

**Molecular mechanisms of HipA-mediated bacterial
persistence in *E. coli* investigated by mass
spectrometry-based proteomics**

Dissertation

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

vorgelegt von
Maja Šemanjski
aus Slavonski Brod/Kroatien

Tübingen
2018

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

Tag der mündlichen Qualifikation:

19.10.2018

Dekan:

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Prof. Dr. Boris Maček

2. Berichterstatter:

Prof. Dr. Doron Rapaport

Statement of contributions

The research work described in this thesis has been carried out under the supervision of Prof. Dr. Boris Mač​ek at the Proteome Center Tübingen at the University of Tübingen.

Around half of the *E. coli* strains and plasmid constructs used in this thesis were provided by Dr. Elsa Germain and Dr. Szabolcs Semsey as described in Methods section. The other plasmid constructs and strains were cloned and engineered by me, including mutants in HipA phosphorylation sites that were constructed in the group of Prof. Kenn Gerdes, University of Copenhagen, Denmark under supervision of Dr. Elsa Germain. The constructs of *hipA* and *hipA7* on pNDM220 and pMG25 plasmids were designed by me and cloned by Kathrin Bratl under my supervision.

All bacterial cell culture work, proteomic experiments and biochemical assays in this thesis were mainly performed by me, with a few exceptions that involved repetition of experiments in form of biological replicates performed by students under my supervision. For that, Andreas Kiessling made one biological replicate of the dynamic phosphoproteome of the HipA-induced growth inhibition and HipB-triggered resuscitation; Jonas Scwickert performed one biological replicate for direct comparison of HipA and HipA7 produced from the pBAD plasmid; Kathrin Bratl performed one biological replicate for comparison of HipA and HipA7 produced in different amounts; Viktor Beke made two biological replicates for HipA and HipB turnover experiments.

Radioactivity assays were performed in the laboratory of Prof. Kenn Gerdes, University of Copenhagen, Denmark. Stine Vang Nielsen performed two replicates of autoradiography assays for *in vitro* phosphorylation dynamics of HipA and HipA7.

All acquired data was processed, analyzed and interpreted by me under supervision of Prof. Boris Mač​ek. All figures and tables were created by me. Dr. Elsa Germain and Prof. Kenn Gerdes participated in the discussions about the results of phosphoproteomic screens.

Table of Contents

Summary.....	1
Zusammenfassung.....	3
1 Introduction.....	7
1.1 Bacterial persistence	7
1.1.1 Clinical relevance of persistence	8
1.1.2 Stochastic and responsive persister formation	9
1.1.3 Toxin-antitoxin (TA) modules	10
1.1.4 The <i>hipBA</i> TA module and HipA kinase	13
1.1.5 The <i>relBE</i> TA module and RelE mRNAse	16
1.1.6 Signaling pathways and other mechanisms involved in persister formation	17
1.1.7 Metabolic activity of persisters and persister resuscitation	19
1.1.8 Eradication of persisters.....	20
1.2 Mass spectrometry-based proteomics.....	21
1.2.1 LC-MS/MS Instrumentation.....	23
1.2.2 Discovery and targeted proteomic acquisition methods.....	26
1.2.3 Data processing and database search	27
1.2.4 Quantitative proteomics	28
1.2.5 Phosphoproteomics.....	31
1.2.6 Protein kinases and phosphoproteomics in bacteria	34
1.3 Protein turnover	36
1.3.1 Strategies for determination of protein turnover	36
1.3.2 Protein turnover studies in bacteria.....	39
2 Aims and objectives of the thesis.....	41
3 Materials and Methods	43
3.1 Materials.....	43
3.1.1 List of chemicals	43
3.1.2 List of materials	45
3.1.3 List of commercial kits and buffers	45
3.1.4 List of enzymes	46
3.1.5 List of instruments.....	46

3.1.6	List of bacterial strains and plasmids.....	47
3.1.7	List of DNA oligonucleotides.....	49
3.2	Methods	51
3.2.1	Cloning strategy.....	51
3.2.2	Plasmid cloning.....	52
3.2.3	Gene replacement.....	54
3.2.4	Media and antibiotics.....	55
3.2.5	SILAC labeling	56
3.2.6	Dynamic SILAC for protein turnover studies	56
3.2.7	Cell lysis and protein extraction	57
3.2.8	Protein in-solution digestion	57
3.2.9	Phosphopeptide enrichment.....	58
3.2.10	High pH reversed-phase peptide fractionation on commercial spin columns.....	58
3.2.11	Offline high pH reversed-phase peptide fractionation	58
3.2.12	SDS-PAGE and in-gel digestion	59
3.2.13	Incorporation and mixing check.....	59
3.2.14	Stage tips	59
3.2.15	LC-MS/MS measurement	60
3.2.16	MS data processing and analysis of phosphoproteomic experiments.....	60
3.2.17	MS data processing and analysis of protein turnover experiments.....	62
3.2.18	Protein expression for His-tag affinity purification	63
3.2.19	His-tag affinity purification.....	64
3.2.20	<i>In vitro</i> kinase assay measured by MS.....	64
3.2.21	<i>In vitro</i> kinase activity of HipA and HipA7 measured by autoradiography	64
3.2.22	Determination of cell viability on SMG plates.....	65
3.2.23	Measurement of persistence	65
4	Results	67
4.1	Identification of <i>in vivo</i> phosphorylation targets of HipA kinase	67
4.1.1	Phosphoproteomic workflow for studying HipA targets.....	67
4.1.2	HipA phosphorylates multiple proteins in addition to GltX	68
4.1.3	HipA autophosphorylation on Ser ³⁵⁹ has no influence on HipA activity	73

4.1.4	RplK phosphorylation on Ser ¹⁰² by HipA has no influence on RelA-dependent persistence	74
4.2	Comparison of HipA and HipA7 (phospho)proteomes to distinguish mechanistically between different phenotypes of two kinase variants	75
4.2.1	HipA7 has fewer <i>in vivo</i> substrates than HipA	75
4.2.2	Overproduction of HipA7, but not HipA, leads to increased abundance of multiple chaperones and proteases	78
4.2.3	Increasing the overproduction of HipA7 partially reproduces the molecular phenotype of mild HipA overexpression	79
4.2.4	Chromosomally encoded HipA7 phosphorylates GltX and phage shock protein PspA	82
4.2.5	A comprehensive map of the <i>E. coli</i> phosphoproteome	85
4.3	Time-resolved analysis of newly synthesized proteins during toxin-induced persistence and antitoxin-mediated resuscitation	87
4.3.1	Establishment of the method for selective labeling of resuscitating persister cells applied to the <i>hipBA</i> TA module	87
4.3.2	High turnover proteins during early phase of resuscitation are involved in amino acid biosynthesis and translation	90
4.3.3	Selective labeling of persister cells during HipA-induced persistence	93
4.3.4	Proteins involved in translation and transcription are newly synthesized during HipA-induced persistence.....	95
4.3.5	Common mechanisms of HipA- and RelE-mediated persistence	96
4.3.6	Estimation of protein half-lives during persistence	99
5	Discussion	101
5.1.1	Phosphoproteome of HipA-induced growth inhibition revealed multiple potential substrates of HipA	101
5.1.2	Comparison of HipA7 and HipA phosphoproteomes revealed lower activity and smaller substrate pool of HipA7.....	102
5.1.3	HipA overproduction is associated with decreased cell motility and HipA7 with increased protein quality control.....	103
5.1.4	HipA-induced persistence is characterized by low metabolic activity and increased stress response.....	104
5.1.5	Persister resuscitation is characterized by energy production	108
6	Conclusions.....	111
7	References.....	113
8	List of Abbreviations.....	131

9	List of Publications.....	133
10	Curriculum Vitae.....	135
11	Acknowledgements	137

Summary

Chronic infections that can relapse after recurrent courses of antibiotic therapy are a major biomedical problem. Bacteria are able to survive prolonged antibiotic treatments not only by acquiring resistance through genetic mutations, but also by exhibiting persistent phenotype in a small subpopulation of genetically uniform cells. These rare phenotypically distinct cells, known as persisters, become transiently tolerant to antibiotics by restraining their growth and entering a dormant-like physiological state. After antibiotic removal, persister cells can sense improved conditions and resume their growth to produce the same phenotypically heterogeneous population that contains a small fraction of drug-tolerant cells. The emergence of persisters is triggered by the activation of so-called toxin-antitoxin (TA) genetic modules, composed of two genes: one encoding a toxin that interferes with essential cellular processes, and another encoding an antitoxin that inhibits toxin activity. One of *Escherichia coli* toxins is a serine/threonine protein kinase HipA that promotes antibiotic tolerance through phosphorylation of the glutamate-tRNA ligase (GltX), causing a halt in translation, inhibition of growth and induction of persistence. Toxic activity of HipA can be counteracted by the corresponding antitoxin protein HipB through a HipA-HipB interaction and by transcriptional repression of the *hipBA* operon. The first gene associated to bacterial persistence was identified by the isolation of a gain-of-function allele *hipA7* from *E. coli*. The *hipA7* gene encodes a HipA variant, HipA7, that increases persistence markedly compared to the wild-type HipA due to two amino acid substitutions (G22S and D291A) and has therefore been widely used as a model for studying persistence. Whereas ectopically expressed *hipA* inhibits growth and induces persistence, *hipA7* expression does not inhibit growth, yet leads to equal increase in persistence as the wild-type HipA. Based on this, it was suggested that growth inhibition and persistence are separate phenotypes caused by two distinct functions of HipA, which could be explained by different substrate specificities of the two kinase variants. Moreover, because HipA affects multiple cellular functions, primarily protein and RNA synthesis, it is likely that this kinase has more than one substrate. Such relaxed substrate specificity is often seen for other bacterial kinases.

To investigate the differences in HipA- and HipA7-mediated growth inhibition and persistence *in vivo*, I employed a stable isotope labeling by amino acids in cell culture (SILAC)-based quantitative phosphoproteomic workflow in combination with high-resolution mass spectrometry. When mildly produced from plasmids, both HipA and HipA7 phosphorylated GltX as the main substrate, which is likely the primary determinant of persistence. Unlike HipA7, HipA phosphorylated several additional substrates involved in translation, transcription, and replication, such as ribosomal protein L11 (RplK) and the negative modulator of replication initiation SeqA. Conversely, HipA7 exhibited reduced kinase activity *in vitro* and phosphorylated only a few additional substrates *in vivo* only when it was highly overproduced from the plasmid. Furthermore, in the model where *hipA7* was expressed from the chromosome, it led to the increase in GltX phosphorylation compared to wild-type *hipA*, providing a direct evidence that HipA7 targets GltX *in vivo*. Under these conditions, phage shock protein PspA was detected as an additional potential substrate of HipA7. To determine the influence of novel phosphorylation events detected in this study on persistence, the function of RplK site-specific

Summary

phosphorylation was further investigated. However, initial testing did not reveal connection between phosphorylated RplK and persistence dependent on the activity of the GTP pyrophosphokinase RelA. Taken together, this study shows that HipA and HipA7 differ substantially in their kinase activities and substrate pools, which may contribute to their distinct phenotypes. Moreover, these results contribute to understanding of molecular mechanisms of HipA and HipA7. In addition, the phosphoproteome data obtained here yielded a comprehensive collection of phosphorylation events in *E. coli* and can thus serve as a valuable resource for further studies of phosphoregulation in bacteria.

Except for investigating signaling events related to HipA-mediated growth inhibition, the second aim of this work was to establish a method for studying turnover of individual proteins during persistence and resuscitation on a system-wide scale. To that end, I performed a time-resolved analysis of protein abundances during HipA-induced persistence and HipB-mediated resuscitation by implementing a dynamic SILAC pulse-labeling approach in conjunction with MS-based proteomic analysis. This method enabled to selectively label persister cells during antibiotic treatment and determine half-lives of several hundred proteins synthesized under growth-inhibited conditions. Accordingly, analysis of newly synthesized proteins revealed that persistence was characterized by reduced metabolism, cell division and cellular respiration. Conversely, proteins involved in general stress response and translation exhibited higher abundance and turnover. The same methodology was then applied to persistence induced by another toxin, mRNAse RelE, in order to investigate common signature of toxin regulation. Indeed, a high overlap between two experiments was observed, yielding a set of proteins that are actively produced during persistence and therefore likely involved in the maintenance of the persistent state. Notably, proteins involved in stress response, protein folding, protein degradation, RNA production and ribosome biogenesis were newly synthesized during persistence in both models. Although mechanisms of persister formation are relatively well understood, much less is known about molecular processes that provoke their resuscitation, a state that is greatly responsible for the reactivation of the disease. Therefore, pulse-labeling was applied to resuscitating persister cells and the resulting data set revealed a number of proteins, the synthesis of which was triggered at the early stage of the wake-up process, including a positive control, antitoxin HipB. In contrast to persistence, resuscitation was characterized by increased metabolism orientated towards energy production and biosynthesis of amino acids. Altogether, this study shows that the dynamic SILAC approach applied to the model of toxin and antitoxin expression could be used as a general strategy for studying newly synthesized proteins in the context of persistence and resuscitation.

Zusammenfassung

Chronische Infektionen, die nach wiederholter Antibiotikatherapie auftreten, stellen ein ernstzunehmendes biomedizinisches Problem dar. Bakterien können andauernde Antibiotikatherapien nicht nur durch erworbene Resistenzen aufgrund genetischer Mutationen überleben, sondern auch durch das Auftreten des persistenten Phänotyps in kleinen Subpopulation genetisch einheitlicher Zellen. Diese vereinzelt, phänotypisch unterschiedlichen Zellen, die als Persister bezeichnet werden, werden durch ein gehemmtes Wachstum und das Eintreten in einen dormanzähnlichen physiologischen Zustand vorübergehend resistent gegen Antibiotika. Nach dem Absetzen der Antibiotika können Persisterzellen die verbesserten Lebensumstände wahrnehmen und ihr Wachstum fortsetzen, um eine phänotypisch gleichartig heterogene Population aufzubauen, die einen kleinen Anteil an antibiotikaresistenten Zellen beinhaltet. Das Auftreten von Persistern wird durch die Aktivierung sogenannter genetischer Toxin-Antitoxin (TA) Module ausgelöst, die aus zwei Genen zusammengesetzt sind: ein toxinkodierendes Gen, das essentielle zelluläre Prozesse stört und ein zweites antitoxinkodierendes Gen, das die Aktivität des Toxins blockiert. Ein Toxin in *Escherichia coli* ist die Serin/Threonin Proteinkinase HipA, die eine Antibiotikatoleranz durch Phosphorylierung der Glutamat-tRNA Ligase (GltX) vermittelt, was einen translationellen Stopp, Wachstumsinhibierung und Persistenzbildung verursacht. Die Toxinwirkung von HipA kann durch das entgegengewirkende Antitoxinprotein HipB mittels HipA-HipB Interaktion und transkriptioneller Repression des *hipBA* Operons blockiert werden. Das erste mit bakterieller Persistenz in Zusammenhang gebrachte Gen wurde durch Isolierung des Funktionserwerb-Allels *hipA7* in *E. coli* identifiziert. Das *hipA7* Gen kodiert die HipA Variante HipA7, die aufgrund zweier Aminosäuresubstitutionen (G22S und D291A) eine deutlich gesteigerte Persistenz im Vergleich zum wildtypischen HipA verursacht und daher häufig als Persistenzmodell dient. Während ektopisch exprimiertes *hipA* das Wachstum inhibiert und Persistenz induziert, steigert die Expression vom *hipA7* die Persistenz wie im Wildtyp, inhibiert das Wachstum jedoch nicht. Darauf basierend wurde vorgeschlagen, dass Wachstumsinhibierung und Persistenz gesonderte Phänotypen darstellen, die durch zwei unterschiedliche Funktionen von HipA ausgelöst werden und durch verschiedene Substratspezifitäten der beiden Kinasevarianten erklärt werden können. Darüber hinaus ist es wahrscheinlich, dass HipA mehr als ein einzelnes Substrat hat, da mehrere zelluläre Funktionen beeinflusst werden, vorrangig die Synthese von RNA und Proteinen. Eine derartig lockere Substratspezifität ist für bakterielle Kinasen gewöhnlich.

Um die Unterschiede zwischen HipA- und HipA7-vermittelter Wachstumsinhibierung und Persistenz *in vivo* zu untersuchen, habe ich eine auf der Markierung von Aminosäuren mit stabilen Isotopen in Zellkultur (SILAC: stable isotope labeling with amino acids in cell culture) basierenden quantitative Phosphoproteome Analyse in Kombination mit hochauflösender Massenspektrometrie angewandt. Wenn HipA und HipA7 mit niedriger Expressionsrate von einem Plasmid exprimiert werden, phosphorylieren beide Varianten GltX als hauptsächliches Substrat, was wahrscheinlich den primären Faktor der Persistenzbildung darstellt. Anders als HipA7 phosphoryliert HipA mehrere zusätzliche Substrate in der Translation, Transkription und Replikation, wie beispielsweise das ribosomale Protein

Zusammenfassung

L11 (RplK) und den negativen Modulator der Replikationsinitiation SeqA. Im Gegensatz dazu hat HipA7 eine reduzierte Kinaseaktivität *in vitro* und phosphorylierte nur wenige zusätzliche Substrate *in vivo*, wenn es von einem Plasmid aus stark überexprimiert wurde. In einem Model, in dem *hipA7* chromosomal exprimiert wurde, führte dies zu einer verstärkten Phosphorylierung von GltX im Vergleich zu wildtypischem *hipA*, was einen direkten Hinweis darauf liefert, dass HipA7 GltX *in vivo* modifiziert. Unter diesen Konditionen wurde das Phagenschockprotein PspA als potenzielles, zusätzliches Substrat von HipA7 identifiziert. Um den Einfluss bisher unbekannter Phosphorylierungsereignisse, die in der vorliegenden Studie identifiziert wurden, in Bezug auf die Persistenz zu untersuchen, wurde die Funktion eines spezifischen RplK Phosphorylierungsereignisses in Bezug auf das Toxin RelA, einer GTP pyrophosphorylase, weiterführend analysiert. Jedoch zeigten erste Versuche keine Verbindung zwischen phosphoryliertem RplK und RelA-abhängiger Persistenz. Zusammen genommen zeigt die vorliegende Studie, dass sich HipA und HipA7 grundlegend in ihrer Kinaseaktivität und Substratspezifität unterscheiden, was zu ihren unterschiedlichen Phänotypen beitragen könnte. Darüber hinaus tragen diese Ergebnisse zum Verständnis der molekularen Mechanismen von HipA und HipA7 bei. Zusätzlich liefert der hier erzeugte Phosphoproteomedatensatz eine umfassende Zusammenstellung von Phosphorylierungsereignissen in *E. coli* und kann deshalb als wertvolle Ressource für weitere Studien zur Phosphoregulation in Bakterien dienen.

Neben der Untersuchung von Signalisierungsereignissen der HipA-vermittelten Wachstumsinhibierung war die Etablierung einer Methode zur Untersuchung des individuellen Proteinumsatzes während der Persistenz und der Regeneration in einem gesamt-systematischen Maßstab das zweite Ziel dieser Arbeit. Hierzu habe ich mittels dynamischer SILAC Puls-Markierung in Verbindung mit MS-basierter Proteomanalytik eine zeitaufgelöste Analyse der Proteinabundanz während HipA-induzierter Persistenz und HipB-vermittelter Regeneration durchgeführt. Diese Methode ermöglichte eine selektive Markierung vom Persistieren bei Antibiotikabehandlung sowie die Bestimmung der Halbwertszeit von mehreren hundert Proteinen, die unter wachstumsinhibierten Bedingungen synthetisiert werden. Dementsprechend ergab die Analyse neu synthetisierter Proteine, dass die Persistenz durch eine Reduktion des Metabolismus, der Zellteilung und der Zellatmung gekennzeichnet ist. Im Gegensatz dazu wiesen Proteine, die mit der generellen Stressantwort und der Translation verbunden sind höhere Abundanzen und Umsätze auf. Dieselbe Methodik wurde angewandt, um die Persistenz die durch ein anderes Toxin induziert wird, die mRNAse RelE, auf eine gemeinsame Signatur der Toxinregulation zu untersuchen. Tatsächlich konnte eine hohe Übereinstimmung zwischen den beiden Experimenten mit einer Vielzahl an Proteinen beobachtet werden, die während der Persistenz aktiv produziert wurden und aufgrund dessen wahrscheinlich in deren Regulation involviert sind. Insbesondere Proteine der Stressantwort, der Proteinfaltung und Degradation, der RNA Synthese und der Ribosomenbiogenese wurden in beiden Experimenten während der Persistenz neu synthetisiert. Obwohl die Mechanismen der Persistenzbildung relative gut verstanden sind, ist noch wenig über die molekularen Prozesse bekannt, welche die Regeneration auslösen und somit hauptverantwortlich für den Wiederausbruch der Erkrankung sind. Um dies zu untersuchen wurde die Puls-Markierung bei regenerierenden Persistenzzellen angewendet. Der resultierende Datensatz zeigte etliche Proteine auf, deren Synthese in einem frühen Stadium des

Zusammenfassung

Aufwachprozesses ausgelöst wurde, unter anderem für die Positivkontrolle, das Antitoxin HipB. Im Gegensatz zur Persistenz war die Regeneration durch einen gesteigerten Energiemetabolismus und eine erhöhte Aminosäurebiosynthese geprägt. Zusammengefasst zeigt diese Studie, dass die dynamische SILAC-Methode als grundlegende Strategie in Bezug auf das Toxin-Antitoxin Expressionsmodell angewandt werden kann, um neu synthetisierte Proteine im Kontext der Persistenz und Regeneration zu analysieren.

1 Introduction

1.1 Bacterial persistence

The antibiotic resistant crisis occurring worldwide is an acute health problem, as the number of recurrent bacterial infections increases drastically. Many bacteria are categorized as serious threats representing a considerable clinical and financial burden (1, 2). This has been connected with the overuse of antibiotics and the lack of new drugs being developed by the pharmaceutical industry. Since several decades, the research was mainly focused on the isolation and genetic characterization of antibiotic-resistant bacterial strains that emerge through mutations in genes associated with the mechanism of drug action or through a horizontal gene transfer (2). Mechanisms of antibiotic resistance are based on the ability of bacteria to either produce the enzymes that directly modify the antimicrobial molecule, change the drug-target site on the proteins, prevent the drug penetration or extrude the drug out of the cell by the activation of antibiotic-efflux pumps (3). In addition to resistance, bacteria also possess a sophisticated innate strategy that enables them to withstand antibiotics without acquiring genetic changes (4). Bacterial persistence is a phenomenon that is being increasingly recognized in the last decade and, unlike resistance, exhibited within a subpopulation of genetically uniform cells (5). Persisters are defined as phenotypic variants of normal bacterial cells that become transiently tolerant to antibiotics by restraining their growth and entering a dormant-like state (6, 7). While bactericidal antibiotics typically require actively growing cells to exploit their function, in the persistent state, cells are not replicating which makes them immune to the lethal action of antimicrobials. In addition to the slow growth, persister phenotype is also associated with low metabolic activity (8), although persisters are not necessarily entirely metabolically inactive (9).

The discovery of bacterial persistence dates back to 1944, when Joseph Bigger observed surviving colonies of penicillin treated *Staphylococcus aureus* that were equally sensitive to the repeated treatment forming a small subpopulation of viable cells (10). He hypothesized that these rare cells survived the antibiotic treatment due to their lack of growth, which was confirmed only 60 years later with microfluidics and live cell imaging (6). This observation was explained by the biphasic killing curve of exponentially growing culture treated with a high dose of ampicillin (**Figure 1**) (6, 7, 11). An initial, rapid drop in bacterial counts explained that a vast majority of cells was sensitive to the antibiotic treatment. The tail of the killing curve that reached the plateau revealed a small subpopulation of surviving cells that were killed at a much slower rate (12). Based on this insight, it became clear that the biphasic killing curve is a hallmark of persistence and time-dependent killing assays are the standard for quantifying persistence (2). Unlike resistant mutants, persister cells do not proliferate in the presence of the drug. However, after antibiotic removal, persisters are able to resume their growth and produce a heterogeneous population that is as drug-sensitive as the initial one, containing mainly antibiotic-susceptible cells and a small fraction of drug-tolerant cells (13).

Introduction

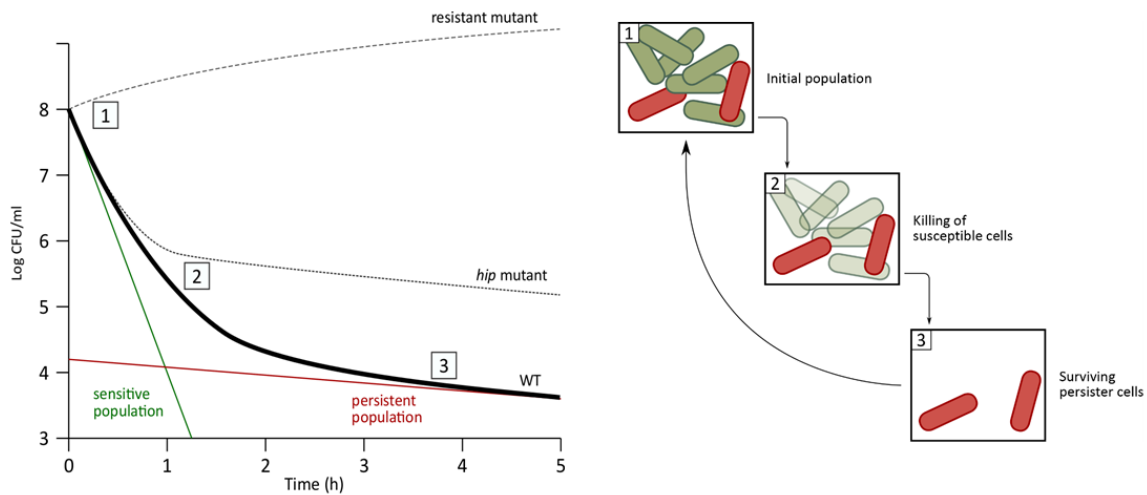


Figure 1. Biphasic killing kinetics of bacterial cultures treated with bactericidal antibiotic. A lethal dose of bactericidal antibiotic added at time point zero rapidly eradicates the drug-sensitive population (green line), while drug-tolerant persister cells are killed at a slower rate (red line). After removal of the antibiotic, surviving persister cells resuscitate and give rise to a population, which is genetically and phenotypically identical to the initial population. Dashed lines show how high-persister (*hip*) and antibiotic-resistant mutants behaved in this model system. Figure from Semanjski and Macek (14).

To summarize, persistence and resistance can be clearly distinguished: persisters are phenotypic variants that cannot pass on their tolerance to their descendants, whereas resistance is hereditary (4). Therefore, re-inoculation of surviving persister cells gives rise to a population with similar small persister fraction (11). Resistance is commonly measured by a minimum inhibitory concentration (MIC) metric. While resistant mutants grown in higher concentrations of antibiotics show an increase in MIC, isolated persisters show no difference. Rather, a minimal duration for killing (MDK) was suggested to be used as a standardized metrics to detect tolerance to drug exposure (2). Mechanistically, persisters are distinct from resistant mutants, as the cellular processes targeted by antibiotics are believed to be inactive in persister cells rendering antibiotics ineffective. Due to the low frequency of persister cells in exponentially growing cultures - typically one in 10^4 - 10^6 - it has been challenging to study this phenomenon. Persister levels generally increase when growth enters stationary phase (11) and they are similar to the levels in slow growing biofilms (15). In recent years, numerous studies begun to unravel the complex molecular mechanisms underlying persister formation and resuscitation. Several laboratory models have been generally used to study the persistent phenotype, including antibiotic treatment of growing and stationary phase cultures, nutrient starvation or rapid nutrient shift to generate persisters, and overexpression of genes involved in the induction of persistence (16).

1.1.1 Clinical relevance of persistence

Persisters are thought to be important in clinical settings and responsible for recalcitrance of chronic infections due to their ability to survive prolonged antibiotic treatments followed by their resuscitation that can cause a therapy to fail. However, only after a discovery of persister cells in biofilms (15), persistence started to draw attention in clinics as a serious medical threat. Biofilm

Introduction

infections are difficult to eradicate because bacterial cells reside in communities enclosed in an exopolysaccharide matrix, where they evade the immune response, but do not strongly restrict the penetration of antibiotics (17). Similar to liquid cultures, planktonic microbial cells showed a biphasic dose-dependent or time-dependent killing with the remaining small population insensitive to increased drug concentrations.

Proving the connection between persistence and therapy failure is a challenging task. Persisters form a small subpopulation that exhibits a temporary phenotype, therefore, it is not possible to introduce them directly into an animal infection model (18). Instead, high doses of bactericidal antibiotics were applied periodically to the clinical isolates of bacterial infections to select for mutants that increase levels of persisters, known as high-persister (*hip*) mutants. This approach was used to determine the presence of persister cells in the isolates of *Pseudomonas aeruginosa* from cystic fibrosis patients (18). Therefore, it was suggested that the emergence of *hip* mutants could be a general feature of recalcitrant infectious diseases. Indeed, *hip* mutants of *hipA* gene were identified in clinical isolates of uropathogenic *E. coli* from urinary tract infections (19). Interestingly, *hip* mutants were also found in the isolates of eukaryotic microorganisms such as the opportunistic pathogenic yeast *Candida albicans* in patients with oropharyngeal candidiasis (20). The use of fluorescent proteins in single-cell imaging enabled first attempts to study persistence in animal models (21). One study examined *Salmonella* Typhimurium persister-cell formation during infection in mice (9). For that, they used a fluorescence dilution method that monitors the extent of bacterial replication at the single-cell level after the termination of green fluorescent protein (GFP) induction combined with the antibiotic treatment. The study revealed the presence of non-replicating bacteria in lymph nodes, which could resume growth *in vitro* in absence of antibiotics. The concept of bacterial persistence can also be applied to other medical problems such as failure of chemotherapy in tumor cells. Similar to bacterial persisters, cancer cell populations are heterogeneous and contain phenotypic variants that are responsive to the same chemotherapeutic agent after remission and can cause a cancer to relapse (21, 22).

1.1.2 Stochastic and responsive persister formation

Although antibiotics enhance the frequency of persister formation, persistence can arise stochastically as an adaptive mechanism of heterogeneous bacterial population to cope with environmental changes. Such evolutionary population-level tactic that is based on the pre-existing phenotypic diversity to increase the survival of an isogenic population in changing conditions, is known as bet-hedging strategy (23, 24). This adaptive strategy involves spontaneous phenotypic switching of a certain number of individuals to different phenotypes, such as dormancy, regardless of the presence of a drug. This concept was for the first time directly linked to persistence in a study, which showed that individual, non-growing *E. coli* persisters form spontaneously before antibiotic exposure (6). It is known that all cellular processes are intrinsically noisy, leading to fluctuations in gene expression in individual cells (25). However, naturally occurring variations are not sufficient to trigger large phenotypic changes unless it involves additional regulatory processes that enhance the signal. In the case of persistence, stochastic persister formation is governed by the fluctuations in the expression of persister genes resulting in coexistence of dormant and growing cells (26).

Besides its spontaneous occurrence, a variety of physiological and environmental factors contributes to the development of persistence and leads to the increase in persister levels (4, 12, 27, 28). These factors include cellular aging (29), nutrient limitation (11), oxidative and acid stress (30), diauxic and rapid nutrient shifts (31-33), heat shock (34), quorum sensing (35), biofilm microenvironment (36) and sublethal concentrations of antibiotics and phagocytosis by immune cells (9). Persister levels depend greatly on the growth phase with very low numbers in the early exponential phase, which increase significantly throughout mid-exponential and stationary phase (11). This can be explained by cellular aging that is characterized by increased protein aggregation, which was shown to correlate with persister levels in *E. coli* (37). The hypothesis of age-induced persistence was also shown in *Mycobacterium smegmatis* treated with isoniazid as older cells showed higher survival probability (29). In addition, decreased nutrient availability in stationary phase cultures can cause starvation that was shown to stimulate persister formation (11). Moreover, reactive oxygen species that are used by the mammalian immune cells to antagonize microbes, such as hydrogen peroxide, lead to an increased persistence in *E. coli* and *S. Typhimurium* (35, 38). Bacterial communication through chemical signaling, a phenomenon known as quorum sensing, can also promote persister formation. Quorum sensing enables bacterial populations to communicate through signaling molecules and coordinate a group behavior by regulation of gene expression in response to fluctuations in cell density. The stationary phase signaling molecule indole, which is normally produced under nutrient-limiting conditions, was shown to induce persister formation in *E. coli* (35). Additionally, when co-cultured with *E. coli*, a non-indole-producing *S. Typhimurium* strain was more tolerant to antibiotics due to indole signal derived from *E. coli*. This suggests that indole signaling between species can induce persistence (38).

1.1.3 Toxin-antitoxin (TA) modules

The genetic basis of persistence was for the first time identified through the selection of chemically mutagenized *E. coli* cells that survived repetitive treatments of high antibiotic doses (39). This led to the isolation of bacterial strains with mutations in specific genes that increased persistence without increasing resistance. In particular, a high-persister mutant of the *hipA* gene, called *hipA7*, was identified as a gain-of-function allele that increased persistence up to 1,000-fold. Since then, several other approaches were used to identify genes involved in persistence, such as screening of deletion mutant libraries (40, 41), transposon insertion libraries (42), overexpression libraries (43), transcriptome analysis of isolated persisters (44, 45) and proteomic analysis of nutrient-shift induced persisters (46).

Since the discovery of *hipA*, a gene that together with *hipB* constitutes a genetic toxin-antitoxin (TA) module (47-49), many other TA modules were found to be involved in persistence, mainly by employing genetic screenings. TA modules consist of two genes: one encoding a toxin that inhibits cell growth, and another encoding an antitoxin that inhibits toxin activity (12, 50). While the product of the toxin gene is always a protein that interferes with essential cellular functions, antitoxin genes encode either noncoding RNAs or small proteins (50). Based on the antitoxin nature and the mechanism by which it neutralizes the activity of the toxin, TA modules are divided in six classes (28).

Introduction

Antitoxins of type I and type III TAs are RNA molecules that inhibit translation of the toxin messenger RNA (mRNA) (type I) or the toxin protein directly (type III). Antitoxins of type II and type IV TA modules are proteins that either form a complex with the toxin (type II) or inhibit the toxin indirectly by reversing its effect on the target (type IV) (28). The antitoxin of the only known type V TA module is an RNase GhoS that cleaves the toxin GhoT mRNA under normal conditions, whereas under stress, the antitoxin mRNA is degraded by the type II antitoxin MqsR, which in turn enables translation of the toxin GhoT (50, 51). The antitoxin of the only type VI TA module is a proteolytic adaptor protein SocA that promotes degradation of its cognate toxin protein SocB by the ClpXP protease system (52). TA modules were initially discovered on bacterial plasmids and had a role in post-segregational killing (53, 54). Since then, TA modules have been found on the chromosomes of many different bacteria (55, 56) including *Mycobacterium tuberculosis* with 88 TA loci (57). The model organism *E. coli* K-12 MG1655 encodes 19 type I, 13 type II and three type IV TA loci (**Figure 2**) (28).

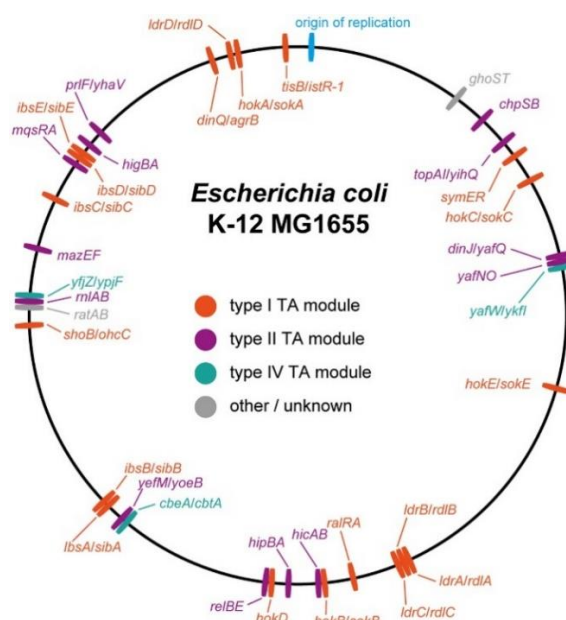


Figure 2. TA modules of *Escherichia coli* K-12. The chromosomal loci of all known TA modules encoded by the *E. coli* K-12 MG1655. Figure from Harms *et al.* (28).

The toxins encoded by TA modules use a variety of molecular mechanisms to inhibit cellular growth (**Figure 3**) (28). Numerous toxins are nucleases that function by inhibiting replication such as DNase RalR (58) or translation, such as ribosome-dependent mRNA endonucleases of the RelE superfamily (59), ribosome-independent mRNA endonucleases (such as MazF and HicA) (60, 61), or VapC family toxins that cleave tRNAs or rRNAs (62). In contrast, some toxins modify their target proteins through posttranslational modification, such as protein kinases HipA and Doc that inhibit translation by phosphorylating their main substrates GltX and EF-Tu, respectively (63, 64). Some other toxins are AMP transferases of the FicT family that adenylate and inactivate DNA gyrase and topoisomerase IV (65) or N-acetyltransferases, such as TacT, that acetylate tRNA leading to translation inhibition (66). The membrane-associated toxin Hok decreases the potential of bacterial membrane leading to a decrease in cellular energy levels and induction of the persistent state (67).

Introduction

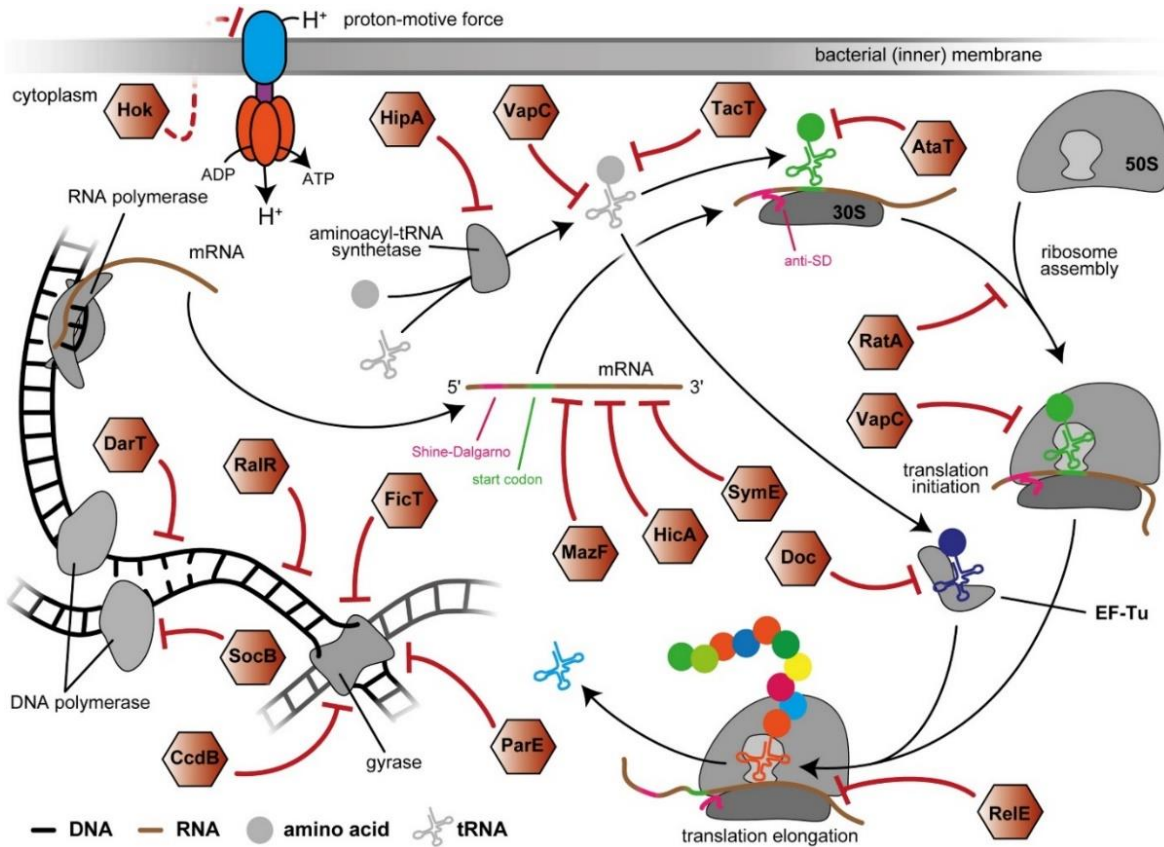


Figure 3. Different molecular mechanisms by which toxins of TA modules interfere with essential cellular processes to inhibit growth. Figure from Harms *et al.* (28).

It is well established that TA modules play a central role in persistence, with thousands of type II TA loci identified across many free-living bacteria (50). Members of type II TA module represent the largest and the best-studied TA class in which both, toxin and antitoxin are proteins. In *E. coli* K-12, 11 of 13 type II TA modules encode toxins which are mRNA endoribonucleases (mRNases) (*relBE*, *yefM/yoeB*, *yafNO*, *dinJ/yafQ*, *higBA*, *prfI/yhaV*, *mqsRA*, *mazEF*, *chpSB*, *hicAB* and *rnlAB*), one is a topoisomerase I specific inhibitor (*topAI/yihQ*) and one is a protein kinase (*hipBA*). The main control mechanism of toxins encoded by type II TA modules is based on the direct interaction of toxin and antitoxin proteins that form a complex, in which the active site of the toxin is blocked. In contrast to the stable toxin, the antitoxin proteins are labile and thus highly susceptible to proteolysis. Therefore, under stress conditions, ATP-dependent cellular proteases selectively degrade antitoxins, which enables a quick release and activation of the toxin. In *E. coli*, antitoxins are predominantly degraded by the Lon protease, while some are recognized by ClpA or ClpX chaperones and degraded by the ClpP protease (68). Elevated amounts of free toxins lead to the generation of persister cells that coexist with the actively growing cells within a genetically uniform population. The regulation of observed phenotypic variability was investigated by a single cell study that established a threshold-based mechanism of TA-mediated bacterial persistence (26). In this mechanism, cells become transiently dormant when the amount of the toxin is higher than a certain threshold of expression, which is specific for each toxin. Fluctuations in toxin amounts above and below this threshold result in co-existence of persistent and growing subpopulations.

Type II antitoxins typically have two protein domains, an N-terminal DNA-binding domain and a C-terminal domain that directly binds and inhibits the cognate toxin (69). Via its DNA-binding domain, an antitoxin can bind to one or more operators located in the promoter region upstream of its respective TA loci to repress the transcription (50). In many cases, toxins can form a complex with antitoxins in variable stoichiometry to regulate transcriptional auto-repression, known as “conditional cooperativity” (70). This transcriptional autoregulation of type II TA modules is based on the ratio of toxin and antitoxin subunits in TA complexes. At low T:A ratios, the TA complex is stable and represses the transcription of TA mRNA, which ensures a minimal expression of the TA locus during normal growth conditions. This enables an immediate release of the toxin from the TA complex upon TA module activation. When the levels of free toxins are elevated due to antitoxin degradation, the T:A ratio increases and destabilizes the binding of the TA complex to DNA, causing transcriptional de-repression of the TA operon. When the signal for antitoxin degradation is gone, expression of TA locus produces high amount of antitoxin, which binds and inhibits the free toxin and restores transcriptional repression (28). Conditional cooperativity was observed for *relBE* TA module (70), but has not yet been shown experimentally for *hipBA*.

1.1.4 The *hipBA* TA module and HipA kinase

The first gene linked to persistence was *hipA* (high persister protein A) gene discovered in *E. coli* by the isolation of a gain-of-function allele *hipA7* in 1983 by Moyed and colleagues (see 1.1.3) (39). It showed a major increase in persistence of up to 1,000-fold due to two amino acid substitutions (G22S and D291A) in the HipA protein (11, 49). This allele has recently been detected in pathogenic and commensal strains of *E. coli* clinical isolates from patients with urinary tract infections (19). Almost a decade after the discovery of the *hipA7* allele, it was found that *hipA* and an upstream located gene *hipB* are organized in the *hipBA* operon (47). Based on the inability to obtain deletion mutants of *hipB* alone (71, 72), it was suggested that HipA is toxic. Indeed, ectopic overproduction of HipA at low levels causes growth inhibition by attenuating protein synthesis, DNA replication and transcription, and strongly increases persistence (73). Growth inhibition can be counteracted by the production of HipB, which interacts directly with HipA and neutralizes its toxicity (72, 73). Therefore, the *hipBA* operon was classified as a type II TA module, which consists of the toxic, 440-residue protein HipA that is co-transcribed with the 88-residue, DNA-binding antitoxin protein HipB (48, 49). Moreover, HipA and HipB can form a protein complex that autoregulates the transcription of the *hipBA* operon by binding to its common promoter (19, 72). Increased activity of HipA during antibiotic treatment is the result of HipB degradation by Lon protease that enables the release of active HipA (74). Unlike other *E. coli* toxins, HipA was suggested to be a protein kinase based on a comparative sequence analysis (75). The predicted kinase activity was supported by the observation that HipA is able to autophosphorylate on Ser¹⁵⁰. The substitution of Ser¹⁵⁰ with alanine yielded an inactive HipA protein, overproduction of which did not abolish cell growth and resulted in decreased persistence to different antibiotics, implicating that the kinase activity of HipA is necessary for persister phenotype (75). Although the HipA toxin was discovered to be involved in persistence more than three decades ago, the core molecular mechanism was unknown until recently. The first structural analysis of HipA revealed the

existence a eukaryotic serine/threonine (Ser/Thr) kinase-like fold, and based on *in vitro* experiments, it was initially proposed that HipA inhibits cell growth by phosphorylating the translation elongation factor EF-Tu (76). However, it was later found that HipA in fact phosphorylates the glutamate-tRNA ligase (GltX, also known as glutamyl-tRNA synthetase GluRS), causing the halt of translation and inducing persistence (63, 77). Phosphorylation of the conserved residue Ser²³⁹ has an inhibitory effect on the aminoacylation activity of GltX (63) preventing the transfer of glutamate to tRNA^{Glu} (**Figure 4**). Phosphorylation of GltX is enhanced upon binding of tRNA^{Glu} to the protein, causing a conformational change of the conserved flexible loop that contains Ser²³⁹, which becomes more exposed in phosphorylated state (63). Consequently, uncharged tRNA^{Glu} accumulates at the ribosomal A-site and triggers the production of (p)ppGpp by RelA (78). Increased amounts of (p)ppGpp inhibit the enzyme exopolyphosphatase (PPX) that degrades polyphosphate (Poly(P)) and causes accumulation of Poly(P) synthesized by polyphosphate kinase (PPK). It has been suggested that Poly(P) stimulates Lon protease to degrade HipB and antitoxins of other type II TAs, which in turn activates the release and production of toxins (79). Most of these toxins are mRNases that inhibit translation further by cleaving mRNA (80).

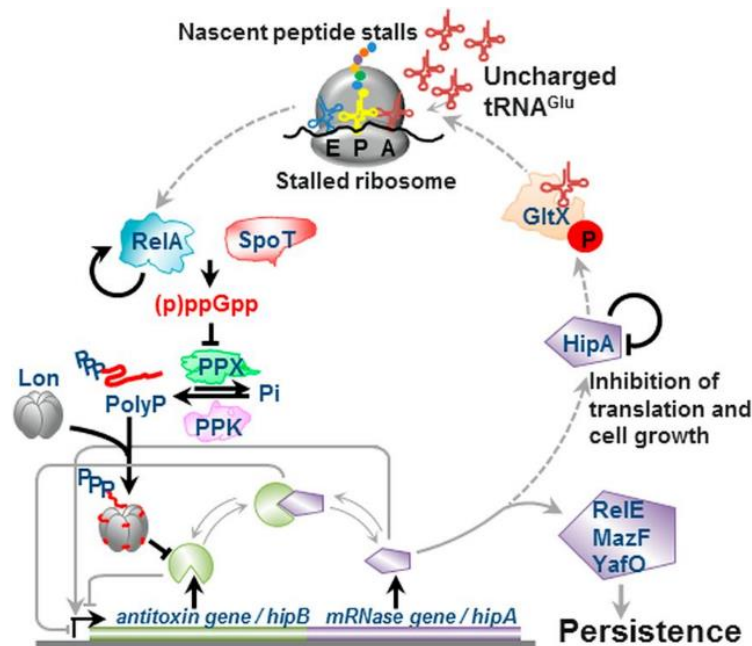


Figure 4. Molecular model of HipA-mediated growth inhibition and persistence and genetic organization of the *hipBA* operon in *E. coli* K-12. Free HipA phosphorylates and inhibits the glutamate-tRNA ligase (GltX); uncharged tRNA^{Glu} accumulates at the ribosome and triggers RelA-dependent synthesis of (p)ppGpp. The (p)ppGpp inhibits the exopolyphosphatase (PPX) and activates polyphosphate kinase (PPK) to synthesize polyphosphate (Poly(P)). Poly(P) stimulates Lon protease to degrade HipB and antitoxins of other TA modules to induce persistence. Figure adapted from Germain *et al.* (79).

HipA is structurally homologous to the human cyclin-dependent kinase 2 (CDK2) and adopts a Hanks-type kinase fold (see 1.2.6) (76, 81), but is distantly related to this type of Ser/Thr kinases (82). It also has all catalytic residues found in other prokaryotic and eukaryotic kinases, including the putative catalytic base Asp³⁰⁹. The first protein structure of HipA in the absence of HipB was obtained by crystallization of the HipA D309Q mutant, in which replacement of Asp³⁰⁹ in the active site prevented HipA to induce growth inhibition and persistence (75). HipA is a 50 kDa globular protein composed of

distinct N- and C-terminal domains, which form a cleft in which ATP binding site is situated. The unusual P loop motif (151-VAGAQEKT-158), that is part of the ATP binding pocket, was found to be disordered in the structure of HipA phosphorylated on Ser¹⁵⁰ (81). Although it was initially suggested that this regulatory site is necessary for the kinase activation (75), structural studies and *in vitro* ATP binding assays showed that phosphorylation of Ser¹⁵⁰ in fact inhibits HipA activity. In the structure of non-phosphorylated HipA, Ser¹⁵⁰ is buried in the protein core and the P loop motif adopts an “in-state” conformation (81). Upon Ser¹⁵⁰ phosphorylation, P loop motif is ejected from the protein hydrophobic core due to steric clashes of Ser¹⁵⁰ and adopts an “out-state” conformation that disrupts the ATP-binding pocket (81). Unexpectedly, the structure of a HipA S150A mutant revealed similar “out-state” conformation of the P loop motif due to inability of Ser¹⁵⁰ to form a hydrogen bond with Leu⁶⁴. Thus, this disordered conformation explains the HipA S150A phenotype (75).

Under normal cellular conditions, HipA activity is neutralized by the antitoxin HipB, which also functions as a transcriptional autoregulator of the *hipBA* operon. HipB is a 10 kDa protein that binds DNA via a helix-turn-helix (HTH) motif (47, 76). Specifically, HipB represses the *hipBA* operon by binding cooperatively to four operators within the *hipBA* promoter region with the palindromic consensus sequence TATCCN₈GGATA, where N₈ refers to a 8 base pair spacer (47). The HipA-HipB₂ complex binds the operators with even higher affinity compared to HipB alone and causes strong *hipBA* operon autorepression (47). In the structure of the HipA-HipB-DNA complex bound to two operators, HipB dimers bind one HipA molecule and induce substantial bending of the DNA strand, enabling dimerization of HipA with other HipA protein and binding of HipB dimers to the second operator (**Figure 5**) (19). This blocks the access of RNA polymerase to the -35 and -10 elements of the *hipBA* promoter, which disables transcription resulting in a low *hipA* and *hipB* expression (19). HipA dimerization blocks the active sites of both HipA molecules, possibly to prevent the ejection of the P loop motif to keep HipA in a non-phosphorylated, active state upon its release. In addition, the sequence of the *hipBA* operators was found in the promoter regions of 33 other genes in *E. coli* genome and it was shown that HipB binds to the promoter region of the *relA* gene (83). Therefore, HipB serves as a transcriptional regulator of multiple other genes beyond the *hipBA* TA module.

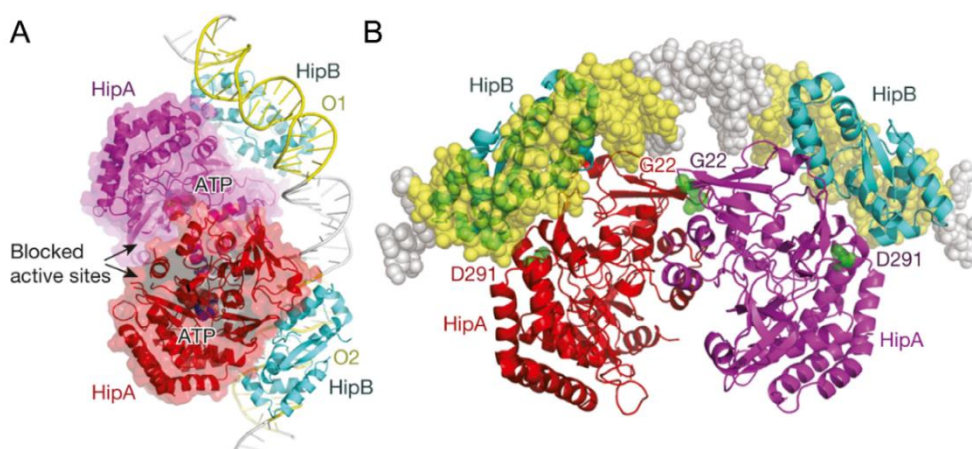


Figure 5. Structure of two HipA-HipB₂-DNA complexes bound to operators of *hipBA* promoter. (A) Active sites of HipA, indicated by ATP molecules, are blocked by HipA dimerization of two HipA-HipB₂-DNA complexes. (B) Locations of HipA7 mutations (G22S, D291A) in two HipA-HipB₂-DNA complexes situated far from the active sites. Gly²² and Asp²⁹¹ are shown as green spheres. Figure adapted from Schumacher *et al.* (19).

The HipA7 mutant (G22S and D291A) is a common model used for studying persistence (5, 6, 26), because bacterial populations carrying the *hipA7* allele exhibit high survivor frequencies of up to 1% when treated with ampicillin (11, 39, 49). Unlike HipA, an ectopic overproduction of the HipA7 variant has only minor effects on cell growth and protein synthesis, but nevertheless increases persistence similarly to the wild-type HipA (73). Although a D291A mutation, located close to the HipA-HipB interface, was initially thought to be important for an observed persistence phenotype of the *hipA7* strain (49), it was later elucidated that this mutation has a dampening effect (19). An increase in persistence was explained by the weakened interaction between two HipA7 molecules in the promoter complex with HipB mainly due to the substitution of Gly²² to serine, located in the HipA dimer interface (**Figure 5**) (19, 26). In this complex, active sites of wild-type HipA are normally blocked, leaving HipA inactive, whereas a G22S substitution impairs HipA7 dimerization and releases HipA7 from the promoter complex. This causes de-repression of the promoter and accumulation of unbound HipA7, which then leads to a higher persistence. Although this model explains a high persister phenotype of *hipA7*, it still remains elusive why HipA7 is seemingly less toxic than HipA. Altogether, a striking difference in the HipA and HipA7 led to the suggestion that persistence and growth inhibition by HipA could be two separate phenotypes caused by two distinct functions of HipA (73).

1.1.5 The *relBE* TA module and RelE mRNAse

The majority of type II toxins are endoribonucleases (RNases) (50), that inhibit translation by RNA cleavage. In *E. coli* K-12, 11 out of 13 type II toxins are mRNAases (28), that either bind to the A site of the ribosome and cleave ribosome-associated mRNA (ribosome-dependent mRNAases), or cleave mRNA in absence of the ribosome (ribosome-independent mRNAases) (84). The *relBE* locus of *E. coli* encodes the RelE toxin, a ribosome-dependent mRNAase that inhibits translation efficiently by mRNA cleavage at the ribosomal A-site of the 30S subunit (**Figure 6**) (85-87). RelE occupies the A site and cleaves mRNA strands after the second nucleotide in specific codons using its catalytic residues Tyr⁸⁷ and Arg⁸¹ (87). Initially, a high codon specificity was observed *in vitro*, with a cleavage preference after the second position of the UAG stop and CAG (glutamine) sense codon, and a lower efficiency for the UAA and UGA stop codons and the UCG (serine) sense codon (88). Later *in vivo* study on a few highly expressed genes and global ribosome profiling data from overproduced RelE revealed a more relaxed specificity with a cleavage preference before G in the third position of the codon (89, 90). Upon RelE cleavage, ribosomes are stalled from damaged mRNA and recognized by transfer-messenger RNA (tmRNA), a small RNA molecule required for rescuing stalled ribosomes and allowing ribosome recycling (91). Indeed, cells lacking tmRNA showed enhanced sensitivity to the presence of RelE and overproduction of tmRNA counteracted RelE toxicity (59). *Trans*-translation is a unique process in bacteria in which tmRNA, together with the SsrA-binding protein SmpB and EF-Tu, adds an alanine from a tRNA-like domain of tmRNA to the unfinished polypeptide. A specific reading frame within tmRNA is translated into a specific short tag sequence (ANDENYALAA) that is added to the polypeptide (92). The resulting polypeptide carries the C-terminal tag, which is recognized and degraded by various cellular proteases (ClpXP, ClpAP, Lon, FtsH and Tsp) in order to protect the cells from non-functional *trans*-translation products (91).

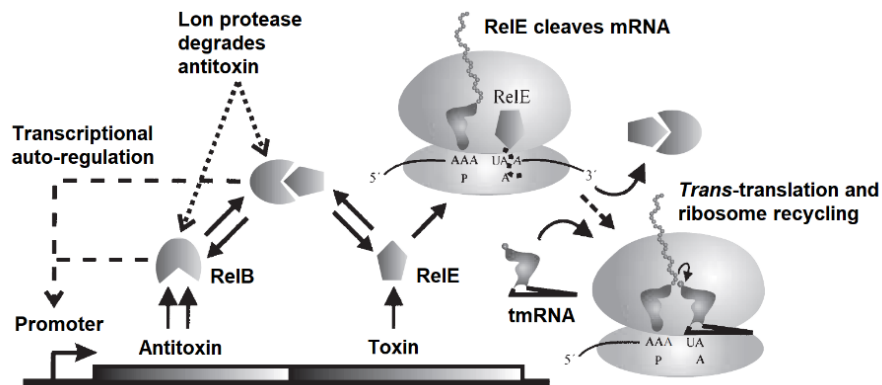


Figure 6. Molecular model of RelE-mediated translation attenuation by mRNA cleavage in *E. coli* and genetic organization of the *relBE* operon. RelB promoter can be transcriptionally auto-regulated by RelB or the RelBE complex. Lon protease degrades RelB and releases RelE toxin, which cleaves mRNA between the second and the third base of the codon in the ribosomal-A site. tmRNA is associated with the ribosome at the A site ready to deliver alanine to the polypeptide chain. Figure adapted from Christensen and Gerdes (59).

Growth inhibition caused by overproduced RelE is counteracted by direct protein-protein interaction with its antitoxin RelB (**Figure 6**) (86). Like HipB, RelB also autoregulates transcription of its TA module by direct binding to the *relBE* promoter region, but RelB alone has only a weak affinity for its operator and the RelB₂-RelE complex represses *relBE* transcription much more efficiently (93). However, an increase in the RelE concentration, caused by the Lon-dependent degradation of RelB (85), changes the RelB:RelE ratio and leads to the formation of an RelB₂-RelE₂ complex with a 1:1 ratio, which cannot bind DNA (71). The destabilization of DNA binding occurs because the additional RelE molecule would clash with the adjacent RelB₂-RelE₂ complex (94). This mechanism of transcriptional autoregulation by the titration of the toxin-antitoxin stoichiometry is known as “conditional cooperativity” (see 1.1.3) and was observed for several other type II TA modules, such as *phd/doc*, *ccdAB*, *vapBC* or *kis/kid* (95-98). Moreover, expression of the *relE* toxin activates transcription of several other TA operons (*mqsRA*, *mazEF*, *dinJ/yafQ*, *hicAB*, *yefM/yoeB*, *prfI/yhaV*) (99). The produced TA transcripts are cleaved and antitoxin-encoding mRNA fragments are rapidly degraded, while toxin-encoding mRNA regions are accumulated. Similarly, transcription of the *relEB* operon was activated in response to the overproduction of other toxins, such as MazF, MqsR, HicA and HipA, but not of YafQ, showing a specificity in transcriptional cross-activation of TAs. Therefore, a mutual cross-interaction of different TA modules and the cleavage of their transcripts towards increase in the toxin products enables a positive feedback loop, which leads to the induction of persistence.

1.1.6 Signaling pathways and other mechanisms involved in persister formation

Various intracellular signaling pathways control formation of persisters, such as stringent and SOS response. The stringent response is a central stress response mechanism that controls persistence through production of the alarmone guanosine tetraphosphate (ppGpp) and guanosine pentaphosphate (pppGpp), collectively known as (p)ppGpp (100). In *E. coli*, amino acid starvation and heat shock activate the GTP pyrophosphokinase RelA to synthesize (p)ppGpp, whereas nitrogen, carbon, iron, phosphate and fatty acid deprivation activate the bifunctional (p)ppGpp

Introduction

synthase/hydrolase SpoT. The (p)ppGpp alarmone is synthesized from ATP and either GDP or GTP by RelA and SpoT and subsequently degraded by SpoT to pyrophosphate and either GTP or GDP (100, 101). (p)ppGpp can also be degraded rapidly by pppGpp pyrophosphatase GppA (102). In exponentially growing cells, (p)ppGpp is present at basal levels and is responsible for the fine-tuning of the metabolism (100). In stress conditions, increased (p)ppGpp levels reprogram the cellular metabolism by altering the transcription of numerous genes and by a direct binding to target proteins (103). During stringent response in *E. coli*, ppGpp interacts directly with the RNA polymerase and, together with the transcription factor DksA, inhibits the transcription from the promoters of stable RNA (rRNA and tRNA) for the expression of amino acid biosynthesis and stress response-related genes (103). For example, a sigma factor RpoS, a master transcription regulator of stationary phase and general stress response, was shown to regulate persister formation (104). Generally, *E. coli* mutants lacking *relA* and *spoT* produce less persisters due to decreased levels of (p)ppGpp (105). High levels of (p)ppGpp are connected to the elevated expression of TA modules and increased activity of toxins, which indicates that TA modules are regulated by the stringent response (4). In favor of this hypothesis, a recent model suggested that increased amounts of (p)ppGpp inhibit exopolyphosphatase PPX and promote synthesis and accumulation of polyphosphate (Poly(P)) by polyphosphate kinase PPK, which stimulates Lon protease to degrade type II antitoxins (**Figure 4**). In contrast to type II protein antitoxins, type I TA antitoxins are RNA molecules, which cannot be targeted by Lon. So far, it has been shown that the expression of toxin HokB is induced independently of Lon or Clp proteases, but by the GTPase Obg, most likely due to direct interaction of Obg with (p)ppGpp (67). The direct interaction of (p)ppGpp with Obg and several other GTPases in *S. aureus* demonstrated their enzymatic inhibition (106).

Because bacterial persistence is based on the inhibition of DNA replication along with other processes, the SOS response was believed to be important for persister formation. Indeed, *E. coli* mutants lacking the SOS response genes *lexA*, *recA* or *recB* showed a reduced ability to form persisters under ciprofloxacin treatment (107). Whereas the SOS response is obviously important for the induction of DNA repair mechanisms in the presence of fluoroquinolones, which are antibiotics that cause direct DNA damage, the SOS regulon was also upregulated in persisters tolerant to β -lactams (108). This is due to the fact that the SOS response activates two type II (*dinJ/yafQ* and *yafNO*) and two type I TA modules (*tisB/istR* and *symE/symR*) (109). For example, SOS response mechanisms are associated with the induction of the toxin TisB to increase persistence by decreasing the proton motive force and ATP levels (110).

Besides TA modules, many other factors have been suggested to be involved in persistence, but their function is much less investigated. For example, the action of aminoglycosides that corrupt the translation can be avoided by ribosome hibernation, a process in which ribosomes form inactive dimers (111). The ribosome modulation factor RMF, a key factor of the ribosome hibernation, was found to be strongly expressed in persister cells (108). Similarly, persisters seem to inactivate cell growth and the main target of β -lactams, peptidoglycan synthesis, in order to escape the effects of this antibiotic (112). Moreover, there was evidence for an increased expression of genes encoding efflux pumps under β -lactam treatment, particularly TolC, which resulted in reduced antibiotic concentrations in persister cells due to an enhanced antibiotic efflux (113).

1.1.7 Metabolic activity of persisters and persister resuscitation

Although persister cells enter a non-dividing dormant state that was assumed to be prerequisite for antibiotic tolerance, it does not mean that persisters are necessarily metabolically inactive. Global transcriptome studies performed either on stochastically formed persister cells isolated by FACS sorting, or on persisters enriched by ampicillin treatment, revealed an increased expression of stress response genes and TA modules, as well as downregulation of flagellar and metabolism associated genes (45, 108). It was later shown that persister cells induced by nutrient depletion are metabolically active (46). The consumption of the available carbon sources for respiratory energy production enables persister cells to sustain the maintenance of the cell associated with growth arrest. Moreover, persisters induced by overexpression of *hipA* and *mazF* were also found to take up glucose and oxygen, and maintain high ATP levels during growth inhibited states, suggesting that persister populations have active metabolism (78, 114). In line with that, enhanced drug efflux activities and an increased expression of the efflux pump *tolC* was associated with a lower drug accumulation and higher persistence in *E. coli* (113). This demonstrated that a passive defense against antibiotics through dormancy and active processes are co-regulated to achieve and maintain persistence. In fact, it was shown by several studies that a lack of growth is not necessary for persistence induction (16). Rapidly growing bacteria can form persisters prior to exposure to antibiotics, and growth inhibition and reduced metabolic activity simply increase the probability of persister formation (115). Likewise, growth inhibition is not a guarantee for persister formation, because only a minority (less than 1%) of a growth arrested population exhibits persistence. Altogether, persistence is triggered by different stress responses occurring in parallel, resulting in a metabolically active, growth-inhibited state (27).

Recently, there have been suggestions that persistence and resistance are more connected than initially assumed. More precisely, persister cells may function as a pool of surviving cells, which are an intermediate state that contributes to the emergence of resistance-conferring mutations (4, 27). The residual metabolic activity and low cell division rates in the presence of antibiotics raised the possibility that persisters accelerate mutagenesis and horizontal gene transfer, due to an induced stress responses. However, further research needs to be done to confer these suggestions.

Persister survival ultimately depends on their ability to resuscitate from dormancy and revert to growth conditions in the absence of antibiotic stress. Persister resuscitation is suggested to be a stochastic event occurring in a temporally random manner, enabling the population to test if the environmental conditions improved (26). The actual molecular mechanisms underlying resuscitation are largely unknown and remain elusive because of the lack of experimental approaches. However, for type II TA modules it has been suggested that conditional cooperativity modulates toxin/antitoxin ratios to promote regrowth and serve for a rapid recovery upon a renewal of the antitoxin pool. It is known that toxin-induced growth inhibition can be rescued by overexpression of the antitoxin or the target (63). A mathematical model proposes that antitoxins buffer increased amounts of free toxin by complex formation and transcriptional autoregulation (116). This results in the increase in antitoxin abundance when the signal for its degradation is absent, and leads to neutralization of the toxin that in turn triggers resuscitation. Other mechanisms could also be involved in a fast response, such as regeneration of corrupted targets in case of the *Salmonella* toxin TacT (66). The acetyltransferase TacT

inhibits translation by acetylation of amine groups of amino acids charged onto their cognate tRNA, whereas the peptidyl-tRNA hydrolase Pth releases uncharged tRNA and promotes growth resumption.

1.1.8 Eradication of persisters

Bacterial persistence is undoubtedly an increasing problem in the public health demanding the development of more efficient therapies to eradicate persisters. The design of new antimicrobial molecules is currently based either on direct killing of persisters or on potentiating conventional antibiotics (4). Despite being dormant and in the state of slow growth, persisters are metabolically active at a low extent with a residual activity, providing a target for new drugs. Therapies that directly kill persisters usually target either the cell envelope or DNA, or involve a combination of antibiotics corrupting multiple targets. For example, antimicrobial dendrimeric peptides with repeating arginine or tryptophan residues effectively eradicated planktonic cells and biofilms of *P. aeruginosa*, probably by disrupting the outer cell membrane (117). Some anti-cancer drugs, such as mitomycin C, were tested in their ability to kill persister cells, although their applicability remains questionable due to strong side effects in patients. Mitomycin C is reduced in the cell cytoplasm where it leads to cross-linking of two DNA strands. An efficient treatment against persisters in planktonic cultures and biofilms of commensal *E. coli* K-12 was proven, as well as against pathogenic strains of *E. coli*, *S. aureus* and *P. aeruginosa* (118). Instead of directly killing persister cells, compounds can also be designed to inhibit their formation. The most obvious way would be to target TA modules by developing toxin-specific inhibitors, which inhibit the transition to the persistence state (50). Due to the high redundancy of TA modules in bacterial genomes, targeting a single TA module would likely not inhibit persistence; rather, a combination of different toxin inhibitors might be required. Several novel HipA inhibitors were discovered using a structure-based virtual screening study. These inhibitors reduced persistence in *E. coli* when combined with ampicillin and kanamycin (119). Furthermore, TA modules can be used as a template for designing drugs that would mimic the activity of toxins and therefore interfere with essential processes in bacteria. Example for this strategy is the ParE toxin, which was used to design and synthesize a series of linear peptides that efficiently inhibited DNA gyrase (120).

Besides the inhibition of persister generation, a promising strategy to eliminate persisters could be the reverting of already existing persister cells to a normal, drug-sensitive state in combination with applying conventional antibiotic treatments. For example, simple sugars like mannitol and fructose can be used to induce the import of aminoglycoside antibiotic like gentamicin by increasing a proton-motive force and thus enhance the antibiotic killing in *E. coli* and *S. aureus* biofilms (121). In addition, fatty acid signaling molecule *cis*-2-deconic acid is able to revert *P. aeruginosa* and *E. coli* persister cells to a metabolically active state and, in combination with ciprofloxacin, efficiently reduce the number of persisters (122). Alternatively to a conventional target-inhibition approach, some targets are rather suitable for corruption in the active state. One example is the acyldepsipeptide antibiotic ADEP4 that activates ClpP-protease and actively reduces its specificity for misfolded proteins (123). This turns ClpP into a non-specific protease causing unregulated protein degradation and self-digestion of the cells. In combination with rifampicin to limit resistance development, ADEP4 efficiently killed *S. aureus* in biofilms and in a mouse infection model.

1.2 Mass spectrometry-based proteomics

Mass-spectrometry (MS)-based technologies have emerged as a powerful method for identification and quantification of proteins and protein posttranslational modifications (PTMs) in biological systems. Proteins are dynamically organized into functional modules and complex networks distributed across different locations in the cell, collectively known as the proteome (124). Overall, MS-based proteomics enables determination of the quantitative state of the proteome through large-scale, systematic analysis of proteins to provide insights into complex phenotypes and biological processes. Current state-of-the-art MS-technology provides high throughput and robust measurements characterized by high sensitivity, accuracy and resolution that are needed for comprehensive studies of entire cellular proteomes. Such complete and reliable analyses on a routine basis are prerequisite for hypothesis-driven investigations and systems biology studies. Tremendous improvements in MS-based proteomics have already been demonstrated by the ability to catalogue the complete proteome of simple organisms such as bacteria or yeast achieved already one decade ago (125). Improvements in instrumentation have allowed comprehensive analysis of the entire yeast proteome is possible in a single measurement (126). Efforts have also been made to map the human proteome by a large collection of experiments from different cell types (127, 128). Recently, the deepest proteome of a human single-cell type has been achieved by the improvements in the chromatographic separation methods and MS instrumentation that enabled identification of more than 14,200 protein isoforms (129).

The most widely used strategy in MS-based proteomics is the bottom-up approach, in which peptides derived from digested proteins are measured by MS and used as proxies for respective proteins (130). In a typical shotgun proteomic workflow proteins are isolated from a biological source such as cells, tissues or body fluids, and enzymatically digested into short peptides (**Figure 7**). Peptides are then separated by liquid chromatography (LC) and subjected to electrospray ionization (ESI) to spray ions directly into an online-coupled mass spectrometer (LC-MS/MS). Separation, ionization, MS measurement and subsequent data interpretation of intact proteins is much more challenging compared to simple peptides due to high complexity of multiply charged protein molecules (131). Therefore, the shotgun proteomic approach is still the method of choice for large-scale studies. Protein digestion with trypsin generates peptides that are typically 10 - 20 amino acids long and their mass-to-charge ratio (m/z) is accurately measured by mass analyzer at the first level (MS1). Following MS1, precursor ion is subjected to a fragmentation and m/z values of specific fragment ions are measured in the MS2 (MS/MS or tandem MS) spectrum. The specific fragment ion pattern of each peptide is matched to the database to retrieve the information about the peptide sequence. In contrast to peptide-based strategy, in the top-down strategy intact proteins are directly subjected to the mass spectrometer where the mass of the entire protein and its fragments are measured (132). This approach is still limited to single proteins and very simple protein mixtures, but has the advantage to achieve complete sequence coverage, full characterization of protein isoforms and quantification of PTMs that coexist on the same protein. However, due to technical challenges regarding proteome coverage, throughput and sensitivity, top-down proteomics is lagging substantially behind bottom-up approaches (133). Recently, a middle-down approaches that analyses 50 – 60 amino acid long N-

terminal peptides have been used to study combinatorial PTMs in the context of heavily modified histone proteins in eukaryotes (134).

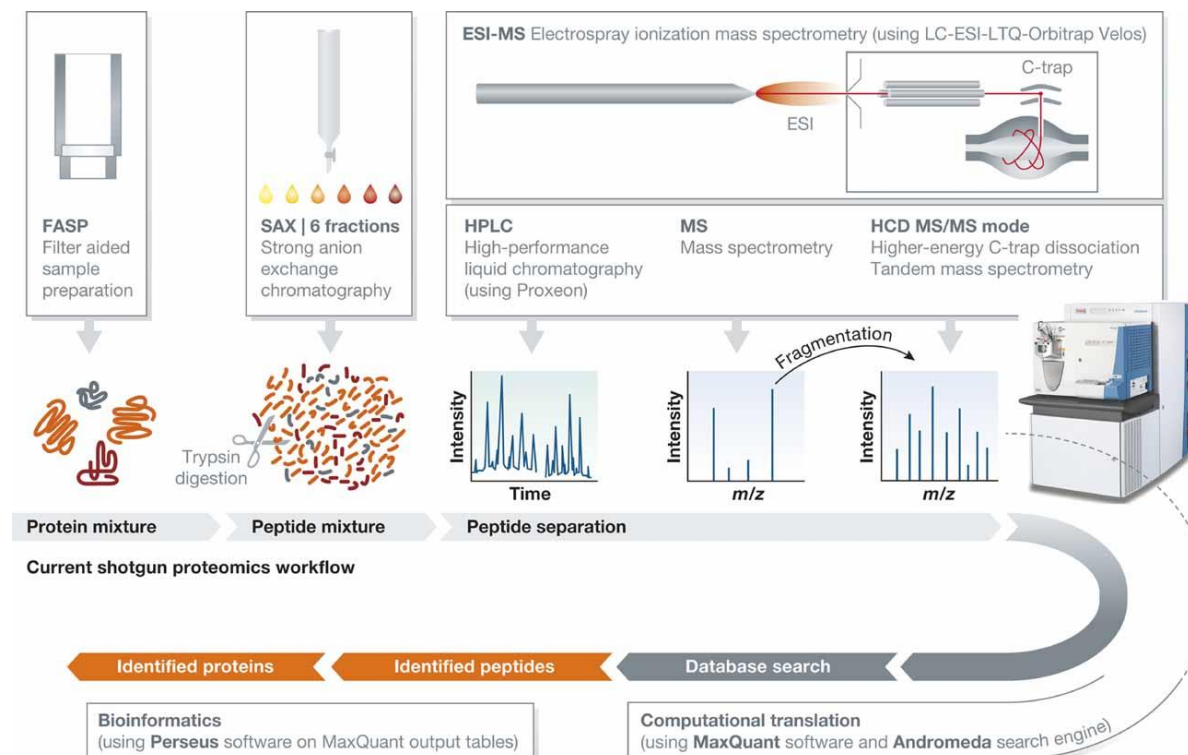


Figure 7. Workflow for high-resolution mass spectrometry-based proteomics. Proteins isolated from cells or tissues are digested into peptides (e.g. by trypsin), which can be first pre-fractionated (e.g. by strong anion exchange chromatography) and then subjected to the nanoscale reverse phase liquid chromatography coupled online to the high-resolution mass spectrometer (LC-MS/MS). As the peptides elute from the column, their accurate mass-to-charge ratio (m/z) is measured in the mass spectrometer, followed by the fragmentation of specific precursor ions to generate characteristic MS/MS spectrum. Figure taken from Mann M. (135).

To achieve deep proteome analysis, the complexity of peptide mixtures can be reduced by introducing a pre-fractionation step prior to LC-MS/MS measurement. Traditionally, two dimensional polyacrylamide gel electrophoresis (2D-PAGE) in combination with matrix-assisted laser desorption/ionization (MALDI) mass spectrometry was the main proteomic technology used before the development of MS-based methods (136). Although 2D gel-based protein separations are outdated and barely used, 1D gel electrophoresis remains a popular approach for protein separation and sample purification, known as GeLC-MS/MS (137). Alternatively, proteins can be separated by a solution-phase isoelectric focusing (IEF) in a gradient of pH. Protein digestion can be carried out either in solution or in the gel by using proteases such as trypsin or endoproteinase Lys-C that cleave specifically C-terminal to lysine and/or arginine residues. Alternatively, other proteolytic enzymes with different specificities, such as chymotrypsin, Asp-N, Glu-C, Arg-C or Lys-N can be used to increase the sequence coverage or to obtain specific peptides that contain a modification at a particular position (136). In addition, digestion with these enzymes ensures that the majority of the peptides are multiply charged following ionization. Protein digestion increases sample complexity further and a standard nano-LC-MS/MS setup, which uses reversed-phase (RP) chromatography to separate peptides under low pH, is not sufficient for extensive fractionation of more complex proteomes. Therefore, efforts

have been made to develop orthogonal off line separation methods with high chromatographic resolution as an additional fractionation step to boost the proteome coverage and dynamic range. Several peptide-base chromatographic techniques that are widely used include strong anion exchange (SAX), strong cation exchange (SCX), hydrophilic SAX (hSAX) (138), hydrophilic interaction (HILIC) or high pH RP chromatography (136, 139, 140). Latter has been increasingly used as the measurement of many fractions in a combination with very short LC gradients and fast scanning MS/MS method enables increased sensitivity and deep proteome coverage (129).

1.2.1 LC-MS/MS Instrumentation

In a standard shotgun LC-MS/MS setup, peptides are first separated based on their hydrophobic properties on an analytical chromatographic column that is commonly packed with reversed-phase octadecyl carbon chain (C18) material. Peptides interact with C18 stationary phase and are resolved according to their hydrophobicity with a gradient of increasing organic solvents compatible with MS such as acetonitrile. The use of long and narrow capillary RP columns with small one micrometer-sized particles improved the sensitivity, resolution and peak capacity (141). However, the backpressure in LC depends on the size of the beads and such long columns combined with small particles increase the backpressure generated by the pumps of the LC system (142). Therefore, ultra-high performance LC (UHPLC) instruments are now routinely used in combination with nanoliter flow rates that allow for low sample amounts and result in high analytical sensitivity due to large concentration efficiency (143). Nanoflow LC (nanoLC) systems are compatible with a soft ionization technique, electrospray ionization (ESI) that revolutionized the field of proteomics (144). ESI can be directly coupled to liquid chromatography allowing for a direct transfer of analyte molecules into the mass spectrometer. Charged analytes are ionized and transferred from the liquid into the gas phase by applying high voltage (2-5 kV). This results in the dispersion of charged droplets into a fine spray followed by the evaporation of solvent molecules and electrostatic repulsion of charged analyte that finally enters the mass spectrometer. In contrast to ESI, MALDI ionization technique uses short laser pulses applied on the analyte that is mixed and crystalized with absorbing matrix (145).

After ESI ionization, charged molecules in the gas phase are guided in the mass spectrometer by pressure differential (vacuum) and by ion optics (RF lens, inject and bent flatapole) using voltages, which also focuses ions and filters neutrals or ions of opposite polarity (**Figure 8**). High vacuum also prevents uncontrolled collision of ions with air molecules that induces their unwanted fragmentation. The essential component of every mass spectrometer is a mass analyzer. The release of the first Orbitrap-based mass spectrometer in 2005, which incorporated an Orbitrap mass analyzer and a linear ion trap (LIT), played a key role in the evolution of MS-based proteomics (**Figure 7**) (146). Since then, a variety of Orbitrap-based instruments have been developed with significant improvements in their performance regarding the resolving power, sensitivity, sequencing speed, dynamic range and mass accuracy (147, 148). The Orbitrap mass analyzer consists of a central spindle and outer barrel-like electrode that enable injected ions to be electrostatically trapped, while rotating around the central electrode and performing axial oscillation (149, 150). Oscillating ions induce an image current into the outer electrodes and these signals are then converted from time to frequency domain by Fourier

transformation. The frequency of axial oscillation is a characteristic of the ion m/z value and therefore frequencies are simply converted into a mass spectrum. The Orbitrap works in combination with an external ion storage device (C-trap), which collects ions and ensures that short ion packages are injected into the Orbitrap (**Figure 8**). Taken together, the Orbitrap mass analyzer enables high mass accuracy of sub-1-ppm, extremely high resolution and fast measurements ideal for high-throughput large-scale analyses. Latest ultra-high-field Orbitrap mass analyzer enables even faster scanning mode and therefore shorter injection times allowing for a high number of MS/MS spectra acquisitions or a shorter total measurement time (147). This improvement also enabled that both, MS and MS/MS, spectra are recorded in the Orbitrap analyzer in the newest generation of Orbitrap instruments (Q Exactive series). Previously, a quadrupole linear ion trap (IT) mass analyzer was used in hybrid MS-systems for ion selection, fragmentation and acquisition of MS/MS spectra due to its higher speed and sensitivity on the cost of the lower resolution and mass accuracy. However, quadrupole analyzers are still used in modern hybrid Orbitrap instruments, as they allow for high ion filtering and isolation efficiency with improved ion transmission, such as in the case of a segmented quadrupole (**Figure 8**) (151).

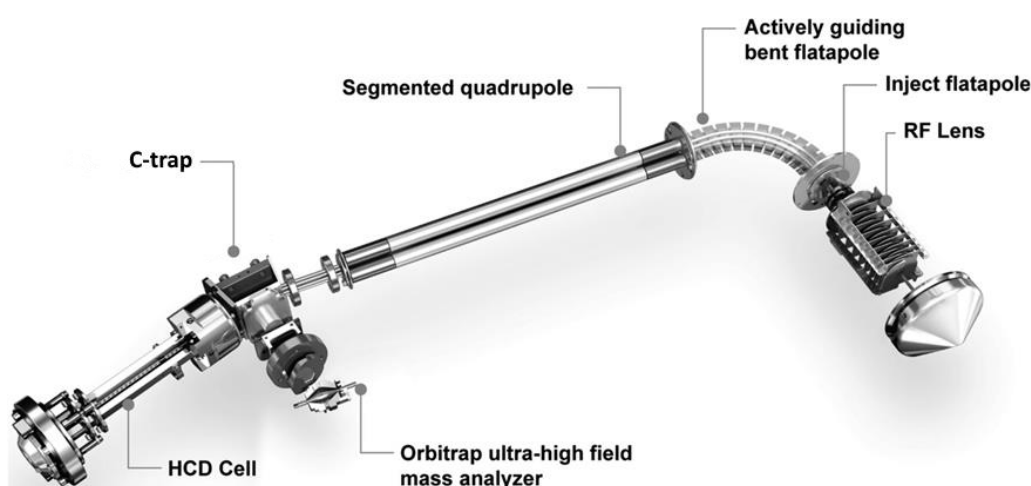


Figure 8. Main components of the Q Exactive HF (Thermo Scientific) mass spectrometer. The instrument consists of a stacked-ring ion guide (S-lens), an injection flatapole, a bent flatapole that ejects solvent droplets and neutral species, a segmented quadrupole mass filter for improved transmission, a C-trap that injects ions into the Orbitrap mass analyzer and the higher energy collisional dissociation (HCD) cell for peptide fragmentation. Figure adapted from Scheltema *et al.* (151).

After MS₁ (full scan) spectrum is recorded in the Orbitrap with high mass accuracy, precursor ions of specific masses are selected in the segmented quadrupole and transferred into the higher energy collisional dissociation (HCD) cell for the fragmentation (**Figure 8**). The HCD cell is an octopole collision cell, in which differential voltages are applied to increase the kinetic energy of the ions which undergo collisions with nitrogen molecules (152). This collision induces fragmentation that occurs primarily on the weakest bonds in the peptide, such as the amide bond of the peptide backbone (131). Predominant fragment ions produced by collision-induced dissociation (CID) in the ion trap or by HCD are called b ions, when the charge is retained by the amino-terminal part of the peptide, and y ions, when the charge is on the carboxy-terminal part (**Figure 9**). Resulting fragment ions are then transferred back to the C-trap and their mass is measured in the Orbitrap. Because the cleavage occurs

Introduction

at different amide bonds in the peptide, fragment ions are of different length and have overlapping sequences, which enables determination of the exact peptide sequence from a generated MS/MS spectrum. In a complementary method, known as the electron transfer dissociation (ETD), electrons are transferred from a reagent anion to the positively charged peptide creating c and z ions (153). Unlike CID and HCD, ETD fragmentation does not cleave labile PTM(s) providing a useful information about the localization of the modification. Whereas ETD fragmentation favors peptides of higher charge states, the combination of ETD and CID/HCD provides more complete fragment ion coverage including peptides of lower charge states (131, 154). Indeed, the combined fragmentation method, called EThcD, is implemented in the novel instruments such as Orbitrap Fusion Lumos that makes use of the deepest sequence coverage and most accurate PTM localization. It also contains a high capacity transfer tube and an electrodynamic ion funnel for improved transmission of incoming ions and an improved peak detection algorithm, named Advanced Peak Determination (APD), which increases the number of assigned precursor charge states in real-time (155). These components are also implemented in the latest Q Exactive HF-X instrument (148), which provides the analysis with high scanning rates of up to 40 Hz and at high resolution of 240,000 defined at 200 m/z (147, 151). Recently, a trapped ion mobility spectrometry (TIMS), an additional ion separation technique, was coupled to the quadrupole time-of-flight (QTOF) mass spectrometer in the timsTOF Pro instrument (Bruker). Using TIMS technique, gas phase ions are accumulated and released sequentially according to their ion mobility in an electrical field, in which ions are separated by their size, shape and mass mobilities (156). This approach enables trapping of precursor ions, which are usually lost during traditional data-dependent analyses during fragmentation of one ion at a time (157). Together with an online scan mode termed parallel accumulation-serial fragmentation (PASEF), which increases sequencing speed (157), timsTOF showed improved sensitivity and enabled measurement of only 10 ng of HeLa digest (158). For comparison, Q Exactive instruments usually require around 500 ng of HeLa sample.

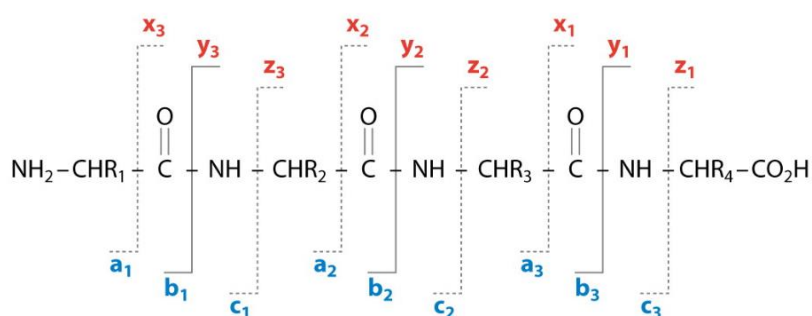


Figure 9. Collision-induced peptide fragmentation along the backbone. Fragments ions that retained positive charge on C-terminus are named a, b, c ions, and ions that retained the charge on N-terminus are named x, y, z ions. Figure from Macek *et al.* (159).

1.2.2 Discovery and targeted proteomic acquisition methods

A mass spectrometer can be operated in different modes to acquire data. The most widely used strategy in discovery proteomics is called data dependent acquisition (DDA), which relies on intensity-based selection of peptides for fragmentation (136). First, the accurate mass of all eluting peptides in a particular time is determined in a survey scan at MS1 level. Afterwards, a specific number of most intense (topN) precursor ions is chosen for further fragmentation. Each of these precursors is individually selected using a narrow m/z window and fragmented into fragments whose mass is recorded in the MS2 spectrum. This cycle of MS1 and topN (usually 10 – 20) MS2 scans is continuously repeated in an automated manner during the entire chromatographic separation. To avoid repeated isolation and fragmentation of same high intense peptides that are typically eluting for a longer time, a dynamic exclusion feature is introduced. It excludes already fragmented precursor ions for a defined period (typically 30 – 60 s) by creating a list of masses in real-time. Despite that, this method suffers from missing low abundance peptides and results in a low dynamic range of detection. Therefore, the proteomic community is now focusing more on the development of the data independent acquisition (DIA) method (160), which circumvents all of these shortcomings as it has a potential to identify all peptides within a sample (161). In DIA analysis, instead of targeting a specific peptide, all peptides within a wide m/z isolation window are fragmented together in a single MS2 event. The isolation windows are overlapping and recorded consecutively to cover the entire m/z range of a full scan. While DIA measurement is relatively simple, the data interpretation is much more challenging because of very complex MS2 spectra that contain fragments of multiple precursor ions (161). Specialized software tools are being developed for processing DIA data and they are typically based on fragment ion spectral libraries. Spectral libraries are sets of refined peptide MS2 spectrum matches recorded *a priori* in a DDA experiment, which are used to match fragmentation pattern of each peptide with DIA derived data. Spectral libraries are also made from synthesized peptides measured by targeted methods (see below) in a method called SWATH MS (162). As a pre-processing step, DIA MS2 spectra derived from co-fragmented precursors is deconvoluted based on the grouping of fragment ions with correlating elution profiles into multiple pseudo-spectra (161) or by more advanced linear deconvolution implemented in Specter software (163).

Because of a stochastic nature of DDA measurements, the reproducibility of such experiments is not sufficient in studies that require quantification of specific proteins in a large number of samples, such as in clinical biomarker validations. Therefore, complementary targeted proteomic workflows that are based on the analysis of a set of selected proteins and peptides were developed (164). Targeted proteomics provides highly reproducible and accurate quantification, which is particularly useful for very low abundant proteins. These assays require the inclusion list that gives the instruction to the mass spectrometer which precursors to choose for the fragmentation and which fragment ions to select for the analysis. Peptides with most optimal sequence and MS characteristics are selected *a priori* from prediction software suites and/or from spectral libraries generated from shotgun experiments or synthetic peptide libraries (165). The precursor ion of interest is first selected in a mass analyzer based on its m/z and retention time, which is followed by its fragmentation in a collision cell and selective detection of predefined precursor-specific fragment ions in the mass analyzer. Several

such precursor/fragment ion pairs (also called transitions) are monitored over time producing a set of chromatographic traces that are defined by the retention time and the intensity of specific transition (166). Single reaction monitoring (SRM) is a classical targeted method performed on a triple quadrupole (QqQ) MS instrument, which enables high selectivity due to two mass filtering steps in the first and the third quadrupole, whereas a second quadrupole serves as a collision cell (164, 166). The extension of SRM to multiple product ions recorded at the same time is known as multiple reaction monitoring (MRM). Targeted assay has also been optimized for high-resolution mass spectrometers in which the Orbitrap mass analyzer serves as a third quadrupole. As it enables scanning of all fragment ions in parallel, it is called parallel reaction monitoring (PRM) (167, 168).

1.2.3 Data processing and database search

Acquired MS proteomic data is extremely complex and thus requires automated data processing methods that automatically assign peptide sequences to the fragmentation spectra. The most often used approach is a database search, in which MS data is searched against the protein sequence database, typically derived from translated gene sequence for the organism of interest. Specialized software packages, such as MaxQuant (169) or Proteome Discoverer (Thermo Fisher Scientific), convert acquired MS and MS/MS spectra into a peak list with accurately determined masses, which are in the MaxQuant additionally corrected for the systematic mass error of the instrument that improves the mass accuracy (170). These programs are used in combination with database search engines, such as Mascot (Matrix Science) or Andromeda (171), that digest *in silico* protein sequences from the database to create theoretical peak lists. First, accurate peptide masses from MS spectra are compared to the theoretical masses within a specified mass tolerance window. Then, the peak list of fragmentation spectra is compared to the peak list of the theoretical spectra predicted by the search engine. Finally, a list of theoretical spectra candidate matches ranked according to the peptide score is created and typically only the top scoring peptide-spectrum-match (PSM) is used for peptide identification and further analysis. The above mentioned search engines use probability-based peptide scoring system. The size of the database search space increases with the number of proteins in the database, the mass tolerance, the number of enzyme missed cleavages specified, labeling state and PTMs allowed. Therefore, statistical tools need to be used to control for true and false peptide matches in the enormous amount of resulting data. Commonly used approach implemented in the MaxQuant is the target-decoy search strategy, in which the target sequences (forward database) are randomized or reversed (decoy database). These databases are searched and decoy hits are used to determine a statistical cutoff for acceptance of PSMs. The most commonly used measure is a false discovery rate (FDR) used as a global cutoff and posterior error probability (PEP) used at the level of individual PSMs (172). For protein identification, peptide hits are assembled into protein hits, in particular, into protein groups that can contain more than one protein if identified based on the common peptide, which is usually the case for protein isoforms.

1.2.4 Quantitative proteomics

To address biological questions, many different quantitative approaches compatible with MS were developed to measure changes in protein and PTM abundance. MS-based quantitation is achieved either by introducing stable isotope labels or by label-free approaches, which can provide the information about relative or absolute protein quantities (**Figure 10**) (136). The most widely used strategy for relative quantification is stable isotope labeling with amino acids in cell culture (SILAC) introduced in 2002 (173). In a typical SILAC experiment, cells are grown in a medium supplemented with either “light” (natural) or “heavy” (isotopically labeled, nonradioactive) form of an amino acid for several cell doublings. After achieving complete proteome labeling, samples are mixed early in the procedure, at the level of cells or proteins minimizing the technical error, and processed together for LC-MS/MS measurement. As isotopes do not change chemical properties of amino acids, differentially labeled peptides co-elute from the LC column and are subsequently analyzed together in the mass spectrometer in the same full scan differentiated only by the mass shift introduced by the label. The relative intensity of peptide peaks corresponding to SILAC pairs in MS1 spectrum can be very accurately quantified to allow for determination of peptide ratios. Multiple MS1 scans can be acquired for each peptide, resulting in highly accurate determination of the peptide ratio that is further extrapolated to the ratio of the corresponding protein. Most commonly used SILAC amino acids are lysine and arginine, which in combination with trypsin digestion ensure that almost every peptide (except for the peptide of protein C-terminus) in a mixture carries the label. SILAC methodology was originally used in mammalian cell culture systems because eukaryotic cells are auxotrophic for lysine and arginine allowing for a complete metabolic labeling by externally added amino acids (173). However, SILAC is also applicable to some autotrophic organisms that can produce their own amino acids, such as bacterial culture. While SILAC labeling with isotopically labeled lysine was successfully applied to *E. coli* grown in defined medium even without the use of lysine auxotrophic mutants (174), other bacteria failed to incorporate it sufficiently to enable quantitative comparison, as in the case of cyanobacteria (175). In some bacteria, such as *Bacillus subtilis*, auxotrophic mutants for lysine can be constructed to obtain optimal SILAC labeling (176). As arginine can be converted to proline in eukaryotes (177), causing difficulties in quantitation, the usage of SILAC amino acids in bacteria is restricted only to lysine (178). Besides cell culture, SILAC labeling was later also applied to human tissues or whole organisms such as mouse or zebrafish (179, 180). The major limitation of SILAC approach is that the maximum number of conditions that can be compared in one experiment is only three. However, multiple SILAC experiments can be compared to each other by using the common sample in one channel, or by using super-SILAC spike-in standard (181, 182).

When the organism or the cell line of choice is incompatible with SILAC labeling, or the experimental design requires multiplexing to more than three conditions, other approaches that are based on chemical, rather than metabolic labeling can be used (**Figure 10**). The most commonly used chemical labels are tandem mass tags (TMT) or its equivalent isobaric tags for relative and absolute quantification (iTRAQ) (183, 184). These chemical tags are composed of a unique mass reporter and a cleavable linker that serves as a mass balancer and carries an amine-reactive group. Labeling is performed on the peptide level after protein digestion ensuring that every peptide is modified on its

Introduction

N-terminus and additionally on lysine residues. All tags have the same total mass and thus are analyzed together in the MS1 scan as one peak, which is subsequently fragmented to release reporter ions of different masses recorded in the MS2 spectrum. The accuracy of the quantification at MS2 level is shown to be compromised by co-eluting peptides with similar m/z that are co-isolated for fragmentation causing underestimation of actual abundance difference, a phenomenon known as ratio compression. The introduction of triple-stage MS (MS3) eliminated interfering ions, although at the cost of scan speed (185). In this approach, fragment ions that carry the tag are isolated at the MS2 level and subjected to MS3 fragmentation (186). This way, precursor identification is performed at MS2 and quantification of reporter ions at MS3 level. Current commercially available kits enable multiplexing of up to 10 conditions in one sample. In addition to TMT and iTRAQ, new labeling reagents have been introduced, such as EASI-tag that fragments at lower collision energy than the peptide backbone enabling generation of peptide-coupled reporter ions at MS2 level (187). Another chemical labeling method that is often used due to a simple and cost-effective procedure is stable isotope dimethyl labeling (188). It is based on chemical modification of peptide amine groups that introduces two isotopically-labeled methyl groups and up to three conditions can be compared in one experiment.

With advances in the speed and resolution of current mass spectrometers, modern quantification approaches are moving from labeling techniques to label-free strategies that overcome limitations of labeling methods, such as time consuming experimental procedure, insufficient labeling, low sample number or quantification of low abundant proteins (136). Label-free quantification (LFQ) is based on the comparison of different samples measured in separate MS runs and can be achieved either by counting the MS/MS spectra corresponding to a particular protein (spectral counting) (189) or by measuring the intensity of the peptide peaks (**Figure 10**) (190). Spectral counting assumes that protein abundance positively correlates with the probability of corresponding peptides to be selected for MS/MS analysis and uses discreet counts of MS/MS spectra as a proxy for protein intensity. Intensity-based approaches are much more accurate as they compare extracted ion currents (XICs) at specific retention times that correlate better with protein abundance. Many algorithms for accurate intensity determination and normalization of label-free data across number of samples were developed, such as MaxLFQ algorithm (191). Although label-free approaches improve the identification rate and the dynamic range due to lower sample complexity compared to labeled samples, they have several disadvantages that include differences in sample preparation, the need for many technical and biological replicates and stable LC-MS/MS systems capable of reproducible measurements across all samples. In addition, label-free approach are more accurate and reliable on protein than on the peptide level, and are therefore less used for quantification of PTMs.

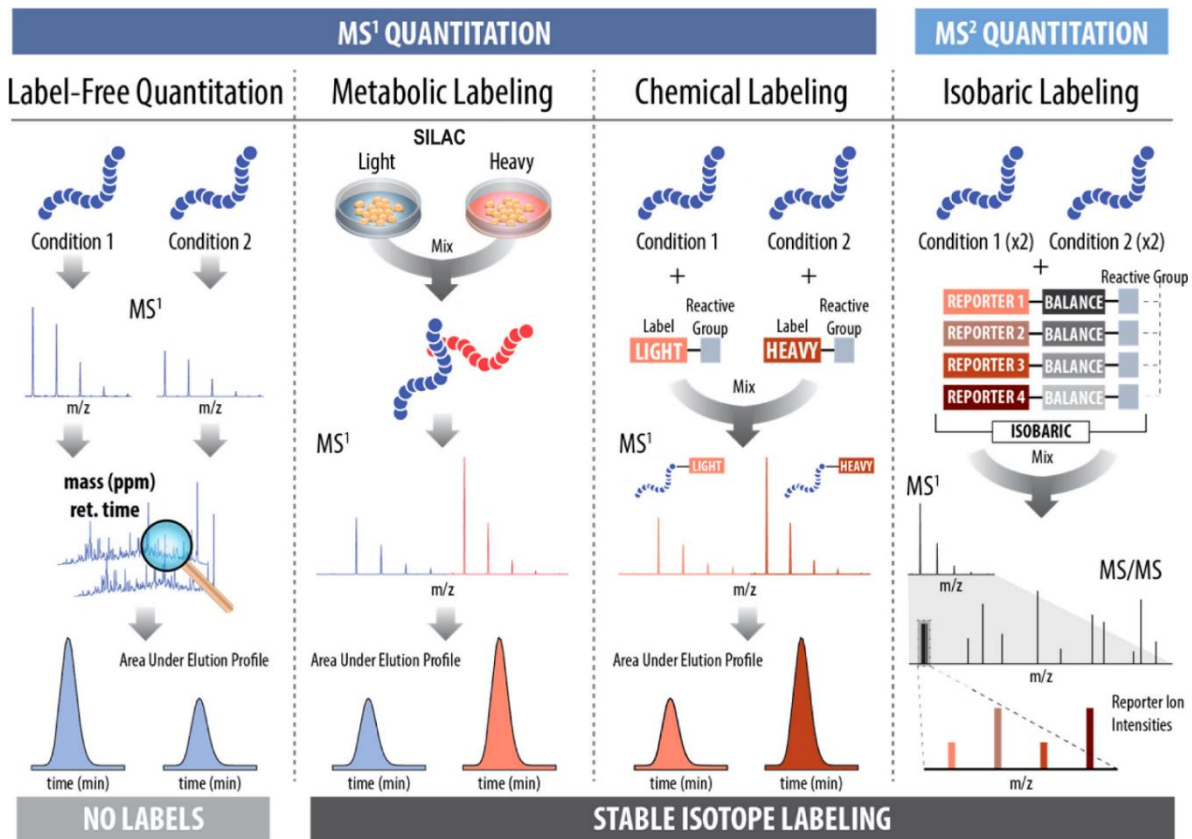


Figure 10. Strategies for quantitative proteomics. Label-free quantitation is performed across separate LC-MS/MS runs using accurate mass and retention time windows for comparison of different samples. Metabolic labeling (e.g. SILAC), chemical labeling (e.g. dimethyl labeling) and chemical isobaric labeling (e.g. TMT or iTRAQ) use stable isotopes for quantitation and samples are mixed after labeling and analyzed together in the same LC-MS/MS run. Figure adapted from Riley and Coon (192).

While relative quantification methods provide valuable information about the altered protein abundances between samples, absolute quantification gives more precise information, as it uses the absolute protein amount described in moles or copy numbers per cell. Most of these methods rely on the use of internal peptide or protein standards of a known concentration, which are often isotopically labeled. For example, absolute quantification strategy (AQUA) uses chemically synthesized stable isotope labeled peptides that are spiked into the peptide sample. The absolute amount of the peptide of interest is determined by comparing its intensity is with the intensity of the spiked-in peptide measured in the same MS spectrum (193). Similarly, isotopically labeled protein can be used as a standard and mixed earlier in the process providing multiple peptides per protein (194). Moreover, SILAC-labeled purified recombinant protein can be added to the cells or tissues in a technique called absolute SILAC (195). Absolute protein amounts can be estimated by using a mixture of unlabeled proteins as an internal standard for other proteins. In particular, the universal proteomics standards (UPS1 and UPS2), that consist of 48 accurately quantified human proteins concentrations of which span six orders of magnitude, are added to the protein sample and quantified by the intensity-based absolute quantification (iBAQ) (196). Briefly, measured protein intensities are normalized to the number of theoretical peptides to calculate IBAQ intensity of each protein. IBAQ intensities of UPS standard proteins are then plotted against their known molar amounts producing a linear calibration curve to which IBAQ intensities of other proteins are correlated to determine their absolute amounts.

1.2.5 Phosphoproteomics

Posttranslational modifications or PTMs control many cellular functions through quick and elegant mechanism. Addition of modification to the protein enables more rapid response to the stimuli compared to gene expression or protein degradation. Considering that PTMs can be studied only at the protein level, MS-based proteomics revolutionized the field of PTMs, not only by allowing the analysis of PTMs in the large-scale, but also by initiating numerous discoveries of novel protein modifications. More than 300 different PTMs have been reported to date among which phosphorylation, acetylation, methylation, ubiquitination and glycosylation have been most extensively studied by MS resulting in identification of tens of thousands sites (197, 198). However, analysis of PTMs is much more challenging, technically and conceptually, compared to the proteome analysis. While one protein can be identified by several different peptides, modification site is unique for one particular peptide that needs to be generated by enzyme digestion and measured. In some cases, tryptic peptides can be too short or too long for MS analysis, thus proteases with different cleavage specificity need to be used to produce modified peptide of optimal size. Due to the low abundance of modified peptides and the substoichiometric occupancy of protein modifications, it is more difficult to identify peptides containing PTMs (199). For that reason, many different strategies were developed for enrichment of PTMs to eliminate the dominant pool of unmodified peptides or proteins. Enrichment methods therefore require much greater starting amounts than classical proteomes. They are based either on metal affinity chromatography, antibodies or other proteins specific for a particular PTM, such as lectins that bind glycosylated proteins. Some PTMs, such as ubiquitin-like modifications pupylation and SUMOylation, are enriched by a pull-down of overexpressed tagged protein that adds the modification (200, 201). Once modified peptides are analyzed in the mass spectrometer, it is important that their MS/MS spectra contain sufficient fragment information, which is needed to accurately localize the modification on a single amino acid. In addition, the sample preparation methods require inhibition of enzymes that remove or add the modification of interest during cell lysis; for example, phosphatase inhibitors must be present in the lysis buffer for phosphoproteomic experiments.

Most common PTM is protein phosphorylation, a fast and reversible modification that is tightly regulated by protein kinases and phosphatases that add or remove the phosphate group, respectively. Phosphorylation can modulate protein activity, its conformation or interaction with other proteins and altering phosphorylation signaling pathways is associated with many diseases such as cancer or chronic inflammatory diseases. Protein phosphorylation is very well studied in eukaryotes where it predominantly occurs on serine (Ser), threonine (Thr) and tyrosine (Tyr) residues via ester bond, as the main switch in many signal transduction cascades (202). Phosphorylation can also occur on other amino acids such as histidine (His), lysine (Lys) or arginine (Arg) through phosphoamidate bonds.

Traditionally, protein phosphorylation has been studied using radioactively labeled ATP (^{32}ATP), Western blot and immunoprecipitation methods with antibodies that target specific phosphorylation sites on particular proteins or phosphorylated amino acids on different proteins. The application of MS allowed the detection and quantification of phosphorylation sites on a large scale with high sensitivity. Unlike other phosphorylated residues, Ser/Thr/Tyr phosphorylation is stable at low pH and

thus most suitable for MS-based analysis. In contrast, phosphorylated His, Lys and Arg are labile and difficult to enrich, hence are much less investigated by MS on the global scale. Aspartate (Asp) and cysteine (Cys) residues can also be phosphorylated. The majority of phosphoproteomic studies use metal ion-based affinity capture methods to enrich for phosphopeptides. Most frequently used techniques are immobilized metal affinity chromatography (IMAC), such as Fe³⁺-IMAC or Ti⁴⁺-IMAC, metal oxide chromatography (MOAC), usually titanium dioxide (TiO₂) or polymer-based metal ion affinity capture (PolyMAC) (**Figure 11**) (203). In these methods, metal ions have high affinity for negatively charged phosphate group, but also for negatively charged amino acid residues such as glutamate and aspartate, which can compete for binding with phosphoryl moiety. However, unspecific binding can be avoided by the usage of very acidic buffers that neutralize the negative charge of peptide carboxyl groups, or by the addition of 2,5-dihydroxybenzoic acid (DHB), which outcompetes acidic peptides while not interfering with phosphopeptide binding (202). These methods are usually performed in multiple sequential enrichment steps to increase phosphopeptide identification or in combination with pre-fractionation strategies to reduce sample complexity. Latter approaches yielded remarkably high coverage of eukaryotic cell line phosphoproteomes, one in a study that identified more than 38,000 phosphorylation sites in HeLa (204) and the other with more than 40,000 phosphosites detected in mice fibroblasts (205). In addition to metal ion-based enrichment approaches that mainly capture Ser and Thr phosphorylated peptides, less abundant Tyr phosphorylation is usually enriched by immunoprecipitation (IP) with specific antibody against phosphorylated Tyr (pTyr). So far, antibody-based approach yielded more than 2,000 Tyr-phosphorylated peptides (204). Due to the negative charge of a phosphate group, that reduces usual positive charge of the peptide (typically from +2 to +1 for singly phosphorylated tryptic peptide), SCX chromatography is also routinely used for phosphopeptide enrichment, often in combination with TiO₂ or IMAC as a second step (206). With recent advances in the speed of current mass spectrometers together with the extensive orthogonal pre-fractionation, it is now possible to obtain comprehensive phosphoproteomes (more than 10,000 phosphosites) even without any enrichment step (129). Finally, most phosphoproteomic studies are performed in quantitative manner to investigate the regulation of phosphorylation signaling pathways in response to stimuli or in disease, and for that, all types of quantitative approaches can be utilized (see 1.2.4).

Introduction

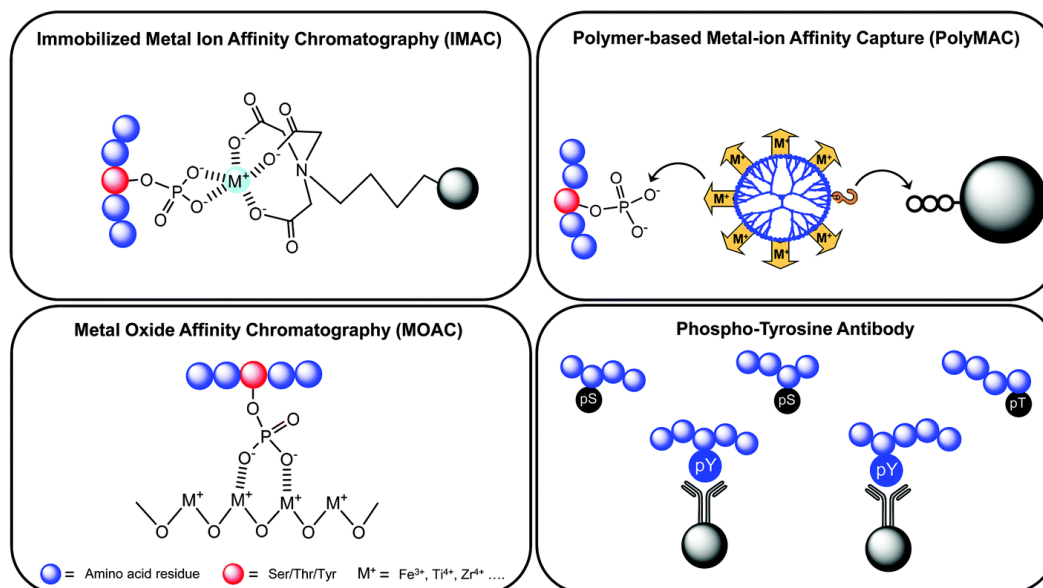


Figure 11. Most common strategies for phosphopeptide enrichment. Negatively charged phosphate group of the phosphopeptide interacts with immobilized metal cations. For Tyr phosphorylation, immunoprecipitation with antibody specific for pTyr is usually used. Figure taken from Arrington *et al.* (207).

Another challenge in PTM studies is the site-specific localization of the modification. The addition of one phosphate group (HPO_3) increases the molecular mass of the peptide by 80 Da. One singly phosphorylated peptide often contains more than one Ser, Thr or Tyr residue that can carry the phosphate group. Therefore, to place the modification on the specific amino acid, a good quality fragmentation spectra needs to be acquired by MS methods optimized for phosphoproteomic experiments. Phosphorylation of Ser and Thr is a labile PTM that is cleaved from the peptide during fragmentation in MS by β -elimination that produces phosphoric acid (H_3PO_4) (208). After this reaction, phosphorylated Ser and Thr are converted to dehydroalanine and dehydro-2-aminobutyric acid, respectively. This leaves a specific mass signature of -98 Da on the fragments that carried the phosphate group prior to fragmentation, called neutral loss. The extent of neutral loss is dependent on the charge of the peptide and the energy imposed for the fragmentation. For example, CID type of fragmentation in the ion trap leads to prominent neutral losses compromising peptide backbone fragmentation, which can cause interference with accurate sequence determination. However, neutral loss fragments can provide useful information when the fragmentation method produces both types of fragment ions (fragments carrying phosphate group and fragments with neutral loss). More sensitive fragmentation technique, HCD, results in a more even cleavage of the peptide backbone, less prevalent neutral loss on γ ions due to consecutive fragmentation, and detects low m/z fragment ions (209). ETD fragmentation is particularly useful for localization of labile PTMs due to its mild nature, which enables direct identification of phosphorylation sites from fragment ions that retain the phosphate group; however the drawback of this technique is longer analysis time. The combination of different fragmentation techniques, such as HCD and ETD (ETHCD) can improve the identification and localization of phosphorylation sites, as it generates different types of fragment ions yielding higher sequence coverage (154). In contrast to Ser and Thr, phosphate group is not eliminated from Tyr, but phosphotyrosine-containing peptides can be identified by the characteristic phosphotyrosine immonium ion at m/z 216 Da (209, 210).

1.2.6 Protein kinases and phosphoproteomics in bacteria

In eukaryotes, protein kinases are usually associated in complex regulatory networks that include multiple cross-phosphorylating kinases with a high substrate specificity. In contrast, bacteria reveal much simpler kinase regulation and it was for long assumed that they do not have such complex signal transduction pathways. However, bacteria use protein phosphorylation extensively and possess several kinase families, including two-component system kinases, Hanks-type kinases, BY kinases and sugar phosphotransferase systems (PTS) (211-213). Two component regulatory systems are hallmark of bacterial sensing and are consisting of a sensor histidine kinase, which autophosphorylates itself on a histidine residue, and a cognate response regulator. The aspartate kinase transfers this phosphate group upon stimulation on the sensor kinase to an aspartate residue on a receiver domain of the response regulator. Most histidine kinases are transmembrane proteins activated by extracellular signals, while response regulators are present in the cytosol and usually modulate expression of a target gene by binding directly to the DNA regulatory elements. Histidine/aspartate kinases were for a long time thought to be the only phosphorylation signaling devices in bacteria (214), until large-scale metagenomics and later proteomic studies revealed that microbial Ser/Thr kinases outnumber histidine kinases (215). Application of MS-based proteomics enable the detection of hundreds of Ser/Thr/Tyr phosphorylation events in bacteria and showed that this form of phosphorylation plays a major role in regulation of diverse processes in bacteria (216, 217). The majority of Ser/Thr kinases in bacteria belong to the family of Hanks-type kinases (82) that was first discovered in eukaryotes (218), but is widely identified across bacterial species with varying frequencies. Hanks-type kinases are usually transmembrane proteins that contain a highly conserved cytosolic kinase catalytic domain and an extracellular sensor domain (219). One example is the *B. subtilis* PrkC kinase that is active in dormant spores and required for spore germination (176, 220). Nevertheless, all Hanks-type kinases contain an activation segment, the most important structural element that regulates kinase activity. The activation loop determines the substrate specificity together with other structural elements and can be phosphorylated on Ser, Thr or Tyr residues by autophosphorylation or transphosphorylation through other kinases (219). In most kinases, phosphorylation of the activation loop leads to a conformational change that allows substrate binding and activation of the kinase. Other important elements are a P+1 loop that interacts with the substrate and determines the kinase specificity, and a glycine-rich consensus motif, known as P loop, which is important for ATP/ADP exchange during the catalytic reaction. Once ATP is positioned in the active site, catalytic Asp residue transfers the γ -phosphate from ATP to specific Ser, Thr or Tyr residue of the substrate.

Whereas all eukaryotic tyrosine kinases are structurally similar to Hanks-type kinases, bacteria have also evolved a unique tyrosine kinase family, the BY kinases (213). These enzymes contain a P loop motif that catalyzes phosphorylation of tyrosine residues on other proteins. Compared to Hanks-type kinases, BY kinases are less frequently encoded in bacterial genomes (221). In *E. coli*, two BY kinases were identified, namely Wzc and Etk. Lastly, PTS system involve a cascade of usually histidine phosphorylated proteins that use phosphoenolpyruvate as a phosphate source to phosphorylate specific sugars for transmembrane import. In addition to the above mentioned kinase families, there are many examples of unusual kinases that are mostly unique to the certain species. One example is

Introduction

a stress response kinase A (SrK_A, previously known as RdoA or YihE) of *E. coli*, which is homologous to the catalytic core of Hanks-type kinases, but does not have all kinase motifs conserved (222).

Current large-scale phosphoproteomic studies identify tens of thousands phosphorylation sites in eukaryotes on a routine basis. In contrast, protein phosphorylation is much less abundant in prokaryotes, making the analysis of bacterial phosphoproteomes a challenging task. Thus, phosphopeptide enrichment workflows require higher amounts of starting material from which very low abundant phosphopeptides can be enriched. Until now, the most comprehensive studies of the *E. coli* phosphoproteome evolved from initial 81 (223) and 150 Ser/Thr/Tyr phosphorylation sites (224) detected from >10 mg of proteins, to 1,088 (225) and 1,883 sites (226) from less than 2 mg of protein input. The distribution of phosphorylation sites detected in *E. coli* usually shows a following pattern: around 70% of phosphorylation sites are located on Ser, 22% on Thr and 8% on Tyr residues. For comparison, phosphorylation in *M. tuberculosis* is biased toward Thr (around 60%) compared to Ser (around 40%) and 3% on Tyr. The largest phosphoproteomic studies in *M. tuberculosis* covered around 500 phosphorylation sites in virulent and avirulent strains of this bacterium, as well as in the clinical Beijing isolate (227-229). Despite large phosphorylation data sets obtained by proteomic studies in bacteria, the knowledge on the link between kinases and their specific physiological substrates is still largely missing. Considering that bacterial genomes encode only a dozen of kinases on average, it is generally accepted that bacterial Ser/Thr kinases have broader substrates specificity. For example, only 4 Ser/Thr and 2 BY kinases are encoded in the *E. coli* genome according to UniProt. Accordingly, it was shown that bacterial kinases phosphorylate a number of different protein substrates such as for the *B. subtilis* Ser/Thr kinase PrkC (176) and the BY kinase PtkA (230). Yet, knowledge about the functional role of discovered phosphorylation events is missing. Moreover, unlike in eukaryotes, no specific kinase consensus motif have so far been attributed to bacterial kinases, suggesting that they possess a less stringent specificity. As recently reported, bacterial kinases are also capable of cross-phosphorylation, making the regulatory network more complex than previously thought (231). One example is the cytosolic two component sensor kinase DegS of *B. subtilis*, which is phosphorylated on a serine residue by the Hanks-type kinase YbdM, followed by a transfer of the phosphate to the response regulator DegU (221).

Due to chemical instability of His and Arg phosphorylation at lower pH values (pH <8), the major focus of global phosphoproteomic studies has been on the acid-stable Ser/Thr/Tyr phosphorylation. However, recent optimizations of sample preparation workflows and applications of more mild fragmentation techniques in mass spectrometers enabled the use of MS to study histidine phosphorylation in a comprehensive manner (226). More importantly, it led to the discovery of arginine phosphorylation in the Gram-positive model bacterium *B. subtilis* with 121 arginine phosphorylation sites identified in a mutant lacking the protein arginine phosphatase (YwIE) (232). Arginine phosphorylation, catalyzed mainly by the protein arginine kinase McsB, was quickly after shown to be used as a degradation signal that is selectively sensed by the ClpC-ClpP proteolytic complex (233). This unique bacterial degradation system regulated through phosphorylation is functionally analogous to the ubiquitin-proteasome system in eukaryotes.

1.3 Protein turnover

Proteomes are much more complex compared to the genomes, due to different protein abundances, isoforms or modifications. A further dimension of the proteome complexity is associated with the protein turnover. Although in steady-state conditions the abundance of proteins is constant, they are in a dynamic state of turnover in which proteins can be turned over rapidly. More precisely, the balance between protein synthesis and degradation determines the total concentration of a protein and controls cellular homeostasis (234). Maintaining protein turnover is energetically expensive for the cell and its main purpose is to alter protein abundances in response to environmental stresses (235). The increase in protein abundance can, for example, be obtained by a faster protein synthesis or a slower degradation, and vice versa (236). Determination of the protein turnover requires direct measurement of proteins and their dynamics, rather than using transcript abundances as a proxy for protein abundance. This is supported by a comprehensive study that showed a poor correlation between protein and mRNA abundance, which led to the conclusion that protein abundance is regulated predominantly at the level of translation (196). Because turnover rate controls the cellular response to metabolic changes, regulatory proteins are likely to have higher turnover rates. Indeed, many transcription factors and proteins involved in signaling have higher turnover rates, while housekeeping proteins tend to be more stable (196).

The simplest model used for protein turnover calculation is based on the assumption that protein synthesis can be described with a zero-order kinetics and degradation with a first-order kinetics (235, 237). As a zero-order process, protein synthesis rate (k_{syn}) is independent on the abundance of a protein (P) and changes linearly with time. Conversely, protein degradation rate (k_{deg}) is proportional to the protein abundance. The change in the protein abundance is a function of time:

$$dP/dt = k_{syn} - k_{deg}[P] \quad (1)$$

In the steady state, protein abundance is constant, therefore $dP/dt = 0$, and the final equation is:

$$[P] = k_{syn}/k_{deg} \quad (2)$$

Usually, a protein half-life ($T_{1/2}$) is often used as measure of the protein turnover and is calculated from the degradation rate constant according to the equation in which $T_{1/2}$ is inversely related to k_{deg} :

$$T_{1/2} = \frac{\ln(2)}{k_{deg}} \quad (3)$$

Therefore, in turnover studies, the degradation rate constant should be ideally the parameter reported which is independent of the initial protein concentration and is defined by the unit of reciprocal time.

1.3.1 Strategies for determination of protein turnover

Early studies on protein turnover were based on the measurement of the bulk protein turnover or the analysis of the stability of single proteins in small-scale studies (238, 239). They included pulse-chase radiolabeling methods that used radioactive reagent such as amino acids (e.g. ^{35}S -methionine or ^3H -

leucine) in combination with immunoprecipitation or 2D gels and detection of radioactivity (240, 241). Other approach is based on chemical inhibition of protein synthesis and detection by Western blot (242). However, translation inhibitors impose stress and cause perturbation of cellular homeostasis, therefore, determined protein half-lives do not reflect the real situation in the intact cells. Tagged proteins have also been used to study dynamics of individual proteins, for example by generating whole-genome libraries of open reading frames (ORFs) tagged with tandem affinity purification (TAP) tag in yeast (243). When containing fluorescent tags, such as the green fluorescent protein (GFP), proteins can be separated by flow cytometry and detected by fluorescent techniques, for example by using fluorescence-activated cell sorting (FACS), as in a study that determined stability of more than 8,000 human proteins (244). Later, tandem fluorescent protein timers (tFTs), an advanced version of fluorescent tags that combines two single-color fluorescent proteins with different maturation kinetics was introduced to analyze protein turnover and mobility in cells (245). Despite many developments, protein tags are not an ideal model for measurement of physiological turnover rates due to artificial tags that need to be cloned individually and can disrupt the structure and physical properties of tagged proteins (238, 239).

Advances in the resolution and throughput of next-generation transcriptomic technologies, such as RNA-Seq, enabled the measurement of RNA amounts on a global scale. Temporal analysis of nascent transcript was possible by metabolic labeling with 4-thiouridine amounts of which can be used to estimate synthesis and degradation rates of proteins (246). However, numerous studies reported a poor correlation of mRNA and protein abundances. Finally, recent advances in MS-based technologies provided the opportunity to measure a native protein turnover of thousands of endogenous proteins in parallel. This was achieved by adapting classic pulse-chase strategies that use metabolic labels that contain stable isotopes, such as amino acids (SILAC), glucose, water or different nitrogen sources, which are labeled with ^{15}N , ^{13}C or ^2H isotopes (235). Most comprehensive studies used SILAC methodology in which a pulse of labeled amino acid is added to the growth medium (247) or introduced through the diet for animal models (179). Protein turnover is then determined by monitoring a time-dependent incorporation of the labeled amino acid into newly synthesized proteins using LC-MS/MS, while pre-existing proteins remain un-labeled. Likewise, cells can be pre-labeled with the “heavy” version of the amino acid and pulse-chased with its “light” version to monitor the loss of the label. This simple approach was originally introduced via a method called “dynamic SILAC” (236, 247) and was later used in many studies (**Figure 12**, right). Another strategy, known as “pulsed SILAC” (pSILAC), used two isotopically labeled versions of the same amino acid to measure relative changes in protein synthesis between two different conditions (**Figure 12**, middle) (248). In this approach, newly synthesized proteins of two differential treatments incorporate “medium-heavy” or “heavy” amino acid version, and their direct comparison reflects the difference in translation under the assumption that degradation occurs at equal rates for both conditions. Standard SILAC-based protein turnover analyses require multiple LC/MS-MS measurements of different time points and therefore can suffer from missing quantitative information due to the DDA mode of measurement. Therefore, a combination of SILAC and TMT-labeling approach was recently implemented for the analysis of multiplexed samples (249). This TMT-SILAC hyperplexing methodology enables measurement of protein turnover kinetics in a single LC-MS/MS run reducing the cost and the number of missing values,

but, on the other hand, has limited quantitation precision and proteome coverage compared to the dynamic SILAC. This methodology was used to study dynamics of protein isoforms for which the quantification is performed on the peptide level (250).

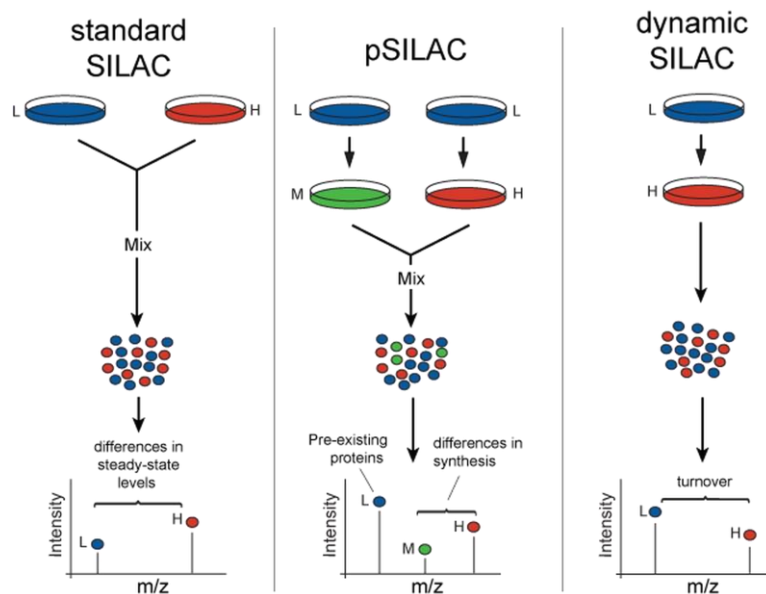


Figure 12. Different SILAC-based workflows for the protein dynamics analysis. Dynamic SILAC is used to determine protein turnover and pSILAC for quantitative comparison of protein translation rates. Figure from https://www.mdc-berlin.de/research/research_teams/intrazellul_re_signalwege_und_massenspektrometrie/project.

Proteins with high and low turnover rates reach their maximum level of label incorporation within a different time window, therefore a time range of sample collection should be adjusted accordingly. It is also preferable to measure multiple time points for one turnover experiment, as they decrease the influence of outliers and enable confident quantification. Moreover, the profiles of high and low turnover proteins will show different kinetics, hence the corresponding curves need to be fitted by appropriate mathematical functions (235). A further complication concerns the recycling of amino acids released from degradation of pre-existing proteins into the medium. This causes a dilution of labeled amino acid pool with un-labeled amino acids and results in apparently lower turnover rates that cannot truly describe actual values. However, these values can be corrected using a recycling factor that can be calculated by measuring label incorporation in partially labeled missed cleaved peptides (196). Another important feature that has to be taken into account when determining protein turnover is the rate of the cell growth, because most of the studies use exponentially growing cells in culture that produce new proteins through cell division. Unlike non-dividing cells, growing cells incorporate the label to much higher extent, which enables more accurate quantification of these large changes. Compared to perturbed systems, in steady-state conditions, protein concentrations are constant and any change in protein abundance can only come from cell division (238). Therefore, cellular growth rate needs to be determined and subtracted from the measured flux (overall turnover). In the case of non-dividing cells, the change in protein abundance reflects true intracellular turnover, as the growth rate correction is not necessary, simplifying the turnover rate calculation. However, the labeling efficiency in these conditions is largely limited making quantification more complicated and

less accurate. Finally, the non-steady-state growth or perturbed systems cannot assume constant protein concentrations, and thus measured values need to be corrected for differences in the abundance of individual proteins across time points (239). This can be achieved by spiking-in a standard, which contains a mixture of known proteins that serve as an internal control. Despite all of these concerns, the simple model undoubtedly provides a useful information about the rank of turnover rates and can be used for comparative studies, in which an absolute value is not crucial.

1.3.2 Protein turnover studies in bacteria

Most of the protein turnover studies were performed in eukaryotic systems and the knowledge on proteome dynamics in bacteria is still quite limited. So far, only several publications investigated bacterial protein turnover on a global-scale by using advanced MS-based approaches. For example, one global study integrated a data on mRNA and protein abundance with protein half-lives of a genetically very simple bacteria *Mycoplasma pneumoniae* using dynamic SILAC (251). Another study used a combination of SILAC and iTRAQ labeling techniques to determine protein turnover rates in the *Streptomyces coelicolor* cultures during their transition from the exponential to the stationary phase (252). Besides growing bacteria in batch cultures, one study analyzed the changes of *S. Typhimurium* protein turnover rates in the cell culture-infection model using dynamic SILAC, which showed increased turnover during initial stage of infection (253). In another study, dynamic SILAC approach was used to identify newly synthesized proteins in the context of antibiotic tolerance, more precisely, in persistent *P. aeruginosa* biofilms after colistin treatment (254). Here, biofilms grown in “light” lysine were treated with antibiotic colistin to kill sensitive population and the remaining antibiotic-tolerant cells were pulse-labeled with “heavy” lysine. This study demonstrated that colistin-tolerant population actively produced ribosomal proteins suggesting that these cells are metabolically active. It also measured a higher production of proteins required for migration and cell-cell signaling (quorum sensing) and showed their importance during antibiotic tolerance, but also in the recovery of biofilms.

Except from natural amino acids, their analogues, such as methionine analogue azidohomoalanine (AHA), can be used to differentiate between newly synthesized and pre-existing proteins (255). The advantage of metabolic labeling with AHA is the possibility to enrich for labeled proteins by affinity purification using chemoselective click-chemistry-based reaction. Such selective enrichment increases the sensitivity of the measurement as it enables the isolation of newly synthesized proteins in the large background of pre-existing unlabeled proteins (239). However, a major drawback of this method is the toxicity of AHA that causes severe cell growth defect shortly after the pulse of AHA, due to the translation inhibition by incorporation of a non-canonical amino acid. Therefore, proteins detected in these conditions cannot provide the true information on protein synthesis during normal, unstressed growth. However, as a proof of principle, AHA-based labeling approach was applied to the *E. coli* culture for a very short time, during which the cell growth was not perturbed (256). The culture was then exposed to a short temperature shock and compared to the unstressed control to identify newly synthesized proteins during the heat shock. Altogether, this short overview shows how the knowledge on protein turnover in bacteria is still very limited and thus demonstrates that further studies are needed to provide basic information on stability of individual proteins in different conditions.

2 Aims and objectives of the thesis

The overall aim of this thesis was to characterize proteome and phosphoproteome of persistence mediated by toxic bacterial kinase HipA in *Escherichia coli*. Because HipA affects multiple essential cellular functions, it was assumed that HipA was likely to have more than one protein substrate. Therefore, the aim was to identify additional phosphorylation targets of HipA beyond GltX using MS-based phosphoproteomic workflows in combination with SILAC labeling. Furthermore, the second aim was to investigate the molecular mechanisms underlying different phenotypes of HipA and its gain-of-function variant HipA7 by identifying phosphorylation targets of the two kinase variants. The third aim was to establish a method to selectively label persister cells and study protein turnover during toxin-induced persistence and antitoxin-triggered resuscitation. The specific objectives were:

1. Identification of *in vivo* phosphorylation targets of HipA
 - a. Optimization of the phosphopeptide enrichment method for detection of HipA-specific phosphorylation events by confirming GltX phosphorylation *in vivo*
 - b. Identification of phosphorylation targets of overproduced HipA and examination of their dynamics during HipA-induced growth inhibition and HipB-mediated resuscitation
 - c. Validation of phosphorylation of selected HipA targets by *in vitro* kinase assay
 - d. Investigation of physiological effect of phosphorylation of a selected HipA substrate
2. Comparison of HipA and HipA7 (phospho)proteomes to distinguish between different phenotypes of two kinase variants
 - a. Identification of phosphorylation events induced by overproduction of HipA and HipA7 in varying amounts
 - b. Analysis of (phospho)proteomes of chromosomally encoded HipA and HipA7
3. Time-resolved analysis of newly synthesized proteins during toxin-induced persistence and antitoxin-mediated resuscitation
 - a. Establishment and optimization of the method for selective labeling of persister cells and resuscitating persister cells using dynamic SILAC pulse-labeling applied to the *hipBA* TA module
 - b. Temporal analysis of newly synthesized proteins during HipA-mediated persistence and upon resuscitation from persistence
 - c. Application of the method to another TA system using mRNAse RelE to find common mechanisms between persistence mediated by HipA and RelE toxins

3 Materials and Methods

3.1 Materials

3.1.1 List of chemicals

Name	Company
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Sigma-Aldrich
β -mercaptoethanol	BDH
[γ - ³² P]-ATP (100 mCi/ml)	PerkinElmer
Acetic acid (glacial)	Sigma-Aldrich
Acetonitrile (ACN)	Merck KGaA
Adenosine 5'-triphosphate (ATP) disodium salt	Sigma-Aldrich
Agar-agar	Roth
Agarose	Roth
Ammonium bicarbonate (ABC)	Merck KGaA
Ammonium chloride (NH ₄ Cl)	Sigma-Aldrich
Ammonia solution 25%	Merck KGaA
Ampicillin, sodium salt	Sigma-Aldrich
L-(+)-arabinose	Sigma-Aldrich
Boric acid	Sigma-Aldrich
Betaine solution (5M)	Sigma-Aldrich
Bovine serum albumin (BSA)	Sigma-Aldrich
Calcium chloride (CaCl ₂)	Merck KGaA
Chloroform	VWR
Chloramphenicol	Sigma-Aldrich
Ciprofloxacin	Sigma-Aldrich
cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail	Roche
Deoxynucleotide (dNTP) mix (10mM)	New England Biolabs
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
Dithiothreitol (DTT)	Merck KGaA
DNA Loading Dye (6x)	Thermo Scientific
DPBS	PAA Laboratories
Ethylendiaminetetraaceticacid (EDTA)	AppliChem
Ethanol	Merck KGaA
Ethanol, absolute for molecular biology	AppliChem
Ethidium bromide	Sigma-Aldrich
Formic acid (FA)	Merck KGaA
D-(+)-glucose	VWR
GeneRuler DNA Ladder Mix	Thermo Scientific

Materials and Methods

GeneRuler 100 bp DNA Ladder	Thermo Scientific
Glycerol	Sigma-Aldrich
Glycerol-2-phosphate	Sigma-Aldrich
Glycine	Sigma-Aldrich
Hydrochloric acid (HCl)	Merck
Imidazole	Sigma-Aldrich
InstantBlue	Expedeon
Iodoacetamide (IAA)	Sigma-Aldrich
Isopropyl- β -D-thiogalactopyranoside (IPTG)	VWR
Kanamycin sulfate	Sigma-Aldrich
LB medium	Roth
L-lysine (Lys0)	Sigma-Aldrich
L-lysine- ² H ₄ hydrochloride (Lys4)	Cambridge Isotope Laboratories
L-lysine- ¹³ C ₆ ¹⁵ N ₂ hydrochloride (Lys8)	Cambridge Isotope Laboratories
Magnesium chloride (MgCl ₂)	Merck KGaA
Magnesium sulfate (MgSO ₄)	Merck KGaA
Methanol	Merck KGaA
L-methionine	Sigma-Aldrich
Midori Green Advance DNA Stain	Nippon Genetics Europe
2-Propanol	Sigma-Aldrich
Poly(ethylene glycol) (PEG) MW=3,350	Sigma-Aldrich
Potassium chloride (KCl)	Merck KGaA
Potassium hydroxide (KOH)	Sigma-Aldrich
Potassium phosphate monobasic (KH ₂ PO ₄)	Sigma-Aldrich
Precision Plus Protein Standards	Bio-Rad Laboratories
SeeBlue Plus2 Prestained Standard	Life Technologies
Sodium dodecyl sulfate (SDS)	Serva
L-serine	Sigma-Aldrich
Sodium chloride (NaCl)	Merck KGaA
Sodium fluoride (NaF)	Sigma-Aldrich
Sodium hydroxide (NaOH)	Merck KGaA
Sodium phosphate dibasic (Na ₂ HPO ₄)	Sigma-Aldrich
Sodium orthovanadate (Na ₃ VO ₄)	Sigma-Aldrich
Sucrose	Sigma-Aldrich
Thiamine hydrochloride	Sigma-Aldrich
Thiourea	Merck KGaA
Trifluoroacetic acid (TFA)	Merck Millipore
Tris(hydroxymethyl)aminomethane (Tris, Trizma Base)	Sigma-Aldrich
Total tRNA from <i>E. coli</i>	Sigma-Aldrich
Urea	Merck KGaA
Yeast extract	Roth

Zeocin	InvivoGen
Zinc sulfate (ZnSO ₄)	Merck KGaA

3.1.2 List of materials

Name	Company
3M C8/C18 solid phase extraction disks	Empore™
Amicon Ultra-4 10 kDa	Merck Milipore
Amicon Ultra-4 30 kDa	Merck Milipore
Centrifuge Columns 0.8 ml	Pierce
Cryogenic vials 2 ml	Greiner Bio-One
HisPur cobalt resin	Thermo Fisher
NuPAGE 4-12% Bis-Tris Gels 1.0 mm (10 and 20 wells)	Life Technologies
ReproSil-Pur C18-AQ 1.9 μm resin	Dr. Maisch
ReproSil-Pur C18-AQ 3 μm resin	Dr. Maisch
Sep-Pak Vac 1cc C18 Cartridges	Waters
Titanium dioxide (TiO ₂) spheres, 300 Å, 5 μm	ZirChrom
Vacuum filter system 1 l, polyethersulfone membrane, 0.22 μm	Corning
Whatman Puradisc 25, polyethersulfone membrane, 0.2 μm	GE Healthcare

3.1.3 List of commercial kits and buffers

Name	Company
Antarctic Phosphatase Reaction Buffer (10x)	New England Biolabs
CutSmart Buffer (10x)	New England Biolabs
High pH Reversed-Phase Peptide Fractionation Kit	Pierce
InstantBlue Protein Stain	Expedeon
GeneJET PCR Purification Kit	Thermo Fisher Scientific
GenElute Bacterial Genomic DNA Kits	Sigma-Aldrich
<i>mi</i> -Gel Extraction Kit	Metabion
<i>mi</i> -PCR Purification Kit	Metabion
<i>mi</i> -Plasmid MiniPrep Kit	Metabion
Phusion GC Reaction Buffer (5x)	New England Biolabs
Protein Assay Dye Reagent Concentrate (5x)	Bio-Rad Laboratories
QIAquick Gel Extraction Kit	Qiagen
QIAprep Spin Miniprep Kit	Qiagen
T4 DNA Ligase Buffer (10x)	New England Biolabs
ThermoPol Reaction Buffer (10x)	New England Biolabs
NuPAGE® LDS sample buffer (4x)	Life Technologies
NuPAGE® MOPS SDS buffer	Life Technologies

3.1.4 List of enzymes

Phusion Hot Start II DNA Polymerase (U/μl)	Thermo Fisher Scientific
<i>Taq</i> DNA Polymerase (5 U/μl)	New England Biolabs
Alkaline Phosphatase Calf Intestinal (CIP) (10 U/μl)	New England Biolabs
Shrimp Alkaline Phosphatase (rSAP) (1 U/μl)	New England Biolabs
FastAP Thermosensitive Alkaline Phosphatase (1 U/μl)	Thermo Fisher Scientific
T4 DNA ligase (400 U/μl)	New England Biolabs
T4 DNA ligase (5 U/μl)	Thermo Fisher Scientific
BamHI-HF (20 U/μl)	New England Biolabs
EcoRI-HF® (20 U/μl)	New England Biolabs
HindIII-HF® (20 U/μl)	New England Biolabs
NotI-HF (20 U/μl)	New England Biolabs
NdeI (20 U/μl)	New England Biolabs
NcoI-HF (20 U/μl)	New England Biolabs
XhoI (20 U/μl)	New England Biolabs
XbaI (20 U/μl)	New England Biolabs
XbaI (10 U/μl)	Thermo Fisher Scientific
SphI-HF (20 U/μl)	New England Biolabs
SphI (PaeI, 10 U/μl)	Thermo Fisher Scientific
GoTaq Green Master Mix	Promega
Lysozyme	Sigma-Aldrich
DNase I	Sigma-Aldrich
Chymotrypsin	Promega
Lysyl endoproteinase (Lys-C)	Wako

3.1.5 List of instruments

Name	Company
Pure Water System arium® 611	Sartorius AG
T100 Thermayl Cycler	Bio-Rad Laboratories
NanoDrop 1000 Spectrophotometer	PeqLab Biotechnologie
Horizontal Electrophoresis System	Bio-Rad Laboratories
PowerPac Power Supply	Bio-Rad Laboratories
XCell Mini-Cell Electrophoresis System	Invitrogen
Electrophoresis power supply 800	Consort
Super-Bright Transilluminator	Vilber Lourmat
Incubator Natural Convec BD	Binder
Incubator B-6120	Heraeus
New Brunswick 44 Incubator Shaker	Innova
Vortex Genie Mixer	VWR

Thermomixer Comfort	Eppendorf
Centrifuge (table-top)	Eppendorf
Centrifuge Fresco 17	Heraeus
Centrifuge Micro Star 12	VWR
Stage-tip Centrifuge	Sonation
Vacuum Centrifuge SpeedVac	Eppendorf
Spinning Carousel	A. Hartenstein
Sonifier 250; Microtip 5	Branson Ultrasonics Corporation
UV-Vis Spectrometer Genesys 6	Thermo Scientific
UV-1600PC Spectrophotometer	VWR
Lyophilizer Alpha 2-4 LD plus	Christ
Technoflow 3F150-II Laminar Flow Hood	Integra Biosciences
Peqlab Phusion SL	Vilber Lourmat
-80 °C Igloo Freezer	Telstar
Chemistry Hybrid Pump RC 6	Vacuubrand
Borosilicate Emitters	Thermo Scientific
EASY-nLC II	Proxeon Biosystems
EASY-nLC 1000	Thermo Fisher Scientific
EASY-nLC 1200	Thermo Fisher Scientific
Column oven	Sonation
Ultimate 3000 UHPLC	Thermo Fisher Scientific
XBridge BEH130 C18 3.5µm 4.6 x 250 mm column	Waters
Electrospray Ionization Source	Proxeon Biosystems
Nano ES Ion Source (ES380)	Thermo Scientific
LTQ-Orbitrap Elite Mass Spectrometer	Thermo Fisher Scientific
LTQ-Orbitrap XL Mass Spectrometer	Thermo Fisher Scientific
Q Exactive HF Mass Spectrometer	Thermo Fisher Scientific
Autoclave	Sauter; Systec
Gel Dryer 583	Bio-Rad Laboratories
Typhoon FLA 7000 laser scanner	GE Healthcare

3.1.6 List of bacterial strains and plasmids

Strains/Plasmids	Genotype	Source
Strains		
MG1655	Wild-type <i>E. coli</i>	(257)
Δ hipBA (EG19)	MG1655 <i>hipBA::FRT</i>	(63)
<i>hipA7</i> (EG398)	MG1655 <i>hipA7 zde264::Tn10 dapA6</i>	This work*
Δ relA (EG109)	MG1655 <i>relA251::aphA</i>	P1 from CF1693 in MG1655 (258)

Materials and Methods

<i>rplK S102A</i>	MG1655 <i>rplK S102A</i>	This work
<i>rplK S102D</i>	MG1655 <i>rplK S102D</i>	This work
BL21(DE3) competent <i>E. coli</i>	<i>fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS</i> <i>λ DE3 = λ sBamHI ΔEcoRI-B</i>	New England Biolabs
NEB 5-alpha competent <i>E. coli</i> (DH5α)	<i>int::(lacI:PlacUV5::T7 gene1) i21 Δnin5</i> <i>fhuA2 (argF-lacZ)U169 phoA glnV44 80</i> <i>(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1</i> <i>hsdR17</i>	New England Biolabs
<i>ΔlysA</i>	MG1655 <i>ΔlysA::FRT</i>	This work*
<i>ΔrelBE</i>	MG1655 <i>ΔrelBE::FRT</i>	Christensen 2001
Plasmids		
pBAD33	p15, cat, araC, P _{BAD} promoter, Cm ^R	(259)
pBAD33:: <i>hipA</i> (pEG5)	pBAD33 P _{BAD} :: sd8 <i>gtg hipA</i>	(63)
pBAD33:: <i>hipA7</i> (pEG9)	pBAD33 P _{BAD} :: sd8 <i>gtg hipA7</i>	This work*
pNDM220	Mini-R1, <i>bla</i> , <i>lacI^q</i> , P _{A1/O4/O3} , Amp ^R	(93)
pNDM220:: <i>hipB</i> (pEG10)	pNDM220 P _{lac} :: <i>sdopt::hipB</i>	This work*
pEG220:: <i>hipB</i>	pEG220 P _{lac} :: <i>sdopt::hipB</i> , Km ^R	This work*
pMG25	pUC, <i>bla</i> , P _{lac} promoter, Amp ^R	Gerdes lab.*
pNDM220:: <i>hipA</i>	pNDM220 P _{lac} :: sd8 <i>atg hipA</i>	This work
pNDM220:: <i>hipA7</i>	pNDM220 P _{lac} :: sd8 <i>atg hipA7</i>	This work
pMG25:: <i>hipA</i>	pMG25 P _{lac} :: sd8 <i>atg hipA</i>	This work
pMG25:: <i>hipA7</i>	pMG25 P _{lac} :: sd8 <i>atg hipA7</i>	This work
pBAD33:: <i>hipA S150A</i> (pEG14)	pBAD33 P _{BAD} :: sd8 <i>atg hipA S150A</i>	This work
pBAD33:: <i>hipA S150D</i> (pEG13)	pBAD33 P _{BAD} :: sd8 <i>atg hipA S150D</i>	This work
pBAD33:: <i>hipA S359A</i> (pEG11)	pBAD33 P _{BAD} :: sd8 <i>atg hipA S359A</i>	This work
pBAD33:: <i>hipA S359D</i> (pEG12)	pBAD33 P _{BAD} :: sd8 <i>atg hipA S359D</i>	This work
pBAD33:: <i>6his hipA</i> (pEG15)	pBAD33 P _{BAD} :: sd8 <i>atg 6his hipA</i>	This work*
pBAD33:: <i>6his hipA7</i>	pBAD33 P _{BAD} :: sd8 <i>atg 6his hipA7</i>	This work
pET28a	T7lac promoter, His tag, T7 tag, Km ^R	New England Biolabs
pET28a:: <i>gltX</i>	pET28a:: <i>6his gltX</i>	This work
pET28a:: <i>rplK</i>	pET28a:: <i>6his rplK</i>	This work
pET28a:: <i>seqA</i>	pET28a:: <i>seqA 6his</i>	This work
pKOV	suicide vector, Psc101 ori, <i>sacB</i> gene, Cm ^R	Addgene
pBAD33:: <i>relE</i> (pKP3035)	pBAD33 P _{BAD} :: <i>relE</i>	(86)
pBR322:: <i>relB</i> (pSEM3152)	pBR322 P _{lac} :: <i>relB</i> , Zeo ^R	This work**
pBR322:: <i>relB</i> (pSEM3152)	pBR322 P _{lac} :: <i>relB</i> , weaker rbs, Zeo ^R	This work**

*unpublished constructs provided by Dr. Elsa Germain, University of Copenhagen, Denmark, for HipA and HipB-related experiments**unpublished constructs provided by Dr. Szabolcs Semsey, University of Copenhagen, Denmark, for RelE-related experiments

3.1.7 List of DNA oligonucleotides

Primers	Sequence 5' to 3'
Primers for cloning	
OMS43*	CCGCCg gatccGTCGACTAAAGGAAAAAAAAAATGCCTAACTTGCTCACTTG
OMS44*	CCGCCgaattcTCACTTACTACCGTATTC
OMS41*	CCGCCgaattcGTCGACTAAAGGAAAAAAAAAATGCCTAACTTGCTCACTTG
OMS42*	CGCCcgatccTCACTTACTACCGTATTC
OEG110*	AGAGAtctagaGTCGACTAAAGGAAAAAAAAAATGcatcaccatcaccatcacCCTAA ACTTGCTCACTTGGATGAAC
OEG111*	AGAGAGcatgcTCACTTACTACCGTATTCTCGGC
OEG57*	CCCCtctagaGTCGACTAAAGGAAAAAAAAAATGCCTAACTTGCTCACTTGGAT
OEG29*	TGACTTTCGCATC GCG GTTGCTGGCGCAC
OEG28*	GTGCGCCAGCAAC CGC GATGCGAAAGTCA
OEG241*	TGACTTTCGCATC GAT GTTGCTGGCGCAC
OEG242*	GTGCGCCAGCAAC ATC GATGCGAAAGTCA
OEG243*	GGGGCTTAACGCA GCG AAAGGCAAAAAAAAA
OEG244*	TTTTTTGCCTTT CGCT GCGTTAAGCCCC
OEG245*	GGGGCTTAACGCA GAT AAAGGCAAAAAAAAA
OEG246*	TTTTTTGCCTTT ATC TGCGTTAAGCCCC
OMS14	GCGCGCcatatgATGAAAATCAAACTCGC
OMS15	GAGAGActcgagTACTGCTGATTTTCGCG
OMS3	AGAGAcatatgATGGCTAAGAAAGTACAAGCC
OMS4	AGAGActcgagTtagTCTCCACTACCAGG
OMS26	GTCCGccatgg CGAT GAAAACGATTGAAG
OMS27	GCATGTtctcgagGATAGTTCCGCAAACCTT
OMS20	AGACTAgcggccgcATGTCTGAAGCTCCTAAA
OMS21	GAGAGAtctagaTtagTTTACAGAAGCGCTCAG
OMS10	AGTGGGTAAAATT GCG CGCGCTCAGCTGC
OMS11	GCAGCTGAGCGCG GCA AATTTTACCCACT
OMS12	AAGTGGGTAAAATT GAT CGCGCTCAGCTGCA
OMS13	TGCAGCTGAGCGCG ATCA AATTTTACCCACTT
Primers for sequencing	
OMS7 (T7 promoter)	TAATACGACTCACTATAGGG
OMS8 (T7 terminator)	GCTAGTTATTGCTCAGCGG
OMS18 (pBAD33 fwd)	CGCAACTCTCTACTGTTTCTC
OMS19 (pBAD33 rev)	CCGCTTCTGCGTTCTG
OMS47 (pNDM220 fwd)	TCTGATGAAGCGTCAGCACGACGTT
OMS48 (pNDM220 rev)	GTAAGGAGAAAATACCGCATCAGGC
OMS32 (pMG25 fwd)	GCGTATCACGAGGCCCTTTC
OMS33 (pMG25 rev)	TTTGCTACCCGTTGTAGCGC

Materials and Methods

OMS22 (pKOV fwd)	CCCGGTCGTCGGTTCAGGGC
OMS23 (pKOV rev)	TTAATGCGCCGCTACAGGGCG
OMS25	ACCGTTTACGCTGACCGTTC
OMS34	TCCGTGGCGGTCAGCGTCGCAA
OMS35	GTCTTTGTTTCGGCTTACCGGAACC
OMS45	CCGCAGTAATGTCACTGATCCTGTGG
OMS46	AACCCGAATATCTCGGATATACGTCTGC
OMS49	ATGGCTAAGAAAGTACAAGCCT
OMS50	TTCTTAGTCCTCCACTACCAGG
OMS51 (S102A)	AGACAAAGTGGGTAAAATTGCG
OMS52	AGACAAAGTGGGTAAAATTTCC

Primers for TSP genotyping

OMS36 (S102A)	CTAAGGGTAAAATT <u>GCG</u>
OMS37	ATGTGTA CTGTTTGTT
OMS38	TATATCACGGGGAGCCTCTCAGA
OMS39	TCGTTACGGTAACGAACCTGGC
OMS40 (S102D)	TCTCTGGGTAAAATT <u>GAT</u>

*oligonucleotides designed by Dr. Elsa Germain, University of Copenhagen, Denmark, for HipA and HipB-related experiments

3.2 Methods

*Parts of the chapters were taken from **Semanjski M**, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B, The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli proteome to promote multidrug tolerance, Science Signaling, 2018, 11(547) (260).*

3.2.1 Cloning strategy

*Plasmid constructs and E. coli strains marked with * and ** in the “List of bacterial strains and plasmids” (see 3.1.6) were provided by Dr. Elsa Germain (*) and Dr. Szabolcs Semsey (**) from the Centre for Bacterial Stress Response and Persistence, University of Copenhagen, Denmark. Others were constructed by me and are described below.*

E. coli strains and plasmids used in this study are listed in 3.1.6. Due to its high toxicity, *hipA* was cloned into expression plasmids together with a mitigate Shine-Dalgarno by changing either the consensus sequence of the SD, the spacer between the SD and the start codon, or the start codon according to (261). In 3.1.7, sd8 indicates a consensus sequence AAGGAA with a spacer of 8 nucleotides to the ATG start codon. Oligonucleotides used are listed in 3.1.7.

pNDM220::*hipA* and pNDM220::*hipA7*. The *hipA* gene was amplified from pEG5 and *hipA7* from pEG9 with primers OMS43/OMS44. The PCR products were digested with EcoRI and BamHI and ligated with pNDM220 digested with same enzymes. The resulting plasmids contain the *hipA* or *hipA7* gene with a mitigate SD (sd8ATG) sequence downstream of the P_{lac} promoter.

pMG25::*hipA* and pMG25::*hipA7*. The *hipA* gene was amplified from pEG5 and *hipA7* from pEG9 with primers OMS41/OMS42. The PCR products were digested with EcoRI and BamHI and ligated with pMG25 digested with same enzymes. The resulting plasmids contain the *hipA* or *hipA7* gene with a mitigate SD (sd8ATG) downstream of the P_{lac} promoter.

pBAD33::*6his hipA7*. The *hipA7* gene was amplified from pEG9 with primers OEG110 and OEG111. The PCR products were digested with XbaI and SphI and ligated with pBAD33 digested with same enzymes. The resulting plasmid contains the *hipA7* gene with a mitigate SD (sd8ATG) sequence downstream of the P_{BAD} promoter and the sequence that encodes 6 histidine residues at protein N-terminus.

pBAD33::*hipA S150A*, pBAD33::*hipA S150D*, pBAD33::*hipA S359A* and pBAD33::*hipA S359D*. Those mutants have been amplified using two step PCR technic which consist in PCR amplification of each fragment upstream and downstream the point mutation and then a third PCR mixing both fragments to obtain *hipA* variant with the external primers OEG57 and OEG111. The *hipA* gene was amplified from pEG5 with primers OEG57/OEG28 and OEG29/OEG111 to construct S150A mutant, with primers OEG57/OEG242 and OEG241/OEG111 to construct S150D mutant, with primers OEG57/OEG244 and OEG243/OEG111 to construct S359A mutant and primers OEG57/OEG246 and OEG245/OEG111 to construct S359D mutant. The final PCR products were digested with XbaI and SphI and ligated with

Materials and Methods

pBAD33 digested with same enzymes. The resulting plasmids contain the *hipA* mutant gene with a mitigate SD (sd8ATG) sequence downstream of the P_{BAD} promoter.

pET28a::*gltX* and pET28a::*rplK*. The *gltX* and *rplK* genes were amplified from MG1655 *E. coli* strain with primers OMS14/OMS15 and OMS3/OMS4, respectively. The PCR products were digested with NdeI and XhoI, and ligated with pET28a digested with the same enzymes. The resulting plasmids contain the *gltX* or *rplK* gene together with the sequence upstream of the *gltX* or *rplK* gene that at protein N-terminus encodes for MGSS, 6 histidine residues and SSGLVPRGSHM.

pET28a::*seqA*. The *seqA* gene was amplified from MG1655 with primers OMS26 and OMS27. The PCR product was digested with NcoI and XhoI, and ligated with pET28a digested with same enzymes. The resulting plasmid contains the *seqA* gene together with the sequence upstream of *seqA* gene that encodes for MA sequence at protein N-terminus and the sequence downstream of *seqA* gene that encodes for LE and 6 histidine residues at protein C-terminus.

The MG1655 *rplK S102A* and *rplK S102D* strains were constructed by replacing a chromosomal *rplK* gene that encodes for ribosomal protein L11 (RplK) with the *rplK* gene that encodes for phosphoablative mutant RplK-S102A or phospho-mimetic mutant RplK-S102D. To construct the mutants, two or three point mutations were introduced into the wild-type *rplK* gene (*gcg* codon for *rplK S102A* and *gat* for *rplK S102D* instead of *tcc* codon) at the position that encodes for Ser¹⁰². Those mutants have been amplified using two-step PCR technic as described above. The external primers used to obtain *rplK* variant were OMS20 and OMS21. *rplK S102A* has been constructed using OMS20/OMS11 and OMS10/OMS21 and for *rplK S102D* variant we used OMS20/OMS13 and OMS12/OMS21. The gene constructs containing *nusG*, *rplK S102A/rplK S102D* and *rplA* genes were cloned into pKOV plasmid using NotI and XbaI restriction enzymes with primers OMS20/OMS21 and allelic replacement was performed as previously described (262). Plasmid pKOV was a gift from George Church (Addgene plasmid #25769). Colonies were screened for mutations using temperature switch PCR genotyping as previously described (263) with primers (OMS36-OMS40) and confirmed by DNA sequencing using chromosomal primers located outside of the cloning region of interest (OMS45 and OMS46) and inside (OMS25, OMS34, OMS35, OMS45, OMS46, OMS49-OMS52).

3.2.2 Plasmid cloning

Genomic DNA was purified from 1.5 ml of *E. coli* K-12 MG1655 wild-type or *hipA7* mutant culture grown overnight in LB medium using GenElute Bacterial Genomic DNA Kit according to the manufacturer instructions. Plasmid DNA was purified from 5 - 10 ml of culture grown overnight in LB or M9 medium supplemented with 0.4% (w/v) glucose (for HipA constructs) using QIAprep Spin Miniprep or *mi*-Plasmid MiniPrep Kit according to the manufacturer instructions. DNA concentration was determined from absorbance measured by Nanodrop spectrophotometer at a wavelength of 260 nm.

Materials and Methods

Genes of interest were amplified either from *E. coli* genomic DNA or from plasmid constructs using Phusion high-fidelity (HF) DNA polymerase. Each polymerase chain reaction (PCR) mixture contained: 1x Phusion GC buffer, 200 μ M dNTPs, 0.44 μ M of each primer, 0.02 U/ μ l Phusion HF DNA polymerase, 2 ng of template plasmid DNA or 200 ng of genomic DNA and nuclease free water. PCR was performed with a standard PCR program for Phusion HF DNA polymerase (**Table 1**). For each PCR independent of the type of the polymerase, the annealing temperature (T_a) was set to 7 °C below the melting temperature (T_m) of the primer with lower T_m . The elongation duration depends on the length of the DNA template.

Table 1. PCR program for Phusion HF DNA polymerase.

PCR step	Temperature	Duration	
Initial denaturation	98 °C	30 s	
Denaturation	98 °C	10 s	
Annealing	T_a	30 s	35 cycles
Extension	72 °C	30 s per kb	
Final extension	72 °C	10 min	
Cooling	4 °C	∞	

PCR products were purified using *mi*-PCR Purification or GeneJET PCR Purification Kit according to the manufacturer procedure and eluted with nuclease free water. The size and the purity of DNA samples determined by agarose gel electrophoresis. Around 1 - 3 μ l of PCR mixture was mixed with 6x loading dye and separated in 1% (w/v) agarose gel containing 7 μ l/100 μ l of Midori Green or 1.2 mg/ml ethidium bromide and 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.2) for 20-40 min at 90 - 110 V. The bands were visualized with UV light.

Site directed mutagenesis was performed with two-step PCR technique (see above). In a second PCR step, two purified PCR products were used as templates in amount of 0.1 pmol each. Final PCR product was purified from the agarose gel using QIAquick Gel Extraction or *mi*-Gel Extraction Kit according to the manufacturer instructions and eluted with nuclease free water.

Purified PCR product (1 - 2 μ g) and plasmid DNA (2 - 4 μ g) were digested with restriction endonucleases (10 U/ μ g of DNA) in 1x CutSmart buffer for 2.5 h at 37 °C. Digested plasmid was then dephosphorylated by adding Alkaline Phosphatase (10 U/ μ g of DNA) to the digestion mixture for 45 min at 37 °C. Restriction enzymes and phosphatase were inactivated for 20 min at 65 °C. Digested PCR products were purified using GeneJET PCR Purification or *mi*-PCR Purification Kit.

Digested DNA and 50 - 200 ng of plasmid were mixed in a molar ratio plasmid:insert of 1:3 in a maximal reaction volume of 8 - 12 μ l, and ligated with 5 U of T4 DNA ligase in 1x T4 ligation buffer for 1h at RT. Ligation mixture was then transformed into chemically competent *E. coli* DH5 α cells. For that, 4 μ l of ligation mixture was added to 50 μ l (around 5×10^9 cells) of thawed chemically competent cells and incubated on ice for 30 min, followed by a 45 s heat shock at 42 °C. Afterwards, the mixture was cooled on ice for 5 min, supplemented with 950 μ l of LB medium and recovered for 1 h by shaking at 37 °C.

Materials and Methods

Transformed cells were plated on LB agar plates containing appropriate antibiotics and incubated overnight at 37 °C. Colonies were screened for the presence of the desired insert by colony PCR. Briefly, colonies were streaked out on LB plate containing antibiotics and the tip was afterwards washed in 12.5 µl of mixture containing 6.25 µl of GoTaq Green Master Mix and 0.4 µM primers. PCR was performed with a standard PCR program for GoTaq polymerase (**Table 2**). The entire PCR mixture was separated in 1% (w/v) agarose gel.

Table 2. PCR program for colony PCR performed with GoTaq polymerase.

PCR step	Temperature	Duration	
Initial denaturation	95 °C	3 min	
Denaturation	95 °C	30 s	
Annealing	T _a	30 s	30 cycles
Extension	72 °C	1 min per kb	
Final extension	72 °C	5 min	
Cooling	4 °C	∞	

DNA sequence analysis was performed by GATC Biotech AG. Samples were prepared according to the company instructions. Primers for sequencing were designed to align at least 50 bp outside the construct of interest. The received sequences were analyzed by the SnapGene or the ApE software.

Plasmid DNA was isolated and transformed into a desired *E. coli* strain by TSB transformation. Briefly, *E. coli* culture was grown to OD_{600nm} of 0.4 - 0.6, cell pellet from 1 ml culture aliquot was re-suspended in TSB medium (10% (w/v) PEG MW=3,350, 5% (v/v) DMSO, 20 mM MgCl₂, LB medium) and incubated with 3 µl of purified plasmid on ice for 30 min. Afterwards, cells were subjected to a heat shock of 42° C for 45 s, cooled on ice for 5 min, supplemented with 900 µl of LB medium and recovered for 1 h by shaking at 37 °C. Transformed cells were plated on LB agar containing appropriate antibiotics and incubated overnight at 37 °C. Finally, *E. coli* strains from overnight cultures grown in LB were permanently stored as 1.5 ml stocks in 40% (v/v) glycerol in cryogenic vials at -80 °C.

3.2.3 Gene replacement

Gene replacement of *rplK* mutants was performed with plasmid pKOV, which was a gift from George Church (Addgene plasmid #25769). The gene constructs containing *nusG*, *rplK S102A* or *rplK S102D* and *rplA* genes were cloned into pKOV plasmid using NotI and XbaI restriction enzymes with primers OMS20/OMS21 and allelic replacement was performed as previously described (262). Briefly, *E. coli* cells transformed with pKOV construct and grown at 30 °C were re-suspended in 10% (w/v) MgSO₄, serially diluted and plated on LB agar plates containing 20 µg/ml chloramphenicol. First, plates were incubated overnight at 43 °C (non-permissive temperature) to select for cells that incorporated pKOV plasmid into the chromosome based on the temperature sensitive origin of plasmid replication. Second, six colonies were re-suspended in LB medium without NaCl (1% (w/v) tryptone, 0.5% (w/v) yeast extract), serially diluted, plated on LB agar-sucrose plates without NaCl (1% (w/v) tryptone, 0.5%

Materials and Methods

(w/v) yeast extract, 1.5% (w/v) agar, 5% (w/v) sucrose) and incubated at 30 °C for 1 - 2 days. Sucrose enables positive selection of cells that underwent second round of homologous recombination and lost the pKOV plasmid, based on the *sacB* gene expression of which is lethal in the presence of sucrose. Colonies were screened for mutations using temperature switch PCR (TSP) genotyping as previously described (263) and confirmed by DNA sequencing. Briefly, two primer pairs with different T_m were used. The locus specific primers (outer primers) with higher T_m were used to amplify 800 - 1,000 bp target locus, while nested locus specific primers (inner primers) with lower T_m were used to amplify 100 – 200 bp allele specific product based on the specific forward primer which contained mutant allele at its 3' end. The PCR mixture for TSP genotyping contained: 1x ThermoPol buffer, 200 μ M dNTPs, 0.1 μ M primers with higher T_m , 0.5 μ M primers with lower T_m , 0.54 M betaine, 1.3 mM DTT, 11 μ g/ml BSA, 1.64% (v/v) DMSO, 0.05 U/ μ l Taq DNA polymerase and nuclease free water. Colonies were re-suspended in 20 μ l nuclease free water and 0.5 μ l was added to 14.5 μ l of PCR mixture. PCR was performed with the TSP genotyping PCR program (Table 3).

Table 3. PCR program for TSP genotyping with Taq DNA polymerase.

PCR step	Temperature	Duration	
Initial denaturation	95 °C	3 min	
Denaturation	95 °C	30 s	
Annealing (outer primers)	58 °C	30 s	15 cycles
Extension	72 °C	1 min	
Denaturation	95 °C	10 s	5 cycles
Annealing (inner primers)	45 °C	30 s	
Denaturation	95 °C	10 s	
Annealing (inner primers + 5' tail)	53 °C	30 s	15 cycles
Product extension	72 °C	5 s	
Cooling	4 °C	∞	

PCR product were separated in 1.5% (w/v) agarose gel containing 7 μ l/100 μ l Midori Green in 1x TBE buffer (90 mM Tris-borate, 2 mM EDTA, pH 8.2) for 30 min at 110 V. GeneRuler 100 bp was used as DNA ladder from 100 - 1,000 bp. Genomic DNA was isolated from *E. coli* strain with positive DNA band at 100 - 200 bp using GenElute Bacterial Genomic DNA Kit according to the manufacturer instructions. DNA sequence analysis was performed by GATC Biotech AG. Mutant clones and double crossover wild-type were stored at -80 °C as glycerol stocks.

3.2.4 Media and antibiotics

Cells were grown in LB (Roth) or in M9 minimal medium (50 mM Na_2HPO_4 , 22 mM KH_2PO_4 , 8.6 mM NaCl, 18.7 mM NH_4Cl , 1 mM MgSO_4 , 0.1 mM CaCl_2 , 0.0001% (w/v) thiamine) supplemented with either 0.5% (w/v) glucose or 0.4% (v/v) glycerol when using strains carrying pBAD33 plasmid constructs. Cultures were grown in batch at 37 °C shaking at 200 rpm. When required, the medium

Materials and Methods

was supplemented with 25 µg/ml chloramphenicol, 25 - 50 µg/ml ampicillin or 50 µg/ml kanamycin. Plasmids carrying P_{BAD} promoter were repressed in pre-cultures by 0.4% (w/v) D-(+)-glucose at OD_{600nm}. The expression of gene constructs on plasmids carrying P_{BAD} promoter was induced with 0.2% (w/v) L-(+)-arabinose at OD_{600nm} of around 0.4. The expression of gene constructs on plasmids carrying P_{lac} promoter was induced with 0.5 - 2 mM β-D-1-thiogalactopyranoside (IPTG); 2 mM IPTG for pNDM220::*hipB* plasmid and 1 mM IPTG for pNDM220 and pMG25 plasmid constructs with *hipA* and *hipA7*.

3.2.5 SILAC labeling

For quantitative (phospho)proteomic experiments (**Table 4**), *E. coli* cells were differentially labeled using following stable isotope-labeled lysine derivatives: 4,4,5,5-D₄ L-lysine (Lys4, “medium-heavy” lysine, K4, Cambridge Isotope Laboratories), ¹³C₆¹⁵N₂ L-lysine (Lys8, “heavy” lysine, K8, Cambridge Isotope Laboratories) or L-lysine (Lys0, “light” lysine, K0, Sigma Aldrich) (178). Both pre-cultures and main cultures were grown in M9 minimal medium containing 0.0025% (w/v) of either Lys0, Lys4 or Lys8.

Table 4. Overview of the SILAC-based phosphoproteomic experiments performed in this study. L = light, M = medium, H = heavy lysine.

Experiment	No. of replicates	SILAC Labeling
HipA and HipB temporal analysis (plasmid)	2	0h (L) + HipA 75min (M) + HipA 3h (H) 0h (L) + HipB 2.5h (M) + HipB 6h (H)
HipA7 vs. HipA (plasmid)	3	empty plasmid (L) + HipA (M) + HipA7 (H) empty plasmids (L) + HipA-low copy (M) + HipA-high copy (H)
HipA7 and HipA, high vs. low copy number plasmid	2	empty plasmids (L) + HipA7-low copy (M) + HipA7-high copy (H)
<i>hipA7</i> vs. <i>wt hipA</i> (chromosomal)	3	<i>wt hipA</i> (L) + Δ <i>hipBA</i> (M) + <i>hipA7</i> (H)
HipA S150 and S359 mutants	1	HipA (L) + HipA S150A (M) + HipA S150D (H) HipA (L) + HipA S359A (M) + HipA S359D (H)

3.2.6 Dynamic SILAC for protein turnover studies

Cells were grown in 100 mL of M9 minimal medium (see 3.2.4) supplemented with 0.4% (v/v) glycerol, 25 µg/ml chloramphenicol for maintenance of pBAD33::*hipA* or pBAD33::*reIE* plasmids and 50 µg/ml kanamycin (for maintenance of pEG220::*hipB* plasmid) or 70 µg/ml zeocin (for maintenance of pBR322::*reIB* plasmid). For resuscitation experiments, strains lacking *lysA* gene (Δ *lysA*) were used and pre-cultures and main cultures were supplemented with 0.025% (w/v) “light” lysine (Lys0) (**Table 5**). All cultures were grown in batch at 37 °C shaking at 200 rpm. Pre-cultures were grown for around 24 hours in a medium supplemented with 0.4% (w/v) D-(+)-glucose to repress the expression of toxins

Materials and Methods

from pBAD33 plasmid. Main cultures were grown starting from OD_{600nm} of around 0.01 – 0.02. Expression of *hipA* and *relE* was induced for 3 hours at OD_{600nm} of around 0.4 with 0.2% and 0.4% (w/v) L-(+)-arabinose, respectively. Cultures were treated with 100 µg/ml ampicillin for 16-20 hours to kill non-persister cells and 0.4% glucose to repress the expression from pBAD33 plasmid. For experiments during persistence phase, a pulse of 0.025% (w/v) “heavy” lysine (Lys8) was added in the presence of ampicillin and 1 – 2 mL of cultures were harvested at specific time points. A time point before Lys8 pulse (0h) was harvested as a control. For experiments during resuscitation phase, cultures were first quickly filtered using pre-warmed 1 L Corning filter system to remove antibiotic, Lys0 and dead cells. Immediately after, cells were quickly resuspended in pre-warmed medium supplemented with appropriate antibiotics, 0.4% glycerol, 0.4% glucose, Lys8 and 2 mM IPTG for expression of *hipB*. Cells were harvested in specific time intervals and a time point before filtering and Lys8 pulse (0h) was harvested as a control.

Table 5. Overview of the dynamic SILAC-based experiments and *E. coli* strains used.

Experiment	No. of replicates	<i>E. coli</i> K-12 strain
HipA-induced persistence	3	MG1655, pBAD33:: <i>hipA</i> , pEG220:: <i>hipB</i>
RelE-induced persistence	3	MG1655 $\Delta relBE$; pBAD33:: <i>relE</i> , pBR322:: <i>relB</i>
HipB-induced resuscitation	3	MG1655 $\Delta lysA$, pBAD33:: <i>hipA</i> , pEG220:: <i>hipB</i>

3.2.7 Cell lysis and protein extraction

Cultures were harvested at specific stages by centrifugation at 4 °C and stored at -80 °C. The cell pellets were re-suspended in a lysis buffer (40 mg/ml SDS, 100 mM Tris-HCl pH 8.6, 10 mM EDTA, 5 mM glycerol-2-phosphate, 5 mM sodium fluoride, 1 mM sodium orthovanadate and Complete protease inhibitors (Roche)) and sonicated at least 5 times for 30 s at 40% amplitude. The cellular debris was pelleted by centrifugation at 13,000 g for 30 min and the crude protein extract precipitated from the supernatant with methanol and chloroform. Protein pellet was re-suspended in a denaturation buffer containing 6 M urea, 2 M thiourea and 10 mM Tris pH 8.0. Protein concentration was measured using standard Bradford assay (Bio-Rad).

3.2.8 Protein in-solution digestion

In each SILAC experiment, differently labeled protein extracts were mixed in equal amounts corrected by the ratios determined by measuring mixing checks (see below) to a total of 12 mg. Proteins were reduced using 1 mM dithiothreitol for 1 h and subsequently alkylated with 5.5 mM iodoacetamide for 1 h. One half of the protein mixture was diluted with 4 volumes of 62.5 mM Tris pH 8.0 and 12.5 mM CaCl₂ and digested with chymotrypsin (1:120 w/w) overnight at room temperature (RT). The other half was predigested with endoproteinase Lys-C (1:100 w/w) for 3 h, then diluted with 4 volumes of 62.5

mM Tris pH 8.0 and supplemented with endoproteinase Lys-C (1:100 w/w) for overnight digestion at RT. The reaction was stopped by acidification with trifluoroacetic acid (TFA) to pH 2. An aliquot of 10 μ g was purified by stage tips (see 3.2.14) and 2 μ g was used for direct proteome measurement with 230 min LC gradient. An additional aliquot of at least 100 μ g intended for further proteome measurements was stored at -80 °C. For protein turnover experiments, 10 μ g of protein extract from each time point collected was separately digested only with endoproteinase Lys-C as described above and purified by stage tips (see 3.2.14).

3.2.9 Phosphopeptide enrichment

Digested peptides were desalted by the solid-phase extraction using Sep-Pak Vac 100 mg C18 column. Briefly, column was activated with methanol and equilibrated with solvent A* (2% (v/v) acetonitrile and 1% (v/v) formic acid). After loading the sample, the column was washed with solvent A (0.1% (v/v) formic acid) and peptides were eluted with 1.8 ml 80% (v/v) acetonitrile and 6% (v/v) TFA. Phosphopeptides were enriched by titanium dioxide (TiO₂) chromatography. Eluted peptides were incubated with TiO₂ spheres (5 μ m, 300 Å, ZirChrom) in 1:10 peptide to bead ratio for 10 min for 5-10 consecutive rounds. TiO₂ spheres were washed twice with 80% (v/v) acetonitrile and 6% (v/v) TFA and loaded onto C8 stage tips. The spheres were washed additionally with 80% (v/v) acetonitrile and 1% (v/v) TFA. Phosphopeptides were first eluted with 50 μ l 1.25% (v/v) ammonium hydroxide of pH 10.5 into 20 μ l 20% (v/v) TFA for 15 min at 1,200 rpm. In the second elution step, phosphopeptides were eluted with 50 μ l 5% (v/v) ammonium hydroxide in 60% (v/v) acetonitrile pH 10.5. Acetonitrile was evaporated from eluates by vacuum centrifugation, samples were acidified to pH 2, if necessary, and purified by stage tips (see 3.2.14).

3.2.10 High pH reversed-phase peptide fractionation on commercial spin columns

Peptides derived from proteome samples of one replicate of each experiment were additionally separated offline using High pH reversed-phase peptide fractionation kit (Thermo Scientific, Cat. No. 84868) according to the manufacturer instructions. Briefly, 50 μ g of SILAC mixture was loaded onto the spin column and peptides were eluted in the gradient of acetonitrile in 9 fractions. The pH of 10 was maintained constant with 10 mM ammonium hydroxide instead of trimethylamine. Eluted fractions were concentrated by vacuum centrifugation and purified by stage tips (see 3.2.14). Fractions were measured separately by LC-MS/MS using optimized LC gradients.

3.2.11 Offline high pH reversed-phase peptide fractionation

Proteome samples digested with chymotrypsin of one replicate in the experiment with chromosomal *hipA7* were fractionated using an offline peptide fractionation at high pH to identify unmodified GlTX peptide. Peptides were loaded onto a reversed phase XBridge BEH130 C18 3.5 μ m 4.6 x 250 mm column installed in an Ultimate 3000 HPLC and detected by UV at λ = 214 nm at 25 °C. The system was

operated under basic conditions using buffer A (5 mM NH₄OH) and buffer B (5 mM NH₄OH in 90% (v/v) ACN) at pH 10. Peptides were eluted using an 80 min gradient at a flow rate of 1 ml/min. The organic portion was ramped from 5% to 25% B in 45 min, to 40% in 10 min and finally to 70% in 5 min followed by column equilibrated. Fractions were collected in 1 min intervals for 60 min and concatenated evenly into 30 pools. Acetonitrile was evaporated by vacuum centrifugation, samples acidified to pH 2 and measured by LC-MS/MS.

3.2.12 SDS-PAGE and in-gel digestion

Protein samples of one replicate of the experiment in which HipA and HipB were induced in temporal manner, were separate via SDS-PAGE and digested. Around 50 µg of protein extracts were separated on a NuPage Bis-Tris 4-12% gradient gel (Invitrogen) and stained with Coomassie stain. Gel lanes containing proteins were cut into 10 equal slices that were washed with 5 mM ammonium bicarbonate (ABC) in 50% (v/v) acetonitrile, reduced with 10 mM dithiothreitol in 20 mM ABC at 56 °C for 1 hour and alkylated with 55 mM iodoacetamide for 30 min in the same buffer. Gel pieces were then washed with 5 mM ABC and dehydrated with acetonitrile and short vacuum centrifugation. Proteins were digested either with endoproteinase Lys-C (1:40 w/w in 20 mM ABC) or chymotrypsin (1:40 in 50 mM Tris pH 8.0 and 10 mM CaCl₂) at 37 °C overnight. Digested peptides were recovered from the gel through a stepwise extraction with 3% (v/v) TFA in 30% (v/v) acetonitrile, 0.5% (v/v) acetic acid in 80% (v/v) acetonitrile and finally with acetonitrile. All supernatants were combined, concentrated by vacuum centrifugation and purified by stage tips (see 3.2.14).

3.2.13 Incorporation and mixing check

The efficiency of SILAC labeling was determined by LC-MS/MS measurement of Lys4- and Lys8-labeled samples. For that, 10 µg of each sample was separately digested with endoproteinase Lys-C, purified by stage tips (see 3.2.14) and measured by LC-MS/MS. In all cases, the labeling efficiencies of Lys4 or Lys8 were ≥ 94%. Prior to the mixing of labeled samples for SILAC experiments, 20 µg of each differentially labeled sample was pre-mixed in equal protein amounts determined by Bradford assay, digested with endoproteinase Lys-C and measured by LC-MS/MS. Median of evidence SILAC ratios was used as a correction factor for mixing the samples to be used in main SILAC experiments.

3.2.14 Stage tips

Prior to each LC-MS/MS measurement, all peptide samples were desalted and purified by C18 stage tips (264). Reversed phase C18 discs (Empore) were activated with methanol and equilibrated with solvent A*. Up to 10 µg of peptides were loaded onto the membrane and washed with solvent A. Peptides were eluted with 50 µL solvent B (80% (v/v) acetonitrile and 0.1% (v/v) formic acid) and concentrated by vacuum centrifugation. The sample volume was adjusted with solvent A and final 10% v/v of solvent A*.

3.2.15 LC-MS/MS measurement

Purified peptide samples were separated by an EASY-nLC 1000 or 1200 system (Thermo Scientific) coupled on-line to a Q Exactive HF mass spectrometer (Thermo Scientific) through a nanoelectrospray ion source (Thermo Scientific). Chromatographic separation was performed on a 20 cm long, 75 μm inner diameter analytical column packed in-house with reversed-phase ReproSil-Pur C18-AQ 1.9 μm particles (Dr. Maisch). The column temperature was maintained at 40 °C using an integrated column oven. Peptides were loaded onto the column at a flow rate of 700 nl/min or 1 $\mu\text{l}/\text{min}$ under maximum back-pressure of 500 or 850 bar, respectively. The peptides were eluted using either 46 min, 76 min, 116 min or 216 min segmented gradient of 10 – 50% solvent B at a constant flow rate of 200 nl/min. When measuring proteome digested with chymotrypsin, the gradient started with 5% of solvent B. For measurements of kinase assays, the peptides were eluted using 33 min segmented gradient of 10 – 50% solvent B at a constant flow rate of 300 nl/min. Samples fractionated by high pH chromatography on column were measured using different 76 min segmented gradients optimized for each fraction were used. For measurement of protein turnover, 116 min gradient were used for each proteome sample starting from 0h to the last time point.

Peptides were ionized by nanoelectrospray ionization at 2.3 kV and the capillary temperature of 275 °C. The mass spectrometer was operated in a data-dependent mode, switching automatically between one full-scan and subsequent MS/MS scans of either 12 (Top12 method) or 7 (Top7 method, for measurement of phosphorylation) most abundant peaks selected with an isolation window of 1.4 m/z. Full scan MS spectra were acquired in a mass range from 300 – 1,650 m/z at a target value of 3×10^6 charges with the maximum injection time of 25 ms and a resolution of 60,000 (defined at m/z 200). The higher energy collisional dissociation (HCD) MS/MS spectra were recorded with the maximum injection time of 45 ms or 220 ms (for measurement of phosphorylation) at a target value of 1×10^5 and a resolution of 30,000 (defined at m/z 200) or 60,000 for phosphoproteome measurement. The normalized collision energy was set to 27% and the intensity threshold was kept at 1×10^5 or 5×10^4 for phosphoproteome measurement. The masses of sequenced precursor ions were dynamically excluded from MS/MS fragmentation for 30 s. Ions with single, unassigned or six and higher charge states were excluded from fragmentation selection. Two phosphoproteome experiments (chromosomal *hipA7*, replicate 1 and 3) were measured on Orbitrap Elite mass spectrometer (Thermo Scientific) with following parameters that differ from Q Exactive HF system: Top15 method, 4 m/z isolation window, 300-2,000 m/z mass range, 1×10^6 full scan target value with 100 ms maximum injection time and 120,000 resolution (defined at m/z 400), 4×10^4 MS/MS scan target value with 150 ms injection time, 15,000 resolution and 35% normalized collision energy, 60s dynamic exclusion.

3.2.16 MS data processing and analysis of phosphoproteomic experiments

Acquired raw data were processed using the MaxQuant software suite (version 1.5.2.8) (169). Raw files of particular experiments were processed separately. In total, 449 raw files were processed, out of which 249 belong to phosphopeptide enrichment fractions. The derived peak list was searched using Andromeda search engine integrated in MaxQuant (171) against a reference *E. coli* K12

Materials and Methods

proteome (taxonomy ID 83333) obtained from UniProt (4,313 protein entries, released in October 2015), protein sequence of HipA7, HipA mutants and a file containing 245 common laboratory contaminants. During the first search, peptide mass tolerance was set to 20 ppm and in the main search to 4.5 ppm. For triple-label SILAC experiments, multiplicity was set to three with Lys4 and Lys8 specified as medium and heavy labels, respectively. Methionine oxidation, protein N-terminal acetylation and Ser/Thr/Tyr phosphorylation (STY) were set as a variable modification, whereas carbamidomethylation of cysteines was set as a fixed modification. The minimum required peptide length was set to seven amino acids with the maximum of two missed cleavages allowed for endoproteinase Lys-C that was set to specifically cleave at lysine C-terminus. Chymotrypsin was set to specifically cleave at phenylalanine, tryptophan, tyrosine, leucine and methionine C-terminus with maximum 5 missed cleavages allowing for maximum of 4 labeled amino acids. All peptide, protein and phosphopeptide identifications were filtered using a target-decoy approach with a false discovery rate (FDR) set to 0.01 (265). Proteins identified by the same set of peptides were combined to a single protein group. For protein quantification, a minimum of two peptide (razor or unique) ratio counts was required, whereas only unique peptides were used for the separate quantification of HipA and HipA7 proteins. To increase the number of quantified features, the “match between runs” option was enabled with a match time window set to 0.7 min. This allows the transfer of peptide identifications across LC-MS/MS runs based on the mass and the retention time of the peptide identified by MS/MS. Re-quantify option was enabled to allow for quantification of SILAC pairs that result in extreme ratio values. All protein groups and phosphorylation site identified in this study across all phosphoproteomic experiments are reported in the Supplementary Tables.

Statistical analysis of MaxQuant output data was performed manually or by using Perseus software (version 1.5.6.0) (266) and figures were edited in Adobe Illustrator. All contaminants and reverse hits were removed. Phosphorylation sites were additionally filtered for PEP scores of < 0.01 . Minimal score of 40 was required for phosphorylation site and 20 for protein identifications. Changes in phosphorylation events were corrected for the mixing error of differentially SILAC-labeled samples and normalized to differences in protein abundances, unless otherwise stated. For that, phosphorylation site SILAC ratios were divided with the protein SILAC ratios of corresponding proteins. Normalized phosphorylation site ratios were \log_2 transformed and plotted against the \log_{10} transformed phosphopeptide intensities summed for each of two SILAC channels observed. Significantly regulated phosphorylation sites were determined by applying an arbitrary ratio threshold of 2 in \log_2 scale (4-fold). In the experiment with *hipA7* on the chromosome in which no plasmids were used, significantly regulated phosphorylation sites were determined by using significance B test with a p-value of 0.01. Significantly regulated proteins were determined by using significance B test with a p-value of 0.001. For Volcano plots, \log_2 transformed ratios of three biological replicates were grouped into one group and compared to the group containing only zero values using t-test with FDR of 0.01 or 0.001 and the minimal fold change S_0 of 1. Phosphorylation site occupancies were determined as the proportion between the phosphorylated peptide and corresponding unmodified peptide using the algorithm implemented in MaxQuant based on the calculation described in Olsen *et al.* (267). The calculation of occupancies requires SILAC ratio of a phosphorylated peptide, the SILAC ratio of the corresponding unmodified peptide and the SILAC protein ratio. For RplK and SeqA,

occupancy values were calculated manually using M/L and H/M ratios giving a and b values between 0 and 1 in three biological replicates. For SeqA in “light” and RplK in “heavy” labeling state, occupancy was determined only from two replicates.

To identify significantly represented temporal protein profiles, we used Short Time-series Expression Miner (STEM) program (p-value of 0.05 after Bonferroni multiple testing correction) (268). Gene-annotation and KEGG enrichment analysis was performed using The Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (version 6.7) with default parameters (269). UniProt IDs were used as an input for the enrichment. Kinase motif analysis was performed using Motif-X software (270) with the reference *E. coli* K-12 MG1655 proteome used as a background and 15 amino acid long sequences (7 amino acids on both sides around phosphorylation site) of all identified phosphorylation sites with the localization probability higher than 0.75 as an input. The parameters of Motif-X analysis were as follows: S or T as a foreground and background central residue, width of 15, 40 occurrences and significance threshold of 0.00000001.

3.2.17 MS data processing and analysis of protein turnover experiments

Acquired raw data (171 raw files) were processed using the MaxQuant and raw files of particular experiments were processed separately, namely 63 raw files for HipB, 54 for HipA and 54 for RelE turnover experiment. The derived peak list was searched using Andromeda search engine against the reference *E. coli* K12 proteome (taxonomy ID 83333) obtained from UniProt (4,313 protein entries, released in October 2015) and a file containing 245 common laboratory contaminants. During the first search, peptide mass tolerance was set to 20 ppm and in the main search to 4.5 ppm. The multiplicity was set to two with Lys8 specified as the heavy label. Methionine oxidation and protein N-terminal acetylation were set as variable modifications, whereas carbamidomethylation of cysteines was set as a fixed modification. The minimum required peptide length was set to seven amino acids with the maximum of two missed cleavages allowed for endoproteinase Lys-C that was set to specifically cleave at lysine C-terminus. All peptide and protein identifications were filtered using a target-decoy approach with a false discovery rate (FDR) of 0.01. Proteins identified by the same set of peptides were combined to a single protein group. For protein quantification, a minimum of two peptide (razor or unique) ratio counts was required. The “match between runs” and re-quantify options were not enabled. Data was additionally processed with IBAQ quantification option enabled to determine protein IBAQ intensities for rough estimation of protein abundances across different experiments. IBAQ quantification normalizes total protein intensity to the number of theoretical tryptic peptides of the respective protein. To calculate recycling factor that was used for the correction of protein turnover rates, data was processed separately with multiplicity of 1 and lysine 8 (Lys8) set as variable modification.

Non-normalized protein H/L ratios from proteingroup.txt were used for protein quantification. All contaminants, reversed hits and proteins identified only by modification were removed. To simplify the data, the data sets from three biological replicate measurements were combined as a union, in which protein H/L ratios in each of the time points were kept if measured only in one replicate (Class

l), a mean was calculated if measured in two (Class II) or in three replicates (Class III). Second, to reduce the number of time points, they were pooled into time. For that, the median protein H/L ratio of respective time points was calculated and assigned to a specific time bin. Proteins were then ranked based on their H/L ratio within each time bin and displayed in a heatmap-like representation created in Excel. Gene-annotation and KEGG enrichment analysis was performed using DAVID tool (version 6.8) with default parameters and the background of all quantified proteins in particular experiment (269). All proteins that were reported in this study as newly synthesized are listed in the Supplementary Tables.

Protein turnover rate (k) of each protein was calculated by linear regression of the linear dependence of natural logarithm of protein SILAC H/L ratio over time as described previously (196) using following equation:

$$k = \frac{\sum_{i=1}^m \log_e(r_{t_i} + 1)t_i}{\sum_{i=1}^m t_i^2}$$

Where m is the number of time points (t_i) and r_{t_i} is protein H/L ratio measured in a time point t_i . The half-life of a protein ($T_{1/2}$) was calculated using following equation:

$$T_{1/2} = \frac{\log_e 2}{k}$$

Protein H/L ratios derived from a union of three replicates were used for the calculation. Only proteins with H/L ratio measured in at least 5/16 time points were taken for turnover rate calculation. As a quality check of linear regression, the coefficient of determination (R^2) had to be higher than 0.70 to ensure good curve fitting and reliable turnover rate estimation. All proteins with determined protein turnover rates are reported in the Supplementary Tables.

3.2.18 Protein expression for His-tag affinity purification

His₆-HipA and His₆-HipA7. Plasmid pBAD33::*6his hipA* was transformed into MG1655 strain. An overnight culture was grown in LB medium containing 25 µg/ml chloramphenicol and 0.4% glucose, washed, diluted 1,000x into 2 l of LB medium containing 25 µg/ml chloramphenicol and grown at 37 °C. The expression of *hipA* and *hipA7* was induced at OD_{600nm} = 0.4 with 0.2% arabinose for 2 h. His₆-HipA7 was expressed from pBAD33::*6his hipA7* in 1 l of LB medium in the same way as His₆-HipA.

His₆-GltX and His₆-RplK. Plasmid pET28a::*gltX* was transformed into BL21(DE3) strain. An overnight culture was grown in LB medium containing 50 µg/ml kanamycin, diluted 1000x into 250 ml of LB medium with kanamycin and grown at 30 °C. The expression of *gltX* was induced at OD_{600nm} = 0.6 with 1 mM IPTG for 2 h. His₆-RplK was expressed from the pET28a::*rplK* plasmid the same way as His₆-GltX. SeqA-His₆ was expressed from the pET28a::*seqA* plasmid in 750 ml of LB medium the same way as His₆-GltX. The expression of *seqA* was induced at OD_{600nm} = 0.5 with 0.2 mM IPTG for 2 h.

3.2.19 His-tag affinity purification

After the protein expression, cultures were harvested by centrifugation and cell pellets were re-suspended in cold lysis buffer (50 mM HEPES/KOH pH 7.4 at 4 °C, 300 mM NaCl, 10 mM MgCl₂, 2 mM β-mercaptoethanol, EDTA-free protease inhibitors (Roche), 5 mM glycerol-2-phosphate, 5 mM sodium fluoride and 1 mM sodium orthovanadate). Each cell lysate was incubated with 0.5 mg/ml lysozyme and 50 U/ml DNase I for 15 min at RT and sonicated at 40% amplitude until clear and centrifuged at 13,000 g. Supernatant containing 10 mM imidazole was incubated with 500 μL HisPur cobalt resin (Thermo) for 1 h at 4 °C. The cobalt resin was washed in buffer A (50 mM HEPES/KOH pH 7.4 at 4 °C, 300 mM NaCl, 10 mM MgCl₂, 2 mM β-mercaptoethanol) containing 10 mM, 20 mM or 30 mM imidazole. Bound proteins were eluted with buffer A containing 150 mM imidazole. Purified proteins were transferred into a storage buffer and concentrated by ultrafiltration using Amicon Ultra centrifugal filter units (Merck) with a pore size of 30,000 Da (for His₆-HipA, His₆-HipA7 and His₆-GltX) or 10,000 Da (for His₆-RplK and SeqA-His₆). Proteins were washed with storage buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT) and transferred into the storage buffer containing 10% glycerol. Protein concentration was measured using standard Bradford assay (Bio-Rad).

3.2.20 *In vitro* kinase assay measured by MS

For *in vitro* kinase assay 1 μM kinase (His₆-HipA or His₆-HipA7) was incubated with 6 μM His-tagged substrate in a kinase buffer (50 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 1 mM DTT, 16 μM ZnSO₄) with or without 5 mM ATP. Each reaction contained 4.5 μg of a total protein amount. Samples were incubated at 37 °C for 45 min and stopped by the addition of 9 volumes of denaturation buffer, followed by the protein digestion using chymotrypsin or Lys-C endoprotease as previously described (see above). Digested peptides were purified using stage tips (see 3.2.14) and 0.2 μg of each sample was measured by LC-MS/MS (see 3.2.15).

3.2.21 *In vitro* kinase activity of HipA and HipA7 measured by autoradiography

For qualitative comparison of HipA and HipA7 autophosphorylation activity and the phosphorylation activity towards GltX, we performed a time-dependent kinase assay. 1 μM kinase (His₆-HipA or His₆-HipA7) was incubated with 54 μM ATP and 12 μM [γ -³²P]-ATP, and with or without 6 μM His₆-GltX and 19 μM of total *E. coli* tRNA in the kinase buffer (50 mM Tris-HCl pH 8.0, 10 mM MgCl₂, 1 mM DTT, 16 μM ZnSO₄). Reactions with HipA and HipA7 were performed simultaneously using the same ATP stock solution. The reactions were incubated at 37 °C for 45 min and stopped by the addition of Laemmli buffer at indicated time points. Reaction mixtures were separated by SDS-PAGE, revealed by phosphorimaging (GE Healthcare) and analyzed using ImageQuant software (GE Healthcare). The intensity of each time point was normalized to the sum of intensities of all time points and presented as a fraction of total intensity. Phosphorylation activity was determined by the linear regression of the linear part of the intensity over time curve.

3.2.22 Determination of cell viability on SMG plates

It is known that $\Delta relA$ mutant exhibits relaxed phenotype when grown in the presence of single carbon amino acids (SMG) (271). *E. coli* MG1555, *rplK* S102A, *rplK* S102D and $\Delta relA$ deletion mutants (see 1.1.4) were grown in M9 medium with glucose for 24 h to stationary phase. Aliquots of cells were serially diluted in M9 medium, plated on M9 agar plates (50 mM Na_2HPO_4 , 22 mM KH_2PO_4 , 8.6 mM NaCl, 18.7 mM NH_4Cl , 1 mM MgSO_4 , 0.1 mM CaCl_2 , 0.0001% thiamine, 0.4% glucose) supplemented with or without 1 mM of serine, methionine and glycine (SMG) and grown for around 40 h at 37 °C.

3.2.23 Measurement of persistence

Cells were grown in 20 ml of M9 medium with glycerol containing 25 $\mu\text{g}/\text{ml}$ chloramphenicol and ampicillin to the exponential phase. At $\text{OD}_{600\text{nm}}$ of around 0.4, *hipA* was induced with 0.2% arabinose for 95 min. Cultures were then treated with 2 $\mu\text{g}/\text{ml}$ ciprofloxacin for 5 h. For determination of colony forming units (CFU), 1 ml aliquots were taken before arabinose addition, 95 min after *hipA* induction and 5 h of ciprofloxacin treatment. Cells were washed with PBS, serially diluted, plated on LB agar plates containing 0.4% glucose and grown for 24 - 40 h at 37 °C. Persistence was calculated by dividing the number of CFU/ml of ciprofloxacin treated culture with the CFU/ml of the culture before antibiotic addition and presented as a frequency of surviving (persister) cells in log₁₀ scale.

4 Results

4.1 Identification of *in vivo* phosphorylation targets of HipA kinase

*Parts of the chapters were adapted from **Semanjski M**, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B, The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli proteome to promote multidrug tolerance, Science Signaling, 2018, 11(547) (260).*

4.1.1 Phosphoproteomic workflow for studying HipA targets

The first goal of this thesis was to characterize proteome and phosphoproteome of the persistent state associated with HipA kinase activity. Given that HipA alters several main cellular processes, we assumed that the kinase likely acts on additional targets along with the well-described GltX (63, 77). Kinase-substrate relationships are usually determined by using deletion mutants (176); however, because *hipA* is expressed in very low amounts and HipA is kept inactive by a direct interaction with HipB in the wild-type cells (47), substrates were screened using an established model in which HipA is ectopically overproduced in growing cells causing growth inhibition and persistence (63). To gain insight into persistent phenotype related to the gain-of-function HipA variant, HipA7, a less artificial model with chromosomally encoded kinase variants was later used.

For all phosphoproteomic screens employed in this thesis, a quantitative proteomic workflow based on SILAC labeling methodology in combination high-resolution mass spectrometry described in **Figure 13** was used. For that, *E. coli* cultures were first metabolically labeled by growing the cells in the minimal medium supplemented with stable isotope-labeled derivatives of lysine: “light” lysine (Lys0), “medium-heavy” lysine (Lys4) or “heavy” lysine (Lys8). Proteins were isolated from three differentially labeled cell extracts, mixed in equal amounts and digested with endoproteinase Lys-C that cleaves proteins specifically after lysine residues. Additionally, the same amount of protein mixture was digested separately with a less specific protease chymotrypsin in order to generate peptides that contain multiple lysine residues in their sequence, which would otherwise be too short for the LC-MS/MS measurement. Digested peptide mixture was either pre-fractionated (e.g. by high pH RP fractionation) or directly subjected to LC-MS/MS measurement to obtain the information about relative protein amounts. To analyze phosphoproteome, digested peptides were first enriched with titanium dioxide beads, which selectively bind negatively charged phosphopeptides, and then analyzed by LC-MS/MS as separate fractions using optimized methods with long injection times and high resolution (see 3.2.15). Finally, the acquired raw MS data was processed using MaxQuant software and the data was analyzed with Excel and Perseus software. Prior to protein mixing, the incorporation of the labels was inspected by separate LC-MS/MS measurements of individual Lys4- and Lys8-labeled samples. In all experiments, the incorporation of Lys4 and Lys8 was > 94% confirming that almost complete labeling of *E. coli* proteome was achieved, which was sufficient for further quantitative comparison.

Results

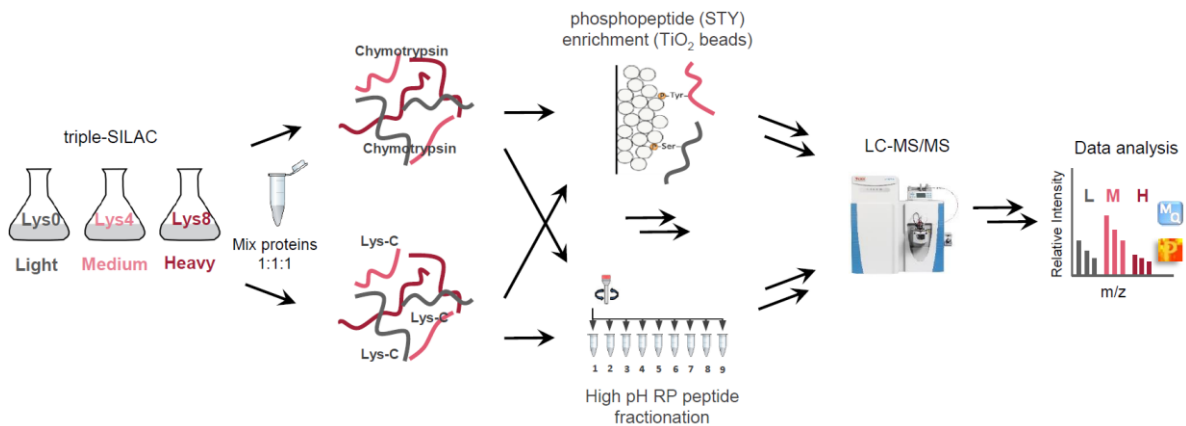


Figure 13. General SILAC-based phosphoproteomic workflow. Bacterial cultures were differentially SILAC-labeled with light¹ lysine (Lys0, L), “medium-heavy” lysine (Lys4, M) or “heavy” lysine (Lys8, H) for their quantitative comparison. Proteins were isolated from cell lysates, mixed in equal amounts 1:1:1 and separately digested with endoproteinase Lys-C or chymotrypsin. Phosphopeptides were enriched with titanium dioxide (TiO₂) beads for 5-10 consecutive rounds and fractions were analyzed by LC-MS/MS (Q Exactive HF mass spectrometer). For proteome measurement, peptides were either pre-fractionated using e.g. high pH reverse phase peptide fractionation or directly subjected to LC-MS/MS. Afterwards, acquired MS data was processed using MaxQuant software and data analysis was performed using different statistical or bioinformatic tools, e.g. Perseus software.

4.1.2 HipA phosphorylates multiple proteins in addition to GltX

To screen for phosphorylation targets of HipA kinase *in vivo*, the expression of *hipA* was induced from a low copy number plasmid pBAD33 in *E. coli* K-12 MG1655 background. The pBAD33 plasmid carries an arabinose-inducible promoter, which can be repressed with glucose during cloning procedures or in pre-cultures to suppress toxic effects of overexpressed *hipA*. The production of HipA inhibited cell growth shortly after the induction (**Figure 14A**) and led to an increase in persistence compared to the empty plasmid (**Figure 14B**), as previously reported (63). Namely, the viability of cells in the presence of ciprofloxacin was around 100-fold higher already 95 minutes after *hipA* induction with arabinose than before the induction and also in comparison to the arabinose-treated cells carrying the empty plasmid. Conversely, overproduction of HipB from a low copy number plasmid pNDM220 with isopropyl- β -D-thiogalactopyranoside (IPTG) counteracted HipA-induced toxicity and growth inhibition and led to the resumption of growth and resuscitation from dormancy (**Figure 14A**). Quantitative phosphoproteome analysis was performed in a form of two triple-SILAC experiments with a common time point labeled with Lys0 and collected just before *hipA* induction at an optical density of 600 nm (OD_{600nm}) of 0.4 (**Figure 14A**). Cultures collected at two time points after *hipA* and *hipB* induction were differentially labeled as follows: 75min (Lys4) and 3h (Lys8) after *hipA*, 2.5h (Lys4) and 6h (Lys8) after *hipB* induction. A combination of these two SILAC experiments with the common time point allowed to follow the relative change in protein abundance and in Ser/Thr/Tyr phosphorylation of individual proteins in a temporal fashion during HipA-induced growth inhibition and HipB-triggered resuscitation (**Figure 14** and **Figure 15**).

Results

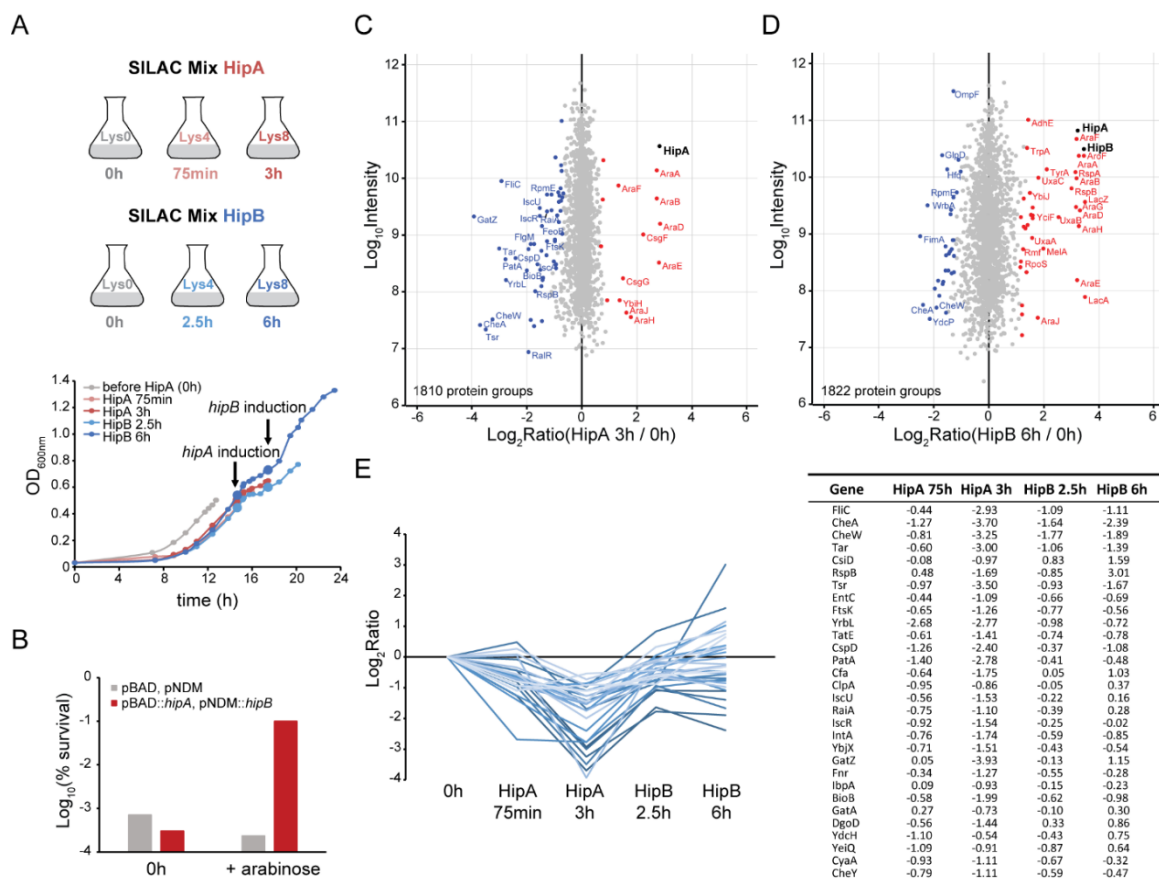


Figure 14. Proteome analysis of HipA-induced growth inhibition and HipB-induced resuscitation. (A) Growth curves of *E. coli* MG1655 strain containing the pBAD::hipA plasmid, in which *hipA* expression is under the control of an arabinose-inducible promoter, and the pNDM::hipB plasmid, in which *hipB* is under the control of an IPTG-inducible promoter. Strains were grown in SILAC-labeled minimal medium containing either “light” lysine (Lys0), “medium-heavy” lysine (Lys4) or “heavy” lysine (Lys8), plus the appropriate antibiotics for the retention of the plasmids. Expression of *hipA* was induced at OD_{600nm} = 0.4 with arabinose, and samples were collected before (Lys0) and 75min (Lys4) and 3h (Lys8) after induction. At the 3h time point, *hipB* expression was induced with IPTG, and samples were collected 2.5h (Lys4) and 6h (Lys8) later. Growth curves are representative of 2 biological replicates. (B) Survival to ciprofloxacin (persistence) of *E. coli* MG1655 after *hipA* expression. Cultures containing pBAD::hipA and pNDM::hipB plasmids or pBAD and pNDM empty plasmids were grown in minimal medium containing appropriate antibiotics for the retention of the plasmids. Expression of *hipA* was induced at OD_{600nm} = 0.4 with arabinose for 95min. Aliquots of cultures were exposed to ciprofloxacin for 5h. Persistence was determined for one biological replicate as described previously (15). (C, D) Distribution of protein SILAC ratios 3h after *hipA* expression (C) and 6h after *hipB* expression (D). Proteins that significantly increased (red) or decreased (blue) in abundance are indicated (significance B test, *p*-value < 0.001). Distributions are representative of 2 biological replicates. (E) STEM clustering analysis of protein dynamics during *hipA* and *hipB* expression. Log₂ transformed protein SILAC ratios of proteins significantly changed during at least one time point (significance B test, *p*-value < 0.001) were used for clustering. The names and the log₂ transformed ratios of proteins of one significant cluster (*p*-value of 9.2E-08) are shown in the table.

Proteome analysis of a total of 2,387 identified and more than 1,800 quantified proteins per replicate confirmed that protein amounts of HipA and HipB increased around 8-fold 3 hours and 6 hours post induction, respectively (Figure 14C, D). Likewise, the significant increase in protein amounts (*p*-value < 0.001) was observed for proteins encoded downstream the arabinose and *lac* promoter confirming the efficient expression from plasmids. More importantly, a number of differentially regulated proteins changed significantly in abundance in response to *hipA* expression of which 15 proteins increased and 50 decreased in abundance. Conversely, 41 proteins showed higher and 31 lower

Results

abundance 6 hours after *hipB* induction. Time-series clustering analysis of proteins that significantly changed in abundance in at least one time point revealed one significant cluster (p -value of $9.2E-08$) of proteins that decreased in abundance during growth inhibition and returned to initial levels upon resuscitation (**Figure 14E**). Those proteins involved, for example, chemotaxis-related proteins that are important for regulation of cell motility, such as CheA, CheW, CheY, Tar and flagellin (FlhC). In contrast, curli production and assembly/transport components CsgF and CsgG were among proteins that increased in abundance upon growth inhibition (**Figure 14C**).

In this phosphoproteomic experiment, a total of 380 phosphorylation sites were identified on 230 proteins (**Figure 15A, B**). Changes in phosphorylation, represented as phosphorylation site SILAC ratios, were corrected for the unequal mixing of differentially SILAC-labeled samples and normalized to the differences in protein abundances to exclude any quantitation bias due to fluctuations of corresponding proteins; this was applied to all phosphoproteomic experiments in the thesis. The reproducibility of phosphorylation site SILAC ratios measured in two biological replicates was good (Pearson correlation coefficient of 0.76 and 0.79) (**Figure 15C, D**). Surprisingly, a strong increase in the global phosphorylation of many phosphorylated proteins was detected as a result of HipA-induced growth inhibition (**Figure 15A**). Because of such binary distribution caused by high perturbation of the system upon *hipA* expression and the lack of sufficient number of data points in the non-changing background, commonly used significance B test was not applicable. Rather, the significance threshold for determination of differentially regulated phosphorylation sites was defined manually as a 4-fold change in phosphorylation site SILAC ratio and applied to all phosphoproteomic experiments in this thesis. As anticipated, phosphorylation of GltX on the Ser²³⁹ residue was one of the phosphorylation events with the highest increase (9-fold) (**Figure 15A**), which confirmed the efficacy of this model in studying targets of HipA kinase. Because Ser²³⁹ is surrounded by three lysine residues in the primary sequence of GltX (²³⁵GKKLSKR²⁴¹), phosphorylated peptide of GltX was detected only in the samples digested with chymotrypsin. In fact, Ser²³⁹ is part of a “KMSKS” motif that is unique to the class I of amino acid-tRNA ligase family involved in the stabilization of the transition state of the amino acid activation reaction (272). As expected, elongation factor Tu (EF-Tu), which was for a long time believed to be the main substrate of HipA, was not detected as phosphorylated in this data set. However, another elongation factor, Ts (Tsf) was found phosphorylated and showed increase in phosphorylation upon *hipA* induction. In addition to GltX, multiple proteins showed an increase in phosphorylation (**Figure 15A**) revealing that HipA kinase likely acts on multiple substrates under the conditions studied. Among those, there were no other amino acid-tRNA ligases phosphorylated confirming the specificity of HipA towards GltX (63). However, many other phosphorylated proteins were found this study, and several of them were components of the ribosome, such as ribosomal proteins S4 (RpsD), S7 (RpsG), S9 (RpsI), S10 (RpsJ), L11 (RplK) and L31 (RpmE). Indeed, translation and DNA metabolic process were enriched (p -value < 0.01) in the gene ontology (GO) functional analysis of up-regulated phosphorylation events upon *hipA* expression (**Figure 16B**) suggesting that, besides GltX, other constituents of the translation machinery might be regulated by HipA in order to efficiently attenuate translation. Other phosphorylated proteins were involved in replication, such as the negative regulator of initiation of replication (SeqA) or in regulation of transcription such as DNA-binding proteins RcsB, Fis and Hns.

Results

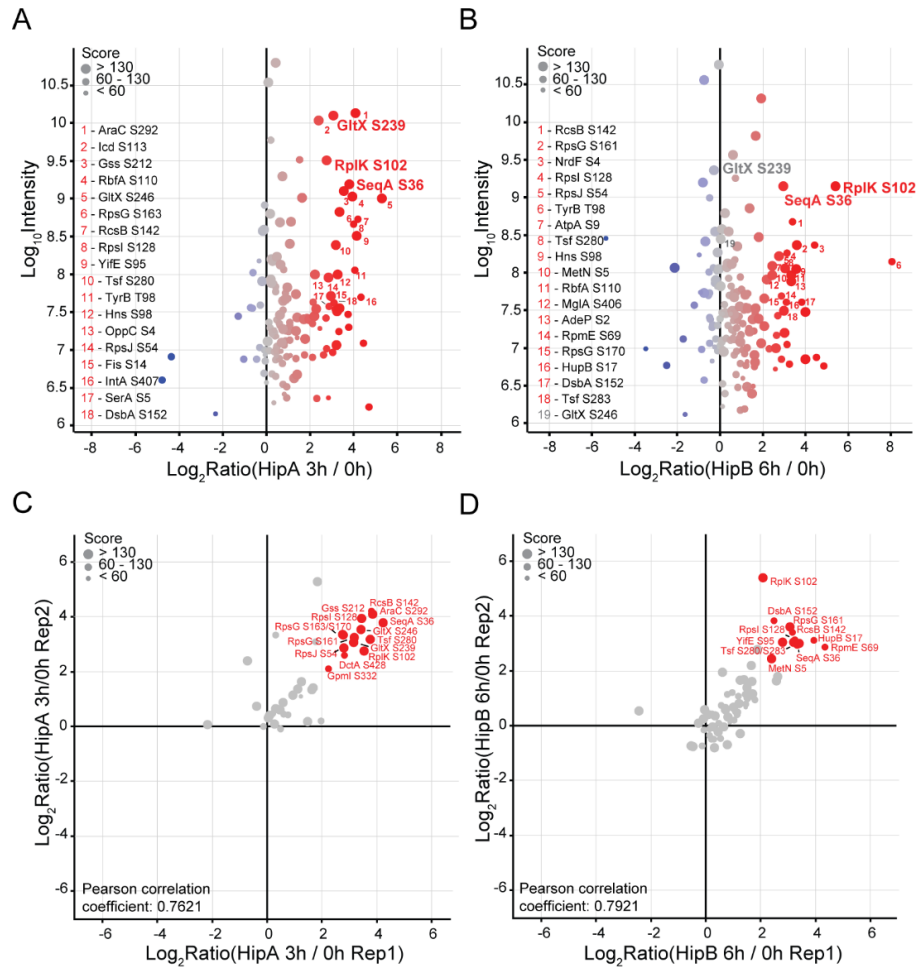


Figure 15. Phosphoproteomic analysis of HipA-induced growth inhibition and HipB-induced resuscitation. For experimental design, see Figure 14A. (A, B) Distribution of phosphorylation site SILAC ratios 3h after *hipA* expression (A) and 6h after *hipB* expression (B). The names of the phosphorylated proteins and the positions of the phosphorylation sites showing at least a 4-fold increase in phosphorylation are indicated. Distributions are representative of 2 biological replicates. (C, D) Correlation of phosphorylation site SILAC ratios 3h after *hipA* (C) and 6h after *hipB* expression (D) for two biological replicates. The names of the phosphorylated proteins and the positions of the phosphorylation sites showing at least a 4-fold change in both biological replicates are indicated.

After triggering resuscitation by inducing the production of the antitoxin HipB, which neutralizes HipA toxicity, phosphorylation of GltX decreased to its initial level. Other phosphorylation sites also followed similar declining pattern, albeit they showed a slower rate of the reduction in phosphorylation and were not dephosphorylated entirely (**Figure 16A**). Considering that a repertoire of potential HipA targets was identified in this study, the obvious next step was to test whether they share a specific kinase target linear motif. For that, overrepresentation of a sequence around Ser and Thr residues was inspected by the Motif-X software. However, no significant sequence similarity around the phosphorylation sites was detected, indicating that HipA kinase does not specifically recognize a common linear sequence motif.

Results

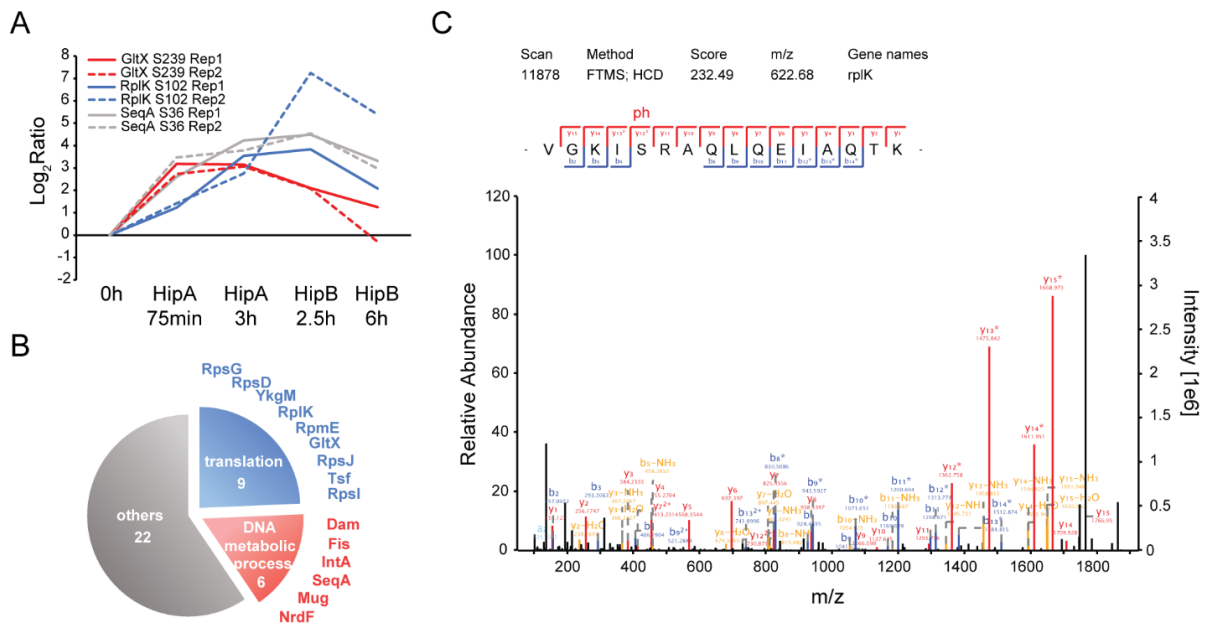


Figure 16. Phosphorylation sites increased in phosphorylation during HipA-induced growth inhibition and HipB-induced resuscitation. (A) Phosphorylation site dynamics of GltX, RplK, and SeqA over four time points during growth inhibition (HipA 75min and HipA 3h) and resuscitation (HipB 2.5h and HipB 6h) in 2 biological replicates. (B) Gene ontology (GO) distribution of phosphoproteins showing at least a 4-fold increase in phosphorylation 3h after *hipA* expression enriched against the background of all identified phosphoproteins (p -value < 0.01). The number of enriched phosphoproteins is indicated below the category name. The distribution is representative of 2 biological replicates. (C) MS/MS fragmentation spectra of RplK peptide phosphorylated on Ser¹⁰² residue.

To additionally confirm that the detected phosphoproteins are indeed direct substrates of HipA, an *in vitro* kinase assays with purified HipA, GltX, and two chosen candidate substrates, SeqA and ribosomal protein L11 (hereafter referred to as RplK) were performed. Purified His₆-GltX, His₆-RplK or SeqA-His₆ were incubated with His₆-HipA kinase and ATP and after the reaction was completed, the samples were digested separately with Lys-C and chymotrypsin and measured by LC-MS/MS (**Figure 17A**). For each substrate, appropriate controls without ATP or kinase were performed adding to a total of 22 LC-MS/MS runs per experiment. Phosphorylated peptides of corresponding substrate proteins was either present or absent in different samples (**Figure 17B**). Indeed, phosphorylation site on Ser¹⁰² of RplK was detected only when purified His₆-RplK was incubated with His₆-HipA and ATP, suggesting that RplK is a substrate of HipA kinase. Phosphorylation of SeqA on Ser³⁶ was detected already in purified SeqA-His₆ without addition of ATP, suggesting that SeqA is phosphorylated endogenously. However, the intensity of the SeqA phosphorylation site significantly increased in the presence of His₆-HipA, which indicated that HipA phosphorylated SeqA *in vitro*. Taken together, this shows that overproduced HipA modifies multiple protein targets in addition to the canonical GltX.

Results

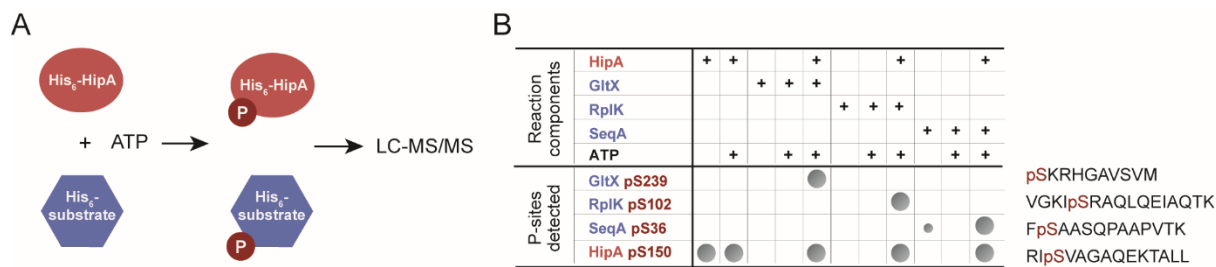


Figure 17. *In vitro* HipA kinase assay. (A) Experimental design of the kinase assay of His₆-HipA with His₆-GltX, His₆-RplK, and SeqA-His₆. Following the phosphorylation reaction, the samples were protease-treated and analyzed by LC-MS/MS. (B) Identified phosphorylation sites represented as circles were detected in 2 independent experiments. Smaller circle depicted for SeqA indicates two orders of magnitude lower intensity of phosphorylation site without His₆-HipA relative to the intensity measured in the presence of His₆-HipA. Sequences of best identified phosphopeptides of particular proteins are shown.

4.1.3 HipA autophosphorylation on Ser³⁵⁹ has no influence on HipA activity

In this data set, an additional phosphorylation site on GltX on Ser²⁴⁶ was detected, but it was modified to much lesser extent and identified by low-quality spectra. This phosphorylation could possibly result from a weak target motif specificity often seen in bacterial kinases. In addition, three phosphorylation sites on HipA itself were identified: on a well-known position Ser¹⁵⁰, and two novel sites on Ser¹⁵⁸ and Ser³⁵⁹. Phosphorylated Ser¹⁵⁰ was previously described as a kinase-inhibiting autophosphorylation site (81) and the mutation of Ser¹⁵⁰ to alanine (S150A) abolished the ability of HipA to induce persistence (75). Other two phosphorylation sites on HipA have not been reported to this day. Phosphorylation of Ser¹⁵⁸ was detected on a peptide that was at the same time phosphorylated on Ser¹⁵⁰, in which the intensity of Ser¹⁵⁸ phosphorylation was much lower than of Ser¹⁵⁰. In contrast, peptide phosphorylated on Ser³⁵⁹ was identified with higher confidence on singly phosphorylated peptide. Considering that Ser³⁵⁹ is situated close to the DNA binding region (Lys³⁷⁹ to Arg³⁸²) of HipA in the structure of HipA-HipB-promoter complex (PDB ID: 4YG7) (19), phosphorylation of Ser³⁵⁹ was likely to be functionally important. Therefore, to investigate the function of Ser¹⁵⁰ and Ser³⁵⁹ phosphorylation sites of HipA, phospho-ablative and phospho-mimetic mutants of Ser¹⁵⁰ and Ser³⁵⁹ residues of HipA were constructed (HipA S150A, HipA S150D, HipA S359A and HipA S359D), in which Ser¹⁵⁰ or Ser³⁵⁹ were substituted with alanine or aspartate to mimic phosphorylated serine, respectively. These HipA mutants were expressed from the pBAD33 plasmid and their phosphoproteomes were compared in two triple SILAC-based experiments (**Figure 18**). Whereas phospho-mimetic and -ablative mutations at Ser¹⁵⁰ residue impaired the activity of HipA observed through decrease in phosphorylation of above identified substrates (**Figure 18A**), the same changes at Ser³⁵⁹ did not affect HipA activity as there was no difference in the phosphoproteomes of HipA S359A and S359D mutants compared to the wild-type HipA (**Figure 18B**). Nevertheless, unexpectedly, it was observed that the phosphorylation of GltX was reduced to lesser extent in HipA S150D mutant compared to the HipA S150A, suggesting that the substitution of Ser¹⁵⁰ by aspartate did not entirely abolish kinase activity of HipA.

Results

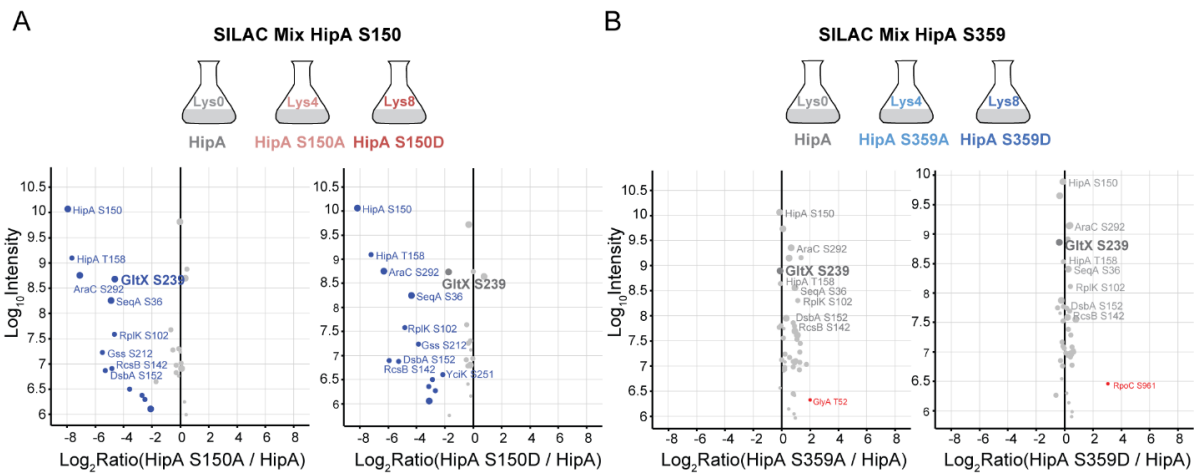


Figure 18. Phosphoproteome analysis of phosphorylation mutants of HipA. *E. coli* MG1655 strains containing pBAD plasmid carrying either a wild-type *hipA* (Lys0 labeled) or phospho-ablative S150A (A) or S359A (B) (both Lys4 labeled), or phospho-mimetic S150D (A) or S359D (B) mutant of *hipA* (both Lys8 labeled), in which Ser¹⁵⁰ or Ser³⁵⁹ of HipA were substituted either with alanine or aspartate were cloned. Strains were grown in SILAC-labeled minimal medium and at OD_{600nm} = 0.4 *hipA* expression was induced with arabinose for 95min. Distribution of phosphorylation site SILAC ratios after expression of phospho-ablative (left) or phospho-mimetic mutant of *hipA* (right) relative to the wild-type *hipA*. The names of the phosphorylated proteins and the positions of the phosphorylation sites showing at least a 4-fold decrease in HipA S150 mutants are indicated.

4.1.4 RplK phosphorylation on Ser¹⁰² by HipA has no influence on RelA-dependent persistence

One of the particularly interesting HipA substrates identified in these phosphoproteomic screens is the ribosomal protein L11 (RplK), phosphorylated on Ser¹⁰² (**Figure 15A**). Phosphorylated peptide of RplK was identified based on several MS/MS fragmentation spectra of a high quality with good sequence coverage (**Figure 16C**). Moreover, phosphorylation of RplK significantly increased upon inhibition of the growth, but continued to rise during early resuscitation, after which it started to decrease (**Figure 16A**). This suggested that phosphorylation of RplK might be involved in both, persistence and resuscitation. It is proposed that RplK coordinates deacyl-tRNA for the activation of GTP pyrophosphokinase RelA at the ribosome A site (273). Ribosome-bound RelA adopts the active conformation in the presence of a deacylated tRNA and synthesizes (p)ppGpp (100). Hence, the hypothesis was that the phosphorylation of RplK might influence the activity of RelA. To test whether the activity of RelA is altered by phosphorylated RplK, chromosomal *rplK* phospho-ablative and phospho-mimetic mutants (*rplK S102A* and *rplK S102D*), were constructed using genome replacement with pKOV vector (262). Gene replacement resulted in replacement of Ser¹⁰² of RplK with alanine or aspartate, in which letter is assumed to mimic phosphorylation. Viability of constructed mutants and the mutant lacking *relA* gene ($\Delta relA$) on the serine-methionine-glycine (SMG) M9 plates was assessed. Presence of only single carbon amino acids induces isoleucine starvation (271) and cell growth under this condition requires (p)ppGpp synthesis to induce isoleucine biosynthesis. As expected, $\Delta relA$ mutant failed to grow on SMG plate, due to its impaired ability to synthesize (p)ppGpp during amino acid starvation; however, *rplK* mutant strains were as viable as the wild-type (**Figure 19**). This implies

Results

that the phosphorylation of RplK is not sufficient to influence RelA-dependent survival under amino acid starvation. However, this does not preclude a possible functional role of this phosphorylation event under different conditions, or in combination with other HipA-induced phosphorylation events.

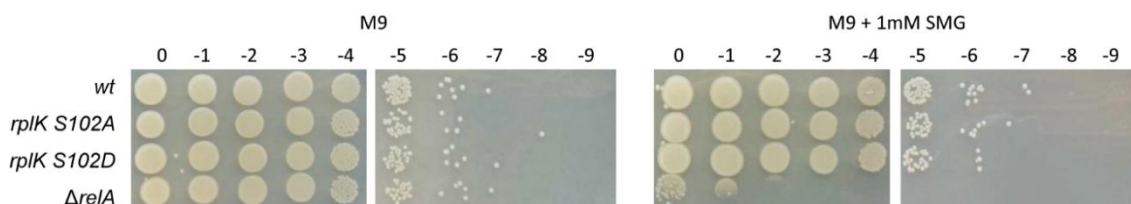


Figure 19. Viability assay of wt MG1655, *rplK S102A*, *rplK S102D* and $\Delta relA$ strains. Stationary phase cultures were grown in minimal medium serially diluted and plated on M9 agar plates supplemented with or without serine, methionine and glycine.

4.2 Comparison of HipA and HipA7 (phospho)proteomes to distinguish mechanistically between different phenotypes of two kinase variants

*Parts of the chapters were adapted from **Semanjski M**, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B, The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli proteome to promote multidrug tolerance, *Science Signaling*, 2018, 11(547) (260).*

4.2.1 HipA7 has fewer *in vivo* substrates than HipA

The second aim of this thesis was to investigate the difference in the kinase activity and the substrate specificity of toxic HipA and less toxic HipA7 kinase, that, despite the differences in toxicity, induce persistence to similar extent, as shown previously (75). For that, *hipA* and *hipA7* genes were expressed from the low copy number pBAD33 plasmid in the *E. coli* MG1655 background. The phosphoproteomes of strain with overproduced HipA (Lys4) and HipA7 (Lys8) were compared directly in one triple-SILAC experiment that also included the empty pBAD33 plasmid as a negative control (Lys0) (**Figure 20A**). Whereas expression of *hipA* caused immediate growth inhibition, *hipA7* induction did not attenuate growth, which is consistent with previous studies (75). In this phosphoproteomic experiment, 173 phosphorylation sites were identified on 130 proteins in three biological replicates that showed good reproducibility (Pearson correlation coefficient of 0.90, 0.90 and 0.89 for HipA7 and HipA comparison) (**Figure 20D, E**). Despite of apparently higher abundance of HipA7 in the cell after induction (**Figure 23A, B**), only phosphorylation of GltX was common to both HipA and HipA7 (**Figure 20B, C**). Surprisingly, no other targets of HipA7 were detected consistently in all three biological replicates. This was in stark contrast to overproduction of HipA, which led to the phosphorylation of several (> 4-fold decrease in HipA7/HipA SILAC ratio) (**Figure 20C**), which were detected in previous phosphoproteomic screen (**Figure 15A**). Namely, RplK, SeqA, RcsB, Tsf and other proteins were

Results

phosphorylated by HipA and not by HipA7 on the same positions as previously detected (**Figure 15B, C**). In addition, this experimental setup also served as a negative control for the initial experiment (**Figure 15A**) by comparing the phosphoproteome of the strain carrying pBAD::*hipA* plasmid induced by arabinose to the strain carrying empty pBAD plasmid also under arabinose-treated conditions. This approach confirmed that increased phosphorylation originated from HipA production and not from arabinose addition.

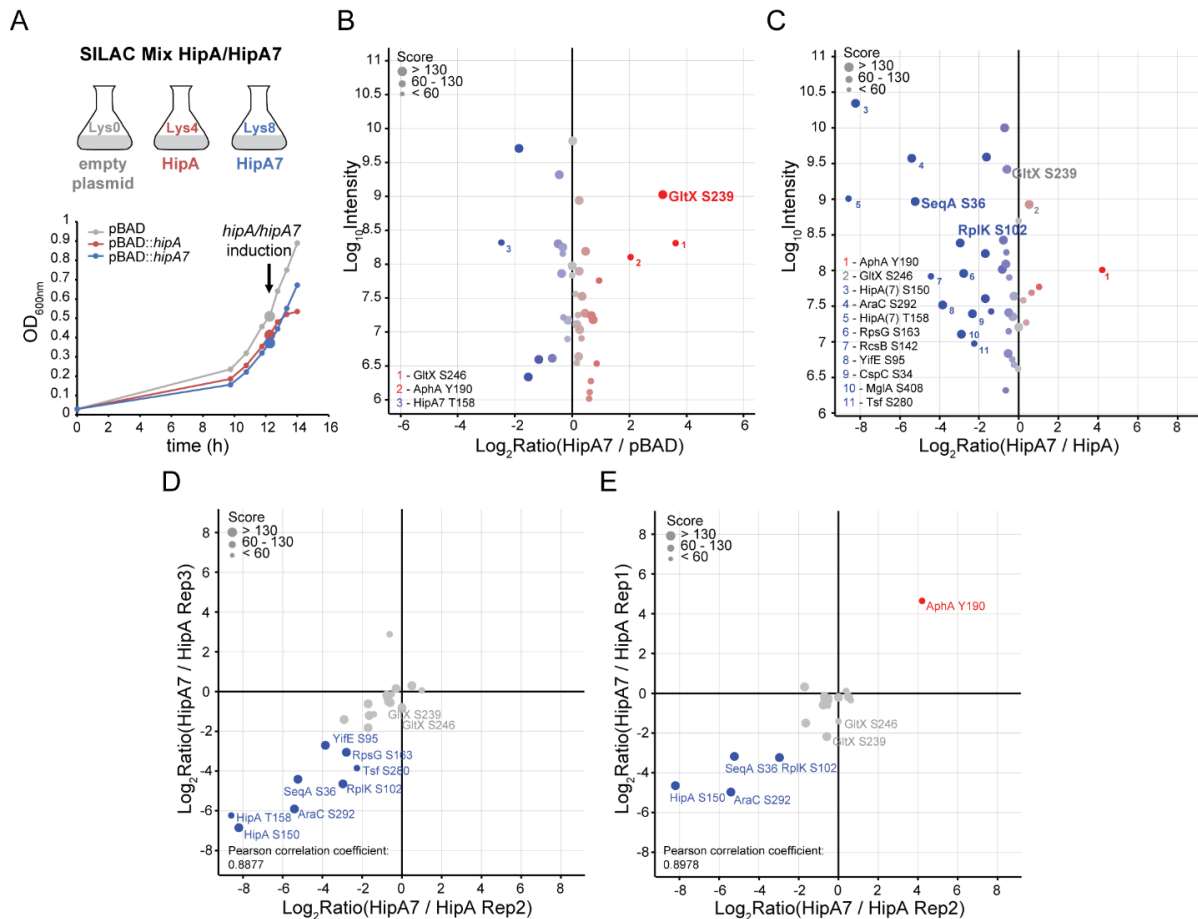


Figure 20. Phosphoproteomic comparison of HipA and HipA7 produced in similar amounts. (A) Growth curves of *E. coli* MG1655 strains containing either pBAD, pBAD::*hipA* or pBAD::*hipA7* plasmid, in which *hipA* and *hipA7* expression is under the control of the arabinose-inducible promoter, and the pNDM220::*hipB* plasmid. Strains were grown in SILAC-labeled minimal medium containing either “light” lysine (Lys0), “medium-heavy” lysine (Lys4) or “heavy” lysine (Lys8) and appropriate antibiotics for the retention of the plasmids. *hipA* and *hipA7* expression was induced at OD_{600nm} = 0.4 with arabinose, and samples were collected 95min later. Growth curves are representative of 3 biological replicates. (B, C) Distribution of phosphorylation site SILAC ratios 95min after *hipA7* expression relative to the empty plasmid (B) and relative to the *hipA* expression (C). The names of the phosphorylated proteins and the positions of the phosphorylation sites showing at least a 4-fold change in phosphorylation are indicated. Distributions are representative of 3 biological replicates. (D, E) Correlation of phosphorylation site SILAC ratios of *hipA7* relative to *hipA* expression for replicate 2 and 3 or (D) replicate 2 and 1 (E). The names of the phosphorylated proteins and the positions of phosphorylation sites showing at least a 4-fold change in phosphorylation in both biological replicates are indicated.

Furthermore, in this experimental setup, we were able to calculate the proportion of phosphorylated proteins at the modification site (phosphorylation site occupancy) using a SILAC ratio of the modified peptide, a SILAC ratio of its unmodified peptide counterpart and a SILAC ratio of the respective protein

Results

(267). Due to the low intensity of the non-phosphorylated version of GltX peptide, we were able to determine the occupancy of GltX phosphorylation site only for one replicate (**Figure 21A**). In the presence of HipA, 76% of GltX molecules were phosphorylated; this level dropped to 48% when HipA7 was produced, which is indicative of a lower kinase activity of HipA7 than HipA towards GltX. The phosphorylation site occupancy of other HipA substrates RplK and SeqA was determined to be around 7% for both proteins when HipA was produced (**Figure 21B, C**).

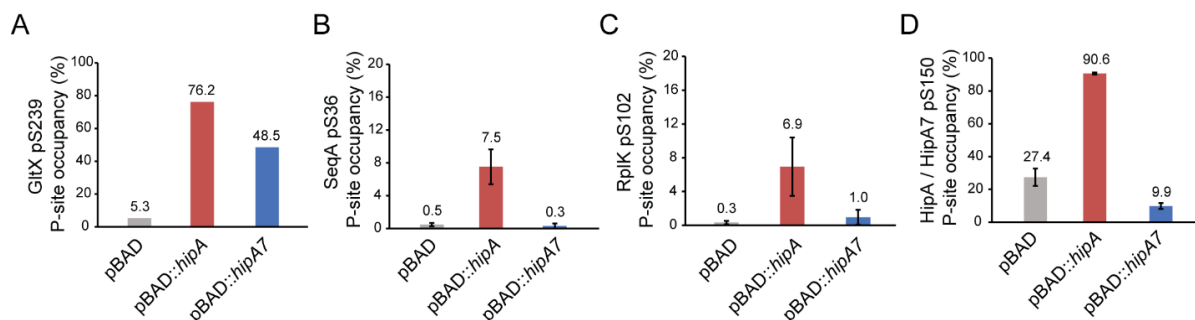


Figure 21. Phosphorylation site occupancy after *hipA* and *hipA7* expression. (A) Occupancy of the GltX Ser²³⁹ phosphorylation site (P-site) determined by MaxQuant in 1 out of 3 biological replicates. (B) Occupancy of SeqA Ser³⁶ and (C) RplK Ser¹⁰² phosphorylation sites calculated manually. (D) Occupancy of the HipA and HipA7 Ser¹⁵⁰ autophosphorylation site determined by MaxQuant. Data in (B, C, D) are means \pm SD from 3 biological replicates.

The observed lower kinase activity of HipA7 towards GltX was additionally tested *in vitro* by time-resolved kinase assay using autoradiography (**Figure 22A**). For that, purified His₆-HipA and His₆-HipA7 were incubated with purified His₆-GltX, total tRNA from *E. coli* and radioactive [γ -³²P]-ATP, and reaction was stopped at several distinct time points to assess dynamics of *in vitro* phosphorylation reaction detected by phosphorimaging. Indeed, HipA7 kinase showed around 100-fold lower phosphorylation activity towards GltX compared to HipA. Together with fewer detected substrates, this implies that two amino acid substitutions in HipA7 (G22S and D291A) led to a decrease in the kinase activity. Furthermore, the phosphoproteomic data also revealed that 90% of the HipA kinase was autophosphorylated on Ser¹⁵⁰, whereas only 10% of HipA7 was autophosphorylated when overproduced (**Figure 21D**). The activity of HipA and HipA7 autophosphorylation was further assessed by *in vitro* autoradiography assay in which His₆-HipA and His₆-HipA7 were incubated with [γ -³²P]-ATP without GltX (**Figure 22B**). The activity of HipA7 autophosphorylation was around 8-fold lower than of HipA additionally confirming lower kinase activity of HipA7 variant.

Results

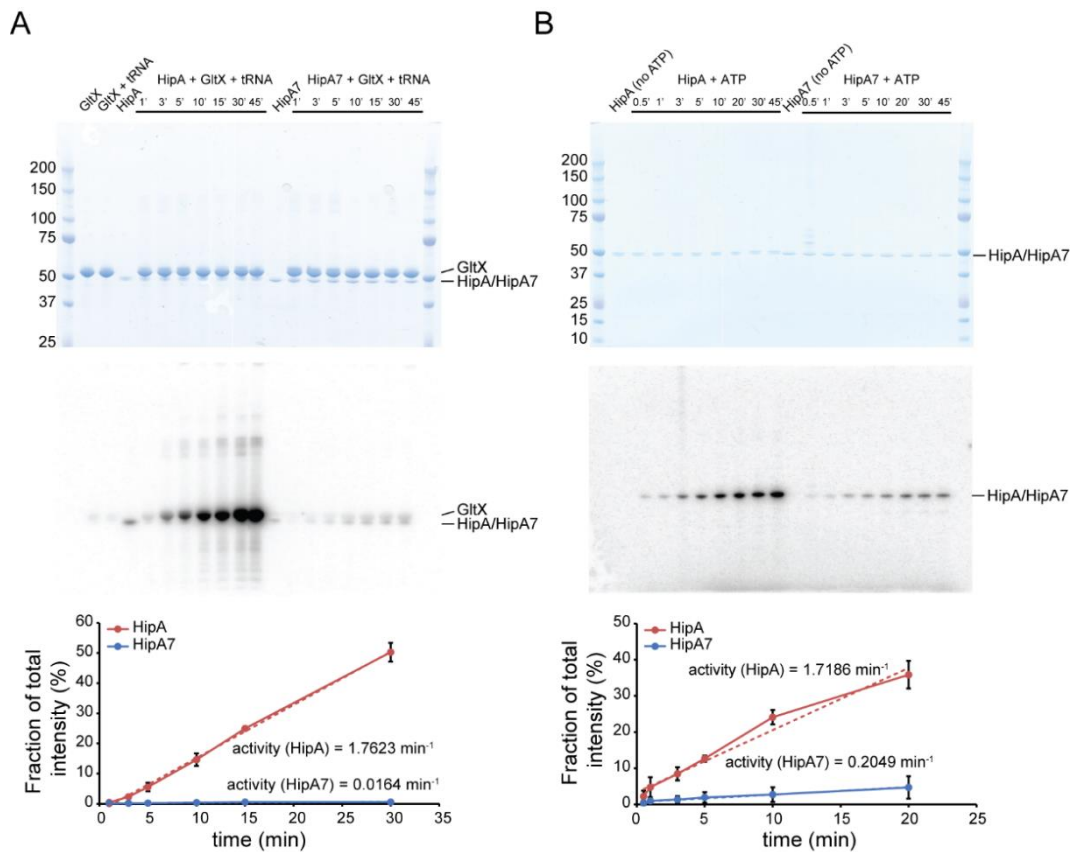


Figure 22. *In vitro* kinase assay with $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$. Kinetics of GltX *in vitro* phosphorylation by HipA or HipA7 (A) or HipA and HipA7 autophosphorylation (B). His₆-HipA or His₆-HipA7 was incubated with or without His₆-GltX, 20 μM of total *E. coli* tRNA extract and $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$. The phosphorylation reactions were stopped at specific time points, proteins were separated by SDS-PAGE and phosphorylation was measured by autoradiography. The gels are representative of 3 independent experiments. The phosphorylation activity was determined by linear regression of the time dependent increase in intensity. Data are means \pm SD from 3 independent experiments.

4.2.2 Overproduction of HipA7, but not HipA, leads to increased abundance of multiple chaperones and proteases

Apart from the obvious differences in the substrate pools of HipA and HipA7, proteome alterations were also observed when different kinase variants were expressed (**Figure 23A, B**). Among 1,458 quantified proteins, 26 showed significant increase in abundance (significance B test, p -value < 0.001) upon *hipA* and 42 upon *hipA7* induction, whereas the abundance of 67 and 47 proteins decreased, respectively. In particular, the components of the stress-induced multi-chaperone system (ClpB, DnaK, DnaJ, GrpE) and other chaperones and chaperonins were significantly increased in abundance only in the presence of HipA7 (**Figure 23B**), but not of HipA (**Figure 23A**). Concomitantly, gene ontology (GO) term analysis of proteins with increased abundance as a result of *hipA7* expression revealed that protein folding was the highest significantly enriched process (**Figure 23C**), which also included proteins of a proteasome-like degradation complex (HslU and HslV). Apart from components of the protein folding machinery, the abundance of small heat shock proteins IbpA and IbpB that associate with aggregated proteins and protect them from proteolysis was highly increased in the presence of

Results

HipA7 (**Figure 23B**). In addition, an autotransporter protein Antigen 43 (Flu) involved in cell aggregation and biofilm formation increased in abundance when HipA7 was overproduced (274).

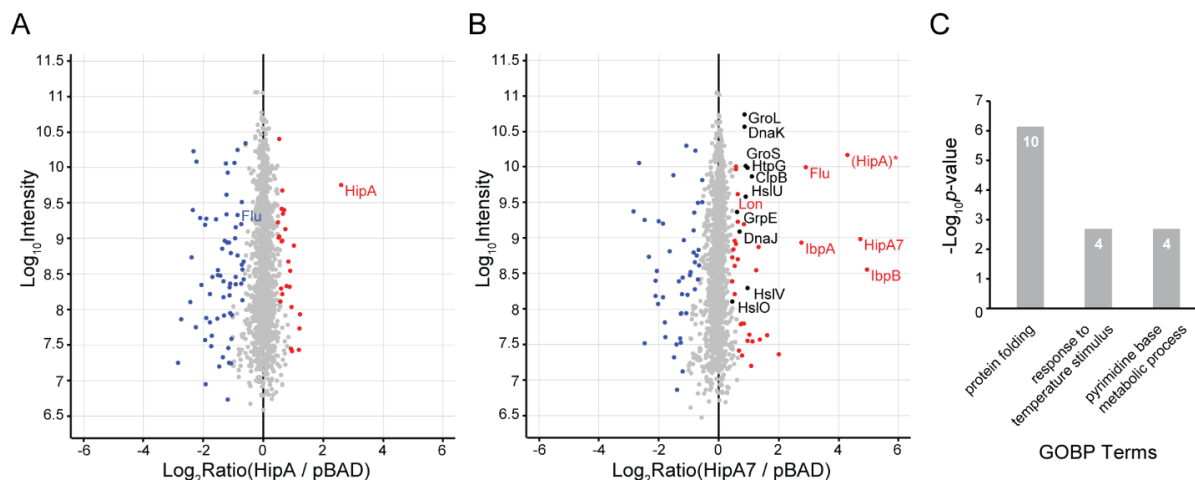


Figure 23. Proteome comparison of HipA and HipA7 produced in similar amounts. (A, B) Distribution of protein SILAC ratios after *hipA* expression (A) or *hipA7* expression (B) relative to the empty pBAD plasmid. Proteins with increased abundances are indicated in red and with decreased abundances in blue. The names of the proteins involved in protein folding (see C) are marked in black. *(HipA) indicates a protein group that is identified by peptides common for HipA and HipA7 sequences; in (B), these peptides belong only to HipA7. HipA7 indicates a protein group identified by peptides unique for HipA7. Distributions are representative of 3 biological replicates. (C) GO analysis of proteins with significantly increased abundances (significance B test, p -value < 0.001) after *hipA7* expression relative to the empty plasmid enriched against the background of all identified proteins. The number of enriched proteins in each category of GO biological processes is indicated within the bars.

4.2.3 Increasing the overproduction of HipA7 partially reproduces the molecular phenotype of mild HipA overexpression

Based on the lower activity of HipA7 determined above, the next aim was to enhance the kinase activity of HipA7 by increasing its amount from low to high copies per cell. This would allow to test whether increased cellular amount of HipA7 can produce a molecular phenotype similar to the one when HipA is produced in low amounts. To that end, *hipA* and *hipA7* genes were cloned into a low (pNDM220) and a high copy number plasmid (pMG25) and their expression was induced with IPTG. To avoid misinterpretation of HipA7 effects on the phosphoproteome by *hipA* expressed from a chromosomal copy, *hipA7* constructs were transformed into the strain lacking *hipA* on the chromosome (Δ *hipBA*). The resulting phosphoproteomes were compared separately for HipA and HipA7 in two triple-SILAC experiments, in which two strains containing one of the empty plasmids were labeled with Lys0 and mixed into one SILAC channel to serve as a negative control (**Figure 24A, B**). Strains containing HipA or HipA7 produced in low or high copies were labeled with Lys4 or Lys8, respectively. As seen previously, *hipA* induction readily inhibited growth (**Figure 24A**), whereas strains with overproduced HipA7 grew undisturbedly (**Figure 24B**). Despite obvious leakage of the high copy number pMG25 plasmid, which caused longer lag time of growth, the proteome did not differ much from the proteome associated with the low copy number plasmid (data not shown).

Results

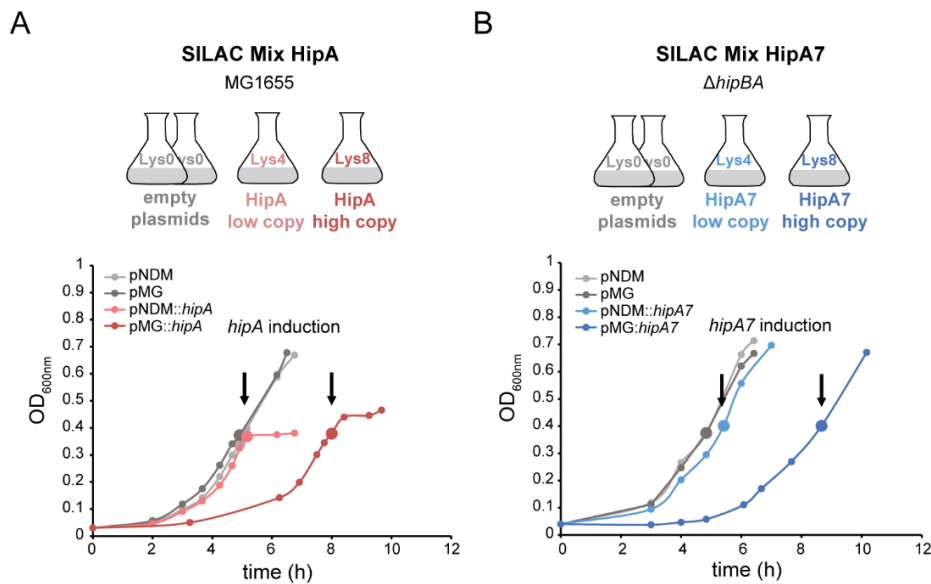


Figure 24. Experimental setup for phosphoproteomic analysis of HipA and HipA7 produced in low and high amounts. (A, B) In two triple-SILAC experiments *hipA* (A) or *hipA7* (B) were expressed from a low copy pNDM (Lys4) and a high copy number plasmid pMG (Lys8), in which *hipA* and *hipA7* are under the control of an IPTG-inducible promoter. Growth curves (below) of *E. coli* MG1655 wt strain (A) or *hipBA* deletion mutant ($\Delta hipBA$) (B) carrying respective plasmids. Strains were grown in SILAC-labeled minimal medium containing either “light” lysine (Lys0), “medium-heavy” lysine (Lys4) or “heavy” lysine (Lys8), and ampicillin for the retention of the plasmids. Expression of *hipA* and *hipA7* was induced at OD_{600nm} = 0.4 with IPTG and samples were collected 95min later.

In this phosphoproteomic experiment, a total of 438 phosphorylation sites were identified on 294 proteins in two biological replicates that correlated well; Pearson correlation coefficient of 0.62 and 0.79 for H/M comparison for HipA and HipA7 experiment, respectively (**Figure 25C, D**). When produced in higher copy numbers, HipA phosphorylated many more proteins, including previously detected targets (**Figure 25A**). Nevertheless, GltX and RplK were phosphorylated by HipA already when present in low amounts. GltX, on the other hand, was phosphorylated at similar levels regardless of HipA amounts present in the cell. In contrast to HipA, HipA7 showed a smaller repertoire of targets even when present in higher copy numbers (**Figure 25B**). However, under those conditions, SeqA and RplK were also phosphorylated, 16-fold and 200-fold when compared to the empty plasmids, and their phosphorylation increased with the increase in HipA7 amount (**Figure 25D**), which implies that HipA7 kinase may resemble the HipA phenotype when produced in higher amounts. In this phosphoproteomic screen, we also detected several phosphorylation sites on HipA7 itself; interestingly, one of them was on Ser²², a residue that was derived from the substitution of Gly²² in HipA (**Figure 25B, D**). Furthermore, a toxin YjjJ was detected phosphorylated on Ser²⁰⁰ when HipA7 was present in high copies (**Figure 25B, D**). Protein YjjJ was only recently reported to act as a novel *E. coli* toxin and based on sequence similarity it was suggested that it could have a kinase activity (275).

Results

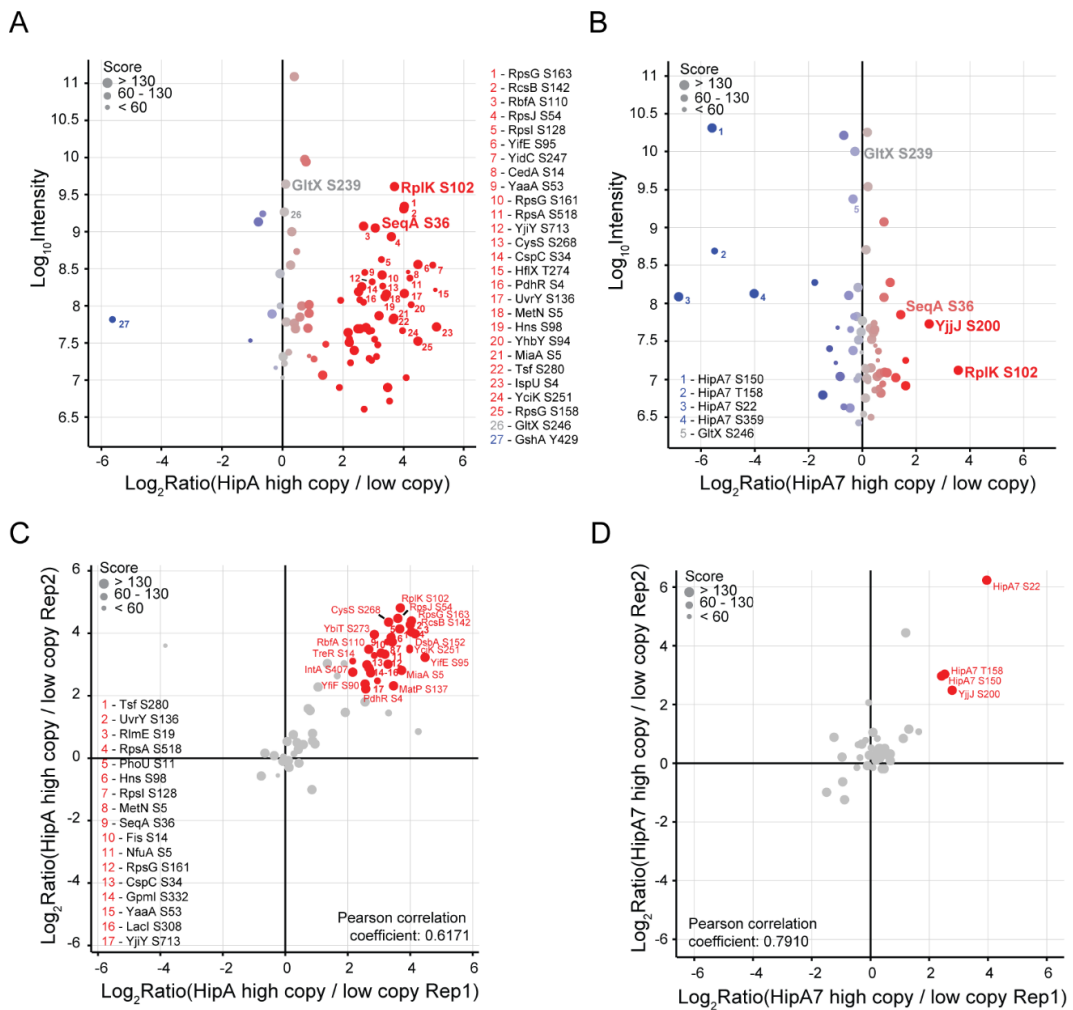


Figure 25. Phosphoproteomic analysis of HipA and HipA7 produced in low and high amounts. (A, B) Distribution of phosphorylation site SILAC ratios of *hipA* (A) and *hipA7* (B) expressed from the high copy number plasmid relative to the expression from the low copy number plasmid. The names of the phosphorylated proteins and the positions of the phosphorylation sites showing at least a 4-fold change in phosphorylation are indicated. Distributions are representative of 2 biological replicates. (C, D) Correlation of phosphorylation site SILAC ratios of *hipA* (C) and *hipA7* (D) expressed from the high copy number plasmid relative to the expression from the low copy number plasmid for 2 biological replicates.

To assess the phosphorylation status of SeqA and RplK by HipA7, an *in vitro* kinase assay was performed. Purified His₆-HipA7 and His₆-GltX, His₆-RplK or SeqA-His₆ were incubated in the presence of ATP, protease-treated and analyzed by LC-MS/MS (**Figure 26A**). Phosphorylation of GltX, but not of SeqA and RplK by HipA7 was detected, suggesting that HipA7 also has lower activity *in vitro* (**Figure 26B**), as observed *in vivo* in phosphoproteomic screens.

Results

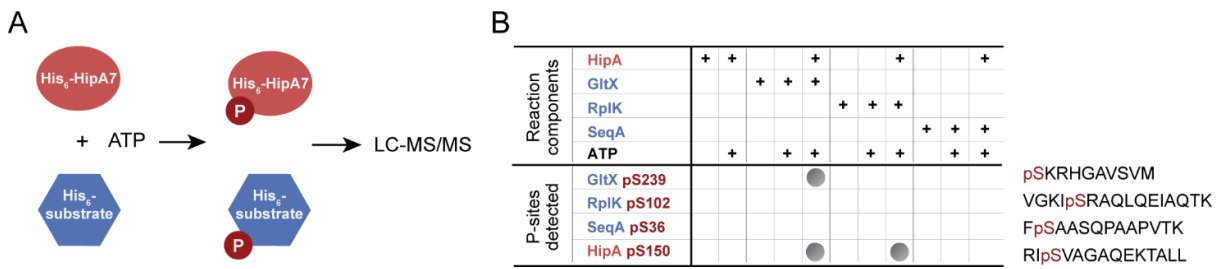


Figure 26. *In vitro* HipA7 kinase assay. (A) Experimental design of the kinase assay of His₆-HipA7 with His₆-GltX, His₆-RplK, and SeqA-His₆. Following the phosphorylation reaction, the samples were protease-treated and analyzed by LC-MS/MS. (B) Identified phosphorylation sites represented as circles were detected in two independent experiments. Note that the SeqA pSer³⁶ site was identified as phosphorylated by HipA7 in one experiment. Sequences of best identified phosphopeptides of particular proteins are shown.

4.2.4 Chromosomally encoded HipA7 phosphorylates GltX and phage shock protein PspA

We next investigated the difference between *hipA* and *hipA7* when expressed from the chromosome, rather than from a plasmid. To that end, we compared phosphoproteomes of stationary phase cultures of *E. coli* K-12 strain MG1655 in which *hipA* was replaced with *hipA7* allele (*hipA7*), to *hipBA* deletion mutant (Δ *hipBA*) and the wild-type strain having a wild-type copy of the *hipA* gene (*wt hipA*) (**Figure 27A**). In total, we identified 2,558 protein groups and 665 phosphorylation sites on 374 proteins in three biological replicates. Phosphorylation site SILAC ratios correlated relatively well; Pearson correlation coefficient of 0.78, 0.40 and 0.56 for *hipA7* and *wt hipA* comparison in three biological replicates, respectively (**Figure 27C**). The poorer correlation compared to previous experiments can be attributed to the lower overall fluctuations in phosphorylation. The results showed that chromosomally encoded HipA7 was able to phosphorylate GltX in the *hipA7* strain without the kinase overproduction (*p*-value < 0.01). In fact, GltX phosphorylation was 12-fold higher in *hipA7* mutant than in the *wt hipA* strain (**Figure 27B**). Accordingly, the *hipA7* strain grew slower and to the lower final OD_{600nm} (**Figure 27A**), which is consistent with the higher phosphorylation of GltX (**Figure 27B**). This apparently higher activity of HipA7 could be explained by the amount of the free active kinase: due to a weakened dimerization of HipA7 molecules, less HipA7 can interact with HipB, leaving more HipA7 available in the cytosol and causing increased *hipBA7* transcription (19). In this proteomic data, we were only able to determine a total cellular abundance of HipA7 relative to HipA. To our surprise, the abundance of HipA7 was slightly lower than the abundance of HipA (**Figure 28C**), which may imply that HipA7 is a less stable protein than HipA, and therefore faster degraded. On the contrary, the abundance of HipB was 4-fold higher in the *hipA7* than in the *wt hipA* strain, whereas GltX abundance did not change significantly (**Figure 28C**). Higher abundance of HipB in the *hipA7* strain could be explained by lower promoter repression due to weakened binding between HipA7 and HipB.

Results

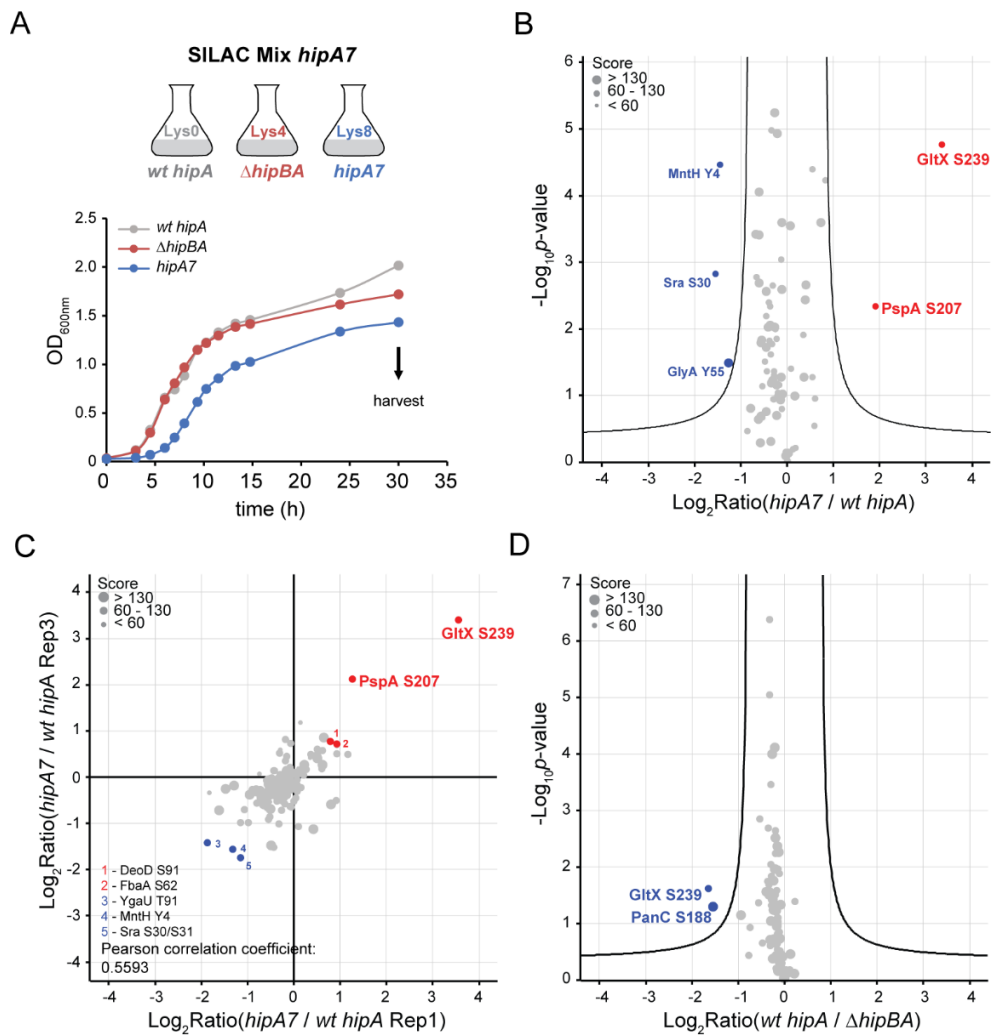


Figure 27. Phosphoproteome comparison of chromosomally expressed *hipA7* and *wt hipA*. (A) Growth curves of three *E. coli* MG1655 strains: strain carrying the wild-type *hipA* gene (*wt hipA*, Lys0), a *hipBA* deletion mutant (Δ *hipBA*, Lys4) and a strain carrying *hipA7* allele (*hipA7*, Lys8) instead of *wt hipA*. Strains were grown in SILAC-labeled minimal medium containing either “light” lysine (Lys0), “medium-heavy” lysine (Lys4) or “heavy” lysine (Lys8), and samples were collected in the late stationary phase after 30h of growth. Growth curves are representative of 3 biological replicates. (B, D) Volcano plot of phosphorylation site SILAC ratios of the *hipA7* relative to the *wt hipA* strain (B) or the *wt hipA* relative to the Δ *hipBA* strain (D) from 3 biological replicates. The black curve indicates significance level with the *p*-value of 0.01 and a minimal fold change 50 of 1. The names of phosphorylated proteins and the positions of the phosphorylation sites significantly increased (red) or decreased (blue) in phosphorylation are indicated. (C) Correlation of phosphorylation site SILAC ratios of the *hipA7* relative to the *wt hipA* strain for biological replicate 1 and 3. The names of the phosphorylated proteins and the positions of the phosphorylation sites significantly increased (red) or decreased (blue) in phosphorylation (*p*-value < 0.01) are indicated.

Besides GltX, phosphorylation of the phage shock protein PspA on Ser²⁰⁷ increased in the *hipA7* strain (*p*-value < 0.01) (Figure 27B), whereas PspA abundance was significantly lower (*p*-value < 0.001) (Figure 28A). This raises the possibility that the abundance of PspA could be reduced by phosphorylation in the presence of *hipA7* on the chromosome. PspA is the negative regulator of a transcription factor PspF, which is associated with the expression of the phage shock protein (*psp*) operon in response to diverse stresses (276) and is known to be implicated in persistence (113); however, the abundances of other members of the *psp* operon were not changed in this data set.

Results

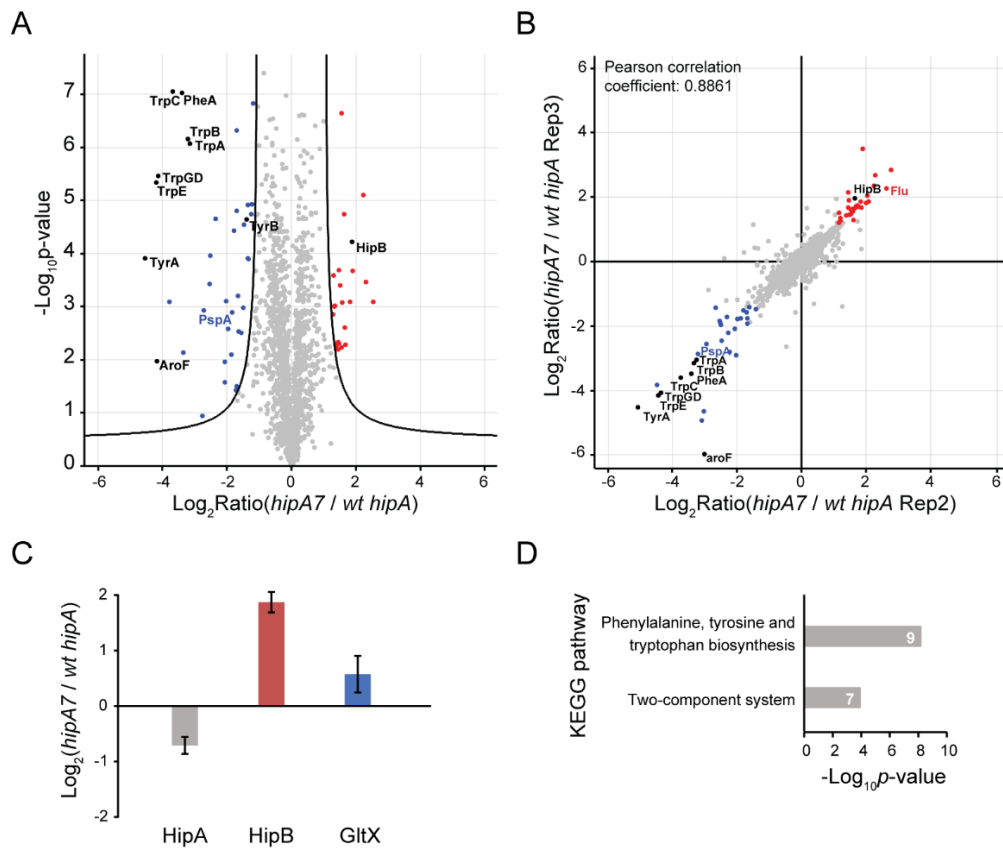


Figure 28. Proteome comparison of chromosomally expressed *hipA7* and *wt hipA*. (A) Volcano plot of protein SILAC ratios of *hipA7* relative to *wt hipA* strain from 3 biological replicates. The black curve indicates significance level with the p -value of 0.001 and the minimal fold change S_0 of 1. Proteins significantly increased in abundance are indicated in red and proteins decreased in abundance in blue. (B) Correlation of the protein SILAC ratios of the *hipA7* relative to the *wt hipA* strain for biological replicate 2 and 3. Significantly regulated proteins (p -value < 0.001) in both replicates are indicated in red (increased in abundance) and blue (decreased in abundance). HipB and proteins of aromatic amino acids biosynthesis pathway are indicated in black (A, B). (C) Relative abundances of HipA, HipB and GltX proteins in the *hipA7* strain in comparison to the *wt hipA* strain represented as means of protein SILAC ratios in log₂ scale \pm SD from 3 biological replicates. Increased abundance of HipB in the *hipA7* strain is significant (significance B test, p -value < 0.001). (D) KEGG pathway analysis of proteins showing a decrease in abundance in the *hipA7* strain compared to the *wt hipA* strain (significance B test, p -value < 0.001) in 3 biological replicates against the background of all identified proteins. The number of enriched proteins for each category is indicated within the bars.

A direct comparison of phosphoproteomes of the *wt hipA* strain and the $\Delta hipBA$ mutant showed that GltX phosphorylation decreased significantly when *hipA* was deleted, which confirmed GltX phosphorylation by endogenous HipA (**Figure 27D**). However, in these conditions, intensity of the GltX phosphopeptide was very low and it is possible that it would be hardly detectable in the absence of the *hipA7* strain in the third SILAC channel. This supports the need for using plasmid-encoded HipA to study its substrates. Except for GltX, a phosphorylation of the pantothenate synthetase (PanC) decreased (p -value < 0.01) on the Ser¹⁸⁸ residue upon *hipA* deletion (**Figure 27D**). Interestingly, Ser¹⁸⁸ residue of PanC is situated in the ATP-binding region of the enzyme (277).

Although the phosphorylation status of *hipA7* strain did not differ substantially from *wt hipA* cells, a direct comparison of their proteome showed a major decrease in the abundance of proteins involved in the biosynthesis of aromatic amino acids (significance B test, p -value < 0.001) (**Figure 28A, B, D**).

This was observed consistently in all three biological replicates that showed good reproducibility (Pearson correlation coefficient 0.80, 0.77, 0.89 for *hipA7* to *wt hipA* comparison) (**Figure 28B**). Those proteins included the products of the entire tryptophan operon, aromatic-amino-acid aminotransferase (TyrB), chorismate mutase and prephenate dehydratases P-protein (PheA) and T-protein (TyrA). All of them are located downstream of the chorismate in the shikimate pathway and they convert chorismate into tyrosine, phenylalanine and tryptophane through several reactions. This observation provides a potential link between HipA7 activity and biosynthesis of aromatic amino acids.

4.2.5 A comprehensive map of the *E. coli* phosphoproteome

Finally, this study produced a substantial, high-quality phosphoproteome data set of *E. coli* K-12 containing a total of 2,727 identified proteins with an estimated false discovery rate (FDR) of around 2-3% per experiment and 1,183 phosphorylation sites (FDR < 1%) on 632 phosphoproteins out of which 919 phosphorylation sites were considered to be confidently localized (localization probability > 0.75) (**Figure 29A**). The proportion of Ser, Thr and Tyr phosphorylation was 64%, 30% and 6%, respectively, which is in agreement with previous studies (**Figure 29B**) (225, 226).

This large phosphoproteome data set enabled to search for characteristic linear phosphorylation motifs of bacterial Ser/Thr kinases, which are still poorly investigated as compared to eukaryotic kinases. The Motif-X software (270) was used to search for overrepresented patterns from 15 amino acid long sequences that contained 7 amino acids on both sides of the phosphorylation site. The analysis was performed individually by defining the central character as S, T or Y against reference *E. coli* K-12 MG1655 proteome using stringent criteria. Motif analysis resulted in three high-confidence sequence patterns containing phosphorylation sites; two for pSer, one for pThr and none for pTyr (**Figure 29C**). All three linear sequence motifs contained a lysine residue and, in two of them, lysine was positioned next to the phosphorylation site at position -1 relative to the pSer or pThr, which is in agreement with a previous study (225).

Furthermore, to determine which cellular functions are affected by phosphorylation, a functional enrichment of all 632 identified phosphoproteins was performed using DAVID software and KEGG pathway annotation. Functional analysis revealed that Ser/Thr/Tyr phosphorylation was spread across many essential cellular processes, such as translation, nucleotide metabolism, glycolysis and gluconeogenesis, pentose phosphate pathway, aminoacyl-tRNA biosynthesis and others (**Figure 29D**). Taken together, this data set can serve as a valuable resource of protein phosphorylation events in *E. coli*.

Results

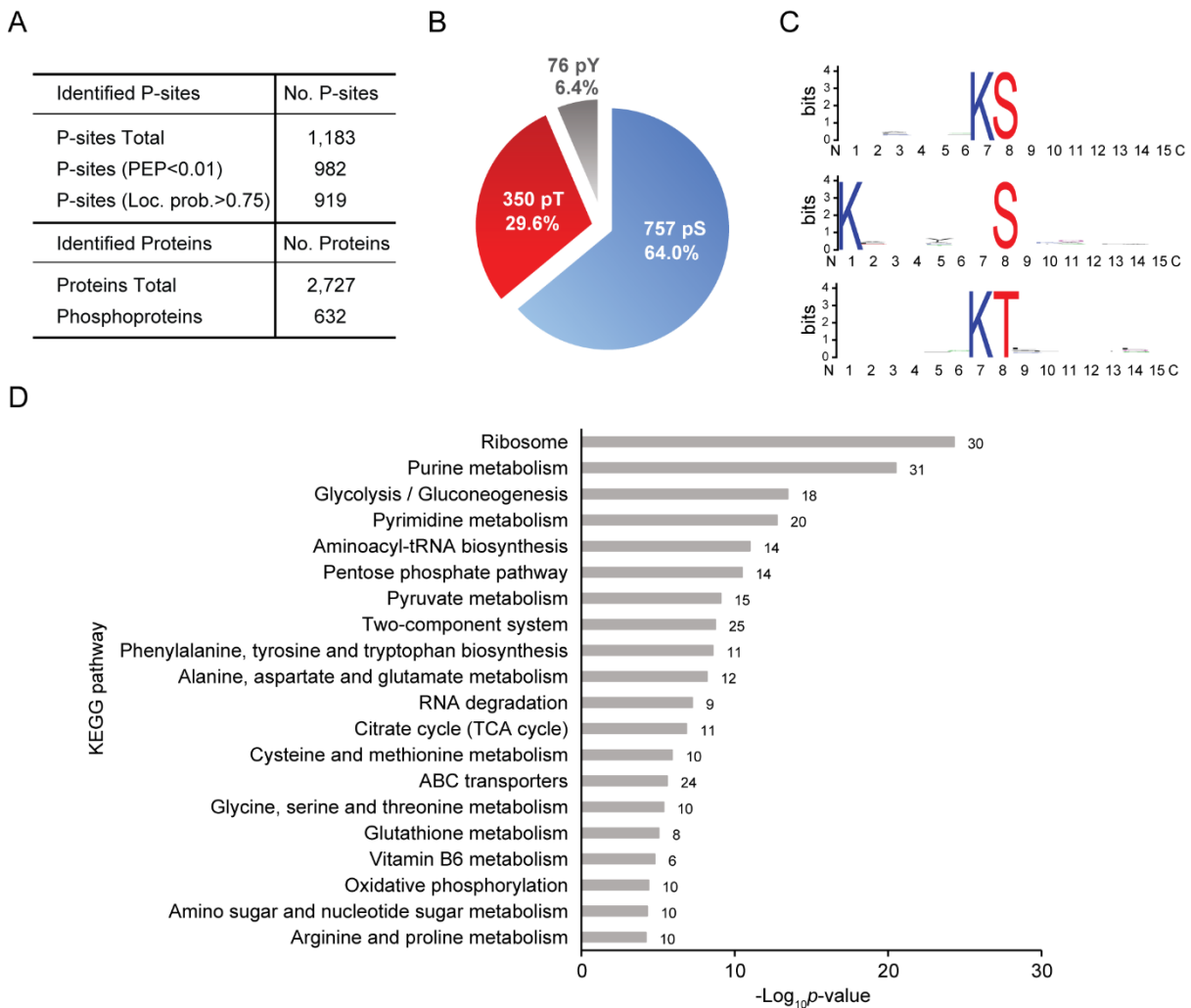


Figure 29. Functional analysis of the *E. coli* phosphoproteome obtained in this thesis. (A) The overview of the number of all identified phosphorylation sites and proteins. PEP stands for posterior error probability and Loc. prob. for localization probability (value between 0 and 1). (B) The distribution of all identified Ser, Thr and Tyr phosphorylation sites. (C) Kinase linear sequence motif enrichment analysis of all identified Ser and Thr phosphorylation sites with localization probability > 0.75 performed with Motif-X software (significance threshold 0.00000001). (D) KEGG pathway enrichment analysis of all identified phosphoproteins performed with DAVID software against the background of *E. coli* K-12 MG1655 genome. The number of enriched phosphoproteins in each category is indicated outside the bars.

4.3 Time-resolved analysis of newly synthesized proteins during toxin-induced persistence and antitoxin-mediated resuscitation

4.3.1 Establishment of the method for selective labeling of resuscitating persister cells applied to the *hipBA* TA module

In addition to investigating cellular phosphorylation signature associated with the kinase HipA, the third aim of this thesis was to study dynamics of newly synthesized proteins in the conditions of persistence and resuscitation. The model based on overexpressed *hipA* and *hipB* was used to establish and optimize a general method for such temporal analysis of protein turnover during toxin-induced persistence and antitoxin-mediated resuscitation, which can be further applied to other TA systems. Ultimately, the same approach was applied to another TA system, *relBE*, which uses different mechanism than *hipBA* to inhibit translation and induce persistence.

First, the method was established for the resuscitation phase, in which cells wake-up from persistence and re-initiate the growth. In this case, cell division enables higher incorporation of the label, allowing for higher sensitivity of the measurements. Furthermore, whereas much more is known about processes that lead to persister formation, the mechanisms that regulate the switch between persistence and resuscitation are not well understood. For all these reasons, the first aim was to implement the method and identify proteins which are rapidly produced during the early phase of the resuscitation and additionally, to characterize the resuscitation process from a system-wide perspective. A time-resolved analysis of protein turnover was performed by employing a pulse-labeling SILAC-based approach known as “dynamic SILAC” (247) in combination with high-resolution MS-based proteomics. For that, low-frequency persister cells were enriched by inducing *hipA* gene from the pBAD33 plasmid in the *E. coli* K-12 MG1655 *lysA* deficient strain ($\Delta lysA$), which is auxotrophic for lysine (**Figure 30**). The strain was therefore grown in the minimal media supplemented with “light” lysine (Lys0). The expression of *hipA* inhibited growth, and 3 hours post induction, the culture was treated with a high dose of ampicillin to lyse non-persister cells and to enrich for low abundant persister cells. Additionally, the expression of *hipA* was repressed during ampicillin treatment to ensure the steady state conditions independent of the constant HipA overproduction. To trigger the regrowth from dormancy, ampicillin, Lys0 and cellular extract of lysed cells were removed by filtration. Intact cells were then transferred into a fresh medium and the expression of the antitoxin-encoding *hipB* gene was induced from pEG220 plasmid by IPTG. At the same time, a pulse of “heavy” lysine (Lys8) was added to metabolically label resuscitating cells and culture aliquots were harvested in intermittent intervals. In this experimental setup, only newly synthesized proteins from living cells (resuscitating persister cells) incorporate Lys8 enabling time-resolved quantification of the abundances during resuscitation simultaneously for each individual protein.

Results

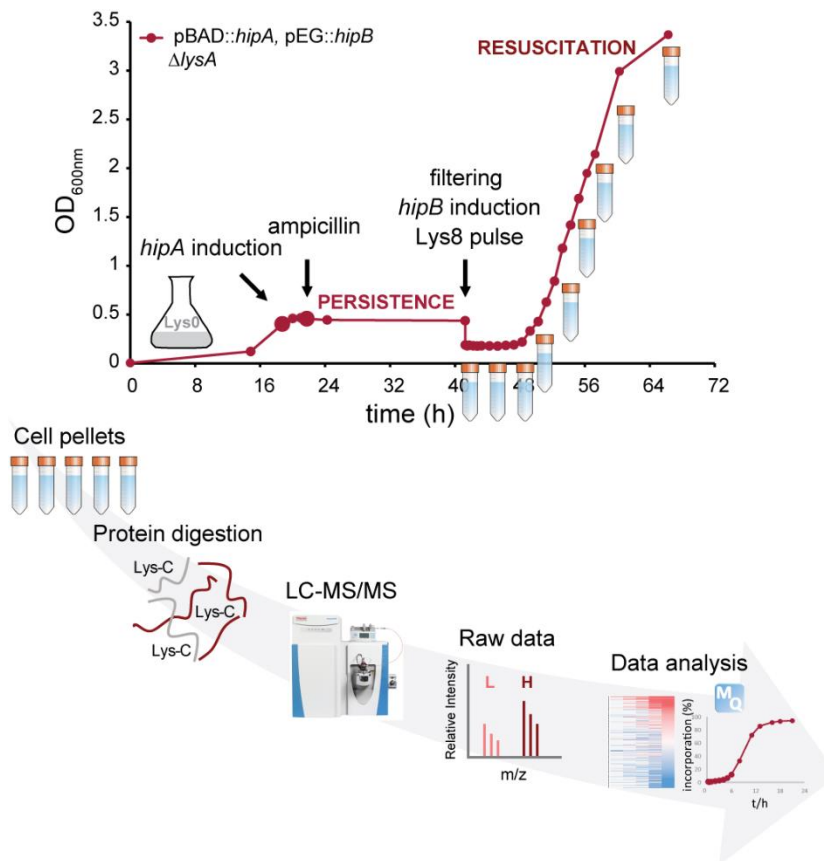


Figure 30. Experimental workflow for measurement of protein turnover during resuscitation from persistence using dynamic SILAC approach. *E. coli* K-12 MG1655 $\Delta lysA$ strain containing pBAD::*hipA* plasmid, in which *hipA* expression is under the control of an arabinose-inducible promoter, and the pEG::*hipB* plasmid, in which *hipB* is under the control of an IPTG-inducible promoter, was grown in the minimal medium containing “light” lysine (Lys0) and chloramphenicol and kanamycin for the retention of the plasmids. Expression of *hipA* was induced at $OD_{600nm} = 0.4$ with arabinose for 3 hours, followed by ampicillin treatment (100 $\mu g/ml$) for next 20 hours and addition of glucose to repress further *hipA* expression. To trigger resuscitation, culture was filtered and cells were transferred into a fresh medium containing IPTG to induce expression of *hipB* and “heavy” lysine (Lys8) for pulse-labeling. Cells were harvested before filtering (0h) and in 19 different time points after. Proteins were isolated, digested with endoproteinase Lys-C and samples were measured by LC-MS/MS in 20 runs. Acquired raw data was processed with MaxQuant and analyzed in Excel. The growth curve is representative of 3 biological replicates.

Intermittent time points were chosen to cover both, a very early phase of resuscitation and different stages during the entire regrowth until stationary phase. In total, cells were harvested in 20 time points during the course of resuscitation in three biological replicates: 0h, 10min, 15min, 30min, 45min, 1h, 1.5h, 2h, 2.5h, 3h, 4h, 5h, 6h, 8h, 10h, 12h, 14h, 16h, 20h and 30h after medium exchange and *hipB* induction. Isolated proteins were digested with protease Lys-C and subjected to the LC-MS/MS measurement. The inspection of raw MS data revealed an increasing incorporation of “heavy” lysine into peptides over time and decreasing abundance of Lys0-labeled (L) peptides (**Figure 31**). Lys8-labeled (H) peptides originate solely from newly synthesized proteins, while peptides that contain Lys0 come from the pre-existing protein pool of persister cells and from dead, intact cells, which were not removed by filtering.

Results

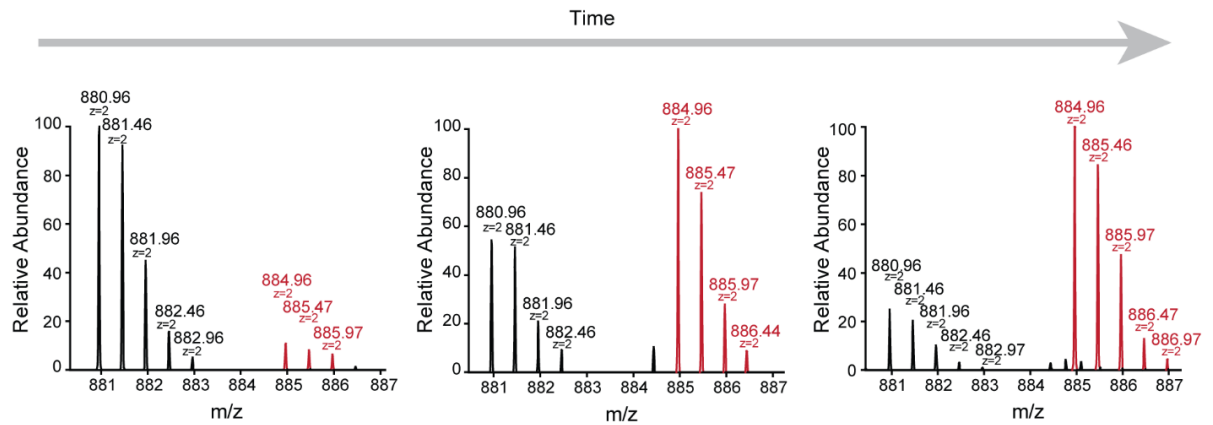


Figure 31. Incorporation of “heavy” lysine (Lys8) over time. MS spectra of the GroL peptide (AIAQVGTIVSANSDETGVGK) show increasing relative abundance of Lys8-labeled (H) peptide (red) and decreasing relative abundance of Lys0-labeled (L) peptide (black) over time.

Incorporation of Lys8 was inspected at the protein level for the whole proteome and revealed that the median of incorporation follows the shape of the growth curve and saturates at $94.6 \pm 1.3\%$ (**Figure 32A**). Already 30 minutes after induction of the resuscitation, more than one hundred proteins were quantified. Although the median of Lys8 incorporation was low during first 4 hours of resuscitation ($< 8\%$), the outliers indicated that a set of proteins incorporated the label more efficiently in each time point. The experiments were performed in three biological replicates that showed good reproducibility of measured protein SILAC ratios across all time points (Pearson correlation coefficient from 0.75 in early time points up to 0.96 in late time points) (**Figure 32B**). On average, $1,880 \pm 40$ protein groups (hereafter referred to as proteins) were identified across time points with an estimated FDR of 1.3% (**Figure 32C**). For a protein to be quantified, two occurrences of the protein H/L ratio were required per time point. The number of quantified proteins increased gradually over time, peaking at 8h time point, in which $1,142 \pm 136$ proteins had H/L ratio assigned, after which the number of quantified proteins decreased. The decrease in the number of quantified proteins can be explained by the fact that proteins with higher turnover rate replace their entire pool composed of pre-existing Lys0-labeled proteins with newly synthesized Lys8-labeled proteins. This results in a low MS signal intensity of Lys0-labeled peptides, which cannot be measured, and therefore, the H/L ratio cannot be determined. Indeed, the intensity of Lys8-labeled proteins increased over time and exceeded the intensity of Lys0-labeled proteins at 8h time point (**Figure 32D**). The almost complete labeling efficiency was already confirmed by the incorporation of Lys8 calculated from protein H/L ratios (**Figure 32A**).

Results

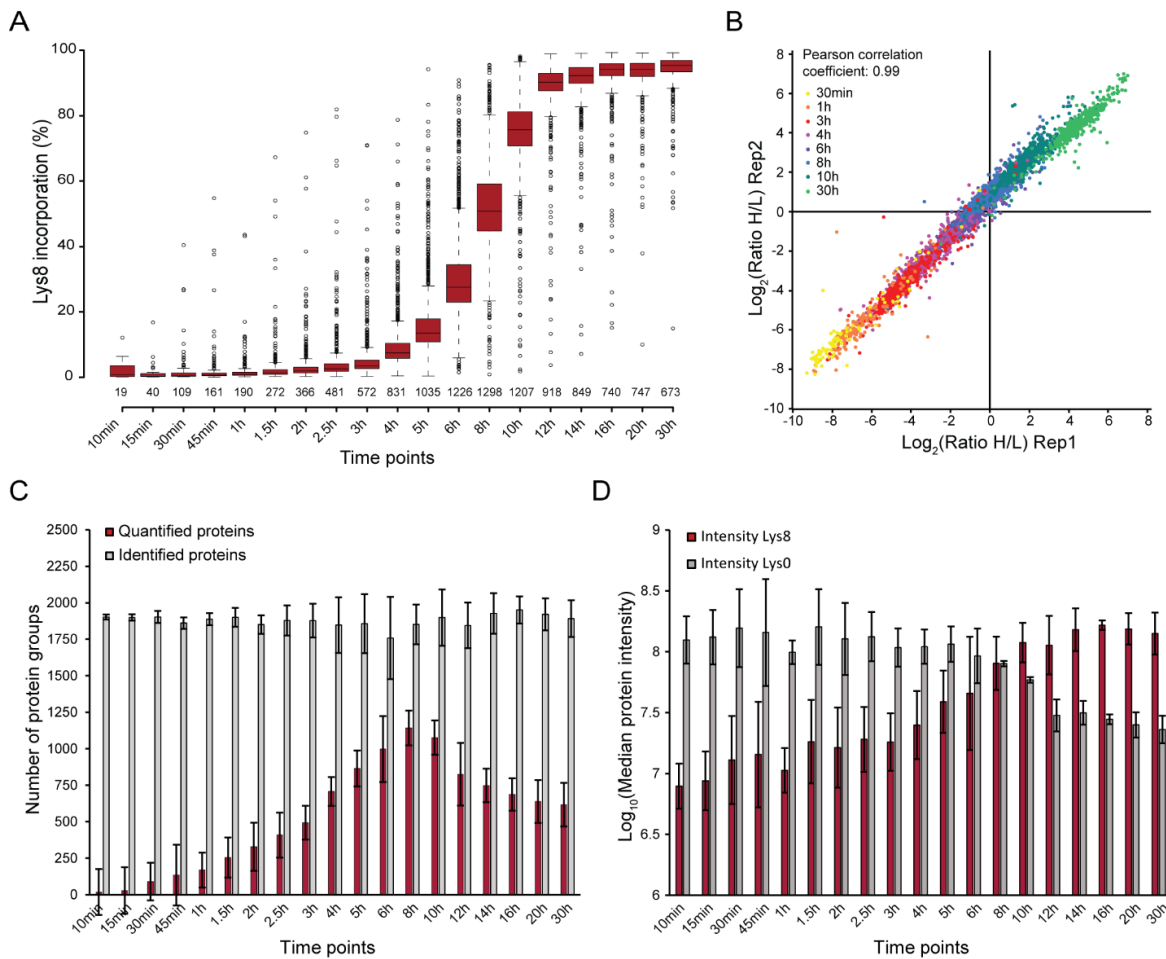


Figure 32. Quality control of dynamic SILAC measurements during resuscitation. (A) Incorporation of Lys8 (H) into proteins across time points calculated from the protein H/L ratio as a proportion of the heavy-labeled protein. The box plots are representative of 3 biological replicates. (B) Correlation of protein H/L ratios for biological replicate 1 and 2 across different time points marked in different colors. (C) Number of protein groups identified and quantified by protein H/L ratio in each time point. (D) Median of protein intensities derived from Lys0- and Lys8-labeled peptides across time points. Data in (C, D) are means \pm SD from 3 biological replicates.

4.3.2 High turnover proteins during early phase of resuscitation are involved in amino acid biosynthesis and translation

The proportion of persister cells in the batch culture is substoichiometric (up to 10%), even when they are enriched by toxin-overexpression (79). Therefore, the MS signal intensity corresponding to proteins of resuscitating persister cells in the early time points was low and masked by the high signal from proteins of dead cells. For this reason, the sensitivity of the measurement was lower and the resulting data contained many missing values (H/L ratios) across time points and replicate measurements. Therefore, such complex data set was simplified for further functional analysis. Because the focus of this study was the early phase of the resuscitation, further data analysis was performed using time points from the period of first 8 hours of resuscitation, in which cells doubled in the optical density. First, measurements from three biological replicates were combined as a union, in which protein H/L ratios from each time point (10min – 8h) were kept if measured only in one

Results

replicate (Class I), a mean was calculated if measured in two (Class II) or in three replicates (Class III). Second, time points were pooled into 5 time bins, namely: 10 – 30min, 45 – 90min, 2 – 3h, 4 – 5h and 6 – 8h, by calculating the median protein H/L ratio per time bin. Proteins were then ranked based on their H/L ratio within each time bin and displayed in a heatmap-like representation (**Figure 33**). According to the KEGG pathway enrichment analysis, proteins exhibiting high turnover (top 25%) during the entire 8-hour course of resuscitation (p -value < 0.05) are involved mainly in the biosynthesis and transport of amino acids, such as arginine and cysteine (through assimilatory sulfate reduction), and in alanine, aspartate and glutamate metabolism (**Figure 33**). Conversely, proteins with the low turnover (bottom 25%) are constituents of the citric acid cycle and the amino sugar and nucleotide sugar metabolism. Altogether, this implies that bacteria are producing proteins necessary to synthesize metabolites that are absent in the medium to accelerate the regrowth from dormancy. At the same time, proteins involved in catabolic pathways that consume metabolites to release energy are turning over at slower rate among all quantified proteins. In addition, proteins exhibiting moderate turnover mainly participate in translation.

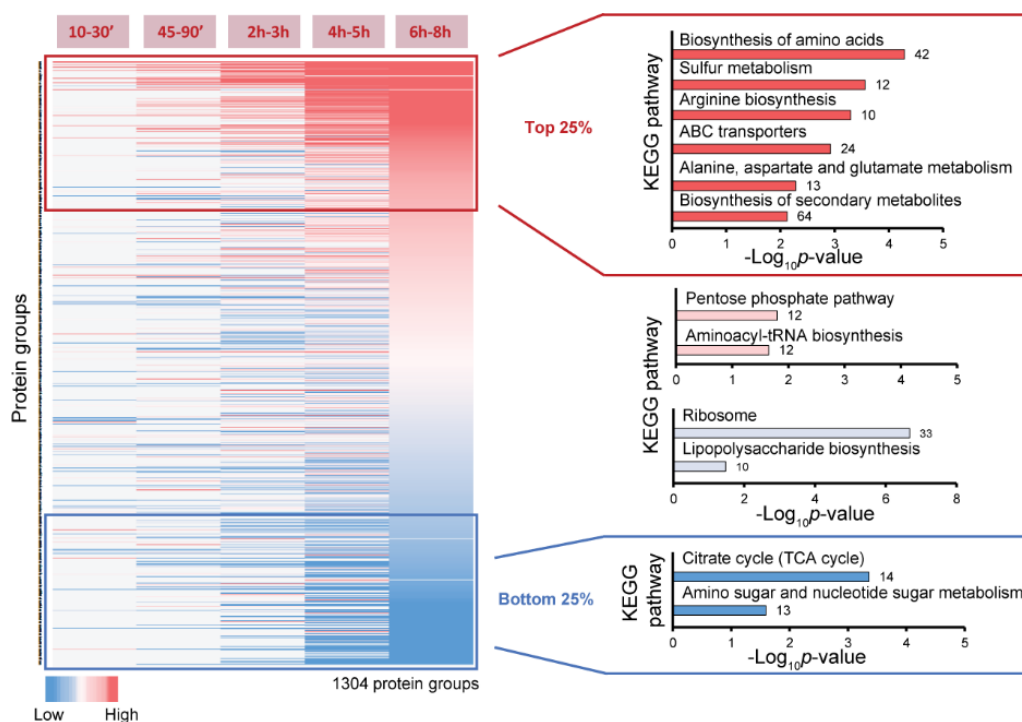


Figure 33. Heatmap of all quantified proteins across time bins during resuscitation. Proteins were combined in time bins and color-coded based on their H/L ratio ranking within each time bin. Missing values are colored in grey. KEGG enrichment analysis of proteins in the 25th percentile (Top 25%, red), 25th – 50th percentile (pink), 50th – 75th percentile (light blue) and 75th percentile (Bottom 25%, blue) across all time bins was performed with DAVID software against the background of all 1,304 quantified proteins. The number of enriched proteins in each category is indicated outside the bars.

The main aim of this study was to identify proteins that are potentially regulating the switch from persistence to resuscitation. Thus, the first time bin (10 – 30min) was analyzed in more detail. GO enrichment analysis revealed that among 153 quantified proteins, 43 of them take part in translation (p -value of $2.75E-27$) out of which 42 are ribosomal proteins and one is the stationary-phase-induced

Results

ribosome-associated protein Sra (**Figure 34A**). Except for the entire translation machinery, metabolism was induced through glycolysis pathway, as a response to freshly provided glucose that can be converted to energy needed for the regrowth. In addition, stress response and protein folding processes were also significantly enriched. The top10 high turnover proteins in the first 30 minutes incorporated the label from 7% to 70% (**Figure 34B**). Among them was the antitoxin HipB, expression of which was induced from the plasmid and therefore served as an internal control. The others participate in various processes, such as the ribonucleotide monophosphatase NagD that dephosphorylates a wide range of (deoxy)ribonucleoside phosphates, or the S-formylglutathione hydrolase FrmB, which converts S-formylglutathione into formate and glutathione to detoxify formaldehyde that can otherwise chemically modify DNA and proteins (278). Moreover, two proteins that are known to play a role in persistence were among high turnover proteins on the onset of resuscitation, namely, the RNA polymerase sigma factor RpoS, which regulates expression of many genes during stress, and the ATP-dependent Clp protease subunit ClpA that directs the ClpAP protease to specific substrates for their degradation.

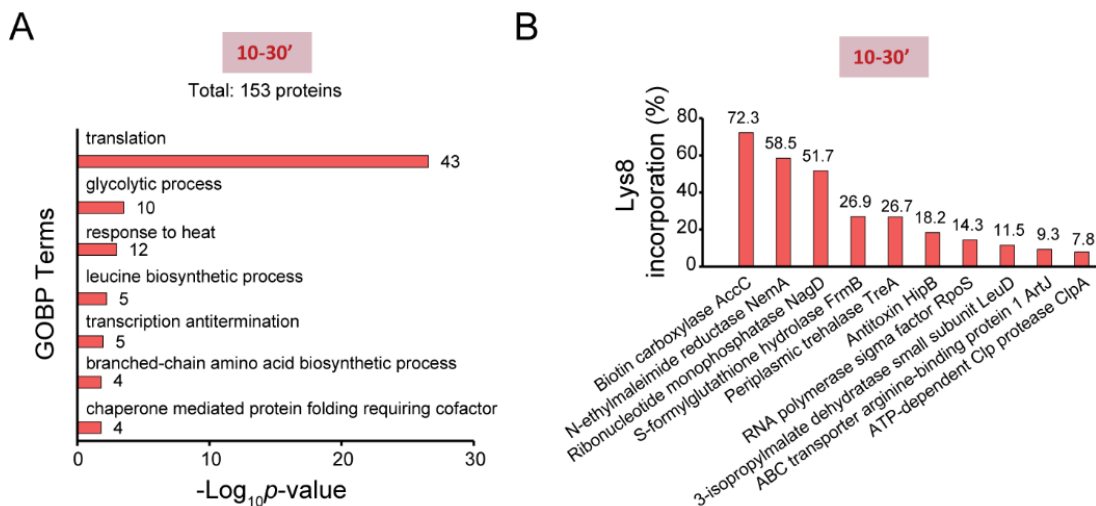


Figure 34. Functional analysis of proteins quantified at the onset of resuscitation. (A) GO analysis of proteins with highest incorporation of Lys8 in the first time bin (10 to 30 min of resuscitation) performed against the background of all identified proteins. The number of enriched proteins in each category of GO biological processes is indicated outside the bars. (B) Top 10 proteins with highest incorporation of Lys8 at the onset of resuscitation (10 - 30 min of resuscitation).

To gain deeper insight into individual proteins that are produced rapidly during specific stages of the early resuscitation, top 25 proteins with highest Lys8 incorporation in each time bin were presented in a map (**Figure 35**). As expected, the antitoxin HipB exhibited high abundance across all time points, as it was produced from the plasmid and served as an internal positive control for the accuracy and sensitivity of measured protein SILAC ratios. Many other proteins with interesting functions showed high turnover, such as the antitoxin DinJ, the multiple stress resistance protein BhsA, cold shock protein CspA, the N-ethylmaleimide reductase Nema that degrades toxic compounds, which can modify proteins and lead to growth inhibition, and many others.

Results

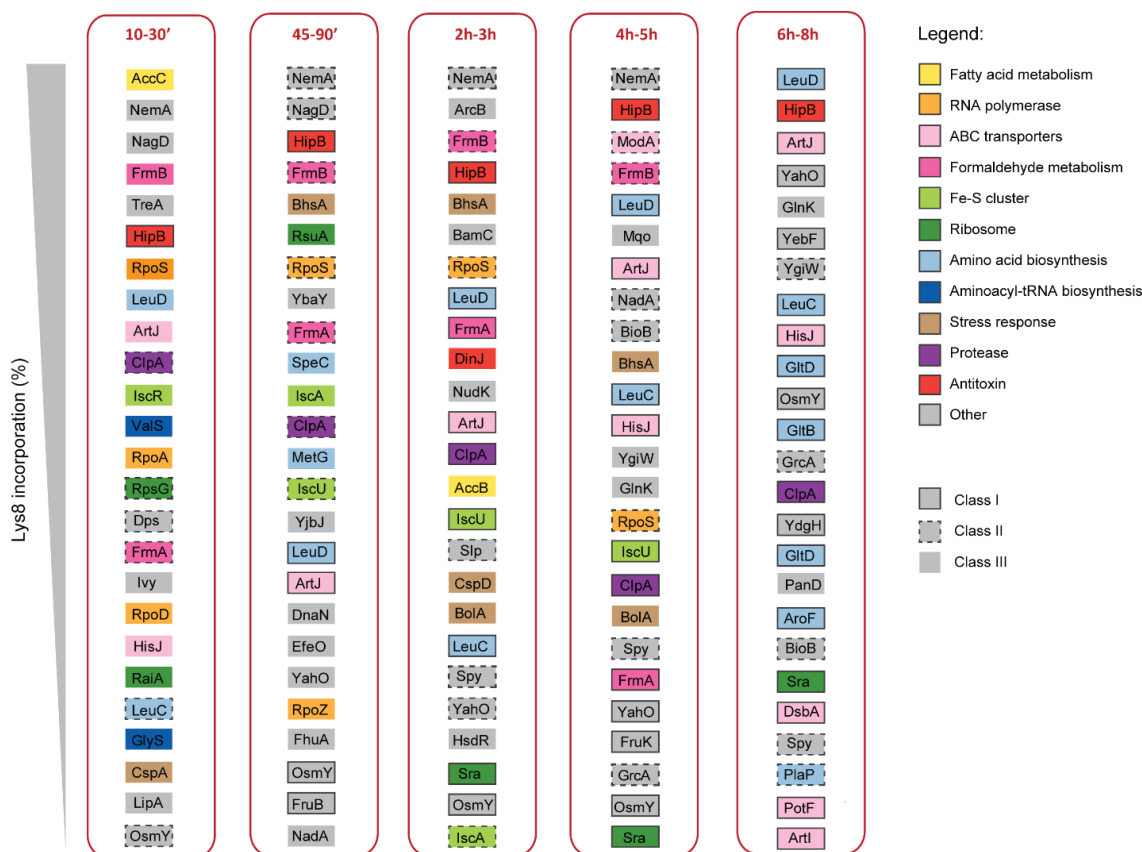


Figure 35. Map of top 25 quantified proteins with highest incorporation of Ly8 in each time bin during resuscitation. Different functional descriptions are color coded according to the legend. Boxes around individual proteins are coded based on the style of the line and indicate the confidence of protein quantification per time bin according to the following classification system: protein is quantified in > 67% of the time points in 3 replicates (Class I), protein is quantified in < 67% and > 33% of the time points in 3 replicates (Class II), protein is quantified in < 33% of the time points in 3 replicates (Class III).

4.3.3 Selective labeling of persister cells during HipA-induced persistence

After establishing the time-resolved labeling of growing cells upon the exit from persistence (see 4.3.1), the possibility of labeling non-dividing, dormant persisters was examined next. For that, the dynamic SILAC pulse-labeling with Lys8 was applied to the culture containing persister cells enriched by *hipA* expression and ampicillin treatment (**Figure 36A**). To ensure that measured protein turnover does not originate from growing cells that were not killed by the antibiotic, the pulse of Lys8 was added around 20 hours of ampicillin treatment, which was sufficient to eradicate sensitive cell and achieve a steady state of persistence (data not shown). In this experiment, a strain with the wild-type *E. coli* K-12 MG1655 background was used instead of $\Delta lysA$ to avoid any changes in the nutrient availability and stress imposed by the filtering process. After the introduction of the label, culture aliquots were harvested in intermittent intervals ranging from 10 min to 22 hours after the pulse. In this experimental setup, only newly synthesized proteins from living cells (persister cells) incorporate Lys8, which enables quantification of protein abundances in temporal fashion.

Results

In total, the experiment was performed in three biological replicates and 17 time points were chosen during the course of persistence: 0h, 10min, 15min, 30min, 45min, 1h, 1.5h, 2h, 2.5h, 3h, 4h, 5h, 6h, 8h, 10h, 12h and 22h after the pulse. Because cell growth was inhibited, incorporation of Lys8 was expectedly low and reached $4.0 \pm 0.7\%$ 22 hours after the start of labeling (**Figure 36B**). However, a number of outliers showed high Lys8 incorporation, which is indicative of the high turnover of these proteins during persistence. The number of identified proteins across time points was consistent, with $1,864 \pm 46$ proteins identified in each point with an estimated FDR of 1.1% (**Figure 36C**). Surprisingly, 470 ± 64 proteins were quantified in three replicates after 22 hours of labeling, even in the high background of Lys0-labeled proteins originating from dead cells. Despite the low incorporation of Lys8, this experiment showed that it is possible to label persister cells enriched by the overproduction of the toxin and that the mass spectrometry is sensitive enough to measure and resolve the fine differences between pre-existing and newly synthesized proteins.

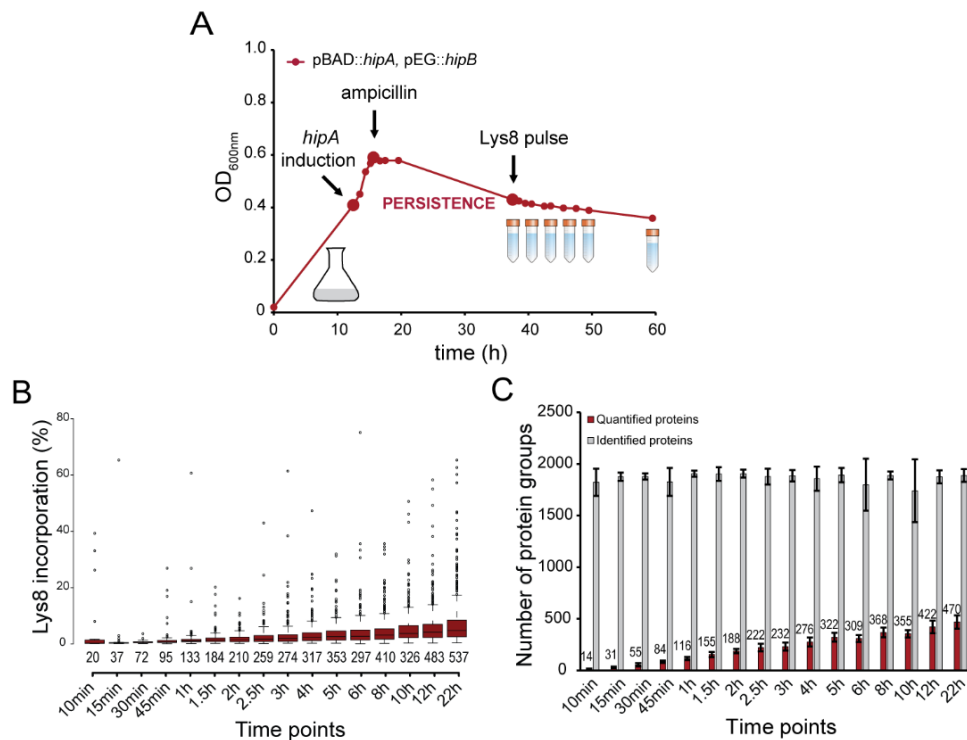


Figure 36. Measurement of newly synthesized proteins during the steady state of HipA-induced persistence using dynamic SILAC. (A) Experimental workflow. *E. coli* K-12 MG1655 strain containing pBAD::*hipA* plasmid, in which *hipA* expression is under the control of an arabinose-inducible promoter, and the pEG::*hipB* plasmid, in which *hipB* is under the control of an IPTG-inducible promoter was grown in minimal medium containing and chloramphenicol and kanamycin for the retention of the plasmids. Expression of *hipA* was induced at $OD_{600nm} = 0.4$ with arabinose for 3 hours, followed by the ampicillin treatment (100 $\mu\text{g}/\text{ml}$) and addition of glucose to repress further *hipA* expression. After around 20 hours, a pulse of “heavy” lysine (Lys8) was added to label newly synthesized proteins. Cells were harvested before the pulse (0h) and in 16 different time points after. The growth curve is representative of 3 biological replicates. (B) Incorporation of Lys8 into proteins across time points calculated from protein H/L ratios. The box plots are representative of 3 biological replicates. (C) Number of protein groups identified and quantified by H/L ratio in each time point. Data are means \pm SD from 3 biological replicates.

4.3.4 Proteins involved in translation and transcription are newly synthesized during HipA-induced persistence

For further analysis, the data from three replicates was combined as described in 4.3.2 and 3.2.17. The time points during the entire 22-hour course of persistence were split into 6 time bins, namely: 10 – 30min, 45 – 90min, 2 – 3h, 4 – 6h, 8 – 12h and 22h. Proteins from each time bin were then ranked based on their H/L ratios and displayed in the heatmap (**Figure 37**). GO analysis of all 602 quantified proteins showed that proteins involved in translation were strongly enriched (p -value of $1.08E-21$). In addition, proteins that play a role in the glycolysis, protein folding, response to heat, regulation of translation and transcription (anti)termination were also significantly enriched (p -value < 0.01). Likewise, ribosomal proteins were overrepresented (p -value of $2.25E-06$) among top 25% proteins that exhibited high turnover across all time points. In contrast, proteins with low turnover (bottom 25%) participate mainly in the metabolic pathways such as biosynthesis of amino acids and other secondary metabolites. This observation is opposite to results from the resuscitation phase, where proteins that synthesize amino acids exhibited high turnover. However, it cannot be excluded that this is a result of the medium exchange, which provided fresh nutrients to the resuscitating cells. Nevertheless, this study showed that dormant persister cells perform active translation and produce proteins that are needed for essential cellular processes. In fact, out of 299 essential genes in *E. coli*, protein products of 144 of those genes incorporated the label during persistence.

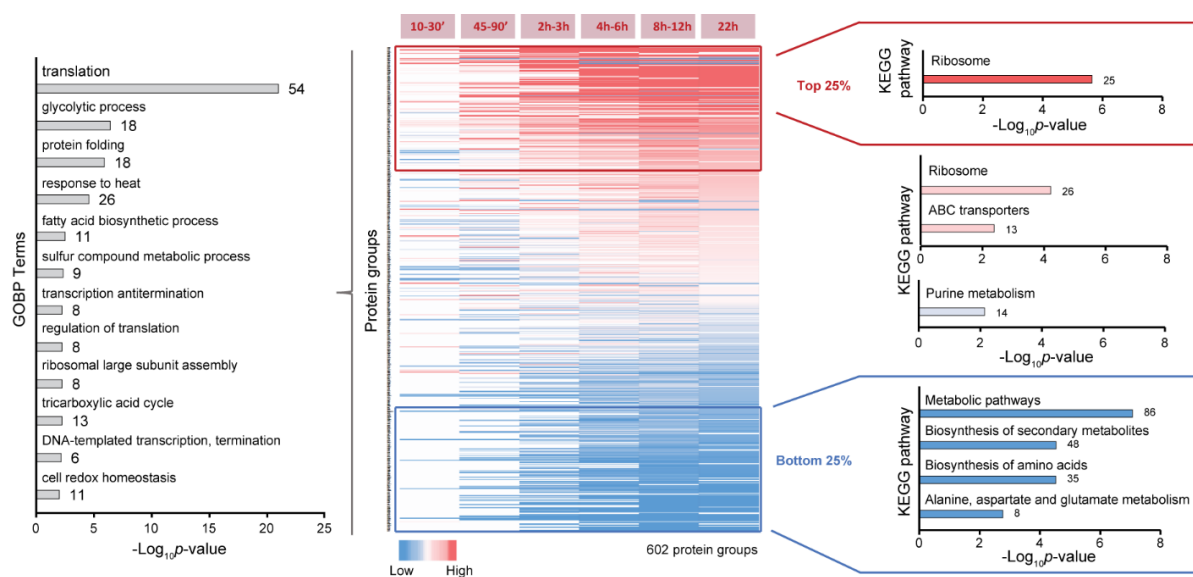


Figure 37. Heatmap of all quantified proteins across time bins during the steady state of HipA-induced persistence. Proteins were combined in time bins and color-coded based on their H/L ratio ranking within each time bin. Missing values are colored in grey. GO biological process (GOBP) analysis (left) of all quantified 602 proteins was performed with DAVID software against the background of all 2,223 identified proteins. KEGG pathway enrichment analysis (right) of proteins in the 25th percentile (Top 25%, red), 25th – 50th percentile (pink), 50th – 75th percentile (light blue) and 75th percentile (Bottom 25%, blue) across all time bins was performed analogously against the background of all 602 quantified proteins. The number of enriched proteins in each category is indicated outside the bars.

Results

To gain a deeper insight into the functions of individual proteins that exhibited high turnover during persistence, specific proteins were manually selected from the set of top 25% proteins with highest H/L ratio and their profiles were shown in a time-resolved representation of Lys8 incorporation (**Figure 38**). Interestingly, all *E. coli* proteases necessary for protein degradation, including degradation of antitoxins necessary to maintain an optimal T:A ratio during persistence, exhibited high turnover. Except for the HipB, two other antitoxins, PrIF and MqsA, also showed high label incorporation. Moreover, several chaperones, which are required for a correct folding of proteins or refolding of denaturated protein aggregates, were among rapidly turning over proteins. Moreover, heat and cold shock proteins important for stress response also showed high Lys8 incorporation. Furthermore, proteins involved in transcription, such as the ribonucleases RNase R (Rnr) and RNase III (Rnc) required for RNA processing and turnover, were also synthesized during persistence at a faster rate including the RNA polymerase sigma factor RpoD. Except from ribosomal proteins, several translation regulatory proteins were also present in this data set, such as the translation initiation factors IF-1 and IF-2 (InfA and InfB), which are essential for the start of protein synthesis.

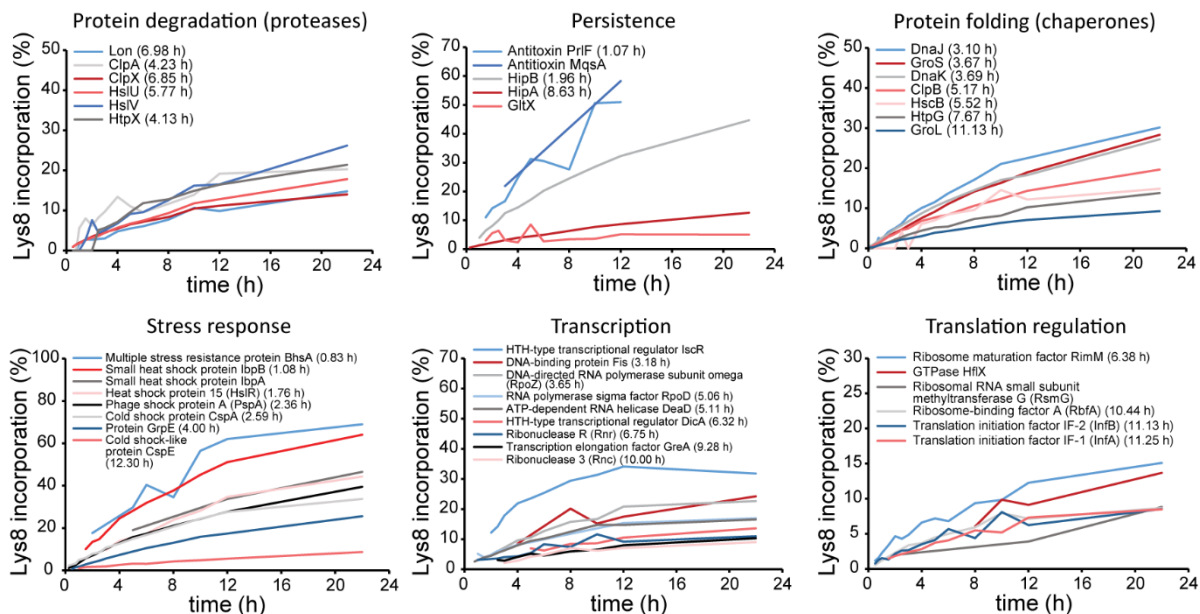


Figure 38. Profiles of selected proteins with highest Lys8 incorporation (Top 25%) during HipA-induced persistence (see Figure 37). Graphs show temporal dependence of Lys8 incorporation calculated from combined protein SILAC H/L ratios from three biological replicates. Proteins are grouped according to their biological functions. Half-life ($T_{1/2}$) values are assigned to each protein if determined (see 4.3.6).

4.3.5 Common mechanisms of HipA- and RelE-mediated persistence

In search for common molecular signature of persistence mediated by other toxins, dynamic SILAC experiment was performed using *relBE* TA system. To that end, protein turnover was investigated during persistence induced by the RelE mRNAse and the experiment was performed analogously to the one with the kinase HipA (see 4.3.3). Expression of *relE* gene was induced from the pBAD33 plasmid followed by the ampicillin treatment and the pulse of Lys8 (**Figure 39A**). Here, the number of

Results

identified proteins was similar to HipA experiment with $1,951 \pm 22$ proteins identified across time points with an estimated FDR of 1.3%. However, the number of quantified proteins was much lower; the number of proteins that incorporated the label during 22 hour time course was 142 ± 15 for RelE compared to 470 ± 64 proteins quantified in the HipA experiment (**Figure 39B**). Likewise, the median incorporation of Lys8 was also lower; it reached $1.22 \pm 0.05\%$. The lower incorporation might be a result of 8-fold lower abundance of RelE compared to HipA determined by their IBAQ intensity in 22 hour time point. Therefore, it is possible that RelE was not produced in a sufficient amount to induce persistence to the same level as it was induced by the higher amount of HipA. Therefore, only the common mechanisms of two toxins were investigated in more detail.

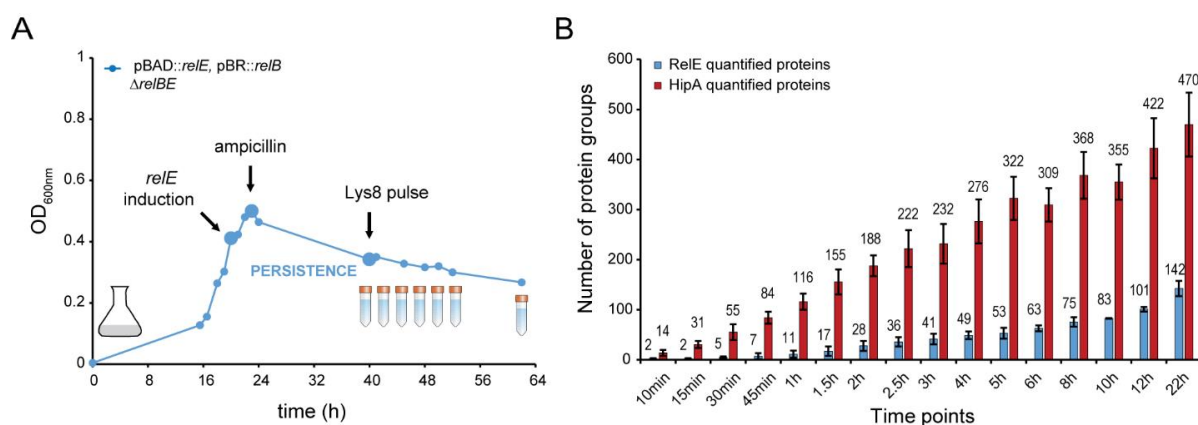


Figure 39. Measurement of newly synthesized proteins during the steady state of RelE-induced persistence using dynamic SILAC. (A) Experimental workflow. *E. coli* MG1655 $\Delta relBE$ strain containing pBAD::relE plasmid, in which *relE* expression is under the control of an arabinose-inducible promoter, and the pBR::relB plasmid, in which *relB* is under the control of an IPTG-inducible promoter was grown in minimal medium containing and chloramphenicol and zeocin for the retention of the plasmids. Expression of *relE* was induced at $OD_{600nm} = 0.4$ with arabinose for 3 hours, followed by the ampicillin treatment ($100 \mu g/ml$) and addition of glucose to repress further *relE* expression. After around 20 hours, a pulse of “heavy” lysine (Lys8) was added to label newly synthesized proteins. Cells were harvested before the pulse (0h) and in 16 different time points after. The growth curve is representative of 3 biological replicates. (B) Number of protein groups identified and quantified by H/L ratio in each time point. Data are means \pm SD from 3 biological replicates.

First, the data from RelE experiment from three replicates was combined as described in 4.3.2 and 3.2.17 and split in 6 time bins as in 4.3.4. In total, 166 proteins were quantified in at least one time bin. GO functional analysis revealed that processes overrepresented in the RelE data set were similar to the HipA data set. Despite the lower expression of *relE*, the overlap of RelE with HipA experiment was high – 155/166 proteins quantified during RelE-mediated persistence were detected in the HipA experiment (**Figure 40**). Surprisingly, the correlation between maximal protein H/L ratio for all time bins was relatively good (Pearson correlation coefficient of 0.48). Measured H/L ratios of individual proteins were consistently lower in RelE than in HipA experiment, which implies that the proportion of persister cells was also lower. Altogether, this suggested that RelE toxin induced persister formation to lesser extent compared to HipA in this experimental setup, but detected proteins can be compared qualitatively.

Results

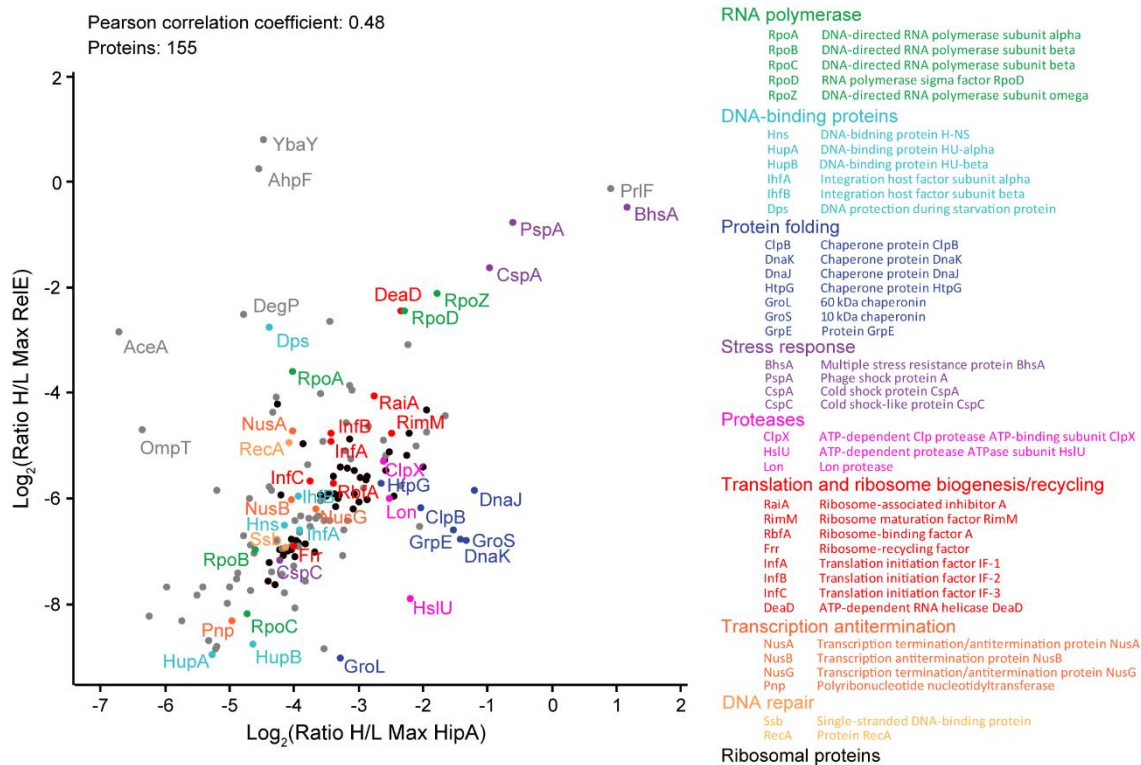


Figure 40. Correlation of maximal protein H/L ratios during HipA- and RelE-induced persistence. Particular proteins are indicated and color-coded according to specific biological functions.

The correlation of H/L ratios from HipA and RelE experiment revealed that proteins, which are actively turning over during persistence are involved in various cellular processes, such as transcription, translation, protein folding and degradation, DNA repair, stress response and others (**Figure 40**). In particular, the antitoxin PrIF and stress-related proteins such as the multiple stress resistance protein BhsA, the phage shock protein PspA and the cold shock protein CspA exhibited high turnover in both experiments. Integration host factors IhfA and IhfB, which are DNA-binding proteins that function in genetic recombination and transcriptional and translational control, were present in this data set and already connected to persistence (45, 46). Moreover, Lon and ClpX protease known to degrade antitoxins were detected in both experiments. Interestingly, proteins of the stress-induced multi-chaperone system, DnaK, DnaJ, ClpB and GrpE, involved in the processing of protein aggregates, were quantified in both experiments and showed a tendency towards higher turnover in HipA than in RelE-induced conditions. Furthermore, many ribosomal proteins and several ribosome-associated proteins (RimM, RaiA, RbfA and Frr) were present in both data sets. Interestingly, together with RNA helicase DeaD, these proteins are involved in the ribosome maturation and biogenesis (279).

Among 11 proteins unique for RelE experiment, two of them exhibited very high incorporation of Lys8, namely, the phage shock protein PspD and the thiosulfate sulfurtransferase PspE. PspD and PspE are expressed from the common phage shock protein (*psp*) operon (*pspABCDE*) important for survival under energy-limited conditions (276). Accordingly, a product of *pspA* gene, which is a part of the same operon, exhibited high turnover in both experiments. In addition, the cold shock protein YdfK and the uncharacterized protein YebG, potentially involved in the DNA repair, were detected only in the RelE experiment.

4.3.6 Estimation of protein half-lives during persistence

The data obtained from these temporally resolved measurements of protein abundances can be used to calculate half-lives of individual proteins. Here, half-lives were determined only for persistence phase during which cells were in a steady-state condition, in which cellular growth was inhibited. In persistence model used in this study, temporal changes in protein abundances reflect true protein turnover, which is independent of the cell division. Therefore, the growth rate does not contribute to the calculation of protein half-lives. Thus, under assumption that there is no cells division and that cells are in a steady-state after ampicillin treatment, protein turnover rate (k) was calculated from the linear dependence of the natural logarithm of the protein H/L ratio and time by linear regression (**Figure 41A**), as described previously (196) (see 3.2.17). Protein H/L ratios derived from a union of three replicates were used for protein turnover rate calculation, from which protein half-lives ($T_{1/2}$) were then determined. To estimate protein turnover rates accurately, two requirements had to be fulfilled: protein H/L ratio had to be measured in at least 5/16 time points and a coefficient of determination (R^2) for linear regression had to be higher than 0.70 to ensure a good curve fitting.

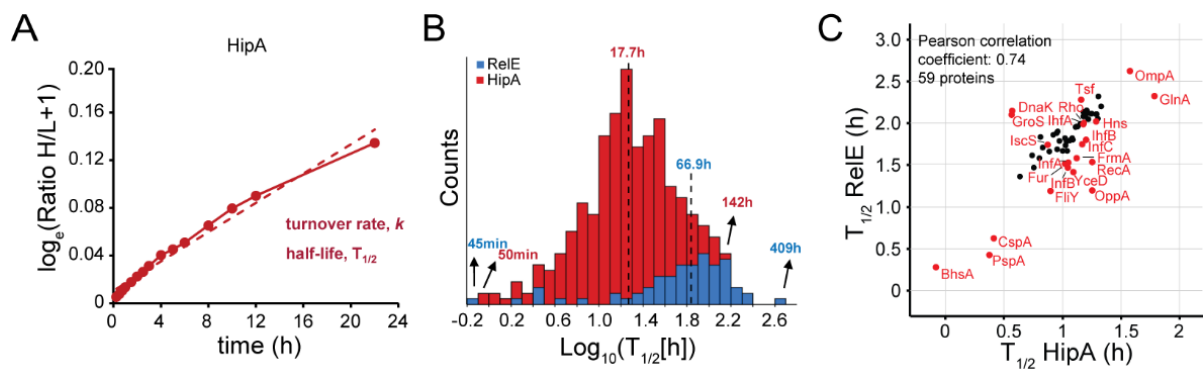


Figure 41. Half-lives of newly synthesized proteins during HipA- and RelE-mediated persistence. (A) Linear regression slope on the example of HipA protein used for calculation of the protein turnover rate (k) from which protein half-lives ($T_{1/2}$) were obtained. (B) Histogram of estimated protein half-lives for HipA (red) and RelE experiment (blue) calculated for 608 and 69 proteins, respectively. (C) Correlation of protein half-lives for HipA and RelE experiment. Ribosomal proteins are indicated in black and other proteins in red.

In total, turnover rates were determined for 609 and 69 proteins in HipA and RelE experiment, respectively, and their half-lives spanned from 45 minutes up to 409 hours (**Figure 41B**). Due to the low Lys8 incorporation measured during RelE-mediated persistence (**Figure 39B**), half-lives of newly synthesized proteins were much longer (median $T_{1/2}$ of 66.9 h) compared to protein half-lives during HipA-mediated persistence (median $T_{1/2}$ of 17.7 h). Although determined protein half-lives differed substantially between experiments on the absolute scale, their trend was very similar as observed by the Pearson correlation coefficient of 0.74 (**Figure 41C**). Except for ribosomal proteins that showed moderate half-lives, several short-lived proteins were common for both studies, BhsA, PspA and CspA. This is in agreement with the results of the data analysis approach used previously (see 4.3.5). In fact, most of top 25% proteins with highest turnover determined by their maximal abundance in HipA experiment had assigned half-lives (**Figure 38**). These proteins were among top 35% proteins with shortest estimated half-lives, which was an additional confirmation of their accurate quantification

Results

across time points that enabled good curve fitting. Furthermore, in contrast to proteins with high turnover rates, two proteins exhibited much longer half-lives relative to the others; those were outer membrane protein OmpA and glutamine synthetase GlnA, which are known to be highly abundant and metabolically stable proteins. Altogether, these results showed that half-lives could be assigned to a considerable set of proteins during HipA-induced persistence and their half-lives correlated with their function; stress response-related proteins exhibited shorter half-lives, while metabolic and membrane proteins were longer lived.

5 Discussion

5.1.1 Phosphoproteome of HipA-induced growth inhibition revealed multiple potential substrates of HipA

*This chapter was adapted from **Semanjski M**, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B, *The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli proteome to promote multidrug tolerance*, *Science Signaling*, 2018, 11(547) (260).*

HipA is the only protein kinase among type II TA-encoded toxins in *E. coli* K-12 (12), shown to strongly inhibit cellular growth and induce bacterial persistence by phosphorylating GltX (63). We hypothesized that HipA could phosphorylate additional protein targets, as shown for other bacterial kinases (217). Using high-resolution mass spectrometry-based proteomics, we were able to confirm that endogenous GltX is phosphorylated when cell growth is inhibited by overproduced HipA, and to detect multiple additional HipA substrates implicated in replication or translation, such as SeqA and ribosomal protein L11 (RplK). Although GltX has been established as a *bona fide* substrate of HipA (63, 77), this is the first time the phosphorylation of GltX on Ser²³⁹ by HipA was confirmed *in vivo* in a whole cell lysate without overexpression and purification of GltX. To detect the phosphorylation site of GltX, the classical phosphoproteomic protocol, which is based on protein digestion with trypsin or LysC-proteases, had to be optimized in order to detect the peptide that contains Ser²³⁹ surrounded by several lysine residues. Phosphorylated GltX peptide was detected by MS only when chymotrypsin was used for protein digestion.

The first phosphoproteomic study revealed that even the mild expression of *hipA* lead to the global increase in protein phosphorylation in *E. coli*. Among detected phosphoproteins, several potential HipA targets seem to be promising candidates for further analysis because of their role in cellular processes associated either with DNA replication, transcription or translation (**Figure 15A**). For example, SeqA negatively regulates DNA replication by binding to newly replicated *oriC* regions thereby preventing premature re-initiation of replication (280). Phosphorylated residue Ser³⁶ is situated in the oligomerization domain of SeqA, which is required for the formation of the filament essential for the function of this protein (281). SeqA is also involved in the stabilization of the DNA during recombination or restart of the replication forks (282). Another potential HipA substrate is the transcriptional regulatory protein RcsB, which regulates the transcription of numerous genes involved in colonic acid capsule synthesis, biofilm formation, cell division and synthesis of outer membrane proteins (283). In addition to RplK, several ribosomal proteins were detected in phosphorylated form, including 30S ribosomal protein S9 (RpsI). RpsI was phosphorylated on the Ser¹²⁸ residue, which is a part of the C-terminal tail important for binding of specific tRNAs at the ribosomal P-site (284), making the detected phosphorylation event potentially relevant for translation inhibition. These multiple substrates of HipA kinase observed in this study imply that HipA may exert pleiotropic effects on essential cellular processes. Although the function of HipA-mediated phosphorylation of RplK in the context of persistence dependent on RelA was not addressed in this study, we believe that the role of

RplK is more complex than initially thought. RplK is a part of the ribosomal stalk that helps the ribosome to interact with GTP-bound translation factors, such as the elongation factor G (EF-G), which catalyzes the GTP-dependent ribosomal translocation step during translation elongation (285). Hence, it would be interesting to analyze the effects of phosphorylation mutants on the translation fidelity and termination. Promiscuous behavior has been observed for other bacterial Ser/Thr kinases, which phosphorylate their substrates less efficiently than histidine kinases to allow for a fine-tuning of the signal (217). In line with that, these data could indicate that several proteins need to be simultaneously phosphorylated in order to integrate the signal that leads to growth inhibition and persistence. Therefore, to observe the effects of these phosphorylation events, a bacterial strain containing mutations in multiple genes that encode here discovered potential HipA substrate should be constructed. Moreover, the lower occupancy of phosphorylation sites of newly identified HipA targets compared to GltX (**Figure 21**) could have an alternative, opposite explanation. Namely, phosphorylation of GltX induces growth inhibition that affects majority of growing population, while phosphorylation of other targets happens only in those cells that became persisters, the proportion of which is more than ten fold lower than the rest of the culture, which is consistent with the low occupancy of phosphorylation sites detected.

5.1.2 Comparison of HipA7 and HipA phosphoproteomes revealed lower activity and smaller substrate pool of HipA7

*This chapter was adapted from **Semanjski M, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B, The kinases HipA and HipA7 phosphorylate different substrate pools in Escherichia coli proteome to promote multidrug tolerance, Science Signaling, 2018, 11(547) (260).***

In contrast to HipA that showed a large pool of targets, its gain-of-function variant HipA7 phosphorylated exclusively GltX. This suggests that phosphorylation of GltX is the main molecular event required for the induction of persistence by both, HipA and HipA7, whereas phosphorylation of other protein targets likely leads to a toxic phenotype observed only for HipA. Although amino acid substitutions in HipA7, G22S and D291A, were previously shown to cause weaker HipA7-HipA7 dimerization in the HipA₂-HipB₄ promoter complex and increased *hipBA7* transcription (19), it was not shown whether the mutations alter the kinase activity of HipA. It was suggested that the catalytic activity is not affected by these substitutions because the active site is located far from Gly²² and Asp²⁹¹ residues (81). However, our *in vivo* phosphoproteomic screens of kinases produced in excess over HipB (**Figure 20, Figure 21**), as well as *in vitro* phosphorylation experiments with purified HipA and HipA7 in the absence of HipB (**Figure 22**), demonstrated that HipA7 is in fact a less active kinase than HipA. Therefore, the less toxic phenotype of HipA7 could also be explained by the lower activity of HipA7 towards GltX, rendering more non-phosphorylated, active GltX available in the cell that is sufficient to sustain the cell growth. The two amino acid substitutions might also impair the substrate specificity of HipA7, leading to a larger pool of substrates phosphorylated by HipA, but not HipA7. Together, this can explain the inability of HipA7 to elicit growth inhibition; however, this hypothesis needs to be further investigated. Our analysis of the phosphorylation site occupancy indicated that a

vast majority (90%) of HipA was autophosphorylated upon the induction from the plasmid (**Figure 21D**), which means that there was only 10% of the active kinase available. Conversely, HipA7 showed much lower phosphorylation of Ser¹⁵⁰ which excludes the possibility that autophosphorylation is responsible for lower activity of HipA7.

In this study we showed a direct evidence that endogenous GltX is phosphorylated *in vivo* by chromosomally encoded HipA7 in the *hipA7* strain, which, to our knowledge, for the first time explains the *hipA7* phenotype. Although HipA7 was less abundant than HipA (**Figure 23**) and showed a weaker kinase activity *in vitro* (**Figure 22**), the increase in GltX phosphorylation by *hipA7* allele could be explained by (i) a disrupted interaction with HipB that leads to a higher abundance of the unbound, active HipA7 and/or (ii) the lower ability of HipA7 to inhibit its kinase activity by autophosphorylation; as a result, a total pool of the weak HipA7 is active and phosphorylates GltX. On the contrary, HipA, which has a more potent kinase activity, is inhibited by its interaction with HipB. With exception of PspA and GltX, phosphorylation of other targets by chromosomally encoded HipA7 was not detected (**Figure 27B, C**). However, higher amounts of HipA7 (produced from a high copy number plasmid) stimulated the phosphorylation of another *E. coli* toxin YjjJ function of which is so far unknown, but is likely to have a kinase activity (**Figure 25B, D**) (275). Autophosphorylation residue Ser¹⁵⁰ of HipA was found to be conserved in YjjJ at the position Ser²⁰⁰ (or Ser²⁰¹) and was detected as phosphorylated in our data. Therefore, it would be interesting to determine the substrate(s) of YjjJ and the link between YjjJ and HipA7.

Finally, the collection of all phosphoproteomic measurements obtained in this study enabled to generate a comprehensive set of phosphorylation events in *E. coli*. In total, 1,183 Ser/Thr/Tyr phosphorylation sites were identified on 632 phosphoproteins. For comparison, the most comprehensive study to date reported 1,687 non-redundant Ser/Thr/Tyr phosphorylation sites by Potel *et al.* (226). However, only 28% of phosphorylation sites identified here were also found in their data set, indicating that a great proportion of detected phosphorylation sites (846) were unique for this study. Thus, this data set can serve as an additional valuable resource for researchers interested in bacterial protein phosphorylation and for further investigation of phosphoregulation in bacteria, especially in the context of persistence.

5.1.3 HipA overproduction is associated with decreased cell motility and HipA7 with increased protein quality control

Except for phosphorylation, this data also provided an extensive resource of proteome changes associated with HipA and HipA7 overproduction. Firstly, several proteins associated with the process of chemotaxis that bacteria use to move in response to the chemical stimulus, decreased in abundance in all experiments in which *hipA* was overexpressed (**Figure 14C, E**). Those proteins included membrane-bound chemotaxis receptors, Tar and Tsr, chemotaxis protein CheW, histidine kinase CheA and the response regulator CheY. Chemotaxis process is used by *E. coli* to navigate towards the concentration gradient of nutrients, or away from toxic metabolites or unfavorable environment (286). Accordingly, the decreased abundance was also observed for the flagellin (FliC), a major

component of flagella that enable bacteria to swim. In contrast, the curli production and assembly/transport components CsgF and CsgG showed increase in abundance after *hipA* expression. Extracellular curli fibers produced by enteric bacteria are used for cell-cell contacts to promote cell aggregation and biofilm formation (287). All this indicated that bacterial motility became restricted as a result of the growth inhibition induced by HipA overproduction and revealed that transient cell aggregation was promoted (288). In contrast, this molecular phenotype was not observed after *hipA7* induction, which is consistent with the inability of HipA7 to inhibit growth.

The unique proteome response was observed in the context of HipA7 overproduction. In these conditions, multiple chaperone proteins (DnaK, DnaJ, ClpB, HtpG), chaperonins (GroL, GroS) and protein GrpE increased in abundance (**Figure 23B**). They are all required for protein quality control by assisting in refolding and folding of polypeptides generated under stress conditions. The multi-chaperone system (ClpB, DnaK, DnaJ and GrpE) prevents the aggregation of denatured proteins for a wide range of substrates (289). Moreover, the increased abundance was also observed for the small heat shock proteins IbpA and IbpB that associate with aggregated proteins to protect them from denaturation and proteolysis, and together with chaperones, enable efficient protein refolding (290). Concomitant with protein quality control process, proteins of a proteasome-like degradation complex (HslU and HslV) were increased in abundance only in the presence of HipA7. This indicates that the increased refolding and unfolding of overproduced HipA7 and other proteins and protein aggregates, followed by their degradation, may be connected to the less toxic phenotype of HipA7 preventing growth inhibition. Interestingly, a connection between HipA7 and the antigen 43 was observed when *hipA7* was expressed from a plasmid or from the chromosome (**Figure 23B, Figure 28B**). This self-associating adhesin that stimulates biofilm formation is produced by the uropathogenic *E. coli* and promotes long-term persistence in the urinary tract infection (291). It is tempting to speculate that the presence of HipA7 may trigger the expression of antigen 43, presence of which in the outer membrane could cause the aggregation of *E. coli* cells and lead to the increase in drug tolerance.

In addition, it is important to note that toxins and antitoxins are in general very low abundant in bacterial proteome even under stress conditions. Therefore, their detection in such complex samples is difficult and requires more extensive pre-fractionation to obtain deeper proteome coverage. However, in this study, several other toxins (RelE, MazG) and antitoxins (DinJ, PrlF, ChpS, MazF) were detected, but they were not regulated upon *hipA* or *hipA7* expression.

5.1.4 HipA-induced persistence is characterized by low metabolic activity and increased stress response

It is generally accepted that persister cells lock themselves into a growth-inhibited state to avoid antibiotics that corrupt active targets. In fact, persisters are assumed to be in a state of slow growth and low metabolic activity, in which a partial cell division was observed (6, 46, 78). This insight prompted us to investigate if persister cells perform translation and which proteins are mainly produced during this process. Therefore, the aim of this study was to establish the method that would enable identification of newly synthesized proteins during persistence on a global scale. To that end,

Discussion

a dynamic SILAC approach based on pulse-labeling of bacterial proteome with isotopically labeled derivative of lysine was used in the context of persistence induced by overproduction of two toxins, HipA and RelE. Dynamic SILAC is a simple and powerful *in vivo* labeling method that has been widely used in eukaryotes for determination of the protein turnover on the global scale in combination with MS (196, 292). The implementation of SILAC approach to bacterial persistence has so far been reported only for selective labeling of drug-tolerant biofilms of *P. aeruginosa* (254).

To our surprise, although the cells were clearly in the growth-inhibited state, the incorporation of the label was indeed measured, and it increased with time, showing that persister cells perform active synthesis of new proteins. Although the average label incorporation was low (around 4%) (**Figure 36B**), this study showed that it is possible to measure such subtle changes in protein abundance. More importantly, a set of proteins exhibited more pronounced label incorporation over time, pointing to the fact that these proteins are actively synthesized in our model of persistence. Therefore, this study provided a resource for identification of proteins that are required to maintain persistence, and ultimately, candidate proteins will be investigated for their functional role.

The major limitation of the experimental setup used in this study is the contamination of the persister cells with the unlabeled pool of proteins that originate from dead cells, which lead to decreased sensitivity of the measurements and lower proteome coverage. Here, we used lytic antibiotic ampicillin to enrich for persisters by physically eliminating sensitive cells; however, the fraction of dead cells with intact membrane was still substantial. This issue could be circumvented by applying an additional enrichment step to isolate persister cells from the batch culture. The method that was described to increase the number of persister cells uses harsh and long enzymatic treatment that targets the cell membrane (293). However, such approach is not compatible with our experiments, as it would impose a significant amount of stress for a longer period, causing proteome changes that would finally lead to data misinterpretation. The alternative would be the usage of flow cytometry based on GFP expression from plasmid carrying specific promoters of genes associated with dormancy, as used previously (19, 45). Ideally, the enrichment should not be performed on living cells, but rather on isolated proteins or peptides, which originate from newly synthesized proteins that incorporated the label. For this, approaches based on click-chemistry could be used, such as the metabolic labeling with methionine analogue L-azidohomoalanine (AHA). Subsequently, affinity tag can be attached to the reactive azide-group and labeled proteins can be enriched by affinity chromatography (255). So far, the application of AHA failed to label the cells without interfering significantly with translation and causing severe growth defect, which is not compatible with our study. However, analogous of other amino acids that would cause less severe effect on translation, could be used, but are so far not developed for such applications.

Consistent with the slow growth attributed to persister cells, proteins required for cell division were either not detected or displayed very low turnover in this data set (FtsH, FtsZ, FtsY and ZipA), suggesting that a residual cell growth is still present, but it is extremely slow. Accordingly, chemotaxis-related proteins that are important for cell motility were not identified, which supports the absence of the cell movement during persistence. The low abundance of proteins involved in metabolism, mainly in biosynthesis of amino acids (**Figure 37**), but also in glycolysis, TCA cycle and pentose

Discussion

phosphate pathway demonstrated that major energy-generating pathways are active in our model of persistence, but their activity is very low. Moreover, cellular respiration was also observed to be strongly decreased during persistence and antibiotic treatments, which is in agreement with previous studies (78, 294). Only several proteins involved in cellular respiration were detected in this data set and exhibited low turnover, namely Ndh, NuoB, NuoC, NuoF, NuoG, NuoI and SdhA. In contrast to reduced metabolism and cell division, we found that ribosomal proteins and general stress response-related proteins exhibited high turnover during persistence. Proteins involved in these processes are discussed below in more detail.

Due to the decreased sensitivity of this experimental setup, proteins quantified in this study are very likely biased towards high abundant proteins. Therefore, toxins of other *E. coli* TA modules were not detected, only two antitoxins, PrIF and MqsA, which are required for maintaining the equilibrium between toxin and antitoxin levels to regulate persistence. Also, proteins involved in the (p)ppGpp synthesis and degradation were not identified as newly synthesized in our data set, which could suggest that they are metabolically stable during persistence. Nevertheless, a number of proteins that have already been connected to persistence were identified here. Those included chaperones DnaK and DnaJ, DNA-binding protein H-NS (Hns), histone-like protein HU (HupA and HupB subunits) and the integration host factor IHF (IhfA and IhfB subunits) (**Figure 40**), that were previously identified to be involved in persister formation using transposon insertion knock-out library of *E. coli* ORFs (40). In that study, the deletion mutants of these genes showed significant decrease in persistence. Integration host factors bind to specific DNA sequences and regulate the transcription of numerous genes (295) and were also found to be elevated in abundance in a proteome study in which persister formation was induced by a nutrient downshift (46). Moreover, H-NS, HU and IHF are nucleoid-associated proteins that are involved in chromosome compaction and gene regulation in *E. coli* (296), suggesting that they might be involved in the regulating the expression of specific genes during persistence.

Furthermore, our study revealed that many chaperones (ClpB, DnaK, DnaJ, HtpG, GroL, GroS, HscB and GrpE) (**Figure 38**) involved in protein folding, were detected as newly synthesized during persistence. Interestingly, DnaK, DnaJ and ClpB chaperones were previously connected to persistence in different studies (40, 108) and deletion of *dnaK* caused a decrease in survival to different antibiotics (40). High turnover of DnaK and other chaperones during persistence observed in our model suggested that they might be required for the maintenance of the proteome homeostasis by enabling correct folding of misfolded client proteins or by targeting unfolded proteins to degradation. Accordingly, several *E. coli* proteases showed high turnover in the HipA experiment (Lon, HslU, HslV and HtpX), as well as the specificity components of the Clp protease (ClpA and ClpX), implying that protein degradation is highly active during persistence. Moreover, SOS response genes are known to be expressed during persistence due to the inhibited DNA replication (108) and one of the most important DNA repair proteins RecA was identified in our study (**Figure 40**).

Except for proteins known to be involved in persistence, our data set also provided a list of previously not reported proteins that potentially play a role in the maintenance of the persistent state. Among proteins that exhibited high turnover, and conversely, shorter half-lives, during persistence mediated by HipA and RelE were antitoxin PrIF and multiple stress resistance protein BhsA (or YcfR) (**Figure 40**).

Discussion

BhsA is an outer membrane-bound protein that has been shown to be highly abundant in biofilms and to increase viability of *E. coli* under stress conditions (297), which is consistent with the stress imposed to cells during persistence. Moreover, the phage shock protein PspA, which is involved in the competition for survival under diverse stress conditions, such as heat or hyperosmotic shock, and in maintaining proton-motive force and membrane integrity in response to extracytoplasmic stress (298), was detected as highly abundant in both experiments. Interestingly, increased abundance of *pspA*, *pspD* and *pspE* transcripts was also observed in persister cells isolated by FACS sorting from an exponentially growing culture (45). The same observation was obtained in transcriptome analysis of the *hipA7* strain treated with ampicillin. There, in addition to *pspA*, *pspD* and *pspE*, other genes of the *psp* operon, *pspB* and *pspC* showed an increase in expression measured by their mRNA levels (108). These observations suggest that PspA plays an important role during persistence, but the exact mechanism is so far not clear. It is known that PspA negatively regulates transcription of the *psp* operon during stress by binding to the activator PspF (298), however, PspD and PspE proteins were detected only in RelE experiments of our study, where they showed higher abundance and turnover compared to PspA. The increased expression of the *psp* operon was also observed during biofilm formation in *E. coli* (299) and in *S. Typhimurium* during macrophage infection (300). Whereas the N-terminal domain of PspA is involved in PspF binding, the C-terminal domain is essential for the formation of a membrane-associated high-order PspA oligomer (36-mer), which functions as an effector that repairs the membrane to prevent proton leakage (301). The increased amounts of PspA does not necessarily shut down the *psp* induction, but can shift the function of PspA to act as the effector. This implies that PspA is activated during ampicillin treatment to maintain the membrane stability, and together with BhsA, PspA could have a function to lower cell membrane stress. It would be interesting to see if the same results would be obtain if we used other type of antibiotic for generation of persister cells, such as fluoroquinolones. Nevertheless, to reveal the function of the *psp* operon encoded proteins during persistence, further functional analysis with *psp* mutants should be made in the context of persistence.

Notably, the major cold shock protein CspA in *E. coli* was detected as a fast turning over protein in both experiments. Increased synthesis of CspA has been detected in *E. coli* treated with the inhibitors of protein synthesis, chloramphenicol and kanamycin (302). CspA is in fact an RNA chaperone that together with other cold shock proteins CspE and CspC, which were also detected in this study, functions as a transcription antiterminator to prevent premature termination of RNA synthesis (303). These cold shock proteins were shown to induce production of the transcription termination/antitermination protein NusA, the translation initiation factor IF-2 (InfB), the 30S ribosome-binding factor RbfA and the polyribonucleotide nucleotidyltransferase Pnp, that are known to be induced during cold shock and were detected in this study (**Figure 40**). NusA, together with the transcription termination factor Rho and transcription antitermination proteins NusB and NusG, form a regulatory complex that associates with RNA polymerase (RNAP) to prevent transcription termination at otherwise functional terminator regions (304). This process is known as processive antitermination and it is important for regulating the efficiency of RNA production. Accordingly, all RNA polymerase (RNAP) subunits (RpoA, RpoB, RpoC and RpoZ) were actively synthesized during persistence in both models used in this study (**Figure 40**). Interestingly, among RNA polymerase sigma

factors, only RpoD was identified in this study, whereas RpoS, which has most often been connected to persistence (40, 100, 104), was not detected. RpoD promotes transcription of ribosomal operons and other protein synthesis-related and other housekeeping genes (305). With this in line, synthesis of ribosomal proteins during persistence induced by both, HipA and RelE, was observed, indicating that translation is important for maintaining the persistent state. Apart from proteins that are primary constituents of the ribosome, several ribosome-associated proteins were showed relatively high turnover, namely, the ribosome-associated inhibitor A (RaiA), the ribosome maturation factor RimM, the ribosome-binding factor A (RbfA) and the ribosome-recycling factor Frr. Together with RNA helicase DeaD and chaperones DnaK and GroEL, RimM and RbfA are involved in the process of the ribosome maturation that is carried out by several different maturation factors (279). In particular, RbfA and RimM assist in processing of 16S rRNA during assembly of the 30S ribosomal subunit. Moreover, when HipA was overproduced, RbfA was found phosphorylated on Ser¹¹⁰, which is located on the C-terminal region of the protein involved in the binding to the 30S ribosomal subunit (306). This phosphorylation event might be an additional level of regulation of the activity of RbfA in triggering persistence formation and during the steady state of persistence. Furthermore, the ribosome modulation factor Rmf that is involved in ribosome hibernation, a process that converts ribosome to an inactive form, has previously been connected to persistence (46, 108); however, it was not identified here. Conversely, another ribosome inhibitor was identified, namely, RaiA, which regulates translation efficiency by blocking the ribosomal A-site during stationary phase (307). In addition, the ribosome-recycling factor Frr, which releases ribosomes from mRNA after translation termination to recycle ribosomes and increase the efficiency of translation was identified in this data set (308). Interestingly, a heat shock-inducible ribosome-dependent GTPase HflX, that rescues stalled ribosomes from deacylated tRNA was found only in the HipA data set, which is consistent with the accumulation of uncharged tRNA due to GltX inhibition by HipA (63). Altogether, it can be concluded that, in our model of persistence, transcription and translation are strictly regulated in order to maintain a basal level of protein synthesis during steady state of persistence. Therefore, above mentioned regulators of transcription antitermination and ribosome maturation and recycling are interesting candidates for further follow up experiments, in which their deletion might lead to decrease in persistence or, alternatively, in prevention of resuscitation.

5.1.5 Persister resuscitation is characterized by energy production

The awakening of persister cells from dormancy is obviously an important process that enables persisters to revert to growth in the absence of stress and ultimately allow their survival. To gain a deeper insight into persister resuscitation, we applied pulse-labeling method to persister cells enriched by *hipA* overexpression and ampicillin treatment during their regrowth from dormancy. In our model, the removal of the antibiotic and addition of the fresh medium initiated the resuscitation process, which was accompanied with the induction of *hipB* expression. The main aim of this study was to generate a list of individual proteins that are rapidly synthesized on the onset of resuscitation and therefore potentially involved in the switch from persistence to growth resumption.

At the very onset of resuscitation, cells restored almost the entire translation machinery and restarted their metabolism by inducing glycolysis to convert glucose from the fresh medium into energy, which can then be used for the regrowth (**Figure 33**). Interestingly, two chaperones (DnaK and GroL) and transcription-related proteins, Rho, DeaD and Pnp, that are involved in RNA degradation, were identified during early response, possibly regulating transcription of specific genes. Because these proteins were also detected to be highly produced during persistence, it is possible that the processes common to persistence are still active during the initial phase of the pulse-labeling. The delay in the wake-up from dormancy was observed previously and depends on several factors, such as the composition of the outgrowth medium or the frequency of persister cells (309, 310). Alternative explanation for high abundance of newly synthesized chaperones during both, persistence and resuscitation might be that these proteins are required for resuscitation, but are produced during dormancy to prepare persister cells for their rapid wake-up when conditions become favorable. Furthermore, the expression of the antitoxin *hipB* from the plasmid was used to assist in triggering resuscitation and to serve as internal positive control. Indeed, HipB was among proteins with the highest label incorporation at the very early phase and also throughout entire resuscitation (**Figure 35**). Notably, RNA polymerase sigma factors RpoS and RpoD showed high label incorporation during early response. RpoS is important for transcriptional reprogramming of many genes that are mainly involved in the metabolism and stress response, whereas RpoD preferentially induces transcription of genes associated with fast growth, such as ribosomal proteins, which is in agreement with our findings (305). Considering that RpoD was also detected during persistence, it can be assumed that it might have a dual role during both processes. Moreover, several other proteins were rapidly synthesized during early resuscitation, such as the N-ethylmaleimide reductase NemaA, which reduces N-ethylmaleimide, that otherwise modifies cysteine residues of cellular proteins and inhibits growth (311). In addition, NemaA degrades toxic compounds to use them as a source of nitrogen (312).

With time, incorporation of the label increased and revealed that a number of proteins involved in metabolism were produced, mainly in the biosynthesis of amino acids, most likely to use them for protein synthesis, as they are not provided in the medium (except for lysine). To resume their growth, cells also produced almost all aminoacyl-tRNA charging enzymes, of which GltX showed the highest label incorporation. Furthermore, proteins involved in the sulfate reduction, which is important for energy production, exhibited high abundance during resuscitation, along with proteins involved in the iron-sulfur cluster assembly, which participate in the electron transport and stabilization of proteins (313). Moreover, protein components of the pentose phosphate pathway increased in abundance as cells continued to resuscitate. Parallel to glycolysis, this pathway is required for generating energy and production of precursors for nucleotides and aromatic amino acids. Altogether, this study demonstrated that the resuscitation from dormancy is characterized by the metabolic switch towards energy production and active translation, which is consistent with a study that investigated awakening of dormant, nitrogen-starved cyanobacteria (314). More importantly, it provided a valuable resource of proteins newly synthesized during resuscitation, of which some are likely involved in the wake-up from persistence. Therefore, this study could help in understanding of the molecular mechanisms underlying both, persister formation and resuscitation, which is important for the development of approaches for their eradication.

6 Conclusions

In this thesis, several approaches were employed to characterize proteome, phosphoproteome and protein turnover during bacterial persistence mediated by the *E. coli* kinase HipA. Based on the obtained results, following conclusions can be drawn:

1. Identification of *in vivo* phosphorylation targets of HipA
 - a. Optimized phosphoproteomic workflow, which involved protein digestion with two different proteases, was applied to study phosphorylation events associated with the kinase HipA. Phosphorylation of a *bona fide* substrate of HipA, GltX, was confirmed by mass spectrometry and for the first time detected *in vivo*.
 - b. Integration of two triple-SILAC experiments enabled detection of phosphorylation events induced by ectopically expressed *hipA* and their temporal characterization during HipA-mediated persistence and HipB-triggered resuscitation. In addition to GltX, a number of previously unknown potential phosphorylation targets of HipA were identified.
 - c. HipA-dependent phosphorylation of two candidate proteins, ribosomal protein L11 (RplK) and negative modulator of initiation of replication SeqA, was confirmed *in vitro*.
 - d. Physiological effect of HipA-specific phosphorylation events detected in phosphoproteomic screens was investigated on the RplK example. However, to that end, phosphorylation of RplK showed no effect on RelA-dependent persistence.
2. Comparison of HipA and HipA7 (phospho)proteomes to distinguish between different phenotypes of two kinase variants
 - a. Mildly produced HipA and HipA7 showed that GltX is the main substrate of both kinase variants. Whereas HipA7 only targeted GltX, HipA phosphorylated several additional unique substrates, which might explain why HipA inhibits growth more efficiently than HipA7. Furthermore, HipA7 showed reduced kinase activity *in vitro* compared to HipA and targeted a substrate pool similar to that of HipA only when highly overproduced *in vivo*.
 - b. Expression of *hipA7* from the chromosome resulted in phosphorylation of GltX, providing a direct evidence that endogenous GltX is phosphorylated by chromosomally encoded HipA7 and suggesting that GltX is the primary determinant of persistence. Additionally, phage shock protein PspA was detected as a potential target unique to HipA7. Although HipA is a more potent kinase, chromosomally encoded HipA did not phosphorylate GltX as strongly as HipA7, due to its inhibition by the interaction with other HipA molecule in the HipA-HipB complex, which is disrupted in the case of HipA7.
 - c. The two kinase variants also differed in autophosphorylation, which was significantly lower in HipA7 when it was produced from the plasmid. The lower ability of HipA7 to inhibit its kinase activity by autophosphorylation could explain increased phosphorylation of GltX by chromosomally encoded HipA7; while HipA can be inhibited by

Conclusions

autophosphorylation and formation of the HipA-HipB complex, a total pool of the weak HipA7 produced from the chromosome is active and phosphorylates GltX.

3. Time-resolved analysis of newly synthesized proteins during toxin-induced persistence and antitoxin-mediated resuscitation
 - a. Dynamic SILAC pulse-labeling approach was established and optimized for the model of HipA-induced persistence and HipB-mediated resuscitation in order to study newly synthesized proteins in respective conditions. It was shown that such pulse-labeling with isotopically labeled amino acids in combination with high resolution mass spectrometry can be applied to different TA modules and used as a general strategy to selectively label persister cells and study protein turnover on a system-wide scale.
 - b. Utilization of the pulse-labeling to HipB-mediated resuscitation revealed that newly synthesized proteins in the early phase of resuscitation are involved in translation, glycolysis and biosynthesis of amino acids to provide energy required for resumption of growth. This data set provides a valuable resource for further validation of proteins that might act as triggers of the resuscitation.
 - c. Application of the pulse-labeling to persistent state of cells enriched by overproduction of HipA revealed that dormant persisters are translationally active and have reduced metabolism, division and cellular respiration. In contrast, they actively synthesized a set of proteins that are involved in general stress response and translation in order to maintain persister survival.
 - d. Analysis of newly synthesized proteins during persistence induced by another toxin, mRNase RelE, revealed very similar set of proteins as observed for the kinase HipA. Comparison of two data sets enabled to investigate common mechanisms of persistence, which are characterized by the production of chaperones necessary for correct protein folding, proteases that degrade unfolded proteins, cold and heat shock involved in general stress response, transcription antitermination factors that regulate efficiency of RNA production and ribosomal and translation-associated proteins required for ribosome maturation.

7 References

1. Ventola CL (2015) The Antibiotic Resistance Crisis: Part 1: Causes and Threats. *Pharmacy and Therapeutics* 40(4):277-283.
2. Brauner A, Fridman O, Gefen O, & Balaban NQ (2016) Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews Microbiology* 14:320.
3. Munita JM & Arias CA (2016) Mechanisms of Antibiotic Resistance. *Microbiology spectrum* 4(2):10.1128/microbiolspec.VMBF-0016-2015.
4. Van den Bergh B, Fauvart M, & Michiels J (2017) Formation, physiology, ecology, evolution and clinical importance of bacterial persisters. *FEMS Microbiology Reviews* 41(3):219-251.
5. Balaban NQ (2011) Persistence: mechanisms for triggering and enhancing phenotypic variability. *Current Opinion in Genetics & Development* 21(6):768-775.
6. Balaban NQ, Merrin J, Chait R, Kowalik L, & Leibler S (2004) Bacterial Persistence as a Phenotypic Switch. *Science* 305(5690):1622-1625.
7. Lewis K (2010) Persister Cells. *Annual Review of Microbiology* 64(1):357-372.
8. Amato SM, Fazen CH, Henry TC, Mok WWK, Orman MA, Sandvik EL, Volzing KG, & Brynildsen MP (2014) The role of metabolism in bacterial persistence. *Frontiers in Microbiology* 5:70.
9. Helaine S, Cheverton AM, Watson KG, Faure LM, Matthews SA, & Holden DW (2014) Internalization of Salmonella by Macrophages Induces Formation of Nonreplicating Persisters. *Science* 343(6167):204-208.
10. Bigger J (1944) Treatment of staphylococcal infections with penicillin by intermittent sterilisation. *The Lancet* 244(6320):497-500.
11. Keren I, Kaldalu N, Spoering A, Wang Y, & Lewis K (2004) Persister cells and tolerance to antimicrobials. *FEMS Microbiology Letters* 230(1):13-18.
12. Maisonneuve E & Gerdes K (2014) Molecular Mechanisms Underlying Bacterial Persisters. *Cell* 157(3):539-548.
13. Lewis K (2006) Persister cells, dormancy and infectious disease. *Nature Reviews Microbiology* 5:48.
14. Semanjski M & Macek B (2016) Shotgun proteomics of bacterial pathogens: advances, challenges and clinical implications. *Expert Review of Proteomics* 13(2):139-156.
15. Spoering AL & Lewis K (2001) Biofilms and Planktonic Cells of *Pseudomonas aeruginosa* Have Similar Resistance to Killing by Antimicrobials. *Journal of Bacteriology* 183(23):6746-6751.
16. Radzikowski JL, Schramke H, & Heinemann M (2017) Bacterial persistence from a system-level perspective. *Current Opinion in Biotechnology* 46:98-105.
17. Lewis K (2001) Riddle of Biofilm Resistance. *Antimicrobial Agents and Chemotherapy* 45(4):999-1007.
18. Mulcahy LR, Burns JL, Lory S, & Lewis K (2010) Emergence of *Pseudomonas aeruginosa* Strains Producing High Levels of Persister Cells in Patients with Cystic Fibrosis. *Journal of Bacteriology* 192(23):6191-6199.
19. Schumacher MA, Balani P, Min J, Chinnam NB, Hansen S, Vulić M, Lewis K, & Brennan RG (2015) HipBA-promoter structures reveal the basis of heritable multidrug tolerance. *Nature* 524:59.
20. LaFleur MD, Qi Q, & Lewis K (2010) Patients with Long-Term Oral Carriage Harbor High-Persister Mutants of *Candida albicans*. *Antimicrobial Agents and Chemotherapy* 54(1):39-44.
21. Michiels JE, Van den Bergh B, Verstraeten N, & Michiels J (2016) Molecular mechanisms and clinical implications of bacterial persistence. *Drug Resistance Updates* 29:76-89.
22. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, Wong K-K, Brandstetter K, Wittner B, Ramaswamy S, Classon M, & Settleman J (2010) A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. *Cell* 141(1):69-80.

References

23. Kussell E & Leibler S (2005) Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. *Science* 309(5743):2075-2078.
24. Veening J-W, Smits WK, & Kuipers OP (2008) Bistability, Epigenetics, and Bet-Hedging in Bacteria. *Annual Review of Microbiology* 62(1):193-210.
25. Elowitz MB, Levine AJ, Siggia ED, & Swain PS (2002) Stochastic Gene Expression in a Single Cell. *Science* 297(5584):1183-1186.
26. Rotem E, Loinger A, Ronin I, Levin-Reisman I, Gabay C, Shoshitaishvili N, Biham O, & Balaban NQ (2010) Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proceedings of the National Academy of Sciences* 107(28):12541-12546.
27. Cohen Nadia R, Lobritz Michael A, & Collins James J (2013) Microbial Persistence and the Road to Drug Resistance. *Cell Host & Microbe* 13(6):632-642.
28. Harms A, Brodersen DE, Mitarai N, & Gerdes K (2018) Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Molecular Cell*.
29. Wakamoto Y, Dhar N, Chait R, Schneider K, Signorino-Gelo F, Leibler S, & McKinney JD (2013) Dynamic Persistence of Antibiotic-Stressed Mycobacteria. *Science* 339(6115):91-95.
30. Wu Y, Vulić M, Keren I, & Lewis K (2012) Role of Oxidative Stress in Persister Tolerance. *Antimicrobial Agents and Chemotherapy* 56(9):4922-4926.
31. Amato Stephanie M, Orman Mehmet A, & Brynildsen Mark P (2013) Metabolic Control of Persister Formation in Escherichia coli. *Molecular Cell* 50(4):475-487.
32. Solopova A, van Gestel J, Weissing FJ, Bachmann H, Teusink B, Kok J, & Kuipers OP (2014) Bet-hedging during bacterial diauxic shift. *Proceedings of the National Academy of Sciences* 111(20):7427-7432.
33. Kotte O, Volkmer B, Radzikowski JL, & Heinemann M (2014) Phenotypic bistability in Escherichia coli's central carbon metabolism. *Molecular Systems Biology* 10(7).
34. Leung V & Lévesque CM (2012) A Stress-Inducible Quorum-Sensing Peptide Mediates the Formation of Persister Cells with Noninherited Multidrug Tolerance. *Journal of Bacteriology* 194(9):2265-2274.
35. Vega NM, Allison KR, Khalil AS, & Collins JJ (2012) Signaling-mediated bacterial persister formation. *Nature Chemical Biology* 8:431.
36. Nguyen D, Joshi-Datar A, Lepine F, Bauerle E, Olakanmi O, Beer K, McKay G, Siehnel R, Schafhauser J, Wang Y, Britigan BE, & Singh PK (2011) Active Starvation Responses Mediate Antibiotic Tolerance in Biofilms and Nutrient-Limited Bacteria. *Science* 334(6058):982-986.
37. Leszczynska D, Matuszewska E, Kuczynska-Wisnik D, Furmanek-Blaszk B, & Laskowska E (2013) The Formation of Persister Cells in Stationary-Phase Cultures of Escherichia Coli Is Associated with the Aggregation of Endogenous Proteins. *PLOS ONE* 8(1):e54737.
38. Vega NM, Allison KR, Samuels AN, Klempner MS, & Collins JJ (2013) Salmonella typhimurium intercepts Escherichia coli signaling to enhance antibiotic tolerance. *Proceedings of the National Academy of Sciences* 110(35):14420-14425.
39. Moyed HS & Bertrand KP (1983) hipA, a newly recognized gene of Escherichia coli K-12 that affects frequency of persistence after inhibition of murein synthesis. *Journal of Bacteriology* 155(2):768-775.
40. Hansen S, Lewis K, & Vulić M (2008) Role of Global Regulators and Nucleotide Metabolism in Antibiotic Tolerance in Escherichia coli. *Antimicrobial Agents and Chemotherapy* 52(8):2718-2726.
41. Ma C, Sim S, Shi W, Du L, Xing D, & Zhang Y (2010) Energy production genes sucB and ubiF are involved in persister survival and tolerance to multiple antibiotics and stresses in Escherichia coli. *FEMS Microbiology Letters* 303(1):33-40.
42. De Groote VN, Verstraeten N, Fauvart M, Kint CI, Verbeeck AM, Beullens S, Cornelis P, & Michiels J (2009) Novel persistence genes in Pseudomonas aeruginosa identified by high-throughput screening. *FEMS Microbiology Letters* 297(1):73-79.

References

43. Spoering AL, Vulić M, & Lewis K (2006) GlpD and PlsB Participate in Persister Cell Formation in *Escherichia coli*. *Journal of Bacteriology* 188(14):5136-5144.
44. Keren I, Minami S, Rubin E, & Lewis K (2011) Characterization and Transcriptome Analysis of *Mycobacterium tuberculosis* Persisters. *mBio* 2(3).
45. Shah D, Zhang Z, Khodursky AB, Kaldalu N, Kurg K, & Lewis K (2006) Persisters: a distinct physiological state of *E. coli*. *BMC Microbiology* 6(1):53.
46. Radzikowski JL, Vedelaar S, Siegel D, Ortega AD, Schmidt A, & Heinemann M (2016) Bacterial persistence is an active σ S stress response to metabolic flux limitation. *Molecular Systems Biology* 12(9).
47. Black DS, Kelly AJ, Mardis MJ, & Moyed HS (1991) Structure and organization of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *Journal of Bacteriology* 173(18):5732-5739.
48. Falla TJ & Chopra I (1998) Joint Tolerance to β -Lactam and Fluoroquinolone Antibiotics in *Escherichia coli* Results from Overexpression of *hipA*. *Antimicrobial Agents and Chemotherapy* 42(12):3282-3284.
49. Korch SB, Henderson TA, & Hill TM (2003) Characterization of the *hipA7* allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Molecular Microbiology* 50(4):1199-1213.
50. Page R & Peti W (2016) Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nature Chemical Biology* 12:208.
51. Wang X, Lord DM, Hong SH, Peti W, Benedik MJ, Page R, & Wood TK (2013) Type II Toxin/Antitoxin MqsR/MqsA Controls Type V Toxin/Antitoxin GhoT/GhoS. *Environmental microbiology* 15(6):1734-1744.
52. Aakre CD, Phung TN, Huang D, & Laub MT (2013) A Bacterial Toxin Inhibits DNA Replication Elongation Through a Direct Interaction with the β Sliding Clamp. *Molecular cell* 52(5):617-628.
53. Ogura T & Hiraga S (1983) Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences of the United States of America* 80(15):4784-4788.
54. Gerdes K, Rasmussen PB, & Molin S (1986) Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences* 83(10):3116-3120.
55. Coray DS, Wheeler NE, Heinemann JA, & Gardner PP (2017) Why so narrow: Distribution of anti-sense regulated, type I toxin-antitoxin systems compared with type II and type III systems. *RNA Biology* 14(3):275-280.
56. Goeders N, Chai R, Chen B, Day A, & Salmond G (2016) Structure, Evolution, and Functions of Bacterial Type III Toxin-Antitoxin Systems. *Toxins* 8(10):282.
57. Ramage HR, Connolly LE, & Cox JS (2009) Comprehensive Functional Analysis of *Mycobacterium tuberculosis* Toxin-Antitoxin Systems: Implications for Pathogenesis, Stress Responses, and Evolution. *PLOS Genetics* 5(12):e1000767.
58. Guo Y, Quiroga C, Chen Q, McNulty MJ, Benedik MJ, Wood TK, & Wang X (2014) RalR (a DNase) and RalA (a small RNA) form a type I toxin-antitoxin system in *Escherichia coli*. *Nucleic Acids Research* 42(10):6448-6462.
59. Christensen SK & Gerdes K (2003) RelE toxins from Bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Molecular Microbiology* 48(5):1389-1400.
60. Zhang Y, Zhang J, Hoeflich KP, Ikura M, Qing G, & Inouye M (2003) MazF Cleaves Cellular mRNAs Specifically at ACA to Block Protein Synthesis in *Escherichia coli*. *Molecular Cell* 12(4):913-923.

References

61. Jørgensen MG, Pandey DP, Jaskolska M, & Gerdes K (2009) HicA of *Escherichia coli* Defines a Novel Family of Translation-Independent mRNA Interferases in Bacteria and Archaea. *Journal of Bacteriology* 191(4):1191-1199.
62. Winther KS & Gerdes K (2011) Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of the National Academy of Sciences* 108(18):7403-7407.
63. Germain E, Castro-Roa D, Zenkin N, & Gerdes K (2013) Molecular Mechanism of Bacterial Persistence by HipA. *Molecular Cell* 52(2):248-254.
64. Castro-Roa D, Garcia-Pino A, De Gieter S, van Nuland NAJ, Loris R, & Zenkin N (2013) The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nature Chemical Biology* 9:811.
65. Harms A, Stanger Frédéric V, Scheu Patrick D, de Jong Imke G, Goepfert A, Glatter T, Gerdes K, Schirmer T, & Dehio C (2015) Adenylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Reports* 12(9):1497-1507.
66. Cheverton Angela M, Gollan B, Przydacz M, Wong Chi T, Mylona A, Hare Stephen A, & Helaine S (2016) A *Salmonella* Toxin Promotes Persister Formation through Acetylation of tRNA. *Molecular Cell* 63(1):86-96.
67. Verstraeten N, Knapen Wouter J, Kint Cyrielle I, Liebens V, Van den Bergh B, Dewachter L, Michiels Joran E, Fu Q, David Charlotte C, Fierro Ana C, Marchal K, Beirlant J, Versées W, Hofkens J, Jansen M, Fauvart M, & Michiels J (2015) Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. *Molecular Cell* 59(1):9-21.
68. Muthuramalingam M, White J, & Bourne C (2016) Toxin-Antitoxin Modules Are Pliable Switches Activated by Multiple Protease Pathways. *Toxins* 8(7):214.
69. Chan WT, Espinosa M, & Yeo CC (2016) Keeping the Wolves at Bay: Antitoxins of Prokaryotic Type II Toxin-Antitoxin Systems. *Frontiers in Molecular Biosciences* 3:9.
70. Overgaard M, Borch J, Jørgensen MG, & Gerdes K (2008) Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Molecular Microbiology* 69(4):841-857.
71. Moyed HS & Broderick SH (1986) Molecular cloning and expression of hipA, a gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *Journal of Bacteriology* 166(2):399-403.
72. Black DS, Irwin B, & Moyed HS (1994) Autoregulation of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *Journal of Bacteriology* 176(13):4081-4091.
73. Korch SB & Hill TM (2006) Ectopic Overexpression of Wild-Type and Mutant hipA Genes in *Escherichia coli*: Effects on Macromolecular Synthesis and Persister Formation. *Journal of Bacteriology* 188(11):3826-3836.
74. Hansen S, Vulić M, Min J, Yen T-J, Schumacher MA, Brennan RG, & Lewis K (2012) Regulation of the *Escherichia coli* HipBA Toxin-Antitoxin System by Proteolysis. *PLOS ONE* 7(6):e39185.
75. Correia FF, D'Onofrio A, Rejtar T, Li L, Karger BL, Makarova K, Koonin EV, & Lewis K (2006) Kinase Activity of Overexpressed HipA Is Required for Growth Arrest and Multidrug Tolerance in *Escherichia coli*. *Journal of Bacteriology* 188(24):8360-8367.
76. Schumacher MA, Piro KM, Xu W, Hansen S, Lewis K, & Brennan RG (2009) Molecular Mechanisms of HipA-Mediated Multidrug Tolerance and Its Neutralization by HipB. *Science* 323(5912):396-401.
77. Kaspy I, Rotem E, Weiss N, Ronin I, Balaban NQ, & Glaser G (2013) HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nature Communications* 4:3001.
78. Bokinsky G, Baidoo EE, Akella S, Burd H, Weaver D, Alonso-Gutierrez J, García-Martín H, Lee TS, & Keasling JD (2013) HipA-triggered growth arrest and β -lactam tolerance in *Escherichia*

References

- coli are mediated by RelA-dependent ppGpp synthesis. *Journal of bacteriology* 195(14):3173-3182.
79. Germain E, Roghanian M, Gerdes K, & Maisonneuve E (2015) Stochastic induction of persister cells by HipA through (p)ppGpp-mediated activation of mRNA endonucleases. *Proceedings of the National Academy of Sciences* 112(16):5171-5176.
 80. Gerdes K & Maisonneuve E (2012) Bacterial Persistence and Toxin-Antitoxin Loci. *Annual Review of Microbiology* 66(1):103-123.
 81. Schumacher MA, Min J, Link TM, Guan Z, Xu W, Ahn Y-H, Soderblom EJ, Kurie JM, Evdokimov A, Moseley MA, Lewis K, & Brennan RG (2012) Role of unusual P-loop ejection and autophosphorylation in HipA-mediated persistence and multidrug tolerance. *Cell reports* 2(3):518-525.
 82. Stancik IA, Šestak MS, Ji B, Axelson-Fisk M, Franjevic D, Jers C, Domazet-Lošo T, & Mijakovic I (2018) Serine/Threonine Protein Kinases from Bacteria, Archaea and Eukarya Share a Common Evolutionary Origin Deeply Rooted in the Tree of Life. *Journal of Molecular Biology* 430(1):27-32.
 83. Lin CY, Awano N, Masuda H, Park JH, & Inouye M (2013) Transcriptional Repressor HipB Regulates the Multiple Promoters in Escherichia coli. *Journal of Molecular Microbiology and Biotechnology* 23(6):440-447.
 84. Masuda H & Inouye M (2017) Toxins of Prokaryotic Toxin-Antitoxin Systems with Sequence-Specific Endoribonuclease Activity. *Toxins* 9(4):140.
 85. Christensen SK, Mikkelsen M, Pedersen K, & Gerdes K (2001) RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National Academy of Sciences* 98(25):14328-14333.
 86. Kim P, K. CS, & Kenn G (2002) Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology* 45(2):501-510.
 87. Neubauer C, Gao Y-G, Andersen KR, Dunham CM, Kelley AC, Hentschel J, Gerdes K, Ramakrishnan V, & Brodersen DE (2009) The Structural Basis for mRNA Recognition and Cleavage by the Ribosome-Dependent Endonuclease RelE. *Cell* 139(6):1084-1095.
 88. Pedersen K, Zavialov AV, Pavlov MY, Elf J, Gerdes K, & Ehrenberg M (2003) The Bacterial Toxin RelE Displays Codon-Specific Cleavage of mRNAs in the Ribosomal A Site. *Cell* 112(1):131-140.
 89. Hurley JM, Cruz JW, Ouyang M, & Woychik NA (2011) Bacterial Toxin RelE Mediates Frequent Codon-independent mRNA Cleavage from the 5' End of Coding Regions in Vivo. *Journal of Biological Chemistry* 286(17):14770-14778.
 90. Hwang J-Y & Buskirk AR (2017) A ribosome profiling study of mRNA cleavage by the endonuclease RelE. *Nucleic Acids Research* 45(1):327-336.
 91. Himeno H, Kurita D, & Muto A (2014) tmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell. *Frontiers in Genetics* 5:66.
 92. Keiler KC, Waller PRH, & Sauer RT (1996) Role of a Peptide Tagging System in Degradation of Proteins Synthesized from Damaged Messenger RNA. *Science* 271(5251):990-993.
 93. Gotfredsen M & Gerdes K (1998) The Escherichia coli relBE genes belong to a new toxin-antitoxin gene family. *Molecular Microbiology* 29(4):1065-1076.
 94. Bøggild A, Sofos N, Andersen Kasper R, Feddersen A, Easter Ashley D, Passmore Lori A, & Brodersen Ditlev E (2012) The Crystal Structure of the Intact E. coli RelBE Toxin-Antitoxin Complex Provides the Structural Basis for Conditional Cooperativity. *Structure* 20(10):1641-1648.
 95. Winther KS & Gerdes K (2012) Regulation of Enteric vapBC Transcription: Induction by VapC Toxin Dimer-Breaking. *Nucleic Acids Research* 40(10):4347-4357.
 96. Hassan A, Noureddine A, Martine C, & Laurence VM (2001) The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system. *Molecular Microbiology* 41(1):73-82.

References

97. Garcia-Pino A, Balasubramanian S, Wyns L, Gazit E, De Greve H, Magnuson RD, Charlier D, van Nuland NAJ, & Loris R (2010) Allosteric and Intrinsic Disorder Mediate Transcription Regulation by Conditional Cooperativity. *Cell* 142(1):101-111.
98. Monti MC, Hernández-Arriaga AM, Kamphuis MB, López-Villarejo J, Heck AJR, Boelens R, Díaz-Orejas R, & van den Heuvel RHH (2007) Interactions of Kid–Kis toxin–antitoxin complexes with the parD operator-promoter region of plasmid R1 are piloted by the Kis antitoxin and tuned by the stoichiometry of Kid–Kis oligomers. *Nucleic Acids Research* 35(5):1737-1749.
99. Kasari V, Mets T, Tenson T, & Kaldalu N (2013) Transcriptional cross-activation between toxin-antitoxin systems of *Escherichia coli*. *BMC Microbiology* 13(1):45.
100. Haurlyuk V, Atkinson GC, Murakami KS, Tenson T, & Gerdes K (2015) Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nature Reviews Microbiology* 13:298.
101. Potrykus K & Cashel M (2008) (p)ppGpp: Still Magical? *Annual Review of Microbiology* 62(1):35-51.
102. Keasling JD, Bertsch L, & Kornberg A (1993) Guanosine pentaphosphate phosphohydrolase of *Escherichia coli* is a long-chain exopolyphosphatase. *Proceedings of the National Academy of Sciences* 90(15):7029-7033.
103. Usheer K, Koji O, & A. HW (2012) Direct binding targets of the stringent response alarmone (p)ppGpp. *Molecular Microbiology* 85(6):1029-1043.
104. Wu N, He L, Cui P, Wang W, Yuan Y, Liu S, Xu T, Zhang S, Wu J, Zhang W, & Zhang Y (2015) Ranking of persister genes in the same *Escherichia coli* genetic background demonstrates varying importance of individual persister genes in tolerance to different antibiotics. *Frontiers in Microbiology* 6:1003.
105. Darija V, Tsuneko O, Keiji M, Heni S, Shizuo K, Katsuhiko H, & Yoichiro M (2006) Functional Analysis of spoT, relA and dksA Genes on Quinolone Tolerance in *Pseudomonas aeruginosa* under Nongrowing Condition. *Microbiology and Immunology* 50(4):349-357.
106. Corrigan RM, Bellows LE, Wood A, & Gründling A (2016) ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria. *Proceedings of the National Academy of Sciences* 113(12):E1710-E1719.
107. Dörr T, Lewis K, & Vulić M (2009) SOS Response Induces Persistence to Fluoroquinolones in *Escherichia coli*. *PLOS Genetics* 5(12):e1000760.
108. Keren I, Shah D, Spoering A, Kaldalu N, & Lewis K (2004) Specialized Persister Cells and the Mechanism of Multidrug Tolerance in *Escherichia coli*. *Journal of Bacteriology* 186(24):8172-8180.
109. Harms A, Maisonneuve E, & Gerdes K (2016) Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science* 354(6318).
110. Dörr T, Vulić M, & Lewis K (2010) Ciprofloxacin Causes Persister Formation by Inducing the TisB toxin in *Escherichia coli*. *PLOS Biology* 8(2):e1000317.
111. McKay SL & Portnoy DA (2015) Ribosome Hibernation Facilitates Tolerance of Stationary-Phase Bacteria to Aminoglycosides. *Antimicrobial Agents and Chemotherapy* 59(11):6992-6999.
112. Amato Stephanie M & Brynildsen Mark P (2015) Persister Heterogeneity Arising from a Single Metabolic Stress. *Current Biology* 25(16):2090-2098.
113. Pu Y, Zhao Z, Li Y, Zou J, Ma Q, Zhao Y, Ke Y, Zhu Y, Chen H, Baker Matthew AB, Ge H, Sun Y, Xie Xiaoliang S, & Bai F (2016) Enhanced Efflux Activity Facilitates Drug Tolerance in Dormant Bacterial Cells. *Molecular Cell* 62(2):284-294.
114. Mok WWK, Park JO, Rabinowitz JD, & Brynildsen MP (2015) RNA Futile Cycling in Model Persisters Derived from MazF Accumulation. *mBio* 6(6).
115. Orman MA & Brynildsen MP (2013) Dormancy Is Not Necessary or Sufficient for Bacterial Persistence. *Antimicrobial Agents and Chemotherapy* 57(7):3230-3239.

References

116. Cataudella I, Trusina A, Sneppen K, Gerdes K, & Mitarai N (2012) Conditional cooperativity in toxin–antitoxin regulation prevents random toxin activation and promotes fast translational recovery. *Nucleic Acids Research* 40(14):6424-6434.
117. Bahar AA, Liu Z, Totsingan F, Buitrago C, Kallenbach N, & Ren D (2015) Synthetic dendrimeric peptide active against biofilm and persister cells of *Pseudomonas aeruginosa*. *Applied Microbiology and Biotechnology* 99(19):8125-8135.
118. Kwan BW, Chowdhury N, & Wood TK (2015) Combatting bacterial infections by killing persister cells with mitomycin C. *Environmental Microbiology* 17(11):4406-4414.
119. Li T, Yin N, Liu H, Pei J, & Lai L (2016) Novel Inhibitors of Toxin HipA Reduce Multidrug Tolerant Persisters. *ACS Medicinal Chemistry Letters* 7(5):449-453.
120. Barbosa LCB, Garrido SS, Garcia A, Delfino DB, Santos LdN, & Marchetto R (2012) Design and synthesis of peptides from bacterial ParE toxin as inhibitors of topoisomerases. *European Journal of Medicinal Chemistry* 54:591-596.
121. Allison KR, Brynildsen MP, & Collins JJ (2011) Metabolite-Enabled Eradication of Bacterial Persisters by Aminoglycosides. *Nature* 473(7346):216-220.
122. Marques CNH, Morozov A, Planzos P, & Zelaya HM (2014) The Fatty Acid Signaling Molecule cis-2-Decenoic Acid Increases Metabolic Activity and Reverts Persister Cells to an Antimicrobial-Susceptible State. *Applied and Environmental Microbiology* 80(22):6976-6991.
123. Conlon BP, Nakayasu ES, Fleck LE, LaFleur MD, Isabella VM, Coleman K, Leonard SN, Smith RD, Adkins JN, & Lewis K (2013) Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature* 503:365.
124. Aebersold R & Mann M (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* 537:347.
125. de Godoy LMF, Olsen JV, Cox J, Nielsen ML, Hubner NC, Fröhlich F, Walther TC, & Mann M (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455:1251.
126. Hebert AS, Richards AL, Bailey DJ, Ulbrich A, Coughlin EE, Westphall MS, & Coon JJ (2014) The One Hour Yeast Proteome. *Molecular & Cellular Proteomics* 13(1):339-347.
127. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese J-H, Bantscheff M, Gerstmair A, Faerber F, & Kuster B (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509:582.
128. Kim M-S, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabudhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LDN, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K, Chatterjee A, Huang T-C, Zhong J, Wu X, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TSK, Hruban RH, Kerr CL, Bader GD, Iacobuzio-Donahue CA, Gowda H, & Pandey A (2014) A draft map of the human proteome. *Nature* 509:575.
129. Bekker-Jensen DB, Kelstrup CD, Batth TS, Larsen SC, Haldrup C, Bramsen JB, Sørensen KD, Høyer S, Ørntoft TF, Andersen CL, Nielsen ML, & Olsen JV (2017) An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell Systems* 4(6):587-599.e584.
130. Schubert OT, Röst HL, Collins BC, Rosenberger G, & Aebersold R (2017) Quantitative proteomics: challenges and opportunities in basic and applied research. *Nature Protocols* 12:1289.
131. Steen H & Mann M (2004) The ABC's (and XYZ's) of peptide sequencing. *Nature Reviews Molecular Cell Biology* 5:699.

References

132. Catherman AD, Skinner OS, & Kelleher NL (2014) Top Down Proteomics: Facts and Perspectives. *Biochemical and biophysical research communications* 445(4):683-693.
133. Zhang H & Ge Y (2011) Comprehensive Analysis of Protein Modifications by Top-Down Mass Spectrometry. *Circulation: Cardiovascular Genetics* 4(6):711-711.
134. Sidoli S & Garcia BA (2017) Middle-down proteomics: a still unexploited resource for chromatin biology. *Expert Review of Proteomics* 14(7):617-626.
135. Mann M (2012) Proteomics for biomedicine: a half-completed journey. *EMBO Molecular Medicine* 4(2):75-77.
136. Zhang Y, Fonslow BR, Shan B, Baek M-C, & Yates JR (2013) Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* 113(4):2343-2394.
137. Rosenfeld J, Capdevielle J, Guillemot JC, & Ferrara P (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Analytical Biochemistry* 203(1):173-179.
138. Ritorto MS, Cook K, Tyagi K, Pedrioli PGA, & Trost M (2013) Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *Journal of Proteome Research* 12(6):2449-2457.
139. Yang F, Shen Y, Camp DG, & Smith RD (2012) High pH reversed-phase chromatography with fraction concatenation as an alternative to strong-cation exchange chromatography for two-dimensional proteomic analysis. *Expert Review of Proteomics* 9(2):129-134.
140. Batth TS, Francavilla C, & Olsen JV (2014) Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics. *Journal of Proteome Research* 13(12):6176-6186.
141. MacNair JE, Lewis KC, & Jorgenson JW (1997) Ultrahigh-Pressure Reversed-Phase Liquid Chromatography in Packed Capillary Columns. *Analytical Chemistry* 69(6):983-989.
142. Thakur SS, Geiger T, Chatterjee B, Bandilla P, Fröhlich F, Cox J, & Mann M (2011) Deep and Highly Sensitive Proteome Coverage by LC-MS/MS Without Prefractionation. *Molecular & Cellular Proteomics* 10(8).
143. Pedro RC (2005) Principles of Nanoflow Liquid Chromatography and Applications to Proteomics. *Current Nanoscience* 1(1):65-71.
144. Fenn J, Mann M, Meng C, Wong S, & Whitehouse C (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246(4926):64-71.
145. Koichi T, Hiroaki W, Yutaka I, Satoshi A, Yoshikazu Y, Tamio Y, & T. M (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* 2(8):151-153.
146. Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, & Horning S (2006) Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Analytical Chemistry* 78(7):2113-2120.
147. Kelstrup CD, Jersie-Christensen RR, Batth TS, Arrey TN, Kuehn A, Kellmann M, & Olsen JV (2014) Rapid and Deep Proteomes by Faster Sequencing on a Benchtop Quadrupole Ultra-High-Field Orbitrap Mass Spectrometer. *Journal of Proteome Research* 13(12):6187-6195.
148. Kelstrup CD, Bekker-Jensen DB, Arrey TN, Högrefe A, Harder A, & Olsen JV (2018) Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *Journal of Proteome Research* 17(1):727-738.
149. Eliuk S & Makarov A (2015) Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Review of Analytical Chemistry* 8(1):61-80.
150. Kalli A, Smith GT, Sweredoski MJ, & Hess S (2013) Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers. *Journal of Proteome Research* 12(7):3071-3086.
151. Scheltema RA, Hauschild J-P, Lange O, Hornburg D, Denisov E, Damoc E, Kuehn A, Makarov A, & Mann M (2014) The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High-performance Quadrupole and an Ultra-high-field Orbitrap Analyzer. *Molecular & Cellular Proteomics* 13(12):3698-3708.

References

152. Diedrich JK, Pinto AFM, & Yates JR (2013) Energy dependence of HCD on peptide fragmentation: Stepped collisional energy finds the sweet spot. *Journal of the American Society for Mass Spectrometry* 24(11):10.1007/s13361-13013-10709-13367.
153. Huang T-Y, Emory JF, O'Hair RAJ, & McLuckey SA (2006) Electron-Transfer Reagent Anion Formation via Electrospray Ionization and Collision-Induced Dissociation. *Analytical Chemistry* 78(21):7387-7391.
154. Frese CK, Zhou H, Taus T, Altelaar AFM, Mechtler K, Heck AJR, & Mohammed S (2013) Unambiguous Phosphosite Localization using Electron-Transfer/Higher-Energy Collision Dissociation (EThcD). *Journal of Proteome Research* 12(3):1520-1525.
155. Hebert AS, Thöing C, Riley NM, Kwiecien NW, Shiskova E, Huguet R, Cardasis HL, Kuehn A, Eliuk S, Zabrouskov V, Westphall MS, McAlister GC, & Coon JJ (2018) Improved Precursor Characterization for Data-Dependent Mass Spectrometry. *Analytical Chemistry* 90(3):2333-2340.
156. Fernandez-Lima F, Kaplan DA, Suetering J, & Park MA (2011) Gas-phase separation using a trapped ion mobility spectrometer. *International journal for ion mobility spectrometry : official publication of the International Society for Ion Mobility Spectrometry* 14(2-3):10.1007/s12127-12011-10067-12128.
157. Meier F, Beck S, Grassl N, Lubeck M, Park MA, Raether O, & Mann M (2015) Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *Journal of Proteome Research* 14(12):5378-5387.
158. Meier F, Brunner A-D, Koch S, Koch H, Lubeck M, Krause M, Goedecke N, Decker J, Kosinski T, Park MA, Bache N, Hoerning O, Cox J, Raether O, & Mann M (2018) Online parallel accumulation – serial fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer. *bioRxiv*.
159. Macek B, Mann M, & Olsen JV (2009) Global and Site-Specific Quantitative Phosphoproteomics: Principles and Applications. *Annual Review of Pharmacology and Toxicology* 49(1):199-221.
160. Venable JD, Dong M-Q, Wohlschlegel J, Dillin A, & Yates Iii JR (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature Methods* 1:39.
161. Hu A, Noble WS, & Wolf-Yadlin A (2016) Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research* 5:F1000 Faculty Rev-1419.
162. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, & Aebersold R (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Molecular & Cellular Proteomics* 11(6).
163. Peckner R, Myers SA, Jacome ASV, Egertson JD, Abelin JG, MacCoss MJ, Carr SA, & Jaffe JD (2018) Specter: linear deconvolution for targeted analysis of data-independent acquisition mass spectrometry proteomics. *Nature Methods* 15:371.
164. Picotti P & Aebersold R (2012) Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods* 9:555.
165. Brusniak M-YK, Kwok S-T, Christiansen M, Campbell D, Reiter L, Picotti P, Kusebauch U, Ramos H, Deutsch EW, Chen J, Moritz RL, & Aebersold R (2011) ATAQS: A computational software tool for high throughput transition optimization and validation for selected reaction monitoring mass spectrometry. *BMC Bioinformatics* 12(1):78.
166. Lange V, Picotti P, Domon B, & Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular Systems Biology* 4:222-222.
167. Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, & Domon B (2012) Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* 11(12):1709-1723.

References

168. Peterson AC, Russell JD, Bailey DJ, Westphall MS, & Coon JJ (2012) Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Molecular & Cellular Proteomics* 11(11):1475-1488.
169. Cox J & Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26:1367.
170. Cox J & Mann M (2009) Computational Principles of Determining and Improving Mass Precision and Accuracy for Proteome Measurements in an Orbitrap. *Journal of the American Society for Mass Spectrometry* 20(8):1477-1485.
171. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, & Mann M (2011) Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *Journal of Proteome Research* 10(4):1794-1805.
172. Käll L, Storey JD, MacCoss MJ, & Noble WS (2008) Posterior Error Probabilities and False Discovery Rates: Two Sides of the Same Coin. *Journal of Proteome Research* 7(1):40-44.
173. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, & Mann M (2002) Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics* 1(5):376-386.
174. Soufi B, Krug K, Harst A, & Macek B (2015) Characterization of the E. coli proteome and its modifications during growth and ethanol stress. *Frontiers in Microbiology* 6(103).
175. Spät P, Mačák B, & Forchhammer K (2015) Phosphoproteome of the cyanobacterium *Synechocystis* sp. PCC 6803 and its dynamics during nitrogen starvation. *Frontiers in Microbiology* 6:248.
176. Ravikumar V, Shi L, Krug K, Derouiche A, Jers C, Cousin C, Kobir A, Mijakovic I, & Macek B (2014) Quantitative Phosphoproteome Analysis of *Bacillus subtilis* Reveals Novel Substrates of the Kinase PrkC and Phosphatase PrpC. *Molecular & Cellular Proteomics* 13(8):1965-1978.
177. Bendall SC, Hughes C, Stewart MH, Doble B, Bhatia M, & Lajoie GA (2008) Prevention of Amino Acid Conversion in SILAC Experiments with Embryonic Stem Cells. *Molecular & Cellular Proteomics* 7(9):1587-1597.
178. Soufi B & Macek B (2014) Stable isotope labeling by amino acids applied to bacterial cell culture. *Methods Mol Biol* 1188:9-22.
179. Krüger M, Moser M, Ussar S, Thievensen I, Lubber CA, Forner F, Schmidt S, Zanivan S, Fässler R, & Mann M (2008) SILAC Mouse for Quantitative Proteomics Uncover Kindlin-3 as an Essential Factor for Red Blood Cell Function. *Cell* 134(2):353-364.
180. Westman-Brinkmalm A, Abramsson A, Pannee J, Gang C, Gustavsson MK, von Otter M, Blennow K, Brinkmalm G, Heumann H, & Zetterberg H (2011) SILAC zebrafish for quantitative analysis of protein turnover and tissue regeneration. *Journal of Proteomics* 75(2):425-434.
181. Geiger T, Wisniewski JR, Cox J, Zanivan S, Krüger M, Ishihama Y, & Mann M (2011) Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nature Protocols* 6:147.
182. Shenoy A & Geiger T (2015) Super-SILAC: current trends and future perspectives. *Expert Review of Proteomics* 12(1):13-19.
183. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, & Pappin DJ (2004) Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics* 3(12):1154-1169.
184. Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, & Hamon C (2003) Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Analytical Chemistry* 75(8):1895-1904.
185. Savitski MM, Mathieson T, Zinn N, Sweetman G, Doce C, Becher I, Pachi F, Kuster B, & Bantscheff M (2013) Measuring and Managing Ratio Compression for Accurate iTRAQ/TMT Quantification. *Journal of Proteome Research* 12(8):3586-3598.

References

186. Ting L, Rad R, Gygi SP, & Haas W (2011) MS3 eliminates ratio distortion in isobaric labeling-based multiplexed quantitative proteomics. *Nature methods* 8(11):937-940.
187. Virreira Winter S, Meier F, Wichmann C, Cox J, Mann M, & Meissner F (2018) EASI-tag enables accurate multiplexed and interference-free MS2-based proteome quantification. *Nature Methods*.
188. Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, & Heck AJR (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols* 4:484.
189. Lundgren DH, Hwang S-I, Wu L, & Han DK (2010) Role of spectral counting in quantitative proteomics. *Expert Review of Proteomics* 7(1):39-53.
190. Andersen JS, Wilkinson CJ, Mayor T, Mortensen P, Nigg EA, & Mann M (2003) Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* 426:570.
191. Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, & Mann M (2014) Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & Cellular Proteomics : MCP* 13(9):2513-2526.
192. Riley NM & Coon JJ (2016) Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. *Analytical Chemistry* 88(1):74-94.
193. Gerber SA, Rush J, Stemman O, Kirschner MW, & Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences* 100(12):6940-6945.
194. Kaiser SE, Riley BE, Shaler TA, Trevino RS, Becker CH, Schulman H, & Kopito RR (2011) Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nature Methods* 8:691.
195. Hanke S, Besir H, Oesterhelt D, & Mann M (2008) Absolute SILAC for Accurate Quantitation of Proteins in Complex Mixtures Down to the Attomole Level. *Journal of Proteome Research* 7(3):1118-1130.
196. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, & Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* 473:337.
197. Olsen JV & Mann M (2013) Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry. *Molecular & Cellular Proteomics* 12(12):3444-3452.
198. Zhao Y & Jensen ON (2009) Modification-specific proteomics: Strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 9(20):4632-4641.
199. von Stechow L, Francavilla C, & Olsen JV (2015) Recent findings and technological advances in phosphoproteomics for cells and tissues. *Expert Review of Proteomics* 12(5):469-487.
200. Hendriks IA & Vertegaal ACO (2016) A comprehensive compilation of SUMO proteomics. *Nature Reviews Molecular Cell Biology* 17:581.
201. Akhter Y & Thakur S (2017) Targets of ubiquitin like system in mycobacteria and related actinobacterial species. *Microbiological Research* 204:9-29.
202. Kasper E-K & R. LM (2013) Technologies and challenges in large-scale phosphoproteomics. *PROTEOMICS* 13(6):910-931.
203. Batalha IL, Lowe CR, & Roque ACA (2012) Platforms for enrichment of phosphorylated proteins and peptides in proteomics. *Trends in Biotechnology* 30(2):100-110.
204. Sharma K, D'Souza Rochelle CJ, Tyanova S, Schaab C, Wiśniewski Jacek R, Cox J, & Mann M (2014) Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Reports* 8(5):1583-1594.
205. Batth TS, Papetti M, Pfeiffer A, Tollenaere MAX, Francavilla C, & Olsen JV (2018) Large-Scale Phosphoproteomics Reveals Shp-2 Phosphatase-Dependent Regulators of Pdgf Receptor Signaling. *Cell Reports* 22(10):2784-2796.
206. Zarei M, Sprenger A, Rackiewicz M, & Dengjel J (2015) Fast and easy phosphopeptide fractionation by combinatorial ERLIC-SCX solid-phase extraction for in-depth phosphoproteome analysis. *Nature Protocols* 11:37.

References

207. Arrington JV, Hsu C-C, Elder SG, & Andy Tao W (2017) Recent advances in phosphoproteomics and application to neurological diseases. *Analyst* 142(23):4373-4387.
208. Annan RS & Carr SA (1996) Phosphopeptide Analysis by Matrix-Assisted Laser Desorption Time-of-Flight Mass Spectrometry. *Analytical Chemistry* 68(19):3413-3421.
209. Olsen JV, Macek B, Lange O, Makarov A, Horning S, & Mann M (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* 4:709.
210. Bateman RH, Carruthers R, Hoyes JB, Jones C, Langridge JI, Millar A, & Vissers JPC (2002) A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. *Journal of the American Society for Mass Spectrometry* 13(7):792-803.
211. Mijakovic I & Macek B (2012) Impact of phosphoproteomics on studies of bacterial physiology. *FEMS Microbiology Reviews* 36(4):877-892.
212. Deutscher J & Saier Jr MH (2005) Ser/Thr/Tyr Protein Phosphorylation in Bacteria – For Long Time Neglected, Now Well Established. *Journal of Molecular Microbiology and Biotechnology* 9(3-4):125-131.
213. Grangeasse C, Nessler S, & Mijakovic I (2012) Bacterial tyrosine kinases: evolution, biological function and structural insights. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1602):2640-2655.
214. Mitrophanov AY & Groisman EA (2008) Signal integration in bacterial two-component regulatory systems. *Genes & Development* 22(19):2601-2611.
215. Kannan N, Taylor SS, Zhai Y, Venter JC, & Manning G (2007) Structural and Functional Diversity of the Microbial Kinome. *PLoS Biology* 5(3):e17.
216. Macek B, Mijakovic I, Olsen JV, Gnad F, Kumar C, Jensen PR, & Mann M (2007) The Serine/Threonine/Tyrosine Phosphoproteome of the Model Bacterium *Bacillus subtilis*. *Molecular & Cellular Proteomics* 6(4):697-707.
217. Kalantari A, Derouiche A, Shi L, & Mijakovic I (2015) Serine/threonine/tyrosine phosphorylation regulates DNA binding of bacterial transcriptional regulators. *Microbiology* 161(9):1720-1729.
218. Hanks S, Quinn A, & Hunter T (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241(4861):42-52.
219. Pereira SFF, Goss L, & Dworkin J (2011) Eukaryote-Like Serine/Threonine Kinases and Phosphatases in Bacteria. *Microbiology and Molecular Biology Reviews : MMBR* 75(1):192-212.
220. Shah IM, Laaberki M-H, Popham DL, & Dworkin J (2008) A Eukaryotic-like Ser/Thr Kinase Signals Bacteria to Exit Dormancy in Response to Peptidoglycan Fragments. *Cell* 135(3):486-496.
221. Jers C, Kobir A, Søndergaard EO, Jensen PR, & Mijakovic I (2011) *Bacillus subtilis* Two-Component System Sensory Kinase DegS Is Regulated by Serine Phosphorylation in Its Input Domain. *PLoS ONE* 6(2):e14653.
222. Jimin Z, Chunhua H, Kumar SV, L. MN, & Zongchao J (2007) Crystal structure of a novel prokaryotic Ser/Thr kinase and its implication in the Cpx stress response pathway. *Molecular Microbiology* 63(5):1360-1371.
223. Macek B, Gnad F, Soufi B, Kumar C, Olsen JV, Mijakovic I, & Mann M (2008) Phosphoproteome Analysis of *E. coli* Reveals Evolutionary Conservation of Bacterial Ser/Thr/Tyr Phosphorylation. *Molecular & Cellular Proteomics* 7(2):299-307.
224. Soares NC, Spät P, Krug K, & Macek B (2013) Global Dynamics of the *Escherichia coli* Proteome and Phosphoproteome During Growth in Minimal Medium. *Journal of Proteome Research* 12(6):2611-2621.
225. Lin M-H, Sugiyama N, & Ishihama Y (2015) Systematic profiling of the bacterial phosphoproteome reveals bacterium-specific features of phosphorylation. *Science Signaling* 8(394):rs10-rs10.

References

226. Potel CM, Lin M-H, Heck AJR, & Lemeer S (2018) Widespread bacterial protein histidine phosphorylation revealed by mass spectrometry-based proteomics. *Nature Methods*.
227. Prisc S, Dankwa S, Schwartz D, Chou MF, Locasale JW, Kang C-M, Bemis G, Church GM, Steen H, & Husson RN (2010) Extensive phosphorylation with overlapping specificity by Mycobacterium tuberculosis serine/threonine protein kinases. *Proceedings of the National Academy of Sciences* 107(16):7521-7526.
228. Fortuin S, Tomazella GG, Nagaraj N, Sampson SL, Gey van Pittius NC, Soares NC, Wiker HG, de Souza GA, & Warren RM (2015) Phosphoproteomics analysis of a clinical Mycobacterium tuberculosis Beijing isolate: expanding the mycobacterial phosphoproteome catalog. *Frontiers in Microbiology* 6:6.
229. Verma R, Pinto SM, Patil AH, Advani J, Subba P, Kumar M, Sharma J, Dey G, Ravikumar R, Buggi S, Satishchandra P, Sharma K, Suar M, Tripathy SP, Chauhan DS, Gowda H, Pandey A, Gandotra S, & Prasad TSK (2017) Quantitative Proteomic and Phosphoproteomic Analysis of H37Ra and H37Rv Strains of Mycobacterium tuberculosis. *Journal of Proteome Research* 16(4):1632-1645.
230. Mijakovic I & Deutscher J (2015) Protein-tyrosine phosphorylation in Bacillus subtilis: a 10-year retrospective. *Frontiers in Microbiology* 6:18.
231. Shi L, Pigeonneau N, Ravikumar V, Dobrinic P, Macek B, Franjevic D, Noirot-Gros M-F, & Mijakovic I (2014) Cross-phosphorylation of bacterial serine/threonine and tyrosine protein kinases on key regulatory residues. *Frontiers in Microbiology* 5(495).
232. Elsholz AKW, Turgay K, Michalik S, Hessling B, Gronau K, Oertel D, Mäder U, Bernhardt J, Becher D, Hecker M, & Gerth U (2012) Global impact of protein arginine phosphorylation on the physiology of Bacillus subtilis. *Proceedings of the National Academy of Sciences* 109(19):7451-7456.
233. Trentini DB, Suskiewicz MJ, Heuck A, Kurzbauer R, Deszcz L, Mechtler K, & Clausen T (2016) Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature* 539:48.
234. White A (1942) The Dynamic State of Body Constituents. *The Yale Journal of Biology and Medicine* 14(6):677-677.
235. Claydon AJ & Beynon R (2012) Proteome Dynamics: Revisiting Turnover with a Global Perspective. *Molecular & Cellular Proteomics* 11(12):1551-1565.
236. Pratt JM, Petty J, Riba-Garcia I, Robertson DHL, Gaskell SJ, Oliver SG, & Beynon RJ (2002) Dynamics of Protein Turnover, a Missing Dimension in Proteomics. *Molecular & Cellular Proteomics* 1(8):579-591.
237. Qingbo L (2010) Advances in protein turnover analysis at the global level and biological insights. *Mass Spectrometry Reviews* 29(5):717-736.
238. Hinkson IV & Elias JE (2011) The dynamic state of protein turnover: It's about time. *Trends in Cell Biology* 21(5):293-303.
239. Hughes C & Krijgsveld J (2012) Developments in quantitative mass spectrometry for the analysis of proteome dynamics. *Trends in Biotechnology* 30(12):668-676.
240. Larrabee KL, Phillips JO, Williams GJ, & Larrabee AR (1980) The relative rates of protein synthesis and degradation in a growing culture of Escherichia coli. *Journal of Biological Chemistry* 255(9):4125-4130.
241. Mosteller RD, Goldstein RV, & Nishimoto KR (1980) Metabolism of individual proteins in exponentially growing Escherichia coli. *Journal of Biological Chemistry* 255(6):2524-2532.
242. Obrig TG, Culp WJ, McKeenan WL, & Hardesty B (1971) The Mechanism by which Cycloheximide and Related Glutarimide Antibiotics Inhibit Peptide Synthesis on Reticulocyte Ribosomes. *Journal of Biological Chemistry* 246(1):174-181.
243. Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, & Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* 425:737.
244. Yen H-CS, Xu Q, Chou DM, Zhao Z, & Elledge SJ (2008) Global Protein Stability Profiling in Mammalian Cells. *Science* 322(5903):918-923.

References

245. Khmelinskii A, Keller PJ, Bartosik A, Meurer M, Barry JD, Mardin BR, Kaufmann A, Trautmann S, Wachsmuth M, Pereira G, Huber W, Schiebel E, & Knop M (2012) Tandem fluorescent protein timers for in vivo analysis of protein dynamics. *Nature Biotechnology* 30:708.
246. Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, Garber M, Gnirke A, Nusbaum C, Hacohen N, Friedman N, Amit I, & Regev A (2011) Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nature Biotechnology* 29:436.
247. Doherty MK, Hammond DE, Clague MJ, Gaskell SJ, & Beynon RJ (2009) Turnover of the Human Proteome: Determination of Protein Intracellular Stability by Dynamic SILAC. *Journal of Proteome Research* 8(1):104-112.
248. Björn S, Manfred G, Gunnar D, & Matthias S (2009) Global analysis of cellular protein translation by pulsed SILAC. *PROTEOMICS* 9(1):205-209.
249. Welle KA, Zhang T, Hryhorenko JR, Shen S, Qu J, & Ghaemmaghami S (2016) Time-resolved Analysis of Proteome Dynamics by Tandem Mass Tags and Stable Isotope Labeling in Cell Culture (TMT-SILAC) Hyperplexing. *Molecular & Cellular Proteomics* 15(12):3551-3563.
250. Zecha J, Meng C, Zolg DP, Samarasinghe P, Wilhelm M, & Kuster B (2018) Peptide Level Turnover Measurements Enable the Study of Proteoform Dynamics. *Molecular & Cellular Proteomics* 17(5):974-992.
251. Maier T, Schmidt A, Güell M, Kühner S, Gavin AC, Aebersold R, & Serrano L (2011) Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Molecular Systems Biology* 7(1).
252. Jayapal KP, Sui S, Philp RJ, Kok Y-J, Yap MGS, Griffin TJ, & Hu W-S (2010) Multitagging Proteomic Strategy to Estimate Protein Turnover Rates in Dynamic Systems. *Journal of Proteome Research* 9(5):2087-2097.
253. Zhe W, Qiang-Qiang H, Mao-Tian Z, Xi C, & Lin G (2016) Protein turnover analysis in *Salmonella Typhimurium* during infection by dynamic SILAC, Topograph, and quantitative proteomics. *Journal of Basic Microbiology* 56(7):801-811.
254. Chua SL, Yam JKH, Hao P, Adav SS, Salido MM, Liu Y, Givskov M, Sze SK, Tolker-Nielsen T, & Yang L (2016) Selective labelling and eradication of antibiotic-tolerant bacterial populations in *Pseudomonas aeruginosa* biofilms. *Nature Communications* 7:10750.
255. Dieterich DC, Lee JJ, Link AJ, Graumann J, Tirrell DA, & Schuman EM (2007) Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nature Protocols* 2:532.
256. Kramer G, Sprenger RR, Back J, Dekker HL, Nessen MA, van Maarseveen JH, de Koning LJ, Hellingwerf KJ, de Jong L, & de Koster CG (2009) Identification and Quantitation of Newly Synthesized Proteins in *Escherichia coli* by Enrichment of Azidohomoalanine-labeled Peptides with Diagonal Chromatography. *Molecular & Cellular Proteomics* 8(7):1599-1611.
257. Guyer M, Reed R, Steitz J, & Low K (1981) Identification of a sex-factor-affinity site in *E. coli* as $\gamma\delta$. *Cold Spring Harbor symposia on quantitative biology*, (Cold Spring Harbor Laboratory Press), pp 135-140.
258. Xiao H, Kalman M, Ikehara K, Zemel S, Glaser G, & Cashel M (1991) Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. *Journal of Biological Chemistry* 266(9):5980-5990.
259. Guzman LM, Belin D, Carson MJ, & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology* 177(14):4121-4130.
260. Semanjski M, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B (2018) The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance. *Science Signaling* 11(547).

References

261. Ringquist S, Shinedling S, Barrick D, Green L, Binkley J, Stormo GD, & Gold L (1992) Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Molecular microbiology* 6(9):1219-1229.
262. Link AJ, Phillips D, & Church GM (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *Journal of Bacteriology* 179(20):6228-6237.
263. Tabone T, Mather DE, & Hayden MJ (2009) Temperature Switch PCR (TSP): Robust assay design for reliable amplification and genotyping of SNPs. *BMC Genomics* 10(1):580.
264. Rappsilber J, Mann M, & Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* 2:1896.
265. Elias JE & Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* 4:207.
266. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, & Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* 13:731.
267. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA, Brunak S, & Mann M (2010) Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Science Signaling* 3(104):ra3-ra3.
268. Ernst J & Bar-Joseph Z (2006) STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* 7(1):191.
269. Huang DW, Sherman BT, & Lempicki RA (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* 4:44.
270. Schwartz D & Gygi SP (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nature Biotechnology* 23:1391.
271. Uzan M & Danchin A (1976) A rapid test for the rel A mutation in *E. coli*. *Biochemical and biophysical research communications* 69(3):751-758.
272. Schmitt E, Panvert M, Blanquet S, & Mechulam Y (1995) Transition state stabilization by the 'high' motif of class I aminoacyl-tRNA synthetases: the case of *Escherichia coli* methionyl-tRNA synthetase. *Nucleic Acids Research* 23(23):4793-4798.
273. Loveland AB, Bah E, Madireddy R, Zhang Y, Brilot AF, Grigorieff N, & Korostelev AA (2016) Ribosome•RelA structures reveal the mechanism of stringent response activation. *eLife* 5:e17029.
274. Klemm P, Hjerrild L, Gjermansen M, & Schembri MA (2004) Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from *Escherichia coli*. *Molecular Microbiology* 51(1):283-296.
275. Maeda Y, Lin C-Y, Ishida Y, Inouye M, Yamaguchi Y, & Phadtare S (2017) Characterization of YjjJ toxin of *Escherichia coli*. *FEMS Microbiology Letters* 364(11):fnx086-fnx086.
276. Brissette JL, Weiner L, Ripmaster TL, & Model P (1991) Characterization and sequence of the *Escherichia coli* stress-induced *psp* operon. *Journal of Molecular Biology* 220(1):35-48.
277. von Delft F, Lewendon A, Dhanaraj V, Blundell TL, Abell C, & Smith AG (2001) The Crystal Structure of *E. coli* Pantothenate Synthetase Confirms It as a Member of the Cytidyltransferase Superfamily. *Structure* 9(5):439-450.
278. Denby KJ, Iwig J, Bisson C, Westwood J, Rolfe MD, Sedelnikova SE, Higgins K, Maroney MJ, Baker PJ, Chivers PT, & Green J (2016) The mechanism of a formaldehyde-sensing transcriptional regulator. *Scientific Reports* 6:38879.
279. Kaczanowska M & Rydén-Aulin M (2007) Ribosome Biogenesis and the Translation Process in *Escherichia coli*. *Microbiology and Molecular Biology Reviews* 71(3):477-494.
280. Lu M, Campbell JL, Boye E, & Kleckner N (1994) SeqA: a negative modulator of replication initiation in *E. coli*. *Cell* 77(3):413-426.

References

281. Guarné A, Brendler T, Zhao Q, Ghirlando R, Austin S, & Yang W (2005) Crystal structure of a SeqA–N filament: implications for DNA replication and chromosome organization. *The EMBO Journal* 24(8):1502-1511.
282. Pedersen IB, Helgesen E, Flåtten I, Fossum-Raunehaug S, & Skarstad K (2017) SeqA structures behind Escherichia coli replication forks affect replication elongation and restart mechanisms. *Nucleic Acids Research* 45(11):6471-6485.
283. Guo X-P & Sun Y-C (2017) New Insights into the Non-orthodox Two Component Rcs Phosphorelay System. *Frontiers in Microbiology* 8:2014.
284. Hoang L, Fredrick K, & Noller HF (2004) Creating ribosomes with an all-RNA 30S subunit P site. *Proceedings of the National Academy of Sciences of the United States of America* 101(34):12439-12443.
285. Agrawal RK, Linde J, Sengupta J, Nierhaus KH, & Frank J (2001) Localization of L11 protein on the ribosome and elucidation of its involvement in EF-G-dependent translocation. *Journal of Molecular Biology* 311(4):777-787.
286. Hansen CH, Endres RG, & Wingreen NS (2008) Chemotaxis in Escherichia coli: A Molecular Model for Robust Precise Adaptation. *PLOS Computational Biology* 4(1):e1.
287. Barnhart MM & Chapman MR (2006) Curli Biogenesis and Function. *Annual Review of Microbiology* 60(1):131-147.
288. Alexandre G (2015) Chemotaxis Control of Transient Cell Aggregation. *Journal of Bacteriology* 197(20):3230-3237.
289. Zolkiewski M (1999) ClpB Cooperates with DnaK, DnaJ, and GrpE in Suppressing Protein Aggregation: a novel multi-chaperone system from Escherichia coli. *Journal of Biological Chemistry* 274(40):28083-28086.
290. Kuczynska-Wisnik D, Kędzierska S, Matuszewska E, Lund P, Taylor A, Lipinska B, & Laskowska E (2002) The Escherichia coli small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock. *Microbiology* 148(6):1757-1765.
291. Ulett GC, Valle J, Beloin C, Sherlock O, Ghigo J-M, & Schembri MA (2007) Functional Analysis of Antigen 43 in Uropathogenic Escherichia coli Reveals a Role in Long-Term Persistence in the Urinary Tract. *Infection and Immunity* 75(7):3233-3244.
292. Claydon AJ & Beynon RJ (2011) Protein Turnover Methods in Single-Celled Organisms: Dynamic SILAC. *Yeast Systems Biology: Methods and Protocols*, eds Castrillo JI & Oliver SG (Humana Press, Totowa, NJ), pp 179-195.
293. Cañas-Duarte SJ, Restrepo S, & Pedraza JM (2014) Novel Protocol for Persister Cells Isolation. *PLOS ONE* 9(2):e88660.
294. Lobritz MA, Belenky P, Porter CBM, Gutierrez A, Yang JH, Schwarz EG, Dwyer DJ, Khalil AS, & Collins JJ (2015) Antibiotic efficacy is linked to bacterial cellular respiration. *Proceedings of the National Academy of Sciences* 112(27):8173-8180.
295. Arfin SM, Long AD, Ito ET, Toller L, Riehle MM, Paegle ES, & Hatfield GW (2000) Global Gene Expression Profiling in Escherichia coli K12: The effects of intergenation host factor. *Journal of Biological Chemistry* 275(38):29672-29684.
296. Liu Y, Xie P, Wang P, Li M, Li H, Li W, & Dou S (2015) A model for chromosome organization during the cell cycle in live E. coli. *Scientific Reports* 5:17133.
297. Zhang X-S, García-Contreras R, & Wood TK (2007) YcfR (BhsA) Influences Escherichia coli Biofilm Formation through Stress Response and Surface Hydrophobicity. *Journal of Bacteriology* 189(8):3051-3062.
298. Lloyd LJ, Jones SE, Jovanovic G, Gyaneshwar P, Rolfe MD, Thompson A, Hinton JC, & Buck M (2004) Identification of a New Member of the Phage Shock Protein Response in Escherichia coli, the Phage Shock Protein G (PspG). *Journal of Biological Chemistry* 279(53):55707-55714.

References

299. Beloin C, Valle J, Latour-Lambert P, Faure P, Kzreminski M, Balestrino D, Haagensen JAJ, Molin S, Prensier G, Arbeille B, & Ghigo J-M (2004) Global impact of mature biofilm lifestyle on *Escherichia coli* K-12 gene expression. *Molecular Microbiology* 51(3):659-674.
300. Eriksson S, Lucchini S, Thompson A, Rhen M, & Hinton JCD (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Molecular Microbiology* 47(1):103-118.
301. Kobayashi R, Suzuki T, & Yoshida M (2007) *Escherichia coli* phage-shock protein A (PspA) binds to membrane phospholipids and repairs proton leakage of the damaged membranes. *Molecular Microbiology* 66(1):100-109.
302. Etchegaray J-P & Inouye M (1999) CspA, CspB, and CspG, Major Cold Shock Proteins of *Escherichia coli*, Are Induced at Low Temperature under Conditions That Completely Block Protein Synthesis. *Journal of Bacteriology* 181(6):1827-1830.
303. Bae W, Xia B, Inouye M, & Severinov K (2000) *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proceedings of the National Academy of Sciences* 97(14):7784-7789.
304. Weisberg RA & Gottesman ME (1999) Processive Antitermination. *Journal of Bacteriology* 181(2):359-367.
305. Maciąg A, Peano C, Pietrelli A, Egli T, De Bellis G, & Landini P (2011) In vitro transcription profiling of the σ S subunit of bacterial RNA polymerase: re-definition of the σ S regulon and identification of σ S -specific promoter sequence elements. *Nucleic Acids Research* 39(13):5338-5355.
306. Xia B, Ke H, Shinde U, & Inouye M (2003) The Role of RbfA in 16S rRNA Processing and Cell Growth at Low Temperature in *Escherichia coli*. *Journal of Molecular Biology* 332(3):575-584.
307. Maki Y, Yoshida H, & Wada A (2000) Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase *Escherichia coli*. *Genes to Cells* 5(12):965-974.
308. Janosi L, Mottagui-Tabar S, Isaksson LA, Sekine Y, Ohtsubo E, Zhang S, Goon S, Nelken S, Shuda M, & Kaji A (1998) Evidence for in vivo ribosome recycling, the fourth step in protein biosynthesis. *The EMBO Journal* 17(4):1141-1151.
309. Jöers A, Kaldalu N, & Tenson T (2010) The Frequency of Persisters in *Escherichia coli* Reflects the Kinetics of Awakening from Dormancy. *Journal of Bacteriology* 192(13):3379-3384.
310. Varik V, Oliveira SRA, Haurlyuk V, & Tenson T (2016) Composition of the outgrowth medium modulates wake-up kinetics and ampicillin sensitivity of stringent and relaxed *Escherichia coli*. *Scientific Reports* 6:22308.
311. Miura K, Tomioka Y, Hoshi Y, Suzuki H, Yonezawa M, Hishinuma T, & Mizugaki M (1997) The effects of unsaturated fatty acids, oxidizing agents and Michael reaction acceptors on the induction of N-ethylmaleimide reductase in *Escherichia coli*: possible application for drug design of chemoprotectors. *Methods Find Exp Clin Pharmacol* 19(3):147-151.
312. Umezawa Y, Shimada T, Kori A, Yamada K, & Ishihama A (2008) The Uncharacterized Transcription Factor YdhM Is the Regulator of the *nemA* Gene, Encoding N-Ethylmaleimide Reductase. *Journal of Bacteriology* 190(17):5890-5897.
313. Ollagnier-de-Choudens S, Mattioli T, Takahashi Y, & Fontecave M (2001) Iron-Sulfur Cluster Assembly: Characterization of IscA and evidence for a specific and functional complex with ferredoxin. *Journal of Biological Chemistry* 276(25):22604-22607.
314. Klotz A, Georg J, Bučinská L, Watanabe S, Reimann V, Januszewski W, Sobotka R, Jendrossek D, Hess Wolfgang R, & Forchhammer K (2016) Awakening of a Dormant Cyanobacterium from Nitrogen Chlorosis Reveals a Genetically Determined Program. *Current Biology* 26(21):2862-2872.

8 List of Abbreviations

°C	Degree Celsius
ABC	Ammonium bicarbonate
ACN	Acetonitrile
AHA	Azidohomoalanine
APD	Advance Peak Determination
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
C18	Octadecyl carbon chain
CFU	Colony forming units
CID	Collision-induced dissociation
DDA	Data-dependent acquisition
DHB	2,5-dihydroxybenzoic acid
DNA	Deoxyribonucleic acid
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
FA	Formic acid
FACS	Fluorescence-activated cell sorting
FDR	False discovery rate
GFP	Green fluorescent protein
GO	Gene-ontology
GTP	Guanosine triphosphate
HCD	Higher energy collisional dissociation
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His ₆	Polyhistidine-tag
HpH	High pH
HPLC	High performance liquid chromatography
IAA	Iodoacetamide
IMAC	Metal ion affinity chromatography
IT	Injection time
iTRAQ	Isobaric tags for relative and absolute quantification
KEGG	Kyoto encyclopedia of genes and genomes
LC	Liquid chromatography
LTD	Linear trap quadrupole
OD _{600nm}	Optical density at 600 nm
M	Molar (mol/l)
m/z	Mass to charge ratio

List of Abbreviations

MALDI	Matrix-assisted laser absorption ionization
MDK	Minimum duration for killing
MeOH	Methanol
MIC	Minimum inhibitory concentration
mRNA	Messenger ribonucleic acid
MOAC	Metal oxide affinity chromatography
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
PEP	Posterior error probability
ppm	Parts per million
PTM	Posttranslational modification
RNA	Ribonucleic acid
RP	Reverse-phase
RT	Room temperature
rRNA	Ribosomal ribonucleic acid
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SILAC	Stable isotope labeling by amino acids in cell culture
SPE	Solid phase extraction
TA	Toxin-antitoxin
TAP	Tandem affinity purification
TCA	Tricarboxylic acid
TFA	Trifluoroacetic acid
TiO ₂	Titanium dioxide
T _m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
UHPLC	Ultra-high performance liquid chromatography
v/v	Volume to volume
w/v	Weight per volume
WT	Wild-type

9 List of Publications

1. **Semanjski M**, Germain E, Bratl K, Kiessling A, Gerdes K, & Macek B
The kinases HipA and HipA7 phosphorylate different substrate pools in *Escherichia coli* to promote multidrug tolerance
Science Signaling, 2018, 11(547)
2. **Semanjski M**, & Macek B
Shotgun proteomics of bacterial pathogens: advances, challenges and clinical implications
Expert review of proteomics, 2016, 13(2), 139-156.
3. Cveticic N*, **Semanjski M***, Soufi B, Krug K, Gruic-Sovulj I, & Macek B
Proteome-wide measurement of non-canonical bacterial mistranslation by quantitative mass spectrometry of protein modifications
Scientific reports, 2016, 6, 28631.
*Equal contribution
4. Soufi B, Täumer C, **Semanjski M**, & Macek B
Phosphopeptide enrichment from bacterial samples utilizing titanium oxide affinity chromatography
Microbial proteomics: Methods and protocols, ed Becher D (Springer New York, New York, NY), 2018, 231-247.
5. Bilus M*, **Semanjski M***, Mocibob M*, Cveticic N, Zivkovic I, Tawfik DS, Toth-Petroczy A, Macek B, & Gruic-Sovulj I
On the mechanism and origin of isoleucyl-tRNA synthetase editing against norvaline
(Manuscript submitted)
*Equal contribution
6. Zhou B, **Semanjski M**, Orlovetskie N, Bhattacharya S, Alon S, Jarrous N, Zhang Y, Macek B, Sinai L, & Ben-Yehuda S
Arginine dephosphorylation propels spore germination in bacteria
(Manuscript submitted)

10 Curriculum Vitae

Personal Information

Name	Maja Semanjski
Date of Birth	18.01.1990
Place of Birth	Slavonski Brod, Croatia
Nationality	Croatian

Education

2014 - 2018	PhD Student in Biochemistry, Proteome Center Tübingen, Interfaculty Institute for Cell Biology, University of Tübingen, Germany
2011 - 2013	Master of Chemistry, University of Zagreb, Croatia Thesis: "Cloning and purification of amino acid:[carrier protein] ligase family representatives"
2008 - 2011	Bachelor of Chemistry, University of Zagreb, Croatia
2004 - 2008	High School for Science and Mathematics "Matija Mesic", Slavonski Brod, Croatia

11 Acknowledgements

First, I would like to thank my supervisor Prof. Boris Maček for introducing me the fantastic world of proteomics and for believing that I can compete with other students that had better scientific background. I am very thankful for the great mentoring, advices and support, and especially for the opportunity to explore the world through so many conferences and scientific visits. I would also like to thank Prof. Doron Rapaport for being the second supervisor of the thesis.

I would like to thank Prof. Kenn Gerdes for a wonderful collaboration and for allowing me to work in his laboratory. Firstly, many thanks to Elsa for a great collaboration, for all the help and advices she gave me. Thanks to Stine and other Copenhagen lab members who welcomed me to their lab. Additionally, thanks to all the members of the research training group RTG 1708.

Many warmest thanks to my lab members for making PCT feel like home. Firstly, to Christoph for all German support and stupid jokes, to Marisa for being a good friend, to Nicolas for all the laughter and to Katarina for enormous help. Special thanks to the old crew members Karsten, Vaishnavi, Bo and Carpy for the best welcome to Germany. Also, many thanks to the new crew Tariq, Philipp and Alen, and my students Andreas, Kathrin, Jonas and Viktor. Special thanks to Nevena whose advices helped me to be a better scientist, and together with Ela to bring me a peace of Croatia when I was homesick. Also, many thanks to other PCT members, Uli, Mirita, Irina, Ana, Silke, Johannes and Jürgen.

Many thanks to my friend Maša for making Tübingen feel a little bit like home. Also, many thanks to my Croatian friends back home, especially to Monika and Tina.

Finally, I would like to thank my family - my mom, my dad and my brother for their greatest support without which this journey would not be possible. Mama i tata, neizmjerno vam hvala.

The greatest thank you goes to my Antonio for being my “oslonac i putokaz” through all of these years – he was my rock. Hvala ti.

