

**Global Analysis of the Serine, Threonine and
Tyrosine Phosphorylation Networks in the
Model Bacterium *Bacillus subtilis***

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Vaishnavi Ravikumar
aus Chennai, Tamil Nadu, Indien

Tübingen
2015

Tag der mündlichen Qualifikation:

30.07.2015

Dekan:

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Prof. Dr. Boris Macek

2. Berichterstatter:

Prof. Dr. Alfred Nordheim

Declaration

The research work described in this thesis has been carried out under the supervision of Prof. Dr. Boris Macek during the period of October 2011 - April 2015 at the Proteome Center Tübingen, Interfaculty Institute of Cell Biology, Eberhard-Karls-University. I hereby declare that all sample preparation, SILAC-based (phospho) proteomics experiments and affinity enrichments have been independently performed by me. Additionally, all molecular biology based experiments were performed by me under the guidance of Lei Shi at the laboratory of Prof. Dr. Ivan Mijakovic, AgroParisTech, INRA, France during a two month lab-exchange, with an exception to those listed below –

- a) All deletion strains used during the course of this research work were generated in the laboratory of Prof. Dr. Ivan Mijakovic.
- b) *In vivo* YkwC enzymatic assay, *in vitro* phosphorylation/dephosphorylation assay with DnaK, heat shock assay and refolding assay were performed by Lei Shi in the laboratory of Prof. Dr. Ivan Mijakovic, Chalmers University, Sweden (Figure 30, 33, 34 and 35).
- c) Circular Dichroism spectra were measured by Kaspar Feldmeier at the Max-Planck Institute of Developmental Biology, Tübingen (Figure 29 and 36).
- d) Figures 20, 21 and 22 and Supplementary Figures 1, 2 and 3 were generated by Karsten Krug, Proteome Center Tübingen.
- e) Synthetic peptides used for the affinity enrichments were synthesized in the laboratory of Dr. Hubert Kalbacher, University of Tübingen, Germany.

Vaishnavi Ravikumar

13th May 2015

Table of Contents

List of Abbreviations	1
Summary	3
Zusammenfassung	6
1. Introduction	10
1.1. Protein Phosphorylation	10
1.1.1. Phosphotransferase System	13
1.1.2. Histidine-Aspartate Signaling and Two-Component Systems.....	14
1.1.3. Serine, Threonine, Tyrosine Phosphorylation in Bacteria.....	15
1.2. <i>Bacillus subtilis</i> as a Model Organism.....	16
1.2.1. Serine, Threonine Kinases and Phosphatases in <i>B. subtilis</i>	17
1.2.2. Tyrosine Kinases and Phosphatases in <i>B. subtilis</i>	19
1.3. Mass Spectrometry Based Proteomics.....	19
1.3.1. Shotgun Proteomics Workflow	20
1.3.2. Instrumentation.....	22
1.3.3. Chromatography	27
1.3.4. Database Search.....	28
1.3.5. Protein Quantification.....	30
1.4. Phosphoproteomics.....	34
1.4.1. Localization of Phosphorylation Events	38
1.4.2. Phosphorylated Peptide Enrichment Strategies	40
1.5. Interaction Proteomics.....	44
2. Aim of Thesis	48
3. Materials and Methods	49
3.1. Materials	49
3.1.1. List of Chemicals and Consumables	49
3.1.2. List of Instruments	51
3.2. Methods.....	52
3.2.1. Polymerase Chain Reaction	52
3.2.2. Gel Electrophoresis	53
3.2.3. DNA Purification.....	54
3.2.4. Restriction Digestion	55
3.2.5. Ligation.....	56
3.2.6. Electro Competent Cell Preparation	56
3.2.7. Transformation.....	57
3.2.8. Plasmid Isolation.....	57
3.2.9. Colony PCR.....	58
3.2.10. Sequencing.....	59
3.2.11. SILAC Labeling of Bacterial Cells	60
3.2.12. Protein Extraction.....	60

3.2.13. Chloroform/Methanol Precipitation	61
3.2.14. Protein Estimation	61
3.2.15. In-solution Digestion	62
3.2.16. Protein Fractionation by Offgel Isoelectric Focusing	63
3.2.17. In Gel Digestion	63
3.2.18. Phosphorylated Peptide Enrichment.....	65
3.2.19. Mass Spectrometric Analysis.....	67
3.2.20. Data Processing.....	68
3.2.21. Bioinformatic Analysis	69
3.2.22. Synthesis and Purification of 6x-His Tagged Proteins	70
3.2.23. <i>In vitro</i> Phosphorylation and Dephosphorylation Assay	71
3.2.24. Enzymatic Assays	72
3.2.25. Circular Dichroism Spectroscopy	72
3.2.26. Strain Construction.....	73
3.2.27. Heat Shock Assays.....	74
3.2.28. Refolding Assay	74
3.2.29. Isolation of Genomic DNA.....	75
3.2.30. Affinity Chromatography	76
4. Results	79
4.1. Quality Control	79
4.1.1. Quality Control of the Knock-out Strains (PCR).....	79
4.1.2. Quality Control of SILAC Incorporation and Mixing.....	80
4.2. Proteome and Phosphoproteome Dynamics of <i>B. subtilis</i> During Growth	82
4.2.1. Experimental Design of the Growth Curve Measurements	82
4.2.2. Proteome and Phosphoproteome Dynamics During Growth in Batch Culture.....	82
4.3. Novel <i>In vivo</i> Substrates of Ser/Thr Kinase PrkC and Phosphatase PrpC	88
4.3.1. Experimental Design of PrkC and PrpC Knock-out Screens.....	88
4.3.2. SILAC-Based Screen of <i>In vivo</i> Substrates of the Kinase PrkC and Phosphatase PrpC.....	88
4.3.3. Validation of YkwC Ser281 as a Substrate of PrkC and PrpC.....	92
4.4. Novel <i>In vivo</i> Substrates of Tyr Kinase PtkA and Phosphatase PtpZ.....	96
4.4.1. Experimental Design of PtkA and PtpZ Knock-out Screens	96
4.4.2. SILAC-Based Screen of <i>In vivo</i> Substrates of the Kinase PtkA and Phosphatase PtpZ	96
4.4.3. Validation DnaK Tyr601 as a Substrate of PtkA and PtpZ	98
4.5. Analysis of YabT, YbdM, YveL, YfkJ Using the Respective Knock-out Strains	102
4.5.1. Experimental Design.....	102
4.5.2. SILAC-Based Screens	102
4.6. Analysis of the Kinase/Phosphatase Network of <i>B. subtilis</i>	107
4.7. A Comprehensive Map of the <i>B. subtilis</i> Phosphoproteome.....	110
4.8. Putative Interactors of Phosphorylated Motifs	114
4.8.1. Quality Control.....	114
4.8.2. Identifying Interactors of Phosphorylation Motifs	116

5. Discussion	124
5.1. Dynamics of the Proteome and Phosphoproteome of <i>B. subtilis</i>.....	124
5.2. Novel Serine, Threonine and Tyrosine Kinase and Phosphatase Substrates	127
5.3. Comprehensive Map of <i>B. subtilis</i> Phosphoproteome	130
5.4. Identifying Putative Interactors of Phosphorylation Motifs using Synthetic Peptides.....	133
6. Conclusions	136
7. References.....	138
8. List of Publications.....	149
9. Curriculum Vitae.....	150
10. Acknowledgements	151

List of Abbreviations

ABC	Ammonium bicarbonate
ACN	Acetonitrile
ATP	Adenosine-5'-triphosphate
BSA	Bovine serum albumin
CID	Collision-induced dissociation
DHB	2,5-dihydrobenzoic acid
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
EtBr	Ethidium bromide
FDR	False discovery rate
GP	Glycerol-6-phosphate
HCD	Higher-energy collisional dissociation
HK	Histidine kinase
HPLC	High performance liquid chromatography
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Lysogeny broth
MMA	High mass measurement accuracy
MS	Mass spectrometry
m/z	Mass-to-charge ratio
NADP	Nicotinamide adenine dinucleotide phosphate
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pI	Isoelectric point
PTM(s)	Post-translational modification(s)
PTS	Phosphotransferase system

rpm	Rotations per minute
RR	Response regulator
RT	Room temperature
SAP	Shrimp alkaline phosphatase
SCX	Strong cationic exchange
SDS	Sodium dodecyl sulfate
SF	Sodium fluoride
SILAC	Stable isotope labeling by amino acids in cell culture
SOV	Sodium orthovanadate
SPE	Solid phase extraction
TEAB	Triethylammonium bicarbonate
TFA	Trifluoroacetic acid
TiO ₂	Titanium dioxide
WT	Wild-type

Summary

Cells constantly have to respond to environmental cues in their natural environment. Many of these cellular responses are mediated by covalent modifications on proteins that are involved in signal transduction cascades. Proteins can be altered by a wide spectrum of post-translational modifications, phosphorylation being one of the most eminent and ubiquitous among them. Reversible phosphorylation on proteins is known to govern discrete functions in a cell, from growth and development to survival and death. In bacteria, as in eukaryotes, loss or an improper function of this otherwise tightly regulated post-translational modification can result in the disruption of signaling networks, causing perturbations in normal cellular and physiological processes. Initially, the classical two-component signaling was considered the only mode of regulation by protein phosphorylation in bacterial systems. Recently, a well-established modification occurring in eukaryotic systems - serine, threonine and tyrosine phosphorylation - was established as highly significant in the biology of microbial communities, influencing processes such as DNA repair, recombination, replication, exopolysaccharide synthesis, stress responses, virulence and pathogenicity. Functional importance of serine/threonine/tyrosine (STY) phosphorylation events has been extensively studied in *Bacillus subtilis*, a Gram-positive model bacterium widely used in basic research and industrial applications. *B. subtilis* possesses nine serine/threonine and tyrosine kinases and seven phosphatases that have been shown to play a crucial role in the sugar phosphotransferase system, spore germination, biofilm formation and resistance to stress factors. Despite of the increasing amount of knowledge on processes regulated by STY phosphorylation, some of the basic mechanistic aspects of this modification in bacteria are still largely unknown. For example, it is not known how many substrates STY-kinases or -phosphatases regulate, whether they are specific or promiscuous in their activity or whether the bacterial proteins contain domains or motifs that specifically recognize phosphorylated residues and can “read” and relay information encoded in them.

The overarching goal of my thesis was to gather experimental data in order to assemble the first network of kinases, phosphatases and their substrates in the model bacterium

B. subtilis. By applying quantitative phosphoproteomics workflow based on SILAC labeling and high accuracy mass spectrometry, the aim was to define potential substrates of *B. subtilis* kinases and phosphatases and understand the global impact and regulation of phosphorylation events during growth of the bacterium in minimal medium under laboratory conditions. Additionally, using a screen based on modified synthetic peptides, the objective was to identify specific interactors of phosphorylated residues and investigate the domain architectures that bind specifically to STY-phosphorylated residues. Structurally conserved elements like SH2 and PTB domains or 14-3-3 proteins that mediate intermolecular interactions have been widely studied and characterized in eukaryotic systems, but almost nothing is known about the existence of such motifs with respect to prokaryotic systems.

In the global proteome analysis during growth of *B. subtilis*, 2,264 proteins were identified covering 55.2 % of the theoretical *B. subtilis* proteome. Proteins were observed to be tightly regulated during the initial stages of growth (logarithmic phase) and fluctuations were more pronounced during later stages of growth (stationary phase). Measurement of the protein dynamics across the growth stages also revealed information about the expression status of specific STY –kinases and –phosphatases of interest, thus assisting in choosing optimum growth time points for the individual phosphoproteomics screens. On the phosphoproteome level, a total of 177 phosphorylation events were identified. Phosphorylation events were observed to be highly upregulated during the late stationary phase of growth which corroborates the recent data reported for other bacteria. The phosphoproteome dataset affirmed that enzymes of the central carbon metabolism and components of the translation machinery were highly phosphorylated in the stationary phase, coinciding with stronger expression of the STY -kinases. A systematic phosphoproteome analysis of PrkC, PtkA, YabT, YbdM and YveL kinase and PrpC, PtpZ and YfkJ phosphatase KO strains detected 416 phosphorylation events. Most of the identified events were quantified and assessed for dependence on specific kinase or phosphatase activity. For example, YkwC, an uncharacterized oxidoreductase, was found to be a direct substrate of PrkC and PrpC, whose activity was abolished upon phosphorylation. Conversely, DnaK, a major

chaperone, was found to be a direct substrate of PtkA and PtpZ, whose chaperon activity was enhanced upon phosphorylation.

Although the kinase/phosphatase screens led to one of the largest bacterial phosphoproteome datasets, the quantification data revealed that the overall number of putative substrates identified in these phosphoproteomics screens was lower than expected, pointing to a high functional kinase and phosphatase specificity. The data obtained from all phosphoproteomics screens were integrated into the first prokaryotic network of kinases, phosphatases and their substrates based solely on experimental data. In several cases, STY –kinases and –phosphatases were observed to act on the same protein substrates highlighting the fact that they act in the same pathways and are likely to be tightly regulated. In addition to identifying specific substrates of the STY –kinases and –phosphatases, the goal was also to build a comprehensive map of the *B. subtilis* phosphoproteome. Processing of all *B. subtilis* MS files gathered in our laboratory over the last four years resulted in detection of 995 non-redundant phosphorylation events, which presents the largest *B. subtilis* phosphoproteome dataset and an important resource for the bacterial phosphoproteomics community. Motif-x analysis of the STY phosphorylated peptides for potential target motifs for kinases resulted in several defined motif patterns pointing to the possible presence of conserved phosphorylation motifs in prokaryotic systems. Finally, a novel interaction proteomics screen using synthetic peptides as baits was applied for the first time in bacterial systems. Although no single protein domain appeared to be enriched in specific interactors of the phosphorylated peptides, large number of ribosomal proteins was observed to bind the PtkA peptide and the PrkC peptide, pointing to their potential involvement in processes mediated by these kinases, such as biofilm formation, sporulation and germination. Overall, the work presented in this thesis is an extensive quantitative analysis of the *B. subtilis* proteome and phosphoproteome that will be of importance for the bacterial phosphoproteomics community. Through identification of potential kinase and phosphatase substrates, the aim was to move one step closer to understanding the significance of this global post-translational modification in a prokaryotic system.

Zusammenfassung

Zellen reagieren ständig auf Umweltreize in ihrer natürlichen Umgebung. Viele dieser zellulären Antworten werden durch kovalente Proteinmodifikationen, die in Signalübertragungskaskaden beteiligt sind, vermittelt. Proteine können durch ein breites Spektrum von posttranslationalen Modifikationen verändert werden. Darunter ist die Phosphorylierung eine der bedeutendsten und allgegenwärtigsten. Die reversible Phosphorylierung von Proteinen steuert diskrete Funktionen einer Zelle, vom Wachstum und der Entwicklung über das Überleben bis zum Tod. In Bakterien, wie auch in Eukaryoten, kann der Verlust oder eine Fehlfunktion dieser ansonsten streng regulierten posttranslationalen Modifikation zu einer Unterbrechung von Signalnetzwerken führen, welches Störungen in normalen zellulären und physiologischen Prozessen auslöst. Ursprünglich wurde das Zweikomponentensystem als die einzige Art der Regulierung von Proteinphosphorylierungen in bakteriellen Systemen erachtet. In den letzten Jahren wurde eine Modifikation, welche gutetabliert in eukaryotischen Systemen ist, als hochsignifikant in der Biologie von mikrobiellen Gemeinschaften eingeführt: die Phosphorylierung von Serin, Threonin und Tyrosin, welche Prozesse wie DNA-Reparatur, Rekombination, Replikation, Exopolysaccharidsynthese, Stressreaktionen, Virulenz und Pathogenität beeinflusst. Die funktionelle Bedeutung von Serin/Threonin/Tyrosin (STY) Phosphorylierungen wurde ausgiebig in *Bacillus subtilis*, einem grampositiven Modelbakterium, welches in der Grundlagenforschung und Industrie Anwendung findet, untersucht. *B. subtilis* besitzt neun Serin-/Threonin- und Tyrosin-Kinasen und sieben Phosphatasen, welchen eine entscheidende Rolle im Phosphotransferase-System sowie bei der Sporenkeimung, Biofilmbildung und der Widerstandsfähigkeit gegen Stressfaktoren spielen. Trotz des zunehmenden Wissens über Prozesse, die durch STY-Phosphorylierung gesteuert werden, sind einige der grundlegenden mechanistischen Aspekte dieser Modifikation in Bakterien noch weitgehend unbekannt. Beispielsweise ist nicht bekannt, wie viele Substrate die Kinasen oder Phosphatasen regulieren, ob sie spezifisch oder unspezifisch in ihrer Aktivität sind oder ob bakterielle Proteine Domänen und Motive enthalten, die spezifisch

phosphorylierte Reste erkennen und „lesen“ können und darin kodierte Informationen weitergeben können.

Das übergreifende Ziel meiner Doktorarbeit war es, experimentelle Daten zu erfassen, um ein erstes Netzwerk von Kinasen, Phosphatasen und ihrer Substrate im Modellobjekt *B. subtilis* zu erstellen. Mit dem Ziel potentielle Substrate von *B. subtilis*-Kinasen und -Phosphatasen zu definieren sowie die umfassende Auswirkung und Regulierung von Phosphorylierungsereignissen während des Wachstums des Bakteriums in Minimalmedium unter Laborbedingungen zu verstehen, habe ich quantitative Phosphoproteomik basierend auf SILAC-Markierung und hochauflösender Massenspektrometrie angewendet. Ein weiteres Ziel war die Identifizierung bestimmter Interaktoren von phosphorylierten Aminosäureresten, sowie die Untersuchung der Domänenstruktur, welche spezifisch STY Aminosäurereste bindet. Die dafür durchgeführten Experimente basierten auf synthetisierten phosphorylierten Peptiden. Während in eukaryotischen Systemen strukturell konservierte, intermolekular wechselwirkende Elemente, wie SH2 und PTB Domänen oder 14-3-3 Proteine, umfassend untersucht und charakterisiert wurden, ist über deren Existenz in prokaryotischen Systemen sehr wenig bekannt. In einer umfassenden Proteomanalyse wurden 2.264 Proteine während des Wachstums identifiziert, was einer Abdeckung von 55,2% des theoretischen *B. subtilis* Proteoms von 55,2% darstellt. Proteine während der Anfangsphase des Wachstums (logarithmische Phase) waren streng reguliert, während Schwankungen in den Expressionsprofilen in den späteren Phasen des Wachstums (stationäre Phase) stärker ausgeprägt waren. Die Bestimmung der Expressionsdynamik der identifizierten Proteine über die Wachstumsphasen hinweg offenbarte auch die Expressionsprofile der spezifischen STY-Kinasen und -Phosphatasen, welche Gegenstand dieser Arbeit sind. Dadurch konnten die optimalen Wachstumszeitpunkte für individuelle Phosphoproteomuntersuchungen gewählt werden. Auf der Phosphoproteomebene wurden insgesamt 177 Phosphorylierungen identifiziert, welche während der späten stationären Wachstumsphase stark hochreguliert waren, was kürzlich publizierte Ergebnisse in anderen Bakterienstämmen bekräftigt. Der Phosphoproteomdatensatz bestätigte, dass Enzyme des zentralen Kohlenstoffwechsels und Komponenten der Translationsmaschinerie in der stationären Phase,

übereinstimmend mit einer stärkeren Expression der STY-Kinasen, stark phosphoryliert waren. Eine systematische Analyse des Phosphoproteoms von PrkC-, PtkA-, YabT-, YbdM- und YveL-Kinase und PrpC-, PtpZ- und YfkJ-Phosphatase KO-Stämmen führte zur Identifizierung von 416 Phosphorylierungen. Die Mehrheit der identifizierten Phosphorylierungsstellen wurde auch quantifiziert und diente zur Untersuchung der Abhängigkeit der Aktivität von spezifischen Kinasen und Phosphatasen. Zum Beispiel wurde YkmC, eine nicht charakterisierte Oxidoreduktase, als direktes Substrat von PrkC und PrpC erkannt, deren Aktivität nach Phosphorylierung aufgehoben wurde. Umgekehrt wurde DnaK, ein wichtiges Chaperon, als direktes Substrat von PtkA und PtpZ identifiziert, dessen Chaperon-Aktivität nach der Phosphorylierung verstärkt wurde. Obwohl die Kinase/Phosphatase-Screens zu einem der größten bakteriellen Phosphoproteomdatensätze führten, offenbarten die quantitativen Daten eine geringere Anzahl von potentiellen Substraten als erwartet, was auf eine hohe Spezifität der Kinasen/Phosphatasen hindeutet. Die aus allen Phosphoproteomuntersuchungen erhaltenen, ausschließlich experimentellen Daten, wurden zu einem ersten prokaryotischen Netzwerk von Kinasen, Phosphatasen und deren Substrate zusammengefügt. In mehreren Fällen wurde beobachtet, wie STY-Kinasen und -Phosphatasen an gleichen Proteinsubstraten agierten, was hervorhebt, das diese in den gleichen Signalwegen aktiv sind und voraussichtlich streng reguliert werden. Neben der Identifizierung spezifischer Substrate der STY-Kinasen und -Phosphatasen war ein weiteres Ziel, eine umfassende Karte des *B. subtilis* Phosphoproteoms aufzubauen. Die Analyse aller *B. subtilis* MS-Daten, welche in den letzten vier Jahren in unserem Labor erhoben wurden, führte zur Identifikation von 995 Phosphorylierungen, was den größten *B. subtilis* Phosphoproteomdatensatz repräsentiert und als wichtige Ressource für die bakterielle Phosphoproteom- Forschungsgemeinschaft darstellt. Eine Analyse des Sequenzmotifs von STY-phosphorylierten Peptiden auf potentielle Ziel motive für Kinasen führte zu mehreren bestimmten Motifprofilen, welche auf eine mögliche Existenz von konservierten Phosphorylierungsmotifen in prokaryotischen Systemen hindeuten. Abschließend wurde eine neue Methode zur Untersuchung von Proteininteraktionen basierend auf der Verwendung synthetischer Peptide erstmalig in einem bakteriellen System angewendet. Obwohl keine Überrepräsentierung einer

einzelnen Proteindomäne in den spezifischen Interaktionspartnern der phosphorylierten Peptide gefunden wurde, konnten eine große Anzahl von ribosomalen Proteinen, die PtkA-Peptide und PrkC-Peptide binden, beobachtet werden. Dies deutet auf eine mögliche Beteiligung der Kinasen in verschiedenen Prozessen, wie zum Beispiel Biofilmbildung, Sporulation und Germination hin.

Zusammenfassend stellt diese Arbeit eine umfangreiche, quantitative Analyse des *B. subtilis* Proteoms und Phosphoproteoms dar, welche von Bedeutung für die bakterielle Phosphoproteom-Forschungsgemeinschaft sein wird. Basierend auf der Identifizierung potenzieller Substrate von Kinasen- und Phosphatasen, war das Ziel dieser Arbeit, einen weiteren Schritt hin zum Verständnis der Bedeutung dieser globalen posttranslationalen Modifikation in einem prokaryotischen System zu erreichen.

1. Introduction

The classical “central dogma of molecular biology” postulates that genetic information within a system is transferred from DNA to mRNA via transcription and is then subsequently translated into proteins. However, this concept needs to be extended to accommodate post-translational modifications that tightly regulate protein structure, localization, interaction and function. Post-translational modifications (PTMs) refer to the enzymatic addition of a chemical moiety or chemical tag to a nascent polypeptide chain or a protein molecule leading to a controlled regulation of cellular signaling events. There are around 300 PTMs detected to date, and most of them are well studied or characterized [1]. PTMs play diverse roles in the cellular milieu. Protein phosphorylation has been implicated to play a key role in cell metabolism and signaling related to cellular growth, division, differentiation and movement, to name a few examples [2]. Protein lysine acetylation is known to play a role in DNA-protein interaction, transcriptional activity, localization, stability, and other processes [3]. Likewise, glycosylation plays a crucial role in protein folding, intracellular trafficking and secretion [4]. Other PTMs such as methylation, ubiquitination, pupylation, lipidation, nitrosylation or thiolation also play significant roles and contribute towards the increase of the functional diversity of the proteome. Thus identifying and characterizing these PTMs that influence diverse cellular pathways and processes is critical in understanding signal transduction networks and help develop therapeutic approaches to treat diseases. The focus of my thesis is on protein phosphorylation networks in prokaryotic systems.

1.1. Protein Phosphorylation

Reversible protein phosphorylation is ubiquitous in nature and is well known to be an integral part of all systems from the three kingdoms of life, namely Prokarya, Archaea and Eukarya. It is one of the most widely spread types of post-translational modification where specific amino acid residues are modified (phosphorylated) by a protein kinase by the addition of a covalently bound phosphate group. All kinases require a divalent metal ion such as Mg^{2+} or Mn^{2+} as a cofactor for their activity, which stabilizes the high-energy bonds of the donor molecule, mostly ATP or GTP and their derivatives, and allows

phosphorylation to occur. The mechanism of phosphorylation (**Figure 1**) involves binding of the ATP and substrate to the kinase, followed by the γ -phosphate transfer from ATP to a specific residue in the substrate and detachment of the substrate after phosphate transfer [2].

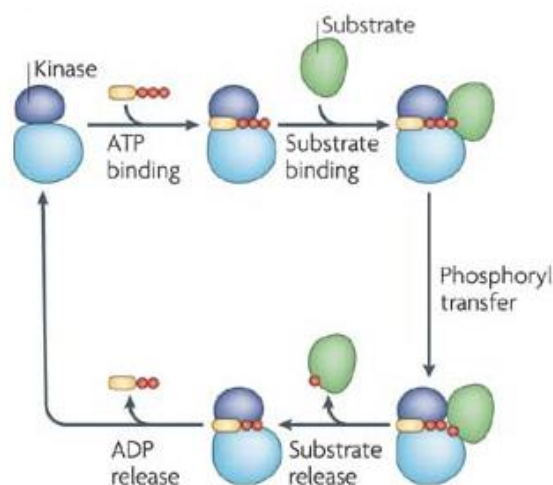


Figure 1: Mechanism of phosphate transfer – The first step in this phospho-transfer mechanism involves binding of the ATP moiety to the active site of the kinase, followed by substrate binding to the same site, transfer of the γ -phosphate to a specific residue on the substrate and finally release of the substrate along with an ADP moiety [2].

The protein kinases, enzymes that catalyze phosphorylation reactions, fall into several classes or groups based on the amino acid residue that they modify (serine/threonine, tyrosine and histidine kinases), as well as based on their sequence homology (tyrosine kinases, CdKs, Ca-dependent, receptor serine kinases and others). The most common types of protein phosphorylation events occur on the side chains of serine, threonine, tyrosine, arginine, histidine, and aspartate residues. Opposite to kinase activity is the action of phosphatases, making phosphorylation a reversible modification. Phosphatases catalyze the removal of the phosphate group from the amino acid by breaking the phospho-ester bond, resulting in formation of a water molecule and release of free inorganic phosphate. Together, the kinases and phosphatases have a steadfast control of cellular signaling events occurring in the cell. In general, in response to external stimuli, kinases undergo autophosphorylation which is subsequently followed by the transfer of the phosphate group to a specific substrate leading to downstream signaling events. The

activity of phosphatases further coordinate these signaling events through a negative or positive feedback loop.

The existence of protein phosphorylation as a regulatory mechanism was first recognized in 1955, where the authors reported the conversion and subsequent activation of glycogen phosphorylase *b* to glycogen phosphorylase *a* in muscle extracts, dependent on the action of a protein kinase together with ATP [5]. Signaling based on serine/threonine/tyrosine (STY) phosphorylation formed the main paradigm and was central to eukaryotic signaling events. However, with the advent of genome sequencing, many genes encoding serine/threonine kinases were discovered in bacterial systems, opening the possibility of a eukaryotic-like phosphoregulation in prokaryotes. With the characterization of the first eukaryotic-like serine/threonine protein kinase, Pkn1, from *Myxococcus xanthus* [6], bacterial protein phosphorylation on serine/threonine residues emerged as a possible major regulatory modification. Similarly, purification and characterization of the first bacterial tyrosine kinase, Ptk from *Acinetobacter johnsonii* [7], led the way to the acceptance of bacterial tyrosine phosphorylation. Since its discovery, through the years, bacterial protein phosphorylation was relatively neglected and pursued only in a few laboratories, largely due to the much more abundant eukaryotic protein phosphorylation. However, in the past decade considerable attention has been devoted to prokaryotic protein STY phosphorylation, leading to deeper understanding of the adaptive and interactive mechanisms mediated by this modification.

The presence of different pathways of regulation via protein phosphorylation (which include the phosphotransferase system, the two-component systems and signaling based on serine/threonine and tyrosine phosphorylation) is indicative of a functional overlap between them. The most prominent example of cross-talk between different phosphorylation systems is the example of catabolite repression. The phosphotransferase system protein HPr is phosphorylated on Ser46 by HPr kinase/phosphorylase. Phosphorylated HPr interacts with the transcriptional regulator CcpA, subsequently binding to the *cre* site leading to carbon catabolite repression [8]. HPr was also recently shown to be phosphorylated on Ser12 by the Hanks-type serine/threonine kinase PrkC [9]. Another striking example is the collective control and

regulation of the *M. xanthus* transcription factor Mrp by the MrpA/MrpB two-component system as well as the Hanks-type serine/threonine kinase Pkn8 and Pkn14 [10]. These examples demonstrate a high degree of connectivity amongst different classes of kinases and phosphorylation systems, functioning co-operatively to govern distinctive aspects of a cell.

1.1.1. Phosphotransferase System

The phosphotransferase system (PTS) is a major mechanism used by bacteria for transport and phosphorylation of various sugars and their derivatives. It is an active transport system that is primarily involved in translocation of sugars (glucose, mannose, fructose, cellobiose) across the cell membrane (**Figure 2**). A phosphate group from a phosphoenolpyruvate molecule, PEP (the energy source) is passed via a cascade of reversible protein phosphorylation reactions, to a sugar moiety. The system was discovered in 1964 by Saul Roseman [11]. The process of PTS transport involves transfer of the phosphoryl group from PEP to a histidine residue on Enzyme EI. Enzyme EI in turn transfers the phosphate to the histidine containing phosphocarrier protein, HPr. From HPr the phosphate moiety is transferred to the histidine on Enzyme EII. Individual Enzyme EII's are specific for a particular sugar molecule (based on the type of sugar to be transported). Subsequently, the phosphate group is transferred to the conserved cysteine of Enzyme EIIB. Finally, Enzyme EIIB phosphorylates the sugar as it crosses the plasma membrane through the transmembrane Enzyme EIIC, forming a sugar-phosphate [12]. The transport of the phosphate onto the translocated sugar prevents it from being recognized by the transporter again thus maintaining a one-way concentration gradient. Phosphorylated forms of some of the multi-components of the PTS are seen to have other regulatory roles in the cell [13]. For example, with the glucose phosphotransferase system, the phosphorylation status of EIIA can have regulatory functions. At low glucose concentrations phosphorylated EIIA gets accumulated and activates the membrane-bound adenylate cyclase. Intracellular cyclic AMP levels rise followed by activation of the catabolite activator protein (involved in the catabolite repression). When the glucose concentration increases, EIIA is dephosphorylated and subsequently inhibits the adenylate cyclase, glycerol kinase, lactose permease, and maltose permease. As such,

along with translocation of substrates into the bacterium, this transport system also regulates other associated proteins [14].

1.1.2. Histidine-Aspartate Signaling and Two-Component Systems

With respect to regulation via phosphorylation in prokaryotic systems, the classical two-component signaling was considered to be the norm up until little more than a decade ago. Two-component systems form the basic response of bacteria to adapt to changing environmental conditions or stress factors. The term two-component system was first coined in 1986 [15, 16]. Two-component systems consist of a histidine sensor kinase (HK) and a response regulator (RR) (**Figure 2**). In response to external stimuli, the sensor kinase undergoes ATP-dependent autophosphorylation on a histidine residue, gets activated and subsequently transfers the phosphate group to the aspartate residue present on the response regulator. Activation of the response regulator via phosphorylation leads to control of gene expression [17]. The phosphorelay is triggered by environmental cues such as nutrient availability, pH stress, salinity, temperature conditions or competition due to the presence of other bacteria. Autophosphorylation of the histidine residue on the sensor kinase is a bimolecular reaction between two homodimers, in which one HK monomer catalyzes the phosphorylation of the conserved His residue onto the second HK monomer [18]. Most often phosphorylation of the response regulator results in transcriptional regulation. However, the activated response regulator also regulates toxin production, cell adhesion, quorum sensing, capsule synthesis, motility, and drug resistance. Many response regulators have autophosphatase activity, with dephosphorylation tuned to act as a negative feed-back loop when required [16]. Environmental stimuli are detected directly by the N-terminal sensing domain of the transmembrane HK. The sensing domain is connected to the cytoplasmic kinase core through a transmembrane helix and a cytoplasmic linker. The catalytic kinase core is composed of a dimerization domain and an ATP/ADP-binding phospho-transfer or catalytic domain [17]. Some of the best studied two-component systems amongst others include bacterial chemotaxis [19], aerobic/anaerobic regulation in *Escherichia coli* [20] and sporulation in *Bacillus subtilis* [21]. High-throughput proteomic analysis do not detect phosphorylation events on histidine-aspartate residues and are solely restricted

to serine, threonine, tyrosine or arginine phosphorylation for the simple reason that the phosphoramidate bond is extremely labile under the acidic conditions generally employed in enrichment of phosphorylated peptides. Nonetheless, in recent years there has been an increasing impetus to develop methods for the analysis of histidine phosphorylation [22]. Though rare, two-components have also been identified in eukaryotic systems [23, 24].

1.1.3. Serine, Threonine, Tyrosine Phosphorylation in Bacteria

Since its discovery in the early 90's, bacterial serine, threonine and tyrosine (STY) phosphorylation was considered synonymous to eukaryotic organisms [25]. cAMP-dependent protein kinase activity in prokaryotic systems was initially reported in 1969 [26], where the authors observed that an enzyme in *E. coli* was responsible for the phosphorylation of histones using ATP. Subsequently, following this, protein kinase activity was also observed in *Streptococcus* and by others in *E. coli* [27, 28]. However there was no concrete proof of the existence of protein kinases, as exogenous proteins were used as substrates in the above mentioned protein kinase assays. As a result, early reports of protein phosphorylation in prokaryotes were found questionable. It took another 10 years for the first protein phosphorylation event to be discovered on a Ser/Thr residue in bacterial systems. Employing pulse-chase along with ³²P labeled orthophosphate, the authors identified Ser/Thr phosphorylated peptides and provided support for the presence of protein kinases in *Salmonella typhimurium* [29]. Around the same time, using a similar strategy, isocitrate dehydrogenase from *E. coli* was shown to be phosphorylated, the first example of an endogenous substrate of a protein kinase in prokaryotes [30].

The existence and significance of STY phosphorylation in prokaryotes has now been realized and ventured into. Like in eukaryotic systems, phosphorylation of a substrate by its cognate kinase directly modulates the function of the substrate or influences downstream signaling (**Figure 2**). STY phosphorylation is seen to have numerous implications in the prokaryotic system. They are found to regulate important pathways and metabolic cycles. In *Streptomyces griseus* it plays an important role in repair, recombination and replication by phosphorylating SSBs [31]. In *B. subtilis* it is involved

in exopolysaccharide synthesis [32]. In *E. coli* it has been reported to play a role in regulation of heat shock response [33]. In *Caulobacter crescentus* it regulates cell morphology and separation [34]. In pathogenic species like *Salmonella*, *Listeria*, *Shigella* it plays a crucial role in virulence [35]. These are a few examples where serine, threonine and tyrosine phosphorylation play an important role. They control various factors and influence growth and development and even survival in different bacterial systems. Thus it is essential to find interactors and or substrates of these kinases and phosphatases to understand their role in the physiology of the organism.

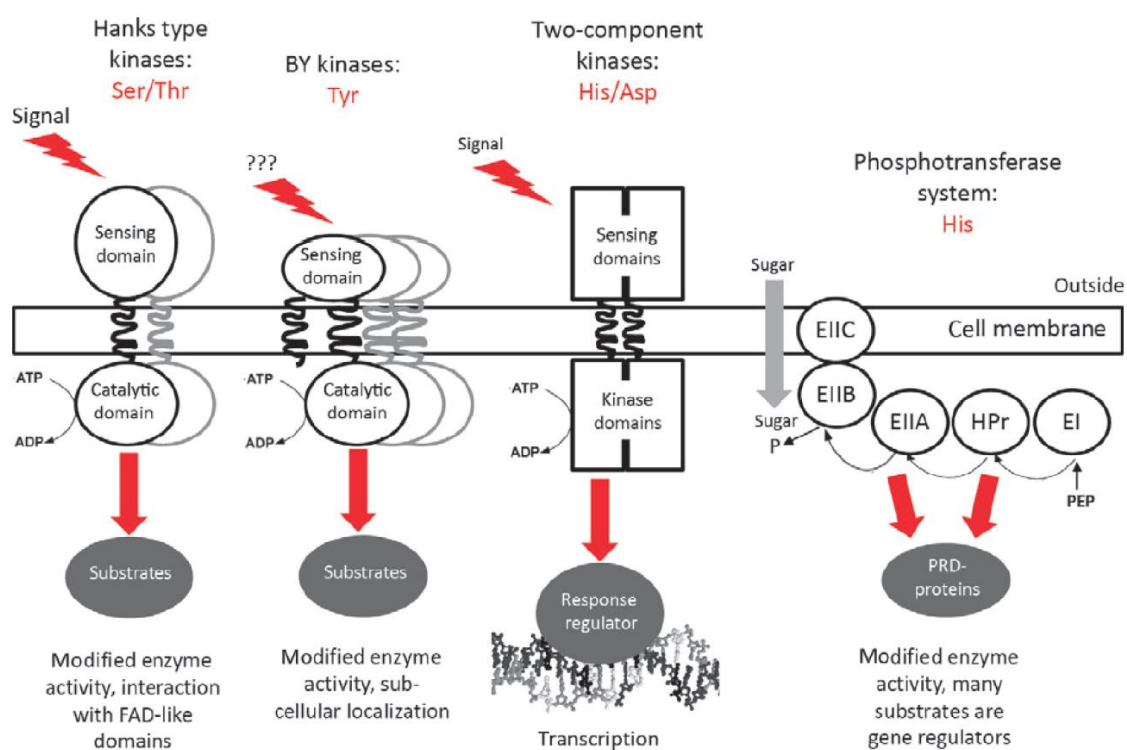


Figure 2: Schematic representation of the major families of bacterial kinases – The four major phosphorylation based regulation of signaling cascades are depicted, namely the Hanks-type kinases, BY-kinase, two-component systems and the PTS system [36].

1.2. *Bacillus subtilis* as a Model Organism

B. subtilis is an aerobic, endospore forming, rod-shaped organism that belongs to the Phylum Firmicutes and Family Bacillaceae. In nature, it is commonly found in soil. It has a relatively small genome of 4,214,810 bp with 4100 protein coding genes [37]. *B. subtilis*

has a rigid cell wall composed of peptidoglycans, teichoic acids, lipoteichoic acids and proteins. It is a Gram-positive model organism that has high commercial value as a heterologous protein expression system and as an important source of industrial enzymes [38]. It also finds its use in microbial fermentation. It has been considered a classic system to study natural phenomena such as bacterial chromosome replication, sporulation, swarming motility, natural competence and carbon catabolite repression. *B. subtilis* also has a characteristic phosphorylation machinery that makes it the best model organism to study STY phosphorylation events. It has nine known STY kinases and seven phosphatases, none of which are classified as essential [39]. Knock-out strains of these kinases and phosphatases are viable and comparable to the wild-type (WT) strain thus making it a suitable organism for our study.

1.2.1. Serine, Threonine Kinases and Phosphatases in *B. subtilis*

Bacterial serine/threonine kinases are Hanks-type kinases that do not exhibit exclusive specificity towards serine or threonine residues and in most cases can phosphorylate both of them. *B. subtilis* has seven Hanks-type kinases, namely, HprK (involved in carbon catabolite repression); PrkC (involved in spore germination); YbdM (protein kinase D); RsbT (controls activation of the SigB sigma factor by environmental stress); RsbW (involved in control of SigB activity); SpoIIAB (involved in septation during sporulation conditions); YabT (involved in control of DNA integrity during spore development) and five serine/threonine phosphatases, namely, PrpC (involved in spore germination); RsbP, RsbU, RsbX (involved in control of SigB activity); SpoIIE (required for normal formation of the asymmetric septum).

The serine/threonine kinases and phosphatases (PrkC, YbdM, YabT, PrpC) analyzed in this study are briefly described here.

PrkC, a Hanks-type kinase and PrpC, a serine/threonine phosphatase are encoded in the same operon. *B. subtilis*, a gram-positive soil bacterium, is able to undergo cell differentiation to form spores [40]. Spore germination occurs in response to cell wall muropeptides. The germination response to muropeptides involves the conserved, eukaryotic-like membrane serine/threonine kinase, PrkC, the extracellular domain of which contains PASTA repeats that are responsible for peptidoglycan binding. During

germination, PrkC phosphorylates the essential translation factor EF-G [41]. EF-G factor is a prokaryotic elongation factor that catalyzes the translocation of the tRNA and mRNA during polypeptide elongation. PrpC dephosphorylates EF-G. Thus, this kinase/phosphatase pair have opposing functions during the stationary phase of growth. PrkC has also been reported to phosphorylate metabolic enzymes involved in carbohydrate metabolism for example transaldolase YwjH, glutamine synthetase GlnA, isocitrate dehydrogenase Icd and acetolacto-decarboxylase AlsD [42]. The goal of my research was to identify additional *in vivo* substrates of PrkC-PrpC, the kinase-phosphatase pair, by employing a quantitative phosphoproteomics approach. Using a SILAC based phosphoproteomics screen, YkwC was identified as a novel *in vivo* substrate of this enzyme pair [43]. YkwC is an uncharacterized oxidoreductase. It belongs to the 3-hydroxyisobutyrate dehydrogenase family (according to Swiss-Prot). It specifically acts on the CH-OH group (donor) with NAD⁺ or NADP⁺ as acceptor [44]. It possibly plays a role in valine metabolism.

YabT is a serine/threonine transmembrane kinase that is implicated to play a role in the control of DNA integrity during spore germination. YabT contains a DNA binding domain that is essential for its activity and lacks the canonical extracellular signal receptor domain. Cells devoid of YabT sporulate slowly and exhibit reduced resistance to DNA damage [45]. Through *in vitro* phosphorylation assays, YabT was shown to phosphorylate DNA-recombinase RecA at Ser2. Phosphorylation of RecA was shown to promote the formation of transient RecA foci [45]. The global transcriptional regulator AbrB was also identified to be phosphorylated *in vitro* by YabT on Ser86. Phosphorylation of AbrB was shown to abolish its DNA-binding activity [46]. YbdM is yet another serine/threonine kinase whose functional significance has not yet been characterized. YbdM was shown to phosphorylate DegS (a two-component sensor kinase involved in competence regulation and flagellar motion [47]) *in vitro* on Ser76 resulting in phospho-transfer to DegU (two-component response regulator) leading to regulation of DegU dependent transcription [48]. YbdM was also shown to phosphorylate AbrB *in vitro* [46].

1.2.2. Tyrosine Kinases and Phosphatases in *B. subtilis*

Bacterial tyrosine kinases or BY-kinases are enzymes with substrate phosphorylation specificities for tyrosine residues. There are two BY-kinases identified in *B. subtilis*, namely, PtkA (implicated to play a role in biofilm formation); YveL (involved in exopolysaccharide synthesis) and two tyrosine phosphatases, namely, PtpZ (implicated to play a role in biofilm formation); YfkJ (plays a role during stress conditions). These are briefly described here.

PtkA, a BY-kinase and PtpZ, a tyrosine phosphatase are also encoded in an operon. During the stationary phase of growth, these two enzymes show opposing growth patterns. While PtkA expression drops, PtpZ protein amounts steadily increase right from onset of stationary phase [43]. This kinase-phosphatase pair is reported to play a role in biofilm formation. Biofilms are antibiotic resistant, sessile bacterial colonies that adhere to the surface. It is composed of extracellular DNA, proteins and polysaccharides. In addition to transcriptional control of initiation of biofilm formation, phosphorylation activity by PtkA also plays a major role and $\Delta ptkA$ strains show defective biofilm formation [49]. UDP-glucosedehydrogenase catalyzes the conversion of UDP-glucose to UDP-glucuronate in the presence of NAD⁺. PtkA has been reported to phosphorylate UDP-glucosedehydrogenase at Tyr70 [50], and to phosphorylate and co-localize with some of its substrates namely, single-stranded DNA exonuclease YorK and aspartate semialdehyde dehydrogenase Asd [51].

YveL is a BY-kinase also implicated to play a role in biofilm formation. It is the second component of a 15-member *eps* operon. The SinR repressor was shown to negatively regulate the transcription of the *eps* operon [52] while RemA, a transcriptional activator, was shown to activate the *eps* operon [53]. YfkJ is a tyrosine phosphatase that forms a part of the *yfkJ-yfkI-yfkH* operon (according to DBTBS). YfkJ gene expression was seen to be induced on stress conditions [54].

1.3. Mass Spectrometry Based Proteomics

“*Proteome refers to the total protein complement able to be encoded by a given genome*” as defined by Ian Humphery-Smith and co-workers in 1995 [55], was the first reference of

characterization of gene products of an organisms, in this case *Mycoplasma genitalium*, using MALDI-TOF followed by sequencing via N-terminal Edman degradation. The classical approaches for mapping a proteome is by employing two-dimensional gel electrophoretic techniques (2-DE). This technique, first introduced in 1975 [56], is based on separation of proteins according to their isoelectric points in the first dimension and according to their molecular weight in the second dimension. Coupling 2-DE to a mass spectrometer helped in the identification of separated proteins, thus providing a global biological insight. Since its conception, the 2-DE proteomics approach has come a long way with technical advances leading to increased sensitivity and reproducibility. However, the ability of 2-DE-MS/MS to characterize proteins of medium or low abundance was insufficient. In addition, underrepresentation of membrane proteins, low resolution of proteins with extreme pI or post-translational modifications and impracticality of quantification, were substantial limitations of this technique [57, 58]. These shortcomings were slowly resolved with the advent of whole genome sequencing in combination with developments in liquid chromatography, improvements in MS instrumentation and expansion of computational power.

The introduction of ionization techniques (Matrix-Assisted Laser Desorption Ionization in 1988 [59, 60] and Electrospray Ionization in 1989 [61]) solved the problem of converting proteins and peptides into gas phase intact ions. Historically, protein fractionation and identification was solely done by in-gel approaches (PAGE [62] and 2-DE [63]) followed by peptide mass fingerprinting [64-66]. These extensive procedures demanded for and created advancement in techniques leading to the next-generation of proteomics, namely, multidimensional protein identification technology, MudPIT [67]. This eventually paved the way for the present day, routinely employed, high-throughput “shotgun” (named after shotgun DNA sequencing) proteomics approaches.

1.3.1. Shotgun Proteomics Workflow

Conventional shotgun proteomics approaches involve identification of proteins from complex mixtures (such as cells, tissues or organelles). This is done by a lysis step, traditionally in buffers containing chaotropes, detergents or reducing agents; additionally this can also be coupled to mechanical cell disruption techniques, such as

sonication, freeze/thawing or homogenization, for improved lysis. Additives such as glass beads, enzymes (lysozyme, endonucleases) and protease inhibitors further facilitate cell lysis and protein extraction. The two most widely used buffer systems by the proteomics community include the RIPA buffer and urea-containing buffer [68]. The basic principle behind cell lysis is denaturation of proteins via solubilization of the membranes and proteins. The next step is to extract proteins from lysates by employing chloroform/methanol or acetone precipitation. Subsequently, the extracted proteome is digested using specific proteases to form a complex peptide mixture. In order to reduce the complexity of the sample mixture, pre-fractionation or enrichment protocols are commonly employed on protein or peptide level. Protein pre-fractionation can be done by SDS-PAGE prior to digestion. Alternatively, peptide pre-fractionation can be done based on the charge, isoelectric point or hydrophobicity of the peptide by Isoelectric Focusing, Reverse Phase chromatography, Ion Exchange chromatography or Hydrophobic Interaction Chromatography. Alternatively, specific modifications (phosphorylation, acetylation, ubiquitylation) can be targeted and enriched for analysis on the LC-MS/MS. The pre-fractionated or enriched samples can then be loaded on a nano-HPLC/UPLC for further fractionation and reduction of sample complexity. This is followed by MS and MS/MS analyses. Most often, signal from the most abundant peptides suppress the low abundant co-eluting ions thus causing ionization, fragmentation and identification of only the high abundant ions at each time point. Each peptide is then ionized, isolated, fragmented and the fragment ions are detected in the mass spectrometer by tandem mass spectrometry (MS/MS or MS²). Frequently used ion fragmentation techniques such as collision-induced dissociation (CID) yield the -b (amino terminal fragment ions) and -y (carboxy terminal fragment ions) ion series. The tandem mass spectra are then interpreted for peptide identification which is largely done by matching peptide spectra to a protein database ("database search") using specialized software and matching algorithms ("search engines"). Inferred peptide sequences are thus assigned to parent proteins whose biological significance can then be further evaluated [69] **(Figure 3)**.

1.3.2. Instrumentation

Mass spectrometry is an indispensable technique that finds many applications in biology. The increase in sensitivity, accuracy and resolution of the instruments, coupled with improvements in sample preparation and chromatography, has opened new dimensions in analyses of complex biological systems such as drug metabolism, lipid analysis, metabolomics, qualitative and quantitative proteomics, direct analysis of intact proteins and live -cell and -tissue imaging [70]. The technical definition of a mass spectrometer is “an instrument that can ionize a sample and measure the mass-to-charge (m/z) ratio of the resulting ions” [71]. The first mass spectrometer (mass spectrograph) was constructed in 1912 by J.J. Thompson and further developed by Aston [72]. Mass spectrometers were primarily used to study the elemental composition and natural abundance of isotopes [73].

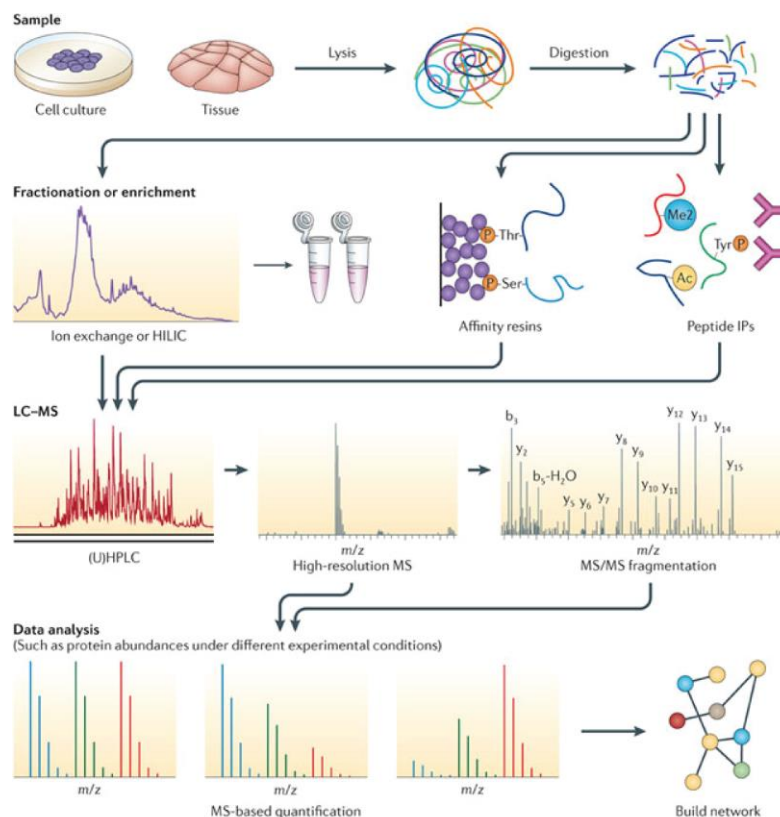


Figure 3: Shotgun proteomics workflow – depicting cell lysis, extraction, digestion, pre-fractionation and/or modification-specific enrichment, injection into the nano-LC system followed by subsequent analysis on the MS instrumentation [69].

Rapid evolution of MS technology led to the 2002 Nobel Prize in Chemistry being awarded to J. Fenn and K. Tanaka for their development of the soft ionization methods, namely the electrospray ionization and matrix assisted laser desorption ionization, for the analysis of biological macromolecules. Constant advances in technology has enlarged the scope of mass spectrometry which is now routinely used for biological applications. All mass spectrometers comprise of three major components – 1) an ion source, 2) a mass analyzer and 3) a detection system (**Figure 4**).

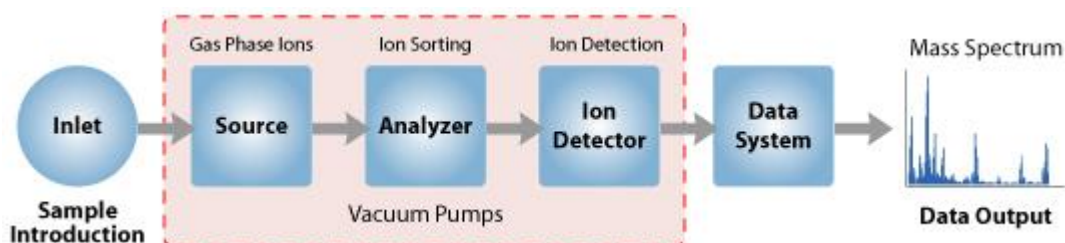


Figure 4: Basic components of a mass spectrometer – include the ion source, mass analyzer and the ion detector. (Source: http://www.premierbiosoft.com/tech_notes/mass-spectrometry.html).

Ion Source - The first step in mass spectrometric analysis is the production of gas phase ions that are amenable for further analysis in an electric field. Proteins and peptides are non-volatile and unstable molecules that need to be ionized into gaseous analytes without inducing extensive degradation [74]. There are many types of ionization methods (hard and soft ionization types) that are used in mass spectrometry – for example classical methods such as electron impact ionization (EI), fast atom bombardment (FAB) and the more modern techniques such as atmospheric pressure chemical ionization (APCI), electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI). MALDI and ESI are the two most prevalent techniques used in modern day proteomics (**Figure 5**). In MALDI, peptides or proteins are mixed with a matrix (often a weak organic acid such as sinapinic acid or 2,5-dihydroxybenzoic acid), deposited on a rigid support and irradiated with a pulsed laser. As the matrix absorbs the laser radiation, the mixture is heated and it expands, transferring a proton to the sample, leading to the ionization of the analyte and its transfer into the gas phase.

MALDI ionization requires series of laser pulses in order to obtain an acceptable signal-to-noise ratio for ion detection [74].

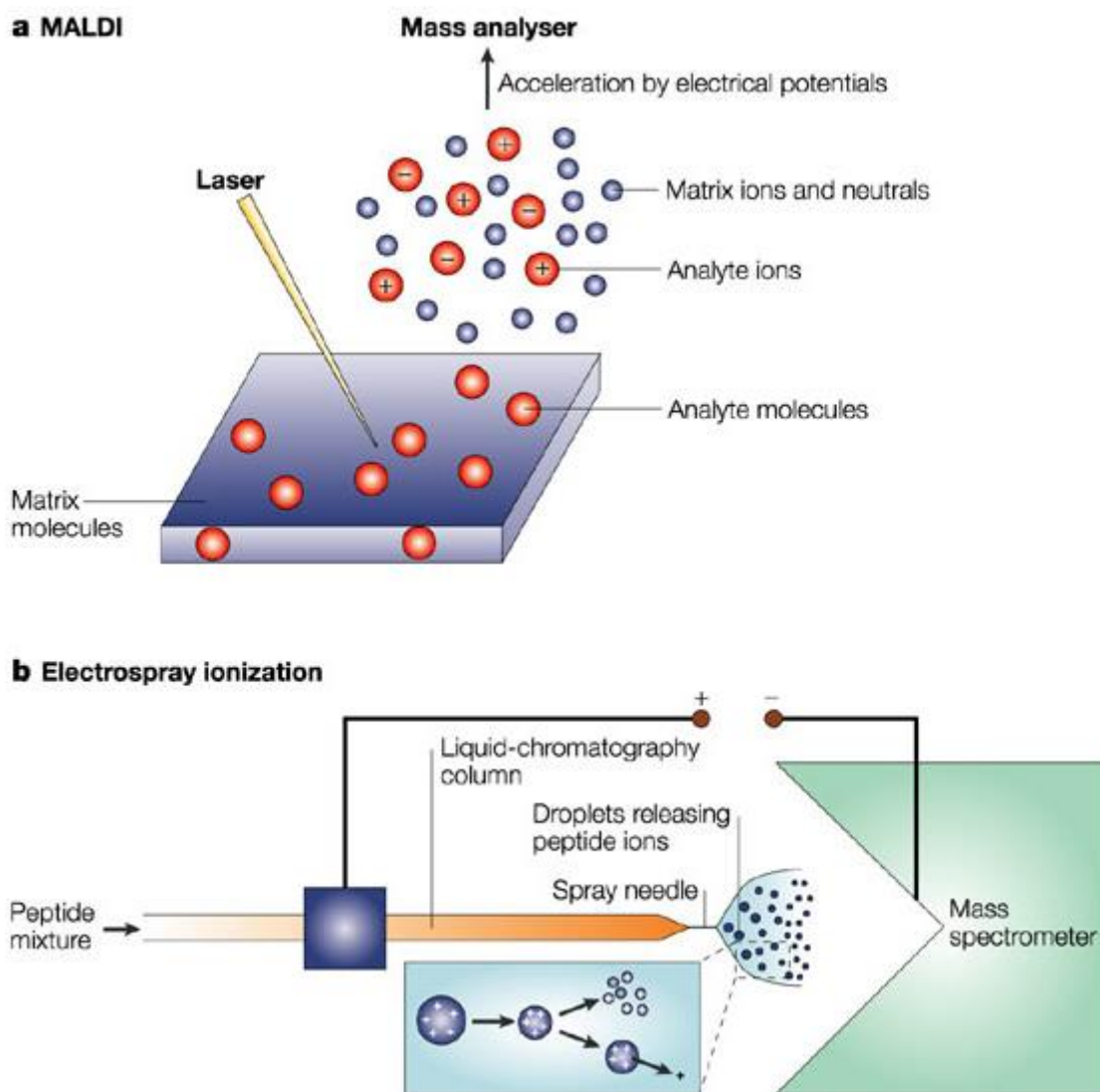


Figure 5: Ionization methods – Two most commonly used soft ionization methods are a) matrix assisted laser desorption ionization and b) electro spray ionization [75].

MALDI-generated ions are predominantly singly charged. MALDI is most often combined with a TOF mass analyzer, and is not easily compatible with liquid chromatography. In contrast to MALDI, the ESI technique produces ions from solution. Electro spray ionization involves application of a high voltage (2–6 kV) between the emitter (e.g. chromatographic column) and the inlet of the mass spectrometer (heated capillary).

Since ion formation depends extensively on solvent evaporation, samples taken for ESI are generally prepared in volatile organic solvents [76]. Compounds that increase conductivity (important for decreasing the initial droplet size) and act as a source of protons such as acetic acid are additionally added to the solvent. The electrospray cone or Taylor cone [77] emanates a jet of charged solvent particles which when above the threshold voltage start evaporating until they reach their Rayleigh limit [78]. At this point electrostatic repulsion of the like charges occur, causing consecutive Coulombic fission processes that are preceded by additional solvent evaporation at each cycle. Droplet formation and desolvation are further aided by using heated capillaries [74]. ESI can be coupled to front-end fractionation by liquid chromatography and mainly generates multiply charged ions thus extending the mass range to accommodate even large molecules or fragment ions (upto kDa or MDa orders of magnitude). Multiple charged ions generated by ESI are also amenable to fragmentation methods such as electron transfer dissociation [79] that has the advantage for analysis of modified peptides.

Mass Analyzer - The second step in mass spectrometry is the transportation of the ions to the mass analyzer by electric or magnetic fields guided by the ion optics system. The mass analyzers form an integral part of the MS instrumentation as they can collect/store the ions, identify them and separate them with high accuracy and sensitivity based on their mass-to-charge ratio. The dynamics of the charged fragments under electric or magnetic fields in vacuum is governed by the differential equation – $(m/Q)a = E + v * B$ where m refers to the mass of the ion, Q refers to the charge of the ion, a refers to the acceleration, E refers to the electric field, v refers to the ion velocity and B refers to the magnetic field. Parameters such as the mass resolution (the ability to distinguish between two peaks with differing m/z ratios), mass accuracy (the degree of deviation of the measured m/z ratio from the theoretical m/z ratio), sensitivity (the number of ions detected with a specific signal-to-noise ratio), dynamic range (difference between the smallest and largest m/z ratio detected) and scan rate (time required to scan all ions to produce a spectrum) need to be taken into consideration as they determine the performance or fitness of the mass analyzer for the purpose of bioanalysis. There are several types of mass analyzers available, used in different instruments, like Time of

Flight (TOF), the quadrupole, the Ion trap quadrupole (IT), the Fourier transform-ion cyclotron resonance (FTICR), the Sector mass analyzer (electric, magnetic or static) and the Orbitrap mass analyzer. Hybrid mass spectrometers (combination of more than one mass analyzer) have been developed to answer specific needs during analysis. The ion trap, Orbitrap and FTICR instruments separate ions based on the m/z resonance frequency, quadrupoles separate ions based on their m/z stability whereas the Time of Flight instruments are based on the flight times of the charged fragment ions [74].

In the Linear ion trap, the ions are trapped radially within a set of quadrupoles and axially by maintaining a potential difference at the electrode ends [80] (**Figure 6**). The mass filter present helps to select ions of a particular m/z ratio. Varying the radio-frequency and the direct current potential applied helps to eject ions of a particular m/z ratio from the high pressure storage cell to the low pressure cell where fragmentation and detection occurs. Detection can be enhanced by filling the ion trap for longer periods. In comparison to the Orbitrap, Linear ion traps have faster scan rates and data acquisition abilities. They can operate as stand-alone instruments and are frequently used as the front end of hybrid instruments. The Orbitrap on the other hand can achieve higher mass resolution, mass accuracy and a good dynamic range and can be stand-alone or interfaced typically to a Linear ion trap in hybrid mass spectrometers to achieve better sensitivity, resolution and speed. Ions are electrostatically trapped around a central spindle shaped electrode around which the ions orbit and oscillate in an axial direction [81]. The oscillation frequency generates an image current which is recorded. Fourier transform algorithms incorporated enable conversion of these signals to m/z spectra (**Figure 6**).

Detectors – The final element of a mass spectrometer is a detector, which typically records the current produced when an ion hits its surface. The most universal type of detector used is the electron multiplier. Signal amplification is required as the number of ions entering the detector at any given point of time is minute. The electron multiplier is capable of multiplying incident charges. A single electron is bombarded on an emissive plate generating electron emission. An electric potential is then applied between consecutive plates causing acceleration of emitted electrons to the next plate subsequent

inducing secondary emission [82]. Other detection systems include the Faraday cups [83], ion-to-photon detectors [84] or microchannel plate detectors [85].

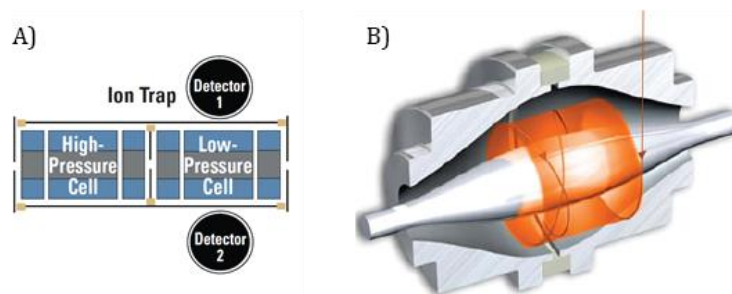


Figure 6: Mass analyzers – Schematic of a – A) 2D-quadrupole linear ion trap, and B) Orbitrap mass analyzer (Source: Thermo.com).

1.3.3. Chromatography

Protein identifications are highly dependent on the separation techniques used. As proteins are identified by the m/z of their peptide fragments, employing chromatographic methods ensures sufficient peptide separation averting unambiguous identifications [74] due to ion suppression. In addition, chromatographic separation increases local concentration of a peptide and thereby facilitates its detection by MS. Reverse-phase liquid chromatography is usually coupled to the front end of mass spectrometers (termed as LC-MS) as an additional step of fractionation for higher detection rates. This process separates peptides according to their hydrophobicity. The chromatographic column is generally packed with small pore size hydrophobic material (stationary phase) and operates at high pressure. The sample is mixed with an appropriate polar mobile phase and is forced at high pressures into the chromatographic column. As a result, the hydrophobic peptides in the sample tend to adhere to the hydrophobic stationary phase causing the hydrophilic peptides to pass through first. The hydrophobic peptides are then eluted by decreasing the polarity of the mobile phase using an organic solvent. The more hydrophobic the peptide, the higher the concentration of the organic solvent required to elute it. The most popular stationary phase is an octadecyl carbon chain bonded silica while the most commonly used mobile phase solvents are acetonitrile, methanol and tetrahydrofuran. A number of factors can

affect the quality of the MS data obtained. In general, flow rates can be up-to 1 mL/min, but columns with smaller IDs require lower flow rates (typically 200 nL/min) and achieve better separation. Buffers containing inorganic ions such as phosphate and acetate could lead to ion suppression and contamination due to formation of MS adducts of sodium or potassium [86]. Elevated temperatures further aid in improved separation of peptides. Increasing column length while decreasing particle size helps to improve peak resolution, sensitivity and analysis time [74]. Optimization of the chromatographic framework is essential for the generation of high resolution spectra with better signal-to-noise ratio.

1.3.4. Database Search

Accurate and precise peptide and protein identification forms a crucial part of the high throughput shotgun proteomics approach. Typical shotgun proteomics workflow involves digestion of proteins into peptides using an appropriate endoprotease followed by nano-LC-MS analysis. The ionized peptide (precursor) ions are subsequently fragmented and give rise to MS/MS spectra which are used to infer peptide sequence and identify proteins. Peptides present in the sample mixture are sequenced and assigned to parent protein molecules using sophisticated downstream bioinformatics. Identifications are made based on the quality of the match between the observed and the predicted spectrum-sequence nature (**Figure 7**) with the assumption that all protein coding regions in the genome are accurately annotated and all gene products are present in the reference protein database [87]. Manual validation of peptide-spectrum matches are not feasible with the current large datasets generated by tandem mass spectrometry. Consequently, peptide sequencing efforts lie on the adeptness of database search engines which are normally used. The three most commonly used search engines in proteomics are Mascot [88], Sequest [89] and MaxQuant/Andromeda [90], which in general match the experimentally acquired spectra to the theoretical spectra generated from *in silico* digestion of peptide sequences. The algorithm integrated in Sequest is an example of a descriptive model for database searching, that uses a cross-correlation score (Xcorr) calculated between the theoretical and experimental MS/MS spectra to discriminate between correct and incorrect peptide assignments. The scoring is additionally

dependent on the mass and charge state of the peptide and the quality of the spectra generated. This scoring was further normalized to the spectral noise and peptide size to improve the scoring assignment. [91]. Mascot or Andromeda based search engines use statistical and probability models for database searching by employing the Empirical Bayes approach. The fundamental principle of this model is to calculate a probability score for each peptide assignment. The best match is considered to be the one with the lowest probability score. The model used in other search engines such as PeptideProphet employ a single discriminant search score which integrates the Xcorr score, ΔC_n (difference between the first and second highest Xcorr score), Δm (difference between mass of the peptide and the assigned peptide) and number of tryptic termini to calculate the probability scores [92]. In addition, there are spectral library search engines (SpectraST [93], Bibliospec [94]) that match the acquired spectra to a library of previously identified spectra and *de novo* search engines (PEAKS [95], Lutefisk [96]) that identify peptides based exclusively on their MS/MS peak pattern without any reference sequences [97]. Database searching algorithms match the acquired mass spectrum to a specific peptide sequence from a protein database and assign a score. This statistically based peptide-sequence match has an inherent risk of being a false positive [98]. Additionally, some degree of ambiguity also exists while attempting to match MS/MS spectra of low quality or low intensity. The widely used target-decoy search strategy helps determine the false positive rate (FDR). The FDR is calculated by searching a decoy database consisting of reversed or scrambled sequences of the identical protein entries as the search database and dividing the false positives obtained by the total proteins identified (**Figure 7**). The FDR is typically fixed at 1 % to 5 % at the protein and peptide level implicating that per 1000 proteins identified, 10 to 50 proteins are false positives [98]. Along with false positive estimations, other parameters can be assessed from the target-decoy strategy that help evaluate the scoring scheme (**Figure 8**). While this strategy works efficiently for large proteome datasets, there is fundamental difficulty with annotating spectra of post-translationally modified peptides (discussed in section 1.4.1). Thus data interpretation faces challenges that need to be evaluated and modified to unleash the complete potential of proteomic research.

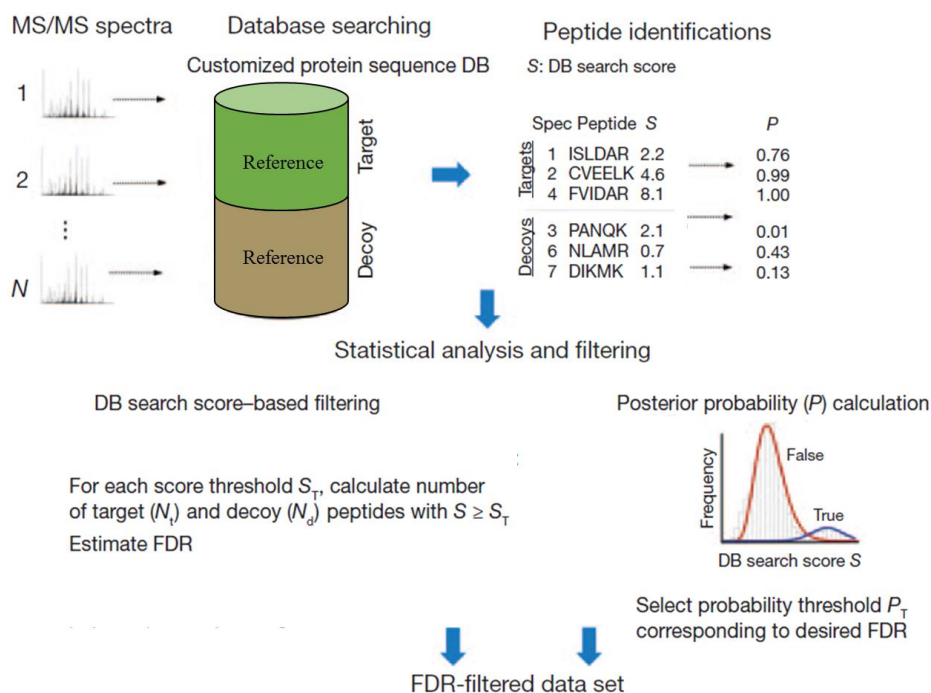


Figure 7: Estimation of identification confidence – MS/MS spectra are searched against a protein database using the target-decoy strategy. A scoring based filter identification is applied and the FDR is calculated based on the number of target and decoy sequences that pass the set threshold (Adapted from [87]).

Measurement	Formulation	Description of estimate
False positive (FP)	$2 \times$ passing decoy assignments	Number of incorrect assignments above score threshold
True positive (TP)	Total passing assignments – number of FPs	Number of correct assignments above score threshold
Total correct (TC)	Maximum TPs for all evaluated score criteria combinations	Number of total correct assignments in the data set
Total incorrect (TI)	Total assignments – TC	Number of total incorrect assignments in the data set
False negative (FN)	TC – TP	Number of correct assignments falling below score threshold
True negative (TN)	TI – FN	Number of incorrect assignments falling below score threshold
Precision	$TP / (TP + FP)$	Fraction of correct assignments above score threshold
FP rate	$FP / (TP + FP)$ or $1 - \text{precision}$	Fraction of incorrect assignments above score threshold
Sensitivity	TP / TC	Fraction of all correct assignments above score threshold
Specificity	$TN / (TN + FP)$	Fraction of all incorrect assignments below score threshold
Accuracy	$(TP + TN) / \text{total assignments}$	Fraction of all assignments correctly classified by score threshold

Figure 8: Measurements derived from target-decoy strategy – The table lists additional parameters that are derived from applying the target-decoy search strategy [99].

1.3.5. Protein Quantification

Mass spectrometry can not only be used to identify peptides and proteins or obtain structural information but can also provide an insight into the differences in the pathways regulated between experimental conditions or a healthy and diseased cell. Classically 2DE based approaches were employed for quantification purposes [100]. But

more and more MS based strategies are being used for the sole purpose of increased sensitivity and reproducibility. Mass spectrometry however does not have inherent quantitative abilities due to differences in ionization efficiencies and peptide detectability between samples. Thus direct comparison between peptides derived from different experimental states was essential. One possibility was the stable isotope dilution theory which was based on the property of similar peptides (in their labeled and unlabeled form) to have similar retention times during chromatographic separation and similar chemical properties during mass spectrometry analysis [101]. In general, comparison between sample conditions could be achieved by two ways – relative quantification which compares protein abundances between samples and yields a quantification ratio or relative changes, and absolute quantification that provides information on the absolute amount or protein concentration in a given sample [102]. Stable isotope labeling, first introduced to proteomics in 1999 [103, 104], is now a widely adopted strategy. Isotopic labels can be introduced into peptides metabolically, chemically, enzymatically or as a spiked in external standard in the form of synthetic peptides [105] (**Figure 9**). The earliest time point possible for the incorporation of a stable isotope into proteins is by metabolic labeling which occurs during cell growth. Classically, ^{15}N was used to achieve complete labeling of cells [104], but as only 1/3rd of the protein backbone is made up of nitrogen there resulted a loss in quantitative information and protein regulation. In addition, there are several inherent complications related to quantitation: 1) the corresponding mass shift depends on the number of nitrogen atoms, resulting in variable length of the stable isotope label, 2) the isotope distribution of the ^{15}N -labeled peptides changes with the number of ^{15}N atoms and is difficult to model, which leads to difficulties in identification and quantification of ^{15}N -labeled peaks. Stable isotope labeling by amino acids in cell culture or SILAC was introduced in 2002 [106] and is frequently implemented as a part of the quantitative proteomics workflow. Cells are grown in medium containing $^{13}\text{C}_6$ -arginine and/or $^{13}\text{C}_6$ -lysine and following tryptic cleavage, every identified peptide can theoretically be quantified due to the presence of the C-terminal labeled amino acid which results in a known and constant mass increment in comparison to its non-labeled counterpart.

Relative quantification is performed at the MS level by comparing the intensities of the isotope clusters of the non-labeled and corresponding labeled peptide.

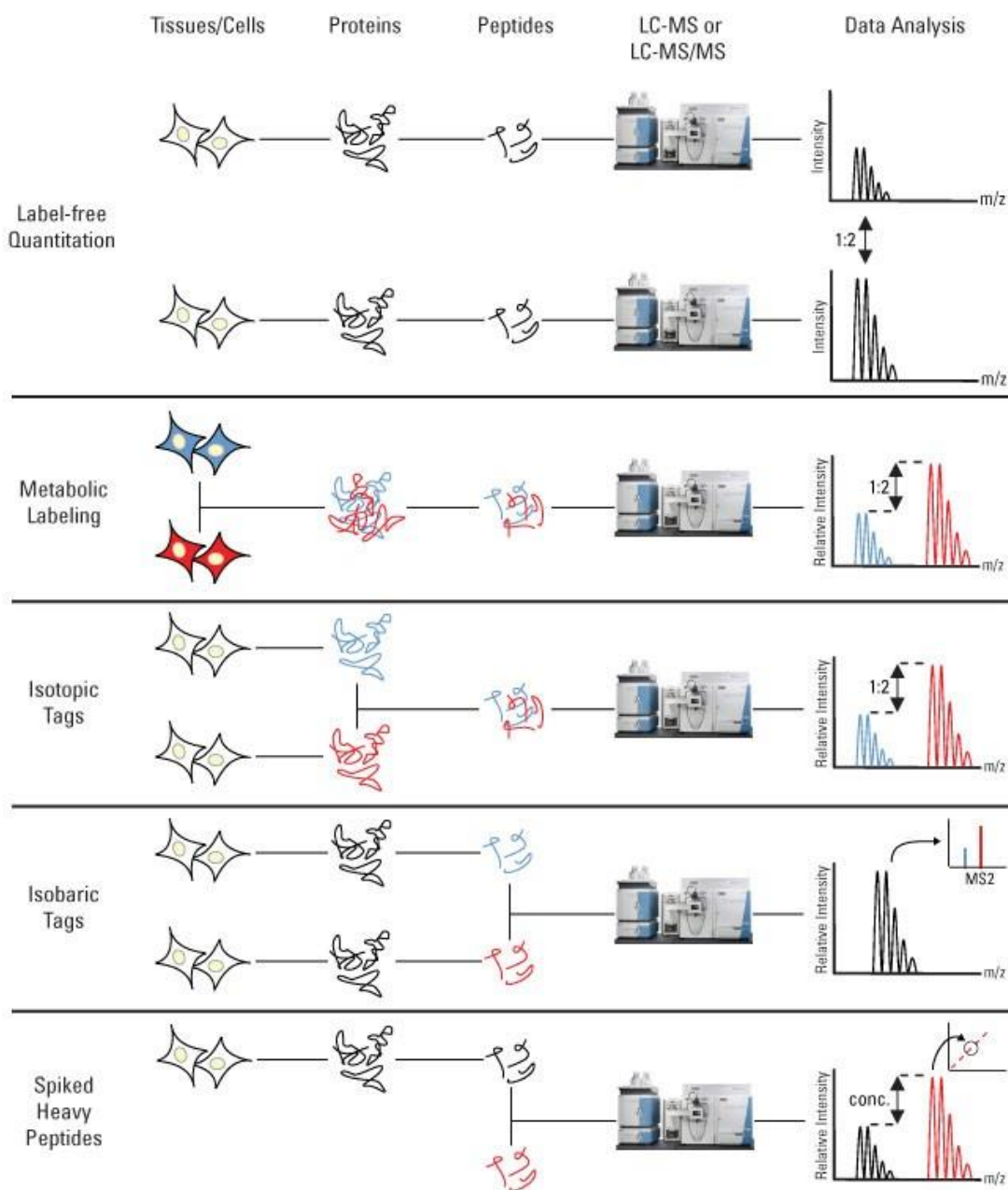


Figure 9: Overview of the methods of quantification – Commonly used labeling techniques in proteomics research are indicated in the schematic above along with the points of sample labeling and analysis on an LC-MS platform; with the exception to label free quantification wherein quantification occurs by spectral counting or comparison of peak intensities (Source – Thermo.com).

Due to fact that samples are combined at the level of intact proteins, metabolic labeling has the advantage of reduced errors due to sample preparation and LC-MS procedures which increases quantification accuracy.

Labeling of peptides and proteins can be achieved by chemical labeling or enzymatic derivatization *in vitro*. A specific way to introduce a label into peptides is by using Glu-C catalyzed incorporation of O^{18} during digestion [107], the drawback being incomplete labeling at most instances due to different label incorporation rates [108]. Isotope-coded affinity tag (ICAT) approach was established in 1999 [103] in which cysteine residues were specifically derivatized with a reagent containing either none or eight deuterium atoms linked to a biotin group for affinity purification followed by MS analysis. As cysteine containing peptides were fewer in number, ICAT reduced the complexity of the mixture but this was also the main disadvantage of this method. Another approach of labeling was to target the peptide N-terminus and the ϵ -amino group of lysine residues e.g. isotope tags for relative and absolute quantification (iTRAQ) [109] and tandem mass tags (TMT) [110]. This method employs N-hydroxy succinimide chemistry to modify the primary amines of peptides with isobaric tags consisting of a reporter and a mass balance group. Differentially labeled peptides co-elute during chromatographic separation and relative quantitation is obtained during fragmentation, which results in dissociation of the reporter groups. Although this concept is suitable for multiplex analysis (simultaneous analysis of up to 10 samples), a general shortcoming of the chemical labeling strategies is that they are often prone to side chain reactions causing formation of unexpected products that hinder quantification [101]. Label free quantification (LFQ) deserves a special mention as it fundamentally differs from other labeling strategies discussed above. LFQ based quantification is done by comparing the signal intensities of the peptide precursor ions or by counting and comparing the number of fragment spectra acquired for all identified peptides [101]. Advantages of LFQ include being a simple, relatively cost-efficient and time-saving procedure with no limit of number experimental conditions that can be compared, but the downside to this strategy is low quantification accuracy due to variations between experimental replicates as a result of manual error and variability of signal intensity during MS measurement. The second type of quantification is absolute quantification for protein abundance using internal standards.

AQUA (absolute quantification of proteins) [111] and iBAQ (intensity based absolute quantification) [112] are the two most commonly used methods. While quantification through AQUA depends on comparison of the mass spectrometric signal of a known amount of a stable isotope-labeled standard peptide spiked in to the signal of the endogenous peptide of interest in the sample, iBAQ relies on summation of all peptide intensities of a protein of interest and subsequent comparison to the total number of identified peptides. Quantitative proteomics thus enables the determination of differences pertaining to changes in protein expression and regulation between distinctive physiological states and thereby facilitates biological understanding.

1.4. Phosphoproteomics

Mass spectrometry based proteomics is a suitable platform to identify, quantify and characterize proteins that have been phosphorylated. It can be employed not only to identify phosphorylation events but also to accurately localize them on peptides and proteins. Additionally at the biological level, phosphoproteomics also provides information about which protein or pathway has been activated due to phosphorylation which in turn reflects protein activity and/or localization. The ability to measure phosphorylation events at a global level makes it a powerful tool to analyze and understand the dynamics of signaling networks. Conventionally, 2D based gel approaches were used to study protein phosphorylation on a proteome-wide scale [113]. Advances in instrumentation and techniques through the years have enabled mass spectrometry to be the prime method for large scale phosphorylation analysis. Phosphoproteomics analysis comprises of four important steps – 1) sample preparation, 2) phosphopeptide enrichment, 3) LC-MS/MS and 4) bioinformatics (**Figure 10**). A brief overview of the important features required for phosphorylated peptide identification and analysis is mentioned here.

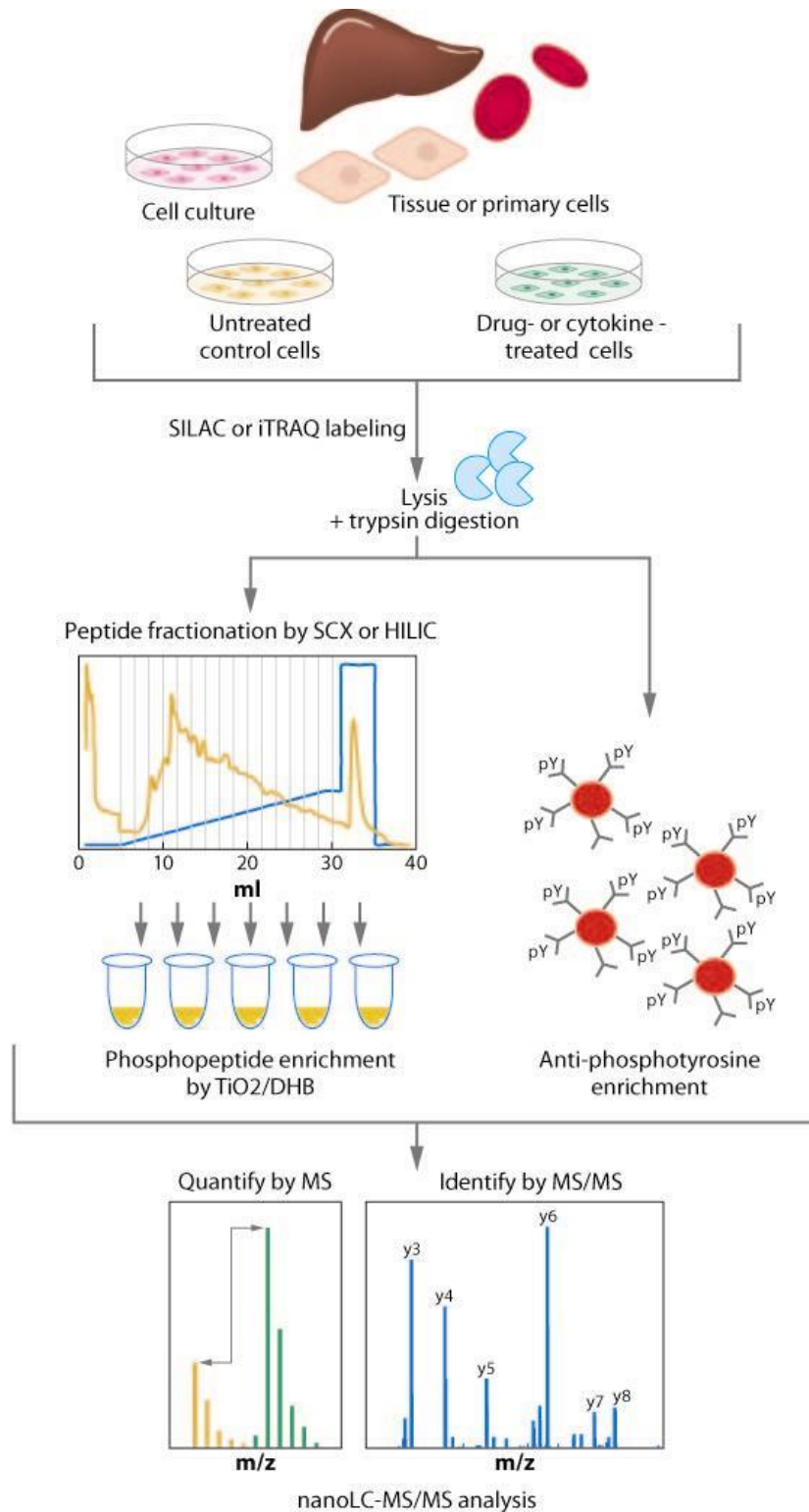


Figure 10: A typical phosphoproteomics workflow – Cells or tissues are labeled with stable isotopes of amino acids, lysed, proteins are extracted and mixed together. The lysates are digested followed by pre-fractionation and phosphorylated peptide enrichment. The enriched phosphorylated peptides are separated on a nano-LC and analyzed on the mass spectrometer [114].

Starting with sample preparation, a straightforward proteome analysis generally requires around 100 μ g of total protein. Conversely, large-scale phosphoproteomics typically requires 10-50 times more of starting material due to the low abundance of phosphorylated peptides. Additionally, it is essential to utilize an appropriate protease and include phosphatase inhibitors during sample preparation procedures. Appropriate enrichment techniques have to be employed to maximize identifications. Different analytical strategies that combine LC fractionation along with phosphorylated peptide enrichments help enhance sensitivity [115]. Precise localization of phosphorylation events using the various *in silico* software tools available is yet another essential factor to be taken into consideration during bioinformatic analysis. Further important step in phosphorylated peptide identification is LC-MS/MS. Collision induced dissociation or CID is routinely employed for peptide ion fragmentation, although lately other fragmentation methods such as high-energy CID (HCD) or electron transfer dissociation (ETD) are increasingly used in protein phosphorylation analysis. Due to the labile nature of phosphorylation on serine and threonine residues, conventional CID results in the loss of phosphoric acid (98 Da) from the precursor ion, which is also referred to as neutral loss and often results in insufficient peptide backbone fragmentation. Due to β -elimination of the phosphoester bond, phosphorylated serine is converted to dehydroalanine and threonine to dehydroaminobutyric acid [116] resulting in modification-specific mass shift (-18 Da), which is used for the localization of modified residues. This limits the information obtained to identification of the phosphorylation event and not localization. Infrared multiphoton dissociation (IRMPD) helps minimize such neutral losses but is restricted to Fourier transform-ion cyclotron resonance (FT-ICR) instruments [117]. Strategies such as MS³ and MultiStage Activation (MSA) are employed which activate the pre-selected neutral loss peaks [118, 119]. During MS³ the peak selected for additional fragmentation is isolated prior to activation resulting in a new set of product ions, while during MSA the second isolation step is eliminated and the spectra contains product ion information from both activation events, making it a pseudo-MS³ approach. Following MS³ or MSA, phosphorylation events on serine and threonine residues are recognized by

the presence of fragment ions containing dehydroalanine and dehydrobutyric acid respectively [117] (**Figure 11**).

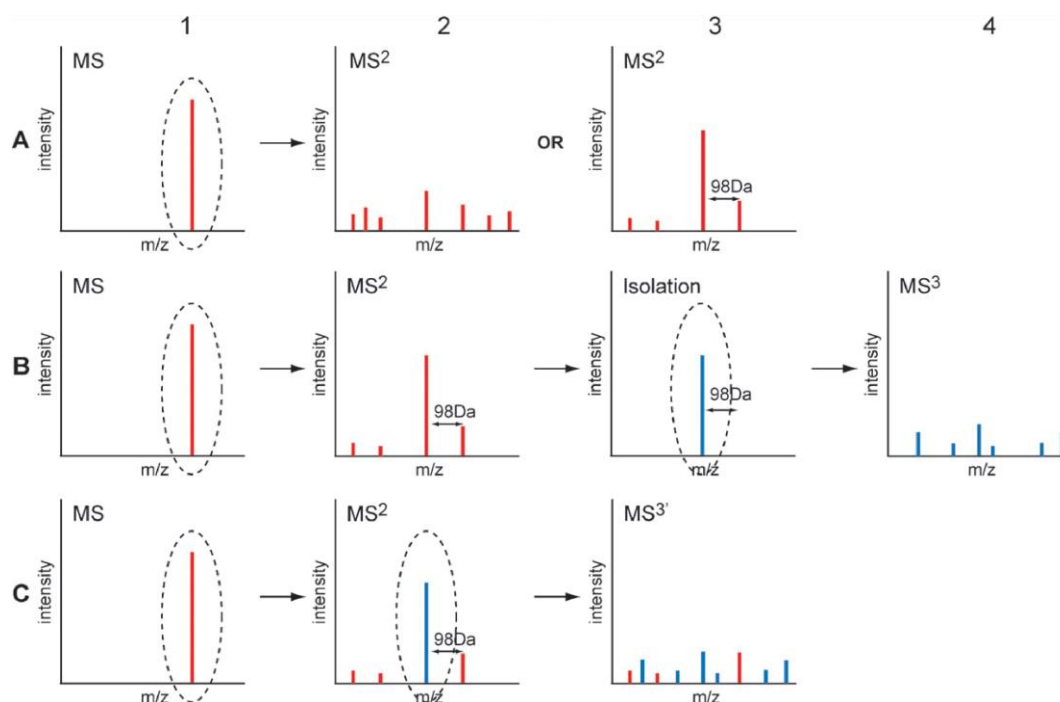


Figure 11: Schematic of MS², MS³ and MultiStage Activation (MSA) during peptide fragmentation – (A) During MS², a phosphorylated peptide is selected for CID and isolated (1A). The resulting fragments consist either of the y- and b-ion series (2A) or a dominant neutral loss peak corresponding to the loss of phosphoric acid (98 Da) (3A). (B) During MS³, a phosphorylated peptide is selected for CID and isolated (1B). If a neutral loss is detected (2B), the neutral loss peak is re-isolated (3B) and additionally fragmented (4B). (C) During MSA, a peptide is selected for CID and isolated (1C). If a neutral loss is detected (2C) the neutral loss peak is re-activated by CID (2C) resulting in a spectra containing fragment ions from both the collision events. Red and blue bars represent parent ions and fragment ions from the first and second fragmentation events, respectively [117].

Alternative fragmentation methods such as electron capture dissociation (ECD) or electron transfer dissociation (ETD) that allow for fragmentation to occur along the peptide backbone while maintaining the phosphate intact [120, 121] can be employed. Assessment of site occupancy helps to obtain a comprehensive overview of the phosphorylation dynamics under various physiological conditions. Phosphorylation site occupancy refers to the stoichiometry of the phosphorylation modification and can be defined as the ratio of the phosphorylated peptide and its non-phosphorylated

counterpart, with respect to an identical site. In 2010, a paper published on assessing phosphorylation site occupancy during mitosis in HeLa S3 cells [122] calculated absolute phosphorylation site occupancies based on the SILAC ratios obtained from a phosphorylated peptide, its non-phosphorylated counterpart and the corresponding total protein ratio (obtained from the phosphorylated and non-phosphorylated peptide). Similarly, another group in 2011 calculated the fractional phosphorylation site occupancy of *Saccharomyces cerevisiae* [123] by comparing the peak intensities of the non-phosphorylated peptide with and without phosphatase treatment. Both strategies for quantification of site stoichiometry are effective and help unravel their significance with regard to biological processes and reflect their importance in regulating such events.

1.4.1. Localization of Phosphorylation Events

This section is a part of "Resources for Assignment of Phosphorylation Sites on Peptides and Proteins" that has been accepted for publication in Methods in Molecular Biology, Springer Journals.

While identification and quantification of the modified protein is important for a deeper insight into signal transduction networks, the knowledge of the precise site of modification on the protein sequence is critical for a mechanistic understanding of the influence of the PTM on individual protein function [36, 115]. Experimental identification of phosphorylation sites is relatively labor-intensive, hence significant efforts have been invested in constructing *in silico* phosphorylation site predictors.

During the past two decades, mass spectrometry has emerged as a method of choice to study PTMs and signal transduction networks. Since the peptide sequence can be easily inferred from the MS/MS fragmentation pattern, these new protocols enable straightforward localization of many modification sites. Protein/peptide identification relies on search engines that match peptide fragmentation patterns to a database of known protein/peptide sequences [75]. PTMs on peptides are detected as a mass increment or deficit relative to the respective unmodified form of the peptide. Presence of a PTM on peptide fragments causes the m/z of the peptide to shift and these modified and unmodified fragments are of special importance in localization; absence of these

fragments results in ambiguity. Algorithms use this information to map the peptide to the best possible spectrum and generate a probability-based localization scoring for phosphorylation assignment. Contrary to assessment of FDR (False Discovery Rate), for site localization there is no similar equivalent measure or FLR (False Localization Rate). An incorrectly localized site would not be considered as a decoy match but have a close match to the correct spectrum [124]. Thus probability based scoring algorithms are used for this purpose.

While determination of the peptide sequence is relatively straightforward, detection and localization of modification sites is more complicated due to the fact that presence of PTMs on peptides increases the number of theoretical spectra that need to be matched to a mass spectrum (search space). The increase in search space that results may also largely escalate false positive identifications during database search. Furthermore, localization of a modification on a peptide requires comprehensive MS fragmentation patterns, but the presence of a phosphorylation modification, for example, often interferes with fragmentation resulting in lower coverage and significance scores. For example, loss of the modification from the peptide bond often occurs during the main fragmentation event, as in case of phosphorylation that occurs on Ser/Thr residues (so-called “neutral loss”). In these cases, CID often results in a prominent unmodified precursor ion peak that is devoid of information on peptide sequence and additional fragmentation needs to be performed on that peak for sequence and modifications site assignment [125]. Since Ser/Thr phosphorylation is “lost” in form of phosphoric acid, the modified serine is converted into dehydroalanine and the resulting mass shift (-18 Da) can be used to localize the modification site. High measurement mass accuracy (MMA) is also important for its generally positive influence on selectivity in case of large database search space, and it is additionally needed to resolve modifications that are very similar in mass, such as tyrosine phosphorylation (79.966 Da) and sulfation (79.956 Da); acetylation (42.011 Da) and tri-methylation (42.047 Da); formylation (27.990 Da) and dimethylation (28.031 Da).

For localization scoring purposes search engines such as Mascot [88] or Sequest [89] have integrated PTM-scoring algorithms, such as Delta-score or A-score, respectively. Software suites such as MaxQuant [90] have integrated scoring algorithms (PTM Score),

in addition to other modules for MS data processing, search engine and statistical post-processing of MS data. Additionally stand-alone programs such as PhosphoRS [126] can also be used which can apply post raw data processing with any software. MS/MS data interpretation can be done with various additional MS vendor tools such as Proteome Discoverer (Thermo Scientific), ProteinLynx (Waters), MassHunter/Spectrum Mill (Agilent) and others.

Dependent peptide analysis [90] incorporated within MaxQuant and Error tolerant search [127] with Mascot can also be used to identify unsuspected modifications on a peptide using the information obtained from their identified unmodified analogues. There are some fundamental strategies/*in silico* predictors widely used by the proteomics community to help localize sites accurately - like the A-score by Beausoleil *et al* [128] and the PTM Score by Olsen *et al* [129, 130] or modified variants of these.

1.4.2. Phosphorylated Peptide Enrichment Strategies

Detection of phosphorylation is relatively simple using gel based systems in combination with γ -P³² labeled ATP, antibodies (western blotting) or Pro-Q Diamond like staining schemes. Whereas quantification or localization of phosphorylation events is more challenging as they are generally more low abundant and information can be easily lost in the vast cellular milieu. Thus, due to the sub-stoichiometric and transient nature of this modification, phosphorylated peptides have to be enriched for in order to eliminate their non-phosphorylated counterparts from complex samples before analysis in the mass spectrometer [114]. Enrichment can be performed on two levels - on intact phosphorylated proteins or on phosphorylated peptides. Furthermore pre-fractionation can be employed prior to enrichment. Various enrichment strategies have been developed (**Figure 12**), some of the most common ones amongst them are described below.

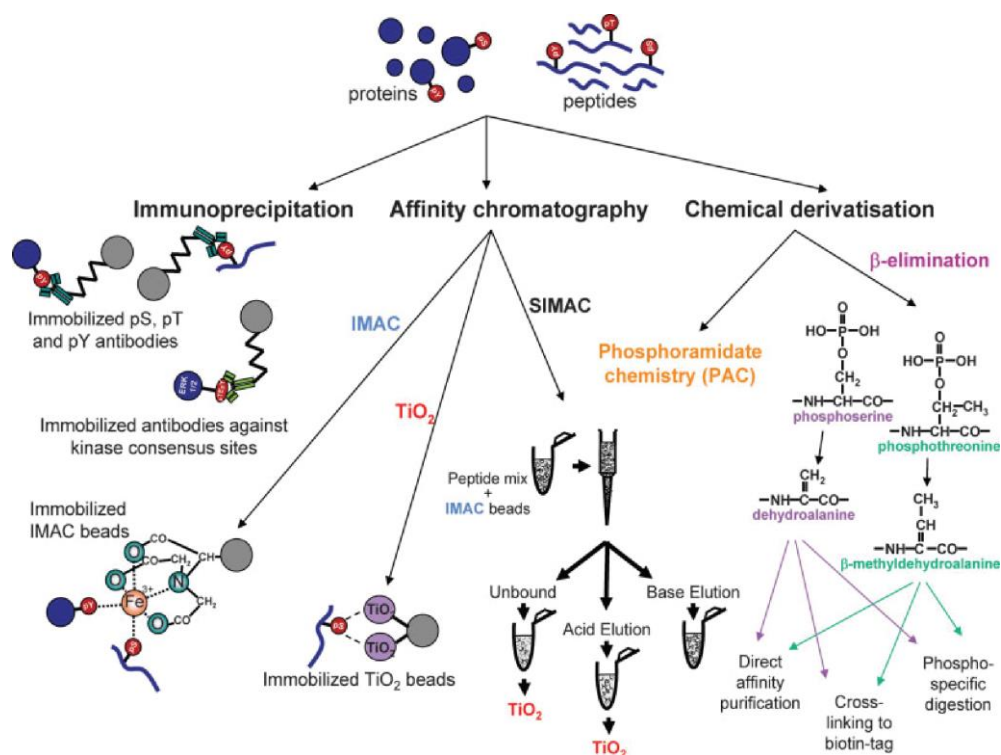


Figure 12: Strategies employed for specific phosphorylation enrichments – The most common amongst all are immunoprecipitation and affinity chromatography [116].

1.4.2.1. Strong Cationic Exchange (SCX)

Strong cationic exchange or SCX is employed to fractionate peptides according to their charge states. A negatively charged chromatographic matrix is used to selectively capture positively charged peptides. SCX is implemented in the presence of an acidic buffer system operating at a pH of 2.7 which helps to protonate the N-terminal amino group and the side chains of lysine and arginine residues [118]. Trypsin digested peptides contain a C-terminal arginine or lysine residue. Consequently, the non-phosphorylated peptides carry an overall net charge of +2. In contrast, the phosphorylated peptides carry a net charge of +1 due to the presence of the negatively charged phosphate group. Peptides bound to the stationary phase are eluted sequentially by increasing the ionic strength and/or pH [131]. Multiply phosphorylated peptides, however, will be present in the flow-through fraction due to a net negative charge or zero and do not bind to the SCX matrix. This fraction is generally collected separately and analyzed. SCX is frequently used in combination with TiO₂ chromatography to increase phosphorylated peptide

identifications. Additionally this method can also not distinguish between phosphorylated peptides and peptides with the same net charge such as N-acetylated peptides [132].

1.4.2.2. Hydrophilic Interaction Liquid Chromatography

Hydrophilic interaction liquid chromatography or HILIC was first introduced in 1990 [133]. Unlike SCX, HILIC fractionates peptides according to their hydrophilicity. The more hydrophilic a peptide is, the longer it is retained on the chromatographic column. Phosphorylated peptides are more hydrophilic thus eluting in the later HILIC fractions, often with the longer acidic peptides. HILIC fractions can be further enriched by employing Fe^{3+} -IMAC for example [134].

1.4.2.3. Immobilized Metal Affinity Chromatography or IMAC

Immobilized metal ion affinity chromatography (IMAC) is based on the high-affinity attraction of the negatively charged phosphate groups to the positively charged trivalent metal ions such as iron, aluminium, gallium or zirconium (**Figure 13**). This concept was first introduced in 1986 [135] wherein the authors show the binding of phosphorylated proteins onto ferric ions immobilized on iminodiacetate-agarose support and are subsequently eluted by increasing the pH. This was later extended to gallium and other metal ions that are more selective and versatile [136, 137]. The drawbacks of this enrichment technique include competitive binding of acidic peptides containing aspartate or glutamate residues to the metal-ion complexes reducing identifications of phosphorylated peptides thus calling for the need to block acidified carboxyl groups by methylesterification for example [138].

1.4.2.4. Titanium Dioxide (TiO_2) Enrichment

Titanium dioxide affinity chromatography, a type of metal oxide affinity chromatography (MOAC), is one of the most widely used enrichment techniques in phosphoproteomics. MOAC matrixes are composed of metal oxides or hydroxides eliminating the need for resin anchoring as in case of IMAC. TiO_2 spheres, first used in 1997 [139], were found to be highly stable, selective and rigid, making it ideal for efficient separation of

phosphopeptides [114]. Phosphate groups are adsorbed to TiO_2 under acidic conditions and can be eluted under alkaline conditions. As in case of IMAC, background interference due to acidic peptides occur. To circumvent this issue 2,5-dihydroxy benzoic acid (DHB) is used as a competitor [140]. DHB has a stronger affinity for TiO_2 as compared to acidic peptides but lower than the phosphorylated peptides themselves. Side effects of DHB include co-elution with phosphorylated peptides during liquid chromatography, reducing identifications. Aliphatic hydroxyl acids such as lactic acid which are more hydrophilic can also be used as an alternative to DHB [141] (**Figure 13**).

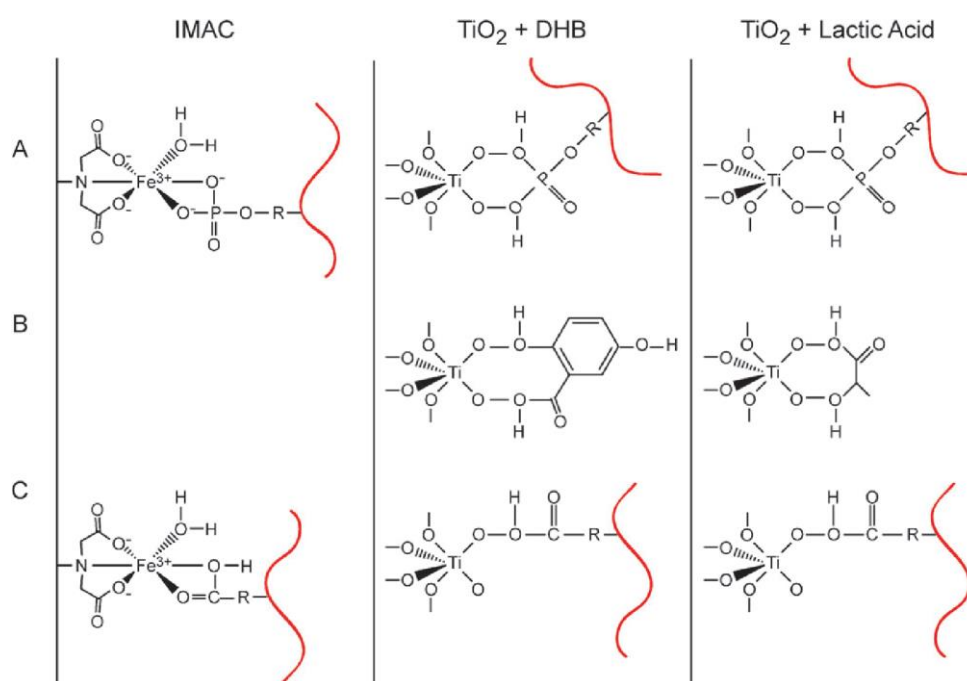


Figure 13 depicts IMAC and MOAC enrichment – (A) The co-ordination of IMAC and TiO_2 resin with a phosphorylated peptide is shown, (B) The co-ordination of TiO_2 resin with DHB (middle) and lactic acid (right) is shown, (C) The co-ordination of IMAC and TiO_2 with a peptide containing an acidic residue (carboxyl group) is shown. Peptides are represented by the red line [117].

1.4.2.5. Immunopurification

Immunoprecipitation (IP) is based on the principle that antibodies bind to various epitopes. IPs can be used for enriching phosphorylated proteins or phosphorylated peptides by employing antibodies that are raised against specific phosphorylated residues, the most common amongst all being phosphorylated tyrosine [142-144]. Many commercially available antibodies against phosphorylated serine/threonine residues are

also available though they do not bind to all pSer/pThr residues with the same affinity as they have been targeted against the phospho-amino acid within the context of its surrounding residues [132], thus giving unsatisfactory results. IPs are unsuitable for large scale phosphoproteomics studies as they are targeted only against specific phospho-amino acids. Nonetheless, they are frequently used while targeting one single type of amino acid like phosphorylated tyrosine [120].

1.5. Interaction Proteomics

Cells use protein interaction domains and the cross-talk between them to largely regulate cellular processes resulting in altered phenotypes. The conserved binding pockets within the modular protein domains are capable of distinguishing between different motifs on peptides with the same type of modifications [145]. Various PTM-dependent interactions such as inducible interactions (e.g. binding of a phosphorylated Tyr residue to the SH2 domain); sequential interactions (e.g. binding of the SH2 domain of E3 ubiquitin ligase to a pTyr residue is required for the transfer of ubiquitin to the target protein), mutually exclusive interactions (e.g. trimethylated histone H3 binds to the chromodomain of heterochromatin protein whereas acetylated histone H3 binds to the bromodomain of TATA-binding protein); antagonist interactions (e.g. binding of the trimethylated histone H3 to the chromodomain of heterochromatin protein is antagonized by phosphorylation of a neighbouring Ser residue) exist that help control subcellular organization [145] (**Figure 14**). Existence of structured mechanisms that enable extracellular or intracellular signals to be relayed to target molecules is what makes signal transduction a precise and well-coordinated process.

Post-translational modifications such as phosphorylation of hydroxyamino acids such as serine, threonine and tyrosine act as important regulatory platforms that initiate activation of signaling pathways and assemblage of protein complexes resulting in regulated gene expression. Crucial protein-protein interactions are most often mediated by “short unstructured sequences” via binding motifs or domains [146]. For example, tyrosyl-phosphorylated peptides like those of RTK's or EGFR's frequently bind to proteins containing the Src homology domain 2 (SH2) [147, 148] or the phosphotyrosyl

binding domain via systematic processes [149, 150] that result in endocytosis or as in the case of peptides with proline rich motifs that bind to the Src homology domain 3 (SH3) regulating the cytoskeletal proteins or other Src kinases [151, 152].

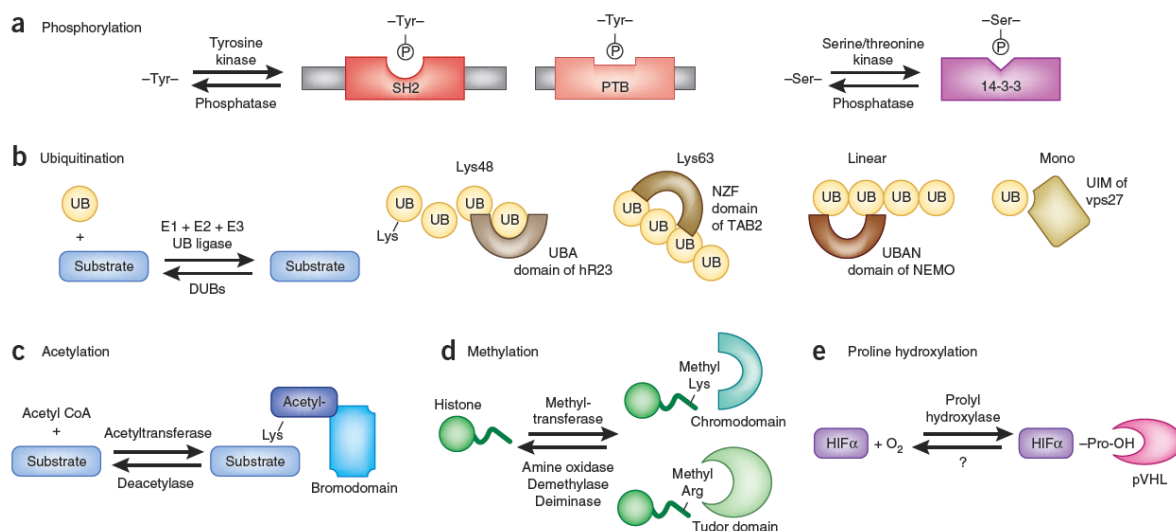


Figure 14: Types of interactions – The figure depicts the different types of regulatory properties of post-translational modifications and their interaction domains [153].

Similarly, binding of the acetylated motifs of histone H3 to the bromodomain of TATA-binding protein associated factor II250 results in the regulation of chromatin structure and gene expression [154]. Knowledge of protein interactions and domain motifs is crucial to gain an insight into the diversity and functionality of the cellular proteome. Protein kinases (S, T or Y) act as catalytic domains that transfer the phosphate group to the target protein, the protein phosphatases act as catalytic domains that remove the phosphate groups making it a reversible process, while SH2 or SH3 domains act as recognition domains that are capable of recognizing and recruiting other macromolecules to the phosphorylation modifications, eliciting an appropriate response. These modules help in information processing at the biological level by operating as a functional triad, namely “writer”, “eraser” and “reader” respectively (**Figure 15**). Combination of these three main modules result in diverse signaling responses [155]. Emergence of high-throughput techniques such yeast two-hybrid [156], phage display

[157], array technologies [158], peptide libraries [159] aid in the identification of physical interactions between proteins and/or the respective domains.

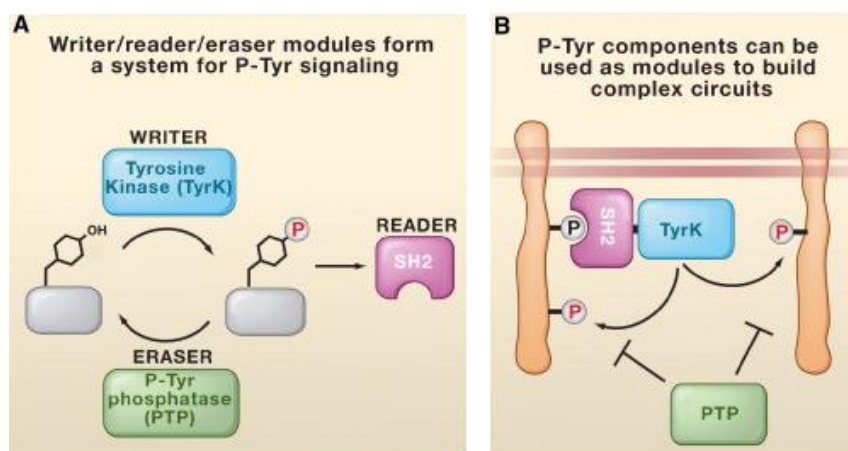


Figure 15: Domains involved during signal transduction – The figure represents (A) the “Writer”, “Reader”, “Eraser” Toolkit in phospho-tyrosine signaling, involving the tyrosine kinase (TyrK), the Src Homology 2 (SH2), and the phospho-tyrosine phosphatase (PTP) domains which form independent signaling platforms. (B) Components of phospho-tyrosine signaling can be used to build complex circuits. For example, recruitment of an SH2-TyrK protein to an initiating pTyr site can lead to amplification of tyrosine phosphorylation through a positive feedback loop [160].

Amongst these, using peptides as baits for affinity chromatography’s is a straightforward and an unbiased mode of determining binders of modified residues or interacting partners. The disadvantages include inability to capture certain transient interactions as the peptides/domains are present outside the context of the whole functional protein and repression of the “true” low abundant interactors due to non-specific interactions [146]. However, mass spectrometry based proteomics can be used as a powerful tool to overcome some of these issues. Using quantitative proteomics methodologies, such as SILAC, can help differentiate between true interactors and non-specific background binders [161].

The concept of affinity pull-downs using synthetic peptides as baits in order to study protein-protein interactions has been established and functional in eukaryotic systems for a decade now (**Figure 16**). In 2003, using a tyrosine phosphorylated peptide of the EGFR, the authors observed binding of Grb2 to EGFR via its SH2 domain and subsequent binding of Sos to Grb2 via the SH3 domain, thus validating their methodology [146]. They

also successfully identified two novel interactors, Pacsin3 and Snx9, of the proline rich motifs present on the Sos peptide, thus further regulating the growth-factor signaling cascade. Similarly, using SILAC in combination with affinity based enrichment, the authors show that the transcription factor TFIID binds directly to the trimethylated histone H3 at lysine 4 via the plant homeodomain (PHD) finger of TAF3. Loss of trimethylation reduces TFIID binding and overall transcription [162]. While the importance of STY phosphorylation events in signaling has been recognized in bacterial systems, the existence of such domains that play a crucial role in recognition, activation and recruitment has not yet been identified. In 1999, James. W and Arthur. L suggest that the P60 invasion protein from *Listeria grayi* contains an SH3 domain on the basis of sequence similarity, structural compatibility and function [163]. Around the same time, Bilwes. A and co-workers reported the presence of two SH3-like domains in the crystal structure of the regulatory domain of *Thermotoga maritima* histidine kinase CheA [164]. Other than these two examples no similar domain motifs with respect to phosphorylation dependent signaling are known to be present in prokaryotic systems till date.

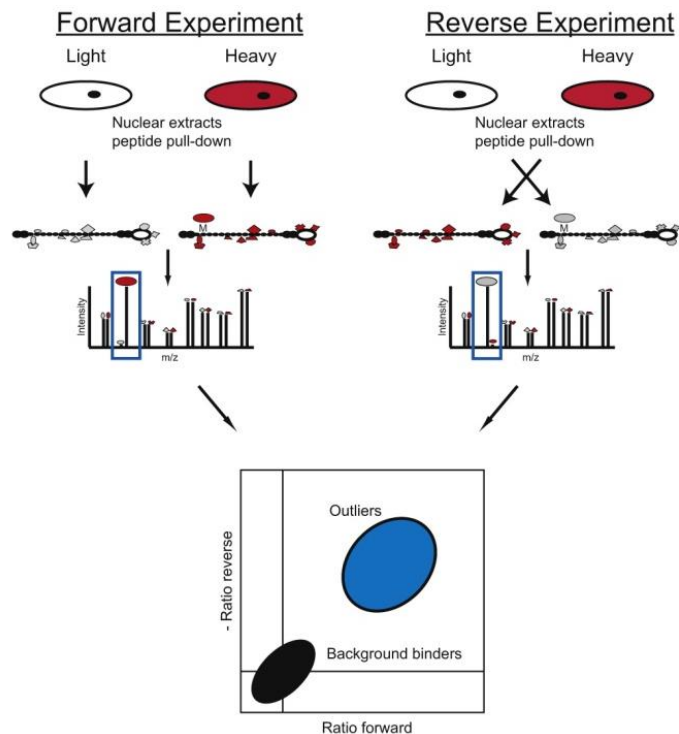


Figure 16: Affinity chromatography – Schematic representation of the experimental approach employed to find binders of methylated histone residues [161].

2. Aim of Thesis

The main aim of this research was to identify, characterize and map phosphorylation events dependent on specific serine, threonine and tyrosine (STY) kinases and phosphatases in the Gram positive model bacterium, *Bacillus subtilis*. A systematic approach, as defined below, was employed in order to address the above described goal. Specific objectives were:

1. SILAC measurement of proteome and phosphoproteome dynamics during growth of *B. subtilis* in minimal medium under laboratory conditions in order to obtain information on the abundance of specific STY kinases and phosphatases and their known substrates during each stage of growth.
2. Development and application of a SILAC-based quantitative proteomics screen for identification of novel potential substrates of specific STY kinases and phosphatases in bacteria. This was achieved by:
 - a) Systematic detection and quantification of regulated phosphorylation events upon kinase or phosphatase knock-out. Five kinases and three phosphatases were investigated.
 - b) Validation of selected regulated phosphorylation sites by biochemical analysis and interpretation of the physiological effect of phosphorylation on selected proteins.
3. Generation of the first network of bacterial kinases, phosphatases and their substrates based on the experimental data obtained in this study.
4. Generation of a comprehensive qualitative map of the *B. subtilis* phosphoproteome based on all detected phosphorylation events in order to obtain further insights into main phosphoproteome features, such as its size, functional distribution and existence of kinase target motifs on detected phosphorylated proteins.
5. Detection of proteins that bind to phosphorylated STY motifs using synthetic peptides as baits in a SILAC format and detection of putative domains that bind specifically to STY phosphorylated residues.

3. Materials and Methods

3.1. Materials

3.1.1. List of Chemicals and Consumables

Note: The composition of all buffers are described along with the respective protocols.

Name	Company	Name	Company
1 kb Generuler	Fermentas	Ex Taq DNA Polymerase	Takara Bio Inc.
2,5-dihydrobenzoic acid	Sigma-Aldrich	Ferrous Sulphate	Merck KGaA
3-hydroxypropionate	Tokyo Chemical Industry UK Ltd.	Formic Acid	Merck KGaA
Acetic Acid	Sigma-Aldrich	Gene Pulser Cuvettes	Bio-Rad
Acetonitrile	Merck KGaA	Glucose	AppliChem
Adenosine triphosphate	Sigma-Aldrich	Glutamic Acid	Merck KGaA
Agarose (Type-D5)	Euromedex	Glycerol	Merck KGaA
Ammonium Bicarbonate	Merck KGaA	Glycerol-2-Phosphate	Sigma-Aldrich
Ammonium Sulphate	Merck KGaA	Green Go Taq Buffer (5x)	Promega
Ampicillin Sodium Salt	Sigma-Aldrich	Go Taq Polymerase	Promega
<i>Bacillus subtilis</i> 168	Kind gift from Prof. Mijakovic	HEPES (potassium salt)	Sigma-Aldrich
<i>Bacillus subtilis</i> BS23	Kind gift from Prof. Mijakovic	Immobiline Drystrips	GE Healthcare
Bovine Serum Albumin	Sigma-Aldrich	Iodoacetamide	Sigma-Aldrich
Bradford Reagent	Bio-Rad Laboratories	Isopropanol	Merck KGaA
C8/C18 Extraction Discs	Empore™	IPTG	Euromedex
Calcium Chloride	Merck KGaA	Lactate Dehydrogenase	Sigma-Aldrich
Chloroform	VWR	Loading Dye (DNA)	Fermentas
Colloidal Staining Kit	Invitrogen	Lysogeny Broth	Sigma-Aldrich
Dithiothreitol	Merck KGaA	Lys-C	Wako
dNTP(s)	NEB	Lysine (Lys0)	Sigma-Aldrich
<i>E. coli</i> NM522	Kind gift from Prof. Mijakovic	Lysine (Lys4, Lys8)	Euriso-Top
<i>E. coli</i> M15(pREP4)	Kind gift from Prof. Mijakovic	Lysozyme	Sigma-Aldrich
Ethanol	Merck KGaA	Magnesium Sulphate	Merck KGaA
Ethidium Bromide	Euromedex	Manganese Sulphate	Merck KGaA

Name	Company	Name	Company
EDTA	Sigma-Aldrich	Methanol	Merck KGaA
Ex Taq Buffer (10x)	Takara Bio Inc.	NAD(P)H	Sigma-Aldrich
NuPAGE® MOPS SDS Running Buffer	Life Technologies	Streptavidin Sepharose Beads	GE Healthcare
PD – 10 Columns	GE Healthcare	Sodium Orthovanadate	Sigma-Aldrich
Phosphate Buffered Saline	PAN Biotech	T ₄ Buffer	Promega
PTMScan® Phospho-Tyrosine Antibody	Cell Signaling	T ₄ DNA Ligase	Promega
Phusion HF Buffer (5x)	NEB	Thiourea	Merck KGaA
Phusion HF DNA Polymerase	NEB	Titanium dioxide Beads	Sachtopore
Pierce Classic IP Kit	Life Technologies	Triethylammonium Bicarbonate	Sigma-Aldrich
Potassium Chloride	Merck KGaA	Trifluoroacetic Acid	Merck Millipore
Potassium Hydrogen Phosphate	Merck KGaA	Tris Base	Sigma-Aldrich
Protease Inhibitor Cocktail (tablets)	Roche	Tryptophan	Sigma-Aldrich
Protein Concentrators	Pierce	Urea	Merck KGaA
pQE30-Xa	Qiagen	Whatman™	GE Healthcare
QIAprep® Spin Miniprep Kit	Qiagen	Wizard® Genomic DNA Purification Kit	Promega
ReproSil-Pur C18-AQ 3 µm resin	Dr. Maisch GmbH	Wizard® SV Gel and PCR Clean-Up System (kit)	Promega
Resource S Column	GE Healthcare	Y-PER Reagent	Thermo Scientific
Restriction Buffers & Enzymes	NEB	Zinc Chloride	Merck KGaA
SeeBlue® Plus2 Prestained Standard	Life Technologies	β-mercaptoethanol	BDH
Sep-Pak C18 Cartridges	Waters		

3.1.2. List of Instruments

Instrument	Company	Instrument	Company
3100 Offgel Fractionator	Agilent Technologies	Gene Pulser II	Bio-Rad
ÄKTA System	GE Healthcare	LTQ-Orbitrap Elite	Thermo Fisher Scientific
Block Thermostat	Kleinfeld Labortechnik	LTQ-Orbitrap XL	Thermo Fisher Scientific
Borosilicate Emitters	Thermo Scientific	Nano ES Ion Source (ES380)	Thermo Scientific
Centrifuge (table-top)	Eppendorf	PCR thermal cycler ¹	Eppendorf
CL-1000 Ultraviolet Crosslinker	UVP	PhosphorImager FLA-7000	Fujifilm Life Science
Column Oven	Sonation	Sonicator	Branson Ultrasonics Corporation
Easy-LC nano-HPLC	Proxeon Biosystems	Sorvall RC 5C	Thermo Scientific
Electrophoresis unit (DNA)	Mupid-One	Spectrophotometer (A ₅₉₅)	Thermo Electron Corporation
Electrophoresis unit (Protein)	Biorad	Spectrophotometer (OD ₆₀₀)	VWR
Electrospray ionization source	Proxeon Biosystems	Stage-tip Centrifuge	Sonation
FLUOstar Omega	BMG Labtech	Vacuum Centrifuge	Eppendorf

3.2. Methods

Some of the methods described in this section are published in “Ravikumar V, et al, Quantitative phosphoproteome analysis of *Bacillus subtilis* reveals novel substrates of the kinase PrkC and phosphatase PrpC. *Mol Cell Proteomics*, 2014. **13**(8): p. 1965-78”.

The *Bacillus subtilis* 168 gene knock-out strains used in this study were generated in Prof. Ivan Mijakovic’s laboratory at AgroParisTech, INRA, France. Markerless gene truncation was done using *pG⁺host8* system [165, 166]. Proteomic experiments described below were carried out with the following generated strains: $\Delta lysA$, $\Delta lysA.\Delta prkC$, $\Delta lysA.\Delta prpC$, $\Delta lysA.\Delta ptkA$, $\Delta lysA.\Delta ptpZ$, $\Delta lysA.\Delta yabT$, $\Delta lysA.\Delta ybdM$, $\Delta lysA.\Delta yveL$, $\Delta lysA.\Delta yfkj$ and $\Delta lysA.\Delta ywIE$. More information can be found in [43].

3.2.1. Polymerase Chain Reaction

Polymerase chain reaction (PCR) was used to amplify the gene(s) of interest. Standard protocols were employed for performing PCR reactions using the *B. subtilis* 168 common laboratory strain, as briefly described below (**Table 1**). Primers were ordered from Eurogentec (<http://www.eurogentec.com/products/custom-oligonucleotides.html>).

Table 1: PCR Reaction Components

Components	Volume for 50 μ L reaction	Final concentration
5x Phusion HF Buffer	10 μ L	1x
100 μ M Forward Primer	0.3 μ L	~ 0.6 μ M
100 μ M Reverse Primer	0.3 μ L	~ 0.6 μ M
10 mM dNTPs	1 μ L	0.05 mM
Template DNA (<i>B. subtilis</i> 168)	1 – 3 μ L	~ 50 - 150 ng
Phusion HF DNA polymerase (2000 units/mL)	0.5 μ L	1 unit
Nuclease-free water	Volume made up to 50 μ L	

The above listed components were mixed on ice and transferred to a thermal cycler. The following thermo-cycling conditions were employed (**Table 2**).

Table 2: PCR Reaction Conditions

Step		Temperature	Time
Initial Denaturation		98 °C	3 min
30 cycles:	Denaturation	98 °C	10 - 20 s
	Annealing	*55 – 65 °C	30 – 45 s
	Extension	72 °C	30 s/kb
Final Extension		72 °C	3 min
Hold		4 °C	∞

*Annealing temperatures (T_a) used for the PCR reaction depends on the melting temperature (T_m) of the designed primer and can be calculated as: $T_a = T_m_lower \pm 5\text{ }^\circ\text{C}$ or from <https://www.neb.com/tools-and-resources/interactive-tools/tm-calculator>.

The genes amplified, along with their primer sequences, are listed in **Table 3**.

3.2.2. Gel Electrophoresis

PCR amplified products were checked by gel electrophoresis. A 1 % agarose gel was cast mixed with 1 μL of ethidium bromide (EtBr). 2 μL of the PCR product was mixed with 1 μL of 6x loading dye and loaded on the gel along with 4 μL of the standard 1 kb Generuler. The gel was run in a 0.5 M TAE buffer (50X TAE: 2 M Tris acetate, 0.05 M EDTA pH 8.2 – 8.4 at 25 °C) at 100 volts for approximately 20 – 30 min (or until the bands of the DNA ladder are seen to be clearly separated). The images were recorded using a BioCapt gel documentation system.

Table 3: List of Primers Used for this Study

Gene	Primer Name	Primer Sequence	Restriction Site	Ref.
<i>prkC</i>	prkC_fwd	5' GAAGATCTATGCTAATCGGCAAG CGGATCAGCGGGCG 3'	<i>BglII</i>	This study
	prkC_trunc_rvs	5' AAAACTGCAGTTACAAAACCCAC GGCCACTTTTTTCTTTTTGCCG 3'	<i>PstI</i>	
<i>prpC</i>	prpC_fwd	5' CGCGGATCCATGTTAACAGC CTTAAAAACAGA 3'	<i>BamHI</i>	This study
	prpC_rvs	5' AAAACTGCAGTTAGCACT GATCTTCACCCT 3'	<i>PstI</i>	
<i>ykwC</i>	ykwC_fwd	5' CGCGGATCCATGTTGAAA AAAACAATTGG 3'	<i>BamHI</i>	This study
	ykwC_rvs	5' AAAACTGCAGTTATTT CACCCAAAGC 3'	<i>PstI</i>	
<i>ykwC_S281D</i>	ykwC_S281D_r vs	5' AAAACTGCAGTTATTTTCACC CAAAGCTTATATATATCCTGTG TTCCGCTGTTTTCTTCACC 3'	<i>PstI</i>	This study
<i>ptkA</i>	ptkA_fwd	5' GTGGGATCCATGGCGCTT AGAAAAACAG 3'	<i>BamHI</i>	[167]
	ptkA_rvs	5' GGCACGCTGCAGGTT ATTTATTTTTGC 3'	<i>PstI</i>	
<i>ptpZ</i>	ptpZ_fwd	5' CGCGGATCCATGATCGATATTCA CTGTCAATTCTTCCC 3'	<i>BamHI</i>	[167]
	ptpZ_rvs	5' AACTGCAGTCAAAAAAACCAAA CAATTTTCTTCTTTGACCGG 3'	<i>PstI</i>	
<i>dnaK</i>	dnaK_fwd	5' CGCGGATCCATGAGTAAAGTT ATCGGAATCGACTTAGGAACAA CAAACATG 3'	<i>BamHI</i>	This study
	dnaK_rev	5' ATTCCCGGGTTATTTTTTGT TTTGGTCGTCGTTTACTTCT TCGTATTCAGCGTCGAC 3'	<i>XmaI</i>	
<i>dnaK_Y601E</i>	dnaK_601E_rev	5' ATTCCCGGGTTATTTTTTGT TTGGTCGTCGTTTACTTCTTC TTCTTCAGCGTCGAC 3'	<i>XmaI</i>	This study
<i>dnaK_Y601F</i>	dnaK_601F_rev	5' ATTCCCGGGTTATTTTTT GTTTTGGTCGTCGTTTACTTCT TTCGAATTCAGCGTCGAC 3'	<i>XmaI</i>	This study

3.2.3. DNA Purification

The PCR products or DNA fragment was then purified using the Wizard® SV gel and PCR clean-up kit as briefly described below:

1. The DNA band of interest was excised from the gel following electrophoresis and transferred to a 1.5 mL Eppendorf tube.

2. 10 μL of the given Membrane Binding Solution was added to the tube for every 10 mg of the gel slice. It was vortexed and incubated at 65 $^{\circ}\text{C}$ until the gel slice had completely dissolved.
3. An SV Minicolumn was inserted into a Collection Tube and the dissolved gel mixture was transferred to the Minicolumn. It was then incubated on the column for 1 min at room temperature (RT).
4. Next centrifugation was carried out at 16,000 $\times g$ for 1 min. The flow-through from the Collection Tube was discarded.
5. 700 μL of the Membrane Wash Solution (pre-mixed with ethanol according to the manufacturer's instructions) was added to the column and spun for 1 min at 16,000 $\times g$. The flow-through was discarded again.
6. Step 5 was repeated with 500 μL of the Membrane Wash Solution and centrifuged for 5 min at 16,000 $\times g$.
7. The Collection Tube was emptied and the Minicolumn assembly was centrifuged for 1 min at 16,000 $\times g$ with the microcentrifuge lid open to allow for evaporation of any residual ethanol.
8. The Minicolumn was next transferred to a fresh 1.5 mL Eppendorf tube and 50 μL of nuclease-free water was added to the column.
9. It was incubated for 1 min at RT followed by centrifugation at 16,000 $\times g$ for 1 min.
10. The Minicolumn was discarded and the purified DNA was stored at - 20 $^{\circ}\text{C}$ until required.
11. Alternatively, the amplified DNA post PCR reaction can be directly purified by adding an equal volume of the Membrane Binding Solution to the tube and continued as for the gel slice.

3.2.4. Restriction Digestion

pQE30-Xa plasmid was digested in a two-step procedure, each at 37 $^{\circ}\text{C}$ for 1 h, with the respective buffers and restriction enzymes in a 100 μL reaction. The restriction enzyme(s) used for each gene is specified along with the respective primer sequence information in **Table 3**. At the end of digestion 1 μL of shrimp alkaline phosphatase (SAP)

was added and the reaction mixture was incubated for an additional half hour at 37 °C to dephosphorylate the 5'/3' ends of the DNA to prevent religation of linearized plasmid. The digested plasmid was next subjected to clean-up or purification as described above in 3.2.3.

A similar procedure was employed for the digestion of the PCR purified DNA fragments (done in a single step) with the respective restriction enzymes, followed by purification. The digested plasmid and DNA fragments were eluted in 25 µL nuclease-free water.

3.2.5. Ligation

The digested and purified vector/plasmid was ligated to the PCR digested and purified DNA fragments at 4 °C overnight along with a control reaction mixture minus DNA (Table 4).

Table 4: Ligation Reaction Mix

Components Volume for	15 µL reaction
pQE30-Xa	1 µL
10x T ₄ buffer	1 µL
T ₄ DNA ligase (1 – 3 units/µL)	0.75 µL
DNA	8 µL (7 - 8 times more than plasmid)
Milli-Q	Volume made up to 15 µL

3.2.6. Electro Competent Cell Preparation

An overnight pre-culture of *Escherichia coli* strain NM522 (genotype: F- *proA*⁺*B*⁺, *lacI*^q, $\Delta(lacZ)M15/\Delta(lac-proAB)$, *glnV*, *thi-1*, $\Delta(hsdS-mcrB)5$) and M15 (pREP4) (genotype: F-, $\Phi80\Delta lacM15$, *thi-*, *lac-*, *mtl-*, *recA*⁺, *KmR*) was inoculated from a glycerol stock (3 µL) into 10 mL Lysogeny broth (LB) and incubated at 37 °C at 200 rpm. The pre-inoculum was used the next day to inoculate 1 L of LB with a 1:100 dilution. Cells were grown at 37 °C at 200 rpm until at OD₆₀₀ of 0.45. The cell culture was then cooled on ice for 15 min. The cells were then spun down in a Sorvall centrifuge for 15 min at 5000 x *g* at 4 °C. The supernatant media was discarded and the cells were washed twice with 500 mL sterile

pre-cooled Milli-Q, and spun each time at 6000 x *g* for 15 min at 4 °C. Cells were next washed with 50 mL sterile ice cold 10 % glycerol and spun for 15 min at 7000 x *g* at 4 °C. The cells were finally resuspended in 2 mL sterile and ice cold 10 % glycerol and 50 µL was aliquoted into 1.5 mL Eppendorf tubes and stored at – 80 °C until required.

3.2.7. Transformation

Bacterial transformation was done by electroporation. Cuvettes with a gap width of 0.2 cm were sterilized in an ultraviolet crosslinker chamber for 5 min at 120,000 microjoules prior to use. 50 µL of the competent cells was thawed and transferred to the cuvette on ice. 1 µL of the plasmid DNA was transferred to the cuvette on ice and mixed well. The set up was incubated on ice for 15 min. Electroporation was carried out using a Gene Pulser II at a voltage of 2.5 volts, 25 µF capacitance and the resistance was set at a range of 200 – 500 Ω. The sample was pulsed for a second and 1 mL of pre-warmed LB media was added immediately to the cuvette, mixed well and incubated at 37 °C for 1 h. After 1 h, 50 µL of the sample was plated on Ampicillin (Amp) plates (concentration: 100 mg/L) for the selection of positive transformants and the plates were incubated overnight at 37 °C.

3.2.8. Plasmid Isolation

QIAprep® Spin Miniprep Kit was used along with the buffers provided, for the purpose of plasmid isolation according to the manufacturer's instructions, as briefly described below:

1. A single colony from the transformed plate was inoculated into 5 mL of LB media containing ampicillin (final concentration: 100 mg/L). The culture was grown overnight at 37 °C at 200 rpm.
2. 2 mL of the overnight grown culture was pelleted at 8000 rpm for 3 min at RT.
3. The supernatant was discarded and the bacterial cells were resuspended in 250 µL of Buffer P1 and transferred to a 1.5 mL Eppendorf tube.

Note: LyseBlue® reagent at a ratio of 1:1000 and the RNase A solution provided was added to Buffer P1 before use, according to the manufacturer's instructions.

4. 250 μ L of Buffer P2 was added to the Eppendorf and mixed thoroughly by inverting the tube 4-6 times until the solution turns blue. Care should be taken not to allow this lysis reaction to proceed more than 5 min.
5. 350 μ L of Buffer N3 was added next and mixed thoroughly by inverting the tube as before until the solution turns colourless.
6. Centrifugation was carried out for 10 min at 13000 rpm.
7. The supernatant from step 6 was transferred into the QIAprep spin column provided by pipetting carefully.
8. It was centrifuged for 60 s at 13000 rpm and the flow through was discarded.
9. The QIAprep spin column was washed with 500 μ L Buffer PB followed by centrifugation for 60 s at 13000 rpm.
10. The flow through was discarded and the QIAprep spin column was next washed with 750 μ L of Buffer PE.

Note: 100 % ethanol was added to Buffer PE before use.

11. It was centrifuged for 60 s at 13000 rpm and the flow through was discarded. An additional centrifugation for 1 min was done to remove any residual wash buffer.
12. The QIAprep spin column was transferred to a fresh 1.5 mL Eppendorf tube and 50 μ L of Buffer EB (10 mM Tris-HCl pH 8.5) or autoclaved Milli-Q water was added to the spin column and let to stand at RT for 1 min, followed by centrifugation at 13000 rpm for 1 min.
13. The eluate was stored at -20 °C and used as and when required.

3.2.9. Colony PCR

Colony PCR was carried out to test for positive transformants, as outlined below. 5 – 10 single colonies were selected from the transformed plates and a part of colony was streaked onto a new LB Amp plate and incubated at 37 °C overnight and a part of the colony was mixed with the PCR reaction mix (**Table 5**).

All reaction components were mixed on ice. The thermo cycling conditions employed are detailed in **Table 6**.

Table 5: Colony PCR Reaction Mix

Components	Volume for 50 μ L reaction	Final concentration
5x Green Go Taq Buffer	10 μ L	1x
100 μ M Forward Primer	0.5 μ L	~ 1 μ M
100 μ M Reverse Primer	0.5 μ L	~ 1 μ M
10 mM dNTPs	1 μ L	0.05 mM
Go Taq Polymerase (5 units/ μ L)	0.5 μ L	2.5 units
Nuclease-free water	Volume made up to 50 μ L	

Table 6: Colony PCR Reaction Conditions

Step	Temperature	Time
Initial Denaturation	95 $^{\circ}$ C	6 min
30 cycles:	Denaturation	95 $^{\circ}$ C
	Annealing	*55 – 65 $^{\circ}$ C
	Extension	72 $^{\circ}$ C
Final Extension	72 $^{\circ}$ C	2 - 3 min
Hold	4 $^{\circ}$ C	∞

*Annealing temperatures used for the PCR reaction depends on the T_m of the designed primer and can be calculated as: $T_a = T_m_lower \pm 5 \text{ }^{\circ}\text{C}$ or from <https://www.neb.com/tools-and-resources/interactive-tools/tm-calculator>

3.2.10. Sequencing

50 – 100 ng of DNA corresponding to the positive transformants were sequenced at GATC Biotech (<http://www.gatc-biotech.com/en/index.html>). Positive clones were taken forward for protein purification and subsequent analysis as described further.

3.2.11. SILAC Labeling of Bacterial Cells

B. subtilis 168 Δ lysA (strain auxotrophic for lysine), referred to as wild-type (WT) henceforth, was streaked from a glycerol stock onto a LB-agar plate and grown in a 37 °C incubator for 10-12 h. All experiments were carried out in the minimal medium containing 15 mM (NH₄)₂SO₄, 2 mM CaCl₂, 1 μM FeSO₄·7H₂O, 8 mM MgSO₄, 10 μM MnSO₄, 27 mM KCl, 0.6 mM KH₂PO₄, 7 mM C₆H₅Na₃O₇·2H₂O, 50 mM Tris-HCl pH 7.5 supplemented with 0.5 % glucose, 0.67 mM glutamic acid and 490 μM tryptophan. SILAC labeling was done by supplementing the medium with 0.025 % of the respective isotopically labeled L-lysine – Light, Lys0: ¹²C₆ ¹⁴N₂; Medium, Lys4: 4,4,5,6-D₄ or Heavy, Lys8: ¹³C₆ ¹⁵N₂. A pre-inoculum in respective minimal media (based on the stage of growth as described below) was then made using a single colony. Cells were grown at 37 °C at 200 rpm and harvested at different stages of growth. The pre-inoculum was grown to an OD₆₀₀ of 0.6 and was then used to inoculate a fresh 200 mL media with a 1:20 dilution. This was grown to an OD corresponding to their respective stages namely, T1 - Exponential Growth Phase (OD₆₀₀ = 0.07); T2 - Entry into Retardation Phase (OD₆₀₀ = 0.6); T3 - Exit from Retardation Phase (OD₆₀₀ = 0.9); T4 - Early Stationary Phase (OD₆₀₀ = 1.1); T5 - Late Stationary Phase (OD₆₀₀ = 1.3). The nomenclature of the growth phases is based on the terms stated by Monod [168]. Two triple SILAC experiments (one with T1, T2 and T3; other with T2, T4 and T5) were conducted, with T2 phase being the common point between both experiments. Cultures grown to T1 and T4 stages were labeled with “Light” lysine, T2 with “Medium” and T3 and T5 with “Heavy” lysine. Cells were harvested by centrifugation at 7000 x g for 10 min. They were flash frozen in liquid nitrogen and stored at -80 °C.

3.2.12. Protein Extraction

Cells were first thawed on ice. Cell lysis was done by either of the two methods:
(1) Cell pellets were resuspended in 1 mL Y-PER Reagent supplemented with 50 μg/mL of lysozyme, 5 mM of phosphatase inhibitors - sodium fluoride (SF) and glycerol-2-phosphate (GP) and protease inhibitor cocktail. Cells were lysed by incubation at 37 °C for 20 min.

(2) Cell pellets were resuspended in a SDS lysis buffer containing 4 % SDS in 100 mM triethylammonium bicarbonate (TEAB) pH 8.6, 5 mM GP, 5 mM SF, 5 mM sodium orthovanadate (SOV), 10 mM ethylenediaminetetraacetic acid (EDTA) and protease inhibitor cocktail. The cell extract was then boiled at 90 °C for 10 min.

Method 3.2.12.(1) or 3.2.12.(2) was then followed by sonication for 30 s at 40 % amplitude in order to degrade DNA. The cell debris was removed by centrifugation at 13.4 rpm for 30 min. The crude protein extract was cleaned up by chloroform/methanol precipitation.

3.2.13. Chloroform/Methanol Precipitation

CHCl₃/MeOH precipitation was used to concentrate and purify the proteins and get rid of contaminants, as follows:

1. Four times the volume of methanol (MeOH) was added to the crude protein extract obtained after cell lysis and vortexed.
2. One time the volume of chloroform (CHCl₃) was added next and vortexed well.
3. Three times the volume of Milli-Q water was added and vortexed well.
4. The sample was spun at 4700 x *g* for 10 min at 4 °C.
5. The aqueous upper phase was discarded carefully ensuring the white interphase containing the proteins was not disturbed.
6. Step 1 with MeOH was repeated again and the sample was centrifuged for 10 min at 4700 x *g* at 4 °C.
7. The supernatant was discarded and the protein pellet was allowed to air dry.
8. The dried protein pellet was dissolved in denaturation buffer containing 6 M urea and 2 M thiourea in 10 mM Tris-HCl pH 8.0.

3.2.14. Protein Estimation

Protein concentration was measured by Bradford protein assay [169]. A standard calibration curve was created using known concentrations of bovine serum albumin, BSA (0.2 mg/mL – 1.4 mg/mL). Standard concentration range was made by mixing 5 – 35 µL

of the BSA stock (2 mg/mL) with denaturation buffer to a total volume of 50 μ L (**Table 7**).

Table 7: Preparation of Standard Curve

Volume of BSA (stock concentration: 2 mg/mL)	Volume of Denaturation Buffer	Final Concentration
5 μ L	45 μ L	0.2 mg/mL
10 μ L	40 μ L	0.4 mg/mL
15 μ L	35 μ L	0.6 mg/mL
20 μ L	30 μ L	0.8 mg/mL
25 μ L	25 μ L	1.0 mg/mL
30 μ L	20 μ L	1.2 mg/mL
35 μ L	15 μ L	1.4 mg/mL

The Bradford dye reagent was diluted 1:5 in purified water and filtered using Whatman™ filters. A total of 7 points were taken. 20 μ L of every standard sample was then mixed with 1 mL of Bradford-solution and incubated for 5 min at RT. The absorbance was measured at a wavelength of 595 nm. The protein extracts were diluted 1:10 so as to fall within the range of the standard calibration curve and were subsequently treated same as the standards. All measurements were performed in duplicates. Concentration of the protein samples were deduced from the standard curve.

3.2.15. In-solution Digestion

10 μ g of labeled protein lysate was digested and analyzed, as described below, for incorporation checks. A total 10 μ g of protein from all labeled and unlabeled lysates were mixed in 1:1:1 ratio and digested and analyzed for mixing error.

For each SILAC experiment, protein extracts labeled with Lys0, Lys4 and Lys8 were mixed in a 1:1:1 ratio (which is highly crucial) and subjected to in-solution digestion. In-solution digestion was carried out as described previously [170] with minor modifications. The protein extract was reduced with 1 mM dithiothreitol (DTT) and

alkylated with 5.5 mM iodoacetamide (IAA) in the dark, for 1 h each at RT. Proteins were then predigested with an endoproteinase, Lys-C (1:100, w/w), for 3 h followed by overnight digestion at RT with Lys-C (1:100, w/w) after a dilution of four times the volume with deionized water. 100 µg of the resulting peptide mixture was taken for isoelectric focusing, while the rest was acidified to a pH of 2.5 with 10 % trifluoroacetic acid (TFA) and subjected to strong cationic exchange (SCX) and titanium dioxide chromatographies.

3.2.16. Protein Fractionation by Offgel Isoelectric Focusing

Peptides obtained from in-solution digestion were separated based on differences in their isoelectric point (pI) using the 3100 Offgel Fractionator. Peptides were separated into 12 fractions using 13 cm Immobiline Drystrips (pH gradient 3-10) using the default settings with a maximum current of 50 µA and potential difference of 20 kVh. Fractionated peptides were acidified using 30 % acetonitrile (ACN), 5 % acetic acid and 10 % TFA in water and further purified by using C18 stage-tips [171]. Briefly, C18 discs were activated with 100 µL MeOH and equilibrated with 200 µL solvent A* (2 % ACN, 1 % TFA). The sample was loaded onto the membrane and washed with 200 µL solvent A (0.5 % acetic acid). Peptides were eluted in 50 µL solvent B (80 % ACN, 0.5 % acetic acid), concentrated in a vacuum centrifuge to a minimum volume of 5 µL, acidified with 10 % solvent A* and subjected to nano-LC-MS/MS measurements on the LTQ-Orbitrap Elite.

3.2.17. In Gel Digestion

3.2.17.1. Electrophoresis

Prior to in-solution digestion, 100 µg of the equally mixed samples were taken for in gel digestion. Samples were separated on a NuPAGE® Bis-Tris 4-12 % gradient gel following the manufacturer's instructions. Briefly, the sample was mixed with 1/4th final volume of 4x NuPAGE® LDS sample buffer, 1/10th the final volume of β-mercaptoethanol and the volume was brought up as required with Milli-Q water. The sample was loaded on the NuPage Bis-Tris 4-12 % precast gel along with 7 µL of SeeBlue® Plus2 Prestained

Standard. Gel electrophoresis was carried out at 200V constant for 1 h in NuPAGE® MOPS SDS running buffer.

3.2.17.2. Coomassie Gel Staining

The gel was stained using the Colloidal Staining Kit according to the manufacturer's instructions. It was first treated with a fixing solution (20 mL Milli-Q water, 25 mL MeOH, 5 mL acetic acid) for 10 min at RT. Next the gel was treated for 10 min at RT with the staining solution (27.5 mL Milli-Q water, 10 mL MeOH, 10 mL Stainer A). After 10 min, 2.5 mL of Stainer B was added and the gel was further stained for 1 h at RT on a rocker. Destaining was carried out for 1 h at RT with Milli-Q water.

3.2.17.3. Digestion and Extraction

The coomassie stained gel was subsequently cut into 16 slices. Resulting gel slices were cut into 1 x 1 mm pieces and transferred to 1.5 mL Eppendorf tubes and destained by washing three times with 200 μ L of 5 mM ammonium bicarbonate (ABC) in 50 % ACN (1:1, v/v). Proteins were then dehydrated in 200 μ L 100 % ACN followed by reduction of the disulfide bonds with 50 – 150 μ L of 10 mM DTT in 20 mM ABC for 45 min at 56 °C and alkylation of cysteines with 50 – 150 μ L of 55 mM IAA in 20 mM ABC for 30 min at RT in the dark. After washing two times with 200 μ L of 5 mM ABC in 50 % ACN, the gel slices were dehydrated once again with 200 μ L of 100 % ACN for 15 min at RT followed by drying in a vacuum centrifuge for 10 min. Proteins were digested with 50 – 100 μ L of Lys-C (12.5 ng/ μ L in 20 mM ABC) at 37 °C over-night. The resulting peptide supernatant was collected in a fresh 1.5 mL Eppendorf tube and the gel pieces were subjected to elution in three steps. Extraction was carried out with 100 - 200 μ L the following solutions, and all supernatants were collected in the same Eppendorf.

Step 1 – 3 % TFA in 30 % ACN

Step 2 - 0.5 % acetic acid in 80 % ACN

Step 3 – 100 % ACN

Samples were then evaporated in a vacuum centrifuge to a final volume of \leq 100 μ L and the peptides were next desalted using stage-tips and analyzed on the LTQ-Orbitrap Elite.

3.2.18. Phosphorylated Peptide Enrichment

Phosphorylated peptides were enriched for by following one of the two methods described below:

Method (1): 11.8 mg of digested proteins (post in-solution digestion) were subjected to phosphorylated peptide enrichment in two stages. In the first stage, peptides were separated by SCX (Strong Cationic Exchange) [170]. The sample was loaded onto a 1 mL Resource S column in 5 mM KH_2PO_4 , 30 % ACN and 0.1 % TFA (buffer pH 2.7) with a flow rate of 1 mL/min. Bound peptides were eluted over a linear gradient of 0-35 % in 350 mM KCl, 5 mM KH_2PO_4 , 30 % ACN and 0.1 % TFA (buffer pH 2.7) over 30 min resulting in 16 2 mL fractions. Multiply phosphorylated peptides that do not bind to the column are retained in the flow-through that is collected separately. The 16 SCX fractions were pooled together according to estimated peptide amounts to form a total of 9 fractions on which a second stage of phosphorylated peptide enrichment was performed, along with the flow-through, using Titanium dioxide (TiO_2) chromatography [172]. TiO_2 spheres of 10 μm were pre-incubated with 2,5-dihydrobenzoic acid (DHB) in 80 % ACN (final concentration 30 mg/mL). 5 mg of TiO_2 beads were added to each of the fractions including the flow-through and incubated for 30 min at RT by end-over-end rotation. The TiO_2 enrichment was repeated twice for each of the pooled SCX fractions and five times in case of the flow-through. Incubated beads were washed once with 1 mL of 30 % ACN, 80 % TFA and a second time with 80 % ACN, 0.1 % TFA for 10 min each. The phosphorylated peptides were eluted from the beads with 3 x 100 μL of 40 % NH_4OH solution in 60 % ACN, pH >10.5. The sample volume was reduced in a vacuum centrifuge at RT, acidified to a pH of 1.5 and desalted using C18 stage-tips and injected on the LTQ-Orbitrap XL.

Method (2): 5 mg of the digest (post in-solution digestion) was acidified to a pH of 2.7 with 10 % TFA. It was incubated at RT for 10 min after which it was centrifuged at 4700 x *g* for 5 min. Solid phase extraction (SPE) was carried out next for the purpose of desalting. Sep-Pak Vac 1cc C18 cartridges were activated with 1 mL MeOH followed by equilibration with 1 mL solvent A* (2 % ACN, 1 % TFA). The acidified protein digest sample was next added to the SPE column. The column was washed with 1 mL solvent A (0.5 % acetic acid) and eluted in 1 mL of loading buffer (80 % ACN, 6 % TFA). Next

phosphorylated peptides were specifically enriched for by using TiO₂ beads. A slurry of TiO₂ beads was made in the loading buffer and the slurry was added to the desalted peptides maintaining a 1:10 bead to peptide ratio. The solution was incubated at RT for 10 min on an end-over-end rotator. This step was repeated five times. The beads from each round of enrichment were subsequently resuspended in 50 µL of loading buffer and transferred to a stage tip containing C8 material. Elution was carried out at 4 °C. The phosphorylated peptides were eluted twice with 50 µL of elution buffer (5 % ammonia in 60 % ACN, pH ≥ 11.0) directly into a fresh Eppendorf containing 5 µL of 10 % TFA. The samples were concentrated in a vacuum centrifuge to a volume of approximately below 50 µL. A final volume of 3 µL of 100 % TFA was added next and vortexed well. The samples were then stage-tipped followed by injection into the LTQ-Orbitrap Elite.

Additionally, for triple-SILAC experiments using BY-kinases or phosphatases, 20 mg of digested peptides were enriched specifically for phosphorylated tyrosines using the PTMScan® Phospho-Tyrosine antibody according to the manufacturer's instructions with slight modifications. Briefly, the digested peptides were acidified (as before) and desalted by solid phase extraction. A Sep-Pak Classic C18 cartridge was activated with 5 mL methanol followed by equilibration with 5 mL solvent A* (2% ACN/1% TFA). The sample to be desalted was next run through the column. This was followed by a wash step with 5 mL solvent A (0.5% acetic acid) and finally elution with 6 mL of 80% ACN/6% TFA. The eluate was evaporated in a vacuum centrifuge and taken for enrichment with the pTyr antibody. The antibody bead slurry was washed four times with 1x PBS (phosphate buffered saline). The eluate from the solid phase extraction procedure was mixed with the antibody slurry and incubated at 4 °C for 2 h by end-over-end rotation. The pTyr peptides bound to the antibody beads were transferred to spin columns (Sigma-Aldrich) and washed twice with IAP (immunoaffinity purification) buffer (50 mM MOPS, 10 mM sodium phosphate pH 7.2, 50 mM sodium chloride). This was followed by three washes with Milli-Q water and elution twice with 0.1% TFA. The eluate was stage-tipped as before and injected on the LTQ-Orbitrap Elite (Thermo Fisher Scientific).

3.2.19. Mass Spectrometric Analysis

3.2.19.1.1. LC-MS/MS on the LTQ-Orbitrap XL

Samples were measured on an Easy-LC nano-HPLC coupled to an LTQ-Orbitrap XL mass spectrometer (MS), as described previously [173, 174]. Chromatographic separation was done on a 15 cm PicoTip fused silica emitter with an inner diameter of 75 μm and a tip diameter of 8 μm , packed in-house with reversed-phase ReproSil-Pur C18-AQ 3 μm resin. A column oven was attached to the column in order to maintain a steady temperature of the column at 30 °C. 4 μL of peptide solution was injected onto the column with solvent A at a flow rate of 500 nL/min and 280 bars. Peptides were then eluted using a 130 min segmented gradient of 5 – 90 % solvent B, the flow rate being maintained constant at 200 nL/min. Separated peptides were ionized by electrospray ionization or ESI in the positive mode. Ion spray voltage was set at 2.35 kV. Capillary temperature was set at 200 °C. The mass spectrometer was operated on a data-dependent mode. Survey full-scans for the MS spectra were recorded in the Orbitrap mass analyzer between 300 – 2000 Thomson at a resolution of 60,000 with a target value of 10^6 charges. The five most intense peaks from the survey scans were selected for fragmentation with collision induced dissociation (CID) in the linear ion trap analyzer using helium gas with a collision energy of 35 %. A target value of 5,000 charges was set for each scan cycle. For phosphoproteomics analysis, ions were fragmented by multi stage activation or MSA with neutral loss occurring at – 97.97, – 48.98 and – 32.66. A 90 s dynamic exclusion was set. Additionally, lock masses option was enabled on the Orbitrap-XL for internal calibration [175] and to improve the mass accuracy.

3.2.19.1.2. Synthetic Peptide Analysis by Direct Infusion

Synthetic peptides synthesized in the laboratory of Dr. Hubert Kalbacher, University of Tübingen, Germany, were injected directly into the LTQ-Orbitrap XL for quality control analysis. 100 μM of lyophilized peptides dissolved in water (stock concentration 1 mg/mL) (sonicated briefly if required for complete dissolution) were mixed with 50 % MeOH and 2.5 % formic acid were loaded onto metal coated borosilicate capillaries (1 μm diameter) and injected directly into the LTQ-Orbitrap XL. The flow rate of injection was maintained at 25 nL/min. The spray voltage was set between 1.0 – 1.5 volts and the

tube lens at 80 volts. All scans were recorded in the Orbitrap mass analyzer at a maximum resolution of 10,000. The scans were recorded continuously for 1 – 5 min. Selected-ion monitoring (SIM) scan mode was enabled to detect the precursor peaks that were not detected in the full scan mode, where in the centroid mass and the width was defined accordingly. Select precursor peaks were taken for fragmentation via CID or HCD at a collision energy of 15 – 35 %. The fragmentation spectra of all analyzed synthesized peptides were annotated manually.

3.2.19.2. LC-MS/MS on the LTQ-Orbitrap Elite

A similar set up as in case for the LTQ-Orbitrap XL was configured for the LTQ-Orbitrap Elite as well [173, 174]. The LTQ-Orbitrap Elite was also coupled to an Easy-LC nano-HPLC. In-house packed 15cm columns, coupled to a column oven (temperature 30 °C), were used with an inner diameter of 75 µm, a tip diameter of 8 µm and packed with ReproSil-Pur C18-AQ 3 µm resin. 2 µL of the peptide solution was injected onto the column. Peptides were eluted using a 90 min (130 min in case of phosphorylated peptide analysis) segmented gradient of 5- 90 % solvent B at a constant flow rate of 200 nL/min. Eluting peptides were ionized by ESI. The MS was operated in the positive mode. Capillary voltage and temperature was set to 2.3 kV and 275 °C respectively. Survey full-scans for the MS spectra were recorded in the Orbitrap mass analyzer between 300 – 2000 Thomson at a resolution of 120,000 with a target value of 10^6 charges. The top 20 most intense peaks from the survey scans were selected for fragmentation with higher-energy collisional dissociation (HCD) in the HCD cell using nitrogen gas with a collision energy of 35 %. A target value of 5,000 charges was set for each scan cycle. A 60 s dynamic exclusion was enabled.

3.2.20. Data Processing

Acquired MS spectra were processed with the MaxQuant software suite [176, 177], integrated with the Andromeda search engine. Database search was performed against a target-decoy database of *B. subtilis* 168 obtained from Uniprot (taxonomy ID 1423), containing 4,195 *B. subtilis* protein entries and 248 commonly observed laboratory contaminant proteins. Endoprotease Lys-C was entered as the protease with a maximum

missed cleavage of two specified. Three isotopic forms of lysine (light, medium and heavy) were stated in the search space. Oxidation of methionines, N-terminal acetylation and phosphorylation on serine, threonine and tyrosine residues was specified as a variable modification. Initial maximum allowed mass tolerance was set to six ppm (for the survey scan) and 0.5 Da for CID fragment ions. Carbamidomethylation on cysteines was defined as a fixed modification. Re-quantify was enabled (except for mixing checks or incorporation checks). A false discovery rate of 1% was applied at the peptide, protein and phosphorylated site level individually and only fragments with a minimum length of seven amino acids were used for SILAC peptide quantification. A minimum of two unmodified peptide counts were required for the respective protein quantification. For the quantification of the modified phosphorylated site precedence was given to the least (preferably singly) modified peptide.

3.2.21. Bioinformatic Analysis

Downstream bioinformatic analysis was performed using the R version 3.0.1 [178].

3.2.21.1. Clustering Analysis

Grouping of similar (phosphorylated) protein expression profiles was achieved by hierarchical clustering analysis. Standard deviations of log₂ transformed ratios of proteins and phosphorylation sites were calculated across the five growth stages. Proteins belonging to the 4th quartile (topmost 25 %) of calculated standard deviations were taken into account for clustering analysis to ensure significant expression changes. Missing values were replaced from the normal distribution via imputation. Hierarchical clustering analysis was performed using “hclust” R-function on Z-score transformed profiles using Euclidean as a distance measure and Ward’s method for linkage analysis.

3.2.21.2. GO Enrichment Analysis

We retrieved Gene Ontology (GO) annotation of *B. subtilis* from Uniprot-GOA (downloaded on 31/03/2012). To test whether specific annotation terms are enriched or depleted within a set of proteins of interest we applied Fisher’s exact test using the

theoretical *B. subtilis* proteome as background. Derived p-values were further adjusted to address multiple hypothesis testing using the method proposed by [179].

3.2.21.3. Analysis of Phosphorylation Events

Stringent acceptance criteria were applied for modified peptides identified and quantified by the above described proteomics analysis. Phosphorylation events with a localization probability of ≥ 0.75 were considered localized on the respective S/T/Y residue. Modified peptides were also filtered for a posterior error probability (PEP) score of ≤ 0.001 . MS/MS spectra of phosphorylated peptides were manually validated for good b- and y- ion series coverage using MaxQuant Viewer (v1.2.2.9). All ratios from each SILAC experiment are relative to the Lys4 labeled T2 phase since this was the common reference point chosen. The SILAC ratios of the phosphorylation sites were further normalized with the respective protein ratios in order to eliminate a bias that could be introduced based on the changing protein abundance. Phosphorylation sites were considered as differentially regulated based on intensity-weighted “Significance B” ($p = 0.05$) [176] calculated after normalization by corresponding protein ratios.

3.2.22. Synthesis and Purification of 6x-His Tagged Proteins

The recombinant plasmids of pQE30-Xa containing the genes *prpC*, *ykwC*, *ykwC_S281D*, *ptkA*, *ptpZ*, *dnaK*, *dnaK_Y601E* and *dnaK_Y601F* respectively were used to transform *E. coli* NM522 cells and the transformants were selected on LB medium containing ampicillin. Cells were grown in LB medium, with the addition of ampicillin (100 mg/L). pQE30-Xa-*prkC* was transformed in *E. coli* M15-pREP4::groESL and subsequently grown on LB medium containing kanamycin (25 mg/L). 1 mM Isopropyl β -D-1-thiogalactopyranoside (IPTG) was used to induce expression of cloned genes. 6xHis - PrpC (30 kDa); -YkwC and -YkwC S281D (33 kDa); -PtkA (28 kDa); -PtpZ (31 kDa); -DnaK, DnaK_Y601E and -DnaK_Y601F (68 kDa) were over-expressed in *E. coli* NM522. 6xHis - PrkC (41 kDa) was over-expressed in *E. coli* M15-pREP4::groESL since it was less soluble. Protein purification was done using Ni-NTA agarose beads by the standard Qiagen protocol with slight changes. Cells were lysed by sonication in a buffer containing 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 10 % glycerol, 5 μ g/mL DNase and 2 mg/mL lysozyme

instead of the phosphate buffer system. Beads were subsequently washed in the same buffer minus DNase and lysozyme containing 30 mM imidazole and proteins were eluted in the same buffer minus DNase and lysozyme supplemented with 300 mM imidazole. Proteins were desalted on PD-10 columns with a buffer containing 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 10 % glycerol. Strep-tagged versions of proteins were obtained using a pQE-30 vector in which the His-tag was replaced by the strep-tag. All purified proteins were confirmed on a denaturing gel.

Proteins expressed for circular dichroism spectroscopy measurements were purified in a similar way but in a different buffer system containing 50 mM NaH₂PO₄ instead of Tris-HCl, 100 mM NaCl and minus glycerol.

3.2.23. *In vitro* Phosphorylation and Dephosphorylation Assay

Protein phosphorylation reactions were performed with 1 μM of the kinase (PrkC or PtkA) and 10 μM of the respective putative substrate (YkwC or DnaK). In addition to the proteins, a typical 30 μL reaction mixture contained 50 μM ATP with 20 μCi/mmol [γ -³²P]-ATP, 5 mM MgCl₂ and 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 10 % glycerol. Samples were incubated for 10, 20 and 40 min respectively at 37 °C after which the reactions were stopped by adding SDS-PAGE sample buffer and heating at 100 °C for 5 min. Proteins were separated by electrophoresis on denaturing gels (0.1 % SDS, 15 % polyacrylamide) which were subsequently treated with boiling 0.5 M HCl for 12 min. After drying the gels, radioactive bands were visualized with a PhosphorImager. Exposure time was 16 h.

For dephosphorylation experiments, [γ -³²P] ATP was first removed from the ³²P-labeled P-Ser/Thr protein samples by using concentrators, before adding 10 μM of the respective phosphatase (PrpC or PtpZ). The reaction mixtures were incubated for 1 and 2 h respectively at 37 °C before dephosphorylation was stopped by adding SDS-PAGE sample buffer and heating at 100 °C for 5 min. PAGE and autoradiography were carried out as described above.

3.2.24. Enzymatic Assays

Enzymatic assays were carried out only for YkwC and YkwC_S281D, the phosphomimetic mutant. The enzyme kinetics of purified YkwC was monitored in a 200 μ L reaction at RT [180]. The reaction mixture contained 100 mM Tris-HCl pH 8.8, 100 μ M MgCl₂, 500 μ M Nicotinamide Adenine Dinucleotide Phosphate (NADP), 50 mM 3-hydroxypropionate (3-HP) and 1 μ g of the enzyme. Measurements were taken every 10 sec for 30 min. The production of NADPH (reduced form) was measured at 340 nm. Similar measurements were made with 1 μ g concentration of YkwC_S281D.

For *in vivo* enzymatic assays, *B. subtilis* WT, Δ *prkC* and Δ *prpC* strain were grown in LB and LB supplemented with 0.3 % glucose [LB + Glucose] at 37 °C with vigorous shaking in flasks. Samples were collected during the transition phase (at OD₆₀₀ approx. 1.4). Cells were lysed by sonication in a buffer containing 50 mM Tris-HCl pH 8.5, 50 mM NaCl, 10 % glycerol and 1 mg/mL lysozyme. Cell debris was removed by centrifugation and the supernatant was desalted using PD-10 columns. The obtained crude extracts were used for the enzymatic assay as described above.

3.2.25. Circular Dichroism Spectroscopy

Circular dichroism (CD) measurements were performed at the Max-Planck Institute for Developmental Biology, Tübingen, by Kaspar Feldmeier. Circular dichroism (CD) measurements for YkwC, YkwC_S281D, DnaK, DnaK_Y601E and DnaK_Y601F were done with the help of Kaspar Feldmeier from Max-Planck Developmental Biology, Tübingen, Germany. CD measurements were made on a Jasco J-810 instrument at RT. Suprasil cells (0.1 cm path length) were used for all the measurements. Baselines were the respective buffer solutions (50 mM NaH₂PO₄ pH 7.5, 100 mM NaCl) in which the proteins were dissolved. At least five repeat scans were obtained for each sample and its respective baseline. The averaged baseline spectrum was subtracted from the averaged sample spectrum. Measurements were made in the range of 240 - 190 nm (far ultraviolet circular dichroism). The scanning speed was set to 100 nm/min with a response time being 1 s. Bandwidth was set at 1 nm and scanning mode was continuous. CD measurements of all mentioned proteins were recorded.

3.2.26. Strain Construction

The mutant or overexpression strains used for the following experiments were generated in the laboratory of Prof. Ivan Mijakovic, Chalmers University, Sweden. Construction of the phosphoablative *B. subtilis dnaK_Y601F* mutant at the natural *dnaK* locus was performed using the modified mutation delivery method of Fabret *et al* [181]. Briefly, the partially overlapping DNA fragments containing parts of the *dnaK_Y601F* gene and the flanking regions were obtained from the oligo pairs dnaKF_1/ dnaKF_4 and dnaKF_3/ dnaKF_2. Using primers dnaKF_1 and dnaKF_2, these fragments were PCR-joined to the insertion cassette containing the phage lambda cI repressor gene and the phleomycin resistance marker. The resulting PCR products were used to transform the competent BS33 cells expressing neomycin resistance gene from lambda promoter negatively controlled by the cI repressor. The transformants were selected for phleomycin resistance and neomycin sensitivity. Finally, counter-selection for neomycin resistance and phleomycin sensitivity was applied to select clones which had lost the insertion cassette from the chromosome via recombination between the flanking direct repeats. To construct *B. subtilis ptpZ* overexpression strain, *gfp* in pSG1729 was replaced with the *ptpZ* ORF through KpnI and XhoI. The *ptpZ* ORF was obtained from the oligos ptpZ_1729_fwd and ptpZ_1729_rev. The resulting plasmid was used to transform BS33. A similar strategy was employed to construct the *dnaK* complementation strain for *B. subtilis dnaK_Y601F*. The *dnaK* ORF was amplified using the oligo dnak_1729_fwd and dnak_1729_rev. The plasmid obtained was used to transform *B. subtilis dnaK_Y601F*. To remove *gfp* from pSG1729, pSG1729 was digested by KpnI and XhoI, the two sticky ends were then made blunt and ligated. The resulting plasmid was transformed into BS33 and *dnaK_Y601F* to obtain the control strain. **Table 8** lists all oligo sequences used for this study.

Table 8: PCR oligos used for strain constructions

Oligo Name	Sequence
dnaKF_1	cagcgtttgaaagacgcagc
dnaKF_2	ccatagcctctgacatttgaac
dnaKF_3	gagctcgaattcactggccgctcgcgacgctgaattgaagaagtaaaccgacgacc
dnaKF_4	cgacctgcaggcatgcaagctgttacttcttcaaattcagcgctcgacaacg
dnaK_1729_fwd	cggggtaccatgagtaaagttatcggaatcgac
dnaK_1729_rev	ccgctcgagttatttttgtttggtcgtcgtttac
ptpZ_1729_fwd	cggggtaccatgatcgatattcactgtcacattc
ptpZ_1729_rev	ccgctcgagttaaaagaaccaacaattttctct

3.2.27. Heat Shock Assays

Heat shock assays were conducted by Lei Shi in the laboratory of Prof. Ivan Mijakovic, Chalmers University, Sweden. *In vivo* heat resistance assay for liquid culture was carried out by splitting the culture into two at the early exponential phase (OD₆₀₀ 0.3). One was grown at 37 °C for 5 min, while the other was treated at 55 °C for 5 min. Both treated and untreated cultures were spread on plates by serial dilutions to obtain the number of cells that survived. For heat-resistance assay done on plates, serial dilutions of culture at OD₆₀₀ 0.3 were spotted onto plates and incubated at 58 °C for 10 h and additionally for 4 h at 37 °C.

3.2.28. Refolding Assay

Refolding assays were performed by Lei Shi in the laboratory of Prof. Ivan Mijakovic, Chalmers University, Sweden. Chemical denaturation and renaturation experiments were performed as described previously [182]. Briefly, 20 μM lactate dehydrogenase (LDH) was denatured in a buffer containing 50 mM potassium-HEPES, pH 7.5, 5 M urea and 5.0 mM DTT for 30 min at 25 °C. For the refolding assay, 2 μL of the denatured LDH was added to 198 μL of refolding solution containing 50 mM potassium-

HEPES, pH 7.5, 150 mM potassium acetate, 10 mM magnesium acetate, 10 mM phosphoenolpyruvate, 0.16 mM NADH, 10 U pyruvate kinase, 10 mM DTT, 4 mM ATP and 1 μ M DnaK. The activity of LDH was revealed by NADH oxidation. The kinetics at A₃₄₀ was measured by FLUOstar Omega.

3.2.29. Isolation of Genomic DNA

Isolation of genomic DNA from Gram-positive cells was using the Wizard® Genomic DNA Purification Kit according to the manufacturer's instructions.

1. An overnight culture was prepared by inoculating 2 μ L of the respective glycerol stock solution in 5 mL of LB. Cultures were incubated at 37 °C at 200 rpm.
2. 1 mL of the overnight culture was centrifuged in a 1.5 mL Eppendorf tube at 13,000 $\times g$ for 2 min to pellet the cells. The supernatant was discarded.
3. The cells were thoroughly resuspended in 480 μ L of 50 mM EDTA pH 8.0.
4. In order to weaken the cell wall, 120 μ L of 10 mg/mL lysozyme was added to the resuspended cell pellet and gently pipetted to mix.
5. The sample was incubated at 37 °C for 60 min. The sample was centrifuged for 2 min at 13,000 $\times g$ and the supernatant was discarded.
6. 600 μ L of Nuclei Lysis Solution provided was added to the sample and gently pipetted to resuspend the cells.
7. The sample was incubated at 80 °C for 5 min for cell lysis to occur and then cooled to RT.
8. 3 μ L of RNase Solution provided was added to the cell lysate. The Eppendorf was inverted 4 - 5 times to mix.
9. The sample was incubated at 37 °C for 60 min. the sample was then cooled to RT.
10. 200 μ L of Protein Precipitation Solution provided was added to the RNase treated cell lysate. The sample was vortexed vigorously for 20 s.
11. The sample was next incubated on ice for 5 min.
12. The sample was centrifuged at 13,000 $\times g$ for 3 min.
13. The supernatant was transferred to a fresh 1.5 mL Eppendorf tube containing 600 μ L of Isopropanol.

14. The Eppendorf tube was gently inverted 3 - 4 times to mix until thread-like strands of the DNA were visible.
15. The sample was centrifuged at this point at 13,000 x *g* for 2 min.
16. The supernatant was carefully discarded. 600 μ L of 70 % ethanol was added to the tube to wash the DNA pellet.
17. The sample was centrifuged at 13,000 x *g* for 2 min. The ethanol was removed by aspiration.
18. The DNA pellet was air-dried for 15 - 20 min.
19. 100 μ L of DNA Rehydration Solution provided was added to the DNA pellet and the sample was rehydrated by incubation at 65 °C for 1 h. The DNA was stored at 4 °C until further use.

3.2.30. Affinity Chromatography

Synthetic peptide pull-downs or affinity chromatographies were carried out in order to find interaction partners and domain structures that bind to phosphorylated STY residues. For this purpose a total of nine pairs of synthetic peptides in their unmodified and modified (phosphorylated) forms coupled to a biotin moiety via an -SGSG- linker were synthesized in the laboratory of Dr. Hubert Kalbacher, University of Tübingen, Germany (**Table 9**). SILAC was employed to differentiate between the true binders and the false positives (unspecific interaction). The same protocol was employed to find interaction partners for all peptide pairs, as described below:

1. 1 mg of the lyophilized unmodified and modified peptides were dissolved in 1 mL of buffer A (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 0.1 % NP-40). The peptides were dissolved by placing them on an end-over-rocker for 20 min. Brief sonication was carried out if required in order to completely dissolve the peptides.
2. 10 μ L of Streptavidin Sepharose beads was taken in two 1.5 mL Eppendorf's separately and washed twice with 1 mL of buffer A.
3. It was briefly vortexed and the beads were spun down each time at 4000 rpm for 2 min at 4 °C.

4. Peptides were next added to the beads at a concentration of 10 µg of peptides for every 10 µL beads. The volume was made up to 500 µL with buffer A and incubated for 30 min at 4 °C on an end-over rocker.
5. The beads were then washed twice as before with 1 mL buffer A at 4 °C to remove excess peptides.
6. 500 µg of SILAC labeled cell lysate in the “heavy” or the “light” form, diluted to 1 mL with dilution buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 0.1 % NP-40, protease inhibitors, 5 mM phosphatase inhibitors GP and SF, 0.5 mM DTT, 10 µM ZnCl₂), was incubated with the beads at 4 °C for 3 h on an end-over-rocker. The unmodified peptide was incubated with “light” lysate and phosphorylated peptide was incubated with “heavy” lysate.
7. The beads were washed thrice with 1 mL dilution buffer containing 350 mM NaCl.
8. The beads were next washed twice with PBS to remove traces of detergent.
9. The “light” and “heavy” beads were mixed together and taken forward for on-bead digestion.
10. The beads were resuspended in 50 µL elution buffer (2M urea in 100 mM Tris-HCl pH 7.5, 10 MM DTT).
11. They were incubated for 20 min at RT at 1000 rpm.
12. IAA was added to a final concentration of 50 mM and the beads were incubated at RT in the dark for 10 min.
13. 0.25 µg of the endoprotease Lys-C was added and the beads were incubated at RT for 2 h at 1000 rpm.
14. The beads were spun down at 2000 x *g* for 2 min at RT.
15. The supernatant was collected and transferred to a fresh 1.5 mL Eppendorf tube. 50 µL of elution was added again to the beads and incubated for 5 min at RT at 1000 rpm.
16. The beads were spun down again at 2000 x *g* for 2 min at RT and the supernatant was collected in the same Eppendorf as above (step 15).
17. 0.25 µg of Lys-C was added to the beads and the bound peptides were digested over night at RT at 1000 rpm.

18. The beads were spun down at 2000 x g for 2 min and the supernatant was added to the same Eppendorf as before (step 15) and the sample was acidified with 5 μ L of 10 % TFA.
19. The sample was stage-tipped as described before using C18 cartridges and analyzed on the LTQ-Orbitrap XL.

Table 9: List of Peptides Synthesized

Peptide Name	Unmodified Sequence	Modified Sequence
HPr	<i>Biotin-SGSG- GKT VNLK<u>S</u>IMGVMSL</i>	<i>Biotin-SGSG- GKT VNLK<u>pS</u>IMGVMSL</i>
PrkC_01	<i>Biotin-SGSG-RIPFDGE<u>S</u>AVSIALK</i>	<i>Biotin-SGSG-RIPFDGE<u>pS</u>AVSIALK</i>
PrkC_02	<i>Biotin-SGSG-IQEDEEM<u>T</u>KAIPIIK</i>	<i>Biotin-SGSG-IQEDEEM<u>pT</u>KAIPIIK</i>
PrkC_03	<i>Biotin-SGSG-SSTTITH<u>T</u>NSVLGSV</i>	<i>Biotin-SGSG-SSTTITH<u>pT</u>NSVLGSV</i>
YkwC	<i>Biotin-SGSG-EENSGTQ<u>S</u>IYKLWVK</i>	<i>Biotin-SGSG-EENSGTQ<u>pS</u>IYKLWVK</i>
YmfM	<i>Biotin-SGSG-NTYHDDV<u>S</u>EKISGMN</i>	<i>Biotin-SGSG-NTYHDDV<u>pS</u>EKISGMN</i>
PtkA	<i>Biotin-SGSG-SKHSEYG<u>Y</u>GTKDNF</i>	<i>Biotin-SGSG-SKHSEYG<u>pY</u>GTKDNF</i>
SunI	<i>Biotin-SGSG-SIKESDK<u>Y</u>GVADNID</i>	<i>Biotin-SGSG-SIKESDK<u>pY</u>GVADNID</i>
DnaK	<i>Biotin-SGSG-DNVVDAE<u>Y</u>EEVNDDQ</i>	<i>Biotin-SGSG-DNVVDAE<u>pY</u>EEVNDDQ</i>

4. Results

4.1. Quality Control

4.1.1. Quality Control of the Knock-out Strains (PCR)

Bacillus subtilis 168 $\Delta lysA$ (strain auxotrophic for lysine), also referred to as wild-type (WT) and STY kinase and phosphatase gene knock-out strains, namely $\Delta prkC$, $\Delta ptkA$, $\Delta yabT$, $\Delta ybdM$, $\Delta yveL$, $\Delta prpC$, $\Delta ptpZ$, $\Delta yfkj$, all generated in a $\Delta lysA$ background, were obtained from the laboratory of Prof. Ivan Mijakovic, INRA, France. All knock-out strains were tested by polymerase chain reaction for the deletion of the respective genes using standard protocols. As seen in the **Figure 17**, PCR performed with genomic DNA isolated from the respective knock-out strains resulted in no gene amplification. Conversely, PCR performed with DNA isolated from the WT strain (positive control for the test reaction conditions) showed amplification of the respective genes of interest.

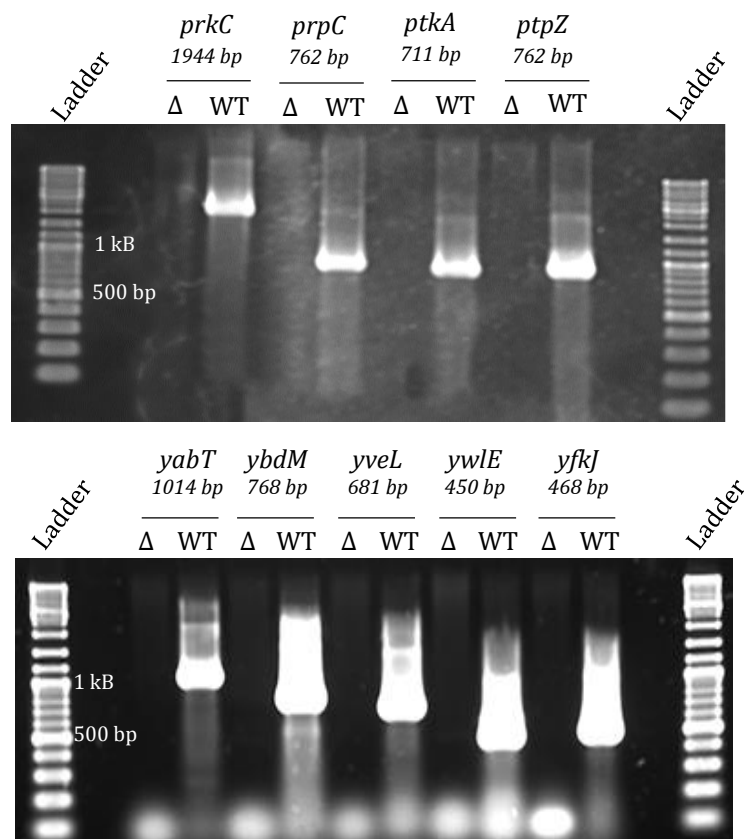


Figure 17: Confirming gene deletions – Deletion of respective genes in the knock-out strains (Δ) was confirmed by polymerase chain reaction. Similar reaction was carried out using “WT” genomic DNA as a positive control.

The strains were additionally tested for the deletion of *lysA*, which is essential for synthesis of lysine, by culturing them in minimal media containing no lysine supplements. The failure to grow in the absence of lysine indicated the deletion of *lysA*.

4.1.2. Quality Control of SILAC Incorporation and Mixing

As a quality control step, Lys4-labeled cultures and Lys8-labeled cultures from all experiments were assessed for the level of incorporation of the lysine label. For this purpose, 10 µg of the labeled protein extract was digested by endoproteinase Lys-C and analyzed on the mass spectrometer (MS). Incorporation level of the label was calculated using the formula

$$\frac{\tilde{X}\left(\frac{H}{L}\right)}{1 + \tilde{X}\left(\frac{H}{L}\right)} * 100$$

where (H/L) refers to the unnormalized ratio of the quantified peptide with the heavy or medium label to the corresponding quantified unlabeled peptide. In all cases, an incorporation of 98 % or higher was observed indicating complete metabolic labeling of the cells (**Figure 18**).

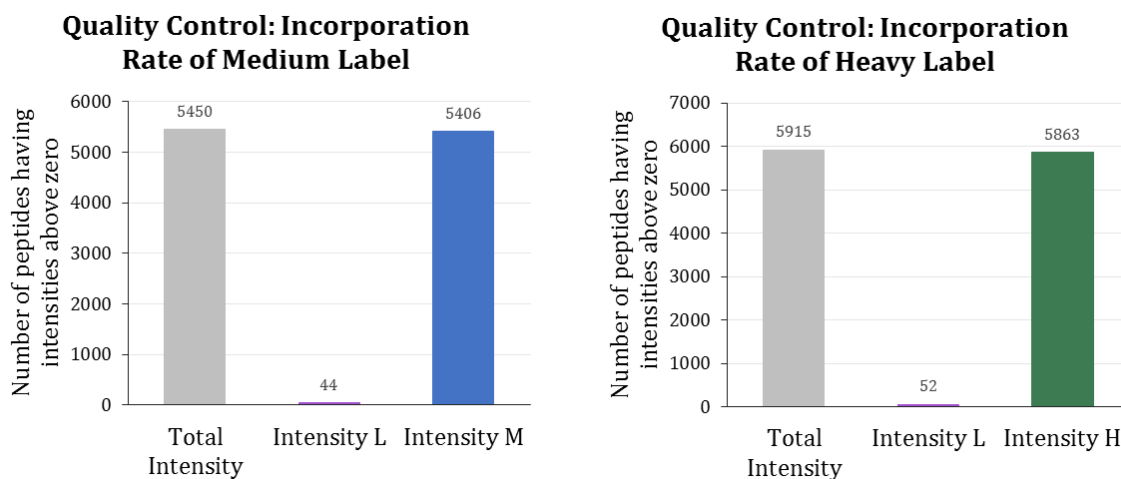


Figure 18: Quality Control – Level of incorporation of the medium and heavy label was checked for all experiments. Shown here, is a representative example, wherein, the left and the right panels represent bar plots showing 99 % incorporation of the medium label and heavy label, respectively.

Labeled protein extracts from individual experiments were also mixed in equal amounts based on Bradford measurements and digested and analyzed for mixing errors. Based on the median of the unnormalized evidences, mixing errors were corrected using the formula:

$$\text{Measured Protein Concentration} * \tilde{X} (H/L) = \text{True Protein Concentration.}$$

A representative example of the mixing check quality control has been shown in **Figure 19**.

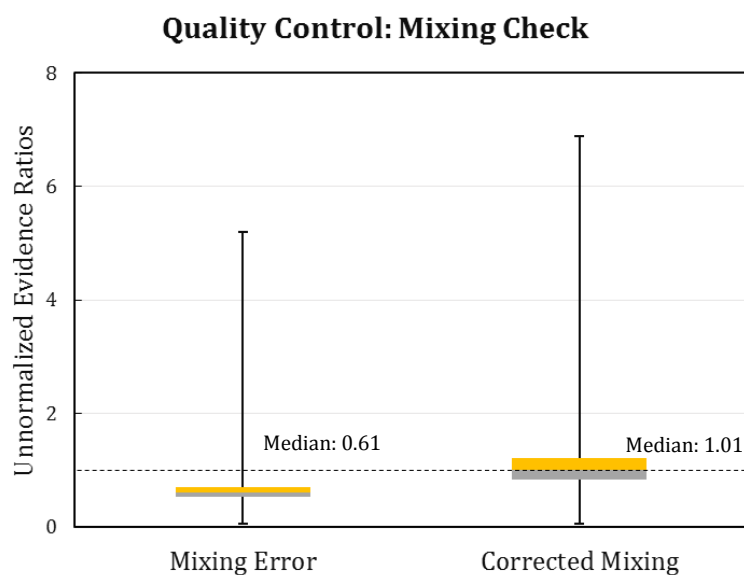


Figure 19: Quality Control – Samples were mixed in equal amounts based on their Bradford measurements and the median of their unnormalized evidence ratios were calculated. The mixing errors were then corrected for, based on their median values. Equal or perfect mixing is considered to be achieved if the median is equal to 1. The above figure shows a boxplot representation of the mixing check done based on the protein concentrations measured by Bradford (left) and mixing check that has been corrected for (right).

4.2. Proteome and Phosphoproteome Dynamics of *B. subtilis* During Growth

This section has been published in "Ravikumar V, et al, Quantitative phosphoproteome analysis of Bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC. Mol Cell Proteomics, 2014. 13(8): p. 1965-78".

4.2.1. Experimental Design of the Growth Curve Measurements

The proteome and phosphoproteome dynamics of the lysine-auxotroph *B. subtilis* Δ lysA strain was investigated by carrying out two triple-SILAC experiments. Cell cultures labeled with different forms of lysine were harvested at specific time points:

T1) exponential growth (Lys0); T2) entry into retardation phase (Lys4); T3) transition to stationary phase (Lys8); T4) early stationary phase (Lys0); T5) late stationary phase (Lys8) [183] (**Figure 20A**). A common time point, T2 (Lys4), was used to integrate the data into a five time point expression curve for each quantified protein. Correlation of intensities of Lys4-labeled peptides was high (Pearson coefficient 0.8) (**Supplementary Figure 1**). Proteins extracted from the corresponding SILAC cultures were mixed in equal ratios and digested with endoproteinase Lys-C. The resulting peptide mixtures were subjected to phosphorylated peptide enrichment using SCX and TiO₂ chromatography; for proteome analysis peptides were fractionated by Offgel isoelectric focusing. All samples were analyzed by nanoLC-MS/MS on Orbitrap mass spectrometers (**Figure 20B**). Two biological replicates were performed; Pearson correlation coefficients between the replicates ranged from 0.6 (T1) to 0.74 (T4).

4.2.2. Proteome and Phosphoproteome Dynamics During Growth in Batch Culture

Analysis of the proteome of *B. subtilis* resulted in identification of 18,699 peptides mapping to 2,264 proteins at false discovery rate (FDR) of 1 % at the peptide and protein level; 2,033 proteins were quantified in one of the five analyzed phases, whereas 1,666 proteins were quantified in all analyzed phases of growth (**Supplementary Figure 2**). The phosphoproteome analysis resulted in identification of 177 phosphorylation events of which 156 were quantified in any of the analyzed phases and 144 were localized to a

specific residue; 64 phosphorylation sites (on 56 proteins) were quantified in all analyzed phases. List of all detected proteins and phosphorylation sites is presented in **Supplementary Table 1**.

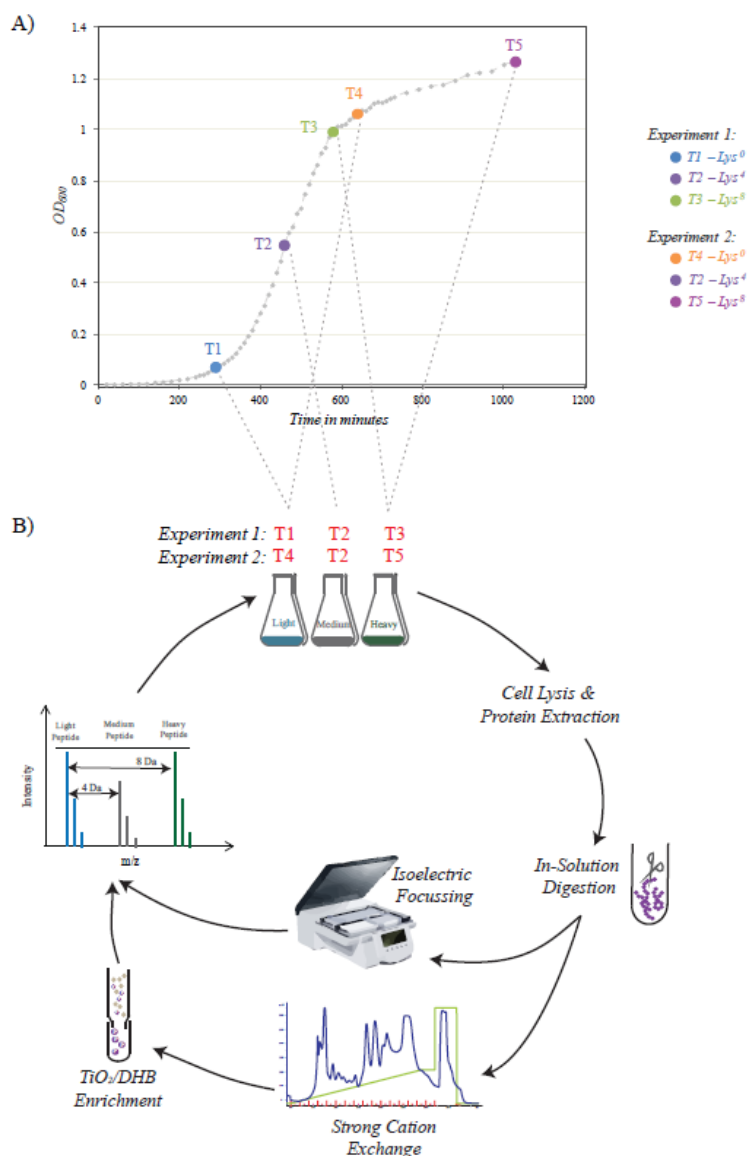


Figure 20: Growth curve profile of *B. subtilis* in minimal medium - A) Depicted here is a composite growth curve of *B. subtilis* grown in batch culture, in minimal medium under laboratory conditions. In each SILAC experiment cultures were labeled with the respective isotope of lysine and cells were harvested at the denoted time points. **Phosphoproteomics workflow** - B) Cells were harvested and lysed at the respective stages in each SILAC experiment. Proteins were extracted, mixed in equal ratios and digested with endoproteinase Lys-C. Digested peptides were fractionated by Offgel isoelectric focussing (for whole proteome analysis) and, separately, by SCX and TiO_2 chromatographies (for phosphorylated peptide enrichment). All samples were analyzed on LTQ-Orbitrap mass spectrometer.

In agreement with previous studies, most of the phosphorylation events were observed on Ser residue (74.6%), followed by Thr (18.6%) and Tyr (7.3%). Normalization of phosphorylated peptide ratios with corresponding protein ratios showed little influence of protein dynamics on phosphorylation site dynamics, especially in early phases of growth.

Dynamics of the proteome and the phosphoproteome across different stages of growth is shown in **Figure 21**. Fluctuation was more pronounced at the phosphorylation level, resulting in higher standard deviation of SILAC ratios in the phosphoproteome dataset (**Figure 21**). As observed previously [184], proteome dynamics was more pronounced at later stages of growth resulting in a wide spread of measured SILAC ratios. The phosphoproteome dynamics followed the same trend, but showed predominantly an increase in phosphorylation levels at later stages of growth (**Figure 21**). Interestingly, recently the same trend was observed in the phosphoproteome of *Escherichia coli* grown under similar conditions [185], pointing to the fact that bacterial protein phosphorylation predominantly occurs at later stages of growth. To define the subset of dynamic (highly fluctuating) and static proteins, they were separated into four bins based on the standard deviation of ratios measured across all growth phases (**Figure 22**). Functional enrichment analysis showed that ribosomal proteins and proteins involved in ATP metabolism were highly fluctuating, whereas enzymes of basic metabolism (e.g. glycolysis) had more constant levels across analyzed growth phases (**Supplementary Figure 3**).

To obtain further insights into functional classes of proteins and phosphorylation sites differentially regulated across the analyzed stages of growth, hierarchical clustering analysis was performed (**Figure 23**). Differentially regulated proteins grouped broadly into 4 clusters: Cluster 1 consisted of proteins that had increased levels at later stages of growth and had functions related to proteolysis and degradation (e.g. D-aminopeptidase, DppA; cell wall-associated protease, WprA) and respiration (e.g. alkyl hydroperoxide reductase, AhpF; cytochrome c-550, CccA). Cluster 2 showed the opposite trend and contained functions related to translation and gene expression (e.g. 30S ribosomal protein S11, RpsK; 50S ribosomal protein L14, RplN) and ATP-binding proteins (e.g. ABC transporters such as cell division protein, FtsE; thiamine import protein, YkoD).

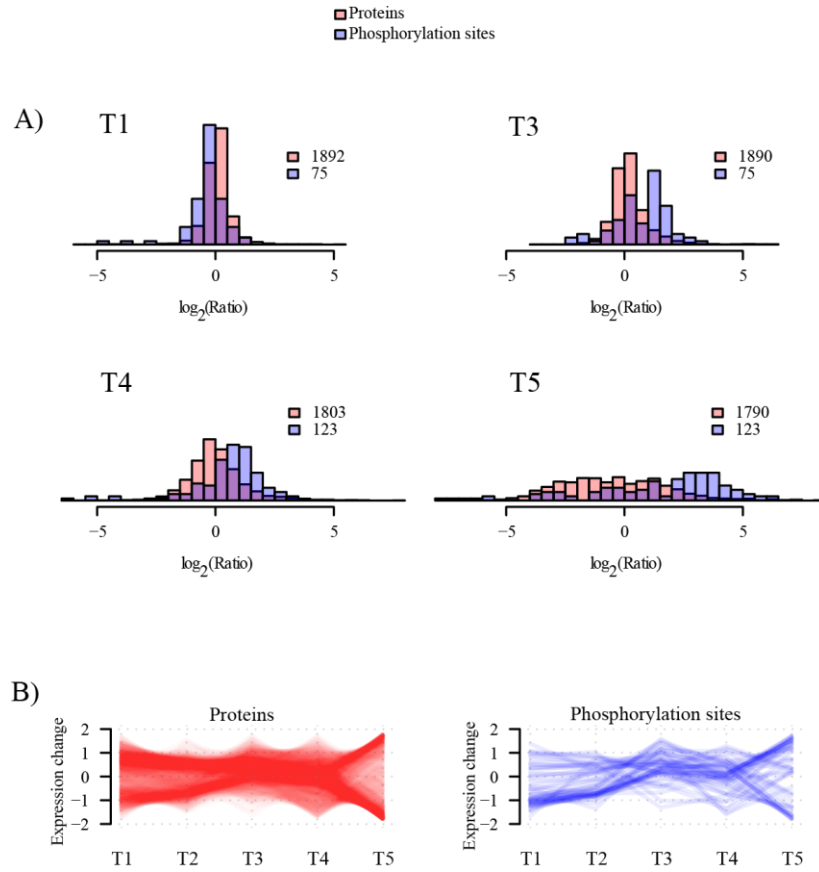


Figure 21: Proteome and phosphoproteome dynamics – A) Histogram depicting measured ratios of proteins and phosphorylation sites in each stage of growth; T1, T3, T4, and T5. All distributions are relative to the common time point, T2. B) Changing expression profiles of proteins and phosphorylation sites detected across all 5 stages of growth.

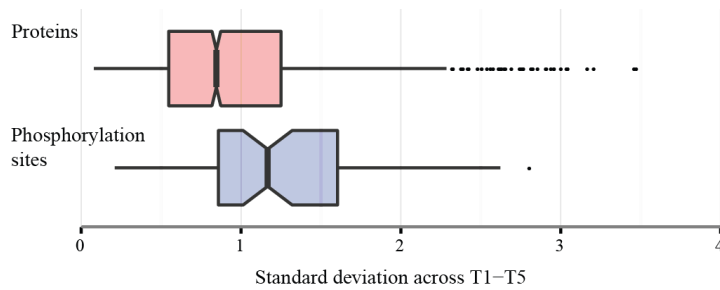


Figure 22: Standard Deviation - Box plot depicting standard deviation across growth phases T1-T5 measured for proteins and phosphorylation sites.

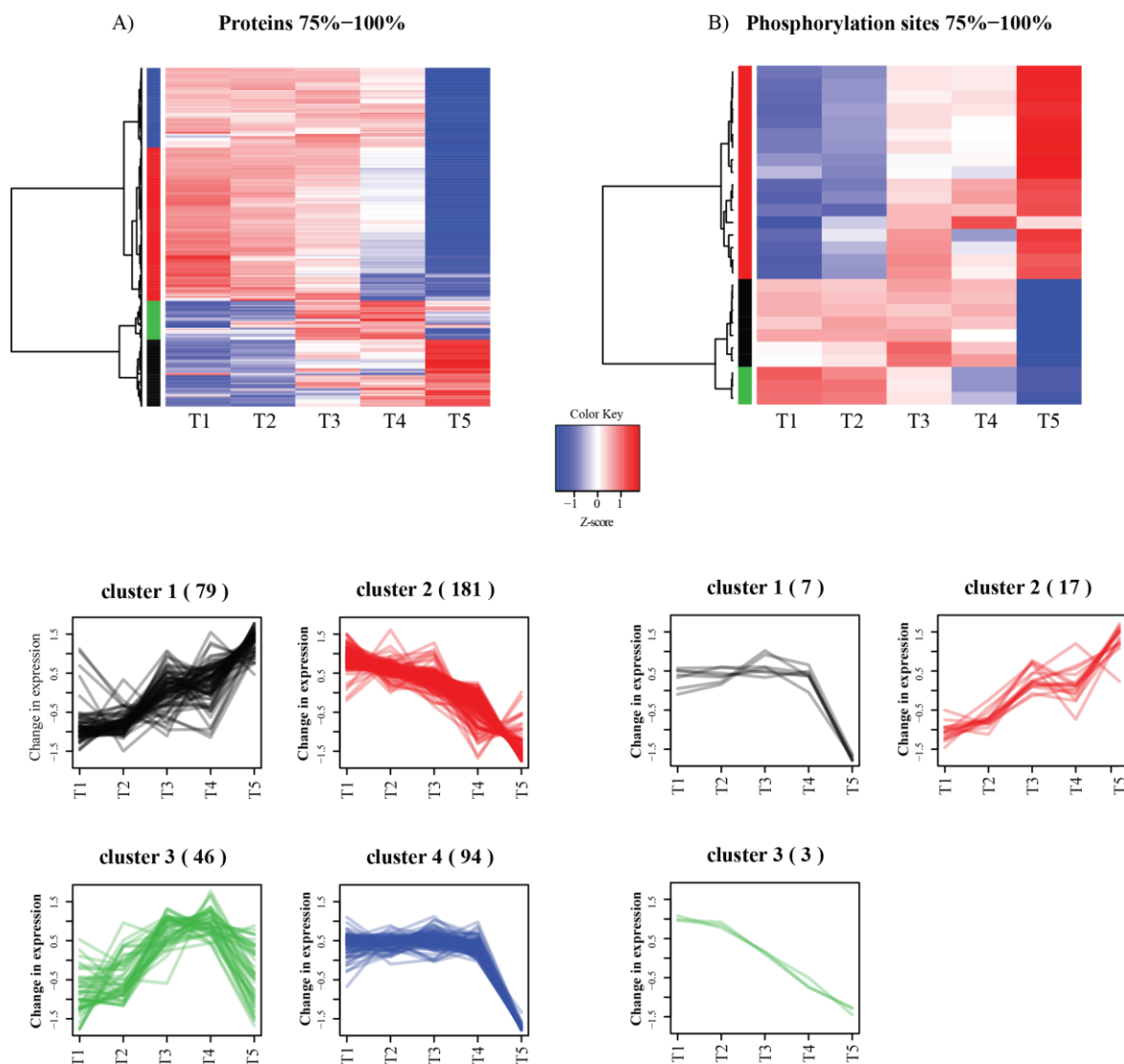


Figure 23: Hierarchical clustering of A) protein groups and B) phosphorylation sites – Fluctuation was defined based on standard deviation, as described in the text.

Cluster 3 contained proteins with expression peaking in transition to stationary phase and was enriched in functions related to antibiotic biosynthesis pathways (e.g. polyketide biosynthesis protein, PksE), fatty acid or carbohydrate metabolism (e.g. teichoic acid biosynthesis protein, GgaA) and response to stress (e.g. general stress protein A, YwaG). Finally, the Cluster 4 consisted of proteins that had abruptly lower levels in the late stationary phase of growth and included transcription factors such as elongation factor Tu, TufA; transcriptional repressor, NrdR and proteins involved in amino acid biosynthesis (e.g. threonine deaminase, IlvA). Differentially regulated phosphorylation

sites grouped into three clusters. The phosphorylation on the first cluster of proteins peaked at the transition phase and abruptly dropped during the stationary phase (e.g. tyrosine-protein kinase PtkB, (EpsB, YveL). The second cluster showed increasing phosphorylation from T1 to T5 and these included proteins involved in carbon metabolism, such as aconitate hydratase, CitB and fructose-1,6-bisphosphate aldolase, Fba. The third cluster had gradually decreasing phosphorylation levels (e.g. chorismate mutase, AroA).

4.3. Novel *In vivo* Substrates of Ser/Thr Kinase PrkC and Phosphatase PrpC

This section has been published in "Ravikumar V, et al, Quantitative phosphoproteome analysis of Bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC. Mol Cell Proteomics, 2014. 13(8): p. 1965-78".

4.3.1. Experimental Design of PrkC and PrpC Knock-out Screens

In order to identify novel *in vivo* kinase/phosphatase substrates and understand their role in bacterial physiology and/or growth, a global and quantitative site-specific phosphoproteomics analysis was performed using the triple-SILAC approach (as described earlier), in the background of the respective kinase/phosphatase knock-outs. The WT ($\Delta lysA$) strain was Light labeled, the $\Delta prkC$ (kinase) strain was Medium labeled and the $\Delta prpC$ (phosphatase) strain was Heavy labeled. Cell cultures labeled with different forms of lysine were harvested during the late stationary phase of growth. Proteins extracted from the corresponding SILAC cultures were mixed in equal ratios and digested with endoproteinase Lys-C. The resulting peptide mixtures were subjected to phosphorylated peptide enrichment using SCX and TiO₂ chromatography; for proteome analysis peptides were fractionated by GeLC and Offgel isoelectric focusing. All samples were analyzed by nanoLC-MS/MS on LTQ-Orbitrap mass spectrometers (**Figure 24**). Three biological replicates were performed; Pearson correlation coefficients between the replicates were 0.35 – 0.55. The low correlation could be attributed to the fact that there is more fluctuation seen at the protein level during the late stationary phase of growth, as described earlier.

4.3.2. SILAC-Based Screen of *In vivo* Substrates of the Kinase PrkC and Phosphatase PrpC

The triple-SILAC workflow established to monitor (phospho) proteome dynamics during growth was extended to identification of novel substrates of the kinase PrkC and phosphatase PrpC. From the preliminary growth curve experiments, stationary phase was identified as the phase of growth where the kinase and phosphatase were most abundant.

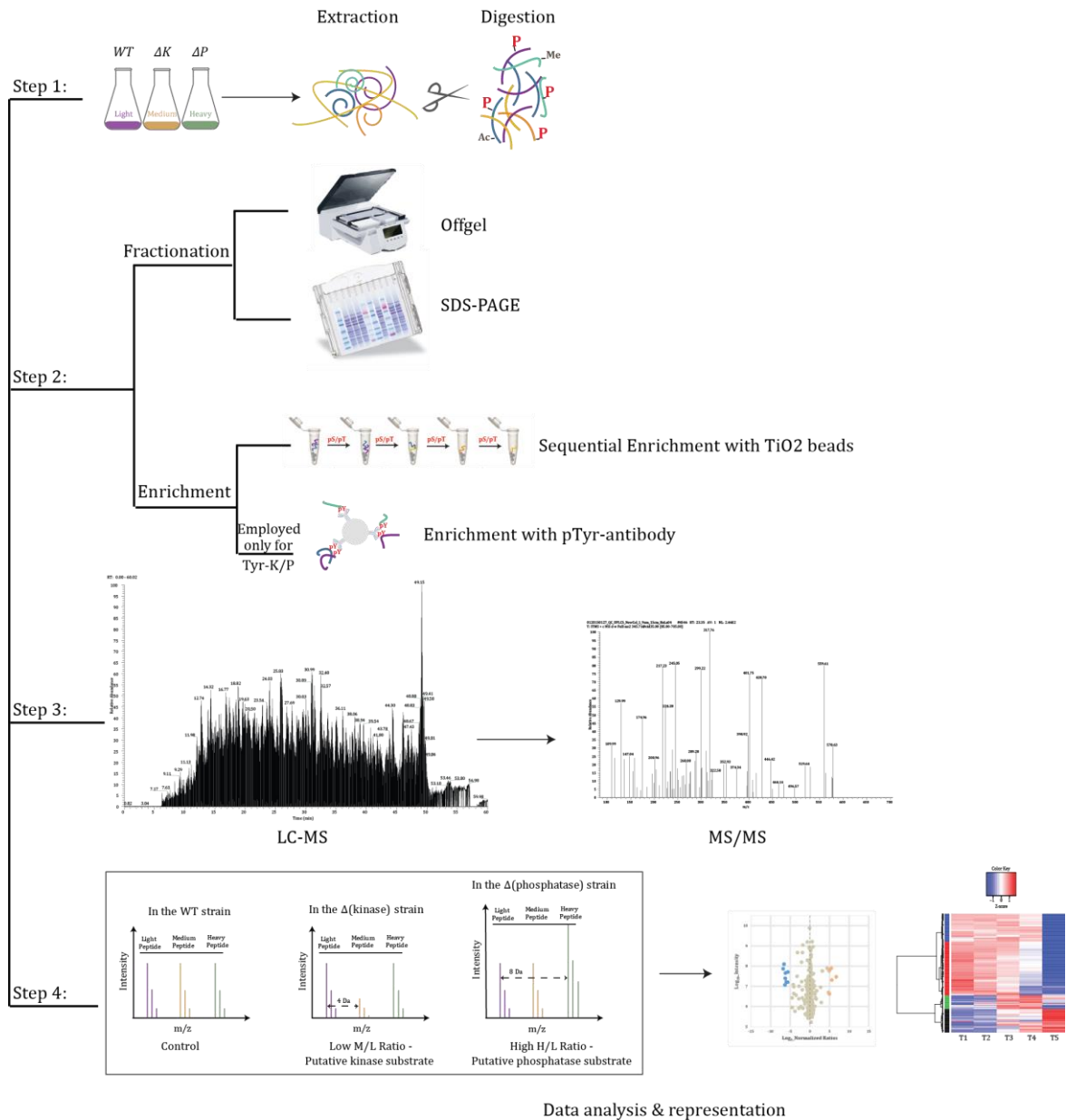


Figure 24: General overview of the proteomics workflow adopted for the identification of novel substrates of the investigated kinases and phosphatases. Briefly, cells were harvested at specific time points of growth. Extracted proteins were digested and fractionated. Portion of the digested peptides was analyzed for phosphorylation sites as before. WT acts as a control and all ratios were normalized to this state. Proteins with a down regulated M/L ratio and/or an up regulated H/L ratio were considered as potential candidates of the kinase and/or phosphatase respectively.

The kinase and phosphatase KO strains, along with the WT (acting as control), were inoculated in respective SILAC minimal media and the cells were harvested in mid-

stationary phase ($OD_{600} = 1.35$). Proteome and phosphoproteome measurements were performed as described above. A total of 177 phosphorylation events were identified of which 124 were quantified. 141 phosphorylation events were localized with a probability of ≥ 0.75 . Ratios obtained from the analysis of the proteome were used to normalize the quantified phosphorylated sites. In these measurements, 2,131 proteins were identified and 2,026 of them were quantified. In addition, the incorporation of the medium and heavy label was measured and found to be above 98 % in all replicates. List of all detected proteins and phosphorylation sites is presented in **Supplementary Table 2**. Global phosphoproteomic analysis during conditions of absence of the kinase or phosphatase led to the detection of seven differentially down regulated events in the $\Delta prkC$ strain (potential substrates of the PrkC kinase) (**Figure 25A**) and eight differentially up regulated events in the $\Delta prpC$ strain (potential substrates of the PrpC phosphatase) (**Figure 25B**).

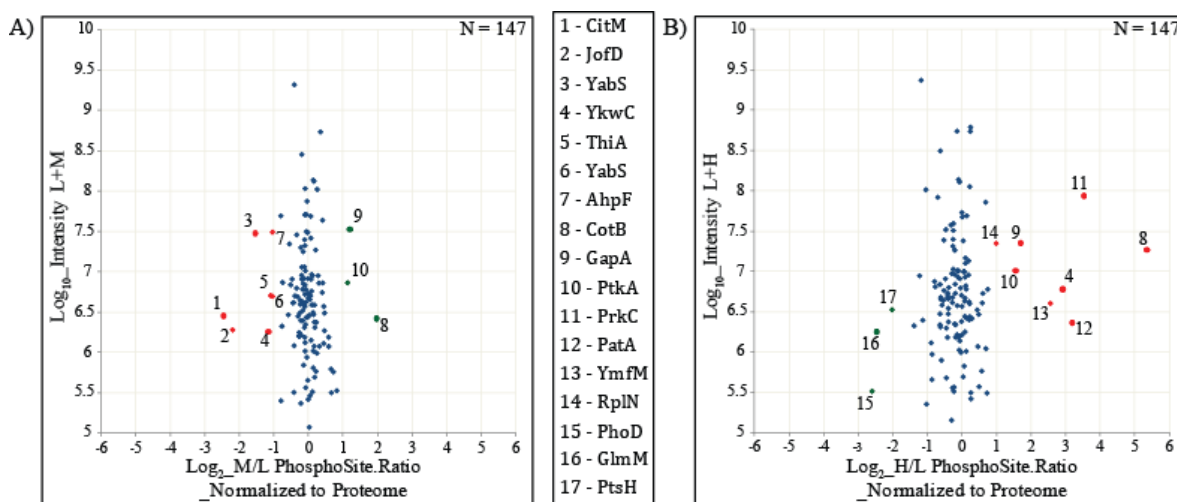


Figure 25: PrkC and PrpC phosphoproteomics screen – Scatter plot showing A) $\Delta prkC$ /WT SILAC ratios and B) $\Delta prpC$ /WT SILAC ratios respectively. \log_2 ratios of the phosphorylation sites are normalized to the corresponding protein and plotted against peptide intensities in \log_{10} scale. Significantly changing ($p = 0.05$) SILAC ratios marked in red, present potential substrates of PrkC or PrpC respectively and those marked in green are differentially regulated but not direct substrates of these enzymes.

Among potential PrkC substrates were CitM, a 2-oxoglutarate dehydrogenase complex component E2 which plays an important role in the Krebs cycle; JofD, subunit of the 30s

ribosomal protein complex; YabS, an uncharacterized protein whose function has not yet been established; ThiA, a thiamine biosynthesis protein; and AhpF, an alkyl hydroperoxide reductase that protects the cell against reactive oxygen species during stress. Potential PrpC substrates included CotB, a spore-coat protein that helps in resistance during adverse conditions; PatA, a putative aminotransferase involved in protein biosynthesis; YmfM, an uncharacterized membrane protein implicated in having a role in cell shape determination; GapA, glyceraldehyde-3-phosphate dehydrogenase enzyme associated with glycolysis; PtkA (YwqD), a tyrosine kinase; and RplN, subunit of the 50s ribosomal protein. An autophosphorylation site of PrkC, known to be dephosphorylated by PrpC, was also detected as highly regulated in the $\Delta prpC$ strain thus acting as a positive control. Interestingly, one phosphorylation site, Ser281 on the uncharacterized oxidoreductase YkwC, was significantly affected by both, PrkC and PrpC, making it a potential substrate of both enzymes. Manual validation of the MS/MS spectrum of the phosphorylated YkwC peptide showed good coverage and annotation of fragmentation ions (**Figure 26**).

Interestingly, in both the kinase and the phosphatase knockout, phosphorylation sites for which the occupancy was affected in the unexpected direction, higher in the kinase knockout (CotB, GapA and PtkA) and lower in the phosphatase knockout (PhoD, GlmM and PtsH), was identified. Although these phosphorylation sites cannot be direct substrates of PrkC and PrpC, they are likely indirectly regulated or positioned further downstream in the signal transduction cascade, a phenomenon commonly observed in eukarya [186]. For example, in the above described dataset, the inactivation of PrkC leads to an increase of phosphorylation of the BY-kinase PtkA at Y228 (autophosphorylation of PtkA). It has recently been established that PrkC phosphorylates PtkA at S223, in the immediate vicinity of Y228 [187]; this phosphorylation event presumably prevents autophosphorylation of PtkA at Y228, leading to an increase of PtkA autophosphorylation in $\Delta prkC$ strain that was observed here. The remaining indirectly affected sites that were detected are likely to be explained by a similar phenomenon of cross-talk between kinases and phosphatases, which seem to be abundant in *B. subtilis* [187].

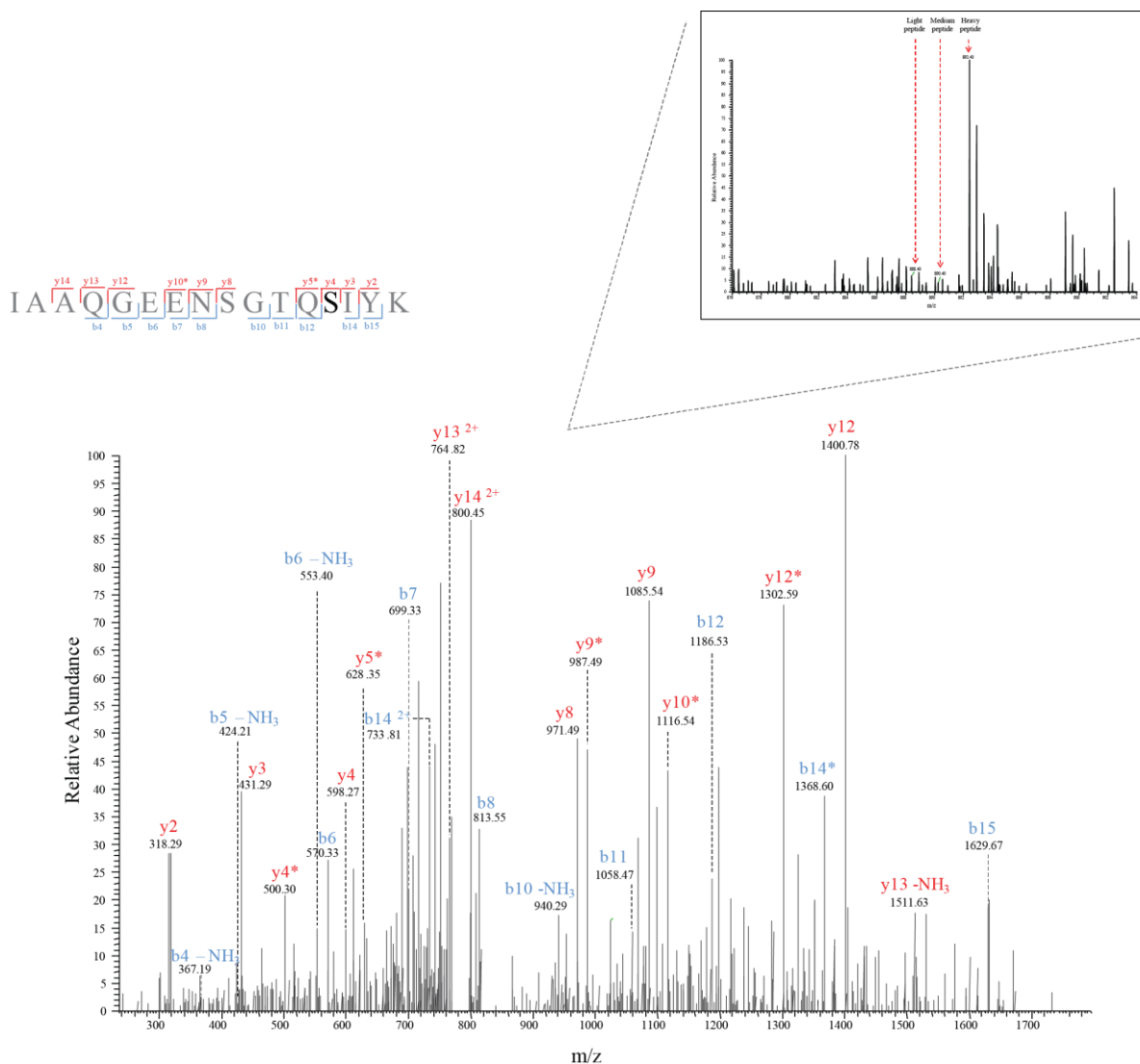


Figure 26: MS/MS spectrum of the YkwC peptide containing phosphorylation at Ser281 - This residue is a potential substrate of both, PrkC and PrpC. Inset shows the corresponding SILAC triplet in a survey scan.

4.3.3. Validation of YkwC Ser281 as a Substrate of PrkC and PrpC

In order to validate Ser281 on YkwC as a substrate of PrkC and PrpC further biochemical experiments were conducted. YkwC, PrkC and PrpC were over-expressed and purified in *E. coli* and *in vitro* phosphorylation and dephosphorylation assays using γ -³²P ATP were performed. The assays showed that YkwC could not autophosphorylate; that PrkC was able to phosphorylate YkwC; and that PrpC was able to dephosphorylate phosphorylated YkwC *in vitro* (**Figure 27**). This confirmed that YkwC was indeed a substrate of both

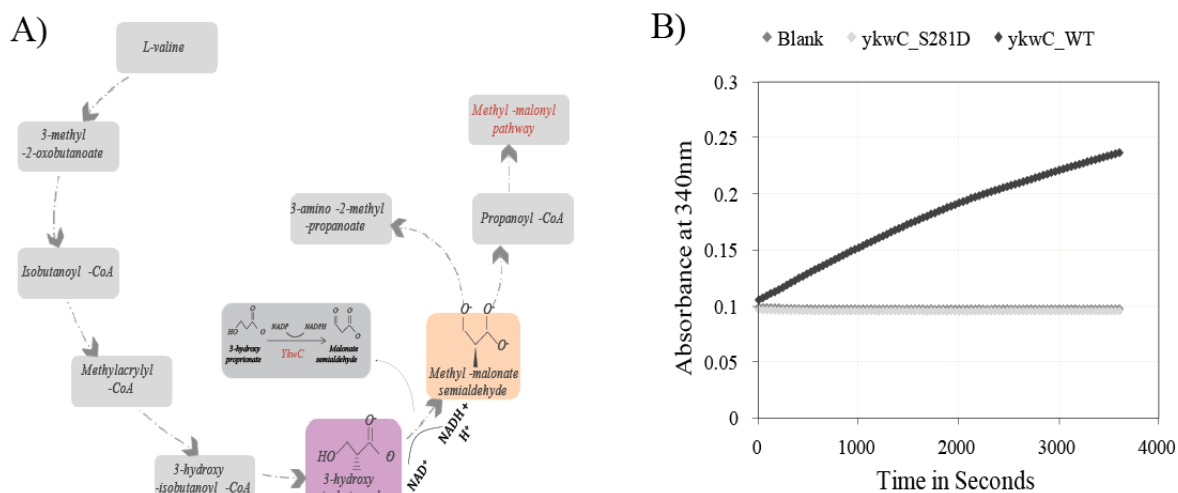


Figure 28: In vitro enzymatic assay – A) The figure represents the valine metabolism pathway in *B. subtilis* and the role of YkwC. B) Enzyme activity of YkwC_WT (1 μ g) and YkwC_S281D (1 μ g), the phosphomimetic mutant, incubated with its substrate (3-HP) was monitored over 1 h at intervals of 30 sec at 340 nm. The rate of formation of NADPH was measured in a 200 μ L reaction.

To ensure that the effect shown by the mutant was not due to protein misfolding induced by the mutation, circular dichroism (CD) spectra of WT and YkwC S821D were compared. The spectra of the two proteins were very similar, indicating that the mutation did not disrupt the protein structure (**Figure 29**). In order to validate our *in vitro* findings *in vivo*, a similar enzymatic assay was repeated with dialyzed total protein extracts obtained from the WT, $\Delta prkC$ and $\Delta prpC$ strains. The results showed increased oxidoreductase activity of YkwC in $\Delta prkC$ strain and loss of oxidoreductase activity in $\Delta prpC$ strain (**Figure 30**). This was fully consistent with *in vitro* findings, and confirmed that PrkC and PrpC control the occupancy of this phosphorylation site *in vivo*. As a negative control, cells were grown without glucose (conditions under which YkwC is not expressed), and YkwC activity could not be detected (**Figure 30**).

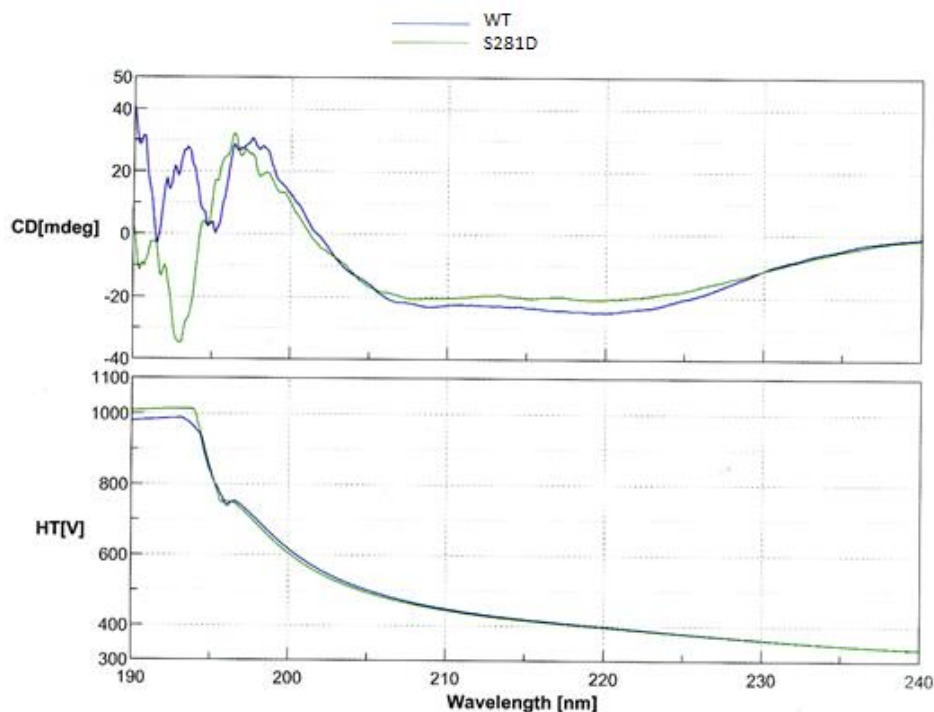


Figure 29: Circular dichroism of YkwC and YkwC_S281D – A far UV CD spectrum of YkwC (WT) and YkwC_S281D, measured in the range of 240 nm to 190 nm, denoted a similar secondary structure pattern of both enzymes.

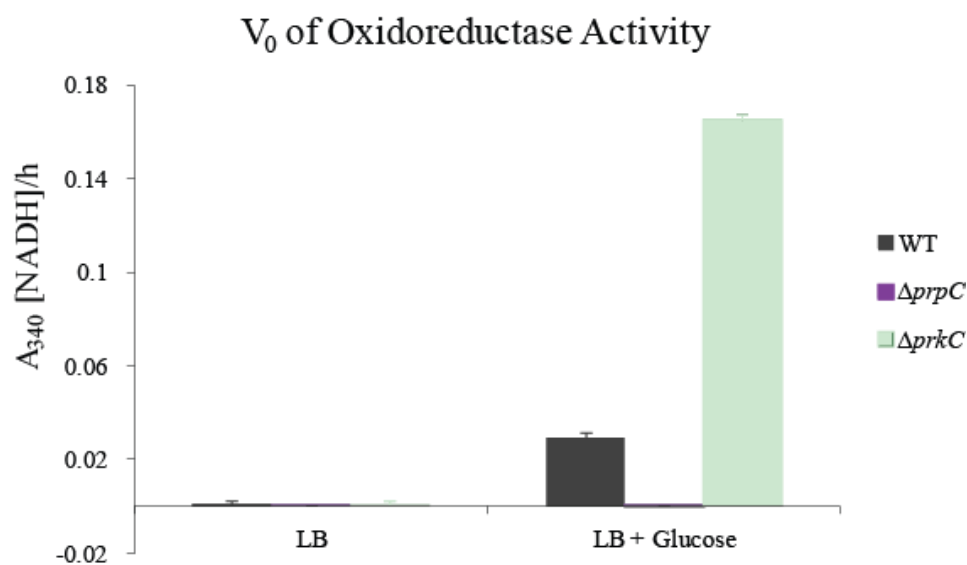


Figure 30: In vivo enzymatic assay – V₀ of oxidoreductase activity of dialyzed total protein extract from *B. subtilis* WT, ΔprpC and ΔprkC cells cultured in LB and LB supplemented with 0.3 % glucose [LB + Glucose] during transition phase. NADPH was detected at A₃₄₀.

4.4. Novel *In vivo* Substrates of Tyr Kinase PtkA and Phosphatase

PtpZ

This section forms a part of “Shi L, Ravikumar V, Derouiche A, Kalantari A, Macek B, Mijakovic I. Phosphorylation of chaperon protein DnaK is crucial for DnaK activity and cellular survival upon heat shock in Bacillus subtilis” (manuscript in preparation).

4.4.1. Experimental Design of PtkA and PtpZ Knock-out Screens

In the background of the kinase knock-out (PtkA) and phosphatase knock-out (PtpZ), a global and quantitative site-specific phosphoproteomics analysis using the triple-SILAC approach (as described earlier) was performed. The WT ($\Delta lysA$) strain was Light labeled, the $\Delta ptkA$ (kinase) strain was Medium labeled and the $\Delta ptpZ$ (phosphatase) strain was Heavy labeled. Since the kinase PtkA was detected to be abundant during mid-logarithmic phase and the phosphatase PtpZ was detected as abundant during the stationary phase of growth, differentially SILAC-labeled cell cultures were harvested either during the late stationary phase of growth (two replicates) or during mid-logarithmic phase (two replicates). Proteins extracted from the corresponding SILAC cultures were mixed in equal amounts and digested with endoproteinase Lys-C. The resulting peptide mixtures were subjected to phosphorylated peptide enrichment using SCX and TiO_2 chromatography. In addition pTyr antibody was used specifically for enriching Tyr phosphorylated peptides. For proteome analysis, peptides were fractionated by GeLC and Offgel isoelectric focusing. All phosphorylated peptide enriched samples were analyzed by nanoLC-MS/MS on LTQ-Orbitrap XL and proteome measurements on the LTQ-Orbitrap Elite mass spectrometers. Processing was done using MaxQuant 1.5.1.0.

4.4.2. SILAC-Based Screen of *In vivo* Substrates of the Kinase PtkA and Phosphatase PtpZ

The incorporation of the medium and heavy label was found to be above 98 % in all of the replicates. . SILAC-based measurements yielded 2,466 protein identifications of which 2,405 of them were quantified. A total of 303 phosphorylation events were identified of which 194 were quantified; 253 phosphorylation events were localized with

a probability of ≥ 0.75 and 121 of them passed the ≤ 0.001 PEP cutoff. List of all detected proteins and phosphorylation sites is presented in **Supplementary Table 3**. Ratios obtained from the analysis of the proteome were used to normalize the quantified phosphorylated sites. Global phosphoproteomic analysis conducted using knock-out strains of the kinase and phosphatase followed by normalization against the proteome, led to the detection of six differentially down regulated events in the $\Delta ptkA$ strain (**Figure 31A**) and six differentially up regulated events in the $\Delta ptpZ$ strain (**Figure 31B**).

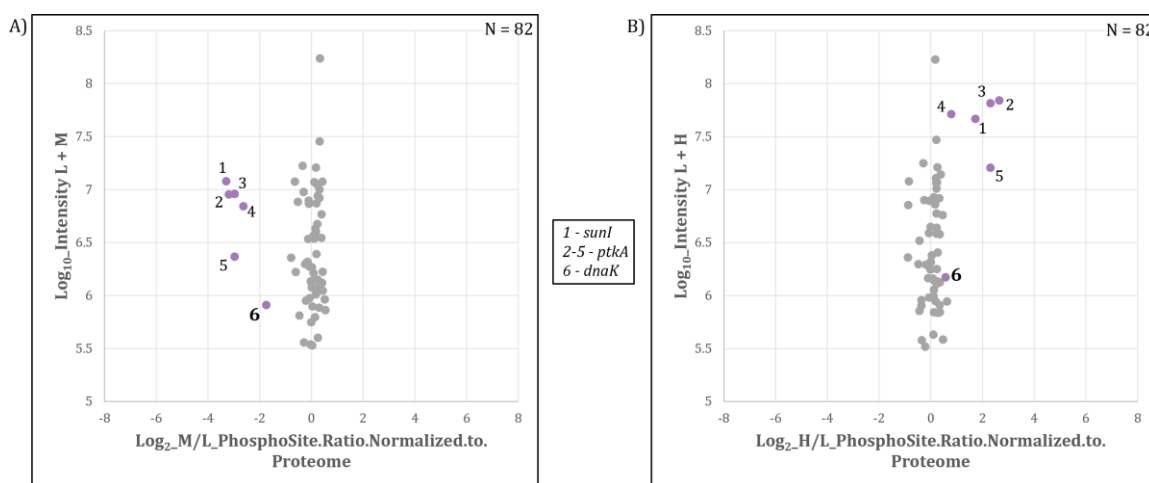


Figure 31: PtkA and PtpZ phosphoproteomics screen – Scatter plot showing A) $\Delta ptkA$ /WT SILAC ratios and B) $\Delta ptpZ$ /WT SILAC ratios respectively. Log_2 ratios of the phosphorylation sites are normalized to the corresponding protein and plotted against peptide intensities in Log_{10} scale. Significantly changing ($p = 0.05$) SILAC ratios marked in purple, present potential substrates of PtkA (M/L Ratios) and PtpZ (H/L Ratios) respectively.

Among potential PtkA and PtpZ substrates were SunI, a bacteriocin producer immunity protein that plays an important role in conferring immunity to the bacterium against sublancin; DnaK, a class-I heat-shock protein that acts as a molecular chaperon and gets recruited under conditions of heat stress or ethanol stress. Both SunI and DnaK were down-regulated in the $\Delta ptkA$ and up-regulated in the $\Delta ptpZ$ strain. Known autophosphorylation events on the kinase, PtkA, were not detected in the PtkA knock-out strain, but are reported as down-regulated as a result of the use of re-quantification algorithm during data processing. Three known autophosphorylation sites of PtkA (Y225, Y227 and Y228), known to be dephosphorylated by PtpZ, were detected as highly up-

regulated in the $\Delta ptpZ$ strain thus acting as a positive control. One phosphorylation site, Tyr601 on the heat-shock chaperon DnaK, regulated by both, PtkA and PtpZ, was chosen as a potential substrate of both enzymes and subjected to further biochemical analyses. Manual validation of the MS/MS spectrum of the phosphorylated DnaK peptide showed good coverage and annotation of fragment ions (**Figure 32**).

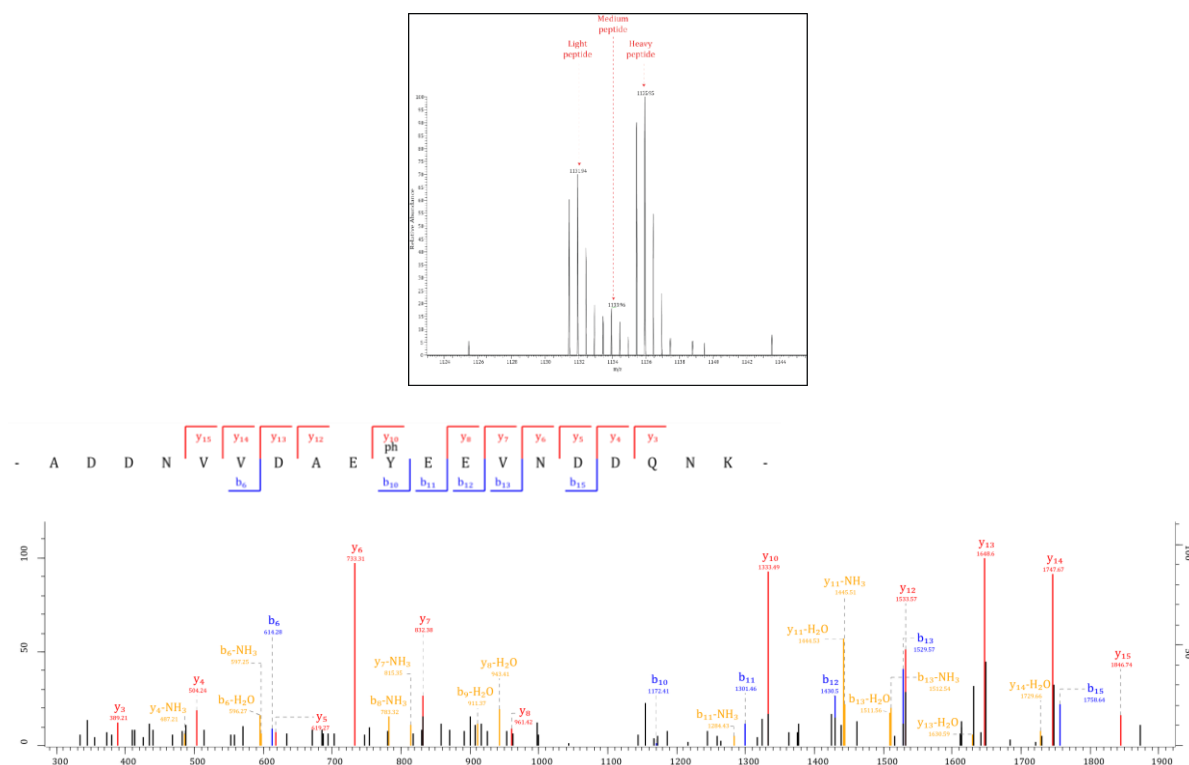


Figure 32: MS² spectrum of DnaK – Survey scan of the corresponding SILAC triplet and MS/MS spectrum of the DnaK peptide depicting the phosphorylation event on Tyr 601.

4.4.3. Validation DnaK Tyr601 as a Substrate of PtkA and PtpZ

In order to validate Tyr601 on DnaK as a substrate of PtkA and PtpZ, further experiments were conducted. DnaK, PtkA and PtpZ were overexpressed and purified in *E. coli* and similar *in vitro* phosphorylation and dephosphorylation assays using γ -³²P ATP were performed, as described above (YkwC). The assays showed that DnaK could not autophosphorylate and that PtkA was able to phosphorylate DnaK and that PtpZ was able to dephosphorylate phosphorylated DnaK *in vitro* (**Figure 33**). This confirmed that DnaK

was indeed a substrate of both enzymes. To determine the influence of phosphorylation of Tyr601 on DnaK activity, the chaperon activity of DnaK was tested. For this purpose the phosphoablative mutant protein DnaK_Y601F was synthesized.

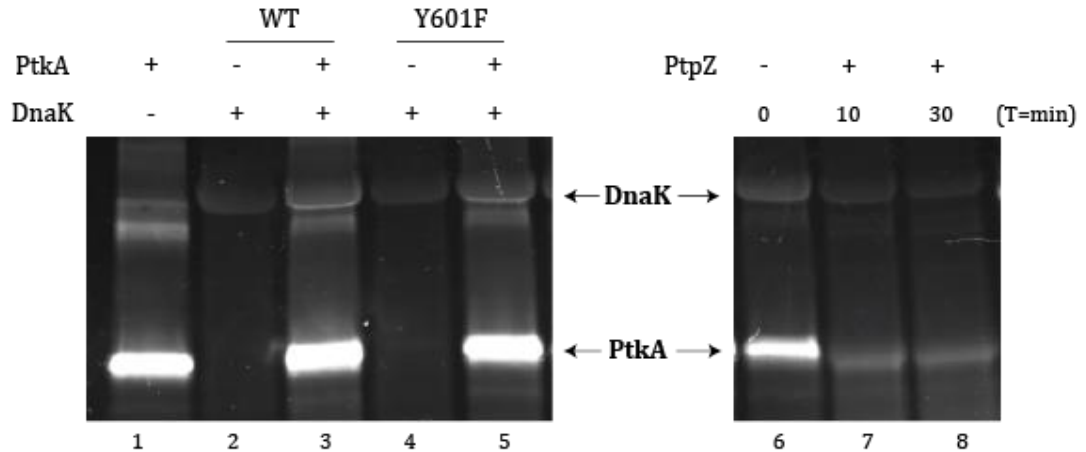


Figure 33: In vitro phosphorylation of DnaK by PtkA and dephosphorylation by PtpZ. Autoradiography showing in vitro phosphorylation/dephosphorylation assays of DnaK labeled with ^{32}P - γ -ATP. Autophosphorylated PtkA and phosphorylated DnaK are indicated by arrows. For the phosphorylation assay, PtkA alone and DnaK and DnaK_Y601F alone were incubated for 30 min (lane 1, 2 and 4); DnaK and DnaK_Y601F was incubated with PtkA (lane 3 and 5). For the dephosphorylation assay, DnaK was incubated with PtkA for 30 min (lane 6), and then with PtpZ for 10 and 30 min respectively (lane 7 and 8).

Denatured-LDH incubated with DnaK and DnaK_Y601F showed diminished ability of the DnaK phosphoablative mutant to refold LDH and restore its activity (**Figure 34**). Additionally, a heat-shock assay conducted at 58 °C with the WT, ΔptkA and DnaK^{Y601F} strains showed decreased cellular survival of the strains lacking the phosphorylation event on Tyr601 of DnaK (ΔptkA and DnaK^{Y601F}) (**Figure 35**) indicating the importance of phosphorylation of DnaK on its activity and overall cell survival under heat shock. To ensure that the effect shown by the mutant was not due to protein misfolding induced by mutation, the circular dichroism (CD) spectra of WT and Y601F were compared. The spectra of both proteins were very similar, indicating that the mutation did not disrupt the protein structure (**Figure 36**).

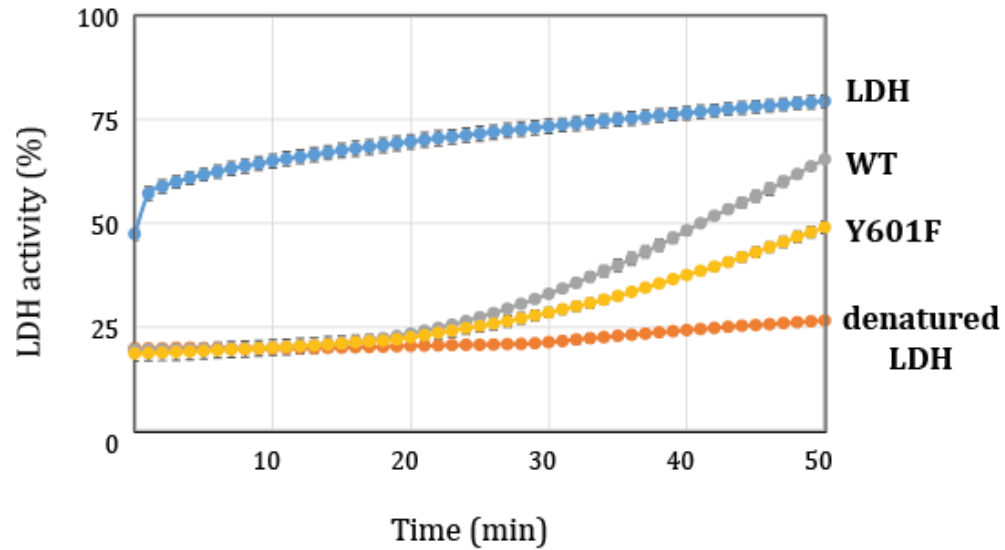


Figure 34: In vitro chaperon activity of DnaK – assayed with LDH (lactate dehydrogenase) depicted decreased heat resistance and diminished folding of denatured LDH by the phosphoablative DnaK mutant, DnaK_{Y601F}.

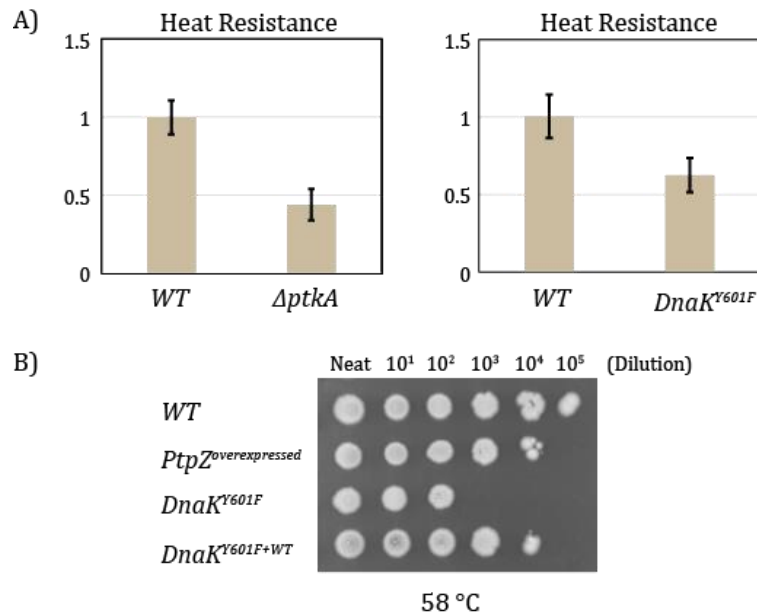


Figure 35: Cellular survival upon heat-shock – A) The $\Delta ptkA$ and DnaK^{Y601F} strains exhibited decreased cellular survival upon heat-shock at 55 °C as compared to the *B. subtilis* WT strain. B) A serial dilution spotting heat-shock assay depicted the PtpZ-overexpression strain and the DnaK^{Y601F} strain to be the most susceptible at 58 °C. The impaired cellular survival of DnaK^{Y601F} under heat-shock could be complemented by the wild copy (DnaK^{Y601F+WT}).

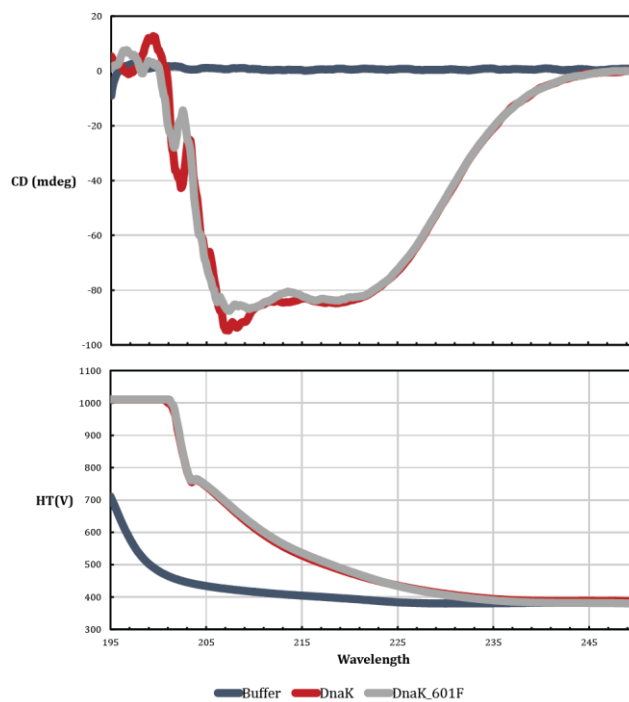


Figure 36: Circular dichroism of DnaK and DnaK_Y601F – A far UV CD spectrum of DnaK (WT) and DnaK_Y601F phosphoablative mutant), measured in the range of 240 nm to 190 nm, denoted a similar secondary structure pattern of all analyzed proteins.

4.5. Analysis of YabT, YbdM, YveL, YfkJ Using the Respective Knock-out Strains

4.5.1. Experimental Design

Phosphoproteomic analysis was done using additional kinase and phosphatase knock-out strains: YabT, YbdM, YveL and YfkJ. Two triple SILAC experiments were conducted – one with the WT, YabT and YbdM (Ser/Thr kinases) and the other with the WT, YveL and YfkJ (Tyr kinase and phosphatase). The WT ($\Delta lysA$) strain was always Light labeled in all the analysis, whereas, the kinase YabT and YveL were Medium labeled and the kinase YbdM and phosphatase YfkJ were Heavy labeled in the respective triple SILAC experiments. Cell cultures labeled with different forms of lysine were harvested during the mid-stationary phase of growth. Proteins extracted from the corresponding SILAC cultures were mixed in equal ratios and digested with endoproteinase Lys-C. The resulting peptide mixtures were subjected to phosphopeptide enrichment using SPE followed by sequential TiO₂ chromatography. In addition pTyr antibody was used specifically for enriching Tyr phosphorylated peptides in the WT, YveL and YfkJ triple SILAC experiment. For proteome analysis, peptides were fractionated by GeLC. All phosphopeptide enriched samples and proteome measurements were analyzed by nanoLC-MS/MS on LTQ-Orbitrap Elite mass spectrometer. Data processing was done using MaxQuant 1.5.1.0.

4.5.2. SILAC-Based Screens

As described earlier the respective WT, YabT, YbdM, YveL and YfkJ strains were inoculated in respective SILAC minimal media and harvested during the mid-stationary phase of growth. Proteome and phosphoproteome measurements were carried out as described earlier.

WT, YabT, YbdM Phosphoproteomics Screen – Quality control check of the incorporation level of the medium and heavy label was found to be above 98 % in both replicates. On the proteome level, 2,238 proteins were identified of which 1,923 were quantified. A total of 122 non-redundant phosphorylation events were identified of which 70 were quantified; 88 sites had a localization probability of ≥ 0.75 and 109 of

them had a PEP ≤ 0.001 . List of all detected proteins and phosphorylation sites is presented in **Supplementary Table 4**. Global phosphoproteomic analysis conducted using knock-out strains of the kinases were normalized against the proteome to identify down regulated events. Five differentially down-regulated events were observed in the $\Delta yabT$ strain and $\Delta ybdM$ strain (**Figure 37**). Among potential substrates of the kinase YabT and YbdM were PtkA, a tyrosine kinase implicated to play a role in biofilm formation; LiaH, a phage shock protein that confers resistance against oxidative stress and cell wall antibiotics. A known phosphorylation event on Ser2 of RecA [45] was identified though not quantified in the $\Delta yabT$ screen, acting as a positive control. Interestingly, in the YbdM kinase knockout, phosphorylation sites were identified which were up-regulated (PtsH, a component of the sugar transferase phosphosystem; DivIVA, a cell division initiation protein) as opposed to being down-regulated upon loss of kinase activity. These phosphorylation events are not direct substrates of the kinase and are likely indirectly regulated by YbdM. Recently, the existence of cross-phosphorylation amongst bacterial kinases was reported [187] and similarly it was observed here that in the absence of the YabT or YbdM kinase, known autophosphorylation events on the kinase PtkA were down-regulated.

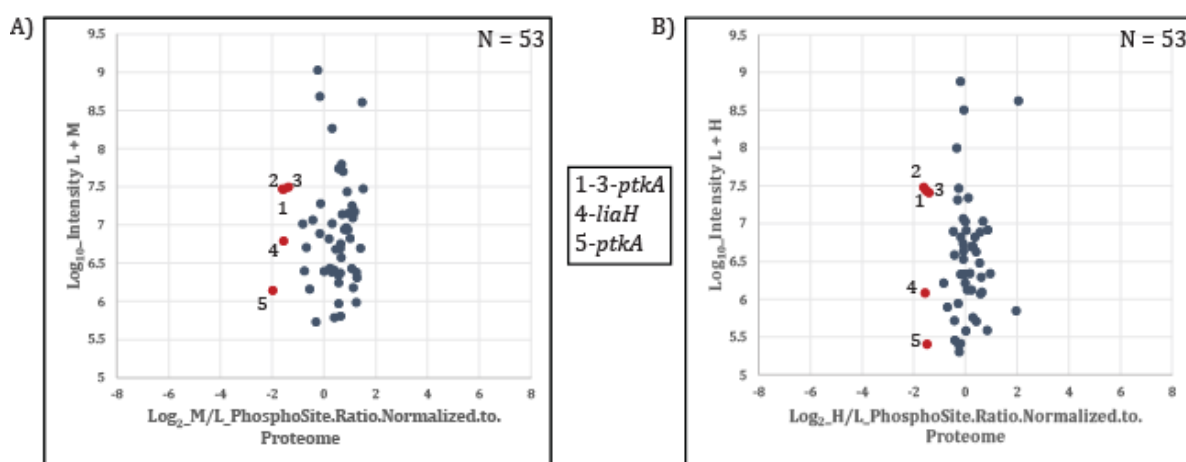
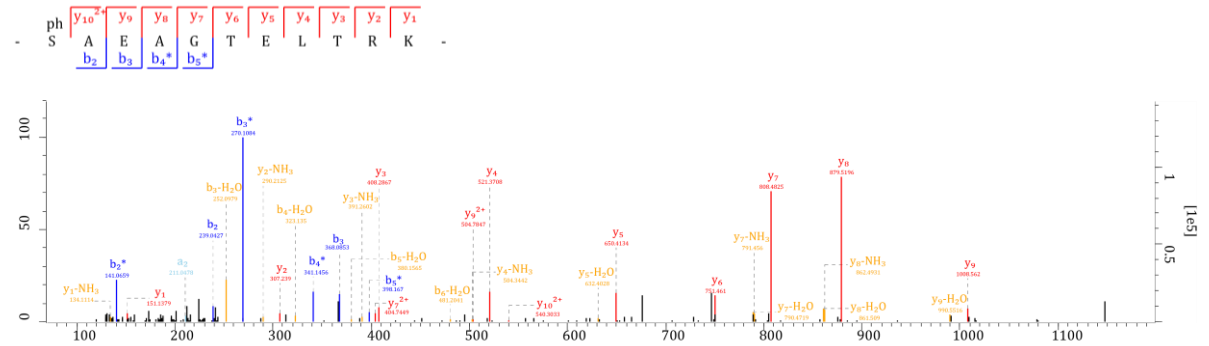


Figure 37: YabT and YbdM phosphoproteomics screen – Scatter plot showing A) $\Delta yabT$ /WT SILAC ratios and B) $\Delta ybdM$ /WT SILAC ratios respectively. \log_2 ratios of the phosphorylation sites are normalized to the corresponding protein and plotted against peptide intensities in \log_{10} scale. Differentially changing SILAC ratios marked in red, present potential substrates of YabT (M/L Ratios) and YbdM (H/L Ratios) respectively.

A novel phosphorylation event on Ser186 on LiaH was seen to be affected in both knock-out screens (YabT and YbdM) making it a putative substrate of both kinases. MS/MS spectrum of the phosphorylated LiaH peptide showed good coverage and annotation of fragment ions (**Figure 38**). Further experiments are required to prove that LiaH is a substrate of both of these kinases.



and ArtQ, an ABC transporter that helps in arginine uptake. Interestingly, in the kinase and phosphatase knockout, phosphorylation sites were identified which had opposing trends - higher in the Δ kinase (SecY, ComGA and ArtQ) and lower (CotG) in the Δ phosphatase respectively.

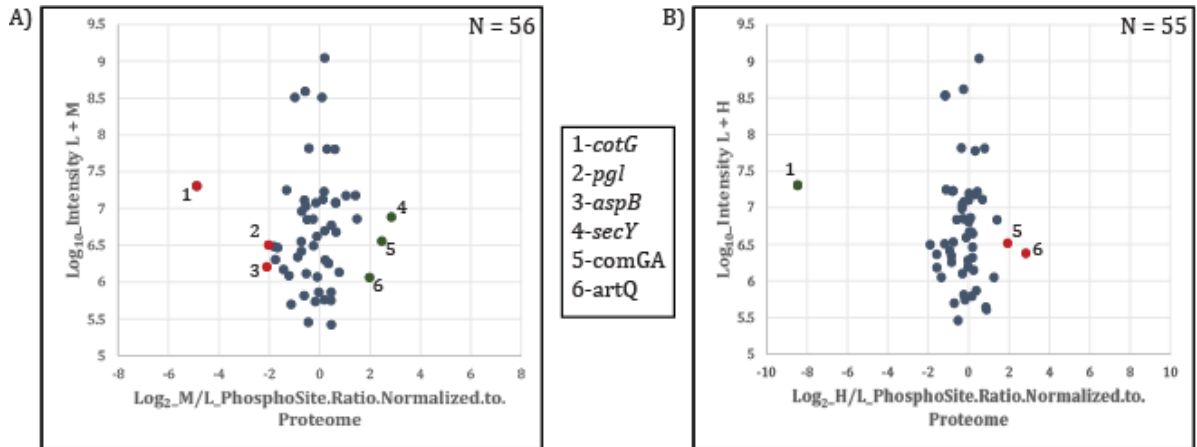


Figure 39: YveL and YfkJ phosphoproteomics screen – Scatter plot showing A) Δ yveL/WT SILAC ratios and B) Δ yfkJ/WT SILAC ratios respectively. Log_2 ratios of the phosphorylation sites are normalized to the corresponding protein and plotted against peptide intensities in Log_{10} scale. Differentially changing SILAC ratios marked in red, present potential substrates of YveL (M/L Ratios) and YfkJ (H/L Ratios) respectively. Those marked in green are differentially regulated but not direct substrates of these enzymes.

These phosphorylation events are probably not direct substrates of the kinase and phosphatase and are likely indirectly regulated by them. An autophosphorylation event Tyr227 was identified on the kinase YveL acting as a positive control. Tyr63 on the AspB protein was seen to be differentially down-regulated in the YveL phosphoproteomics screen making it a potential putative substrate of the kinase. MS/MS spectral annotation showed good coverage of the sequence and fragment ions (**Figure 40**). Similarly, Tyr196 on ComGA was affected in the YfkJ phosphoproteomics screen making it a potential putative substrate of the phosphatase. MS/MS spectral annotation showed good coverage of the sequence and fragment ions (**Figure 41**). Further biochemical experimentation is required to prove that both these phosphorylated proteins are substrates of the respective enzymes.

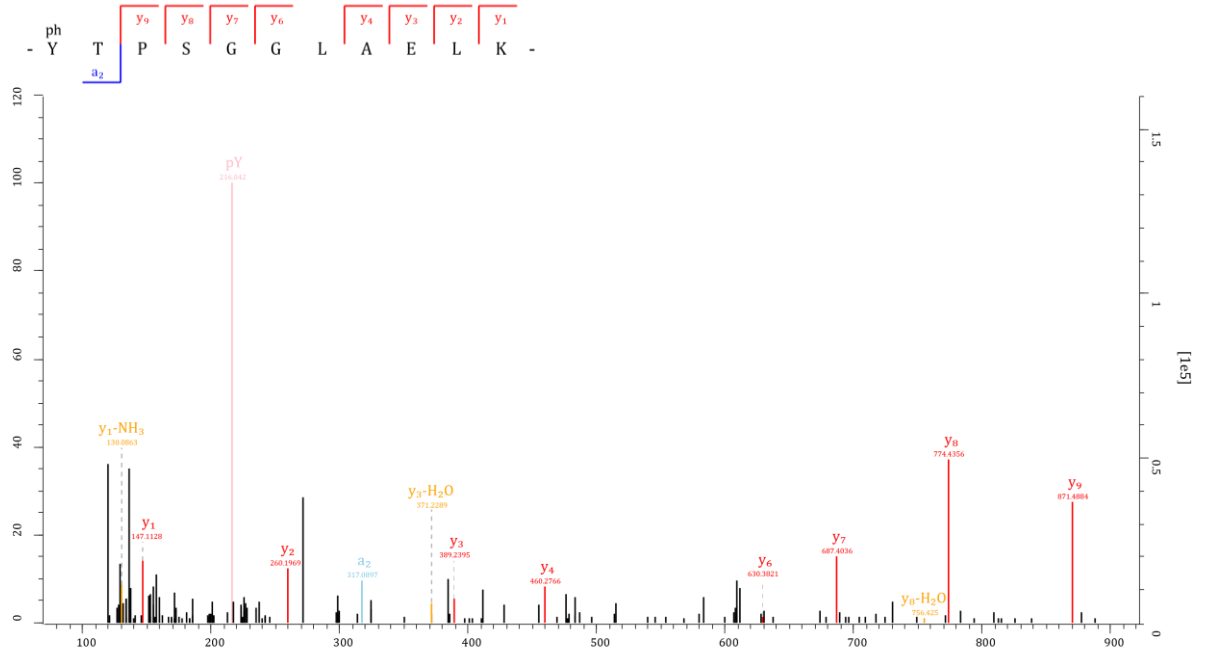


Figure 40: MS/MS spectrum of the AspB peptide phosphorylated on Tyrosine 63.

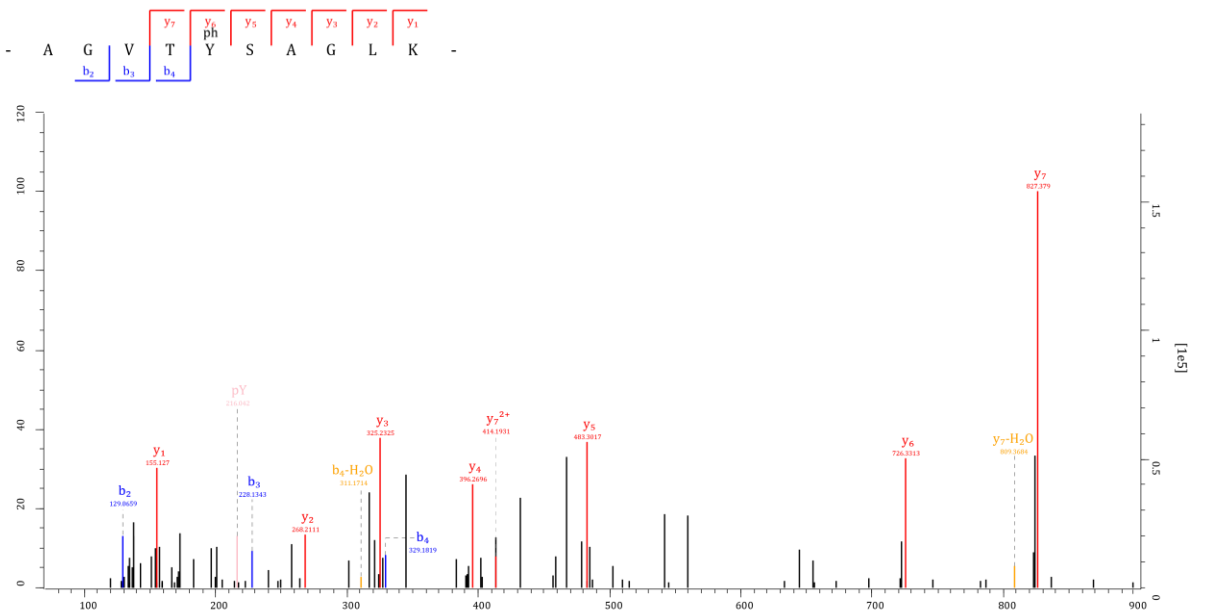


Figure 41: MS/MS spectrum of the ComGA peptide phosphorylated on Tyrosine 196.

4.6. Analysis of the Kinase/Phosphatase Network of *B. subtilis*

In order to get an overview and be able to map phosphorylation events dependent on all known STY kinases and phosphatases in *B. subtilis*, data from all the kinase/phosphatase SILAC triplet screens were processed together with the MaxQuant software suite (version 1.5.1.0). A total of 416 non-redundant phosphorylated sites were identified on 227 proteins and 278 sites were quantified (**Figure 42**). 319 of them had a localization probability ≥ 0.75 and 204 of them had a PEP score ≤ 0.001 . Distribution of all phosphorylation events on serine, threonine and tyrosine residues was found to be in the ratio of 76% : 18% : 6% respectively. Gene ontology analysis, using the KEGG Mapper (<http://www.genome.jp/kegg/mapper.html>), categorized the 227 proteins broadly into 14 groups (**Figure 43**), the most enriched being proteins involved in metabolic processes – carbon, amino acid and fatty acid metabolism. ABC transporters or proteins involved in secretion of enzymes or toxins were also seen to be phosphorylated. Since all the screens were completed using stationary phase cultures, proteins involved in polysaccharide biosynthesis, stress response, motility and chemotaxis were also found to be phosphorylated.

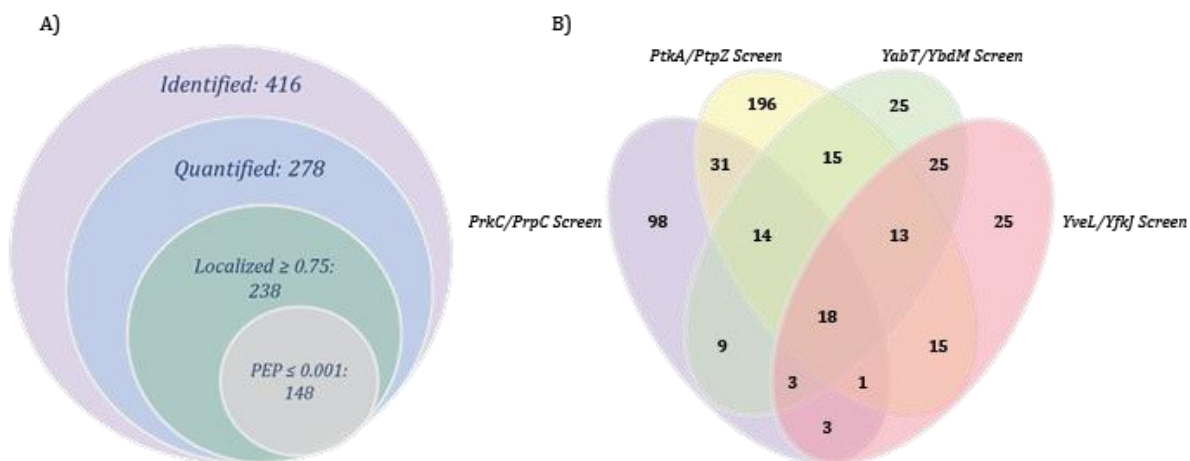


Figure 42: Venn diagram depicting intersection of data – A) A total of 416 phosphorylated sites were identified of which 278 were quantified, 238 were localized with a probability of ≥ 0.75 and 148 passed the PEP threshold of ≤ 0.001 . B) Overlap between identified phosphorylated sites amongst the different kinase/phosphatase screens.

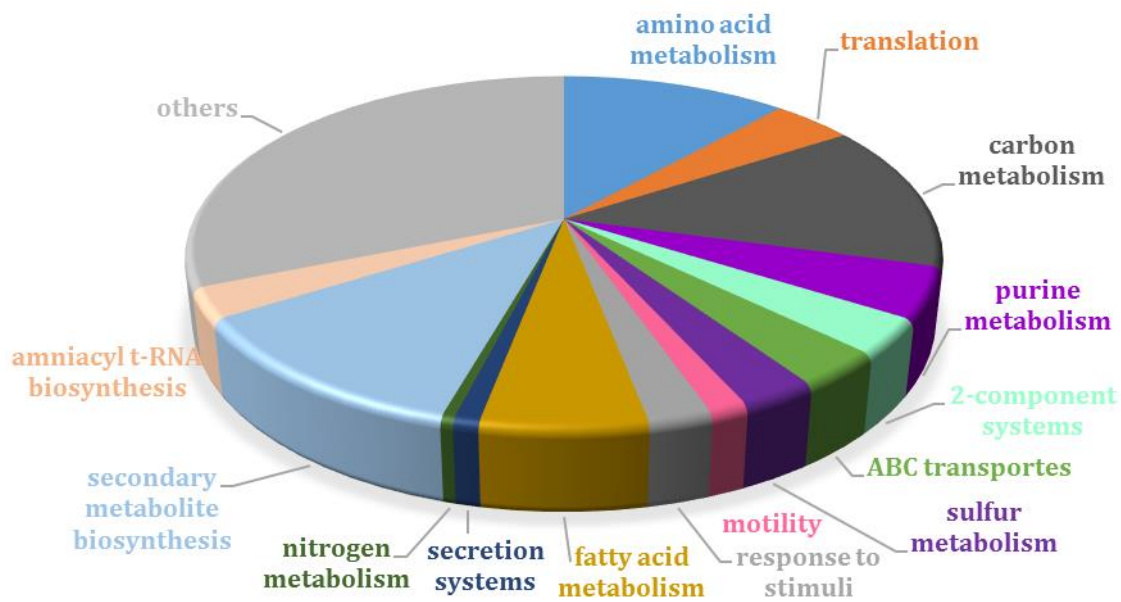


Figure 43: GO-enrichment – Pie-chart representing GO annotation of phosphorylated proteins from kinase/phosphatase screens.

A cutoff of 1.5 on the Log₂ scale was chosen and phosphorylation events that were up or down regulated in each of the SILAC screens were analyzed. A total of 18 sites were found to be commonly affected in more than one kinase or phosphatase knock-out condition and 27 phosphorylation sites were exclusively regulated in either one of the knock-out conditions. Spore coat proteins CotB and CotG were seen to be multiply phosphorylated on serine/threonine residues and phosphorylation events on two sites S253 and S254 of CotB were found to be affected in all knock-out conditions. A network was mapped based on the all the interactors of the kinases and phosphatases analyzed which were within the set threshold. All regulated interactors were further analyzed using STRING (<http://string-db.org/>) (**Supplementary Table 6**) and a composite functional network is represented in **Figure 44**. A large number of proteins amongst the most regulated interactors were involved in metabolic processes. 39 % of the phosphorylation events were on uncharacterized or putative proteins whose function is unknown. Some known phosphorylation events such as Y227, Y225 and Y228 on PtkA; S46 on HPr; S680 on FusA; S100 on GlmM were identified as strongly regulated.

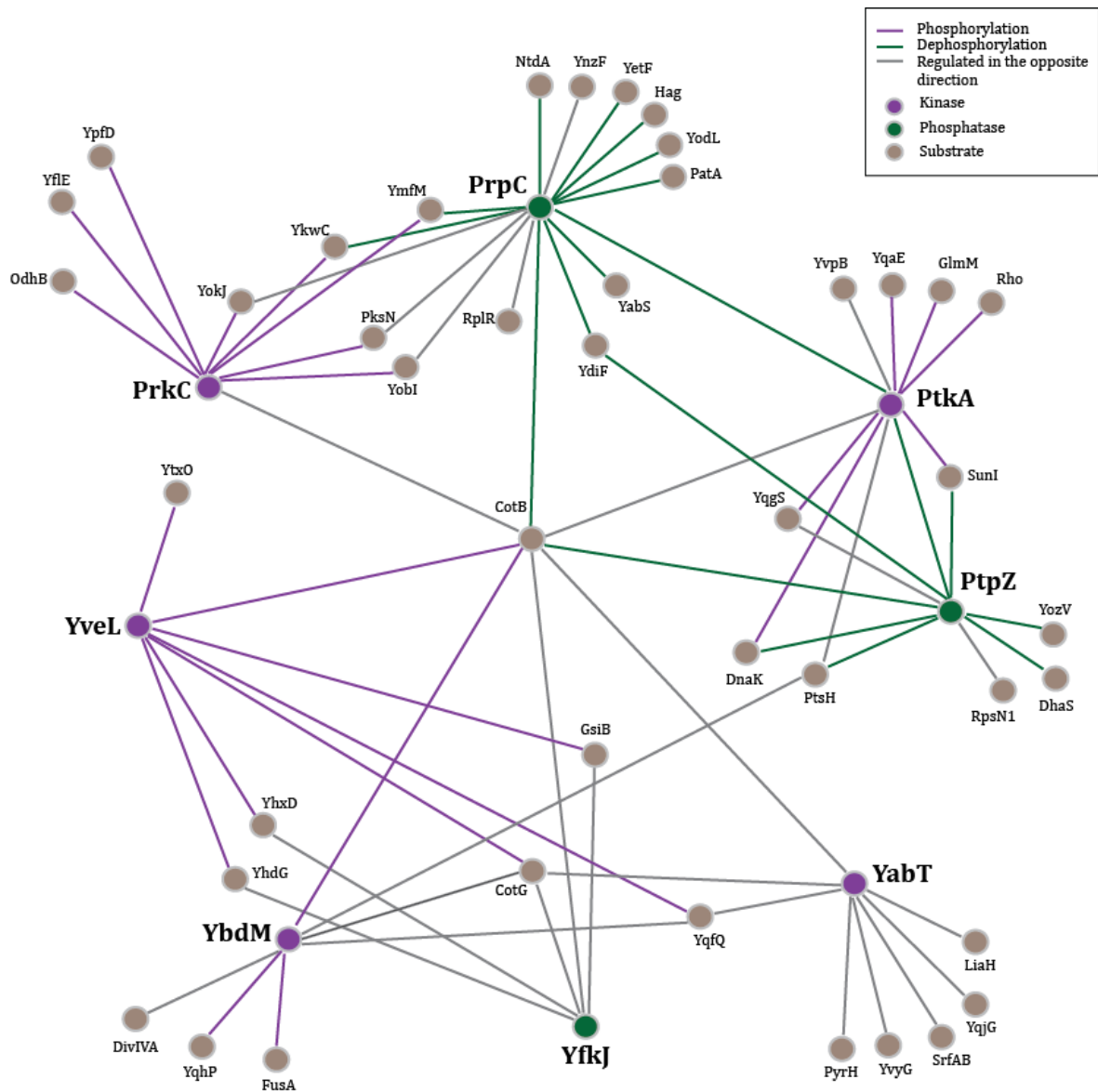


Figure 44: An interactome of all regulated putative substrates (direct or indirect) of all analyzed kinases and phosphatases. Kinases and their phosphorylation events are depicted in purple and phosphatases and their dephosphorylation events are depicted in dark green. Proteins that are regulated by the respective kinases and phosphatases in the opposite direction are depicted in light grey. Respective substrates are depicted in brown.

4.7. A Comprehensive Map of the *B. subtilis* Phosphoproteome

To construct the most comprehensive map of the *B. subtilis* phosphoproteome to date, data from all phosphorylation enrichment experiments (a total of 7 experiments and 632 raw files) performed in course of this thesis were processed with the MaxQuant software suite 1.5.1.0. Global analysis of phosphorylation events in *B. subtilis* led to the identification of 995 phosphorylation sites, present on 484 proteins. 793 of them had a localization probability of ≥ 0.75 and 470 of them had a PEP ≤ 0.001 . 270 of these 995 sites come from the dataset from Soufi *et al* (unpublished). The distribution of serine : threonine : tyrosine sites was in the ratio 58.3% : 21.5% : 20.2%. 397 sites on 225 proteins that were localized and had a low PEP score, were loosely categorized into 15 protein classes using PANTHER (<http://pantherdb.org/>) (**Figure 45**). The largest fraction were transferases or oxidoreductases followed by hydrolases, transcription factors and nucleic acid binding proteins.

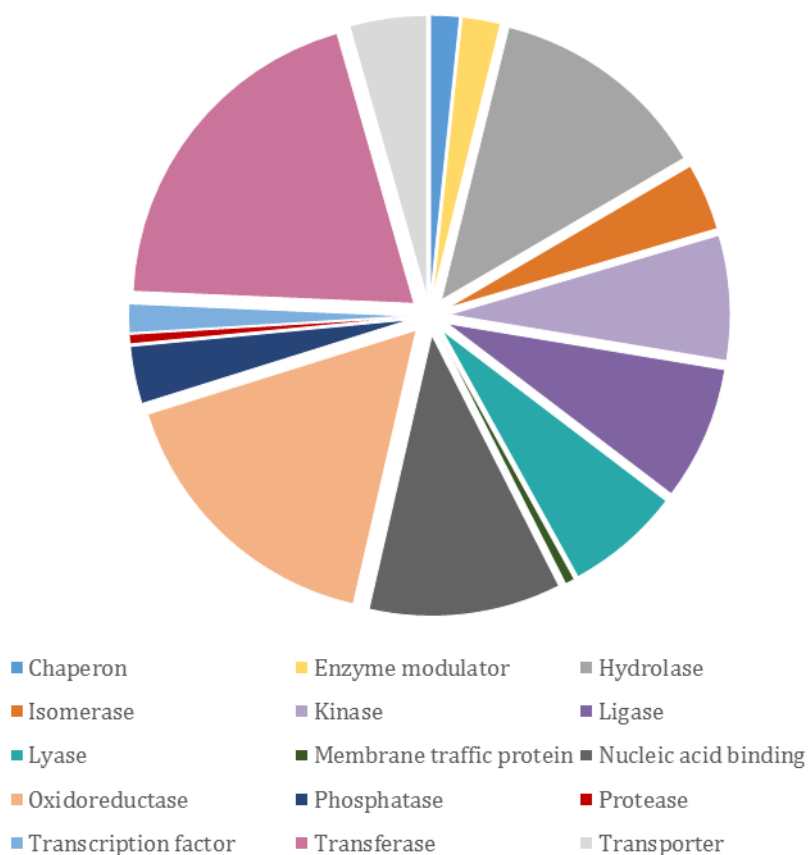


Figure 45: GO-enrichment – Pie-chart representing 15 protein classes of the 397 localized and high confidence sites.

PANTHER GO_Slim analysis ($P < 0.05$) (**Table 10**) showed a >5 fold enrichment of phosphorylated proteins from the carbon metabolic processes and >4 fold enrichment of proteins from cellular amino acid metabolism - 27 from carbon metabolism, 27 from amino acid metabolism, 12 from purine metabolism and 10 from fatty acid metabolism. Phosphorylated proteins identified were further mapped onto KEGG pathways (<http://www.genome.jp/kegg/mapper.html>) and an overview is represented in **Figure 46**.

Table 10: Statistical Overrepresentation Test using PANTHER

PANTHER GO-Slim Biological Process	<i>B. subtilis</i> Reference Proteome	Analyzed <i>B. subtilis</i> Phosphoproteome			
	#	#	expected	Fold Enrichment	P value
tricarboxylic acid cycle	22	6	0.84	> 5	1.95E-02
cellular amino acid catabolic process	54	9	2.06	4.37	2.26E-02
translation	98	16	3.74	4.28	1.17E-04
monosaccharide metabolic process	59	9	2.25	4	4.27E-02
cellular amino acid biosynthetic process	150	20	5.72	3.49	1.20E-04
generation of precursor metabolites and energy	113	14	4.31	3.25	1.09E-02
carbohydrate metabolic process	370	34	14.12	2.41	1.13E-04
cellular amino acid metabolic process	401	36	15.3	2.35	8.80E-05
primary metabolic process	1266	99	48.31	2.05	1.16E-14
metabolic process	1504	114	57.39	1.99	8.80E-18
protein metabolic process	394	29	15.03	1.93	3.96E-02
Unclassified	2565	40	97.88	0.41	0.00E+00

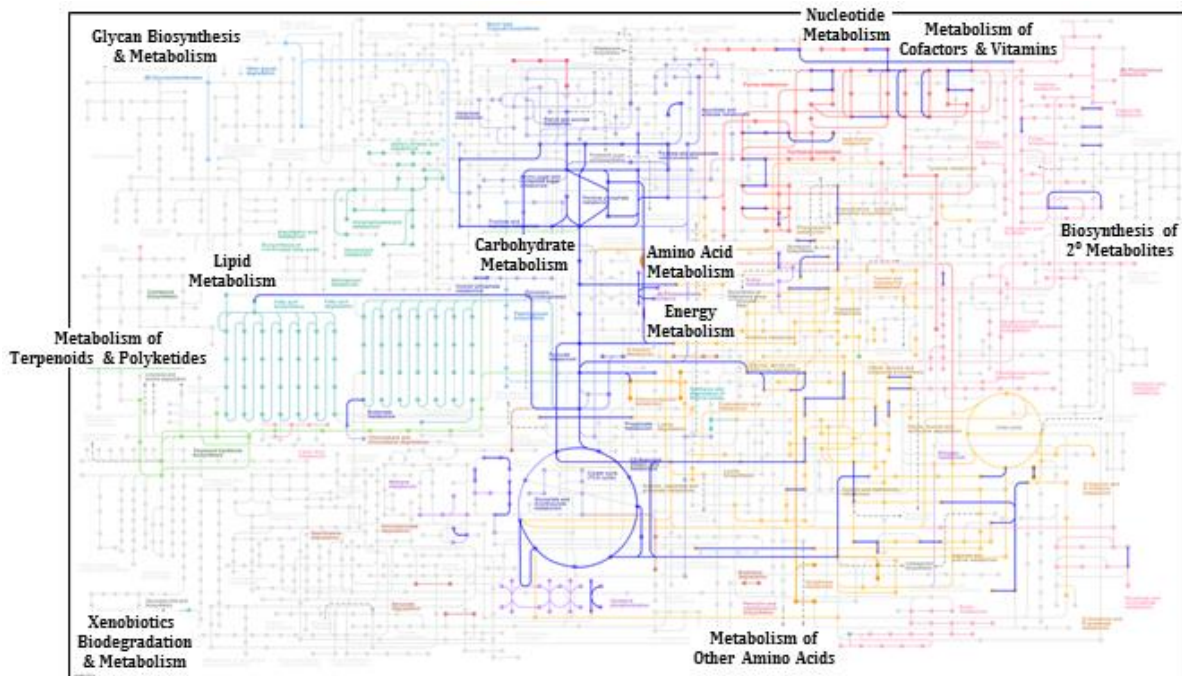


Figure 46: Phosphoproteins mapped onto KEGG pathways. The blue lines indicate the proteins identified during the comprehensive analysis of the *B. subtilis* phosphoproteome that play a role in various metabolic pathways.

Motif-x [188, 189] was used for the purpose of motif extraction in order to analyze the identified phosphorylated peptides for the presence of characteristic motifs. Only localized sites were chosen for the analysis. A sequence window containing 6 amino acids upstream and downstream of the phosphorylated residue was chosen as data input. An occurrence threshold of 20 was set and a default P-value threshold (for the binomial probability) of 10^{-6} was maintained. The sequences were tested against a background *B. subtilis* database (Reference ID 1423) obtained from UniProt. Peptides phosphorylated on serine, threonine and tyrosine residues were analyzed individually by defining the central character as S, T or Y respectively. Motif analysis conducted for threonine phosphorylated peptides resulted in two preferential patterns (**Figure 47**) with low motif scores. No significant motif was found amongst tyrosine phosphorylated peptides. Similar preliminary analysis conducted for serine phosphorylated peptides displayed a preference for Lys (K) or Ser (S) at close proximity to the phosphorylated residue (**Figure 47**). A total of 10 motif patterns were detected during the analysis of the 475 localized phosphorylation serine peptides. **Table 11** outlines the details of the results obtained.

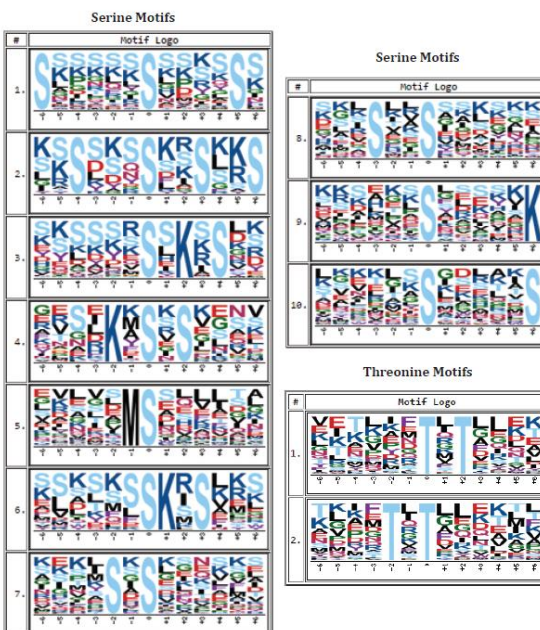


Figure 47: Motif analysis for serine and threonine phosphorylated peptides – Overrepresented motifs surrounding the phosphorylated serine and threonine peptides were detected using Motif-x.

Table 11: Motif extraction results obtained from the analysis for threonine motifs, where the motif score corresponds to statistically significant motifs; foreground matches (FM) and background matches (BM) refer to number of peptides containing the given motif in the analyzed dataset and the *B. subtilis* proteome respectively; foreground size (FS) and background size (BS) indicated the total number of peptides contained in the analyzed dataset and the *B. subtilis* proteome respectively; fold increase refers to the enrichment of the extracted motifs in comparison to the background.

Residue	#	Motif	Motif Score	FM	FS	BM	BS	Fold Increase
Serine	1	S.....S.....S	31.48	48	937	408	73808	9.27
	2	..S...S..S..S	34.07	20	889	66	73400	25.02
	3S.K.S..	30.95	41	869	359	73334	9.64
	4K.S.S....	19.67	24	828	332	72975	6.37
	5MS.....	11.33	54	804	1652	72643	2.95
	6SK.S...	17.73	20	750	255	70991	7.42
	7S.S.....	9.04	98	730	4977	70736	1.91
	8	...S..S.....	10.23	90	632	4508	65759	2.08
	9S.....K	7.39	72	542	4129	61251	1.97
	10S.....S	6.96	66	470	4049	57122	1.98
Threonine	1T.T....	7.90	41	269	3664	63998	2.66
	2T.T.....	9.61	40	228	3435	60334	3.08

4.8. Putative Interactors of Phosphorylated Motifs

4.8.1. Quality Control

4.8.1.1. Peptide Check

Nine pairs of peptides in their unmodified, as well as phosphorylated form were synthesized in the laboratory of Dr. Hubert Kalbacher, University of Tübingen, Germany. Lyophilized peptides were dissolved in water to a concentration of 1 mg/mL and sonicated whenever required to bring the peptides completely into solution. A quality control check of the synthesized peptides was carried out by direct infusion (as described in the Methods section 3.2.19.1.2.) to verify the amino acid sequence of the respective peptides and to ensure the presence of the phosphate moiety on the correct residues in the respective phosphorylated peptides. **Figure 48** depicts a representative example of the verification conducted. **Supplementary Figure 4(B-R)** represent the rest of the eight annotated spectra of the synthesized peptides.

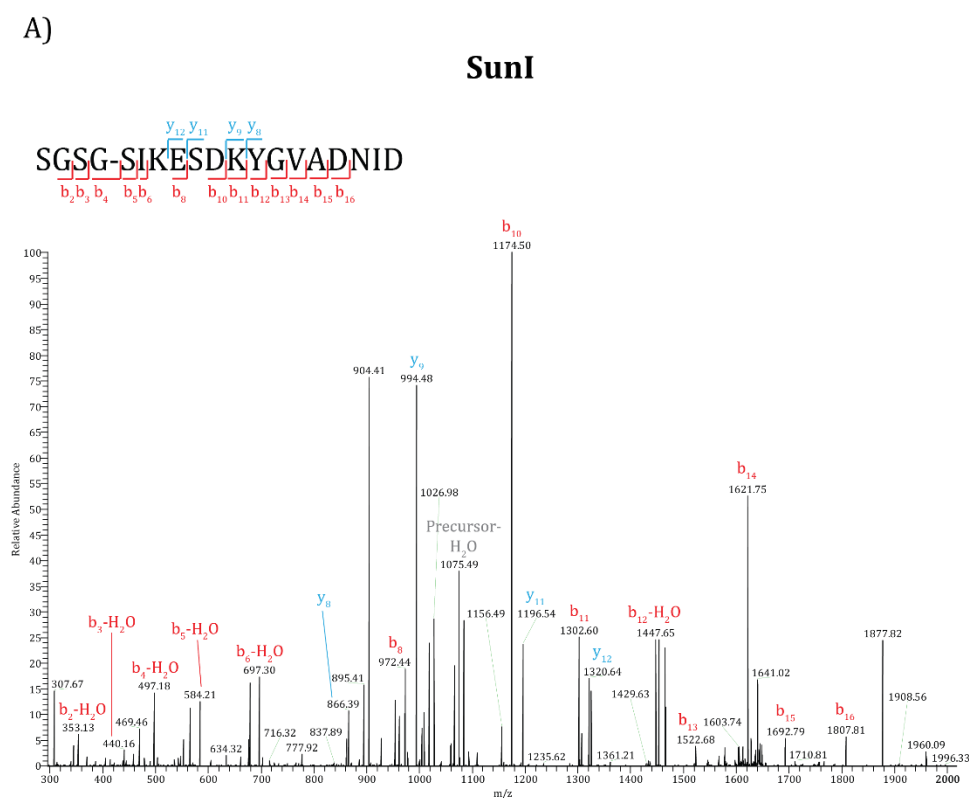


Figure 48: Quality Control – Annotated MS/MS spectra of the unmodified SunI peptide.

4.8.1.2. Mixing and Incorporation Check

All experiments were conducted using the WT ($\Delta lysA$) strain inoculated in Light and Heavy SILAC minimal media and harvested during the mid-stationary phase of growth. As a quality control step, Lys8-labeled cultures were digested with Lys-C and analyzed for the incorporation of the heavy labeled amino acid. An incorporation rate of $\geq 98\%$ was observed (**Figure 49**). Additionally, unlabeled and labeled protein extracts from each experiment were mixed in equal amounts and analyzed on the mass spectrometer for mixing errors. Errors were then corrected for based on their median. A representative example is depicted in **Figure 50**.

4.8.1.3. Retention of the Phosphate moiety

Due to the possible presence of active phosphatases in the lysate, it was also important to ensure the retention of the phosphate group on the modified synthetic peptides after interaction with the cell lysate. Preservation of the phosphate moiety on the phosphorylated peptide was observed in all cases.

Table 12 shows peptide evidences of the retention of the phosphorylation event (in comparison to the unmodified peptide counterpart) on a serine, threonine and tyrosine residue, chosen as a representative example.

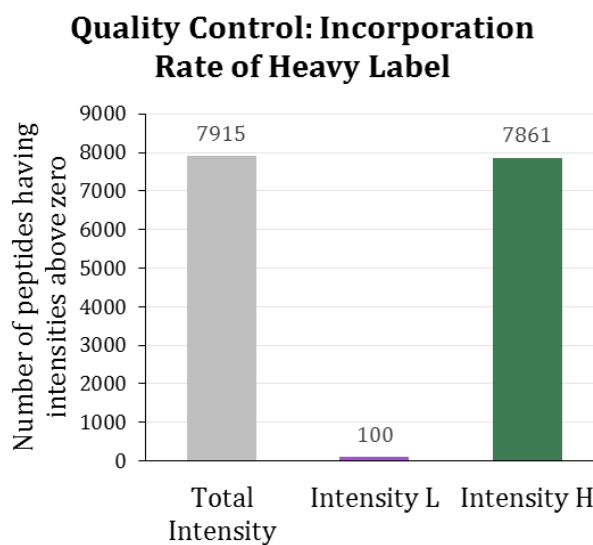


Figure 49: Quality Control – Level of incorporation of the heavy label was checked for all experiments. Shown here, is a single representative example of a bar plot showing 98 % incorporation of the heavy label.

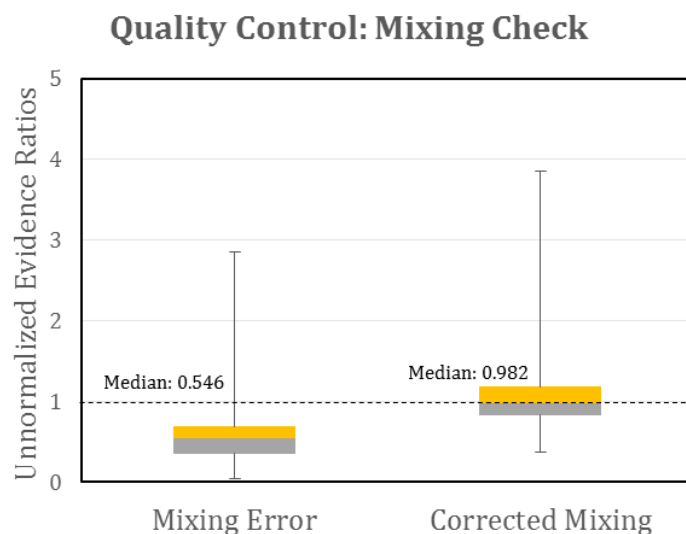


Figure 50: Quality control – The figure depicts a representative boxplot of the mixing check done based on the protein concentrations measured by Bradford (left) and mixing check that has been corrected for (right).

Table 12: Peptide evidences showing retention of phosphorylation on STY residues post pull-down. Mod (modifications) are denoted by “-” for unmodified peptides and “+” for modified peptides.

RESIDUE	SEQUENCE	MOD (-/+)	CHARGE	m/z	INTENSITY	
					Light	Heavy
Serine	_TVNLKSIMG VM SL_	-	2	696.88	8.13E+08	8.0E+06
	TVNLKS(ph)I MGVMSL	+	2	736.87	5.74E+08	1.3E-07
Threonine	_EEMTKAIIPII K_	-	2	636.76	4.37E+07	3.99E+06
	EEMT(ph)KA IPIIK	+	2	676.78	3.51E+06	1.23E+03
Tyrosine	_ESDKYGVAD NID_	-	2	663.30	2.09E+09	0
	ESDKY(ph)G VADNID	+	2	703.28	1.45E+09	0

4.8.2. Identifying Interactors of Phosphorylation Motifs

In order to detect and explore putative interactors that bind specifically to STY phosphorylated residues synthetic peptides were used as affinity baits (**Figure 51**). A total of nine pairs of peptides (unmodified peptides and their respective phosphorylated

counterparts) were analyzed (**Table 9** in section 3.2.30 lists all the synthesized peptides along with their sequences). The unmodified as well as the modified (phosphorylated) peptides synthesized were attached to a biotin moiety through an SGSG-linker. Thus immobilized peptide was interacted with cell lysate either in the “Light” or “Heavy” SILAC state. At every instance, a “Forward” experiment was carried out wherein the unmodified peptide was interacted with “Light” labeled cell lysate and the modified peptide was interacted with “Heavy” labeled cell lysate and vice versa in case of the “Reverse” experiment. Subsequently, affinity purification of interacting partners was carried out using Streptavidin beads followed by analysis on the LTQ-Orbitrap XL.

Interactors of the phosphorylated residue are represented in **Figure 52(A-I)** in the form of a scatter plot. Protein ratios of the “Forward” experiment are plotted against the protein ratios of the corresponding “Reverse” in the Log₂ scale. SILAC was employed to differentiate the non-specific binders, which form a cluster around the zero on the axis, from the specific binders (outliers).

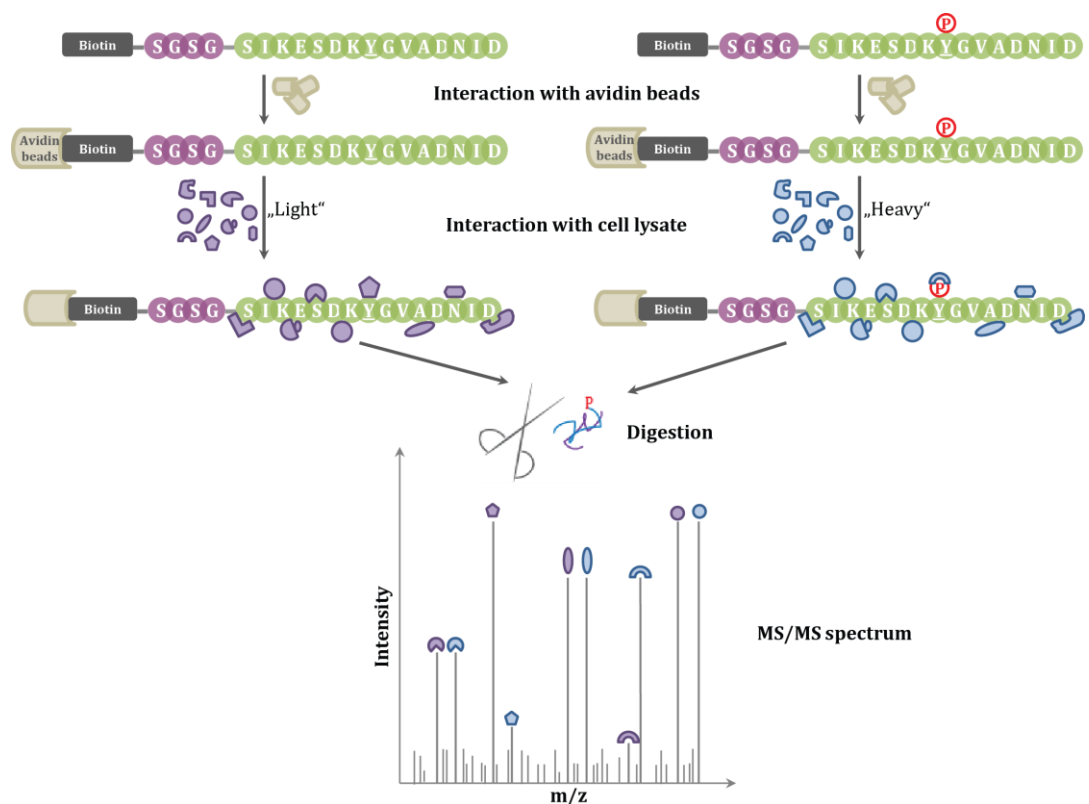


Figure 51: The peptide-protein interaction screen employed in order to detect interactors of *STY* phosphorylated residues.

A) PrkC_01 (RIPFDGESAVSIALK)

A total of 613 proteins were identified of which 375 were quantified in both experiments. Analysis of the affinity enrichment experiments revealed no significant outliers in case of PrkC_01, a serine/threonine kinase involved in sporulation (**Figure 52A**).

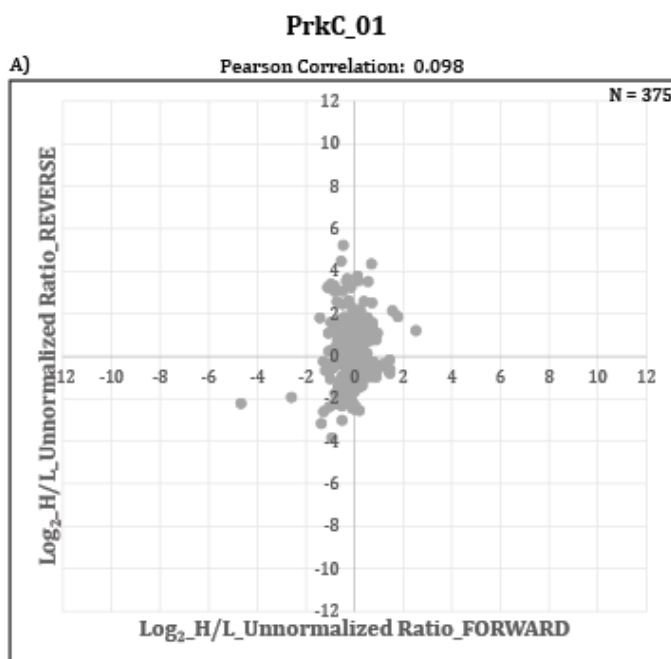


Figure 52A: PrkC_01 interaction screen – Scatter plot depicting the H/L unnormalized ratio of the “Forward” experiment plotted against the H/L unnormalized ratio of the “Reverse” experiment in the Log_2 scale, where N denotes the number of protein groups quantified. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be 0.098.

B) YmfM (NTYHDDVSEKISGMN)

A total of 990 proteins were identified of which 800 were quantified in both experiments. Analysis of the affinity enrichment also revealed no significant outliers in case of YmfM, an essential membrane protein involved in cell shape determination (**Figure 52B**).

C) DnaK (DNVVDAEYEEVNDDQ)

A total of 479 proteins were identified of which 299 were quantified in both experiments. Similarly, analysis of the affinity enrichment revealed no significant outliers in case of DnaK, a heat-shock chaperon (**Figure 52C**).

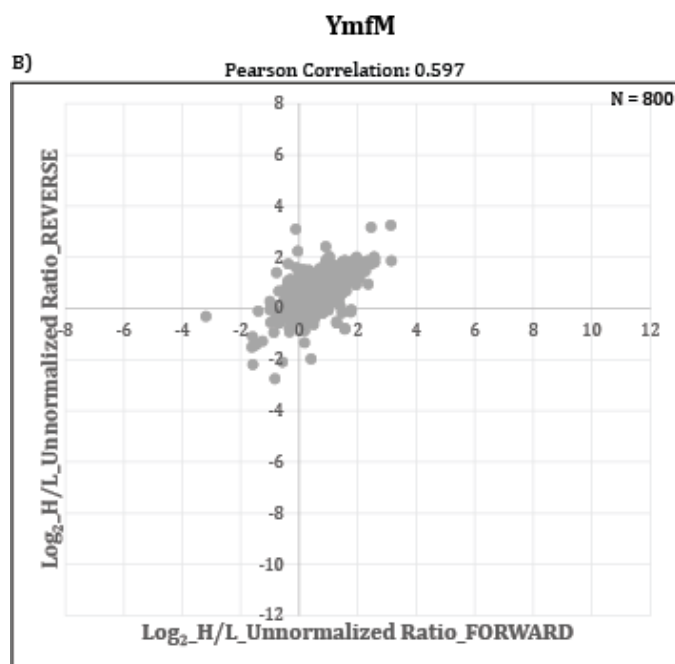


Figure 52B: YmfM interaction screen – Scatter plot depicting the H/L unnormalized ratio of the “Forward” experiment plotted against the H/L unnormalized ratio of the “Reverse” experiment in the Log_2 scale, where N denotes the number of protein groups quantified. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be 0.597.

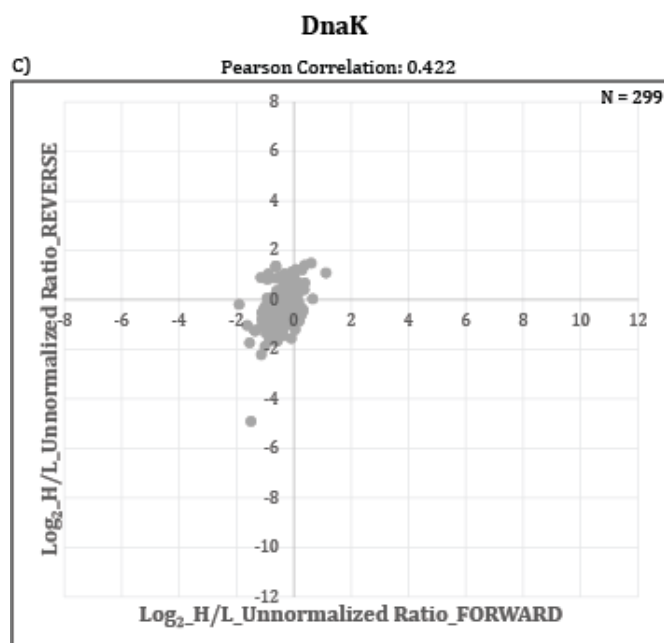


Figure 52C: DnaK interaction screen – Scatter plot depicting the H/L unnormalized ratio of the “Forward” experiment plotted against the H/L unnormalized ratio of the “Reverse” experiment in the Log_2 scale, where N denotes the number of protein groups quantified. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be 0.422.

D) YkwC (EENSGTQSIYKLWVK)

Synthetic peptide interaction screens conducted with the YkwC peptide, an oxidoreductase involved in valine metabolism, yielded 1085 protein identifications of which 819 were quantified. Of the quantified proteins, 15 of them bound more to the phosphorylated serine residue (pS218) and one to the unmodified peptide (**Figure 52D**). Nine of the 15 binders were uncharacterized proteins with six of them having DUF domains (Domain of Unknown Function).

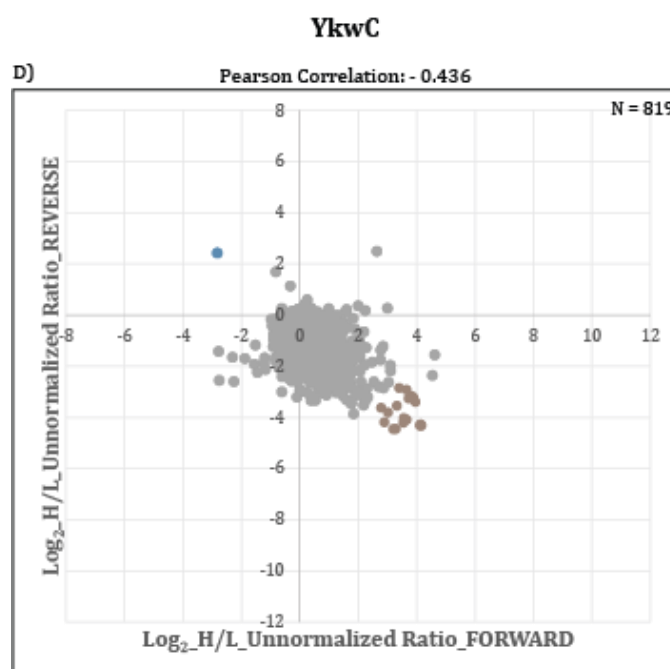


Figure 52D: YkwC interaction screen – Scatter plot showing binders of the unmodified peptide (in blue) and binders of the phosphorylated peptide (in brown). H/L unnormalized ratios of the “Forward” experiment is plotted against the H/L unnormalized ratios of the “Reverse” experiment in the Log_2 scale, where N denotes the number of protein groups. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be - 0.436.

E) HPr (GKTVNLKSIIMGVMSL) –

A total of 851 proteins were identified in the HPr, a component of the sugar phosphotransferase system, interaction screen and 367 of them were quantified. Amongst the interactors of the Ser46 peptide of HPr were PycA, a pyruvate carboxylase and AccB, acetyl-CoA carboxyl protein. There was a general shift in the plot towards the left indicating more binding to the unmodified peptide of HPr (**Figure 52E**). In general,

proteins involved in stress response, NAD biosynthesis and protein folding bound more to the unmodified peptide.

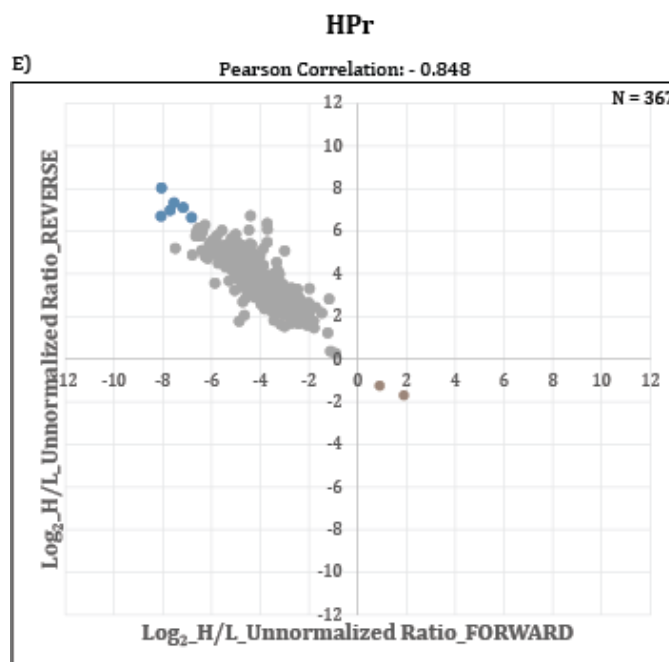


Figure 52E: HPr interaction screen – Scatter plot showing binders of the unmodified peptide (in blue) and binders of the phosphorylated peptide (in brown). H/L unnormalized ratios of the “Forward” experiment is plotted against the H/L unnormalized ratios of the “Reverse” experiment in the Log_2 scale, where N denotes the number of protein groups. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be - 0.848.

F) SunI (SIKESDKYGVADNID) –

Affinity enrichments with SunI, a bacteriocin immunity protein, resulted in 586 identifications and 447 quantifications. Five of them were observed to bind more to the unmodified peptide which included proteins involved in metabolic processes such as carbohydrate (*pdhA*) or amino acid metabolism (*leuA*), an uncharacterized protein with a DUF domain (*ytxG*), a chemotaxis protein (*mcpB*) and a peptidase (*pepT*). Additionally, PtkA, a tyrosine kinase that was found to phosphorylate SunI on Y57 in our previous studies, was observed to bind more to the unmodified peptide of SunI, thus acting as a positive control and validating our strategy (**Figure 52F**).

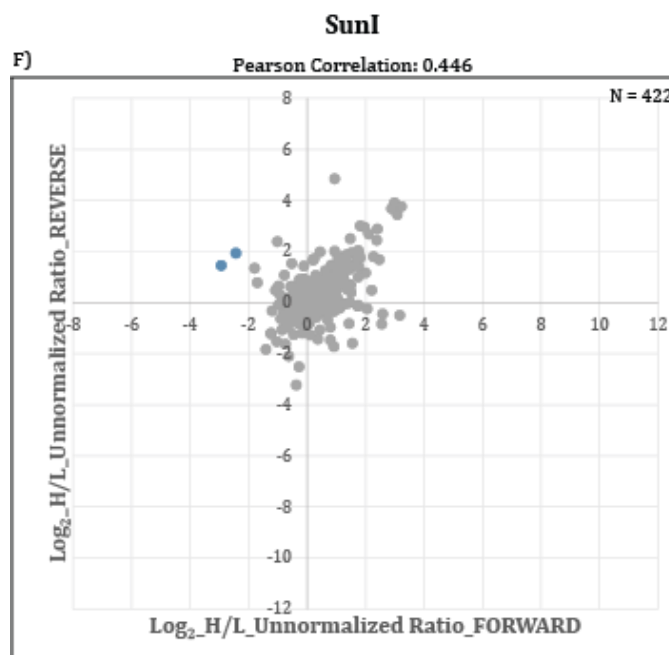


Figure 52F: SunI interaction screen – Scatter plot showing binders of the unmodified peptide (in blue) and binders of the phosphorylated peptide (in brown). H/L unnormalized ratios of the “Forward” experiment is plotted against the H/L unnormalized ratios of the “Reverse” experiment in the Log₂ scale, where N denotes the number of protein groups. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be 0.446.

G) PtkA (SKHSEYGYGTKDNF) –

811 proteins were identified during the PtkA interaction screen and 421 of them were quantified. In case of the PtkA peptide, interestingly, a large number of proteins seemed to bind to the unmodified peptide of PtkA as compared to the phosphorylated tyrosine peptide (pY227) causing a shift of the distribution towards the left top quadrant (**Figure 52G**). 80 % of the binders of the unmodified peptide were ribosomal proteins – 50s, 30s and associated proteins.

H - I) PrkC_02 (IQEDEEMTKAIIK) and PrkC_03 (SSTTITHTNSVLGSV) –

A total of 968 and 1107 proteins were identified by the PrkC_02 (**Figure 52H**) and PrkC_03 (**Figure 52I**) interaction screen respectively while 722 and 790 proteins were quantified in each of the screens respectively. In contrast to PtkA, a large number of ribosomal proteins (~50 %) bound to the phosphorylated threonine peptide of PrkC_02

and PrkC_03 (pT290 and pT167, respectively) causing a shift in the distribution towards the right bottom quadrant.

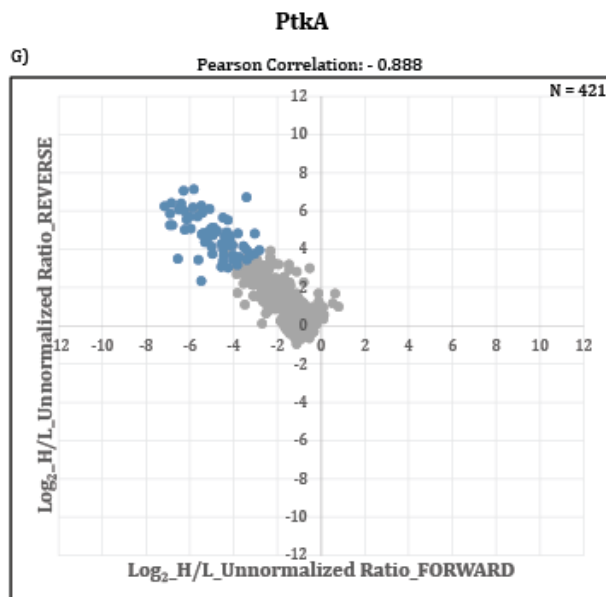


Figure 52G: PtkA interaction screen – Scatter plot showing binders of the unmodified peptide (in blue) and binders of the phosphorylated peptide (in brown). H/L unnormalized ratios of the “Forward” experiment is plotted against the H/L unnormalized ratios of the “Reverse” experiment in the Log₂ scale, where N denotes the number of protein groups. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be - 0.888.

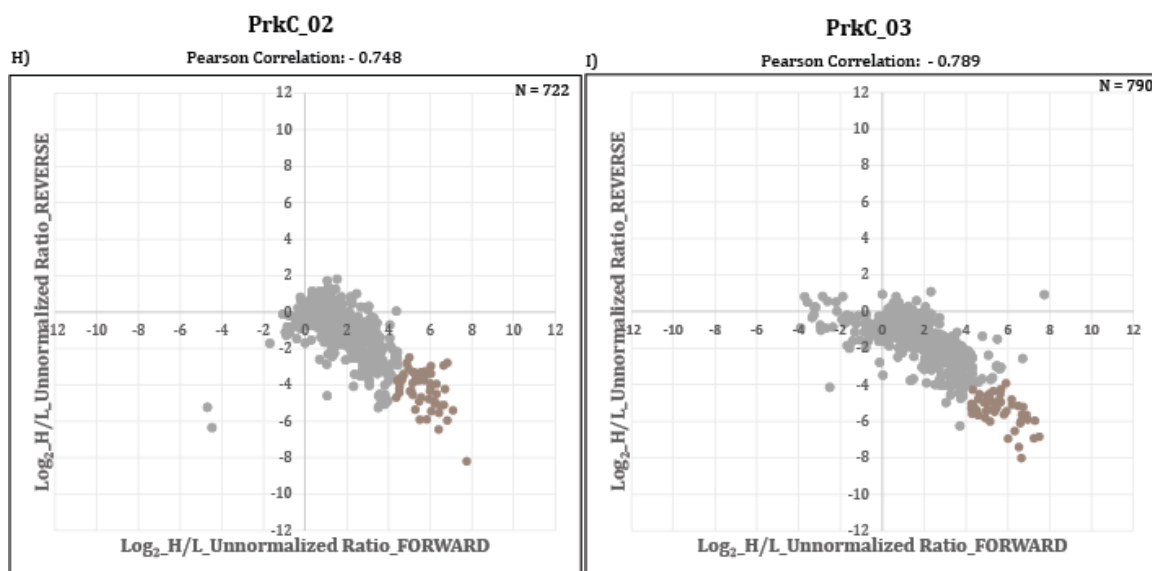


Figure 52H-I: PrkC_02 & PrkC_03 interaction screen – Scatter plot showing binders of the unmodified peptide (in blue) and binders of the phosphorylated peptide (in brown). H/L unnormalized ratios of the “Forward” experiment is plotted against the H/L unnormalized ratios of the “Reverse” experiment in the Log₂ scale, where N denotes the number of protein groups. The Pearson Correlation between the “Forward” and the “Reverse” experiment was found to be - 0.748 and - 0.789 respectively.

5. Discussion

5.1. Dynamics of the Proteome and Phosphoproteome of *B. subtilis*

This section has been published in "Ravikumar V, et al, Quantitative phosphoproteome analysis of Bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC. Mol Cell Proteomics, 2014. 13(8): p. 1965-78".

The main objective of this study was to analyze the global dynamics of proteins and phosphorylation events as the *B. subtilis* culture undergoes different stages of growth. In the course of this study 2,264 proteins (55.2% of the *B. subtilis* theoretical proteome) were detected, which represents one of the most exhaustive proteomic datasets of this model organism.

In order to validate the quantitative proteome dataset, the expression correlation of subunits of known protein complexes was analyzed, as they are expected to be regulated in a similar fashion (**Supplementary Figure 5**). Expression of proteins involved in sporulation initiation (SpoOKD - SpoOKB - SpoOKE - SpoOKC - SpoOKA) and proteases belonging to the same proteasomal complex (ClpC - ClpX - ClpP) were not significantly altered from T1 through T5, and remained highly correlated. Proteins belonging to the flagellar hook assembly system (FlgE - FlgL - FlgK) were highly correlated and up regulated in T5. This is coherent with nutrient deprivation at T5, which induces chemotaxis in order to look for an additional source of subsistence. The respiratory complex (CtaC - CtaF) was up regulated during T3 after which it steadily dropped down towards T5 in order to conserve energy during adverse conditions.

As expected, the phosphoproteome of *B. subtilis* was highly dynamic during growth. Phosphorylation of glycolytic enzymes and other enzymes of the central carbon metabolism increased strongly in early and late stationary phase. These highly phosphorylated enzymes included: Eno (enolase), Fba (aldolase), Pgi (glucose-6-phosphate isomerase), GapA (glyceraldehyde-3-phosphate dehydrogenase), CitK (2-oxoglutarate dehydrogenase), Ykjl (6-phosphogluconate dehydrogenase), CitB (citrate hydro-lyase), SerA (3-phosphoglycerate dehydrogenase), CitC (isocitrate dehydrogenase) SucC (succinyl-CoA ligase) and TktA (transketolase). It has been reported by Pietack *et al.* [9] that phosphorylation of several glycolytic enzymes did not

affect their enzyme activity directly. However, glycolytic flux is known to be controlled in some cases at the level of enzyme assemblies that channel substrates [190]. Such multi-enzyme assemblies have been shown to exist in humans, and their assembly is regulated by protein phosphorylation [191]. The high correlation of phosphorylation dynamics between different glycolytic enzymes in *B. subtilis* opens up the possibility that a similar regulation may exist in bacteria. Phosphorylation of a number of house-keeping proteins involved in translation also increased towards the stationary phase. These include ribosomal proteins (JofD, RpsH, RplL, RplN, RplR, RpmD), chaperones (GroEL) and elongation factors (FusA, Tsf, TufA). The competence protein ComJ (Nin) known to inhibit the DNase activity of NucA [192] exhibited a significant drop of phosphorylation occupancy in the early stationary phase, corresponding to the “time window” for competence development. This suggests that phosphorylation may have a functional consequence. An interesting differential phosphorylation pattern was detected in the purine synthesis pathway, where PurA and PurM were more strongly phosphorylated towards stationary phase, while phosphorylation of PurH diminished. Since these proteins act in concert, and PurH and PurM are encoded by the same operon, these phosphorylation events may be expected to have opposite effects on the target enzymes. Proteins involved in amino acid biosynthesis were strongly phosphorylated in the late stationary phase. Among them, MetC and ThrC are involved in biosynthesis of methionine and threonine and IlvA, IlvC, LeuA and LeuB are involved in the biosynthesis of branched-chain amino acids and are induced during amino acid starvation [193]. IlvC has already been found to be phosphorylated at residue Ser338 in a previous study [184]; here new phosphorylated residues were detected. Of particular interest, proteins involved in phosphate metabolism exhibited strong differential phosphorylation. Phosphate binding ABC transporter protein PstS has previously been identified in the phosphoproteome of *B. subtilis* [194], and is known to be strongly induced by phosphate starvation [195]. The current study indicates that PstS is very strongly dephosphorylated in early and late stationary phase (where phosphate depletion is expected), suggesting that the dephosphorylated state may be more active in phosphate import. Similarly, in the case of PhoA, alkaline phosphatase 4, it is strongly induced on phosphate starvation [196] and

was seen to be dephosphorylated in stationary phase, indicating that the non-phosphorylated form of this enzyme may be more active.

Phosphorylation of the recombinase RecA at Ser2, which has been shown to contribute to a check-point in spore development [45], was depleted in the transition phase and non-detectable throughout the stationary phase. This is consistent with the fact that this phosphorylation is catalyzed by the kinase YabT, whose expression is tightly shut until the onset of spore development [45]. All these examples testify to the capacity of quantitative phosphoproteomics to detect relevant signaling and regulatory events *in vivo*.

The phosphoproteome dataset is particularly interesting from the perspective of *B. subtilis* protein-tyrosine kinases, which catalyze the least abundant of phosphorylation events detected here. Phosphorylation of the putative BY-kinase PtkB (EpsB, YveL) was detected, for which the kinase activity has been previously suggested but never demonstrated *in vitro* [167, 197]. This enzyme exhibited a very interesting probable autophosphorylation profile, with the site occupancy peak in the transition phase. This could be correlated to the involvement of PtkB in biofilm development, which has been previously suggested [198]. The main *B. subtilis* BY-kinase PtkA was found to be phosphorylated by an unknown Ser/Thr kinase and dephosphorylated by the phosphatase PrpC, which is a novel example of cross talk between different families of bacterial kinases. In terms of BY-kinase substrate phosphorylation, phosphorylation of the global gene regulator AbrB at Tyr39 was detected, and the occupancy of this phosphorylation site increased towards stationary phase. Phosphorylation site of AbrB is in its DNA-binding domain and phosphorylation could hence be expected to impair DNA binding. This would be consistent with abolishing AbrB-mediated repression concomitant to the exit from exponential phase [199]. A similar mechanism has already been described for another PtkA substrate, the transcription regulator FatR [200]. Phosphorylation of known PtkA substrates (Eno, Asd and YjoA [51]) on Ser or Thr residues was also detected, indicating a cross-talk between BY-kinases and Hanks-type kinases not only on the kinase but also on the substrate level.

5.2. Novel Serine, Threonine and Tyrosine Kinase and Phosphatase Substrates

A part of this section has been published in "Ravikumar V, et al, Quantitative phosphoproteome analysis of Bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC. Mol Cell Proteomics, 2014. 13(8): p. 1965-78".

SILAC experiments with kinase and phosphatase knockouts reveal physiological substrates for these enzymes, by pin-pointing phosphorylation sites for which the occupancy changes dramatically in the respective knockout mutants [36]. In this study, the validity of this principle was demonstrated for the first time in bacteria. The phosphorylation of YkwC Ser281 dropped significantly in $\Delta prkC$ and increased in $\Delta prpC$ suggesting that it is a substrate for both these enzymes. This hypothesis was validated by showing *in vitro* phosphorylation and dephosphorylation of YkwC by PrkC and PrpC, respectively. Moreover, PrkC dependent phosphorylation was shown to regulate the activity of YkwC *in vivo* and *in vitro* [43]. YkwC belongs to the 3-hydroxyisobutyrate dehydrogenase family that is involved in the breakdown of L-valine. L-valine has been found to play an important role in spore germination [201, 202]. Phosphorylation of YkwC by PrkC inhibits its oxidoreductase activity, thus possibly maintaining a suitable L-valine concentration required for spore germination amongst other factors. Similarly, the phosphorylation of DnaK on Tyr601 by PtkA and dephosphorylation by PtpZ, helps regulate DnaK activity *in vivo* and *in vitro*. DnaK belongs to the Class I heat-shock proteins that act as molecular chaperons that bind to partially unfolded proteins and promote either renaturation or degradation through specific mechanisms [203]. The involvement of heat-shock proteins in biofilm formation has been shown previously [204, 205]. Phosphorylation of DnaK by PtkA enhances its chaperon activity, conceivably promoting biofilm formation (manuscript in preparation).

SILAC-based phosphoproteomics screens conducted with other STY kinases and phosphatases, performed in this study, led to the identification of respective putative substrates. LiaH, involved in resistance against oxidative stress and cell wall antibiotics, was identified as a putative substrate of both Ser/Thr kinases YabT and YbdM. AbrB, a transcriptional regulator playing a crucial role in regulating gene expression during

transition from growth phase to stationary phase in *B. subtilis* [206], represses the promoter P_{liaI} , preventing transcription of the *liaIH* operon [207]. YabT and YbdM phosphorylate AbrB *in vitro*, thus abolishing DNA-binding [46]. It was shown before that the promoter P_{liaI} is induced by phosphorylated LiaR (a 2-component response regulator), leading to the expression of the *liaIH-liaGFSR* locus, with *liaIH* being the most induced genes [208]. Additionally, upon envelop stress, LiaH is induced in large amounts and recruited to the membrane subsequently adopting a protective role by maintaining the integrity of the cytoplasmic membrane [209]. Taking into account all these facts, it can be contemplated that YabT and YbdM phosphorylate LiaH on Ser186, either directly or indirectly through phosphorylated LiaR, leading to LiaH induction and antibiotic resistance. Similarly, AspB, an aspartate transaminase involved in the biosynthesis of aspartate, was identified as a putative substrate of the tyrosine kinase YveL, confers resistance against oxidative damage. It was shown in *Corynebacterium glutamicum* and in *Escherichia coli*, that L-methionine and L-cysteine are derived from the amino acid aspartate. Aspartate is transferred into homoserine and this is followed by integration of reduced sulphur into homoserine and subsequent formation of homocysteine. Homocysteine is then either methylated leading to methionine production or converted to cysteine by reverse transsulfuration [210]. In *Candida albicans*, methionine and cysteine pathways were seen to be up-regulated during biofilm formation. Disruption of genes involved in this pathway was shown to diminish biofilm formation [211]. Thus it can be conjectured that YveL, a tyrosine kinase involved in biofilm formation, phosphorylates AspB and activates it, leading to aspartate production and subsequently increasing cysteine and methionine concentration, aiding in biofilm formation. Likewise, ComGA, a late competence protein that binds and transports transforming DNA, was identified as a potential substrate of the tyrosine phosphatase YfkJ. ComGA has been previously reported to be phosphorylated on arginine residues [212]. During oxidative stress, a Spx-dependent global transcriptional control is induced [213]. Spx was also found to be a negative effector of competence by forming a quaternary complex with ClpC, MecA and ComK preventing transcription of the genes involved in DNA uptake [214]. In this regard, two regulation-via-inhibition models can be hypothesized during the conditions of oxidative stress – 1) Phosphorylated ComGA on Tyr186 could be

antagonistic to phosphorylation on arginine due to a conformational change, thus preventing competence; 2) During oxidative stress, the cell undergoes rapid changes in physiological properties leading to reduced cell membrane permeability [215], which would hamper DNA uptake. Phosphorylation of ComGA on tyrosine could inhibit its activity thus preventing competence and helping in energy conservation. While LiaH, AspB and ComGA were identified as potential targets of the respective STY kinases and phosphatases (YabT, YbdM, YveL and YfkJ) in the phosphoproteomics screens and the effect of phosphorylation could be speculated with respect to their role on the bacterial physiology, further biochemical experimentation is required to prove the above discussed hypothesis.

Overall, these screens demonstrate the tremendous potential of quantitative phosphoproteomics in deciphering regulatory networks in bacteria. Bacterial kinases and phosphatases are thought to be promiscuous, and the SILAC experiments with $\Delta prkC$, $\Delta prpC$, $\Delta ptkA$, $\Delta tptZ$, $\Delta yabT$, $\Delta ybdM$, $\Delta yveL$, $\Delta yfkJ$ yielded perhaps less substrates than expected. However it must be mentioned that this experiment was performed only at one time point on the growth curve, and one can reasonably expect to detect more substrate by changing the growth conditions. In conclusion, this study reveals important emerging features of bacterial regulatory networks based on protein phosphorylation and opens numerous leads for physiological validation of the dataset.

5.3. Comprehensive Map of *B. subtilis* Phosphoproteome

We were interested in mapping the entire phosphoproteome of *B. subtilis* in order gain an insight into the metabolic pathways regulated due to phosphorylation of key serine, threonine and tyrosine residues in the bacterium. A total of 995 non-redundant sites on 483 proteins were identified, of which 397 passed the stringent localization probability (≥ 0.75) and PEP (≤ 0.001) threshold set. This dataset significantly increases the number of previously reported phosphorylation events in *B. subtilis*. Keeping in trend with other studies, 58.3 % of the identified sites were on serine residues. Interestingly, in contrast to the general distribution of phosphorylation on threonine and tyrosine residues, 21.5 % and 20.2 % phosphorylation events were detected on threonine and tyrosine residues respectively (214 sites on threonine and 201 sites on tyrosine). This speaks in favor of conducting additional enrichments of phosphorylated peptide using specific antibodies in order to increase detection of phosphorylated tyrosine. Enrichment of the 397 localized sites present on 225 proteins using PANTHER displayed three main classes of biological processes (P value < 0.05): a) cellular amino acid metabolic processes (8.80×10^{-5}); b) carbohydrate metabolism (1.13×10^{-4}); c) generation of precursor metabolites and energy (1.09×10^{-2}). Regulation via phosphorylation of vital proteins that form a part of the carbohydrate, amino acid, fatty acid or purine metabolism pathways are known to essential [9]. Several examples of crucial members from these metabolic pathways were observed to be phosphorylated in our dataset. A large number of enzymes involved in glycolysis or gluconeogenesis and Krebs's cycle such as glyceraldehyde 3-phosphate dehydrogenase, lactate dehydrogenase, glucose 6-phosphate isomerase, isocitrate dehydrogenase, succinyl-CoA synthetase, 2-oxoglutarate dehydrogenase were identified to be phosphorylated on critical residues. Similarly, enzymes involved in the biosynthesis of amino acids such as tyrosine, methionine, aspartate, glycine, arginine, glutamate and threonine were found to be phosphorylated. Novel phosphorylation events on methionine synthase, MetE; aspartate transaminase, AspB; N-acetylglutamate synthase, ArgJ; glutamate synthase, GltA; threonine synthase, ThrC were identified. Furthermore, in addition to the two known phosphorylation events on Ser2 and Thr4 on chorismate mutase (AroA), three other sites – Ser91, Ser72 and Tyr82 were identified.

Maintenance of the redox balance in a cell is essential and occurs via controlled intracellular ATP concentrations. F-type ATPases drive the interconversion between ADP and ATP. Eight members of the F1-ATPase complex were identified as phosphorylated indicating the importance of phospho-regulation of the operon. Synthesis of purine nucleotides is regulated by a 12-gene cluster [216], three of which (*purA*, *purH*, *purM*) were found to be phosphorylated in our dataset. Other regulatory proteins involved in catabolism or interconversion of purine nucleotides such as xanthine phosphoribosyltransferase (*xpt*) or adenine deaminase (*adeC*), previously not reported to be phosphorylated were identified to be modified on STY residues in our dataset. STY phospho-regulation of two-component and phosphotransferase systems was also observed. During conditions of phosphate limitation, the OmpR regulon becomes activated and is involved in the acquisition of phosphate. Phosphate assimilation occurs via PhoR, the two-component sensor kinase that in turn recruits PhoB, PhoP and other members of the alkaline phosphatase gene family, namely PhoA, PhoD and PstS during phosphate starvation. PstS and PhoB have been previously reported to be phosphorylated [194]. In this study, PhoA and PhoD were additionally found to be phosphorylated on serine and threonine residues. Other proteins involved in two-component system regulations such as hag (part of the flagella regulon involved in motility and chemotaxis), KinC (two-component sensor kinase involved in sporulation) and MalR (two-component response regulator involved in malate utilization) were also found to be phosphorylated. The current dataset supports the emerging concept of cross-talk among different modification systems and establishes the importance of STY phosphorylation in gene regulation as well. Similarly, proteins that form a part of the phosphotransferase system and are involved in fructose or cellobiose uptake and phosphorylation were also found to be phosphorylated. Interestingly, a high number of phosphorylated proteins (N = 40) that were detected were unclassified with no known or a putative function. The largest bacterial phosphoproteomics dataset reported to date is from *Mycobacterium tuberculosis*, wherein the authors detected 516 phosphorylation events on 301 proteins [217]. The data presented here boosts the overall number of phosphorylation events detected in bacterial systems, pointing towards the existence of equally complex and well diversified phosphorylation-regulated pathways as seen in

higher organisms. Motif analysis with the 475 localized phosphorylated serine peptides resulted in 10 potential patterns that passed the set P-value threshold of 10^{-6} . Many of the motif patterns included a lysine or additional serine residues in close proximity to the central phosphorylated serine residue. In some instances the presence of arginine was also noted. These results illustrate the importance of choice of the endoprotease during sample preparation. Commonly used endoprotease trypsin could be suboptimum with respect to digestion of proteins due to the frequency of lysine or arginine residues in proximity of phosphorylated residues in bacterial proteome. Thus use of an appropriate combination of proteases may be required to improve identifications. Similarly, analysis of the 150 localized threonine phosphorylated peptides yielded two potential patterns, suggesting a preference for an additional threonine residue either in +2 or -2 position. The presence of serine or threonine residues often in close proximity to the modified residue, points towards a loose “structural” specificity (as opposed to the known high “functional” specificity of bacterial kinases). While no specific motif pattern was determined amongst the 168 localized tyrosine phosphorylated peptides, in concordance with a recent motif analysis done in *E. coli* [218], lysine was observed to be in close proximity ($\leq \pm 6$) to the phosphorylated tyrosine residue in majority of the sequences.

5.4. Identifying Putative Interactors of Phosphorylation Motifs using Synthetic Peptides

In the present study, synthetic peptides were constructed in their unmodified and corresponding phospho-modified form and used as baits during affinity chromatography. The aim of this study was to identify novel interactors of the respective proteins and possibly detect and explore putative domains and proteins that bind specifically to serine, threonine or tyrosine phosphorylated residues. According to a recent report [219], around 20% of bacterial proteins have “domains of unknown functions” (denoted as DUFs). Unlike in eukaryotes, where phosphorylated residues often serve as docking sites for specific adapter proteins and domains (e.g. SH2 domain), very little is known about specific binders of bacterial phosphorylation events and their role in signal transduction. Bacterial domains are however, seen to be conserved and are of great interest as they are implicated to be important from a functional and evolutionary perspective. Thus detection of protein domains that bind to specific phosphorylated residues via pull-down assays using synthetic peptides could possibly shed some light on their taxonomic distribution.

The protocol for the bait-prey affinity based interaction was adapted from eukaryotic systems to study the interactomes of detected phosphorylation sites in bacteria. Nine pairs of peptides were synthesized for this purpose. No significant outliers (interactors) were observed using the PrkC_01, YmfM and DnaK peptide pairs. While the methodology used is robust in design and extremely effective, an inherent disadvantage of this technique is the use of peptides outside the context of the whole folded protein. As a result, this technique could evade detection of certain interactions, as in case of PrkC_01, YmfM or DnaK. Additionally, certain technical issues were encountered during the affinity enrichments due to unequal amounts of peptides used for the pull-downs with the light and heavy lysates or the loss of sepharose beads during washes, thus resulting in a shifted distribution as observed in case of HPr or SunI. Further optimization of the protocol is required in order to be able to address such technicalities. Amongst the other interaction screens, binders of the PtkA and PrkC (threonine phosphorylated) peptide were the most interesting. A large number of ribosomal proteins were observed to bind

more to the unmodified peptide of PtkA. The BY-kinase PtkA is known to play a role in biofilm formation [49] and its auto-phosphorylation prevents its ATPase activity [167]. Recently, the ribosomal genes S21 and S11 were shown to inhibit biofilm formation in *B. subtilis* [220]. The inactivation of these two proteins hampers cell motility and promotes biofilm formation. Additionally, the role of ribosomal proteins L1, L22 and L34 with respect to sporulation was also realized and their mutations were reported to reduce sporulation frequency [221]. Based on this information, it can be speculated that binding of the ribosomal proteins to the unmodified (non-auto-phosphorylated) PtkA peptide could bring about their phosphorylation and interaction leading to biofilm formation during adverse conditions. Formation of structured biofilms is followed by cellular differentiation and spore development. Mutants incapable of producing extracellular matrix form unstructured biofilms and are deficient in sporulation [222]. On return of favorable conditions, spore germination is triggered in response to nutrients. It has been shown previously in *Streptomyces granaticolor* that ribosomal proteins undergo post-translational alterations and play a crucial role during spore germination [223]. The importance of ribosomes in *B. subtilis* has also been outlined in the past and in recent reports [224]. The authors specify the state of the spore reservoir that contains elements required for initial protein synthesis and other energy sources facilitated by translational components such as RpmE and Tig. The mechanism of activation of ribosomes during spore germination however was not elucidated. The auto-phosphorylation of the Hanks-type kinase PrkC is reported to considerably increase its kinase activity [225]. Thus it can be hypothesized that the ribosomal proteins and their associated factors get activated by auto-phosphorylated PrkC and help prepare the cell for conditions of germination.

As stated earlier, the interaction screen with the YkwC peptide pair yielded nine uncharacterized proteins that bound more to the phosphorylated serine peptide. Multiple sequence alignment analysis depicted seven of the nine proteins to have a [V/I] XXXXXX [V/I] motif and a conserved negatively charged residue (E or D) towards the C-terminus. This could depict a preferential motif pattern. Due to the lack of an overall large number of interactors of the phosphorylated serine residues and the absence of any site-directed mutational studies, the presence of this structural motif pattern can only be conjectured. Further mutagenic assays followed by functional assays are essential to

substantiate and establish domains that specifically bind to phosphorylated serine residues. Similarly, despite adopting a systematic approach to identify motifs that bind to phosphorylated threonine or tyrosine residues it remained elusive. Deeper structural motif predictions are imperative in order to delineate the functional relationships of protein domains and discern their influence on cellular physiology.

6. Conclusions

Phosphorylation events dependent on specific serine, threonine and tyrosine (STY) kinases and phosphatases in *Bacillus subtilis* were identified, characterized and subsequently mapped to central metabolic pathways. The conclusions drawn from this study are:

1. Measurement of the proteome and phosphoproteome dynamics during growth of *B. subtilis*:
 - a) Integration of two triple-SILAC experiments is an effective approach to encompass the entire bacterial growth curve and measure the changes occurring at the protein and phosphorylation site level.
 - b) Specific stages during the growth in which the kinases and phosphatases of interest were abundantly expressed were identified; with an exemption of PtkA, all of the quantified kinases and phosphatases were upregulated in the stationary phase of growth.
2. Quantitative phosphoproteomics screen for the identification of putative STY –kinase and –phosphatase substrates:
 - a) Application of the triple-SILAC approach in the background of the respective kinase or phosphatase gene knock-outs resulted in identification of putative novel *in vivo* substrates.
 - b) Selected putative candidates, such as YkwC and DnaK, were confirmed to be regulated by the respective kinases and phosphatases and the effect of phosphorylation on substrate activity was determined.
3. A prokaryotic network of kinases, phosphatases and their respective substrates was generated based on data obtained from the quantitative phosphoproteomics screens. The network contains multiple nodes potentially regulated by both, the kinases and the phosphatases pointing to the fact that they act in the same pathways and are likely to be tightly regulated.

4. Comprehensive map of *B. subtilis* phosphoproteome:

- a) 995 phosphorylation events were identified making it the most comprehensive phosphorylation specific map of *B. subtilis* to date.
- b) Novel phosphorylation events on proteins involved in regulation of carbohydrate, amino acid, lipid and purine metabolism; two-component systems; phosphotransferase system; stationary phase processes such as sporulation, biofilm formation, competence; stress responses were identified.
- c) Motif analysis of the phosphorylated peptides resulted in identification of 10 potential phosphorylated serine motifs and two phosphorylated threonine motifs, shedding light on the potential motif patterns that could be involved in phosphorylation in prokaryotes.

5. Identification of putative interactors of phosphorylated residues

- a) A SILAC-based bait peptide approach, applied for the first time in bacterial systems, resulted in the identification of putative interactors of phosphorylation motifs.
- b) Identification of ribosomal proteins binding to proteins involved in spore germination (PrkC) and biofilm formation (PtkA) confirmed transcriptional regulation of these processes.
- c) While no specific protein domain was identified as binding to phosphorylation motifs, a potential motif pattern - [V/I]XXXXXX[V/I] followed by a conserved negatively charged residue towards the C-terminus - was determined to bind the phosphorylated oxidoreductase peptide of YkwC, possibly indicating a motif pattern that interacts with phosphorylated serine residues.

7. References

1. Zhao, Y. and O.N. Jensen, *Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques*. Proteomics, 2009. **9**(20): p. 4632-41.
2. Ubersax, J.A. and J.E. Ferrell, Jr., *Mechanisms of specificity in protein phosphorylation*. Nat Rev Mol Cell Biol, 2007. **8**(7): p. 530-41.
3. Kim, S.C., et al., *Substrate and functional diversity of lysine acetylation revealed by a proteomics survey*. Mol Cell, 2006. **23**(4): p. 607-18.
4. Ohtsubo, K. and J.D. Marth, *Glycosylation in cellular mechanisms of health and disease*. Cell, 2006. **126**(5): p. 855-67.
5. Fischer, E.H. and E.G. Krebs, *Conversion of phosphorylase b to phosphorylase a in muscle extracts*. J Biol Chem, 1955. **216**(1): p. 121-32.
6. Munoz-Dorado, J., S. Inouye, and M. Inouye, *A gene encoding a protein serine/threonine kinase is required for normal development of M. xanthus, a gram-negative bacterium*. Cell, 1991. **67**(5): p. 995-1006.
7. Grangeasse, C., et al., *Characterization of a bacterial gene encoding an autophosphorylating protein tyrosine kinase*. Gene, 1997. **204**(1-2): p. 259-65.
8. Deutscher, J., C. Francke, and P.W. Postma, *How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria*. Microbiol Mol Biol Rev, 2006. **70**(4): p. 939-1031.
9. Pietack, N., et al., *In vitro phosphorylation of key metabolic enzymes from Bacillus subtilis: PrkC phosphorylates enzymes from different branches of basic metabolism*. J Mol Microbiol Biotechnol, 2010. **18**(3): p. 129-40.
10. Nariya, H. and S. Inouye, *Identification of a protein Ser/Thr kinase cascade that regulates essential transcriptional activators in Myxococcus xanthus development*. Mol Microbiol, 2005. **58**(2): p. 367-79.
11. Bramley, H.F. and H.L. Kornberg, *Sequence homologies between proteins of bacterial phosphoenolpyruvate-dependent sugar phosphotransferase systems: identification of possible phosphate-carrying histidine residues*. Proc Natl Acad Sci U S A, 1987. **84**(14): p. 4777-80.
12. Lengeler, J.W., G. Drews, and H.G. Schlegel, *Biology of prokaryotes*. 1999, Blackwell Science: Stuttgart, Germany.
13. Deutscher, J., et al., *P-Ser-HPr--a link between carbon metabolism and the virulence of some pathogenic bacteria*. Biochim Biophys Acta, 2005. **1754**(1-2): p. 118-25.
14. Meadow, N.D., D.K. Fox, and S. Roseman, *The bacterial phosphoenolpyruvate: glycolate phosphotransferase system*. Annu Rev Biochem, 1990. **59**: p. 497-542.
15. Nixon, B.T., C.W. Ronson, and F.M. Ausubel, *Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC*. Proc Natl Acad Sci U S A, 1986. **83**(20): p. 7850-4.
16. Ninfa, A.J. and B. Magasanik, *Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in Escherichia coli*. Proc Natl Acad Sci U S A, 1986. **83**(16): p. 5909-13.
17. Stock, A.M., V.L. Robinson, and P.N. Goudreau, *Two-component signal transduction*. Annu Rev Biochem, 2000. **69**: p. 183-215.
18. Surette, M.G., et al., *Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis*. J Biol Chem, 1996. **271**(2): p. 939-45.

19. Falke, J.J., et al., *The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes*. *Annu Rev Cell Dev Biol*, 1997. **13**: p. 457-512.
20. Iuchi, S. and L. Weiner, *Cellular and molecular physiology of Escherichia coli in the adaptation to aerobic environments*. *J Biochem*, 1996. **120**(6): p. 1055-63.
21. Hoch, J.A., *Regulation of the phosphorelay and the initiation of sporulation in Bacillus subtilis*. *Annu Rev Microbiol*, 1993. **47**: p. 441-65.
22. Besant, P.G. and P.V. Attwood, *Detection and analysis of protein histidine phosphorylation*. *Mol Cell Biochem*, 2009. **329**(1-2): p. 93-106.
23. Maeda, T., S.M. Wurgler-Murphy, and H. Saito, *A two-component system that regulates an osmosensing MAP kinase cascade in yeast*. *Nature*, 1994. **369**(6477): p. 242-5.
24. Chang, C., et al., *Arabidopsis ethylene-response gene ETR1: similarity of product to two-component regulators*. *Science*, 1993. **262**(5133): p. 539-44.
25. Bakal, C.J. and J.E. Davies, *No longer an exclusive club: eukaryotic signalling domains in bacteria*. *Trends in Cell Biology*, 2000. **10**(1): p. 32-38.
26. Kuo, J.F. and P. Greengard, *An adenosine 3',5'-monophosphate-dependent protein kinase from Escherichia coli*. *J Biol Chem*, 1969. **244**(12): p. 3417-9.
27. Khandelwal, R.L., T.N. Spearman, and I.R. Hamilton, *Protein kinase activity in cariogenic and non-cariogenic oral streptococci: Activation and inhibition by cyclic AMP*. *FEBS Lett*, 1973. **31**(2): p. 246-250.
28. Rahmsdorf, H.J., et al., *Protein kinase induction in Escherichia coli by bacteriophage T7*. *Proc Natl Acad Sci U S A*, 1974. **71**(2): p. 586-9.
29. Wang, J.Y. and D.E. Koshland, Jr., *Evidence for protein kinase activities in the prokaryote Salmonella typhimurium*. *J Biol Chem*, 1978. **253**(21): p. 7605-8.
30. Garnak, M. and H.C. Reeves, *Phosphorylation of Isocitrate dehydrogenase of Escherichia coli*. *Science*, 1979. **203**(4385): p. 1111-2.
31. Mijakovic, I., et al., *Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine*. *Nucleic Acids Research*, 2006. **34**(5): p. 1588-1596.
32. Mijakovic, I., et al., *Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases*. *Embo Journal*, 2003. **22**(18): p. 4709-4718.
33. Klein, G., C. Dartigalongue, and S. Raina, *Phosphorylation-mediated regulation of heat shock response in Escherichia coli*. *Molecular Microbiology*, 2003. **48**(1): p. 269-285.
34. Shapland, E.B., et al., *An Essential Tyrosine Phosphatase Homolog Regulates Cell Separation, Outer Membrane Integrity, and Morphology in Caulobacter crescentus*. *Journal of Bacteriology*, 2011. **193**(17): p. 4361-4370.
35. Cozzone, A.J., *Role of protein phosphorylation on serine/threonine and tyrosine in the virulence of bacterial pathogens*. *J Mol Microbiol Biotechnol*, 2005. **9**(3-4): p. 198-213.
36. Mijakovic, I. and B. Macek, *Impact of phosphoproteomics on studies of bacterial physiology*. *FEMS Microbiol Rev*, 2012. **36**(4): p. 877-92.
37. Kunst, F., et al., *The complete genome sequence of the gram-positive bacterium Bacillus subtilis*. *Nature*, 1997. **390**(6657): p. 249-56.
38. Harwood, C.R., *Bacillus subtilis and its relatives: molecular biological and industrial workhorses*. *Trends Biotechnol*, 1992. **10**(7): p. 247-56.
39. Kobayashi, K., et al., *Essential Bacillus subtilis genes*. *Proc Natl Acad Sci U S A*, 2003. **100**(8): p. 4678-83.
40. Madec, E., et al., *Characterization of a membrane-linked Ser/Thr protein kinase in Bacillus subtilis, implicated in developmental processes*. *Mol Microbiol*, 2002. **46**(2): p. 571-86.
41. Gaidenko, T.A., T.J. Kim, and C.W. Price, *The PrpC serine-threonine phosphatase and PrkC kinase have opposing physiological roles in stationary-phase Bacillus subtilis cells*. *J Bacteriol*, 2002. **184**(22): p. 6109-14.

42. Pietack, N., et al., *In vitro Phosphorylation of Key Metabolic Enzymes from Bacillus subtilis: PrkC Phosphorylates Enzymes from Different Branches of Basic Metabolism*. Journal of Molecular Microbiology and Biotechnology, 2010. **18**(3): p. 129-140.
43. Ravikumar, V., et al., *Quantitative phosphoproteome analysis of bacillus subtilis reveals novel substrates of the kinase PrkC and phosphatase PrpC*. Mol Cell Proteomics, 2014. **13**(8): p. 1965-78.
44. Rougraff, P.M., et al., *Purification and characterization of 3-hydroxyisobutyrate dehydrogenase from rabbit liver*. J Biol Chem, 1988. **263**(1): p. 327-31.
45. Bidnenko, V., et al., *Bacillus subtilis serine/threonine protein kinase YabT is involved in spore development via phosphorylation of a bacterial recombinase*. Mol Microbiol, 2013. **88**(5): p. 921-35.
46. Kobir, A., et al., *Phosphorylation of Bacillus subtilis gene regulator AbrB modulates its DNA-binding properties*. Mol Microbiol, 2014. **92**(5): p. 1129-41.
47. Cairns, L.S., et al., *A mechanical signal transmitted by the flagellum controls signalling in Bacillus subtilis*. Mol Microbiol, 2013. **90**(1): p. 6-21.
48. Jers, C., et al., *Bacillus subtilis two-component system sensory kinase DegS is regulated by serine phosphorylation in its input domain*. PLoS One, 2011. **6**(2): p. e14653.
49. Kiley, T.B. and N.R. Stanley-Wall, *Post-translational control of Bacillus subtilis biofilm formation mediated by tyrosine phosphorylation*. Mol Microbiol, 2010. **78**(4): p. 947-63.
50. Petranovic, D., et al., *Activation of Bacillus subtilis Ugd by the BY-Kinase PtkA Proceeds via Phosphorylation of Its Residue Tyrosine 70*. Journal of Molecular Microbiology and Biotechnology, 2009. **17**(2): p. 83-89.
51. Jers, C., et al., *Bacillus subtilis BY-kinase PtkA controls enzyme activity and localization of its protein substrates*. Mol Microbiol, 2010. **77**(2): p. 287-99.
52. Kearns, D.B., et al., *A master regulator for biofilm formation by Bacillus subtilis*. Mol Microbiol, 2005. **55**(3): p. 739-49.
53. Winkelman, J.T., et al., *RemA is a DNA-binding protein that activates biofilm matrix gene expression in Bacillus subtilis*. Mol Microbiol, 2013. **88**(5): p. 984-97.
54. Hoper, D., U. Volker, and M. Hecker, *Comprehensive characterization of the contribution of individual SigB-dependent general stress genes to stress resistance of Bacillus subtilis*. J Bacteriol, 2005. **187**(8): p. 2810-26.
55. Wasinger, V.C., et al., *Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium*. Electrophoresis, 1995. **16**(7): p. 1090-4.
56. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins*. J Biol Chem, 1975. **250**(10): p. 4007-21.
57. Gygi, S.P., et al., *Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology*. Proc Natl Acad Sci U S A, 2000. **97**(17): p. 9390-5.
58. Wu, C.C. and M.J. MacCoss, *Shotgun proteomics: tools for the analysis of complex biological systems*. Curr Opin Mol Ther, 2002. **4**(3): p. 242-50.
59. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Anal Chem, 1988. **60**(20): p. 2299-301.
60. Tanaka K, W.H., Ido Y et al, *Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry*. Rapid Commun Mass Spectrom, 1988. **2**(8): p. 151-153.
61. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. Science, 1989. **246**(4926): p. 64-71.
62. Mann, M., P. Hojrup, and P. Roepstorff, *Use of mass spectrometric molecular weight information to identify proteins in sequence databases*. Biol Mass Spectrom, 1993. **22**(6): p. 338-45.

63. Henzel, W.J., et al., *Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases*. Proc Natl Acad Sci U S A, 1993. **90**(11): p. 5011-5.
64. Pappin, D.J., P. Hojrup, and A.J. Bleasby, *Rapid identification of proteins by peptide-mass fingerprinting*. Curr Biol, 1993. **3**(6): p. 327-32.
65. James, P., et al., *Protein identification by mass profile fingerprinting*. Biochem Biophys Res Commun, 1993. **195**(1): p. 58-64.
66. Yates, J.R., 3rd, et al., *Peptide mass maps: a highly informative approach to protein identification*. Anal Biochem, 1993. **214**(2): p. 397-408.
67. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, *An automated multidimensional protein identification technology for shotgun proteomics*. Anal Chem, 2001. **73**(23): p. 5683-90.
68. Ngoka, L.C., *Sample prep for proteomics of breast cancer: proteomics and gene ontology reveal dramatic differences in protein solubilization preferences of radioimmunoprecipitation assay and urea lysis buffers*. Proteome Sci, 2008. **6**: p. 30.
69. Altelaar, A.F., J. Munoz, and A.J. Heck, *Next-generation proteomics: towards an integrative view of proteome dynamics*. Nat Rev Genet, 2013. **14**(1): p. 35-48.
70. Guengerich, F.P., *Thematic minireview series on biological applications of mass spectrometry*. J Biol Chem, 2011. **286**(29): p. 25417.
71. Finehout, E.J. and K.H. Lee, *An introduction to mass spectrometry applications in biological research*. Biochem Mol Biol Educ, 2004. **32**(2): p. 93-100.
72. Münzenberg, G., *Development of mass spectrometers from Thomson and Aston to present*. International Journal of Mass Spectrometry, 2013. **349-350**: p. 9-18.
73. Willard, H.H., et al., *Instrumental methods of analysis*. 1988, Wadsworth Publishing Co., Belmont, CA. p. 465-507.
74. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, *Proteomics by mass spectrometry: approaches, advances, and applications*. Annu Rev Biomed Eng, 2009. **11**: p. 49-79.
75. Steen, H. and M. Mann, *The ABC's (and XYZ's) of peptide sequencing*. Nat Rev Mol Cell Biol, 2004. **5**(9): p. 699-711.
76. Olumee, Z., J.H. Callahan, and A. Vertes, *Droplet dynamics changes in electrostatic sprays of methanol-water mixtures*. J Phys Chem A, 1998. **102**(46): p. 9154-60.
77. Taylor, G., *Disintegration of water droplets in an electric field*. Proc Royal Soc Lond, 1964. **280**(1382): p. 383.
78. Rayleigh, L., *On the equilibrium of liquid conducting masses charged with electricity*, in *Philosophical Magazine*. 1882. p. 184-86.
79. Udeshi, N.D., et al., *Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry*. Nat Protoc, 2008. **3**(11): p. 1709-17.
80. Douglas, D.J., A.J. Frank, and D. Mao, *Linear ion traps in mass spectrometry*. Mass Spectrom Rev, 2005. **24**(1): p. 1-29.
81. Makarov, A., *Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis*. Anal Chem, 2000. **72**(6): p. 1156-62.
82. Allen, J.S., *An improved electron multiplier particle counter*. Review of Scientific Instruments 1947. **18**(10): p. 739.
83. Brown, K.L. and G.W. Tautfest, *Faraday-cup monitors for high-energy electron beams*. Review of scientific instruments, 1956. **27**(9): p. 696-702.
84. Dubois, F., R. Knochenmuss, and R. Zenobi, *An ion-to-photon conversion detector for mass spectrometry*. Int J Mass Spectrom Ion Processes, 1997: p. 89-98.
85. Wiza, J., *Microchannel plate detectors*. Nuclear Instruments and Methods, 1979. **162**(1-3): p. 587-601.

86. Pitt, J.J., *Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry*. Clin Biochem Rev, 2009. **30**(1): p. 19-34.
87. Nesvizhskii, A.I., *Proteogenomics: concepts, applications and computational strategies*. Nat Methods, 2014. **11**(11): p. 1114-25.
88. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. **20**(18): p. 3551-67.
89. Eng, J.K., A.L. McCormack, and J.R. Yates, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database*. J Am Soc Mass Spectrom, 1994. **5**(11): p. 976-89.
90. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification*. Nat Biotechnol, 2008. **26**(12): p. 1367-72.
91. Sadygov, R.G., D. Cociorva, and J.R. Yates, 3rd, *Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book*. Nat Methods, 2004. **1**(3): p. 195-202.
92. Thomas, S.N., et al., *Redox Proteomics - From protein modifications to cellular dysfunction and diseases*. Wiley-Interscience Series on Mass Spectrometry, ed. D.M. Desiderio and N.M. Nibbering. 2006: John Wiley & Sons, Inc.
93. Lam, H., et al., *Development and validation of a spectral library searching method for peptide identification from MS/MS*. Proteomics, 2007. **7**(5): p. 655-67.
94. Frewen, B.E., et al., *Analysis of peptide MS/MS spectra from large-scale proteomics experiments using spectrum libraries*. Anal Chem, 2006. **78**(16): p. 5678-84.
95. Ma, B., et al., *PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry*. Rapid Commun Mass Spectrom, 2003. **17**(20): p. 2337-42.
96. Taylor, J.A. and R.S. Johnson, *Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry*. Anal Chem, 2001. **73**(11): p. 2594-604.
97. Shteynberg, D., et al., *Combining results of multiple search engines in proteomics*. Mol Cell Proteomics, 2013. **12**(9): p. 2383-93.
98. Paulo, J.A., *Practical and Efficient Searching in Proteomics: A Cross Engine Comparison*. Webmedcentral, 2013. **4**(10).
99. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry*. Nat Methods, 2007. **4**(3): p. 207-14.
100. Gorg, A., W. Weiss, and M.J. Dunn, *Current two-dimensional electrophoresis technology for proteomics*. Proteomics, 2004. **4**(12): p. 3665-85.
101. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review*. Anal Bioanal Chem, 2007. **389**(4): p. 1017-31.
102. Nikolov, M., C. Schmidt, and H. Urlaub, *Quantitative mass spectrometry-based proteomics: an overview*. Methods Mol Biol, 2012. **893**: p. 85-100.
103. Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags*. Nat Biotechnol, 1999. **17**(10): p. 994-9.
104. Oda, Y., et al., *Accurate quantitation of protein expression and site-specific phosphorylation*. Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6591-6.
105. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative*. Nat Chem Biol, 2005. **1**(5): p. 252-62.
106. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics*. Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
107. Yao, X., et al., *Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus*. Anal Chem, 2001. **73**(13): p. 2836-42.

108. Johnson, K.L. and D.C. Muddiman, *A method for calculating 160/180 peptide ion ratios for the relative quantification of proteomes*. J Am Soc Mass Spectrom, 2004. **15**(4): p. 437-45.
109. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*. Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
110. Thompson, A., et al., *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS*. Anal Chem, 2003. **75**(8): p. 1895-904.
111. Gerber, S.A., et al., *Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS*. Proc Natl Acad Sci U S A, 2003. **100**(12): p. 6940-5.
112. Schwanhaussner, B., et al., *Global quantification of mammalian gene expression control*. Nature, 2011. **473**(7347): p. 337-42.
113. Schmidl, S.R., et al., *The phosphoproteome of the minimal bacterium Mycoplasma pneumoniae: analysis of the complete known Ser/Thr kinome suggests the existence of novel kinases*. Mol Cell Proteomics, 2010. **9**(6): p. 1228-42.
114. Macek, B., M. Mann, and J.V. Olsen, *Global and site-specific quantitative phosphoproteomics: principles and applications*. Annu Rev Pharmacol Toxicol, 2009. **49**: p. 199-221.
115. Roux, P.P. and P. Thibault, *The coming of age of phosphoproteomics--from large data sets to inference of protein functions*. Mol Cell Proteomics, 2013. **12**(12): p. 3453-64.
116. Thingholm, T.E., O.N. Jensen, and M.R. Larsen, *Analytical strategies for phosphoproteomics*. Proteomics, 2009. **9**(6): p. 1451-68.
117. Rogers, L.D. and L.J. Foster, *Phosphoproteomics--finally fulfilling the promise?* Mol Biosyst, 2009. **5**(10): p. 1122-9.
118. Beausoleil, S.A., et al., *Large-scale characterization of HeLa cell nuclear phosphoproteins*. Proc Natl Acad Sci U S A, 2004. **101**(33): p. 12130-5.
119. Schroeder, M.J., et al., *A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry*. Anal Chem, 2004. **76**(13): p. 3590-8.
120. Molina, H., et al., *Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2199-204.
121. Stensballe, A., et al., *Electron capture dissociation of singly and multiply phosphorylated peptides*. Rapid Commun Mass Spectrom, 2000. **14**(19): p. 1793-800.
122. Olsen, J.V., et al., *Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis*. Sci Signal, 2010. **3**(104): p. ra3.
123. Wu, R., et al., *A large-scale method to measure absolute protein phosphorylation stoichiometries*. Nat Methods, 2011. **8**(8): p. 677-83.
124. Chalkley, R.J. and K.R. Clauser, *Modification site localization scoring: strategies and performance*. Mol Cell Proteomics, 2012. **11**(5): p. 3-14.
125. Martin, D.B., et al., *Investigation of neutral loss during collision-induced dissociation of peptide ions*. Anal Chem, 2005. **77**(15): p. 4870-82.
126. Taus, T., et al., *Universal and confident phosphorylation site localization using phosphoRS*. J Proteome Res, 2011. **10**(12): p. 5354-62.
127. Creasy, D.M. and J.S. Cottrell, *Error tolerant searching of uninterpreted tandem mass spectrometry data*. Proteomics, 2002. **2**(10): p. 1426-34.
128. Beausoleil, S.A., et al., *A probability-based approach for high-throughput protein phosphorylation analysis and site localization*. Nat Biotechnol, 2006. **24**(10): p. 1285-92.
129. Olsen, J.V. and M. Mann, *Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation*. Proc Natl Acad Sci U S A, 2004. **101**(37): p. 13417-22.

130. Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks*. Cell, 2006. **127**(3): p. 635-48.
131. Gruhler, A., et al., *Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway*. Mol Cell Proteomics, 2005. **4**(3): p. 310-27.
132. Fila, J. and D. Honys, *Enrichment techniques employed in phosphoproteomics*. Amino Acids, 2012. **43**(3): p. 1025-47.
133. Alpert, A.J., *Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds*. J Chromatogr, 1990. **499**: p. 177-96.
134. McNulty, D.E. and R.S. Annan, *Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection*. Mol Cell Proteomics, 2008. **7**(5): p. 971-80.
135. Andersson, L. and J. Porath, *Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography*. Anal Biochem, 1986. **154**(1): p. 250-4.
136. Posewitz, M.C. and P. Tempst, *Immobilized gallium(III) affinity chromatography of phosphopeptides*. Anal Chem, 1999. **71**(14): p. 2883-92.
137. Kweon, H.K. and K. Hakansson, *Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis*. Anal Chem, 2006. **78**(6): p. 1743-9.
138. Ficarro, S.B., et al., *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae*. Nat Biotechnol, 2002. **20**(3): p. 301-5.
139. Ikeguchi, Y. and H. Nakamura, *Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania*. Anal Sci, 1997. **13**: p. 479-85.
140. Larsen, M.R., et al., *Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns*. Mol Cell Proteomics, 2005. **4**(7): p. 873-86.
141. Sugiyama, N., et al., *Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications*. Mol Cell Proteomics, 2007. **6**(6): p. 1103-9.
142. Imam-Sghiouar, N., et al., *Subproteomics analysis of phosphorylated proteins: application to the study of B-lymphoblasts from a patient with Scott syndrome*. Proteomics, 2002. **2**(7): p. 828-38.
143. Pandey, A., et al., *Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors*. Proc Natl Acad Sci U S A, 2000. **97**(1): p. 179-84.
144. Rush, J., et al., *Immunoaffinity profiling of tyrosine phosphorylation in cancer cells*. Nat Biotechnol, 2005. **23**(1): p. 94-101.
145. Seet, B.T., et al., *Reading protein modifications with interaction domains*. Nat Rev Mol Cell Biol, 2006. **7**(7): p. 473-83.
146. Schulze, W.X. and M. Mann, *A novel proteomic screen for peptide-protein interactions*. J Biol Chem, 2004. **279**(11): p. 10756-64.
147. Schulze, W.X., L. Deng, and M. Mann, *Phosphotyrosine interactome of the ErbB-receptor kinase family*. Mol Syst Biol, 2005. **1**: p. 2005 0008.
148. Hunter, T., *Signaling-2000 and beyond*. Cell, 2000. **100**(1): p. 113-27.
149. Bradshaw, J.M. and G. Waksman, *Molecular recognition by SH2 domains*. Adv Protein Chem, 2002. **61**: p. 161-210.
150. Yaffe, M.B., *Phosphotyrosine-binding domains in signal transduction*. Nat Rev Mol Cell Biol, 2002. **3**(3): p. 177-86.
151. Musacchio, A., et al., *SH3--an abundant protein domain in search of a function*. FEBS Lett, 1992. **307**(1): p. 55-61.

152. Mayer, B.J., *SH3 domains: complexity in moderation*. J Cell Sci, 2001. **114**(Pt 7): p. 1253-63.
153. Deribe, Y.L., T. Pawson, and I. Dikic, *Post-translational modifications in signal integration*. Nat Struct Mol Biol, 2010. **17**(6): p. 666-72.
154. Jacobson, R.H., et al., *Structure and function of a human TAFII250 double bromodomain module*. Science, 2000. **288**(5470): p. 1422-5.
155. Pincus, D., et al., *Evolution of the phospho-tyrosine signaling machinery in premetazoan lineages*. Proc Natl Acad Sci U S A, 2008. **105**(28): p. 9680-4.
156. Fromont-Racine, M., J.C. Rain, and P. Legrain, *Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens*. Nat Genet, 1997. **16**(3): p. 277-82.
157. Kay, B.K., J. Kasanov, and M. Yamabhai, *Screening phage-displayed combinatorial peptide libraries*. Methods, 2001. **24**(3): p. 240-6.
158. Espejo, A., et al., *A protein-domain microarray identifies novel protein-protein interactions*. Biochem J, 2002. **367**(Pt 3): p. 697-702.
159. Elia, A.E., L.C. Cantley, and M.B. Yaffe, *Proteomic screen finds pSer/pThr-binding domain localizing Plk1 to mitotic substrates*. Science, 2003. **299**(5610): p. 1228-31.
160. Lim, W.A. and T. Pawson, *Phosphotyrosine signaling: evolving a new cellular communication system*. Cell, 2010. **142**(5): p. 661-7.
161. Vermeulen, M., et al., *Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers*. Cell, 2010. **142**(6): p. 967-80.
162. Vermeulen, M., et al., *Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4*. Cell, 2007. **131**(1): p. 58-69.
163. Whisstock, J.C. and A.M. Lesk, *SH3 domains in prokaryotes*. Trends Biochem Sci, 1999. **24**(4): p. 132-3.
164. Bilwes, A.M., et al., *Structure of CheA, a signal-transducing histidine kinase*. Cell, 1999. **96**(1): p. 131-41.
165. Maguin, E., et al., *Efficient insertional mutagenesis in lactococci and other gram-positive bacteria*. J Bacteriol, 1996. **178**(3): p. 931-5.
166. Monedero, V., et al., *Regulatory functions of serine-46-phosphorylated HPr in Lactococcus lactis*. J Bacteriol, 2001. **183**(11): p. 3391-8.
167. Mijakovic, I., et al., *Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases*. EMBO J, 2003. **22**(18): p. 4709-18.
168. Monod, J., *The Growth of Bacterial Cultures*. Annual Review of Microbiology, 1949. **3**: p. 317-394.
169. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 1976. **72**: p. 248-254.
170. Macek, B., et al., *The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis*. Mol Cell Proteomics, 2007. **6**(4): p. 697-707.
171. Ishihama, Y., J. Rappsilber, and M. Mann, *Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics*. J Proteome Res, 2006. **5**(4): p. 988-94.
172. Macek, B., et al., *Phosphoproteome analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation*. Mol Cell Proteomics, 2008. **7**(2): p. 299-307.
173. Franz-Wachtel, M., et al., *Global detection of protein kinase D-dependent phosphorylation events in nocodazole-treated human cells*. Mol Cell Proteomics, 2012. **11**(5): p. 160-70.
174. Krug, K., et al., *Deep coverage of the Escherichia coli proteome enables the assessment of false discovery rates in simple proteogenomic experiments*. Mol Cell Proteomics, 2013. **12**(11): p. 3420-30.

175. Olsen, J.V., et al., *Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap*. Mol Cell Proteomics, 2005. **4**(12): p. 2010-21.
176. Cox, J., et al., *A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics*. Nat Protoc, 2009. **4**(5): p. 698-705.
177. Cox, J., et al., *Andromeda: a peptide search engine integrated into the MaxQuant environment*. J Proteome Res, 2011. **10**(4): p. 1794-805.
178. Team, R.D.C., *R: A language and environment for statistical computing*. 2014.
179. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society. Series B (Methodological), 1995. **57**: p. 289-300.
180. Yao, T., et al., *The catalytic property of 3-hydroxyisobutyrate dehydrogenase from Bacillus cereus on 3-hydroxypropionate*. Appl Biochem Biotechnol, 2010. **160**(3): p. 694-703.
181. Fabret, C., S.D. Ehrlich, and P. Noirot, *A new mutation delivery system for genome-scale approaches in Bacillus subtilis*. Mol Microbiol, 2002. **46**(1): p. 25-36.
182. Nakamoto, H., et al., *Physical interaction between bacterial heat shock protein (Hsp) 90 and Hsp70 chaperones mediates their cooperative action to refold denatured proteins*. J Biol Chem, 2014. **289**(9): p. 6110-9.
183. Monod, J., *The Growth of Bacterial Cultures*. Annual Review of Microbiology, 1949. **3**(1): p. 371-394.
184. Soufi, B., et al., *Stable isotope labeling by amino acids in cell culture (SILAC) applied to quantitative proteomics of Bacillus subtilis*. J Proteome Res, 2010. **9**(7): p. 3638-46.
185. Soares, N.C., et al., *Global dynamics of the Escherichia coli proteome and phosphoproteome during growth in minimal medium*. J Proteome Res, 2013. **12**(6): p. 2611-21.
186. Knight, Z.A., H. Lin, and K.M. Shokat, *Targeting the cancer kinome through polypharmacology*. Nat Rev Cancer, 2010. **10**(2): p. 130-7.
187. Shi, L., et al., *Cross-phosphorylation of bacterial serine/threonine and tyrosine protein kinases on key regulatory residues*. Front Microbiol, 2014. **5**: p. 495.
188. Chou, M.F. and D. Schwartz, *Biological sequence motif discovery using motif-x*. Curr Protoc Bioinformatics, 2011. **Chapter 13**: p. Unit 13 15-24.
189. Schwartz, D. and S.P. Gygi, *An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets*. Nat Biotechnol, 2005. **23**(11): p. 1391-8.
190. Kholodenko, B.N., et al., *Strong control on the transit time in metabolic channelling*. FEBS Lett, 1996. **389**(2): p. 123-5.
191. Campanella, M.E., H. Chu, and P.S. Low, *Assembly and regulation of a glycolytic enzyme complex on the human erythrocyte membrane*. Proc Natl Acad Sci U S A, 2005. **102**(7): p. 2402-7.
192. van Sinderen, D., R. Kiewiet, and G. Venema, *Differential expression of two closely related deoxyribonuclease genes, nucA and nucB, in Bacillus subtilis*. Mol Microbiol, 1995. **15**(2): p. 213-23.
193. Tojo, S., et al., *Molecular mechanisms underlying the positive stringent response of the Bacillus subtilis ilv-leu operon, involved in the biosynthesis of branched-chain amino acids*. J Bacteriol, 2008. **190**(18): p. 6134-47.
194. Levine, A., et al., *Analysis of the dynamic Bacillus subtilis Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes*. Proteomics, 2006. **6**(7): p. 2157-73.
195. Qi, Y., Y. Kobayashi, and F.M. Hulett, *The pst operon of Bacillus subtilis has a phosphate-regulated promoter and is involved in phosphate transport but not in regulation of the pho regulon*. J Bacteriol, 1997. **179**(8): p. 2534-9.

196. Antelmann, H., C. Scharf, and M. Hecker, *Phosphate starvation-inducible proteins of Bacillus subtilis: proteomics and transcriptional analysis*. J Bacteriol, 2000. **182**(16): p. 4478-90.
197. Mijakovic, I., et al., *Protein-tyrosine phosphorylation in Bacillus subtilis*. J Mol Microbiol Biotechnol, 2005. **9**(3-4): p. 189-97.
198. Chu, F., et al., *Targets of the master regulator of biofilm formation in Bacillus subtilis*. Mol Microbiol, 2006. **59**(4): p. 1216-28.
199. O'Reilly, M. and K.M. Devine, *Expression of AbrB, a transition state regulator from Bacillus subtilis, is growth phase dependent in a manner resembling that of Fis, the nucleoid binding protein from Escherichia coli*. J Bacteriol, 1997. **179**(2): p. 522-9.
200. Derouiche, A., et al., *Interaction of bacterial fatty-acid-displaced regulators with DNA is interrupted by tyrosine phosphorylation in the helix-turn-helix domain*. Nucleic Acids Res, 2013. **41**(20): p. 9371-81.
201. Atluri, S., et al., *Cooperativity between different nutrient receptors in germination of spores of Bacillus subtilis and reduction of this cooperativity by alterations in the GerB receptor*. J Bacteriol, 2006. **188**(1): p. 28-36.
202. Hyatt, M.T. and H.S. Levinson, *Conditions affecting Bacillus megaterium spore germination in glucose or various nitrogenous compounds*. J Bacteriol, 1962. **83**: p. 1231-7.
203. Schumann, W., *Regulation of the heat shock response in Escherichia coli and Bacillus subtilis*. J Biosci, 1996. **21**(2): p. 133-148.
204. Becherelli, M., J. Tao, and N.S. Ryder, *Involvement of heat shock proteins in Candida albicans biofilm formation*. J Mol Microbiol Biotechnol, 2013. **23**(6): p. 396-400.
205. Kuczynska-Wisnik, D., E. Matuszewska, and E. Laskowska, *Escherichia coli heat-shock proteins IbpA and IbpB affect biofilm formation by influencing the level of extracellular indole*. Microbiology, 2010. **156**(Pt 1): p. 148-57.
206. Banse, A.V., et al., *Parallel pathways of repression and antirepression governing the transition to stationary phase in Bacillus subtilis*. Proc Natl Acad Sci U S A, 2008. **105**(40): p. 15547-52.
207. *Two-component systems in bacteria*, ed. R. Gross and D. Beier. 2012: Caister Academic Press.
208. Schrecke, K., S. Jordan, and T. Mascher, *Stoichiometry and perturbation studies of the LiaFSR system of Bacillus subtilis*. Mol Microbiol, 2013. **87**(4): p. 769-88.
209. Dominguez-Escobar, J., et al., *Subcellular localization, interactions and dynamics of the phage-shock protein-like Lia response in Bacillus subtilis*. Mol Microbiol, 2014. **92**(4): p. 716-32.
210. *Amino acid biosynthesis - Pathways, regulation and metabolic engineering*, ed. A. Steinbuchel. Vol. 5. 2007: Springer.
211. Li, D.D., et al., *ECM17-dependent methionine/cysteine biosynthesis contributes to biofilm formation in Candida albicans*. Fungal Genet Biol, 2013. **51**: p. 50-9.
212. Elsholz, A.K., et al., *Global impact of protein arginine phosphorylation on the physiology of Bacillus subtilis*. Proc Natl Acad Sci U S A, 2012. **109**(19): p. 7451-6.
213. Nakano, S., et al., *Spx-dependent global transcriptional control is induced by thiol-specific oxidative stress in Bacillus subtilis*. Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13603-8.
214. Nakano, M.M., S. Nakano, and P. Zuber, *Spx (YjbD), a negative effector of competence in Bacillus subtilis, enhances ClpC-MecA-ComK interaction*. Mol Microbiol, 2002. **44**(5): p. 1341-9.
215. Lushchak, V.I., *Adaptive response to oxidative stress: Bacteria, fungi, plants and animals*. Comp Biochem Physiol C Toxicol Pharmacol, 2011. **153**(2): p. 175-90.

216. Ebbole, D.J. and H. Zalkin, *Cloning and characterization of a 12-gene cluster from Bacillus subtilis encoding nine enzymes for de novo purine nucleotide synthesis*. J Biol Chem, 1987. **262**(17): p. 8274-87.
217. Pristic, S., et al., *Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases*. Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7521-6.
218. Hansen, A.M., et al., *The Escherichia coli phosphotyrosine proteome relates to core pathways and virulence*. PLoS Pathog, 2013. **9**(6): p. e1003403.
219. Goodacre, N.F., D.L. Gerloff, and P. Uetz, *Protein domains of unknown function are essential in bacteria*. MBio, 2013. **5**(1): p. e00744-13.
220. Takada, H., et al., *Cell motility and biofilm formation in Bacillus subtilis are affected by the ribosomal proteins, S11 and S21*. Biosci Biotechnol Biochem, 2014. **78**(5): p. 898-907.
221. Akanuma, G., et al., *Inactivation of ribosomal protein genes in Bacillus subtilis reveals importance of each ribosomal protein for cell proliferation and cell differentiation*. J Bacteriol, 2012. **194**(22): p. 6282-91.
222. Vlamakis, H., et al., *Control of cell fate by the formation of an architecturally complex bacterial community*. Genes Dev, 2008. **22**(7): p. 945-53.
223. Mikulik, K., et al., *RNA and ribosomal protein patterns during aerial spore germination in Streptomyces granaticolor*. Eur J Biochem, 1984. **145**(2): p. 381-8.
224. Sinai, L., et al., *The molecular timeline of a reviving bacterial spore*. Mol Cell, 2015. **57**(4): p. 695-707.
225. Madec, E., et al., *Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from Bacillus subtilis*. J Mol Biol, 2003. **330**(3): p. 459-72.

8. List of Publications

1. Shi L*, **Ravikumar V***, Derouiche A, Kalantari A, Macek B, Mijakovic I.
Phosphorylation of chaperone protein DnaK is crucial for DnaK activity and cellular survival upon heat shock in *Bacillus subtilis*.
(manuscript in preparation).
2. **Ravikumar V**, Macek B, Mijakovic I.
Resources for assignment of phosphorylation sites on peptides and proteins.
Methods in Molecular Biology, 2015 (in press).
3. Shi L, Pigeonneau N, **Ravikumar V**, Dobrinic P, Bidnenko V, Macek B, Franjevic D, Noirot-Gros MF, Mijakovic I.
Cross-phosphorylation of bacterial serine/threonine and tyrosine protein kinases on key regulatory residues.
Frontiers in Microbiology, Sept 2014, **5**:495.
4. **Ravikumar V***, Shi L*, Krug K, Derouiche A, Jers C, Cousin C, Kobir A, Mijakovic I, Macek B.
Quantitative phosphoproteome analysis of *Bacillus subtilis* reveals novel substrates of the kinase PrkC and phosphatase PrpC.
Molecular and Cellular Proteomics, Aug 2014, **13**(8):1965-1978.
5. Soufi B*, Soares NC*, **Ravikumar V**, Macek B.
Proteomics reveals evidence of cross-talk between protein modifications in bacteria: focus on acetylation and phosphorylation.
Current Opinion Microbiology, Jun 2012, **15**(3):357-363.

9. Curriculum Vitae

PERSONAL INFORMATION

Name: Vaishnavi Ravikumar

Date of Birth: 16th October 1986

Nationality: Indian

EDUCATIONAL QUALIFICATIONS

09/2011 - Present	PhD Student in Biology Proteome Center Tübingen, Eberhard Karls Universität, Tübingen, Germany
07/2007 – 06/2009	Master of Science in Life Sciences Mount Carmel College, Bangalore, India
07/2004 – 06/2007	Bachelor of Science in Chemistry, Botany, Zoology St. Joseph's College of Arts and Science, Bangalore, India

RESEARCH EXPERIENCE

11/2009 – 08/2011	Research Assistant Indian Institute of Science, Department of Molecular Reproduction, Development and Genetics, Bangalore, India
08/2009 – 10/2009	Industrial Internship Avesthagen Private Limited, Bangalore, India
05/2008 – 06/2008	Industrial Internship V.B Medicare Private Limited, Hosur, India

10. Acknowledgements

First and foremost I would like to express my gratitude to my supervisor Prof. Dr. Boris Macek for believing in me and giving me the opportunity to work at the Proteome Center Tübingen on such an interesting topic. My research would not have been completed successfully without his constant guidance and support.

I would like to thank the directors of the IMPRS “From Molecules to Organisms” program – Dr. Elisa Izaurralde and Prof. Dr. Alfred Nordheim, for the organization of a structured program that provided ample opportunities and enabled me to grow as a scientist. I would additionally like to thank Prof. Dr. Alfred Nordheim for taking the time to evaluate my thesis.

I would also like to thank my Thesis Advisory Committee members - Prof. Dr. Alfred Nordheim, Prof. Dr. Oliver Kohlbacher and Dr. Silke Hauf for their valuable insights on my project.

I would like to give a special thanks to Dr. Sarah Danes, the IMPRS “From Molecules to Organisms” program coordinator, for all her help from the time I started, right until the end. I would have been completely lost if not for her various inputs, be it regarding formalities required while starting off as a new student in a new country or during the finishing stages of my thesis and submission procedures.

I would like to thank all members of my lab, past and present – Dr. Boumediene Soufi, Dr. Alejandro Carpy, Dr. Karsten Krug, Dr. Nelson Soares, Johannes Madlung, Katarina Matic, Philipp Spät, Christoph Täumer, Maja Semanjski, Timo Eninger, Silke Wahl, Irina Droste-Borel, Dr. Mirita Franz, Dr. Ana Velic, Uli Grammig, Elias Probst, Bianca Kühn and Boris Günthner for a wonderful working environment and for all the useful discussions during the lab meetings and the PhD meetings that helped me in my research work. I would specially like to thank Gesa Behrends for all the support and for the fruitful discussions during our cherished coffee breaks.

I would like to thank Prof. Dr. Ivan Mijakovic for a wonderful collaboration, for inviting me to INRA, France and allowing me to work in his laboratory. I would like to thank members of his laboratory – Dr. Lei Shi, Dr. Abderahmane Derouiche, Aida Kalantri, Dr.

Charlotte Cousin and Ms. Sandrine Poncet for welcoming me into their laboratory, making my two month stay comfortable and memorable and guiding me at every step.

I am extremely grateful to Dr. Birte Höcker and Kaspar Feldmeier from the Max-Planck Institute of Developmental Biology, Tübingen for their valuable time and help with the circular dichroism measurements.

I am thankful to all my friends here in Tübingen (Prateek, Bala, Janani, Siva, Anurag, Anjali, Martina, Prajwal, Ruchika, *etc*) and back home for all the wonderful times and memories. I would specially like to thank Abhiroop, Priya, Siva and Aayush for patiently listening to me vent my frustration anytime at all and constantly encouraging me at every step; and additionally for all the fun trips and joyous moments.

And last but not the least, I am extremely thankful to my family, especially mom, dad, Maiths, Vasanthi chithi and Raj, for their never-ending support throughout the course of my PhD.