAGGCCACCA
3721    CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
3781    TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
3841    TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC
3901    GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA
3961    AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
4021    GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG
4081    ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG
4141    CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC
4201    TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
4261    TCGCCGCAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC
```

```
4321     AATACGCAAA CCGCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG
4381     GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA
4441     TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG
4501     CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGCCAAGG TCGACTTAAC
4561     CCTAGAAAGA TAATCATATT GTGACGTACG TTAAAGATAA TCATGCGTAA AATTGACGCA
4621     TGTGTTTTAT CGGTCTGTAT ATCGAGGTTT ATTTATTAAT TTGAATAGAT ATTAAGTTTT
4681     ATTATATTTA CACTTACATA CTAATAATAA ATTCAACAAA CAATTTATTT ATGTTTATTT
4741     ATTTATTAAA AAAAAACAAA AACTCAAAAT TTCTTCTATA AAGTAACAAA ACTTTTAGCA
4801     GTGAAAAAAA TGCTTTATTT GTGAAATTTG TGATGCTATT GCTTTATTTG TAACCATTAT
4861     AAGCTGCAAT AAACAAGTTA ACAACAACAA TTGCATTCAT TTTATGTTTC AGGTTCAGGG
4921     GGAGGTGTGG GAGGTTTTTT AAAGCAAGTA AAACCTCTAC AAATGTGGTA TGGCTGATTA
4981     TGATCCTCTG GAGATCCTAG GCGCTCAGAA GAACTCGTCA AGAAGGCGAT AGAAGGCGAT
5041     GCGCTGCGAA TCGGGAGCGG CGATACCGTA AAGCACGAGG AAGCGGTCAG CCCATTCGCC
5101     GCCAAGCTCT TCAGCAATAT CACGGGTAGC CAACGCTATG TCCTGATAGC GGTCCGCCAC
5161     ACCCAGCCGG CCACAGTCGA TGAATCCAGA AAAGCGGCCA TTTTCCACCA TGATATTCGG
5221     CAAGCAGGCA TCGCCATGGG TCACGACGAG ATCCTCGCCG TCGGGCATGC GCGCCTTGAG
5281     CCTGGCGAAC AGTTCGGCTG GCGCGAGCCC CTGATGCTCT TCGTCCAGAT CATCCTGATC
5341     GACAAGACCG GCTTCCATCC GAGTACGTGC TCGCTCGATG CGATGTTTCG CTTGGTGGTC
5401     GAATGGGCAG GTAGCCGGAT CAAGCGTATG CAGCCGCCGC ATTGCATCAG CCATGATGGA
5461     TACTTTCTCG GCAGGAGCAA GGTGAGATGA CAGGAGATCC TGCCCCGGCA CTTCGCCCAA
5521     TAGCAGCCAG TCCCTTCCCG CTTCAGTGAC AACGTCGAGC ACAGCTGCGC AAGGAACGCC
5581     CGTCGTGGCC AGCCACGATA GCCGCGCTGC CTCGTCCTGC AGTTCATTCA GGGCACCGGA
5641     CAGGTCGGTC TTGACAAAAA GAACCGGGCG CCCCTGCGCT GACAGCCGGA ACACGGCGGC
5701     ATCAGAGCAG CCGATTGTCT GTTGTGCCCA GTCATAGCCG AATAGCCTCT CCACCCAAGC
5761     GGCCGGAGAA CCTGCGTGCA ATCCATCTTG TTCAATCATG GGGCCGGGGT TCTCCTCCAC
5821     GTCACCGGCC TGCTTCAGCA GGCTGAAGTT GGTGGCGCCG CTGCCCCCGG GGAGCATGTC
5881     AAGGTCAAAA TCGTCAAGAG CGTCAGCAGG CAGCATATCA AGGTCAAAGT CGTCAAGGGC
5941     ATCGGCTGGG AGCATGTCTA AGTCAAAATC GTCAAGGGCG TCGGTCGGCC CGCCGCTTTC
6001     GCACTTTAGC TGTTTCTCCA GGCCACATAT GATTAGTTCC AGGCCGAAAA GGAAGGCAGG
6061     TTCGGCTCCC TGCCGGTCGA ACAGCTCAAT TGCTTGTTTC AGAAGTGGGG GCATAGAATC
6121     GGTGGTAGGT GTCTCTCTTT CCTCTTTTGC TACTTGATGC TCCTGTTCCT CCAATACGCA
6181     GCCCAGTGTA AAGTGGCCCA CGGCGGACAG AGCGTACAGT GCGTTCTCCA GGGAGAAGCC
6241     TTGCTGACAC AGGAACGCGA GCTGATTTTC CAGGGTTTCG TACTGTTTCT CTGTTGGGCG
6301     GGTGCCGAGA TGCACTTTAG CCCCGTCGCG ATGTGAGAGG AGAGCACAGC GGTATGACTT
6361     GGCGTTGTTC CGCAGAAAGT CTTGCCATGA CTCGCCTTCC AGGGGGCAGG AGTGGGTATG
6421     ATGCCTGTCC AGCATCTCGA TTGGCAGGGC ATCGAGCAGG GCCCGCTTGT TCTTCACGTG
6481     CCAGTACAGG GTAGGCTGCT CAACTCCCAG CTTTTGAGCG AGTTTCCTTG TCGTCAGGCC
6541     TTCGATACCG ACTCCATTGA GTAATTCCAG AGCAGAGTTT ATGACTTTGC TCTTGTCCAG
6601     TCTAGACATC TTATCGTCAT CGTCTTTGTA ATCCATGGTG GCGGATCCCG CGTCACGACA
6661     CCTGTGTTCT GGCGGCAAAC CCGTTGCGAA AAAGAACGTT CACGGCGACT ACTGCACTTA
6721     TATACGGTTC TCCCCCACCC TCGGGAAAAA GGCGGAGCCA GTACACGACA TCACTTTCCC
6781     AGTTTACCCC GCGCCACCTT CTCTAGGCAC CGGTTCAATT GCCGACCCCT CCCCCCAACT
6841     TCTCGGGGAC TGTGGGCGAT GTGCGCTCTG CCCACTGACG GGCACCGGAG CCACTCGAGT
6901     GGAATT
```

8.1.11. Overview of experiments

Table 35:Overview of the performed experiments and assignments to the described sections and laboratory journal numbers.

Schemes & Figures		Com-pound	Section (Methods and Materials)	Section (Results and Discussion)	Laboratory journal No.
Scheme 1		14	6.4.15.	3.1.1.	hessox40, hessox44
Scheme 2		2	6.4.1.	3.1.1.	hessox4, hessox14
		3	6.4.2.		hessox5.4, hessox15
		4	6.4.3.		hessox16
		5	6.4.4.		hessox17
		6	6.4.5.		hessox20
		7	6.4.6.		hessox29
Scheme 3		9	6.4.7.	3.1.1.	hessox26
		10	6.4.8.		hessox28
Scheme 4		11	6.4.9., 6.4.10., 6.4.11.	3.1.1.	hessox30, hessox32, hessox43
Scheme 5		12	6.4.12.	3.1.1.	hessox41
Scheme 6		13	6.4.13., 6.4.14.	3.1.1.	Hessox39, hessox65, hessox74
		14	6.4.15.	3.1.1.	hessox40, hessox44
Scheme 7		15	6.4.16.	3.1.1.	hessox47
		16	6.4.17., 6.4.18.		hessox53
		17	6.4.19.		hessox67
Scheme 8		18-NHS	6.4.22.	3.1.2.1.	BisBG preactivation with DIC
Scheme 9		19	6.4.23.	3.1.3.	hessox54
Scheme 10/ Figure 10+11			6.6.6, 6.6.7.	3.1.3.	OH186
Scheme 11/ Figure 12		17-NHS	6.4.20.	3.1.4.	OH182
Figure 13			6.7.	3.2.1.	OH3, OH6-OH9
Figure 14			6.8.3.2.	3.2.1.	OH43
Figure 15	a		6.8.2.1.	3.2.2.1.	OH97, OH99, OH100
	b		6.8.2.1.		OH118
	c, d		6.8.3.3.		OH117

Table 35: Continued.

Schemes & Figures		Com-pound	Section (Methods and Materials)	Section (Results and Discussion)	Laboratory journal No.
Figure 16			6.8.3.4., 6.8.3.5., 6.8.4.3.	3.2.2.1.	OH161
Figure 17	a		6.8.2.2.	3.2.2.2.	OH119
	b, c		6.8.3.4.		OH120
Figure 18			6.8.3.5.	3.2.2.2.	OH148
Figure 19			6.8.2.3., 6.8.3.5., 6.8.4.3.	3.2.2.2.	OH181
Figure 20			6.8.4.4.	3.3.1.1.	OH164, OH194
Figure 21			6.8.4.5., 6.8.4.6.	3.3.1.2.	OH172.4
Figure 22			6.8.4.7., 6.8.4.8.	3.3.1.3.	OH162.2, OH179
Figure 23			6.8.4.9.	3.3.2.1.	OH140, OH143
Figure 24			6.8.4.10.	3.3.2.2.	OH167, OH173, OH177, OH191
Figure 25	a, b		6.8.4.11.	3.3.2.2.	OH183, OH187, OH189, OH190, OH193
Figure 26	c, d		6.8.4.12., 6.8.4.13.	3.3.2.3.	OH162, OH180
	e				OH162, OH188
	f				OH155, OH162
Figure S19	A549, Huh7		6.8.4.14.	8.1.1.	OH176
	U2OS, U87				OH184

8.2. Conference and retreat contributions

Talks

Oliver Heß, Intracellular Delivery – Overcoming a billion years evolutionary defense, IFIB Retreat 2019, Freudenstadt, Germany, May 09-10, 2019.

Poster presentations

Oliver Heß, Thorsten Stafforst, ASGP Receptor mediated uptake of oligonucleotides, IFIB Retreat 2018, Hechingen, Germany, August, 26-27, 2018.

Oliver Heß, Thorsten Stafforst, SNAP®-ADAR mediated RNA-Editing in HepG2 cells, 15th Annual Meeting of the Oligonucleotide Therapeutics Society, Munich, Germany, October 13-16, 2019