# Mass spectrometry based proteomics approaches unraveling dynamic processes in the entire bacterial cell

Inauguraldissertation

zur

Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer.nat.)

an der Mathematisch-Naturwissenschaftlichen Fakultät

der

Ernst- Moritz- Arndt- Universität Greifswald

vorgelegt von

Andreas Willy Otto

geboren am 21.06.1979

in Kiel

Greifswald, den 6.12.2010

Dekan:           Prof. Dr. K. Fesser

1. Gutachter      Prof. Dr. M. Hecker

2. Gutachter      Prof. Dr. J. M. van Dijl

Tag der Promotion:   11.02.2011

_____

# *Content*

_____

**Published Manuscripts**

_____

# *Summary*

Deciphering the entire protein complement of a living cell together with the elucidation of dynamic processes on protein level are the main goals of proteomics as it is used today. To achieve this goal, namely the elucidation of dynamic processes of the entire bacterial cell, we have developed strategies and distinct workflows to cover the most proteins in different subcellular localizations in bacteria together with a stable isotopes labeling approach to follow temporal and spatial changes in different proteomic subfractions.

The first study contributing to this ambitious goal was a study that mainly pointed out to a comparison of gel-free and gel-based techniques on deciphering changes in the proteome of heat stressed cells of *Bacillus subtilis*. Subcellular fractionation was performed on exponential growing *B. subtilis* cells to get access to the cytosolic and the enriched membrane fraction. It became clear that aside from the classical 2D gel based quantitative proteomic approach, gel-free techniques together with alternative labeling methods (in this case based on *in vitro* chemical labeling by iTRAQ (isotope tags for relative and absolute quantitation)) were promising in getting qualitative and quantitative access to proteins that were not accessible before (Wolff et al. 2006).

As it could be stated in the successive review, having improved the accessibility of certain protein groups by gel free techniques (allowing the determination of hydrophobic, small and low abundant proteins), identification of 34% of the theoretical proteome of growing *B. subtilis* cells was feasible now. By application of different subcellular fractionation methods and either gel based and gel free proteomics techniques the qualitative and quantitative description of the entire protein complement came into reach as pointed out by Wolff and coworkers 2007.

During the thesis, it turned out that *in vitro* labeling techniques were not equally suited for cytosolic and other subcellular fractions in *B. subtilis* proteomics. Therefore, an important prerequisite for further determination of the entire protein inventory and its changes during different physiological states was the establishment of an in vivo labeling method that was compatible with the proteomics workflows that have been used before. We could prove the applicability of *in vivo* $^{15}$N-metabolic labeling in a study comparing both SILAC (stable isotope labeling with amino acids) and $^{15}$N-metabolic labeling (Dreisbach et al. 2008). The combination of GeLCMS and in vivo labeling techniques showed up to be easily applicable for cytosolic and membrane proteomics giving reproducible results allowing comprehensive and simultaneous studies on multiple cellular sub fractions. Altogether we could raise the number of integral membrane proteins to be identified to about 25% of the predicted integral membrane proteome. Furthermore, the membrane enrichment approach is very specific: 60% of the identified proteins belonged either to the group of integral membrane proteins or to proteins functionally/structurally related to the membrane. Comparing both labeling techniques, the study revealed the metabolic labeling being superior to SILAC in terms of 50% more quantified proteins most probably due to the advantage of a complete labeling for the $^{15}$N- labeling approach.

_____

*In vivo* $^{15}$N-metabolic labeling and proteome analysis of the membrane proteome was accordingly applied in studies on *E. coli* (Price et al. 2010) and *B. subtilis* (Hyyryläinen et al. 2010 and Saller et al. 2010). In these studies we were able to address defined biological questions with the methodologies that were set up before. Furthermore, a study on the surface proteome in *Staphylococcus aureus* proved the *in vivo* $^{15}$N-labeling technique to be powerful even for surface exposed proteins in a study by Hempel et al.. (2010)

Having paved the way for comprehensive studies on multiple subcellular fractions, the $^{15}$N-metabolic labeling technique has been used in proteomics studies targeting complete proteomes and the changes within. In a study on a comparison of growing and non growing *S. aureus* Col, it could be shown that a coverage of 65% of the proteome in *S. aureus* Col is feasible by subcellular fractionation and the analysis of five different subproteomes (Becher et al. 2009).

Besides the comparison of two physiological situations, application of $^{15}$N-metabolic labeling enables for the visualization of dynamic processes in a cell. This was achieved for five time points with concurrent sampling of transcriptome data in a study on osmotic upshift in *B. subtilis* revealing strict time dependent transcription of stress related ECF- sigma factor dependent genes and its correspondence in the membrane proteome (Hahne et al. 2010).

Finally, in a study on glucose starvation in *B. subtilis*, a most comprehensive view on the dynamics covering five time points over five subcellular localizations combining proteomic, transcriptomic and metabolomic data was accomplished. Comparable in proteome coverage to the study in *S. aureus*, 53% of the proteome complement could be identified serving in the global quantitative study as the basis for an unmatched global overlook on processes taking place in the starved cell (Otto et al. 2010).

To conclude, it has been shown that the use of mass spectrometry based in vivo quantitation techniques and the application of subcellular and chromatographic fractionation has lead to a new level of qualitative and quantitative proteomics data. Emphasizing on the studies revealing the dynamics of the bacterial physiology on a time resolved base, both spatial and temporal processes can be monitored to obtain knowledge on physiological processes in a depth that has not been reached before in comparable global studies.

_____

# *Zusammenfassung*

Das Ziel heutiger Proteomstudien liegt in der Aufklärung des Proteinbestands ganzer Zellen und der Verfolgung dynamischer Prozesse auf der Proteinebene. In der vorliegenden Arbeit wurden zu diesem Zweck Strategien und zielgerichtete Arbeitsabläufe entwickelt, um eine möglichst hohe Abdeckung des theoretisch exprimierten Proteoms in verschiedenen subzellulären Lokalisationen zu erreichen. Weiterhin wurde die Methode der Proteinmarkierung mit stabilen Isotopen etabliert, um zeitliche und örtliche Veränderungen in verschiedenen proteomischen Unterfraktionen zu verfolgen.

Eine erste Studie dazu war der Vergleich von gelfreien und gelbasierten Analysetechniken in exponentiell wachsenden Zellen von *Bacillus subtilis*, in der durch Hitzestress hervorgerufene Veränderungen im Proteom untersucht wurden. Um sowohl das Cytosol als auch die Membranfraktion zu analysieren, wurden hier beide Subproteome über subzelluläre Fraktionierung voneinander getrennt. Im Ergebnis wurde deutlich, dass die Kombination des klassischen, 2D Gel basierten, quantitativen Proteomansatzes mit gelfreien Techniken, wie der *in vitro* iTRAQ-Markierung (iTRAQ - isotope tags for relative and absolute quantitation), einen experimentellen Ansatz bieten, um qualitativ und quantitativ eine Vielzahl von Proteinen beschreiben zu können, welche zuvor nicht erfassbar waren (Wolff et al. 2006).

In der darauf folgenden Übersichtsarbeit konnte dargestellt werden, dass durch die analytisch neu zu erreichenden Proteingruppen (hydrophobe, kleine und niedrig abundante Proteine) eine Proteomabdeckung von 34% in *B. subtilis* erreicht werden konnte. Weiterhin wurde klar, dass durch die Kombination von subzellulären Fraktionierungsmethoden die Möglichkeit theoretisch gegeben ist, die Gesamtheit aller Proteine einer Zelle qualitativ und quantitativ zu analysieren (Wolff et al 2007). Während der Durchführung der vorliegenden Dissertationsarbeit stellte sich heraus, dass *in vitro* Markierungstechniken nicht gleichermaßen für cytosolische und andere subzellulären Fraktionen in Proteomstudien zu *B. subtilis* geeignet sind. Somit wurde als Voraussetzung für eine umfassende Zusammenstellung von Proteombibliotheken und der Erfassung ihrer Dynamik die Etablierung einer *in vivo* Markierungsmethode notwendig, welche möglichst mit bereits vorhandenen Analyseprotokollen kompatibel sein sollte. Die Vorteile der Anwendung von *in vivo* Markierungsmethoden, basierend auf dem Einbau von schwerem $^{15}$N Stickstoff oder von Aminosäuren mit schwerem Isotopen (SILAC) während der Proteinsynthese,, wurde in der folgenden Studie von Dreisbach et al. (2008) gezeigt. Eine Kombination der GeLCMS Technik (1D SDS PAGE kombiniert mit LC-MS Analysen) mit der *in vivo* Markierung erwies sich als zielführend für Analyse von cytosolischen und Membranproteinen. Gekennzeichnet ist diese Methode durch die hohe Reproduzierbarkeit der Ergebnisse und die Möglichkeit, umfassende Untersuchungen gleichzeitig in unterschiedlichen subzellulären Fraktionen durchzuführen. Zusammengenommen konnte die Zahl der identifizierten Membranproteine auf etwa 25% des theoretisch vorhergesagten Membranproteoms gesteigert werden. Weiterhin kann festgestellt werden, dass der experimentelle Ansatz der

_____

Vorfraktionierung einzelner Subproteome zu fraktionsspezifischen Ergebnissen führt: So gehörten 60% der identifizierten Proteine entweder zu den integralen Membranproteinen oder sind funktioneller oder struktureller Teil der Membran. Der Vergleich der unterschiedlichen metabolischen Markierungsmethoden mit $^{15}$N oder SILAC zeigte, dass in über die $^{15}$N Markierung eine signifikant höhere Anzahl an Proteinen quantifiziert wurde. Dies ist mit hoher Wahrscheinlichkeit auf die vollständige Markierung aller Aminosäuren eines Proteins während des $^{15}$N Ansatzes zurückzuführen.

Im Folgenden wurden *in vivo* $^{15}$N Markierungsansätze auf Membranproteomstudien in *E. coli* (Price et al. 2010) und *B. subtilis* (Hyyryläinen et al. 2010 und Saller et al. 2010) angewandt. In diesen Studien wurden die erarbeiteten Methoden eingesetzt, um definierte biologische Fragestellungen zu beantworten. Weiterhin konnte die $^{15}$N Markierung als Quantifizierungsgrundlage auch in einer Analyse des Oberflächenproteoms von *Staphyloccocus aureus* erfolgreich eingesetzt werden (Hempel et al. 2010).

Der so ausgereifte Ansatz - metabolische $^{15}$N Markierung in Kombination mit subzellulären Fraktionierungstechniken - wurde hiernach in umfangreichen Studien, die verschiedene subzelluläre Fraktionen umfassen, verwendet.

In einer vergleichenden Studie von wachsenden und nicht- wachsenden *S. aureus* COL Zellen konnte gezeigt werden, dass durch die Verwendung von subzellulären Fraktionierungsmethoden und die Analyse der fünf resultierenden Subproteome eine Abdeckung von 65% des Gesamtproteoms erreichbar ist (Becher et al. 2009).

Neben dem direkten Vergleich von physiologisch unterschiedlichen Zellzuständen kann die $^{15}$N Markierung auch dazu verwendet werden, dynamische Veränderungen in der Zelle sichtbar zu machen. So wurde diese Methode in einer Proteom und Transkriptom basierten Studie der Zellantwort von *B. subtilis* auf osmotischen Stress für die Analyse von fünf Zeitpunkten vor und während der Stressadaptation angewendet. Hierbei wurden insbesondere für ECF Faktor abhängige Gene und ihren Entsprechungen im Membranproteom eng begrenzte und zeitabhängige Veränderungen gefunden (Hahne et al. 2010).

Schließlich konnte in einer Studie von *B. subtilis* ein sehr umfängliches Bild der stattfindenen dynamischen Prozesse des durch Glucosehunger induzierten Überganges von wachsenden Zellen in die stationäre Phase gezeichnet werden. Hierzu wurden proteomische Daten von fünf verschiedenen subzellulären Lokalisationen, transkriptomische und metabolomische Daten zusammengefasst. Es konnte so, vergleichbar zu der *S. aureus* Studie von Becher et al., 53% des Gesamtproteoms identifiziert werden. Dies diente als solide Basis für einen umfangreichen Überblick über Prozesse in der hungernden Zelle (Otto et al. 2010).

Zusammenfassend konnte in dieser Arbeit gezeigt werden, dass die Anwendung Massenspektrometrie basierter *in vivo* Markierungstechniken und subzellulärer/ chromatographischer Fraktionierungsmethoden zu einer neuen Güte von qualitativen und quantitativen Proteomdaten geführt hat. Insbesondere in den proteomischen Studien, welche die zeitaufgelöste Dynamik der bakteriellen Physiologie zum Gegenstand hatten, und welche örtliche und zeitliche Veränderungen (subzelluläre Lokalisation; Zunahme/ Abnahme) verfolgten, konnte ein *bis dato* unerreichter Umfang an Informationen in physiologischen Prozessen erzielt werden.

_____

## *Introduction*

## *1. Proteomics*

Understanding processes of life has been the incitement of biological and biomedical research since the beginning of modern times. Today, broad knowledge on key molecular and cellular processes of life at different levels has been acquired. To achieve this goal, sophisticated methodologies are invented and applied to gain an even more comprehensive understanding on cell physiology. For a complete description of biological systems, information on DNA, RNA, proteins and involved metabolites are integrated to understand the interplay of the entities that allow the complex and dynamic functioning of life.

Proteomics depicts the experimental approach that determines the entire compliment of proteins expressed by a cell at a point in time. Within proteomic studies, it is possible to either study proteins qualitatively (identification, distribution, post translational modifications, interactions, structure and function of proteins) or quantitatively (abundance, distribution within different localizations, degradation/synthesis) [1].

This scientific field had its beginning as early as 1975 starting with first proteomic studies accomplished by using the revolutionary technique of two dimensional SDS polyacrylamide gel electrophoresis (2D SDS PAGE) [2]. This was 20 years before the first genome was sequenced and marked a starting point to a whole new field of protein analytics – long before the term 'proteomics' was coined in the year 1994 [3]. Today, progress led this field to a highly diversified ground of a mass spectrometry aided or even dominated scientific domain [4].

Proteomics in general is technically challenging: several thousands of proteins are found even in rather simple prokaryotic organisms ranging up to ten thousands of proteins in a eukaryotic cell [5]. Analysis of inherently complex proteomic samples requires techniques that are of effective of separating proteins according to their physical properties or biological function.

### *1.1.     2D GE based proteomics*

O'Farrell and Klose introduced the landmark technology of high resolution two dimensional gel electrophoresis (2D GE) and laid the ground for simultaneous separation of hundreds of proteins that was accomplished by a highly orthogonal and

_____

reproducible method [2, 6]. In two dimensional gel electrophoresis proteins are separated according to two different characteristics: in the first dimension, today consisting of immobilized pH gradients (IPG) having replaced the initially used carrier ampholytes [7-9], the denatured proteins are separated according to their isoelectric point (p*I*). In the second dimension, a separation step via SDS PAGE is implemented depending on the molecular weight (MW) of the protein [10]. With 2D GE it is possible to separate 2000 proteins on a single gel with up to 20 gels in parallel [10, 11] allowing for a global view on cellular activity with high resolution and high sensitivity. Advantageous of 2D GE is the possibility of a global view on most main cellular functions on protein level that makes this method a basis for a plethora of high- throughput comparative studies [12]. Within these studies it is possible to examine the physiological state and alterations imposed by nutrient starvation and physical stresses of known and even of unknown organisms. Shifts in p*I* may be a hint for phosphorylation events that can be analyzed further providing the stability of the phosphorylation under examination [12, 13]. Further examples are glycosylation assays [14] or even limited proteolysis that can be visualized by multiple spots with differing molecular weight on the 2D gel [15]. An important advantage of the 2D gel based proteomics approaches is the possibility to display protein synthesis vs. accumulated proteins in a fast and efficient way compared to mass spectrometry based approaches by pulse chase experiments [12].

Despite other technologies that came up, later termed "second generation proteomics"[16], 2D GE still remains a pivotal methodology in proteomics that can be routinely applied for parallel quantitative expression profiling.

## 1.2.    *Mass spectrometry based proteomics*

The rapid development of proteomics is technology driven [10]. Proteomics perceived a steep rise in the late 1970s by invention of the 2D GE and despite the major improvements introduced with IPGs in the first dimension a decline was seen in the late 1980s through the inability to identify the gel separated proteins on a large scale. This gap in the analytical process was closed in the last decade of the millennium.

Until the early 1990s, mass spectrometry on proteins and peptides was rather seldom [16] and classical protein biochemistry techniques like Edman sequencing [17] and SDS PAGE [18] for determination of molecular weight were used to identify protein species. Publications dealing with the concept of peptide mass fingerprinting (PMF) [19-22] and

_____

identification of peptides by pattern comparison of CID spectra [23-25] together with EST databases [26, 27] (built up with the data produced by fast, partial cDNA sequencing) paved the road to a successful application of mass spectrometry based proteomics. Finally in 1995, release of the first complete sequenced genome of *Hemophilus influenzae* was groundbreaking for the new era of functional genomics [28]. New techniques allowing high throughput identification of proteins together with sophisticated quantitation methods relying on different isotopes define the second generation of proteomics- a mass spectrometry based proteomics era [4].

### *1.2.1. Technological and conceptual prerequisites for success of mass spectrometry in life sciences*

Mass spectrometry in general is an analytical technique of determination of the mass of a charged molecule. Individual molecules are "weighed" by transforming them into ions in a vacuum and then measuring the response of their trajectories to electric and magnetic fields or both [29]. The success of mass spectrometry in the field of life science as we experience today was made possible by two major developments both on the technical side and on the side of data evaluation [30].

Invention of soft ionization methods like ESI (electrospray ionization) [29] and MALDI (matrix assisted laser desorption/ionization) [31] revolutionized the analysis of ions in the gas phase previously limited for heat stable analytes and made polypeptides and even proteins accessible for mass spectrometry. The technical and conceptual advances in this area led to the 2002 Nobel prize awarded to John Fenn and Koichi Tanaka (amongst others) for the development of methods for identification and structure analyses of biological macromolecules.

Another trigger for the success of mass spectrometry based proteomics was the availability of large databases with genomic information, represented first by EST databases and since 1995 by complete genome sequences.

In principle, a mass spectrometer consists of an ion source, a mass analyzer and a detector. The ion source is the prerequisite of ionizing a sample so that the analyte is brought into gas phase. ESI and MALDI, most used ion sources to volatize and ionize proteins and peptides for mass spectrometric analyses are suited differently for different types of analytes (Fig.1). ESI ionizes analytes out of a mobile phase, typically in

_____

combination with liquid chromatography (LC). In opposition to that, MALDI sublimates and ionizes the analyte out of a crystalline matrix with laser pulses in a vacuum [32].

The mass analyzer subsequently measures the ratio of the mass over charge (m/z) of the particular analyte. To add quantitative information to the process, the detector registers the number of ions of a specific m/z ratio [4].

Additionally to the basic analytical power of mass spectrometry to identify peptides and proteins from biological samples introduced in the analytical workflow, the advent of high resolution and high mass accurate mass spectrometers in the field of proteomics had a tremendous impact in quantitative proteomics. Methodologies that utilize isotope labeling techniques by *in vivo* metabolic labeling or *in vitro* chemical/ enzymatic labeling inherently depend on high mass accuracy and resolution in order to get valid results. Furthermore, label free quantitation approaches not only depend on exact masses to distinguish between near- isobaric masses but also on highly reliable liquid chromatography systems for a differential analysis .

_____
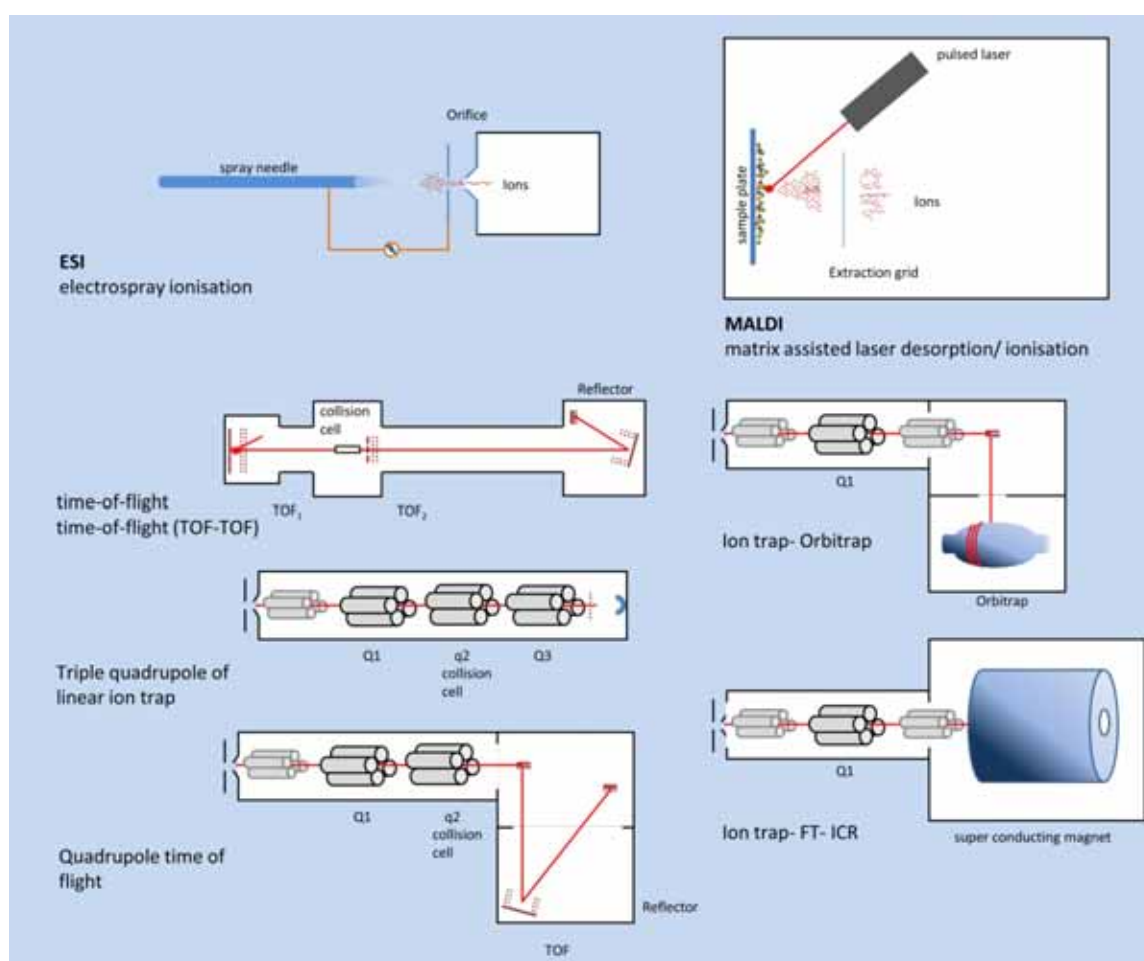


**Fig.1 Ion sources and most common mass analyzer used in proteomic research (adapted from Aebersold et al. [4])**

Mass analyzer used in mass spectrometry are central to the technology. Key parameters of judgment over a specific instrument are its limit of detection (sensitivity), the ability to resolve two adjacent masses and hence nearly isobaric peptide species with different composition (resolution) as well as the mass accuracy and the ability to generate information rich spectra from ion fragments enabling for highly specific database searches with the decisive assignment of fragment spectra to theoretical fragment patterns.

Mass analyzer in proteomics are typically of the ion trap (IT), time of flight (ToF), quadrupole (Q) or Fourier Transform- MS (FT- MS) (Fourier Transform- Ion cyclotron resonance or Fourier Transform Orbitrap) type. The different kinds of analyzers are either used alone or in tandem configuration.

_____

In time of flight (ToF) instruments, the ions are accelerated to high kinetic energy and are separated along a flight tube as a result of their different velocities. ToF- instruments are either linear mode instruments or reflectron mode instruments. By the use of an electro/ nic mirror, the reflector, the ions are guided towards the detector simultaneously being compensated for slight differences in kinetic energy [33] and therefore enhancing the resolution markedly. ToF-ToF instruments are tandem configurations of two flight tubes with an added collision cell to be able to generate fragment ions of particular masses (Fig. 1).

Quadrupole mass analyzer consist of four rods that generate time- varying electric fields which permit a stable trajectory only for ions of a particular desired m/z. Quadrupoles are used as linear ion traps [34] with a two dimensional quadrupole field or as quadrupole ion trap [35] that work as three dimensional traps or as collision cells. In triple Quadrupole mass analyzers, ions are first selected in the quadrupole that is used as mass filter (Q1), then the ions are fragmented in the second (q2) and then the fragments are separated in the third quadrupole (Q3) prior to detection of ion intensities (fig 1).

Qq- ToF instruments are tandem instruments combining the mass filter/collision features of two quadrupoles with the mass accuracy of a TOF instrument (Fig 1). Qq- ToF instruments can be regarded as a triple quadrupole instrument where the third quadrupole is replaced by an orthogonal ToF.

Fourier transform mass spectrometry instruments (FT MS) are basically of the ion trap type that trap ions with the help of strong magnetic fields. The FT- Orbitrap [36] and the FT- ICR [37] are of unsurpassed high mass accuracy, highly sensitive and of high dynamic range. Both are coupled to a linear ion trap (LTQ) as tandem instruments.

As pointed out before, proteomics is highly technology driven and therefore the development of new types of mass spectrometric instrumentation is influencing research and vice versa. Today's instrumentation is mostly driven by the need for fast and sensitive analyzers that should either be highly accurate in determination of masses (LTQ FT- ICR or LTQ Orbitrap/ LTQ Orbitrap Velos), yield high throughput analyses with high mass accuracy (MALDI ToF-ToF) or that enable for specific modes of acquisition like triple Quadrupole instruments enabling for targeted analyses (MRM) or more specific instrumentation adding a gas- phase separation via ion mobility.

_____

## *2. Towards the comprehensive proteomic description of the gram positive model organism Bacillus subtilis*

Since the beginning of proteomics, bacteria were intensively studied according to their relatively low complexity in comparison to higher organisms. The sequencing of genomes beginning with important model organisms and bacterial pathogens started in 1995 with the sequencing of *Haemophilus influenzae* [28] and is increasing nowadays exponentially through the up come of next generation DNA sequencing techniques [38]. Meanwhile, it is clear, that even relatively small genomes (e.g. *Mycoplasma pneumoniae*) are bearing all information necessary for cell growth and proliferation to form the protein complement of a living cell. Therefore even low complexity organisms make reasonable model organisms to address crucial and basic problems of processes determining life by using proteomics [12].

*Bacillus subtilis* as the second most intensively studied bacterium beside *Escherichia coli* [39] is serving as a model organism for gram positive bacteria being assigned to this role through both inherent biological as well as economic features. As from the beginning of the 1950s, the possibility to study basic cellular processes (gene regulation and metabolism) as well as cell differentiation (this organism is capable of forming spores in times of less favorable conditions) together with the possibility of genetic manipulation led to the importance of this bacterium in the research community. Furthermore, *B. subtilis* is of economic importance as an industrial production strain owing to its secretion capacities [40].

The proteomics based research aiding in understanding of this organism is still under way despite this long time of research. As reviewed in Wolff et al. (2007), until today one third of the proteome is still of unknown function and parts of the proteome is still inaccessible for experimental or technical reasons [41]. A plethora of subfractionation techniques on cellular level have been implemented in *Bacillus* research in order to get to a most comprehensive coverage of the *B. subtilis* proteome. In the following, the relevance and accessibility of *B. subtilis* subfractions is described that will lead towards a comprehensive proteomic description when taken together.

_____

## 2.1. Cytosolic proteins

In *B. subtilis*, the soluble proteins are the most studied species of proteins by means of accessibility and importance for the main metabolic pathways. These proteins are ideally targeted by classical 2D GE approaches enabling for a display of complete physiological pathways. Cytosolic proteins often have to be depleted in other subcellular fractionation approaches based on their high abundance compared to proteins residing e.g. in the cell membrane. Cytosolic proteins are mostly addressed by analysis of the cell crude extract or in the supernatant after depletion of the cell membranes.

## 2.2. Extracellular proteins

Especially for *B. subtilis* being able to excrete a large number of proteins even on an industrial scale [40], the extracellular proteins are of scientific and economic interest. *Hirose* et al. introduced a method of getting access to proteins being secreted by *B. subtilis* during growth by analysis of these proteins in the growth medium [42]. Generally, proteins that are secreted into the medium are precipitated by acidification and subsequent centrifugation. As trichloroacetic acid is used, intensive washing has to be conducted. Special attention has to be paid to cytosolic contamination in the extracellular fraction obtained by acid depletion: lysed cells release their intracellular contents and the high abundant cytosolic proteins are found in the extracellular fraction. With new mass spectrometry based proteomic approaches with high sensitivity, this poses a major disadvantage on deciphering the "genuine" extracellular proteome. The phenomenon of released cell content is discussed to have a biological meaning: in cells of stationary phase *B. subtilis*, cell death is induced in a fraction of the population [43] and cannibalism is taking place [44].

## 2.3. Membrane proteins

The membrane proteome represents one of the most difficult subproteomes to get access to in terms of physical properties and the low abundance of the proteins [45]. Proteomic analyses of membrane proteins of *B. subtilis* developed over a period of time. First attempts were carried out by Bunai et al. to map the membrane bound solute binding proteins of ABC transporter [46]. The authors used different detergents to tackle the hydrophobic membrane proteins and adapted protocols of 2D GE to get access to the hydrophobic fraction of the proteome. Turning away from the classical 2D GE inherently

_____

unsuitable for very hydrophobic proteins, A. Dreisbach in Eymann et al. proposed a membrane protein enrichment protocol that consists of different steps [47]. Cell membranes obtained by ultracentrifugation are washed consecutively with high salt buffers and carbonate buffers with high pH to deplete cytosolic contaminants. Finally, membrane proteins are solubilized with the detergent *n*-dodecyl-β-D-maltoside and separated via a 1D SDS PAGE. This method proved to be very successful to deplete high abundant soluble proteins with a concurrent relative enrichment of proteins with transmembrane domains and proteins being attached to the membrane [41, 47, 48]. Wolff and coworkers then came up with a protocol in *S. aureus* being highly complementary to the membrane protein enrichment protocol: membrane proteins with a higher number of transmembrane domains could be identified by mass spectrometry giving this field a new dimension by means of accessibility [49]. Hahne et al. managed to transfer this method to *B. subtilis* [50]. In the two preceding studies it was shown that the membrane enrichment protocol and the newly introduced protocol by Wolff et al. yield highly complementary results. The first is predominantly targeting the soluble loops and domains of membrane spanning proteins whereas the latter gets hold of the transmembrane helices.

## 2.4. Surface proteins

The surface of a cell is most important for communication and interaction with its environment. In *B. subtilis,* this subproteome can be addressed by high salt extraction protocols as shown in a study aimed at the elucidation of the cell wall proteome by a method using LiCl extraction and subsequent proteomic analysis via 2D gel electrophoresis [51]. Tjalsma et al. succeeded in using a "shedding and shaving" approach in *B. subtilis*: incubation of intact bacteria with trypsin in order to cleave off surface proteins with consecutive identification of these "shaved" tryptic peptides led to the finding of 41 proteins being specifically shaved off this subcellular location. Verification of these results was carried out with proteolysis by bead- bound trypsin that is unable to penetrate the bacterial cell wall demonstrating a genuine surface-exposed localization [52]. Despite the use of a gel- free approach, this study only gave limited information on surface bound proteins and implies strict criteria for localization verification.

For the human pathogen *S. aureus*, this subcellular location is even more essential than for its non- pathogenic gram positive counterpart *B. subtilis*: virulence is mainly characterized by cell wall bound/extracellular factors and hence this subproteome is of

_____

high importance to viability in the host environment of the cell. Other studies than the above mentioned successfully applied different methods for isolation of surface proteins with serial extraction of proteins by different, highly concentrated salts, enzymatic extractions and biotinylation with subsequent affinity purification in the gram positive *Listeria monocytogenes* [53]. Another study on *S. aureus* has been recently published by Hempel et al. [54] accordingly based on the biotinylation of proteins exposed on the cell surface of intact cells and the following affinity purification of these protein species. This study exemplifies the integration of a gel-free quantitation approach using $^{15}$N metabolic labeling with the biotinylation of surface exposed proteins and yields an unsurpassed number of surface bound protein species.

_____

## 3. Fractionation techniques applied in mass spectrometry based proteomics

Despite the fact that mass spectrometry based quantitative proteomics is experimentally simple and results in accurate quantitative data, it mostly suffers from a limited dynamic range. This limitation in dynamic range comes with an inherent drawback of biological samples that contain proteins in a wide range of concentrations. Generally, concentration of the most abundant analytes define the range that lower abundant proteins can still be identified in a given sample. The mass ratio between the least and the most abundant proteins ranges from four orders of magnitude in prokaryotic samples, six in eukaryotic cells [55] and 12 in complex biological fluids such as plasma [56]. Several authors have questioned the capacity of the current mass spectrometry based shotgun approaches to be alone powerful enough to determine a comprehensive coverage of any proteome [57-59]. Therefore, Nielsen et al. propose the use of extensive separation steps prior to digestion and any further fractionation to overcome this limitation [58].

Prefractionation prior to mass spectrometric analysis may be conducted at subcellular, protein or peptide level. Fractionation techniques may be used either for the elucidation of spatial changes in the proteome of specific subcellular localizations or towards the reduction of complexity of a given biological sample.

Intact eukaryotic cells consist mostly of different organelles that possess different physical properties so that these behave differently in separation techniques that are applied to resolve these subcellular structures. Besides organelles, physical properties of subcellular structures are exploited: the solubility in separating cytosolic proteins and hydrophobic membrane proteins, the density leading to separation in centrifugation or density gradient centrifugation.

### 3.1. Chromatographic/electrophoretic fractionation methods

#### 3.1.1. GeLCMS

The most common protein prefractionation method is the GeLCMS method [60]. Denatured protein samples are electrophoretically separated by 1D SDS PAGE and the resulting gel slab is cut to yield protein fractions differing in molecular weight. Subsequent to *in gel* proteolytic digestion, peptides are eluted and subjected to mass spectrometric LC-MS/MS

_____

analysis [61, 62]. The advantage of this approach is the fast and easy separation of the sample on protein level with a significant reduction in sample complexity effectively increasing the dynamic range of the analysis. Simultaneous cleaning of the sample to avoid interference with salts or detergents used in sample preparation facilitates the analysis after the electrophoretic separation. Furthermore, limited proteolysis may be seen by analysis of the different fractions for multiple appearance of a specific protein. Through the prefractionation of high abundant proteins over the whole lane on the gel, systematic cutting and digestion of every part of the band provides analysis of even very low abundant proteins that are not visually detected on such a gel. This systematic approach would not be feasible in large 2D GE with subsequent MS analysis. Therefore GeLCMS is most efficient for a cartography of proteins in a sample and in combination with stable isotope labeling for systems wide proteomic protein profiling studies [60].

In combination with standard cell fractionation methods, the GeLCMS approach gives reliable results assured by the orthogonality of the two fractionation methods [58]. To date, this method is besides the IEF RP one of the multidimensional methods that has proved to be most successful. [63, 64].
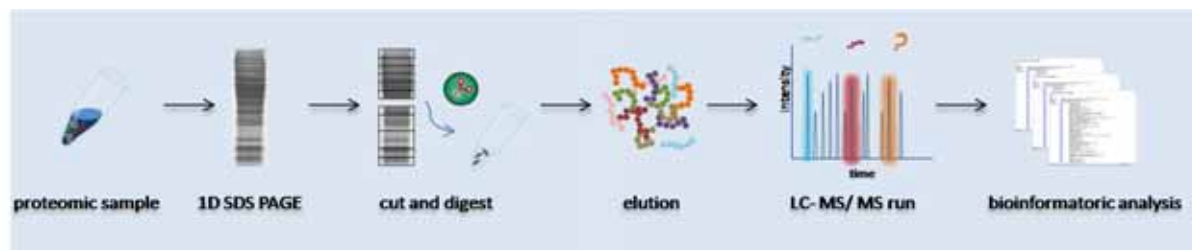


**Fig.3 schematic representation of the GeLCMS workflow**

### 3.1.2. IEX-RP

Ion exchange chromatography coupled to reversed phase chromatography is the most common on- line two dimensional chromatographic prefractionation method for peptides used in proteomics [65]. In ion exchange chromatography, the analytes are separated by the interaction of the analyte with the static phase based on the effective charge of the analyte under specific mobile phase conditions. IEX is carried out mostly as strong cation exchange chromatography (SCX) either being on line or off- line coupled to the reversed phase (RP) separation step [1].

Online coupling was introduced in a concept called MudPit (multidimensional protein identification technology) [66]. In this approach, reversed phase C18 material is packed into

_____

a fused silica capillary. Subsequently, SCX material is packed directly behind the first packed resin to get hold of a biphasic column enabling for two dimensional separations. Consecutive chromatographic runs with salt gradients and gradients containing an organic solvent elute the peptides from the SCX columns and subsequently separate the analytes on the RP column prior to on- line LC-MS/MS analysis. Later studies published demonstrated improvements like triphasic columns of this type even enhancing the applicability to samples not being desalted prior to analysis [67]. An advantage of MudPIT is that the entire system is directly coupled to mass spectrometric analysis and hence no loss of sample is encountered. The problem of salts used in SCX for the step wise elution of peptides from the resin is solved by the use of mass spectrometry compatible salts that evaporate during the ionization step at the interface of LC and MS (mostly salts containing ammonium used as part of the weak acid-base pair is used). Disadvantages lie especially in the limit of organic solvent that may be used on applying the salt step acting on the SCX part due to the second dimension of RP material. Furthermore, any repetition of sample analysis is not given for a single salt/elution step [16].

Offline SCX is slower than the on line coupled approach but is easy to perform with no specific column set up. Applying larger column sizes and higher organic solvents in elution buffer systems increase the amount of loading and recovery of peptides of the step wise elution of peptides from the SCX column [16].

### 3.1.3. IEF RP chromatography

Isoelectric focusing is a fractionation technique that is used since the beginning of proteomics studies in separation of proteins or peptides mainly focused on the 2D GE. In mass spectrometric based workflows the IEF RP methods are applied to proteins or to peptides resulting from enzymatic digestions of proteins. The proteins/ peptides are focused in an electric field according to their p*I* and subsequently the resulting fractions are analyzed by RP LC-MS/MS (with an additional digestion step in the case of proteins fractionated by IEF).

The isoelectric focusing step may be carried out in different ways based on the same principle: either in solution, in free flow electrophoresis, in capillary electrophoresis or by in gel electrophoresis using an IPG strip [60, 68, 69].

The method of applying IEF based on IPG strips as the first dimension prior to mass spectrometric analyses was first introduced by Loo et al. who used IPG- IEF with

_____

subsequent offline LC- MALDI or online LC- ESI MS analysis in a method called "virtual 2D" [70]. IPG- IEF has proven to be most powerful in several studies [71-74]. IPG- IEF is superior to both GeLCMS and SCX based studies in terms of identified peptides [72, 75]. This leads to an increased proteome coverage in less time regarding sample preparation and analysis time with a high resolving power [64]. Cargile et al. showed that IPG IEF improved the sensitivity by up to 100% compared to SCX [71]. Furthermore the use of accurate mass and peptide isoelectric point (pI) as identification criteria substantially increase the number of validated peptides of over 20% with unchanged false positive rate [76]. As the p*I* is easy to calculate for given proteins and peptides, this information may be used to corrobate results obtained from mass spectrometric analyses.

Besides Off- gel approaches set up with non commercial solutions, equipment for Off-gel IEF is commercially available and is based on a comparable concept as the original IPG IEF [77]. Off- gel IEF was shown to be reliable in peptide/proteins analysis in several studies [78, 79]. A drawback in using commercially available systems is the fixed number of fractions to be obtained in a single run and the predefined shape and size of the wells.

An additional feature of the IEF based on IPG strips or the Off-gel IEF is the applicability in PTM analysis regarding phosphorylation prior only applicable in 2D GE [80].

# 4. Protein profiling in proteomic analyses

## 4.1.    2D GE coupled with MS

A great problem of proteomic analyses based on mass spectrometry is the fact that mass spectrometry itself is poorly quantitative. The classical proteomic approach of 2D GE coupled to subsequent identification of proteins circumvents this disadvantage of mass spectrometry, as protein abundances are determined by comparison of spot intensities in 2D gels relative to each other.

Quantitative information from staining intensities may be used to elucidate changes in complete protein amount, synthesis rate in combination with radio labeled protein samples or change of phosphorylation in certain protein samples. Identification of proteins in this classical approach is mostly carried out by MALDI TOFTOF instruments capable of MS/MS. The use of specific stains allows for accurate quantitation using

_____

internal standards encoded by differentially absorbing dyes [81, 82] or specific staining of modifications like phosphorylations and glycosylations [83, 84].

## *4.2.* *Mass spectrometry based quantitation*

Besides the classical 2D GE based approach, quantitative analyses based on the concept of isotope dilution [85] or even without any label are applied in proteomics [86].

In quantitative analyses using the stable isotope labeling strategy, stable isotope tags are introduced by metabolic labeling, enzymatic labeling and chemical labeling or by spiking labeled synthetic peptides into the sample. Isotope labeling was introduced in its present form in 1999 [87-89]. In principle, pairs of chemically identical analytes of different stable isotope compositions can be differentiated in a mass spectrometer owing to their mass difference. Comparison of the intensities of the masses relative to each other gives quantitative information that allows conclusion on relative quantities and abundance ratios for the compared analytes. [90]

Label free techniques hold the promise to yield relative quantitative results within proteomics experiments to be compared simply based on LC- MS/MS runs without any additional labeling. Basically, acquired spectra for a specific peptide matching to a protein are used for a comparison between two or more experiments in order to give relative quantitative information. Furthermore it is possible to compare mass spectrometric runs and relative abundances by the use of spectral counting that is relying on the frequency of occurrence of spectra belonging to each peptide.

### *4.2.1. Metabolic labeling*

Metabolic labeling in its two forms, complete metabolic labeling with isotopically labeled salts containing $^{15}$N or $^{13}$C [88] and SILAC (stable isotope labeling of amino acids in cell culture) [91] is the most accurate quantitative mass spectrometric method in terms of the overall experimental process (biochemical and mass spectrometric procedures) [92]. The sample is labeled *in vivo* either to yield proteins containing nitrogen practically exclusively as $^{15}$N or to yield proteins with only a subgroup of amino acids being isotopically labeled. Inherently, the earliest possible time is used to label the complete proteome (see Fig. 4). Mixing of labeled samples may be accomplished as early as at cell level. This makes this methodology most widely applied

_____

in quantitative proteomics: all sample preparation methods and all existing protocols on pre- fractionation may be used with samples that are already metabolically labeled and mixed.
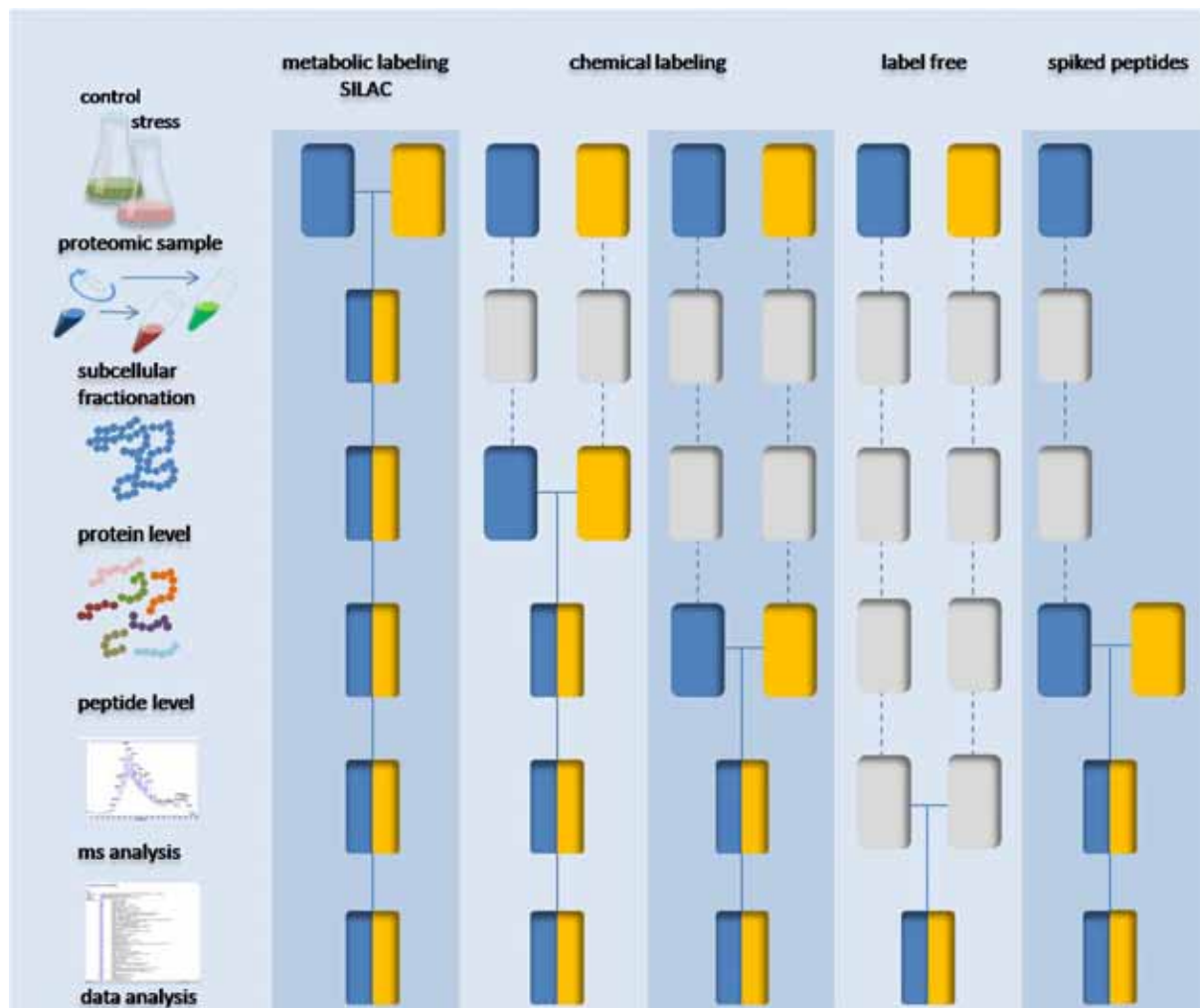


**Fig.4 schematic representation of mass spectrometry based quantitation workflows** Boxes in blue and orange represent two different physiological states/experimental conditions. Horizontal lines depict the stage of sample combination. Grey boxes indicate steps in the workflow prone to experimental variations and thus to quantitation errors. (Adapted from Bantscheff et al.[86] )

Dependent on the experimental prerequisites, metabolic labeling with heavy $^{15}$N salts is

_____

mostly used in prokaryotic systems as all available nitrogen sources must be exchanged in the labeled cell culture.

Metabolic labeling with eukaryotes or even higher eukaryotes such as fruit flies [93], rats [94] or plants [95] have been carried out but are scarce due to the cost of the complete labeling dependent on the effort to be conducted and time required establishing a complete labeling. The greatest bottle neck of the complete metabolic labeling is the increase in complexity that is introduced by the theoretical doubling of peptide species in a labeled proteome. Furthermore, suitable software was not available for a long time compared to other stable isotope methods that rely on defined mass shifts in comparison to mass shifts determined by the primary structure of the peptide.

SILAC, introduced by Mathias Mann and coworkers, is the most popular metabolic labeling approach used in functional genomics. Studies based on SILAC methodology cover protein profiling on proteome level [96], post translational modification analysis on large scale [97, 98], examination of phosphorylation pattern changes [99, 100] and large scale interaction studies [101]. To assure an almost complete labeling of all tryptic peptides, most commonly used amino acids are $^{13}$C6 arginine, $^{13}$C6 lysine and $^{15}$N4 arginine [86, 91]. A maximum of three different stages may be compared in one sample. Like the complete metabolic labeling described before, SILAC is carried out in cell cultures *in vivo* with the advantage that purity over 90% may be achieved within 6-8 passages[91].

A disadvantage over metabolic labeling with $^{15}$N enriched salts may be the interconversion/ degradation of amino acids that complicates the bioinformatic analysis of the quantitative data [102]. Contrary to the $^{15}$N metabolic labeling, post- acquisitional bioinformatoric processing resembles that of other isotopic labeling techniques relying on a fixed mass difference.

### *4.2.2. Enzymatic labeling*

Another possibility of the introduction of isotope labels is the *in vitro* utilization of an enzymatic reaction. Labeling is carried out either during the proteolytic digestion or after the proteolysis in a second incubation step (Fig. 4). Enzymatically introduced $^{18}$O isotopes will result in a mass shift of 2 to 4 Da dependent on the exchange of one or two atoms and in theory practically no side reactions are to be found during the reaction. Commonly used enzymes for this task are congruent with the proteases that are utilized for digesting proteins prior to mass spectrometric analysis like trypsin [103, 104]. In general,

_____

the possibility of quantitative bioinformatic tools to compare differentially labeled peptides in a spectrum relies on a mass difference that is higher than the width of the isotopomer cluster of the light mass. Therefore, [18]O labeling is applicable to quantitative mass spectrometry with the disadvantages that a full labeling is practically never achieved and that the incorporation rate depends on the nature of the peptides [105, 106].

### *4.2.3. Chemical labeling*

Chemical labeling is a method that yields a great versatility in terms of derivatization of functional groups found on the primary structure of a peptide and the features the chemical tag exhibits. Most commonly derivatized side chains in peptides are these of cysteine, peptide N-termini and the ε amino group of lysine [86]. In this *in vitro* method, labeled and unlabeled samples are combined either at protein or at peptide level (fig. 4) and mass shifts between labeled and unlabeled peptides are fixed.

One of the first methods applied in quantitative mass spectrometry based proteomics was ICAT (Isotope coded affinity tag) by Gygi et al. [87]. The mass tag consists either of deuterated or unlabeled biotin affinity tags that bind to cysteine residues and yield the possibility to affinity purify the peptides after labeling and thus reducing the complexity of the sample analysed. The concept of labeling a peptide mix with subsequent reduction of complexity is advantageous with inherent major drawbacks: this method is only usable for cysteine containing peptides which are relatively rare. Furthermore limitations apply that the tag cannot be used in more specific studies regarding post translational modifications or studies that need a most comprehensive protein coverage e.g. in the analysis of splice variants make this concept only applicable in certain studies. Despite these disadvantages, the ICAT method was used in a wide range of studies [107-110].

Labels targeting N-termini and the ε- amino group of lysine are based on N-hydroxysuccinimide chemistry or bonds mediated by other esters/acid anhydrides. This class of label comprises a plethora of different systems e.g. ICPL (isotope coded protein label) [111], iTRAQ (isotope tags for relative and absolute quantitation) [112], TMT (tandem mass tags) [113], acetic acid/succinimid acid [114] and formaldehyde labeling [115]. Within this group it is important to distinguish between methods that rely on quantitation on survey scan level and those that identify and quantify peptides at the same time in the fragment ion scans. TMT and iTRAQ for example are based on isobaric tags that differ only in the fragmentation pattern in the mass spectrometer (**fig. 5**). Isobaric labeling should in theory

_____

not affect the chromatographic behavior of the peptides and no increase in complexity on survey scan level should be seen. Another great advantage is the possibility with multiplexing of iTRAQ: up to eight samples that may be compared with each other in a single run.
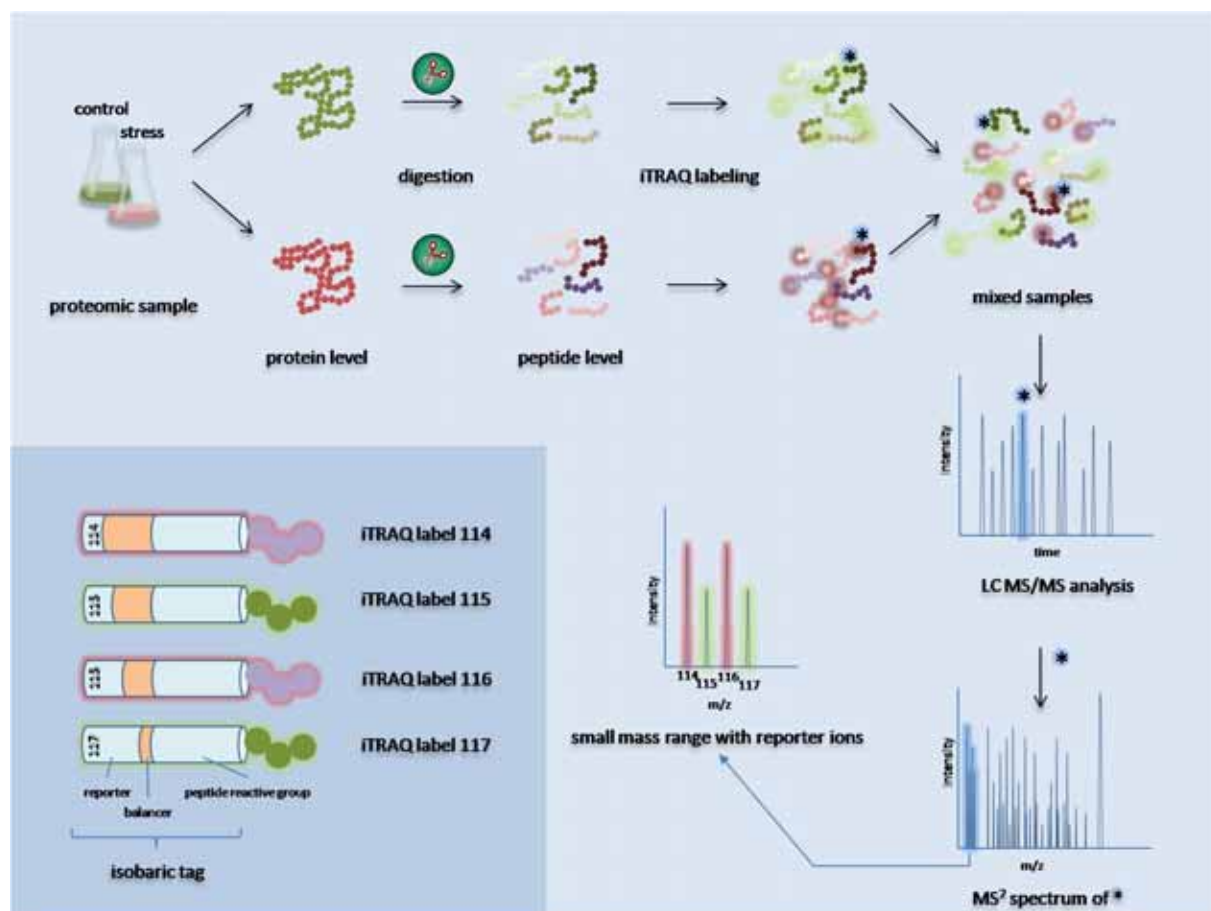


**Fig.5 schematic representation of relative quantitation workflow based on the iTRAQ reagent** proteins or peptides (schematic workflow for peptides is shown) of proteomic samples are labeled with up to eight isobaric label (4-plex is shown) and mixed. Relative quantitation takes place only at fragmentation level without increased sample complexity

### 4.2.4. Spiking

Spiking of defined analytes to a complex sample is already known and applied for a long time [116, 117]. In today's proteomics, addition of a stable isotope labeled standard peptide to a protein digest with subsequent comparison of the mass spectrometric signal to the endogenous peptide in the sample was coined AQUA (absolute quantitation of proteins) [118, 119]. In principle, the methodology is based on the concept that with the information on the relative abundance of a synthetic reference peptide with vs. its endogenous

_____

counterpart in the sample it is possible to trace back the quantities of the corresponding protein in the sample and therefore in the proteomic sample if the absolute quantity of the spiked reference peptide is given. Contrary to relative quantitation, the addition of synthetic peptides to a proteome sample focuses on the determination of the quantity of only a small number of proteins of interest. Difficulties within this concept arise in highly complex samples in which the labeled/unlabeled peptides are hard to detect due to co-eluting isobaric peptides that might be more abundant. A combination of information on retention time, peptide mass and fragment pattern eliminates these ambiguities, as it is done in targeted approaches like multiple reaction monitoring (MRM) [120]. MRM is a tandem MS scan mode unique to triple- quadrupole MS instruments that are capable of rapid, sensitive, and specific quantitation of analytes in highly- complex sample matrices [121]. With targeted approaches, it is possible to achieve a quantitation range of 4-5 orders of magnitude [122]. Targeted approaches will gain more and more importance in mass spectrometry based proteomics – biomarker research and biomedical studies will profit of this emerging field in proteomics [123]. A prerequisite for this will be large databases of proteomics identification data [124] in order to chose the right synthetic peptide that is suitable as a standard in targeted approaches [125]. Availability of peptides in precisely determined concentrations is a limitation to AQUA. To partially overcome this problem, the use of artificial genes that are translated into labeled proteins has been invented. These concatamers of peptides are contained in an artificial QconCAT (quantitation concatamers) protein that is used to be spiked into a sample. Having determined the accurate protein amount before, the concentration of the QconCAT peptides is used to determine absolute amounts of all peptides/ proteins after digestion of the sample and following mass spectrometric analysis. [126, 127]. The clear cut advantage of such an approach compared to the AQUA technique is the limited cost coming with the necessity only to purify and label a single protein which can be expressed in a suitable biological system.

### 4.2.5. Label free mass spectrometry based quantitation

Label free mass spectrometry based quantitation methods do not rely on any labeling of the analyzed samples (Fig. 4). Initially, at times where only low resolution mass spectrometry based e.g. Ion trap mass analyzers were available, peptide and hence protein

_____

identification data was used for a rough indication of protein quantity. Basically, the frequency that peptides belonging to a specific protein are fragmented by CID served as a proxy for the relative abundance of the protein in the sample. This methodology was termed spectral counting [66] and is used as a fast, convenient and intuitive quantitation strategy [128, 129]. The disadvantage of spectral counting lies within the dependency of this method on the quantity of the peptide identifications. Furthermore, quantitation is impossible for peptides with multiple assignments to different proteins. The most intriguing disadvantage is the fact that proteins that are determined with less than three peptides are failing to be correctly relatively quantified. The reason for that lies within the underrepresentation of proteins of low abundance in the stochastically MS/MS sampling process in the mass spectrometer [130].

With the advent of high accurate and high resolving mass spectrometry in combination with very reproducible runs of on- line coupled liquid chromatography, label free quantitation based on the integration of peptide elution peaks on the survey scan level became possible. Peptide elution patterns are tracked using the retention time characteristics in combination with the correct high accurate monoisotopic mass trace. Due to variations in the LC-MS runs, sophisticated alignment algorithms have to be applied [131]. With these characteristics, a chromatographic elution profile of every peptide may be reconstructed and be distinguished from chemical noise. Peptide abundances are calculated based on intense bioinformatoric analysis by comparison of the integrated peptide elution peaks [132]. Advantageous to this method is the increased dynamic range and the lack of the undersampling problems to be encountered in the spectral counting approach. Mandatory for this approach are the consistency of the LC- MS/MS runs to yield reproducible results of peptide identification with consistent masses and retention time coordinates. Furthermore, besides the almost infinite applicability of this method to all proteomic samples, comparison of samples needs substantial replicates for a robust statistical validation and a preferably low number of sample preparation steps should be provided prior to the mass spectrometric measurement to keep the experimental error at a low level [86].

# *5. References*

_____

1.     Issaq, H.J., Chan, K.C., Janini, G.M., Conrads, T.P. & Veenstra, T.D. Multidimensional separation of peptides for effective proteomic analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* **817**, 35-47 (2005).

2.     O'Farrell, P.H. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **250**, 4007-4021 (1975).

3.     Wilkins, M.R*., et al.* Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology & genetic engineering reviews* **13**, 19-50 (1995).

4.     Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207 (2003).

5.     Görg, A., Weiss, W. & Dunn, M.J. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **4**, 3665-3685 (2004).

6.     Klose, J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26**, 231-243 (1975).

7.     Bjellqvist, B*., et al.* Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* **6**, 317-339 (1982).

8.     Görg, A., Postel, W. & Gunther, S. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **9**, 531-546 (1988).

9.     Görg, A*., et al.* The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**, 1037-1053 (2000).

10.    Görg, A., Drews, O., Luck, C., Weiland, F. & Weiss, W. 2-DE with IPGs. *Electrophoresis* **30 Suppl 1**, S122-132 (2009).

11.    Anderson, N.L. & Anderson, N.G. Analytical techniques for cell fractions. XXII. Two-dimensional analysis of serum and tissue proteins: multiple gradient-slab gel electrophoresis. *Anal Biochem* **85**, 341-354 (1978).

12.    Hecker, M., Antelmann, H., Buttner, K. & Bernhardt, J. Gel-based proteomics of Gram-positive bacteria: a powerful tool to address physiological questions. *Proteomics* **8**, 4958-4975 (2008).

13.    Eymann, C*., et al.* Dynamics of protein phosphorylation on Ser/Thr/Tyr in Bacillus subtilis. *Proteomics* **7**, 3509-3526 (2007).

14.    Zong, C*., et al.* Two-dimensional electrophoresis-based characterization of post-translational modifications of mammalian 20S proteasome complexes. *Proteomics* **8**, 5025-5037 (2008).

15.    Gerth, U*., et al.* Clp-dependent proteolysis down-regulates central metabolic pathways in glucose-starved Bacillus subtilis. *J Bacteriol* **190**, 321-331 (2008).

16.    Canas, B., Pineiro, C., Calvo, E., Lopez-Ferrer, D. & Gallardo, J.M. Trends in sample preparation for classical and second generation proteomics. *J Chromatogr A* **1153**, 235-258 (2007).

17.    Matsudaira, P. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J Biol Chem* **262**, 10035-10038 (1987).

18.    Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).

19.    Patterson, S.D. & Aebersold, R. Mass spectrometric approaches for the identification of gel-separated proteins. *Electrophoresis* **16**, 1791-1814 (1995).

_____

20. Henzel, W.J.*, et al.* Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci U S A* **90**, 5011-5015 (1993).

21. James, P., Quadroni, M., Carafoli, E. & Gonnet, G. Protein identification by mass profile fingerprinting. *Biochem Biophys Res Commun* **195**, 58-64 (1993).

22. Yates, J.R., 3rd, Speicher, S., Griffin, P.R. & Hunkapiller, T. Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem* **214**, 397-408 (1993).

23. Griffin, P.R.*, et al.* Direct database searching with MALDI-PSD spectra of peptides. *Rapid Commun Mass Spectrom* **9**, 1546-1551 (1995).

24. Yates, J.R., 3rd, Eng, J.K. & McCormack, A.L. Mining genomes: correlating tandem mass spectra of modified and unmodified peptides to sequences in nucleotide databases. *Anal Chem* **67**, 3202-3210 (1995).

25. Mann, M. & Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem* **66**, 4390-4399 (1994).

26. Boguski, M.S., Lowe, T.M. & Tolstoshev, C.M. dbEST--database for "expressed sequence tags". *Nat Genet* **4**, 332-333 (1993).

27. Adams, M.D.*, et al.* Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651-1656 (1991).

28. Fleischmann, R.D.*, et al.* Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* **269**, 496-512 (1995).

29. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. & Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71 (1989).

30. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212-217 (2006).

31. Henzel, W.J., Watanabe, C. & Stults, J.T. Protein identification: the origins of peptide mass fingerprinting. *J Am Soc Mass Spectrom* **14**, 931-942 (2003).

32. Zenobi, R. & Knochenmuss, R. Ion formation in MALDI mass spectrometry. *Mass Spectrometry Reviews* **17**, 337-366 (1998).

33. Wollnik, H. Time-of-flight mass analyzers. *Mass Spectrometry Reviews* **12**, 89-114 (1993).

34. Douglas, D.J., Frank, A.J. & Mao, D. Linear ion traps in mass spectrometry. *Mass Spectrom Rev* **24**, 1-29 (2005).

35. Jonscher, K.R. & Yates, J.R. The quadrupole ion trap mass spectrometer: a small solution to a big challenge. *Analytical biochemistry* **244**, 1-15 (1997).

36. Perry, R.H., Cooks, R.G. & Noll, R.J. Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass spectrometry reviews* **27**, 661-699 (2008).

37. Comisarow, M.B. & Marshall, A.G. Frequency-sweep fourier transform ion cyclotron resonance spectroscopy* 1. *Chemical physics letters* **26**, 489-490 (1974).

38. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat Biotechnol* **26**, 1135-1145 (2008).

39. Sonenshein, A.L. Control of key metabolic intersections in Bacillus subtilis. *Nat Rev Microbiol* **5**, 917-927 (2007).

40. Zweers, J.C.*, et al.* Towards the development of Bacillus subtilis as a cell factory for membrane proteins and protein complexes. *Microb Cell Fact* **7**, 10 (2008).

_____

41. Wolff, S.*, et al.* Towards the entire proteome of the model bacterium Bacillus subtilis by gel-based and gel-free approaches. *J Chromatogr B Analyt Technol Biomed Life Sci* **849**, 129-140 (2007).

42. Hirose, I.*, et al.* Proteome analysis of Bacillus subtilis extracellular proteins: a two-dimensional protein electrophoretic study. *Microbiology* **146 ( Pt 1)**, 65-75 (2000).

43. Butcher, B.G. & Helmann, J.D. Identification of Bacillus subtilis W dependent genes that provide intrinsic resistance to antimicrobial compounds produced by Bacilli. *Molecular microbiology* **60**, 765-782 (2006).

44. Gonzalez-Pastor, J.E., Hobbs, E.C. & Losick, R. Cannibalism by sporulating bacteria. 510 (AAAS, 2003).

45. Santoni, V., Molloy, M. & Rabilloud, T. Membrane proteins and proteomics: un amour impossible? *Electrophoresis* **21**, 1054-1070 (2000).

46. Bunai, K.*, et al.* Profiling and comprehensive expression analysis of ABC transporter solute-binding proteins of Bacillus subtilis membrane based on a proteomic approach. *Electrophoresis* **25**, 141-155 (2004).

47. Eymann, C.*, et al.* A comprehensive proteome map of growing Bacillus subtilis cells. *Proteomics* **4**, 2849-2876 (2004).

48. Wolff, S.*, et al.* Gel-free and gel-based proteomics in Bacillus subtilis: a comparative study. *Mol Cell Proteomics* **5**, 1183-1192 (2006).

49. Wolff, S., Hahne, H., Hecker, M. & Becher, D. Complementary analysis of the vegetative membrane proteome of the human pathogen Staphylococcus aureus. *Mol Cell Proteomics* **7**, 1460-1468 (2008).

50. Hahne, H., Wolff, S., Hecker, M. & Becher, D. From complementarity to comprehensiveness--targeting the membrane proteome of growing Bacillus subtilis by divergent approaches. *Proteomics* **8**, 4123-4136 (2008).

51. Antelmann, H., Yamamoto, H., Sekiguchi, J. & Hecker, M. Stabilization of cell wall proteins in Bacillus subtilis: a proteomic approach. *Proteomics* **2**, 591-602 (2002).

52. Tjalsma, H., Lambooy, L., Hermans, P.W. & Swinkels, D.W. Shedding & shaving: disclosure of proteomic expressions on a bacterial face. *Proteomics* **8**, 1415-1428 (2008).

53. Schaumburg, J.*, et al.* The cell wall subproteome of Listeria monocytogenes. *Proteomics* **4**, 2991-3006 (2004).

54. Hempel, K.*, et al.* Quantitative cell surface proteome profiling for SigB-dependent protein expression in the human pathogen Staphylococcus aureus via biotinylation approach. *J Proteome Res* (2010).

55. Lu, P., Vogel, C., Wang, R., Yao, X. & Marcotte, E.M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* **25**, 117-124 (2007).

56. Anderson, N.L. & Anderson, N.G. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* **1**, 845-867 (2002).

57. Domon, B. & Aebersold, R. Challenges and opportunities in proteomics data analysis. *Mol Cell Proteomics* **5**, 1921-1926 (2006).

58. Nielsen, M.L., Savitski, M.M. & Zubarev, R.A. Extent of modifications in human proteome samples and their effect on dynamic range of analysis in shotgun proteomics. *Mol Cell Proteomics* **5**, 2384-2391 (2006).

59. Cox, J. & Mann, M. Is proteomics the new genomics? *Cell* **130**, 395-398 (2007).

_____

60.     Rabilloud, T*., et al.* Power and limitations of electrophoretic separations in proteomics strategies. *Mass Spectrom Rev* **28**, 816-843 (2009).

61.     Schirle, M., Heurtier, M.A. & Kuster, B. Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics* **2**, 1297-1305 (2003).

62.     Lasonder, E*., et al.* Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. *Nature* **419**, 537-542 (2002).

63.     Ishihama, Y. Proteomic LC-MS systems using nanoscale liquid chromatography with tandem mass spectrometry. *Journal of Chromatography A* **1067**, 73-83 (2005).

64.     Graumann, J*., et al.* Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol Cell Proteomics* **7**, 672-683 (2008).

65.     Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J. & Gygi, S.P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* **2**, 43-50 (2003).

66.     Washburn, M.P., Wolters, D. & Yates, J.R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**, 242-247 (2001).

67.     McDonald, W.H., Ohi, R., Miyamoto, D.T., Mitchison, T.J. & Yates, J.R. Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *International Journal of Mass Spectrometry* **219**, 245-251 (2002).

68.     Krijgsveld, J., Gauci, S., Dormeyer, W. & Heck, A.J.R. In-gel isoelectric focusing of peptides as a tool for improved protein identification. *J. Proteome Res* **5**, 1721-1730 (2006).

69.     Cargile, B.J., Bundy, J.L., Freeman, T.W. & Stephenson, J.L., Jr. Gel based isoelectric focusing of peptides and the utility of isoelectric point in protein identification. *J Proteome Res* **3**, 112-119 (2004).

70.     Loo, J.A*., et al.* High sensitivity mass spectrometric methods for obtaining intact molecular weights from gel-separated proteins. *Electrophoresis* **20**, 743-748 (1999).

71.     Cargile, B.J., Talley, D.L. & Stephenson, J.L., Jr. Immobilized pH gradients as a first dimension in shotgun proteomics and analysis of the accuracy of pI predictability of peptides. *Electrophoresis* **25**, 936-945 (2004).

72.     Krijgsveld, J., Gauci, S., Dormeyer, W. & Heck, A.J.R. In-gel isoelectric focusing of peptides as a tool for improved protein identification. 1721-1730 (2006).

73.     Scherl, A*., et al.* Exploring glycopeptide-resistance in Staphylococcus aureus: a combined proteomics and transcriptomics approach for the identification of resistance-related markers. *BMC Genomics* **7**, 296 (2006).

74.     Lengqvist, J., Uhlen, K. & Lehtio, J. iTRAQ compatibility of peptide immobilized pH gradient isoelectric focusing. *Proteomics* **7**, 1746-1752 (2007).

75.     Essader, A.S., Cargile, B.J., Bundy, J.L. & Stephenson Jr, J.L. A comparison of immobilized pH gradient isoelectric focusing and strong cation exchange

_____

chromatography as a first dimension in shotgun proteomics. *Proteomics* **5**, 24-34 (2005).

76.    Cargile, B.J. & Stephenson, J.L., Jr. An alternative to tandem mass spectrometry: isoelectric point and accurate mass for the identification of peptides. *Anal Chem* **76**, 267-275 (2004).

77.    Ros, A*., et al.* Protein purification by Off-Gel electrophoresis. *Proteomics* **2**, 151-156 (2002).

78.    Heller, M*., et al.* Added value for tandem mass spectrometry shotgun proteomics data validation through isoelectric focusing of peptides. *J Proteome Res* **4**, 2273-2282 (2005).

79.    Horth, P., Miller, C.A., Preckel, T. & Wenz, C. Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Mol Cell Proteomics* **5**, 1968-1974 (2006).

80.    Xu, C.F., Wang, H., Li, D., Kong, X.P. & Neubert, T.A. Selective enrichment and fractionation of phosphopeptides from peptide mixtures by isoelectric focusing after methyl esterification. *Anal. Chem* **79** (2007).

81.    Ünlü, M., Morgan, M.E. & Minden, J.S. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. 2071-2077 (Wiley Subscription Services, Inc., A Wiley Company Hoboken, 1997).

82.    Alban, A*., et al.* A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* **3**, 36-44 (2003).

83.    Schulenberg, B., Beechem, J.M. & Patton, W.F. Mapping glycosylation changes related to cancer using the Multiplexed Proteomics technology: a protein differential display approach. *Journal of Chromatography B* **793**, 127-139 (2003).

84.    Steinberg, T.H*., et al.* Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology. *Proteomics* **3**, 1128-1144 (2003).

85.    Bowers, G.N., Jr., Fassett, J.D. & White, E.t. Isotope dilution mass spectrometry and the National Reference System. *Anal Chem* **65**, 475R-479R (1993).

86.    Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* **389**, 1017-1031 (2007).

87.    Gygi, S.P*., et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994-999 (1999).

88.    Oda, Y., Huang, K., Cross, F.R., Cowburn, D. & Chait, B.T. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci U S A* **96**, 6591-6596 (1999).

89.    Jensen, P.K*., et al.* Probing proteomes using capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* **71**, 2076-2084 (1999).

90.    Ong, S.E. & Mann, M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* **1**, 252-262 (2005).

91.    Ong, S.E*., et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386 (2002).

92.    Sechi, S. & Oda, Y. Quantitative proteomics using mass spectrometry. *Curr Opin Chem Biol* **7**, 70-77 (2003).

_____

93.    Krijgsveld, J.*, et al.* Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. *Nat Biotechnol* **21**, 927-931 (2003).

94.    Wu, C.C., MacCoss, M.J., Howell, K.E., Matthews, D.E. & Yates, J.R., 3rd. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal Chem* **76**, 4951-4959 (2004).

95.    Gruhler, A., Schulze, W.X., Matthiesen, R., Mann, M. & Jensen, O.N. Stable isotope labeling of Arabidopsis thaliana cells and quantitative proteomics by mass spectrometry. *Mol Cell Proteomics* **4**, 1697-1709 (2005).

96.    de Godoy, L.M.*, et al.* Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biol* **7**, R50 (2006).

97.    Ong, S.E., Mittler, G. & Mann, M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat Methods* **1**, 119-126 (2004).

98.    Andersen, J.S.*, et al.* Nucleolar proteome dynamics. *Nature* **433**, 77-83 (2005).

99.    Olsen, J.V.*, et al.* Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635-648 (2006).

100.   Mann, M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol* **7**, 952-958 (2006).

101.   de Godoy, L.M.*, et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251-1254 (2008).

102.   Ong, S.E., Kratchmarova, I. & Mann, M. Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res* **2**, 173-181 (2003).

103.   Yao, X., Freas, A., Ramirez, J., Demirev, P.A. & Fenselau, C. Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal Chem* **73**, 2836-2842 (2001).

104.   Reynolds, K.J., Yao, X. & Fenselau, C. Proteolytic 18O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *J Proteome Res* **1**, 27-33 (2002).

105.   Johnson, K.L. & Muddiman, D.C. A method for calculating 16O/18O peptide ion ratios for the relative quantification of proteomes. *J Am Soc Mass Spectrom* **15**, 437-445 (2004).

106.   Ramos-Fernandez, A., Lopez-Ferrer, D. & Vazquez, J. Improved method for differential expression proteomics using trypsin-catalyzed 18O labeling with a correction for labeling efficiency. *Mol Cell Proteomics* **6**, 1274-1286 (2007).

107.   Goodchild, A., Raftery, M., Saunders, N.F., Guilhaus, M. & Cavicchioli, R. Cold adaptation of the Antarctic archaeon, Methanococcoides burtonii assessed by proteomics using ICAT. *J Proteome Res* **4**, 473-480 (2005).

108.   Chen, R.*, et al.* Quantitative proteomics analysis reveals that proteins differentially expressed in chronic pancreatitis are also frequently involved in pancreatic cancer. *Mol Cell Proteomics* **6**, 1331-1342 (2007).

109.   Cho, S.H., Goodlett, D. & Franzblau, S. ICAT-based comparative proteomic analysis of non-replicating persistent Mycobacterium tuberculosis. *Tuberculosis (Edinb)* **86**, 445-460 (2006).

110.   Leichert, L.I.*, et al.* Quantifying changes in the thiol redox proteome upon oxidative stress in vivo. *Proc Natl Acad Sci U S A* **105**, 8197-8202 (2008).

111.   Schmidt, A., Kellermann, J. & Lottspeich, F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**, 4-15 (2005).

_____

112.  Ross, P.L.*, et al.* Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **3**, 1154-1169 (2004).

113.  Thompson, A.*, et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**, 1895-1904 (2003).

114.  Ji, J.*, et al.* Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J Chromatogr B Biomed Sci Appl* **745**, 197-210 (2000).

115.  Hsu, J.L., Huang, S.Y. & Chen, S.H. Dimethyl multiplexed labeling combined with microcolumn separation and MS analysis for time course study in proteomics. *Electrophoresis* **27**, 3652-3660 (2006).

116.  Barr, J.R.*, et al.* Isotope dilution--mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. *Clin Chem* **42**, 1676-1682 (1996).

117.  Desiderio, D.M., Kai, M., Tanzer, F.S., Trimble, J. & Wakelyn, C. Measurement of enkephalin peptides in canine brain regions, teeth, and cerebrospinal fluid with high-performance liquid chromatography and mass spectrometry. *J Chromatogr* **297**, 245-260 (1984).

118.  Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W. & Gygi, S.P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A* **100**, 6940-6945 (2003).

119.  Gerber, S.A., Kettenbach, A.N., Rush, J. & Gygi, S.P. The absolute quantification strategy: application to phosphorylation profiling of human separase serine 1126. *Methods Mol Biol* **359**, 71-86 (2007).

120.  Kirkpatrick, D.S., Gerber, S.A. & Gygi, S.P. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **35**, 265-273 (2005).

121.  Kondrat, R.W., McClusky, G.A. & Cooks, R.G. Multiple reaction monitoring in mass spectrometry/mass spectrometry for direct analysis of complex mixtures. *Analytical Chemistry* **50**, 2017-2021 (1978).

122.  Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D.A. & White, F.M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc Natl Acad Sci U S A* **104**, 5860-5865 (2007).

123.  Gergov, M., Ojanperä, I. & Vuori, E. Simultaneous screening for 238 drugs in blood by liquid chromatography-ionspray tandem mass spectrometry with multiple-reaction monitoring. *Journal of Chromatography B* **795**, 41-53 (2003).

124.  Deutsch, E.W., Lam, H. & Aebersold, R. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. 429 (Nature Publishing Group, 2008).

125.  Mallick, P.*, et al.* Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol* **25**, 125-131 (2007).

126.  Pratt, J.M.*, et al.* Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat Protoc* **1**, 1029-1043 (2006).

127.  Rivers, J., Simpson, D.M., Robertson, D.H.L., Gaskell, S.J. & Beynon, R.J. Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. 1416 (ASBMB, 2007).

_____

128.	Liu, H., Sadygov, R.G. & Yates, J.R., 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* **76**, 4193-4201 (2004).

129.	Gilchrist, A*., et al.* Quantitative proteomics analysis of the secretory pathway. *Cell* **127**, 1265-1281 (2006).

130.	Old, W.M*., et al.* Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* **4**, 1487-1502 (2005).

131.	Listgarten, J. & Emili, A. Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. 419 (ASBMB, 2005).

132.	May, D*., et al.* A platform for accurate mass and time analyses of mass spectrometry data. *J. Proteome Res* 6, 2685-2694 (2007).

_____

*Gel- Free and Gel- based proteomics in Bacillus subtilis*

# Gel-free and Gel-based Proteomics in *Bacillus subtilis*

## A COMPARATIVE STUDY·Ⓢ

**Susanne Wolff‡, Andreas Otto‡, Dirk Albrecht‡, Jianru Stahl Zeng§, Knut Büttner‡, Matthias Glückmann§, Michael Hecker‡, and Dörte Becher‡¶**

The proteome of exponentially growing *Bacillus subtilis* cells was dissected by the implementation of shotgun proteomics and a semigel-based approach for a particular exploration of membrane proteins. The current number of 745 protein identifications that was gained by the use of two-dimensional gel electrophoresis could be increased by 473 additional proteins. Therefore, almost 50% of the 2500 genes expressed in growing *B. subtilis* cells have been demonstrated at the protein level. In terms of exploring cellular physiology and adaptation to environmental changes or stress, proteins showing an alteration in expression level are of primary interest. The large number of vegetative proteins identified by gel-based and gel-free approaches is a good starting point for comparative physiological investigations. For this reason a gel-free quantitation with the recently introduced iTRAQ™ (isobaric tagging for relative and absolute quantitation) reagent technique was performed to investigate the heat shock response in *B. subtilis*. A comparison with gel-based data showed that both techniques revealed a similar level of up-regulation for proteins belonging to well studied heat hock regulons (SigB, HrcA, and CtsR). However, additional datasets have been obtained by the gel-free approach indicating a strong heat sensitivity of specific enzymes involved in amino acid synthesis. *Molecular & Cellular Proteomics 5:1183–1192, 2006.*

Two-dimensional polyacrylamide gel electrophoresis invented by O'Farrell (1) and Klose (2) is undoubtedly still the gold standard to separate complex protein mixtures. This technique allows the mass spectrometric identification of hundreds of proteins after their separation on a two-dimensional (2D)[1] gel, thereby covering an essential portion of "low complexity" proteomes such as those of bacteria. However, there are limitations to 2D PAGE that make certain classes of proteins inaccessible. Gel-critical properties include extremes in pI and molecular mass, but the most significant shortcoming is certainly the poor separation of proteins showing a pronounced hydrophobicity. The dynamic range in protein concentration that can be covered by 2D PAGE regarding non-radioactive staining methods spans 3–4 orders of magnitude, whereas the protein concentration in human blood serum extends at least 9 orders (3). For this reason a simple 2D gel approach is insufficient to analyze entire proteomes, including very low abundance proteins. The challenges inherent to a gel-based approach point to a demand for alternative techniques.

Although the characterization and quantitation of stained protein spots on 2D gels was introduced decades ago and further developed to date (4–6), attempts of protein identification and concurrent quantitation exclusively based on mass spectrometry have emerged over the last years (7). Initially the combination of multidimensional chromatography and tandem mass spectrometry that became known as shotgun proteomics was used to identify hundreds of proteins out of highly complex peptide mixtures (8–12). Very soon the first gel-free methods arose allowing a relative quantitation of proteins from different samples. A common technique is the use of stable isotope labeling of proteins or peptides, mostly realized by chemical linking of tag and biomolecule. Here one of the sample sets is provided with a "light" tag, whereas the others are linked to heavy isotope-enriched variants of the tag. Although almost all of the previously established methods (13–15) make use of a quantitation procedure based on ion signal intensity observed at the MS level, the recently introduced isobaric tagging for relative and absolute quantitation (iTRAQ™) supports a quantitation based on reporter ion signals observed at the MS/MS level that is linked with several advantages (16). First of all, the differential labeling of peptides does not challenge scan rates of mass spectrometers because the complexity at the MS level is not increased. Second, the detection of peptides originating from low abundance proteins is facilitated by the addition of ion currents of equal but differentially labeled peptides in MS spectra. On the one hand the requirement of MS/MS experiments only allows a quantitation of peptide signals exceeding a given threshold, but on the other hand the unambiguous identification of a peptide becomes more likely because it is not only based on

[1] The abbreviations used are: 2D, two-dimensional; 1D, one-dimensional; C.I., confidence interval; SCX, strong cation exchange; iTRAQ, isobaric tagging for relative and absolute quantitation.

_____

## Gel-free and Gel-based Proteomics in *B. subtilis*

the determination of its peptide mass fingerprint. Moreover iTRAQ reagent technology surpasses other gel-free quantitation methods with the capability of performing multiplex experiments in which up to four different conditions can be compared.

In this study we used shotgun proteomics to investigate the cytosolic proteome and a semigel-based approach to identify membrane proteins of vegetative *Bacillus subtilis* cells. Specifically iTRAQ reagents were used to analyze the heat shock response in *B. subtilis* as a model. Because heat shock is the best characterized stress of this organism and is associated with a substantial change in proteome signatures (17–19), it was chosen for the initial application of iTRAQ reagent technology. Resultant quantitative datasets were verified by a simultaneous protein quantitation based on 2D gels.

### EXPERIMENTAL PROCEDURES

*Sample Preparation for a Gel-free Identification of Proteins—B. subtilis* 168 wild type (20) was grown aerobically at 37 °C in a synthetic medium (21). Cells were harvested in exponential growth phase at an $A_{500}$ of 0.5. After centrifugation (8000 × *g* for 10 min at 4 °C) cell pellets were washed twice with water and again centrifuged. Cell lysis was performed in a French press (minicell, SLM Aminco, Rochester, NY). The lysate was centrifuged (20,000 × *g* for 30 min at 4 °C), and protein concentration of the supernatant was determined using Roti-Nanoquant (Roth, Karlsruhe, Germany). For the removal of impurities, proteins were extracted with phenol, precipitated by the addition of ice-cold acetone, and incubated overnight at −20 °C. Proteins were pelleted by centrifugation (20,000 × *g* for 30 min at 4 °C), air-dried, resuspended in 30 mM $NH_4HCO_3$, and digested with trypsin (Promega, Madison, WI), which was reconstituted prior to use as suggested by the manufacturer. After digestion for 16 h at 37 °C, peptides were subjected to ultracentrifugation (100,000 × *g* for 16 h at 4 °C). Aliquots of the supernatant were stored at −20 °C.

*Preparation of Membrane Proteins with Subsequent 1D SDS-PAGE Separation—*The purification of membrane proteins, their separation via 1D SDS-PAGE, and their in-gel digestion was carried out as described by Eymann *et al.* (22). Peptides were separated via reverse phase chromatography in a 3-h gradient and analyzed by an LTQ FTICR mass spectrometer (see "2D LC" and "MS/MS Analysis").

*Sample Preparation with Subsequent iTRAQ Reagent Labeling—B. subtilis* cells grown in the synthetic medium to an $A_{500}$ of 0.5 were stressed by a sudden temperature shift to 52 °C. Cell harvest occurred shortly before (control) and at 10, 30, and 60 min after continuous heat shock. Cell pellets were washed twice with ice-cold 10 mM triethylammonium bicarbonate buffer (pH = 8.5, Fluka, Taufkirchen, Germany). Lysis of *B. subtilis* cells was performed in the same buffer using a RiboLyser (HYBAID, Ashford, UK). Cell lysates were subjected to ultracentrifugation (100,000 × *g* for 2 h at 4 °C) before the protein content was determined as described above. The iTRAQ reagent labeling was performed following the protocol given by the supplier (Applied Biosystems, Foster City, CA) and according to Ross *et al.* (16). Samples were labeled as described in Table I. The mixture of iTRAQ reagent-labeled peptides was aliquoted, lyophilized, and stored at −70 °C for further analyses.

*2D LC—*The peptide mixture was separated by off-line strong cation exchange (SCX) chromatography to lower the complexity of the mixture. Removal of volatile buffer salts ($NH_4HCO_3$ or triethylammonium bicarbonate) was accomplished by drying the peptide solution via vacuum centrifugation. An amount of 3 μg of peptides was resuspended in SCX running buffer (25% (v/v) ACN, 0.1% formic acid)

### TABLE I
#### Allocation of iTRAQ reagents

| Unit mass of iTRAQ reagent reporter ion in Da | Sample |
|---|---|
| 114 | Control |
| 115 | 10 min, 52 °C |
| 116 | 30 min, 52 °C |
| 117 | 60 min, 52 °C |

prior to loading onto a SCX column (μ-Precolumn™ cartridge, BioX-SCX, 500-μm inner diameter × 15 mm, LC Packings, Amsterdam, Netherlands) using the Ettan™MDLC system (GE Healthcare). Elution of peptides off the SCX column was achieved by the injection of salt plugs of the following concentrations: 7.5, 12.5, 17.5, 25, 37.5, 50, 75, 150, 300, and 500 mM $NH_4Cl$. With a flow rate of 20 μl/min, collection time per salt fraction was set to 4 min. The SCX fractionation of iTRAQ reagent-labeled peptides had to be adjusted due to a different elution performance of peptides carrying a label. For this reason concentrations of $NH_4Cl$ salt plugs were changed to 50, 100, 150, 200, 300, 400, 500, and 750 mM. The SCX fractions of which each contained 200–300 ng of peptides were subjected to a reverse phase separation carried out on an EttanMDLC system. Prior to MS analysis peptides were loaded onto a trap column (nano-Precolumn™, PepMap™, $C_{18}$, 300-μm inner diameter × 5 mm, LC Packings) that was washed for 15 min with buffer A (0.05% (v/v) acetic acid). Elution onto the analytical column (PepMap, $C_{18}$, 75-μm inner diameter × 15 cm, LC Packings) was achieved by formation of a binary gradient (3 h) of buffer A and buffer B (90% (v/v) acetonitrile, 0.05% (v/v) acetic acid) with a flow rate of 250 nl/min.

For MALDI MS/MS analyses buffer A contained 0.1% (v/v) TFA. Buffer B consisted of 90% (v/v) acetonitrile and 0.1% (v/v) TFA. Here the binary gradient was shortened to 70 min.

*MS/MS Analysis—*LC-ESI MS experiments were performed using the Qstar® Pulsar system (Applied Biosystems MDS Sciex) and an LTQ (linear ion trap) FTICR mass spectrometer (Thermo Electron Corp., San Jose, CA).

The Qstar system was used to carry out a survey scan in the mass range of *m/z* 230–2000 in the first step. Exercising dynamic exclusion, up to four precursor ions exceeding a total ion current of 10 counts were selected for a fragmentation in MS/MS experiments. Product ions were detected in the range of *m/z* 70–2000. The LTQ FTICR mass spectrometer was used to acquire a full FT survey scan in the range of *m/z* 300–2000. Subsequently MS/MS experiments of the three most abundant precursor ions were carried out in the LTQ instrument. Meanwhile the masses of the precursor ions were determined with high accuracy via single ion mode scans in the ICR cell of the mass spectrometer.

For MALDI-TOF/TOF (4700 Proteomics Analyzer, Applied Biosystems MDS Sciex) mass spectrometry a Probot™ microfraction collector (LC Packings) was used to spot LC-separated peptides onto a MALDI target with a rate of 20 s/spot. The LC flow of 250 nl/min was mixed with matrix consisting of 2 mg/ml α-cyano-4-hydroxycinnamic acid in 70% (v/v) acetonitrile and 0.1% (v/v) TFA in a ratio of 1:5. The 4700 Proteomics Analyzer acquired MS spectra in a window of *m/z* 900–3700. The three most abundant precursor ions having a signal-to-noise ratio higher than 150 were chosen for MS/MS fragmentation, which was performed using medium collision energy.

*Data Analysis—*For an identification of proteins from LTQ FTICR data, the SEQUEST algorithm Version 27.12 (Thermo Electron Corp. (23)) was used to perform database searches against a *B. subtilis* database extracted from SubtiList (genolist.pasteur.fr/SubtiList/). A mass deviation of 0.01 Da for precursor ions as well as for fragment ions and one missed cleavage site of trypsin were allowed. The

oxidation of methionine was considered during the search. The search result was filtered using BioWorks 3.2 (Thermo Electron Corp.). A multiple threshold filter applied at the peptide level consisted of the following criteria: (*a*) peptide sequence length: 7–30 amino acids; (*b*) RSp ≤ 4; (*c*) percentage of ions: 70 (70% of all theoretical b- and y-ions of a peptide are experimentally found); (*d*) Xcorr *versus* charge state: 1.90 for singly charged ions, 2.20 for doubly charged ions, and 3.75 for triply charged ions. If the same MS/MS spectrum from the same scan was matched to different sequences, identification was assigned only to one protein that gave the better hit. Different forms of a peptide (charge state and modification) were counted as a single peptide hit. At the time this work was performed, a quantitation of iTRAQ reagent-labeled samples using the LTQ FTICR mass spectrometer was not yet possible. Therefore, only data for the purpose of protein identification were obtained.

iTRAQ reagent experiments on the Qstar system were evaluated with the software packages Pro QUANT 1.0 and Pro GROUP 1.0.2 (Applied Biosystems MDS Sciex). A mass deviation for precursor and fragment ions of 0.2 Da was permitted during the search against the *B. subtilis* database. Only top hit peptides within a confidence interval (C.I.) of 95% and a quantitation error factor below 2 were taken into account. Thereby the error factor defines the quantitative 95% C.I. of a ratio, which is the range within which the true protein ratio is 95% likely to fall. The 95% C.I. for quantitation is calculated as follows.

$$\text{lower border of 95% C.I.} = \text{protein ratio/error factor} \quad (\text{Eq. 1})$$

$$\text{upper border of 95% C.I.} = \text{protein ratio} \times \text{error factor} \quad (\text{Eq. 2})$$

For proteins quantitated with one peptide in the Qstar system analyses, error factor calculation was based on the occurrence of this peptide in different experiments.

The processing of spectra obtained with the 4700 Proteomics Analyzer was carried out with the software GPS Explorer™ Version 3.5 (Applied Biosystems). The database search was mediated by the Mascot® Version 2.0 search engine (Matrix Science Ltd., London, UK). Mass errors of 150 ppm for precursor ions and 0.2 Da for fragment ions were allowed. Only proteins that had been identified and quantified with at least two different top hit peptides within a C.I. of 95% were used for further statistical filtering. To achieve uniform error estimation with Qstar system results, standard deviations given by GPS Explorer were converted to the quantitative 95% C.I. using the formula

$$95\% \text{ C.I.} = (\mu - 2\sigma, \ \mu + 2\sigma) \quad (\text{Eq. 3})$$

where $\mu$ = experimental protein ratio and $\sigma$ = standard deviation. Proteins whose 95% C.I. showed an error factor larger than 2 were omitted. In database searches of 4700 Proteomics Analyzer data and also of Qstar system results, one missed cleavage site of trypsin, the oxidation of methionine, and a possible iTRAQ reagent labeling of tyrosine residues were considered. Samples of two independent heat shock experiments were measured in 2D LC-ESI Q-TOF as well as 2D LC-MALDI-TOF/TOF analyses. For protein quantitation the arithmetic mean of technical and biological replicates was calculated. Pro QUANT 1.0 as well as GPS Explorer Version 3.5 only determined ratios if all four iTRAQ reporter ions could be detected and if the total peak area exceeded a value of 40 counts.

*2D Gel Electrophoresis*—2D gel electrophoresis was carried out as described by Büttner *et al.* (24). To prevent a saturation of spots of high abundant proteins in the gel-based quantitation, 100 μg of protein were loaded onto IPG strips (pH range 4–7, Amersham Biosciences) in the first dimension. Gels were stained with colloidal Coomassie Brilliant Blue G-250 (Amersham Biosciences). For the processing of gel images and gel-based relative quantitation of pro-
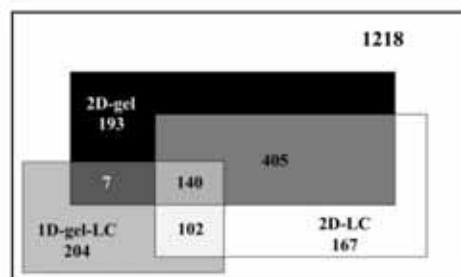


FIG. 1. **A total of 1218 protein identifications of exponentially growing *B. subtilis* cells has been gained through a combination of datasets from previous 2D PAGE experiments, 2D LC analyses of the cytosolic proteome, and 1D gel-LC studies of membrane proteins.** There is a significant number of protein identifications unique to each of the techniques used. Only 140 proteins of the 1218 could be detected by all three approaches. Because in 2D PAGE as well as in 2D LC analyses the cytosolic proteome was under investigation, there is a broad overlap of the proteins identified. The overlap of proteins demonstrated by 1D gel-LC experiments and those identified by the 2D LC and particularly the 2D PAGE approach is rather small because there membrane proteins were the focus.

tein spots the software Delta2D Version 3.3 (Decodon, Greifswald, Germany) was used. 2D PAGE was carried out for two independent heat shock experiments running a replicate for each sample. Quantitative data of this work are the arithmetic mean of technical and biological replicates.

## RESULTS

*Identification of Proteins Using 2D LC-MS/MS and 1D Gel-LC-MS/MS in Comparison with Standard 2D PAGE-MS/MS*—The cytosolic protein fraction of exponentially growing *B. subtilis* cells was analyzed using a non-gel-based approach. Two-dimensional chromatography coupled to tandem mass spectrometry resulted in the identification of 814 proteins of which 535 were identified with two or more different peptides (see Supplemental Table 3, A and B, for the list of proteins and their physical properties and their proof of identification, respectively).

Because the probability to assign false positive protein hits is much higher for protein identifications based on one peptide, the false positive rate was estimated by performing a reverse database search. Applying the aforementioned filter criteria this search revealed 50 proteins (Supplemental Table 3C), all of them identified by one peptide, leading to an error rate of 6.1%. Furthermore if an identification by two or more peptides was required a reverse database search gave no result. In conclusion the 535 proteins based on at least two peptide identifications represent highly reliable hits. Thirty-nine of the one-peptide identifications could be verified by additional peptides found in 1D gel-LC-MS/MS experiments of the membrane proteome fraction of growing *B. subtilis* cells. The remaining 240 one-hit wonders detected by 2D LC-MS/MS will require further analyses to ensure their certain identification.

The 1D gel-LC-MS/MS analysis of membrane proteins as a

## Gel-free and Gel-based Proteomics in *B. subtilis*

TABLE II

*Proteins altered in amount after heat shock and their relative quantities obtained by iTRAQ and 2D PAGE*

Proteins increased in amount were assigned to the known classes of heat shock proteins. The ones with quantity decline were allocated to functional groups and transcriptional units if possible. Protein names in boldface are subject to quantitative changes of a factor of 4 or higher. Protein names in italic have not been identified via 2D PAGE until now. Function of the proteins is according to the SubtiList database. For proteins decreased in amount a trend of transcriptional regulation is given in the column "Array" (19) with the following symbols: $--/++$ for repression/induction by a factor of 3 or higher; $-/+$ for repression/induction by a factor of at least 2; $+/-$ means no significant change in gene expression.

| Protein | iTRAQ | | | Gel | | | Number of modified forms | | Function | Array |
|---|---|---|---|---|---|---|---|---|---|---|
| | 10 min | 30 min | 60 min | 10 min | 30 min | 60 min | Charge | Mass | | |
| **Proteins increased in quantity** | | | | | | | | | | |
| HrcA-controlled genes | | | | | | | | | | |
| DnaK | 1.62 | 1.98 | 1.84 | 1.58 | 1.85 | 2.30 | 2 | 2 | Class I heat shock protein (chaperonin) | |
| GroEL | 2.14 | 2.85 | 3.04 | 3.07 | 4.32 | 5.20 | 2 | 4 | Class I heat shock protein (chaperonin) | |
| GroES | 2.33 | 3.38 | 3.29 | | | | 2 | 0 | Class I heat shock protein (chaperonin) | |
| SigmaB-dependent genes | | | | | | | | | | |
| **Ctc** | 4.31 | 5.54 | 6.23 | 3.01 | 2.87 | 3.00 | 1 | 1 | General stress protein | |
| **GsiB** | 3.39 | 5.90 | 5.27 | 11.93 | 16.57 | 17.40 | 0 | 0 | General stress protein | |
| **GspA** | 4.62 | 4.44 | 3.92 | 7.93 | 11.54 | 7.61 | 1 | 0 | General stress protein | |
| RsbV | 2.39 | 2.66 | 3.00 | | | | 1 | 0 | Positive regulator of SigmaB activity (anti-anti-sigma factor) | |
| RsbW | 2.14 | 2.05 | 2.28[a] | 2.06 | 2.22 | 1.09 | 0 | 0 | Negative regulator of SigmaB activity | |
| SodA | 1.37 | 1.59 | 1.85 | 1.48 | 1.54 | 1.97 | 1 | 0 | Superoxide dismutase | |
| TrxA | 1.60 | 1.87 | 2.04 | | | | 0 | 0 | Thioredoxin | |
| *Xpf* | 1.25 | 1.40 | 1.60 | | | | | | RNA polymerase PBSX sigma factor-like | |
| **YdaG** | 4.06 | 5.71 | 7.11 | | | | 0 | 0 | Unknown; similar to general stress protein | |
| **YflT** | 4.36 | 5.31[a] | 4.74 | | | | 0 | 0 | Unknown; similar to general stress protein | |
| **YkzA** | 2.67 | 3.73 | 4.51 | 4.54 | 8.84 | 7.56 | 0 | 0 | Unknown; similar to general stress protein | |
| YraA | 1.41 | 1.49 | 2.42 | 2.12 | 3.50 | 3.90 | 0 | 0 | Unknown; similar to general stress protein | |
| YtxH | 1.85 | 2.43 | 2.14 | 1.62 | 5.81 | 3.85 | 0 | 0 | Unknown; similar to general stress protein | |
| YvgN | 1.44 | 1.79 | 2.25 | 1.37 | 1.71 | 2.29 | 0 | 0 | Unknown; similar dehydrogenase | |
| **YvyD** | 3.29 | 4.05 | 3.85 | 5.59 | 7.43 | 7.51 | 0 | 0 | Unknown; similar to ribosomal protein S30AE family | |
| *YwdJ* | 1.23 | 1.23 | 1.55 | | | | | | Unknown; similar to unknown proteins | |
| CtsR-controlled genes | | | | | | | | | | |
| ClpC[b] | 2.23 | 2.84 | 3.06 | 3.48[c] | 6.26[c] | 4.03[c] | 3 | 0 | Class III stress response-related ATPase | |
| **ClpE** | 3.72 | 4.76 | 1.59 | 23.31 | 2.32 | 3.41[c] | 1 | 0 | ATP-dependent Clp protease-like | |
| **ClpP**[b] | 1.83 | 3.09 | 4.25 | 2.79 | 6.11 | 7.40 | 1 | 1 | ATP-dependent Clp protease proteolytic subunit | |
| Other mechanisms | | | | | | | | | | |
| AhpC | 1.25 | 1.42 | 1.52 | 1.68 | 1.87 | 1.86 | 2 | 2 | Alkyl hydroperoxide reductase (small subunit) | |
| *CopZ (YvgY)* | 1.51 | 2.48 | 2.81 | | | | | | Metallochaperone, binds Cu(I), important for copper resistance (41) | |
| HtpG | 1.65 | 2.24 | 2.73 | 1.24 | 0.79 | 3.66 | 2 | 0 | Heat shock protein (chaperonin) | |
| IolS | 1.32 | 1.65 | 1.96 | 1.33 | 1.83 | 2.23 | 0 | 0 | *myo*-Inositol catabolism | |
| *YitW* | 1.44 | 1.61 | 1.67 | | | | | | Unknown; similar to unknown proteins | |
| YjoA | 1.26 | 1.66 | 2.32 | | | | 0 | 0 | Unknown | |
| *YlbO* | 1.33 | 1.71 | 1.51 | | | | | | Involved in coat protein formation, similar to RsfA (42) | |
| *YtpP* | 1.20 | 1.38 | 1.56 | | | | | | Thioredoxin-like protein, under Spx control (43) | |
| YugJ | 1.28 | 1.50 | 1.71 | | | | 0 | 0 | Unknown; similar to NADH-dependent butanol dehydrogenase | |

TABLE II—*continued*

| Protein | iTRAQ | | | Gel | | | Number of modified forms | | Function | Array |
|---|---|---|---|---|---|---|---|---|---|---|
| | 10 min | 30 min | 60 min | 10 min | 30 min | 60 min | Charge | Mass | | |
| **Proteins decreased in quantity** | | | | | | | | | | |
| Branched-chain amino acid synthesis | | | | | | | | | | |
| *ilvB* operon | | | | | | | | | | |
| **LeuA** | 0.76 | 0.50 | 0.43 | 0.81 | 0.63 | 0.37 | 1 | 1 | 2-Isopropylmalate synthase | + |
| **LeuC** | 0.40 | 0.14[a] | 0.12[a] | 0.50[c] | 0.13[c] | 0.08 | 1 | 0 | 3-Isopropylmalate dehydratase (large subunit) | + |
| **LeuD** | 0.49 | 0.22[a] | 0.19[a] | 0.55 | 0.12 | 0.08 | 0 | 0 | 3-Isopropylmalate dehydratase (small subunit) | +/− |
| IlvA | 0.89 | 0.65 | 0.50 | | | | 0 | 0 | Threonine dehydratase | − |
| **IlvE (YwaA)** | 0.62 | 0.18 | 0.17 | 0.70 | 0.48 | 0.38 | 1 | 0 | Unknown, similar to branched-chain amino acid aminotransferase | + + |
| YkwC | 0.64 | 0.32[a] | 0.30[a] | | | | 1 | 0 | Unknown; similar to 3-hydroxyisobutyrate dehydrogenase | + + |
| S box regulon | | | | | | | | | | |
| *cysH* operon | | | | | | | | | | |
| CysC | 0.85 | 0.53 | 0.40[a] | | | | 0 | 1 | Probable adenylylsulfate kinase | − − |
| Sat | 1.04 | 0.79 | 0.35 | 0.78 | 0.08 | 0.12 | 1 | 0 | Probable sulfate adenylyltransferase | − − |
| *mtnK* operon | | | | | | | | | | |
| **MtnK (YkrT)** | 0.50 | 0.19 | 0.17 | 0.52 | 0.30 | 0.17 | 1 | 0 | Methylthioribose kinase (37) | +/− |
| MtnA (YkrS) | 0.80 | 0.47 | 0.42 | | | | 1 | 1 | Methylthioribose-1-phosphate isomerase (37) | +/− |
| MetC (YjcJ) | 0.93 | 0.62 | 0.42 | | | | 0 | 1 | Cystathionine β-lyase (44) | + + |
| **MetE** | 0.32 | 0.14 | 0.11 | 0.20 | 0.05[c] | 0.03[c] | 5 | 3 | Cobalamin-independent methionine synthase | + + |
| **MtnD (YkrZ)** | 0.50 | 0.19 | 0.16 | 0.76 | 0.17 | 0.06 | 0 | 0 | Aci-reductone dioxygenase (37) | + + |
| AhrC regulon | | | | | | | | | | |
| ArgG | 0.85 | 0.52 | 0.41 | 0.89 | 0.41 | 0.21 | 1 | 0 | Argininosuccinate synthase | + + |
| ArgJ | 0.92 | 0.64 | 0.43 | | | | 0 | 0 | Ornithine acetyltransferase/amino-acid acetyltransferase | + + |
| Nucleotide metabolism | | | | | | | | | | |
| **NadA** | 0.46 | 0.19 | 0.14 | 0.65 | 0.22 | 0.12 | 0 | 0 | Quinolinate synthetase | +/− |
| *NifS* | 0.79 | 0.50 | 0.37[a] | | | | | | Probable cysteine desulfurase, required for NAD biosynthesis | +/− |
| **PurC** | 0.55 | 0.20 | 0.13[a] | 0.76 | 0.41 | 0.34 | 1 | 0 | Phosphoribosylaminoimidazole-succinocarboxamide synthetase | − − |
| PyrF | 0.52 | 0.45[a] | 0.42[a] | 1.18 | 1.11 | 0.49[c] | 0 | 0 | Orotidine-5′-phosphate decarboxylase | − − |
| ThiC | 0.86 | 0.61 | 0.36 | 0.65 | 0.14[c] | 0.16 | 1 | 0 | Biosynthesis of the pyrimidine moiety of thiamin | +/− |
| Other | | | | | | | | | | |
| **CspC** | 0.43 | 0.18 | 0.11 | | | | 0 | 0 | Cold shock protein | − |
| DivIVA | 0.74 | 0.46 | 0.39[a] | 1.18 | 0.43 | 0.21 | 1 | 0 | Cell-division initiation protein (septum placement) | +/− |
| Efp | 0.82 | 0.57 | 0.47 | 0.78 | 0.31 | 0.17 | 1 | 0 | Elongation factor P | +/− |
| FusA | 0.69 | 0.31 | 0.32 | 0.70 | 0.56 | 0.40 | 2 | 2 | Elongation factor G | − |
| PpiB | 0.81 | 0.38 | 0.26 | 1.59 | 0.37 | 0.37 | 1 | 0 | Peptidyl-prolyl isomerase | − |
| RplM | 0.48 | 0.69 | 1.33 | | | | 0 | 0 | Ribosomal protein L13 | +/− |
| **YciC** | 0.34 | 0.13 | 0.12 | 0.49 | 0.10[c] | 0.11[c] | 1 | 1 | Unknown; similar to unknown proteins | +/− |
| YjlD | 0.81 | 0.65 | 0.49 | 0.51[c] | 0.11[c] | 0.25 | 3 | 0 | Unknown; similar to NADH dehydrogenase | − |
| CspB | 0.88 | 0.59 | 0.35 | | | | 1 | 0 | Major cold shock protein | − − |
| PdxT (YaaE) | 0.82 | 0.47 | 0.37 | | | | 1 | 0 | Glutamine amidotransferase subunit pdxT | − |

[a] iTRAQ ratios showed an error factor larger than 2 due to high signal-to-noise values caused by strong alterations in quantity. The 95% C.I. of all iTRAQ datasets is given in Supplemental Table 4A.
[b] Also under control of SigmaB.
[c] Gel-based data with a relative standard deviation higher than 30%.

semigel-based technique resulted in 453 protein identifications, 265 of them based on two or more peptides (Supplemental Table 4, A and B, give proteins and proof of identification, respectively). A reverse database search of the datasets from 1D gel-LC-MS/MS experiments led to 16 protein hits, leaving the 453 protein identifications with a false

FIG. 2. **Quantitation profile of enzymes involved in the synthesis of methionine (A), arginine (B), and branched-chain amino acids (C) that displayed a decrease in quantity after heat shock.** A, MetC, cystathionine β-lyase; MtnA, methylthioribose-1-phosphate isomerase; CysC, probable adenylylsulfate kinase; Sat, probable sulfate adenylyltransferase; MtnK, methylthioribose kinase; MtnD, aci-reductone dioxygenase; MetE, cobalamin-independent methionine synthase. B, ArgJ, ornithine acetyltransferase; ArgD, N-acetylornithine aminotransferase; CarB, carbamoylphosphate transferase-arginine (subunit B); ArgF, ornithine carbamoyltransferase; ArgG, argininosuccinate synthase. C, IlvB, acetolactate synthase (large subunit); IlvH, acetolactate synthase (small subunit); IlvC, ketol-acid reductoisomerase; IlvA, threonine dehydratase; LeuA, 2-isopropylmalate synthase; LeuC, 3-isopropylmalate dehydratase (large subunit); LeuD, 3-isopropylmalate dehydratase (small subunit).
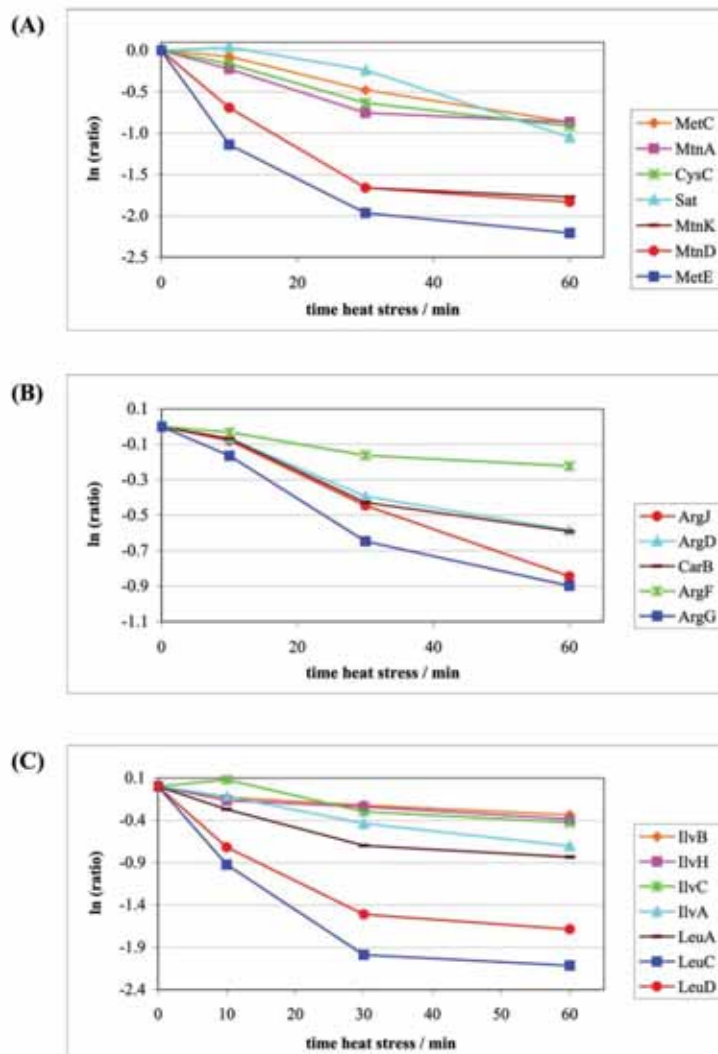
positive rate of 3.5% (Supplemental Table 4C). A fraction of 51.2% of the 453 proteins possessed transmembrane domains (25) indicating that not all cytosolic proteins could be removed in the course of the isolation procedure of membrane proteins. Still this semigel-based study revealed 204 proteins that were not accessible by standard 2D PAGE-MS/MS or 2D LC-MS/MS. Combining current protein identifications of all three approaches gave rise to a number of 1218 proteins. Fig. 1 shows the allocation of protein identifications to the single techniques as well as the overlap between them.

*Proteins Altered in Quantity after Heat Shock*—Having the capability of identifying hundreds of proteins from complex peptide mixtures provides the basis to carry out global quantitative proteome studies, which is one major concern in terms of unraveling cellular physiology. The recently introduced

iTRAQ reagent technique was used to investigate the cytosolic proteome fraction in the heat shock response of *B. subtilis* as a well characterized physiological model (17–19).

Most of the proteins showing increases in quantity in response to heat shock are members of well characterized classes of heat shock proteins (Table II) and are therefore not extensively discussed at this point. They are either under control of the transcriptional repressor HrcA (26–28), belong to the SigmaB regulon (29–33), are repressed by CtsR (34, 35), or are subject to another, still unknown mechanism of regulation. Although the number of up-regulated proteins reported here only represents a fraction of the entirety of all heat-induced proteins (18), the fact that the iTRAQ approach assessed their increased concentration verifies the reliability of the gel-free method. The non-gel-based technique further justified its potential by the identification of six heat-induced proteins (CopZ, Xpf, YitW,

FIG. 3. **Correlation of relative iTRAQ reagent- and gel-based quantitation for proteins with significant changes in amount after continuous heat shock for the three time points 10, 30, and 60 min.**

$$y = 1.209x$$
$$R^2 = 0.817$$

ln(ratio gel)

ln(ratio iTRAQ™)

- 10 min heat
- 30 min heat
- 60 min heat
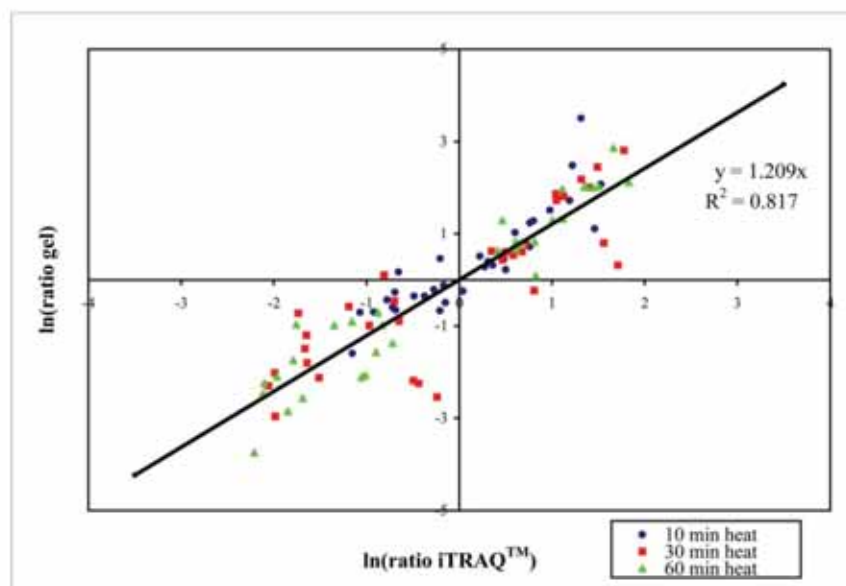
YlbO, YtpP, and YwdJ) that to our knowledge have not been accessible by 2D PAGE to date. Xpf and YwdJ are known to be under SigB control, but the heat induction mechanisms of the remaining four proteins are still unknown.

Nearly all of the proteins that were found to be degraded after heat shock assume "housekeeping" functions in the cell (Table II). In addition to enzymes associated with nucleotide metabolism, ribosomal proteins, and elongation factors, the number of proteins involved in the biosynthesis of important amino acids is most striking. These enzymes could be assigned to three major groups, the S box regulon (36, 37), arginine biosynthesis, and synthesis of branched-chain amino acids. Their quantitation profiles are given in Fig. 2, A–C, respectively.

*iTRAQ Reagent-based Quantitation of Proteins Confirmed by 2D PAGE*—To evaluate iTRAQ reagent-obtained data, a comparison with the well established quantitation via 2D PAGE was drawn. For this purpose extracts originating from the same heat shock experiments were subjected to separation on 2D gels, and protein spots were subsequently quantitated using the software Delta2D. Applying qualitative and statistic filters 2D LC-MS/MS analyses of the iTRAQ reagent-labeled samples resulted in the reliable quantitation of 292 proteins (see Supplemental Table 5, A and B, for relative protein ratios and their proof of identification, respectively). Only proteins that were subject to an alteration in their amount by a ratio outside of 0.5–1.5 at one of the time points (determined by iTRAQ technology) will be discussed below. Regarding proteins with no change in quantity, datasets of both approaches were in good correspondence (data not shown). The criterion applied to 63 proteins, revealing either an increment or reduction in their amount after heat shock (Table II). Thirty-two of the proteins could be localized on 2D gels ac-

quired in parallel. Overall protein quantities determined at the peptide level via iTRAQ reagents correlated very well with gel-based quantitation at the protein level (Fig. 3). Proteins showing discrepancies in quantity need to be looked at in greater detail, for instance at their appearance on the 2D gel. In the following a few representative examples will be given that demonstrate the sound agreement of the two different quantitation methods but also the pitfalls of which one has to be aware.

The gel-based quantitation of proteins showing only one distinct spot on the gel agreed mostly with the data obtained by the iTRAQ reagent method. YvgN (unknown; possible dehydrogenase; SigmaB-controlled (29)) with an increased amount and NadA (quinolinate synthetase) with a reduced amount are members of this group of proteins. In contrast, the chaperone GroEL and elongation factor G (FusA) scattered over more than one spot on the gel posing a challenge to the gel-based quantitation. Proteins with a change in concentration by more than a factor of 4 were categorized as "on/off" proteins (38). At one of the time points they appeared only with a very low signal-to-noise ratio on the 2D gel. The consequence was a distorted gel-based quantitation result leading to protein ratios that were much higher ("on" for protein GsiB (general stress protein)) or lower ("off" for protein MetE (cobalamin-independent methionine synthase)) than gel-free derived data. Comparative quantity profiles for the exemplary proteins are given in Fig. 4, A–C.

DISCUSSION

The proteome of exponentially growing *B. subtilis* cells was extensively investigated by Eymann *et al.* (22). By the use of the conventional 2D PAGE these authors were able to identify 745 proteins. The present study describes the expansion of the comprehensive vegetative proteome map of *B. subtilis* by
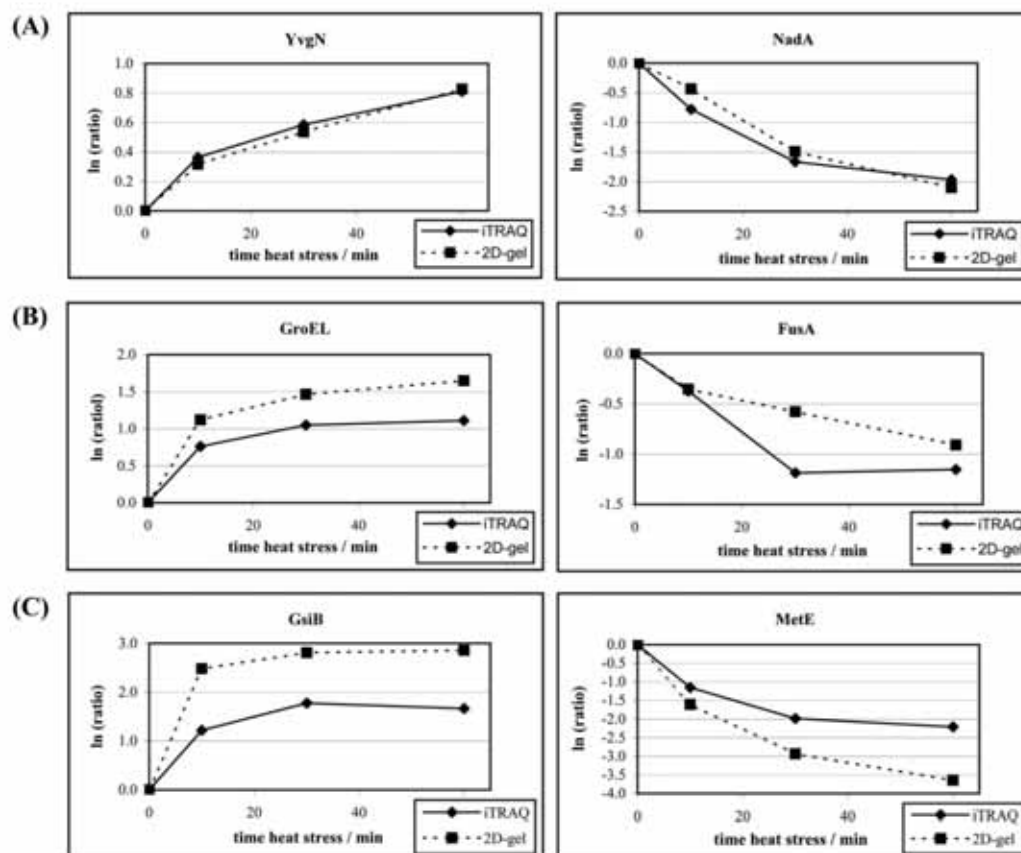
43

_____

**Gel-free and Gel-based Proteomics in *B. subtilis***
_____



FIG. 4. *A*, gel-based and non-gel-based quantitation of YvgN (unknown; similar dehydrogenase) and NadA (quinolinate synthetase). Both proteins appear as single spots on a 2D gel. *B*, correlation of iTRAQ reagent quantitation and relative protein amounts obtained by 2D PAGE for the proteins GroEL (chaperone) and FusA (elongation factor G), which form multiple spots on a 2D gel. *C*, gel-based and gel-free quantitation of the general stress protein GsiB and the cobalamin-independent methionine synthase MetE, both of which showed an extreme alteration in amount after heat shock.

473 proteins. A gel-free analysis of cytosolic proteins and a semigel-based technique for the dissection of the membrane proteome enabled the identification of proteins obviously not accessible by 2D PAGE. With an assumption of 2500 genes being transcribed under exponential growth conditions (22) half of the *B. subtilis* vegetative proteome has been covered now and is therefore ready for physiological exploration.

iTRAQ reagent technology was used to carry out quantitative studies on the heat shock response in *B. subtilis*. A simultaneously performed 2D gel quantitation as well as earlier published data (18) confirmed the correctness of the results obtained by the gel-free approach. Through the direct comparison of 2D gel and iTRAQ reagent quantitation, the strength and weaknesses of both approaches were disclosed. The iTRAQ reagent technique is most suitable to determine the sum of all subspecies of a protein. Consequently the problem of multiple protein spots or comigration of proteins inherent to a 2D gel quantitation is avoided. Aiming at a quantitation of single variants of a protein, 2D PAGE is cur-

rently the appropriate technique because modifications that result in a change of pI or molecular mass come to light on a 2D gel. A major concern in the relative quantitation of proteins is the question of dynamic range of the technique used. An exceedingly narrow range will cause difficulties in the study of proteins showing massively altered quantities. Although we found the iTRAQ reagent technique to have a wider dynamic range than a colloidal Coomassie-stained 2D gel, the problem of an incorrect quantitation due to very low signal-to-noise ratios for one of the samples remains a challenge of both the gel-based and the gel-free approach.

In addition to the expected up-regulation of the well known classes of heat shock proteins, we could provide new physiological insights on the heat shock response of *B. subtilis*. Despite intensive gel-based studies, the heat-induced large scale degradation of enzymes involved in amino acid anabolism has not been reported for this organism yet. Biran *et al.* (39) found MetA (homoserine transsuccinylase) of *Escherichia coli* to be inactivated upon heat shock and interpreted the

heat sensitivity of this important enzyme in methionine bio-synthesis as a control element of growth rate during heat stress. The same purpose can be presumed regarding the heat susceptibility of essential enzymes in amino acid synthesis in *B. subtilis*.

Most interesting is a comparison between the protein quantitation of this survey and results from mRNA profiling experiments by Helmann *et al.* (19). Although the concentration of many proteins involved in methionine, arginine, and branched-chain amino acid synthesis was decreasing, Helmann *et al.* (19) determined several of the corresponding genes to be induced after heat shock, some of them even in a dramatic fashion (Table II). One has to note that expression of these genes is controlled by a feedback mechanism in which the metabolic product represents the molecular effector. Therefore, a decreasing quantity of the above mentioned enzymes with the resulting deprivation of certain amino acids most likely caused an increased transcription of the corresponding genes after heat shock. The comparison drawn between array data and relative protein amounts proved once more that mRNA profiling alone is not sufficient to unravel cellular and physiological adaptation mechanisms upon environmental changes (40).

There remains the question about the fate of proteins found to be reduced in quantity. Most likely those proteins are denatured upon heat shock and are either prone to proteolysis or to sedimentation during ultracentrifugation after cell lysis. To answer this question a comparable experiment with a mutant deficient in ClpP, the major cytosolic protease in *B. subtilis*, was set up, and a qualitative analysis of the ultracentrifugation sediment was carried out (data not shown). Many proteins reduced in quantity in the wild type but stable in a *clpP* mutant after heat shock could be assigned as ClpP substrates under chosen conditions. Several of the proteins whose amounts decreased in the wild type as well as in the *clpP* mutant at 52 °C could be identified from the ultracentrifugation sediment, indicating their heat-induced denaturation and subsequent sedimentation.

Overall results of this work demonstrated that the gel-based approach in which 745 proteins were identified, the gel-free studies revealing 814 proteins, and the 1D gel-LC-MS/MS experiments resulting in 453 protein identifications of which 232 proteins possess transmembrane domains represent complementary techniques, mutually compensating for each other's limitations. Through the parallel implementation of these different approaches state of the art technology will be utilized to decipher the details and unresolved aspects in physiological proteomics.

¶ To whom correspondence should be addressed: Inst. for Microbiology, Ernst-Moritz-Arndt-University Greifswald, Friedrich-Ludwig-Jahn-Str. 15, D-17487 Greifswald, Germany. Tel.: 49-3834-864208; Fax: 49-3834-864202; E-mail: dbecher@uni-greifswald.de.

REFERENCES

1. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250,** 4007–4021
2. Klose, J. (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26,** 231–243
3. Adkins, J. N., Varnum, S. M., Auberry, K. J., Moore, R. J., Angell, N. H., Smith, R. D., Springer, D. L., and Pounds, J. G. (2002) Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteomics* **1,** 947–955
4. Bossinger, J., Miller, M. J., Vo, K. P., Geiduschek, E. P., and Xuong, N. H. (1979) Quantitative analysis of two-dimensional electrophoretograms. *J. Biol. Chem.* **254,** 7986–7998
5. Lipkin, L. E., and Lemkin, P. F. (1980) Data-base techniques for multiple two-dimensional polyacrylamide gel electrophoresis analyses. *Clin. Chem.* **26,** 1403–1412
6. Aittokallio, T., Salmi, J., Nyman, T. A., and Nevalainen, O. S. (2005) Geometrical distortions in two-dimensional gels: applicable correction methods. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **815,** 25–37
7. Tao, W. A., and Aebersold, R. (2003) Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. *Curr. Opin. Biotechnol.* **14,** 110–118
8. Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., III (1999) Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17,** 676–682
9. Wolters, D. A., Washburn, M. P., and Yates, J. R., III (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **73,** 5683–5690
10. Davis, M. T., Beierle, J., Bures, E. T., McGinley, M. D., Mort, J., Robinson, J. H., Spahr, C. S., Yu, W., Luethy, R., and Patterson, S. D. (2001) Automated LC-LC-MS-MS platform using binary ion-exchange and gradient reversed-phase chromatography for improved proteomic analyses. *J. Chromatogr. B Biomed. Sci. Appl.* **752,** 281–291
11. Washburn, M. P., Wolters, D., and Yates, J. R., III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19,** 242–247
12. Lipton, M. S., Pasa-Tolic, L., Anderson, G. A., Anderson, D. J., Auberry, D. L., Battista, J. R., Daly, M. J., Fredrickson, J., Hixson, K. K., Kostandarithes, H., Masselon, C., Markillie, L. M., Moore, R. J., Romine, M. F., Shen, Y., Stritmatter, E., Tolic, N., Udseth, H. R., Venkateswaran, A., Wong, K. K., Zhao, R., and Smith, R. D. (2002) Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 11049–11054
13. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999
14. Chakraborty, A., and Regnier, F. E. (2002) Global internal standard technology for comparative proteomics. *J. Chromatogr. A* **949,** 173–184
15. Schmidt, A., Kellermann, J., and Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5,** 4–15
16. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces*

_____

## Gel-free and Gel-based Proteomics in *B. subtilis*

cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154–1169

17. Hecker, M., Schumann, W., and Völker, U. (1996) Heat-shock and general stress response in *Bacillus subtilis*. *Mol. Microbiol.* **19**, 417–428

18. Bernhardt, J., Büttner, K., Scharf, C., and Hecker, M. (1999) Dual channel imaging of two-dimensional electropherograms in *Bacillus subtilis*. *Electrophoresis* **20**, 2225–2240

19. Helmann, J. D., Wu, M. F., Kobel, P. A., Gamo, F. J., Wilson, M., Morshedi, M. M., Navre, M., and Paddon, C. (2001) Global transcriptional response of *Bacillus subtilis* to heat shock. *J. Bacteriol.* **183**, 7318–7328

20. Anagnostopoulos, C., and Spizizen, J. (1961) Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* **81**, 741–746

21. Stülke, J., Hanschke, R., and Hecker, M. (1993) Temporal activation of beta-glucanase synthesis in *Bacillus subtilis* is mediated by the GTP pool. *J. Gen. Microbiol.* **139**, 2041–2045

22. Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam le, T., Büttner, K., Buurman, G., Scharf, C., Venz, S., Völker, U., and Hecker, M. (2004) A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics* **4**, 2849–2876

23. Ducret, A., Van Oostveen, I., Eng, J. K., Yates, J. R., III, and Aebersold, R. (1998) High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein Sci.* **7**, 706–719

24. Büttner, K., Bernhardt, J., Scharf, C., Schmid, R., Mäder, U., Eymann, C., Antelmann, H., Völker, A., Völker, U., and Hecker, M. (2001) A comprehensive two-dimensional map of cytosolic proteins of *Bacillus subtilis*. *Electrophoresis* **22**, 2908–2935

25. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580

26. Schmidt, A., Schiesswohl, M., Volker, U., Hecker, M., and Schumann, W. (1992) Cloning, sequencing, mapping, and transcriptional analysis of the groESL operon from *Bacillus subtilis*. *J. Bacteriol.* **174**, 3993–3999

27. Yuan, G., and Wong, S. L. (1995) Isolation and characterization of *Bacillus subtilis* groE regulatory mutants: evidence for orf39 in the dnaK operon as a repressor gene in regulating the expression of both groE and dnaK. *J. Bacteriol.* **177**, 6462–6468

28. Schulz, A., and Schumann, W. (1996) hrcA, the first gene of the *Bacillus subtilis* dnaK operon encodes a negative regulator of class I heat shock genes. *J. Bacteriol.* **178**, 1088–1093

29. Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J. D., Völker, U., and Hecker, M. (2001) Global analysis of the general stress response of *Bacillus subtilis*. *J. Bacteriol.* **183**, 5617–5631

30. Binnie, C., Lampe, M., and Losick, R. (1986) Gene encoding the sigma 37 species of RNA polymerase sigma factor from *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5943–5947

31. Petersohn, A., Bernhardt, J., Gerth, U., Höper, D., Koburger, T., Völker, U., and Hecker, M. (1999) Identification of sigma(B)-dependent genes in *Bacillus subtilis* using a promoter consensus-directed search and oligo-nucleotide hybridization. *J. Bacteriol.* **181**, 5718–5724

32. Petersohn, A., Antelmann, H., Gerth, U., and Hecker, M. (1999) Identification and transcriptional analysis of new members of the sigmaB regulon in *Bacillus subtilis*. *Microbiology* **145**, 869–880

33. Price, C. W., Fawcett, P., Ceremonie, H., Su, N., Murphy, C. K., and Youngman, P. (2001) Genome-wide analysis of the general stress response in *Bacillus subtilis*. *Mol. Microbiol.* **41**, 757–774

34. Krüger, E., and Hecker, M. (1998) The first gene of the *Bacillus subtilis* clpC operon, ctsR, encodes a negative regulator of its own operon and other class III heat shock genes. *J. Bacteriol.* **180**, 6681–6688

35. Derre, I., Rapoport, G., and Msadek, T. (1999) CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in gram-positive bacteria. *Mol. Microbiol.* **31**, 117–131

36. Grundy, F. J., and Henkin, T. M. (1998) The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Mol. Microbiol.* **30**, 737–749

37. Sekowska, A., Denervaud, V., Ashida, H., Michoud, K., Haas, D., Yokota, A., and Danchin, A. (2004) Bacterial variations on the methionine salvage pathway. *BMC Microbiol.* **4**, article 9

38. Kolkman, A., Dirksen, E. H., Slijper, M., and Heck, A. J. (2005) Double standards in quantitative proteomics: direct comparative assessment of difference in gel electrophoresis and metabolic stable isotope labeling. *Mol. Cell. Proteomics* **4**, 255–266

39. Biran, D., Brot, N., Weissbach, H., and Ron, E. Z. (1995) Heat shock-dependent transcriptional activation of the metA gene of *Escherichia coli*. *J. Bacteriol.* **177**, 1374–1379

40. Griffin, T. J., Gygi, S. P., Ideker, T., Rist, B., Eng, J., Hood, L., and Aebersold, R. (2002) Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **1**, 323–333

41. Gaballa, A., and Helmann, J. D. (2003) *Bacillus subtilis* CPx-type ATPases: characterization of Cd, Zn, Co and Cu efflux systems. *Biometals* **16**, 497–505

42. Kuwana, R., Okumura, T., Takamatsu, H., and Watabe, K. (2005) The ylbO gene product of *Bacillus subtilis* is involved in the coat development and lysozyme resistance of spore. *FEMS Microbiol. Lett.* **242**, 51–57

43. Nakano, S., Kuster-Schock, E., Grossman, A. D., and Zuber, P. (2003) Spx-dependent global transcriptional control is induced by thiol-specific oxidative stress in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13603–13608

44. Auger, S., Yuen, W. H., Danchin, A., and Martin-Verstraete, I. (2002) The metIC operon involved in methionine biosynthesis in *Bacillus subtilis* is controlled by transcription antitermination. *Microbiology* **148**, 507–518

46

*Towards the entire proteome of the model bacterium Bacillus subtilis by gel-based and gel-free approaches*

Review

# Towards the entire proteome of the model bacterium *Bacillus subtilis* by gel-based and gel-free approaches☆

Susanne Wolff[a], Haike Antelmann[a], Dirk Albrecht[a], Dörte Becher[a], Jörg Bernhardt[a],
Sierd Bron[b], Knut Büttner[a], Jan Maarten van Dijl[c], Christine Eymann[a],
Andreas Otto[a], Le Thi Tam[a], Michael Hecker[a,*]

[a] *Institut für Mikrobiologie, Ernst-Moritz-Arndt-Universität Greifswald, Friedrich-Ludwig-Jahn-Str. 15,
D-17487 Greifswald, Germany*
[b] *Department of Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, Kerklaan 30,
9751 NN Haren, the Netherlands*
[c] *Department of Medical Microbiology, University Medical Centre Groningen and University of Groningen, Hanzeplein 1,
P.O. Box 30001, 9700 RB Groningen, the Netherlands*

## Abstract

With the emergence of mass spectrometry in protein science and the availability of complete genome sequences, proteomics has gone through a rapid development. The soil bacterium *Bacillus subtilis*, as one of the first DNA sequenced species, represents a model for Gram-positive bacteria and its proteome was extensively studied throughout the years. Having the final goal to elucidate how life really functions, one basic requirement is to know the entirety of cellular proteins. This review presents how far we have got in unraveling the proteome of *B. subtilis*. The application of gel-based and gel-free technologies, the analyses of different subcellular proteome fractions, and the pursuance of various physiological strategies resulted in a coverage of more than one-third of *B. subtilis* theoretical proteome.
© 2006 Elsevier B.V. All rights reserved.

*Keywords:* Proteome coverage; Stress and starvation signatures; Subproteomic fractions; Gel-based and gel-free proteomics

## Contents

_____

☆ This paper is part of a special volume entitled "Analytical Tools for Proteomics", guest edited by Erich Heftmann.
* Corresponding author. Tel.: +49 3834864200; fax: +49 3834864202.
  *E-mail address:* hecker@uni-greifswald.de (M. Hecker).

## 1. Introduction

In the year 1995 proteomics [1] was coined to study the diverse properties of proteins combined with the elucidation of their cellular function. With the achievement to decipher entire genome sequences [2–4] protein researchers were supplied with the blueprint of every possible gene product of an organism. One initial goal in proteomics was the identification of all proteins expressed by a cell as the equivalent to the genome. Meanwhile also the exact determination of amino acid sequences, post-translational modifications, interaction partners, protein quantity, synthesis, stability, activity, localization and structure have been focused on, whereby the protein identification still forms the groundwork for such kind of studies. There is no single technology that could be utilized to answer all these versatile problems, but certainly these issues will be addressed profoundly by the still ongoing development of complementary technology platforms.

Whereas the complete sequencing of a genome is a finite undertaking, it is virtually impossible to determine an organisms proteome as a whole due to the significantly higher complexity mainly caused by protein targeting, post-translational modifications as well as mRNA splicing and protein processing. The invention of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) by O'Farrell [5] and Klose [6] made an extensive separation of hundreds or even thousands of protein species possible and therefore brought the goal of a preferably complete proteome dissection closer. Due to the development of mild ionization methods [7,8] mass spectrometry was entering protein science. Combined with the introduction of database search algorithms [9–13], proteins separated on 2D-gels could now be identified in a significantly higher throughput compared with the hitherto performed N-terminal sequencing via Edman degradation [14].

Although 2D-PAGE proved to be a powerful separation technique, it has its downsides in resolving proteins with extremes in isoelectric point (p$I$) or molecular mass, and most unfavorable is certainly the poor resolution of hydrophobic proteins such as those spanning the cellular membrane. To gain access also to this gel-critical fraction of the proteome alternative techniques arose based on a gel-free sample separation at peptide rather than at protein level [15–17]. This approach became famous as shotgun proteomics and allowed the unbiased identification of hundreds and even thousands of proteins from highly complex protein mixtures [18,19]. Thus, a remarkable fraction of the proteome for several "low complexity" organisms could already be covered [20–24].

One of these model organisms is the well-studied soil bacterium *Bacillus subtilis*. Already by the mid-20th century some species of the genus *Bacillus* were known more generally for their roles in human and animal infections and others were valued as producers of antibiotics and industrial enzymes. Having become the paradigm for Gram-positive bacteria, *B. subtilis* is an ideal experimental system for studying mechanisms of gene regulation, metabolism and cell differentiation also in view of investigations of highly related microorganisms, such as the human pathogen *Staphylococcus aureus*. Because of the far-reaching history of research on *B. subtilis*, it was among the first species being fully DNA sequenced [25]. With the whole genome sequence available not only mRNA profiling of gene expression could be performed, but also the gate towards high-throughput protein identification was pushed open. Though we can look back on 20 years of proteome research in *B. subtilis* [26,27], still one-third of *B. subtilis* 4100 genes are not assigned to a defined function yet [28], and only for about one-third of all genes the corresponding protein could be demonstrated so far [24]. This review gives an overview on the strategies we explored with the objective to increase proteome coverage of *B. subtilis*. The results clearly evince the necessity to employ different technology platforms, but also tell us that it is still a long way to the implementation of complete proteome coverage, which will take us closer to the intrinsic goal of understanding how a single cell actually functions.

## 2. Cytosolic proteome map for growing and non-growing cells of *B. subtilis*

Two-dimensional gel electrophoresis with its high-resolution power is an obvious method to detect as many proteins as possible and, therefore, provides the basis for comprehensive physiological proteome studies. The technique was used successfully for a far-reaching proteome coverage in several bacteria, e.g. *Haemophilus influenzae* [29]. The introduction of proteomic approaches based on liquid chromatography coupled to mass spectrometry gave an important stimulus in the further increase of proteome coverage. Lipton et al. published the identification of 61% of the theoretical proteome of *Deinococcus radiodurans* [21]. With 1480 proteins identified, 35% of the *Escherichia coli* proteome could be covered [23], and in the "minimal" organism *Mycoplasma mobile* even 88% of the 635 open reading frames (ORFs) have been detected on protein level [30]. Combining 2D-PAGE to study cytosolic proteins and a gel-free approach for an analysis of membrane proteins Oesterhelt and coworkers were able to identify 34% of the theoretical proteome of the halophilic archaeon *Halobacterium salinarum* [31,32].

However, the depiction of an entire proteome does not solely depend on the analytical tools and their continuous improvement. A considerable part of the genome remains more or less silent under standard growth conditions and will be expressed only in response to specific stimuli. Therefore, two major classes of proteomes have to be defined: the proteome of growing cells (vegetative proteome) and the proteome of non-growing cells suffering from stress or starvation.

### 2.1. Cytosolic proteome map of growing cells

Macroarray analyses revealed that approximately 2500 of *B. subtilis* 4107 genes are actively transcribed under exponential growth conditions [33]. A first comprehensive map of the vegetative *B. subtilis* proteome was published by Büttner et al. [34], who were able to identify more than 300 cytosolic proteins by standard 2D-PAGE located in an analytical window of p$I$ 4–7. Three years later Eymann et al. [33] expanded the cytosolic

proteome map of growing cells to a total number of 693 proteins in the region p*I* 4–7. This substantial increase in protein identifications can be ascribed to refined experimental protocols, to further sophistication in mass spectrometric technology and to the utilization of several ultra zoom gels (e.g. p*I* 4.5–5.5) which improved the resolution of protein separation in the first dimension on a 2D-gel. Proteins having a p*I* beyond 7, which theoretically applies to one-third of the *B. subtilis* proteome, are not covered by conventional 2D-PAGE (p*I* 4–7). Consequently, protocols have been worked out to enable the electrophoretic separation of alkaline proteins [35,36]. Ohlmeier et al. [37] used immobilized pH gradients (IPGs) covering a pH range from 4 to 12 to establish an alkaline master gel for *B. subtilis*. A remarkable fraction of the alkaline proteome is made up of ribosomal and highly hydrophobic proteins such as membrane proteins. From the alkaline 2D-gels 41 proteins with p*I* > 7 could be identified, most of which were ribosomal proteins. The goal to embed membrane proteins on alkaline 2D-gels failed. Their extreme hydrophobicity causes precipitation in IPG strips, which prevents their separation in the second dimension. Therefore, more attention has to be directed to membrane proteins since they seem to necessitate specific preparation and separation techniques. Additional 19 alkaline proteins were identified by Eymann et al. [33], who continued the analysis of the alkaline protein fraction on 2D-gels covering the pH range from 6 to 11.

At the time Ohlmeier et al. [37] started their advent of an alkaline mastergel there were basically no alternatives to 2D-PAGE for a separation of such complex protein samples. Only after the emergence of gel-free and direct measurement techniques it became possible to fractionate complex samples of proteins with extreme physico-chemical properties, because such techniques allow separation steps at peptide instead of protein level [18,19].

A gel-free approach, in which the cytosolic proteome fraction of growing *B. subtilis* cells was dissected via two-dimensional liquid chromatography, yielded a total of 814 identified proteins of which 140 possessed a p*I* beyond 7 [24]. Surprisingly 129 of the 814 proteins are characterized by a p*I* between 4 and 7 and have no transmembrane domains but still they were not detected via standard 2D-PAGE. The incapability to detect these proteins on the gel could be caused by their low abundance, as 10 of the 129 proteins are transcriptional regulators that are most likely difficult to visualize by current staining methods. There are also proteins smaller than 10 kDa or larger than 100 kDa which will probably bottom out the gel or will not migrate into the 2D-gel, respectively. Despite the high-resolution power of 2D-PAGE one cannot rule out that some proteins could not be identified from the gel due to protein comigration, which results in masking of small protein spots by higher abundant proteins.

Combining the results from 2D-PAGE analyses with those obtained by a peptide fractionation via 2D-LC followed by MS/MS (tandem mass spectrometry) acquisition, it was possible to identify 1014 cytosolic proteins from exponentially growing cells. Most of the proteins of the central carbohydrate metabolism (glycolysis, pentose phosphate shunt and citric acid cycle), of almost all amino acid synthesis pathways, of purine and pyrimidine metabolism, of fatty acid metabolism, and of the main cellular functions like replication, transcription, translation and cell wall synthesis have been detected in 2D-PAGE studies already (Figs. 1 and 3). However, gel-free approaches allowed the identification of additional proteins, as alkaline, hydrophobic and low-abundant ones (Fig. 2). Thus, the foundation of a comprehensive monitoring of *B. subtilis* cellular physiology has been layed (Fig. 3).

## 2.2. Cytosolic proteome map of non-growing cells

The continuous requirement for adaptation of bacteria to physical stress and starvation in natural ecosystems has forced the development of a very complex adaptational network of gene regulation. In contrast to growing cells, which are characterized by a stable "vegetative core proteome" the proteome of non-growing cells is additionally determined by a specific stress or starvation condition which is reflected as "proteome signature" [38–41]. Tam et al. [42] complemented the vegetative proteome map of *B. subtilis* by a proteome map in response to stress and starvation using the 2D-gel-based approach. For that purpose the synthesized proteins of cytoplasmic proteomes of [$^{35}$S]-methionine labeled *B. subtilis* cells were analyzed in response to heat, salt, hydrogen peroxide and paraquat stress as well as after ammonium, tryptophan (*B. subtilis* carries an auxotrophy for tryptophan), glucose and phosphate starvation. The 2D-gel images of all stress and starvation experiments were combined by an image fusion approach to a stress (Fig. 4a and b) and starvation (Fig. 5a and b) proteome map, respectively [43]. In those maps all marker proteins induced specifically by one stimulus or generally by multiple stimuli were labeled with a defined color code of 15 different colors indicating their stress/starvation respondent induction profile. Such 2D-proteome maps with their color codes according to protein subsets with similar regulation have been used to classify stress and starvation proteins into specific and general regulons [42]. For example, the heat shock signature is characterized by an induction of heat-specific HrcA-dependent chaperones and the general induction of the SigB, CtsR and Spx regulons [44]. The treatment with hydrogen peroxide and paraquat resulted in the induction of the oxidative stress-specific PerR regulon, the iron starvation Fur regulon and the Spx regulon as general indicator for protein damage by non-native disulfide bond formation [45–48]. In contrast, the SOS regulon induction is specific for DNA damage caused by peroxides, only [49]. The specific adaptative function of stress specific regulons (e.g. HrcA, PerR) is to accomplish resistance against the stressor by neutralizing it, by adaptation to the specific stressor or by repair of the damage it caused [41,50]. In contrast, the SigB-dependent general stress regulon encodes proteins with different functions conferring a multiple, non-specific and preventive stress resistance to non-growing *B. subtilis* cells in anticipation of future stress possibly encountered during long-term stationary growth stages [41,51].

The proteome map summarizing the protein patterns of ammonium, tryptophan, glucose and phosphate starvation provided the tool to define specifically induced starvation regulons and more generally induced transition phase regulons. Starvation-specific proteins are required for high-affinity uptake of the limiting substrate or for the utilization of alternative
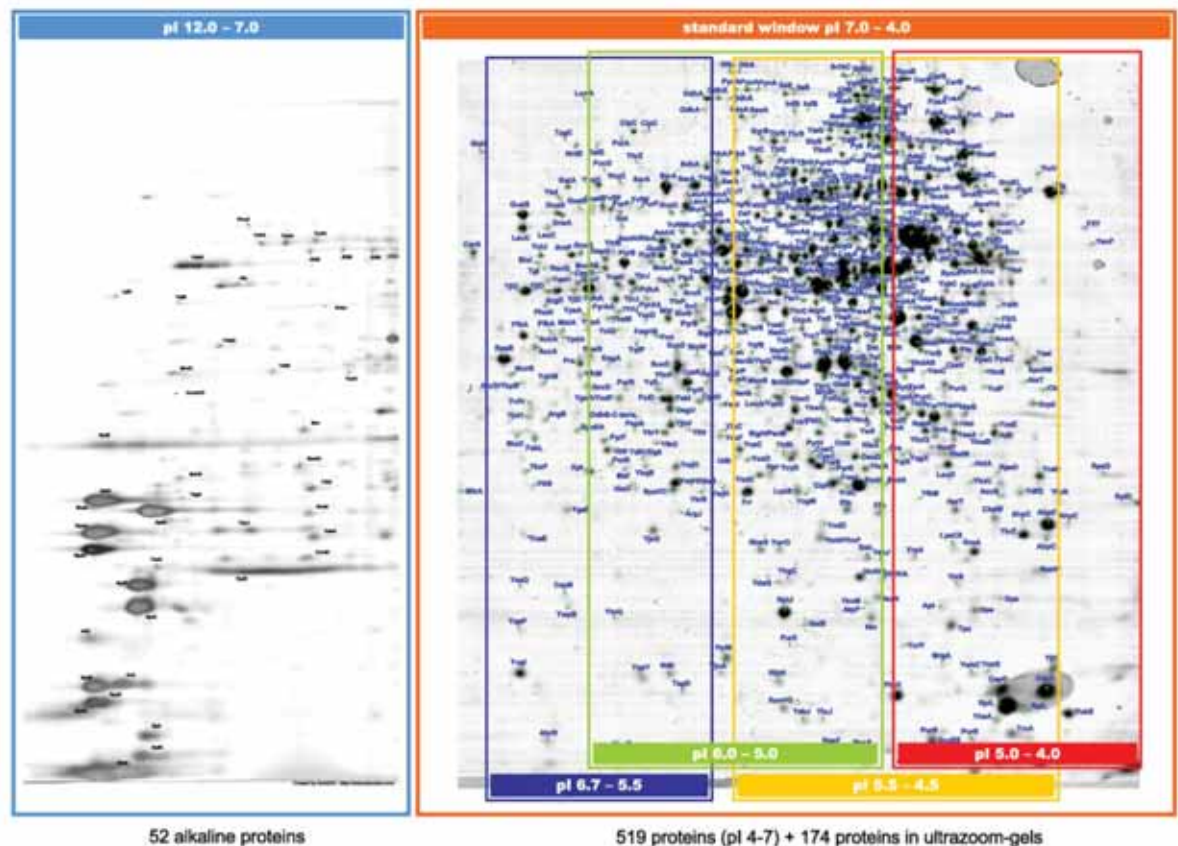
Fig. 1. The cytoplasmic vegetative proteome map of *B. subtilis* in the standard pH range (4–7), in the narrow pH ranges (4.0–5.0, 4.5–5.5, 5.0–6.0, 5.5–6.7) and in the alkaline pH range (7–12). Cytoplasmic protein extracts were separated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and stained with Colloidal Coomassie Blue [33]. Spot identification was performed using matrix assisted laser desorption ionization tandem time of flight (MALDI-TOF-TOF) mass spectrometry.

substrates when the preferred one is limited or exhausted. The induction of the TnrA regulon involved in uptake and utilization of alternative nitrogen sources and the SigL/BkdR regulon for degradation of branched chain amino acids indicates an ammonium starvation-specific proteome signature [52–55]. In contrast, an induction of the TRAP-regulated tryptophan biosynthesis enzymes is the specific proteome signature for tryptophan starvation [56,57]. In response to glucose starvation several carbon catabolite-controlled marker proteins are specifically activated in the absence of glucose and repressed in the presence of glucose by CcpA, CcpN and AcoR [58–62]. Finally, the induction of the PhoPR regulon is the specific proteome signature for phosphate starvation [63–66]. Starvation regulons such as TnrA, TRAP, CcpA or PhoPR were shown to be specifically involved in the uptake and utilization of alternative nutrient sources in response to a particular nutrient limitation. The proteome signature for different kinds of nutrient limitation is mediated by the CodY, SigB and SigH transition phase regulon. These general starvation regulons that are induced in response to carbon, ammonium or phosphate starvation are required for the adaptation of the cell to post-exponential stationary phase processes such as survival under non-growing conditions, competence
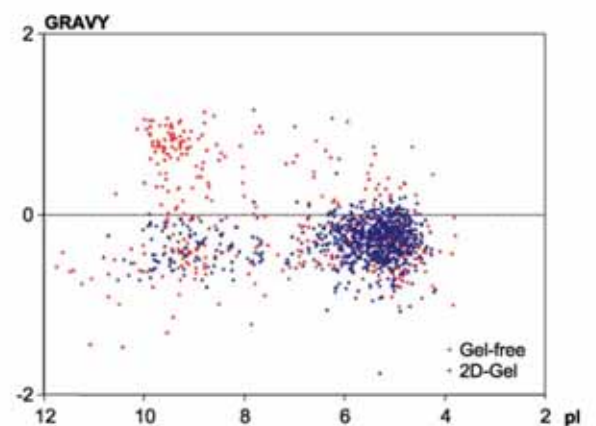


Fig. 2. Comparison of the subsets of proteins identified by a two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) approach and gel-free methods [24]. The plot GRAVY (grand average of hydropathy), which is a value for the hydrophobicity of proteins, versa isoelectric point clearly indicates the power of 2D-gels to separate proteins in an analytical window of pI 4–7. The gel-free approach could significantly increase the number of alkaline proteins and hydrophobic proteins which possess a more positive GRAVY value.

51

_____
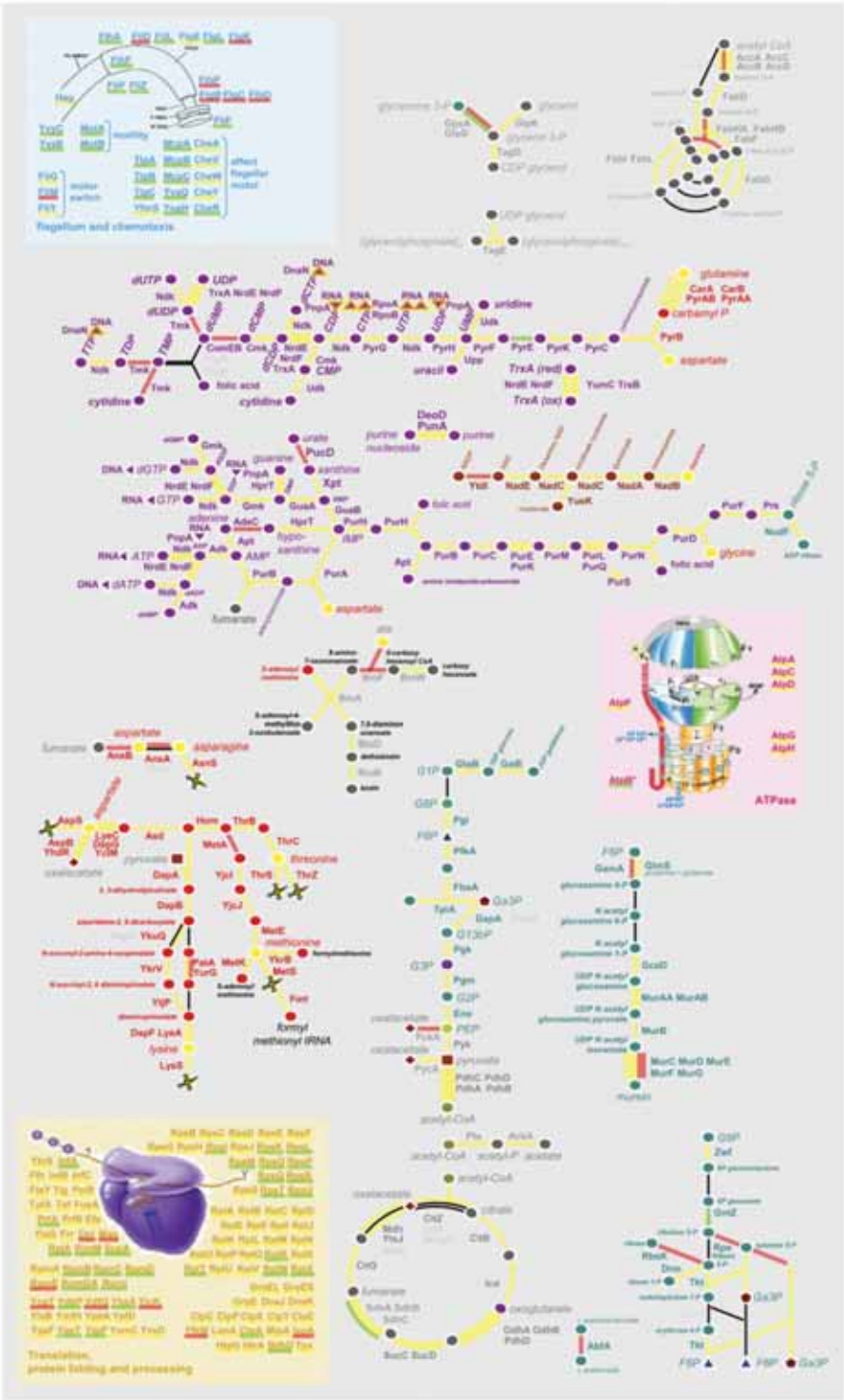
Fig. 3. Assignment of proteins identified to the different branches of cellular metabolism. Proteins that have not been identified in the 2D-gel images thus far are shaded light grey. Components of the carbohydrate metabolism such as glycolysis (center of the left page), pentose phosphate shunt (right side of the left page) and aminosugar synthesis for murein synthesis (right side of the left page) are indicated in dark cyan. Citric acid cycle (lower center of the left page), biotin (center of the left page) and fatty acid metabolism (upper right corner of the left page) are indicated by dark grey. Amino acid metabolism is colored red.
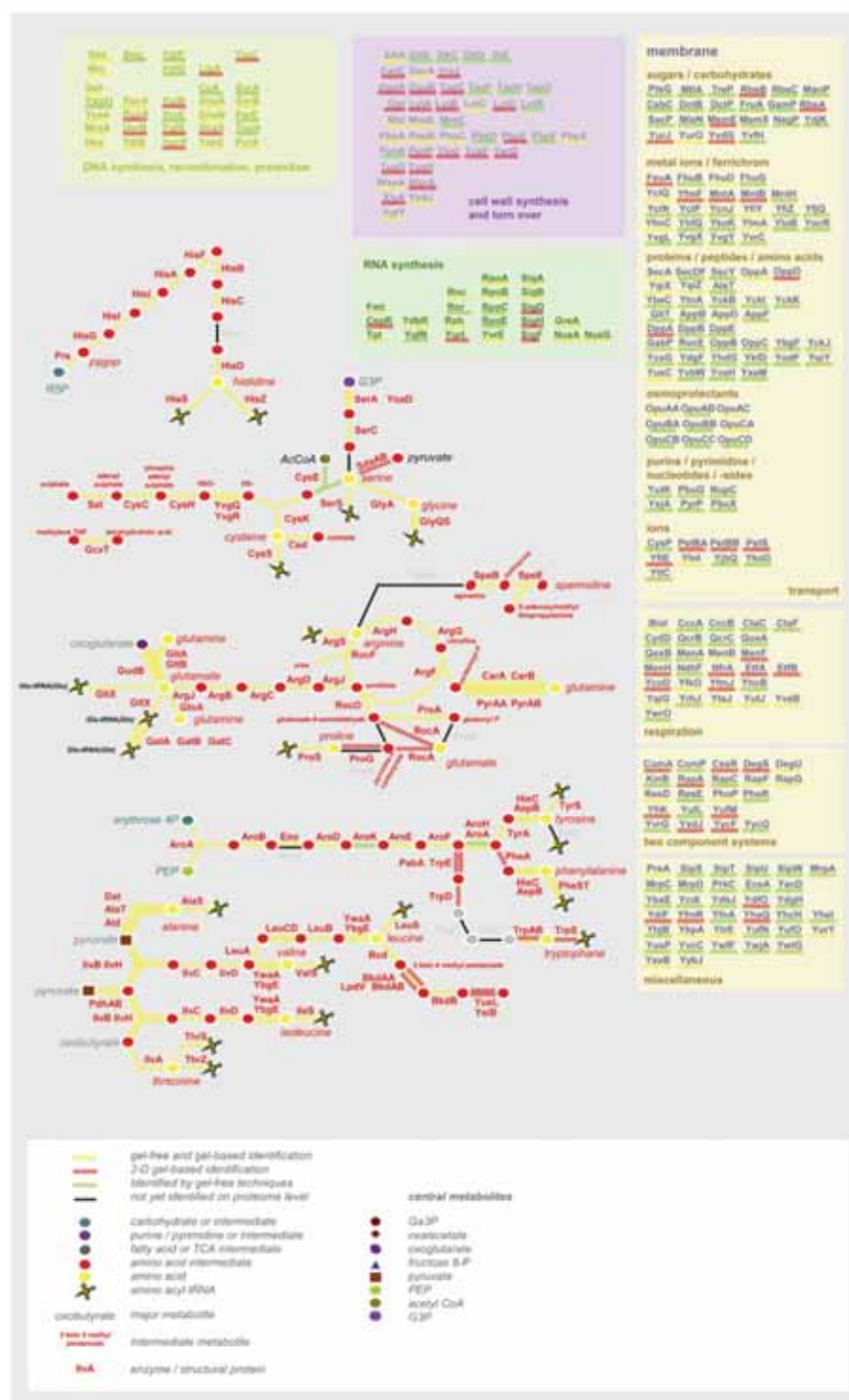
Fig. 3. (*Continued*) The essential amino acids are highlighted in yellow. Purin and pyrimidine (upper left page) metabolism are encoded in purple and the nicotinate metabolism (upper left page) in brown. Components not directly involved in metabolic pathways but in essential cell structures are presented in boxes. DNA-related functions in yellowish green (upper left corner right page), flagellum and chemotaxis-related components in azure (upper left corner left page) and ATPase components in pink (right side left page). Membrane proteins are highlighted cream-colored, components of the transcriptional machinery in green, and the ribosome and other components of the translational apparatus are encoded in ochre.

135
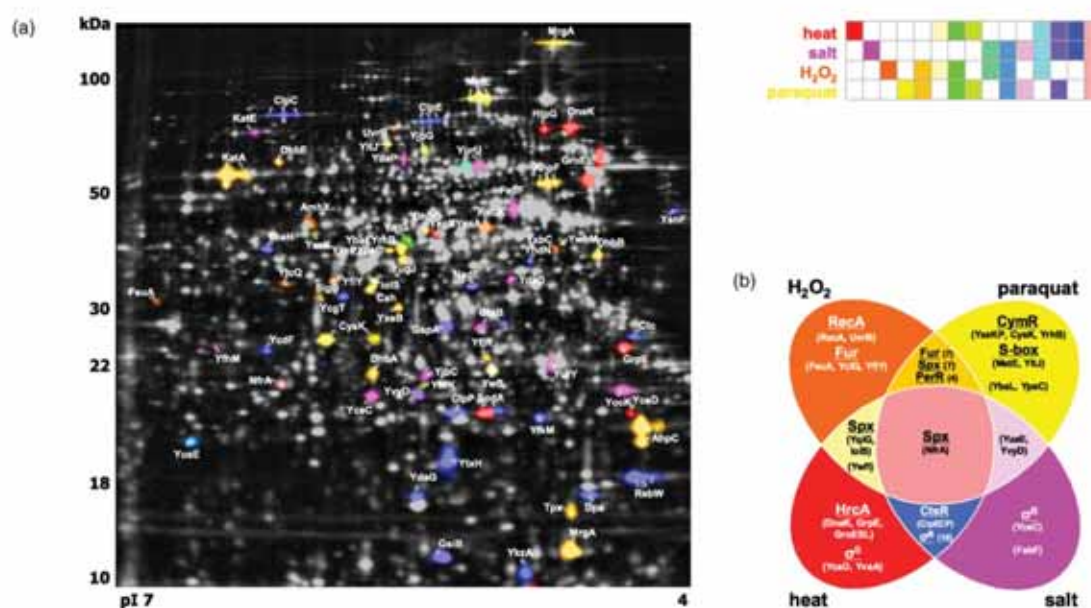


Fig. 4. (a) Cytoplasmic proteome map of *B. subtilis*. The realistically looking 2D-pattern is a composite image from five positionally corrected 2D-autoradiographs derived from exponentially grown as well as heat, salt, hydrogen peroxide and paraquat stressed *B. subtilis* cells. These autoradiographs showed the protein synthesis patterns from 5 min of cultivation incorporated [35S]-methionine. After applying the union fusion algorithm of the Delta2D software ([42], DECODON) to this stack of autoradiograms the resulting composite image summarizes the spot information from the whole gel series including all stress proteins ever induced. This made it possible to detect a spot consensus containing all stress proteins and valid for the whole gel set. This spot consensus was transferred and adapted to any original autoradiograph. The quantitation was done according to the consensus' spot shapes. For color coding, each spot set on any autoradiograph was divided into a non-changed and a more than three-fold compared to the control induced fraction. The latter fraction was considered as a group of stress proteins. This binarization of any spot set made it possible to highlight all induced proteins according to the shown color legend. The labels indicate which protein had been identified behind the spots. (b) The specific and general stress regulons in *B. subtilis*. Commonly shared (generally induced) and unique (specifically induced) stress regulons and proteins in the proteome of *B. subtilis* exposed to heat, salt, hydrogen peroxide and paraquat stress according to the fused stress proteome map in (a) [42]. The specific and general stress regulons are underlined and the encoded specific or general stress proteins are listed in parentheses.

or sporulation. The CodY regulon encodes proteins that allow broader adaptation to nutrient depletion including extracellular degradative enzymes, transporter proteins, catabolic enzymes, factors involved in genetic competence, antibiotic synthesis pathways, chemotaxis proteins and sporulation proteins [67]. In addition, the SigB regulon is the general proteome signature for glucose and phosphate starvation and required for the stationary phase survival upon starvation [41,61]. SigH directs the transcription of several genes that function in the transition from exponential growth to stationary phase, including the initiation of spore formation, genetic competence and general adaptation to nutrient depletion [68]. The color coded fused proteome map for starvation categorized a subset of six SigH-dependent proteins (YvyD, YtxH, YisK, YpiB, Spo0A, YuxI and the CodY-dependent RapA) as general starvation proteins indicating the transition to stationary phase caused by nutrient limitation [42]. Thus, the fused stress/starvation proteome maps and the color code approach provide important leads for future research on protein function of the novel identified general stress/starvation proteins during stationary phase survival in *B. subtilis*.

In total more than 200 stress or starvation induced proteins have been identified [42,45,58,64,69]. Eighty-three of them are absent in the vegetative proteome map, whereas the remaining stress and starvation induced proteins exhibit a basal expression level already in vegetative growing cells, indicating a basic physiological role in growing and non-growing cells.

## 3. The extracytoplasmic proteome of *B. subtilis* dissected by 2D-PAGE

As a Gram-positive bacterium *B. subtilis* lacks an outer membrane. It is therefore able to secrete large amounts of extracellular proteins directly into the environment. These proteins perform several very important functions, such as tapping of nutrients, environmental detoxification, cell-to-cell communication, combating competitors and, with regard to pathogenic bacteria, extracellular proteins play a critical role in virulence [70]. According to previous studies there are at least four distinct secretion pathways for approximately 300 proteins predicted to be located extracellularly [71,72]. Most proteins are secreted via the Sec pathway whereas the twin arginine translocation pathway, the pseudolipin export pathway and pathways using ATP-binding cassette transporters are employed for special purposes for only a few specific proteins [73]. The way a protein is translocated and its final destination are determined by the presence of particular signal peptides and retention signals. The first proteome analyses on secreted proteins in *B. subtilis* were performed in minimal medium with different carbon sources [74] as

54

Fig. 5. (a) Cytoplasmic proteome map of *B. subtilis* after ammonium, tryptophan, glucose and phosphate starvation. The protein synthesis pattern (autoradiograms) of *B. subtilis* in response to starvation for ammonium, tryptophan, glucose and phosphate were combined to generate a fused starvation proteome map of *B. subtilis* using the union image fusion approach of the Delta2D software [42]. The induced marker proteins were color coded according to their expression profiles. All spots induced specifically after one starvation condition or generally also by other starvation conditions more than 3-fold are labeled in the fused proteome map by a defined color code that indicates the induction profile. (b) The specific and general starvation regulons in *B. subtilis*. Commonly shared (generally induced) and unique (specifically induced) starvation regulons and proteins in the proteome of *B. subtilis* after ammonium, tryptophan, phosphate and glucose starvation according to the fused starvation proteome map in (a) [42]. The specific and general starvation regulons are underlined and the encoded specific or general starvation proteins are listed in parentheses.
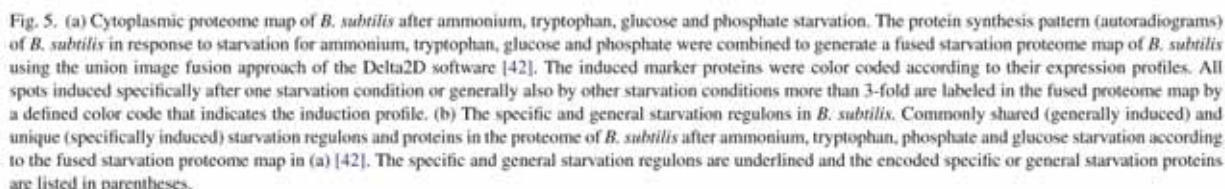
well as under phosphate starvation conditions [64]. However, the highest level of protein secretion in *B. subtilis* is observed when the cells have entered stationary growth phase in rich medium [75]. A 2D-PAGE approach was used to study the *B. subtilis* extracytoplasmic proteome which can be divided into three subproteomes: the true extracellular proteome, the lipoproteome and the cell wall associated proteome.

### 3.1. The extracellular proteome of B. subtilis

Most of the proteins that are completely transported across the cytoplasmic membrane are synthesized with an N-terminal signal peptide that is cleaved by a signal peptidase I (SPase I) after translocation which allows the protein to be released into the growth medium.

When *B. subtilis* is grown under conditions of phosphate starvation PhoPR-dependent proteins comprising extracellular phosphatases and phosphodiesterases (PhoA, PhoB, PhoD), the glycerophosphoryl phosphodiesterase GlpQ, the 5'nucleotidase YfkN, the ribonuclease YurI, the binding component of the high-affinity phosphate-specific transporter PstS and the lipoprotein YdhF are exceptionally strong induced [64,66,76,77]. These phosphate starvation-specific proteins remarkably account for 30% of the total phosphate starvation extracellular proteome (Fig. 6a). A master gel for the secreted proteome of *B. subtilis* was defined in Luria Broth (LB) medium during the stationary

phase (Fig. 6b). In total, 113 different proteins of the *B. subtilis* secretome could be identified. Fifty-four of them were predicted to be secreted because of an N-terminal signal peptide with a SPase I cleavage site and the lack of retention signals [71,73]. The remaining proteins are unpredicted secretory proteins which either possess signal peptides and retention signals (18 lipoproteins, 6 cell wall-binding proteins and 6 membrane proteins) or lack signal peptides (17 cytoplasmic proteins, 6 phage-related proteins and 7 Flagella-related proteins). The identified extracellular proteins are involved in the uptake, transport or utilization of carbohydrates, proteins, nucleotides, lipids and phosphate, in the cell wall metabolism, in environmental detoxification or they show flagella- and phage-related functions [73,75].

### 3.2. The "lipoproteome" of B. subtilis

The lack of an outer membrane has, most likely, forced Gram-positive bacteria to modify several extracellular proteins with lipids to adhere them as "lipoproteins" at the membrane surface, as the high-affinity substrate-binding proteins of ABC-transporters. Lipoproteins in *B. subtilis* are Sec-dependently secreted, lipid modified by the diacyl glyceryl transferase Lgt and cleaved by the type II signal peptidase LspA [78]. To describe the "lipoproteome" fraction of *B. subtilis*, a washed cell membrane fraction was prepared that was insoluble in the non-detergent sulfobetaine (NDSB) and then extracted with

_____

PhoPR-dependent phosphate starvation-induced proteins (underlined)
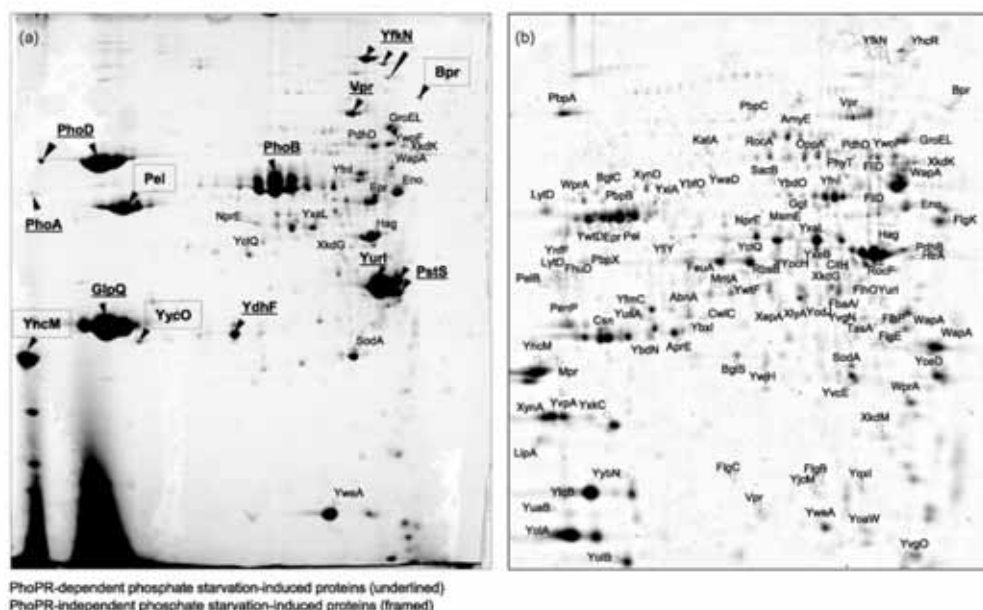PhoPR-independent phosphate starvation-induced proteins (framed)

Fig. 6. The extracellular proteome of *B. subtilis* 168 under conditions of phosphate starvation (a) and in complete medium (b). Cells of *B. subtilis* 168 were grown in minimal medium under the conditions of phosphate starvation (a) and in Luria Broth (b). Proteins in the growth medium were harvested 1 h after entry into the stationary phase. After precipitation with trichloric acid (TCA), the extracellular proteins were separated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and stained with Sypro Ruby [63,74].

the detergent *n*-dodecyl-β-D-maltoside as described previously [79]. A separation of this lipoprotein containing fraction by 2D-PAGE yielded 10 lipoproteins which have not been identified in the cytoplasmic proteome [76]. Interestingly, nine of these lipoproteins are also components of the extracellular proteome. They are most likely released into the growth medium by proteolytic shaving [80].

### 3.3. The cell wall proteome of B. subtilis

Several enzymes involved in cell wall turnover contain a variable number of repeat domains [71] which show an affinity for components of the cell wall [81]. These repeats are thought to facilitate the direction of enzymes to distinct sites in the cell wall. In Gram-positive organisms also a covalent anchoring of surface proteins to the cell wall is known [82]. It requires a C-terminal cell wall sorting signal that will be recognized by a specific transpeptidase, the sortase A, which covalently binds the protein to the cell wall [83]. The LiCl-extracted cell wall proteome of *B. subtilis* consists of 12 abundant non-covalently linked cell wall proteins including the WapA processing products (CWBP105 and CWBP62), the processing products of the major cell wall protease WprA (CWBP52 and CWBP23) and the autolysins LytC and LytB [84]. Interestingly, some cell wall proteins have been identified exclusively in the extracellular proteome (e.g. LytD, an autolysin), but not in the cell wall fraction. This indicates that the presence of cell wall binding repeats is not a guarantee for retention in the cell wall. Cell wall proteins might be released into the medium due to proteolytic shedding and cell wall turnover [73].

In total the comprehensive analyses of the *B. subtilis* extracytoplasmic proteome by a 2D-gel-based approach lead to the identification of 119 additional proteins which are either secreted, lipid anchored or associated to the cell wall.

### 4. The membrane proteome of vegetative cells

Membrane proteins constitute an important facet in physiological proteomics due to their importance in maintaining cell integrity, signal sensing, transport processes, energy conservation and their role in virulence regarding pathogenic bacteria. Despite their functional relevance the study of membrane proteins is notoriously difficult [85] and thus lagging behind analyses on other subproteomic fractions as cytosolic [86], cell wall-associated [76,84] and secreted [73,75,76] proteins. Similar to the extracellular proteins membrane proteins possess signal peptides marking them for cellular export. However, in most cases these signal peptides are not removed by a SPase and, consequently, they serve in the anchoring of proteins in the membrane. Furthermore, membrane proteins may contain additional membrane-spanning domains, which determine the topology of these proteins in the cytoplasmic membrane. According to the TMHMM algorithm to predict transmembrane domains [87] more than one quarter of all *B. subtilis* proteins is characterized by one or more membrane-spanning domains, emphasizing the need to focus on this barely characterized proteome fraction.

With a profiling of ABC-transporter solute-binding proteins, Bunai et al. initiated the work on the *B. subtilis* membrane protein fraction [79]. Dreisbach and coworkers established a technique
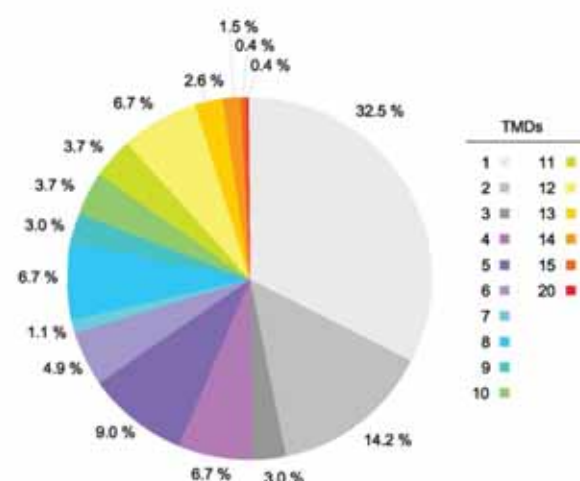
Fig. 7. Depiction of the number of transmembrane domains in the 268 membrane proteins identified by semi-gel-based proteomics.

in which the cytoplasmic membrane is washed in high-salt and alkaline buffers before the membrane proteins are solubilized with the detergent *n*-dodecyl-β-D-maltoside [33]. The membrane proteins are subsequently separated by one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (1D-SDS-PAGE), enzymatically digested and their peptides resolved via reverse phase chromatography followed by MS/MS analysis. With a result of 268 identified proteins which contain membrane-spanning domains this technique was successfully applied to start the dissection of the *B. subtilis* membrane proteome of exponentially growing cells [24,33]. Remarkably, 134 proteins possess four or more transmembrane domains (TMDs), which is consistent with their strong hydrophobicity and explains why this semi-gel-based approach is needed for their detection. Fig. 7 depicts the numbers of membrane-spanning domains in the 268 identified membrane proteins. Most of the detected membrane proteins are still of unknown function. The functionally defined proteins can be assigned to several essential categories, such as permeases and transporters, components of respiratory chains and ATP-synthase, redoxactive enzymes (e.g. oxidoreductases and dehydrogenases), two component systems, proteins involved in cell motility, cell division, autolysis, chemotaxis, osmoregulation, penicillin-binding proteins, parts of translocation systems and signal peptidases, as well as extracellularly acting proteases. Of the 268 membrane proteins, 16 are predicted lipoproteins according to Tjalsma et al. [71], 11 of those have not been detected in 2D-gel-based studies.

## 5. Conclusion

The genome of *B. subtilis* contains more than 4100 genes. In our proteome study 1395 proteins were identified so far. Most of them are vegetative proteins which are synthesized in growing cells to put on house-keeping functions (950 cytosolic proteins, 268 membrane proteins, 12 cell wall-bound proteins).

Some vegetative proteins are characterized by a basal expression level indicating basic cell functions even in growing cells, but they are strongly induced in response to stress stimuli, like chaperones as prominent members of this class. In response to nutrient starvation 155 proteins were determined to be induced and can be used as proteomic signatures for carbon, nitrogen or phosphate starvation. Finally, our work resulted in the detection of 113 extracellular and 21 lipid-anchored proteins, mostly expressed in non-growing cells. The total proteome information we obtained covers more than one-third of *B. subtilis* theoretical proteome, now ready for physiological application. Most proteins of the main metabolic pathways, predominantly located in the proteomic window of p*I* 4–7, were covered by our proteomic approaches. Many enzymes, already known for a long time, have been visualized for the first time. What can we do with this new kind of information? First of all we have the chance now to follow the regulation of entire metabolic pathways on a proteome-wide scale, as it was already done for the carbon core metabolism [42,58,60,61].

Nevertheless, there is a long way to go to the entire proteome of *B. subtilis* with more than 4100 predicted gene products. To reach this final goal we have two major approaches in mind, first of all an analytical approach: relying on gel-free techniques we were able to add many more proteins to the "gel-based proteome map", particularly to the membrane proteome fraction which still escapes detection by gel-based proteomics. This is, however, only a beginning. By a consequent extension of the application of gel-free techniques we are convinced to increase the proteome coverage in a substantial way already in the near future.

The second approach follows a physiological consideration: a substantial part of the genome is expressed only under definite physiological circumstances. Concurrent transcriptome and proteome studies revealed that even in cells that were exposed to various stress or starvation stimuli only a subfraction of the induced genes, in most cases between 30 and 60%, were covered by proteomics [33]. The still missing genes code for membrane proteins, low abundant proteins and many others.

Furthermore, we almost ignored the proteome of sporulating cells which may consist of more than 300 or even 400 proteins. And finally, one has to bear in mind that we exposed the cells to laboratory stress and starvation stimuli which can only simulate growth-restricting conditions in nature. The list of these "natural stimuli" is far off from being complete. We have only limited information on the proteome of cells grown under nitrate respiration or fermentation conditions, and we do not know much about the proteome of cells under cell wall stress, alkaline or acidic stress, to mention some of them. Furthermore, many stimuli cells typically have to face in their natural habitat were ignored. Cell growth in nature is characterized by a characteristic "multicellular behaviour" including the formation of microcolonies and biofilms, for example. The growth of *B. subtilis* adjacent to plant roots in the soil is probably very important regarding its natural life style. It is an attractive and challenging task for future studies to analyze the proteome of cells grown under environmentally related conditions in order to increase the coverage of the proteome on the one side and to gain new and essential information on its ecophysiology on the other.

## Acknowledgements

## References

[1] M.R. Wilkins, J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, K.L. Williams, Biotechnol. Genet. Eng. Rev. 13 (1996) 19.

[2] R.D. Fleischmann, M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, J.F. Tomb, B.A. Dougherty, J.M. Merrick, et al., Science 269 (1995) 496.

[3] A. Goffeau, R. Aert, M.L. Agostini-Carbone, A. Ahmed, M. Aigle, L. Alberghina, K. Albermann, M. Albers, M. Aldea, D. Alexandraki, et al., Nature 387 (1997) 5.

[4] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, et al., Science 291 (2001) 1304.

[5] P.H. O'Farrell, J. Biol. Chem. 250 (1975) 4007.

[6] J. Klose, Humangenetik 26 (1975) 231.

[7] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.

[8] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.

[9] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, C. Watanabe, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 5011.

[10] M. Mann, P. Hojrup, P. Roepstorff, Biol. Mass. Spectrom. 22 (1993) 338.

[11] D.J. Pappin, P. Hojrup, A.J. Bleasby, Curr. Biol. 3 (1993) 327.

[12] P. James, M. Quadroni, E. Carafoli, G. Gonnet, Biochem. Biophys. Res. Commun. 195 (1993) 58.

[13] J.R. Yates 3rd, S. Speicher, P.R. Griffin, T. Hunkapiller, Anal. Biochem. 214 (1993) 397.

[14] K. Heirwegh, P. Edman, Biochim. Biophys. Acta 24 (1957) 219.

[15] D.F. Hunt, R.A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A.L. Cox, E. Appella, V.H. Engelhard, Science 255 (1992) 1261.

[16] R.A. Henderson, H. Michel, K. Sakaguchi, J. Shabanowitz, E. Appella, D.F. Hunt, V.H. Engelhard, Science 255 (1992) 1264.

[17] D.F. Hunt, H. Michel, T.A. Dickinson, J. Shabanowitz, A.L. Cox, K. Sakaguchi, E. Appella, H.M. Grey, A. Sette, Science 256 (1992) 1817.

[18] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates 3rd, Nat. Biotechnol. 17 (1999) 676.

[19] D.A. Wolters, M.P. Washburn, J.R. Yates 3rd, Anal. Chem. 73 (2001) 5683.

[20] M.P. Washburn, D. Wolters, J.R. Yates 3rd, Nat. Biotechnol. 19 (2001) 242.

[21] M.S. Lipton, L. Pasa-Tolic, G.A. Anderson, D.J. Anderson, D.L. Auberry, J.R. Battista, M.J. Daly, J. Fredrickson, K.K. Hixson, H. Kostandarithes, C. Masselon, L.M. Markillie, R.J. Moore, M.F. Romine, Y. Shen, E. Stritmatter, N. Tolic, H.R. Udseth, A. Venkateswaran, K.K. Wong, R. Zhao, R.D. Smith, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 11049.

[22] J. Peng, J.E. Elias, C.C. Thoreen, L.J. Licklider, S.P. Gygi, J. Proteome Res. 2 (2003) 43.

[23] M. Taoka, Y. Yamauchi, T. Shinkawa, H. Kaji, W. Motohashi, H. Nakayama, N. Takahashi, T. Isobe, Mol. Cell Proteomics 3 (2004) 780.

[24] S. Wolff, A. Otto, D. Albrecht, J. Stahl Zeng, K. Büttner, M. Glückmann, M. Hecker, D. Becher, Mol. Cell Proteomics 21 (2006) 21.

[25] F. Kunst, N. Ogasawara, I. Moszer, A.M. Albertini, G. Alloni, V. Azevedo, M.G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, et al., Nature 390 (1997) 249.

[26] U.N. Streips, F.W. Polio, J. Bacteriol. 162 (1985) 434.

[27] A. Richter, M. Hecker, FEMS Microbiol. Lett. 36 (1986) 69.

[28] A.L. Sonenshein, J.A. Hoch, R. Losick (Eds.), *Bacillus subtilis* and its Closest Relatives, ASM Press, Washington, DC, USA, 2001, p. 4.

[29] H. Langen, B. Takacs, S. Evers, P. Berndt, H.W. Lahm, B. Wipf, C. Gray, M. Fountoulakis, Electrophoresis 21 (2000) 411.

[30] J.D. Jaffe, N. Stange-Thomann, C. Smith, D. DeCaprio, S. Fisher, J. Butler, S. Calvo, T. Elkins, M.G. FitzGerald, N. Hafez, C.D. Kodira, J. Major, S. Wang, J. Wilkinson, R. Nicol, C. Nusbaum, B. Birren, H.C. Berg, G.M. Church, Genome Res. 14 (2004) 1447.

[31] A. Tebbe, C. Klein, B. Bisle, F. Siedler, B. Scheffer, C. Garcia-Rizo, J. Wolfertz, V. Hickmann, F. Pfeiffer, D. Oesterhelt, Proteomics 5 (2005) 168.

[32] C. Klein, C. Garcia-Rizo, B. Bisle, B. Scheffer, H. Zischka, F. Pfeiffer, F. Siedler, D. Oesterhelt, Proteomics 5 (2005) 180.

[33] C. Eymann, A. Dreisbach, D. Albrecht, J. Bernhardt, D. Becher, S. Gentner, T. Tam le, K. Büttner, G. Buurman, C. Scharf, S. Venz, U. Völker, M. Hecker, Proteomics 4 (2004) 2849.

[34] K. Büttner, J. Bernhardt, C. Scharf, R. Schmid, U. Mäder, C. Eymann, H. Antelmann, A. Völker, U. Völker, M. Hecker, Electrophoresis 22 (2001) 2908.

[35] A. Görg, C. Obermaier, G. Boguth, A. Csordas, J.J. Diaz, J.J. Madjar, Electrophoresis 18 (1997) 328.

[36] A. Görg, G. Boguth, C. Obermaier, W. Weiss, Electrophoresis 19 (1998) 1516.

[37] S. Ohlmeier, C. Scharf, M. Hecker, Electrophoresis 21 (2000) 3701.

[38] J. Bernhardt, U. Völker, A. Völker, H. Antelmann, R. Schmid, H. Mach, M. Hecker, Microbiology 143 (1997) 999.

[39] H. Antelmann, J. Bernhardt, R. Schmid, H. Mach, U. Völker, M. Hecker, Electrophoresis 18 (1997) 1451.

[40] R.A. VanBogelen, E.E. Schiller, J.D. Thomas, F.C. Neidhardt, Electrophoresis 20 (1999) 2149.

[41] M. Hecker, U. Völker, Proteomics 4 (2004) 3727.

[42] L.T. Tam, H. Antelmann, C. Eymann, D. Albrecht, J. Bernhardt, M. Hecker, Proteomics 6 (2006) 4565.

[43] S. Luhn, M. Berth, M. Hecker, J. Bernhardt, Proteomics 3 (2003) 1117.

[44] M. Hecker, U. Völker, Adv. Microb. Physiol. 44 (2001) 35.

[45] J. Mostertz, C. Scharf, M. Hecker, G. Homuth, Microbiology 150 (2004) 497.

[46] J.D. Helmann, M.F. Wu, A. Gaballa, P.A. Kobel, M.M. Morshedi, P. Fawcett, C. Paddon, J. Bacteriol. 185 (2003) 243.

[47] J. Ollinger, K.B. Song, H. Antelmann, M. Hecker, J.D. Helmann, J. Bacteriol. 188 (2006) 3664.

[48] S. Nakano, E. Kuster-Schock, A.D. Grossman, P. Zuber, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 13603.

[49] S. Fernandez, S. Ayora, J.C. Alonso, Res. Microbiol. 151 (2000) 481.

[50] M. Hecker, U. Völker, Mol. Microbiol. 29 (1998) 1129.

[51] T.A. Gaidenko, C.W. Price, J. Bacteriol. 180 (1998) 3730.

[52] S.H. Fisher, Mol. Microbiol. 32 (1999) 223.

[53] L.V. Wray Jr., A.E. Ferson, K. Rohrer, S.H. Fisher, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 8841.

[54] K. Yoshida, H. Yamaguchi, M. Kinehara, Y.H. Ohki, Y. Nakaura, Y. Fujita, Mol. Microbiol. 49 (2003) 157.

[55] M. Debarbouille, R. Gardan, M. Arnaud, G. Rapoport, J. Bacteriol. 181 (1999) 2059.

[56] P. Babitzke, P. Gollnick, J. Bacteriol. 183 (2001) 5795.

[57] P. Gollnick, P. Babitzke, A. Antson, C. Yanofsky, Annu. Rev. Genet. 39 (2005) 47.

[58] J. Bernhardt, J. Weibezahn, C. Scharf, M. Hecker, Genome Res. 13 (2003) 224.

[59] H.M. Blencke, G. Homuth, H. Ludwig, U. Mäder, M. Hecker, J. Stülke, Metab. Eng. 5 (2003) 133.

[60] K. Yoshida, K. Kobayashi, Y. Miwa, C.M. Kang, M. Matsunaga, H. Yamaguchi, S. Tojo, M. Yamamoto, R. Nishi, N. Ogasawara, T. Nakayama, Y. Fujita, Nucleic Acids Res. 29 (2001) 683.

[61] T. Koburger, J. Weibezahn, J. Bernhardt, G. Homuth, M. Hecker, Mol. Genet. Genomics 274 (2005) 1.

[62] P. Servant, D. Le Coq, S. Aymerich, Mol. Microbiol. 55 (2005) 1435.

[63] C. Eymann, H. Mach, C.R. Harwood, M. Hecker, Microbiology 142 (1996) 3163.

[64] H. Antelmann, C. Scharf, M. Hecker, J. Bacteriol. 182 (2000) 4478.

[65] F.M. Hulett, Mol. Microbiol. 19 (1996) 933.

[66] N.E. Allenby, N. O'Connor, Z. Pragai, A.C. Ward, A. Wipat, C.R. Harwood, J. Bacteriol. 187 (2005) 8063.

[67] V. Molle, Y. Nakaura, R.P. Shivers, H. Yamaguchi, R. Losick, Y. Fujita, A.L. Sonenshein, J. Bacteriol. 185 (2003) 1911.

[68] R.A. Britton, P. Eichenberger, J.E. Gonzalez-Pastor, P. Fawcett, R. Monson, R. Losick, A.D. Grossman, J. Bacteriol. 184 (2002) 4881.

[69] D. Höper, J. Bernhardt, M. Hecker, Proteomics 6 (2006) 1550.

[70] P.R. Jungblut, U.E. Schaible, H.J. Mollenkopf, U. Zimny-Arndt, B. Raupach, J. Mattow, P. Halada, S. Lamer, K. Hagens, S.H. Kaufmann, Mol. Microbiol. 33 (1999) 1103.

[71] H. Tjalsma, A. Bolhuis, J.D. Jongbloed, S. Bron, J.M. van Dijl, Microbiol. Mol. Biol. Rev. 64 (2000) 515.

[72] J.M. van Dijl, P.G. Braun, C. Robinson, W.J. Quax, H. Antelmann, M. Hecker, J. Müller, H. Tjalsma, S. Bron, J.D. Jongbloed, J. Biotechnol. 98 (2002) 243.

[73] H. Tjalsma, H. Antelmann, J.D. Jongbloed, P.G. Braun, E. Darmon, R. Dorenbos, J.Y. Dubois, H. Westers, G. Zanen, W.J. Quax, O.P. Kuipers, S. Bron, M. Hecker, J.M. van Dijl, Microbiol. Mol. Biol. Rev. 68 (2004) 207.

[74] I. Hirose, K. Sano, I. Shioda, M. Kumano, K. Nakamura, K. Yamane, Microbiology 146 (2000) 65.

[75] H. Antelmann, H. Tjalsma, B. Voigt, S. Ohlmeier, S. Bron, J.M. van Dijl, M. Hecker, Genome Res. 11 (2001) 1484.

[76] I. Humphery-Smith, M. Hecker (Eds.), Microbial Proteomics, Wiley-VCH, Weinheim, 2006, pp. 179–208.

[77] A.L. Sonenshein, J.A. Hoch, R. Losick (Eds.), *Bacillus subtilis* and its Closest Relatives, ASM Press, Washington, DC, USA, 2001, pp. 193–201.

[78] S. Leskela, E. Wahlstrom, V.P. Kontinen, M. Sarvas, Mol. Microbiol. 31 (1999) 1075.

[79] K. Bunai, M. Ariga, T. Inoue, M. Nozaki, S. Ogane, H. Kakeshita, T. Nemoto, H. Nakanishi, K. Yamane, Electrophoresis 25 (2004) 141.

[80] H. Tjalsma, J.M. van Dijl, Proteomics 5 (2005) 4472.

[81] J.M. Ghuysen, J. Lamotte-Brasseur, B. Joris, G.D. Shockman, FEBS Lett. 342 (1994) 23.

[82] O. Schneewind, D. Mihaylova-Petkov, P. Model, Embo J. 12 (1993) 4803.

[83] W.W. Navarre, O. Schneewind, Mol. Microbiol. 14 (1994) 115.

[84] H. Antelmann, H. Yamamoto, J. Sekiguchi, M. Hecker, Proteomics 2 (2002) 591.

[85] V. Santoni, M. Molloy, T. Rabilloud, Electrophoresis 21 (2000) 1054.

[86] R. Schmid, J. Bernhardt, H. Antelmann, A. Völker, H. Mach, U. Völker, M. Hecker, Microbiology 143 (1997) 991.

[87] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, J. Mol. Biol. 305 (2001) 567.

59

_____

*Monitoring of changes in the membrane proteome during*

*stationary phase adaption of Bacillus subtilis using in vivo*

*labeling techniques*

Monitoring of changes in the membrane proteome during stationary phase adaption of *Bacillus subtilis* using *in vivo* labeling techniques

_____

RESEARCH ARTICLE

# Monitoring of changes in the membrane proteome during stationary phase adaptation of *Bacillus subtilis* using *in vivo* labeling techniques

Annette Dreisbach[1]*, Andreas Otto[2]*, Dörte Becher[2], Elke Hammer[1], Alexander Teumer[1], Joost W. Gouw[3], Michael Hecker[2] and Uwe Völker[1]

[1] Interfakultäres Institut für Genetik und Funktionelle Genomforschung, Ernst-Moritz-Arndt-Universität, Greifswald, Germany
[2] Institut für Mikrobiologie, Ernst-Moritz-Arndt-Universität, Greifswald, Germany
[3] Department of Biomolecular Mass Spectrometry, Utrecht University, The Netherlands

*Bacillus subtilis* has been developed as a model system for physiological proteomics. However, thus far these studies have mainly been limited to cytoplasmic, extracellular, and cell-wall attached proteins. Although being certainly important for cell physiology, the membrane protein fraction has not been studied in comparable depth due to inaccessibility by traditional 2-DE-based workflows and limitations in reliable quantification. In this study, we now compare the potential of stable isotope labeling with amino acids (SILAC) and $^{14}$N/$^{15}$N-labeling for the analysis of bacterial membrane fractions in physiology-driven proteomic studies. Using adaptation of *B. subtilis* to amino acid (lysine) and glucose starvation as proof of principle scenarios, we show that both approaches provide similarly valuable data for the quantification of bacterial membrane proteins. Even if labeling with stable amino acids allows a more straightforward analysis of data, the $^{14}$N/$^{15}$N-labeling has some advantages in general such as labeling of all amino acids and thereby increasing the number of peptides for quantification. Both, SILAC as well as $^{14}$N/$^{15}$N-labeling are compatible with 2-DE, 2-D LC-MS/MS, and GeLC-MS/MS and thus will allow comprehensive simultaneous interrogation of cytoplasmic and enriched membrane proteomes.

## 1 Introduction

The Gram-positive rod *Bacillus subtilis* constitutes a well-established model system for physiological proteomics [1]. Combining the well-established genetic tools and the extensive biochemical and physiological knowledge with the potential of genome-wide proteomic approaches new insights into stress adaptation reactions, basic physiology as well as protein secretion and targeting have been gained [2–5]. Such physiological studies essentially depend on comparative analysis of extended sample series and the implementation of labeling technologies which allow quantitative comparisons of the data generated in time course experiments or in comparisons of wild-type bacteria and their isogenic mutant strains. Most of these studies have been performed applying traditional 2-DE-based proteomics allowing easy comparisons of multiple samples. Unfortunately, this

**Correspondence:** Professor Uwe Völker, Interfakultäres Institut für Genetik und Funktionelle Genomforschung, Ernst-Moritz-Arndt-Universität, Friedrich-Ludwig-Jahn-Str. 15A, 17487 Greifswald, Germany
**E-mail:** voelker@uni-greifswald.de
**Fax:** +49-3834-8680012

**Abbreviations: MSD,** membrane spanning domain; **SILAC,** stable isotope labeling with amino acids

* Both authors contributed equally to this work.

InterScience®

www.proteomics-journal.com

approach is confined to the analysis of the cytoplasmic, cell wall-bound and secreted proteins and excludes membrane proteins which escape detection by traditional 2-DE due to their extreme hydrophobicity [6, 7]. These limitations of 2-DE have meanwhile been overcome by application of MS-centered proteomics approaches that do not rely on 2-DE for separation of complex protein mixtures. These methods offer even greater sensitivity and are generally based on separation of complex tryptic-digested samples by reverse phase chromatography in combination with mass spectrometric analysis (LC-MS/MS) [8]. To improve coverage of the proteome one can prefractionate the sample by a number of methods like 1-D SDS-PAGE or multidimensional LC [9, 10]. These alternative strategies have also been used to address the membrane proteome of *B. subtilis*, but have mainly been limited to proof of principle studies rather than specific physiological issues [11–14] mainly due to the fact that quantification of membrane proteins in different samples has not been rigorously performed. In principle, a number of labeling strategies are available, which all rely on the introduction of a certain mass shift by labeling of samples with isotopologic compounds and a subsequent combination of differentially labeled samples. After mass spectrometric analysis the quantitative data can then be obtained from the integrated peak areas. Labeling can either be performed at the peptide or protein level and the latter even *in vivo*. The *in vivo* approach would clearly be the preferred choice for the quantitative analysis of membrane proteins because it would allow mixing of samples before the membrane preparation and thus dramatically reduce the bias introduced during the different preparation steps. Two different methods for the *in vivo* labeling of proteins the stable isotopic labeling of amino acids (SILAC) in cell culture [15] and the $^{14}N/^{15}N$-metabolic labeling [16, 17] have already been published. Whereas the SILAC technique is based on the integration of isotopologous amino acids into proteins, the $^{14}N/^{15}N$-metabolic labeling relies on the incorporation of different N-isotopes into all amino acids and subsequently into proteins. While the latter can be applied to all bacterial strains growing in N-containing minimal medium the SILAC approach requires an auxotrophy for at least one abundant amino acid.

In the study presented here, we applied both labeling methods to evaluate the changes of the membrane proteome of *B. subtilis* that occur during the transition of growing cells to the stationary phase. Both methods are compared regarding the amount and the quality of the quantitative data gained for the membrane protein fraction.

# 2 Materials and methods

## 2.1 Bacterial strains and culture

For SILAC experiments the *B. subtilis* wild-type strain IS58 (*trpC2 lys2*) [18] was cultured over night in Spizizen minimal medium [19] supplemented with 0.86 mM tryptophan (Sigma–Aldrich, Taufkirchen, Germany) and 0.86 mM lysine (Sigma–Aldrich). This culture was used to inoculate a pre-culture in Spizizen minimal medium with reduced levels of lysine (0.2 mM) to $OD_{540}$ of 0.05 with lysine either as the light $^{12}C_6$-$^{14}N_2$-lysine or its heavy isotopic counterpart $^{13}C_6$-$^{15}N_2$-lysine (98 at.% $^{13}C$ and $^{15}N$, Sigma–Aldrich). After reaching $OD_{540}$ of 0.5 this culture was rediluted into the same prewarmed medium to $OD_{540}$ of 0.1. Exponentially growing cells which have been grown with light or heavy lysine, respectively, for at least six generations were harvested by centrifugation (16 900 × g, 10 min, 4°C) at $OD_{540}$ of 1 and stationary phase cells were harvested 2 h after transition into stationary phase. Finally, aliquots (equivalent OD-units) of cultures grown with light or heavy lysine were combined and stored at −80°C.

For metabolic labeling experiments using $^{14}N/^{15}N$ the *B. subtilis* wild-type strain 168 (*trpC2*) was grown aerobically for at least three passages (minimum 15 generations) at 37°C in a synthetic minimal medium [20] supplemented with either $^{15}N$-ammonium sulfate/$^{15}N$-L-tryptophan (0.078 mM, 98 at. % excess, Cambridge Isotope Laboratories, Andover, USA) or $^{14}N$-ammonium sulfate and $^{14}N$-L-tryptophan. Furthermore, glucose was added to a final concentration of 0.05% to induce a glucose starvation after growth to an $OD_{500}$ of 1.0. Cells were harvested either in exponential phase or 2 h after transition into stationary phase and equivalent OD-units of bacteria grown in the presence of $^{15}N$ or $^{14}N$ were combined and stored at −80°C.

## 2.2 *In silico* analysis of the theoretical proteome of *B. subtilis*

The theoretical proteome of the *B. subtilis* strain 168 comprises 4106 different proteins. To obtain a reliable prediction of *B. subtilis* membrane proteins four different prediction algorithms were used for an *in silico* analysis: TMHMMv.2/ ALOM2 [21–23], TMHMMv.2 [23], HMMTOP2 [24], and SOSUI [25]. A protein was considered as integral membrane protein only when at least two algorithms predicted membrane spanning domains (MSDs). The number of MSDs which is given in this work is the mean value of the four prediction models. Furthermore, secreted proteins and lipoproteins [26] as well as proteins which are retained in the membrane despite secretion or lipoprotein signal peptides [27] were considered for a reliable membrane protein prediction.

## 2.3 Sample preparation

Protoplasts were prepared according to Chang and Cohen [28] followed by a preparation of membrane proteins as described previously [11]. Briefly, the membrane proteins were enriched by differential centrifugation after cell disruption. Further purification of the membrane fraction was achieved by washing with high-salt and $Na_2CO_3$-buffer. Finally, the membrane proteins were solubilized using a 15 mM solution of the nonionic detergent n-dodecyl-β-D-

Monitoring of changes in the membrane proteome during stationary phase adaption of
*Bacillus subtilis* using *in vivo* labeling techniques

_____

maltoside (Glycon Biochemicals, Luckenwalde, Germany).
Membrane proteins (50 μg) were separated using SDS-
PAGE (10%T, 7 cm length) and stained with a colloidal Coo-
massie staining procedure (GE Healthcare, Munich, Ger-
many). Fifteen protein bands were excised from the gel and
destained twice using 200 μL of 200 mM $NH_4HCO_3$/50%
ACN for 30 min at 37°C and dried in a vacuum centrifuge.
The gel pieces were soaked with a trypsin (Promega, Man-
nheim, Germany) containing solution (10 ng/μL of 20 mM
$NH_4HCO_3$) and incubated over night at 37°C. For peptide
extraction the gel pieces were covered with 0.1% acetic acid
and subsequently 0.05% acetic acid in 50% ACN and incu-
bated in an ultrasonic water bath for 15 min. The eluates of
both extractions were combined and concentrated in a
vacuum centrifuge.

## 2.4 Mass spectrometric analysis

### 2.4.1 MS/MS analysis

Separation of the complex peptide solutions was achieved by
RP chromatography (PepMap, 75 μm id × 250 mm, LC
Packings) operated on an MDLC nano-HPLC (GE Health-
care) with a nonlinear 180 min gradient from 9 to 54% ACN
in 0.05% acetic acid at a constant flow rate of 250 nL/min.
MS data were acquired with an LTQ-FT mass spectrometer
(Thermo Electron, Bremen, Germany) equipped with a
nanoelectrospray ion source. After a first survey scan in the
FTICR ($r = 25\,000$) high accuracy precursor masses were
determined by a single ion mode scan ($r = 50\,000$) and $MS^2$
data were recorded in the linear IT at a collision induced
energy of 35% [29].

### 2.4.2 Data analysis

Proteins were identified *via* an automated database search
using the SEQUEST algorithm (Bioworks v.3.2, Thermo
Electron) with a *B. subtilis* database extracted from SubtiList
(genolist.pasteur.fr/SubtiList/). In addition to the SubtiList
annotations, Swiss-Prot accessions (release 54.6) are also
provided in the result files to include annotation information
that has accumulated since the last update of the SubtiList
database.

Initial mass tolerances for protein identification by MS
were set to 0.01 Da. Up to two missed tryptic cleavages were
considered and methionine oxidation was set as variable
modification. Peptide identifications by MS/MS were con-
sidered significant at the following parameters: charge de-
pendent Xcorrelation score for one-fold charged peptides
>1.9; for two-fold charged peptides>2.2; three-fold charged
peptides >3.75; RSp value = 4; ion percent >70%, amino
acid sequence length: 7–30.

A multiconsensus search was performed using both a
forward and a reversed database to estimate the false-positive
rate ($2n_{reverse}/[n_{reverse} + n_{forward}]$).

### 2.4.3 Extraction of the quantitative data of the SILAC-approach

The incorporation of either light or heavy lysine during
growth introduced a mass shift of 8.015 *per* lysine. The
XPRESS tool, which is implemented in the Bioworks-
Browser v.3.2 was used to detect light peptides and their iso-
topologues, integration of peak areas and calculation of
ratios. For the quantitative analysis the mass tolerance was
set to 0.5 Da and the intensity minimum threshold adjusted
to $1 \times 10^6$ within a retention time window of 60 s before and
after MS/MS-acquisition. The data were manually validated
to ensure that only perfectly matching isotopologue peak
pairs were used for quantification.

### 2.4.4 Extraction of the quantitative data from the $^{14}N/^{15}N$-metabolic labeling approach

Database searches were performed using the MASCOT
search engine (Matrix Science, London, UK) against the
above-mentioned *B. subtilis* database. For the quantification
of datasets the search parameters were set to an MS accuracy
of 10 ppm and an MS/MS accuracy of 0.2 Da with a max-
imum of one missed cleavage site of trypsin. Modifications
such as the oxidation of methionine and phosphorylation of
serine, threonine, and tyrosine were also taken into account
during the search. Peptides with an individual MASCOT
score below 30 were filtered out. The datasets retrieved by the
MASCOT search engine independently for the heavy and the
light peptides were processed with MSQuant [30]. The ver-
sion used in this work is capable of relative quantification
based on $^{14}N/^{15}N$ datasets generated with the MASCOT
search engine.

The quantification algorithm of MSQuant parses the
spectra of the assigned peptides based on the peptide sum-
mary report of MASCOT in the Finnigan-raw files and cal-
culates the masses of the corresponding isotopologues. For
the determination of the peak area of the $^{14}N$- and $^{15}N$-con-
taining peptides a retention time window of 90 s before and
after the acquisition of the MS/MS spectrum was selected.
The ratios of the peak areas were calculated and finally
manually confirmed for consistency and used for further
analysis.

### 2.5 Evaluation of the quantitative data

Peptide ratios of specific proteins, which lay outside the 1.5-
fold interquartile distance below and above the 25 and 75%
quartile respectively, were defined as outliers and thus ex-
cluded from the analysis. From the remaining peptide ratios
the mean value was calculated. The filtering and calculation
processes were carried out using an in-house perl-script. The
relative difference of the ratios from the two technical repli-
cates was calculated by dividing the difference of the ratios by
the mean. A protein was termed differentially expressed when
the results of two technical replicates differed less than 20%

Monitoring of changes in the membrane proteome during stationary phase adaption of
*Bacillus subtilis* using *in vivo* labeling techniques

_____

and the change in protein amount was at least 1.5-fold. Due to the higher variability of the measurement of extreme ratios, proteins were considered differentially expressed, when the quantitative changes were larger than three-fold in each of the technical replicates analyzed even if the independent samples varied more than 20%. Protein ratios, which were determined in adjacent gel slices, were combined and the same criteria that were described above were applied.

## 3    Results and discussion

### 3.1 Identification of membrane proteins

Qualitative and quantitative analysis of complex proteomes independent of 2-DE separation has meanwhile become a routine procedure for soluble protein fractions of microbial cells. However, quantitative analyses of the cytoplasmic membrane fraction are still limited probably due to the fact that quantification has to closely interface with and avoid any bias of the multistep-purification of membrane fractions. However, a genome-wide comprehensive monitoring of bacterial membrane proteins is highly desirable for a better understanding of bacterial physiology because this cellular subfraction is important for signaling, energy generation, supply of nutrients, disposal of waste, and last but not least for maintaining cell integrity.

In this study, we used an optimized membrane purification protocol in combination with *in vivo* stable isotopic labeling techniques to quantitatively monitor changes in the membrane proteome of *B. subtilis* during transition from exponential growth to starvation induced stationary phase. In particular, we analyzed the adaptation of *B. subtilis* 168 to glucose limitation and the response of *B. subtilis* IS58 to amino acid limitation induced stationary phase.

An *in silico* analysis of the theoretical proteome of *B. subtilis* 168 revealed 1142 integral membrane proteins (Supporting Information Table S1). In the study presented here, in total 666 unique proteins were identified, of which 327 were integral membrane proteins (Supporting Information Tables S2–S4). The high proportion of proteins with at least four MSDs (44%) documents that highly hydrophobic membrane protein species are also covered by the method (Fig. 1). The studies of the *B. subtilis* membrane proteome, which have been published until now identified 112 [11], 158 [13], and 232 [12] integral membrane proteins, respectively. Thus, the integral membrane proteins identified in this study, which represent 25% of the predicted ones, constitute the most extensive collection of the membrane fraction of *B. subtilis* presented so far. The number of detected integral membrane proteins is in the same range as investigations of other bacteria like *Halobacterium salinarum* (101 integral membrane proteins detected; 20% of predicted) [31], *Pseudomonas aeruginosa* (333 integral membrane proteins detected; 32% of predicted) [32], and *Corynebacterium glutamicum* (326 integral membrane proteins detected; 50% of predicted) [33].
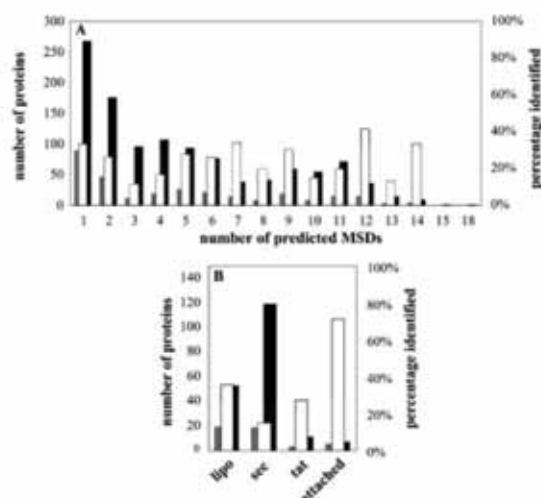


**Figure 1.** Predicted and observed membrane proteins of *B. subtilis*. The numbers of theoretically predicted (black columns) and identified proteins (gray columns) and the percentage of predicted proteins that were actually observed (white columns) are presented. (A) Integral membrane proteins were presented in relation to the number of predicted MSDs. (B) Summary for lipoproteins, proteins secreted *via* Sec- or Tat-systems and membrane attached proteins.

Our membrane protein preparation was also covering membrane-associated proteins in addition to integral membrane proteins and thus, besides the transmembrane domains also the grand average of hydrophobicity (GRAVY) [34] was calculated for the observed proteins, and protein sequences were searched for lipid anchors and secretion signal peptides [26]. A secretion signal peptide was predicted for 21 (18 Sec-system, 3 Tat-system) and a lipid anchor for 19 of the proteins detected by MS (Fig. 1B). For an extra five proteins without MSDs retainment in the membrane was already predicted by Tjalsma and van Dijl (Fig. 1B) [35]. Furthermore, 64 of the identified proteins were classified as membrane associated based on their function, which indicates that they either directly associate with the membrane or with membrane proteins by hydrophobic or protein–protein interaction. In summary, we were able to assign 400 of the 666 identified proteins (60%) to the membrane fraction either due to their biochemical properties (336) or function (64). The remaining primarily cytoplasmic proteins are either high abundance proteins that were still copurifying or they are constituents of high molecular weight protein complexes that copurified with the membrane fraction.

One hundred twenty-four of all identified proteins exhibited a GRAVY value above 0.25 and thus would not have been accessible to the traditional 2-DE analysis [6]. A functional annotation of the identified membrane proteins is complicated by the fact that 45% of them have not been assigned a biochemical function yet. Grouping of the mem-

brane proteins into functional groups according to the SubtiList database revealed that important categories such as membrane bioenergetics (17 of 39 proteins), signal transduction/sensors (10 of 32 proteins), membrane transporters/binding proteins (100 of 312 proteins), protein secretion apparatus (9 of 18 proteins), and motility/chemotaxis (14 of 20 proteins) are well covered (Fig. 2).

Of the 666 proteins which were identified in total in the membrane fractions 415 proteins were shared by the analyses of the adaptation to glucose and amino acid limitation of *B. subtilis* strains 168 and IS58, respectively, which emphasizes the reproducibility and the robustness of the membrane protein enrichment procedure. In the future, the coverage of the membrane proteome can very likely still be increased for *B. subtilis* if membrane fractions from cells grown under various conditions are included in the analysis or coverage of MSDs themselves is improved by utilization of organic solvents as introduced by Fischer *et al.* [33].

### 3.2 Quantification of membrane proteins using SILAC

A basic requirement for the SILAC technique is the auxotrophy of the bacterium for at least one abundant amino acid. For this reason, the established *B. subtilis* strain IS58, which displays auxotrophies for tryptophan and lysine, was chosen for this experiment. Because of its higher abundance in the theoretical proteome of *B. subtilis* lysine (7% of total amino acids) was better suited for the labeling procedure than tryptophan (1% of total amino acids). When *B. subtilis* cells of the strain IS58 are cultured in minimal medium containing either "light" ($^{12}C,^{14}N$-) or "heavy" ($^{13}C,^{15}N$-) lysine the different isotopes are incorporated into the proteins allowing mixing of cells of different physiological states, *e.g.*, exponentially growing *versus* starving cells, directly after harvest. Then, membrane proteins can be purified and analyzed simultaneously to reduce technical errors and quantitative information is obtained by the integration of the peak areas from the MS-spectrum. This approach will greatly reduce artifacts during processing. However, varying recovery effi-

ciencies of membrane proteins from cells of different physiological states (*e.g.*, growing *vs.* starving cells) can still not be compensated for.

As a prerequisite for proving the suitability of the SILAC approach it had to be demonstrated that complete replacement of light lysines by their heavy counterparts was achieved in the experimental setup. For this purpose cells were cultivated in the presence of $^{13}C,^{15}N$-lysine as described above and a gel slice of an SDS-PAGE-separated crude protein extract was subjected to tryptic digestion. The resulting peptide mixture was analyzed by MS. When $^{13}C,^{15}N$-lysine was chosen as fixed modification, 165 different peptides all containing heavy lysine were annotated. However, when the database search was conducted without that modification 52 different peptides were detected, none of which contained lysine, thus proving complete labeling of the sample.

The mass spectrometric analysis of two technical replicates of the SILAC-labeled membrane fraction resulted in the reproducible identification of 456 different proteins (Supporting Information Tables S2 and S3) with a false positive rate of 0.8%. Quantitative data were acquired for 234 unique proteins of which 121 (52%) contained MSDs (see Table 1 and Supporting Information Table S5).

The intraexperimental variation [36] was checked by calculating the average RSD for both experiments. The results for the two datasets were 10 and 11% respectively after the removal of outliers. The removal of outliers did not really influence the mean relative deviation, because only a minority of peptide ratios was removed (2.7 and 3.4% in the two datasets, respectively). Thus, the majority of the protein ratios were based on reproducible determination of peptide ratios. The interexperimental variation which was calculated from the relative difference of the two protein ratios was 18%, allowing us to define 20% relative difference as a threshold for reproducibility. The fact that single peptide quantifications also exhibited only an average relative difference of 18.6% prompted us to include this subgroup of proteins into our analysis. Applying these criteria a list of 192 proteins that yielded highly confident quantification data could be used for the analysis of starvation mediated changes in the membrane protein profile.
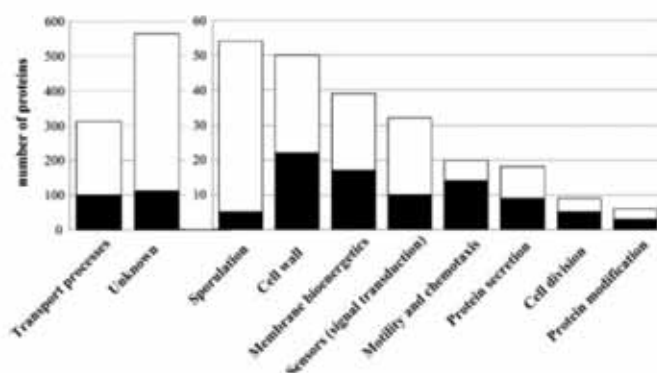


**Figure 2.** Assignment of identified membrane proteins into functional categories according to SubtiList. Given are theoretical numbers derived from the *in silico* membrane proteome analysis (white bars) and numbers of membrane proteins identified in this study (black bars).

Sixty-eight (35%) of these proteins displayed a growth phase dependent level (Table 1). The largest group consisted of 44 proteins which were considered as differentially expressed based on at least two peptide ratios in each analysis; data for eight proteins were based on at least two peptides in one and one peptide in the other technical replicate, and for 16 additional proteins one peptide in each of both technical replicates was used. In general, ratios of peptides obtained from adjacent gel slices were in agreement supporting the notion that membrane proteins were detected in more than one gel slice due to smearing rather then degradation.

An exception to this general statement is the protein YueB which was detected in different gel sections corresponding to a mass difference of approximately 50 kDa and displayed increased or decreased levels in stationary phase cells, respectively (Table 2). Both factors of YueB were determined with high confidence being based on at least ten peptide ratios. YueB is the prototype of phage-receptor proteins of Gram-positive bacteria [37], which consists of a long extracellular loop flanked by N- and C-termini that are likely integrated into the cytoplasmic membrane. The protein appears to be a dimer and the elongated extracellular loop reaches through cell-wall providing a binding sites for phage SPP1 [37]. The two forms of YueB that we detected probably constitute the extracellular domain and the full-length protein. Both forms were probably tightly bound to each other during the membrane preparation and then separated by SDS-PAGE. The two ratios likely reflect an increase in the level of full-length protein during entry into stationary phase and a decrease during entry into stationary phase of the fraction of the extracellular domain alone that was cleaved of their N- and C-terminal tails.

## 3.3 Quantification of membrane proteins after $^{14}$N/$^{15}$N-metabolic labeling

The *B. subtilis* wild-type strain 168 is the type strain in *B. subtilis* research but has only an auxotrophy for the low abundance amino acid tryptophan and is thus not suitable for quantification using the SILAC approach. Therefore, it was important to establish metabolic $^{14}$N/$^{15}$N labeling for *B. subtilis* 168. This approach can then easily be implemented in well-established workflows in many laboratories. Minimal media are commonly used to study the impact of various stresses onto bacterial physiology.

For quantification, peptide ratios were calculated with the MSQuant software which builds upon database searches performed with the MASCOT search engine. In order to be able to compare the data with those obtained with the SILAC approach, we had to inspect the comparability of SEQUEST and MASCOT database searches, because the first is based on heuristic and the latter on probability models. Therefore, false positive rates for SEQUEST and MASCOT database searches were compared. When MASCOT threshold was set to a score of 30 the false positive rate (0.98 and 0.87%) was slightly higher compared to using SEQUEST (0.86 and 0.65%) but in the same range allowing comparisons of both approaches.

Using metabolic labeling we were able to generate data for relative abundance for 579 quantification pairs representing 300 unique proteins (Table 1, Supporting Information Table S6). The intraexperimental variation of the $^{14}$N/$^{15}$N labeling experiments was determined to be 11.5 and 11.2% after removal of outliers (7.8 and 7.1%, respectively). The average difference between all ratios of both technical repli-

**Table 1.** Summary of data of SILAC- and $^{14}$N/$^{15}$N-labeling approaches for protein identification, quantification, and growth phase dependent expression

| | SILAC | | | $^{14}$N/$^{15}$N metabolic labeling | | |
|---|---|---|---|---|---|---|
| | Unique proteins | Proteins with MSDs[a] | Membrane associated proteins[a] | Unique proteins | Proteins with MSDs[a] | Membrane associated proteins[a] |
| **Number of proteins accessible to identification and quantification** | | | | | | |
| Number of proteins identified | 456 | 217 (48%) | 46 (10%) | 488 | 288 (59%) | 31 (6%) |
| Number of proteins accessible to quantification[b] | 234 | 121 (52%) | 24 (10%) | 300 | 180 (60%) | 29 (10%) |
| Number of proteins fulfilling significance criteria for quantification | 192 | 105 (54%) | 22 (11%) | 281 | 173 (61%) | 26 (9%) |
| ≥2 peptides in both analyses | 130 | 73 (56%) | 17 (13%) | 153 | 94 (61%) | 16 (10%) |
| ≥2 peptides in one analysis and one in the other | 24 | 15 (63%) | 2 (8%) | 40 | 23 (58%) | 2 (5%) |
| One peptide in both analyses | 38 | 17 (45%) | 3 (8%) | 88 | 56 (64%) | 8 (9%) |
| **Number of proteins displaying growth phase dependent differential regulation** | | | | | | |
| Total | 68 | 40 (59%) | 6 (9%) | 115 | 64 (56%) | 9 (8%) |
| ≥2 peptides in both analyses | 44 | 26 (59%) | 5 (11%) | 57 | 32 (56%) | 6 (11%) |
| ≥2 peptides in one analysis and one in the other | 8 | 4 (50%) | 0 | 15 | 8 (53%) | 0 |
| One peptide in both analyses | 16 | 10 (63%) | 1 (6%) | 43 | 24 (56%) | 3 (7%) |

a) Percentage indicates the fraction of the total proteins displaying MSDs or being associated with the membrane.
b) Unique proteins with peptides delivering quantitative data.

**Table 2.** Summary of proteins displaying growth-phase dependent variation in their level in the SILAC experiment (*B. subtilis* IS58, amino acid limitation in minimal medium)

| Protein (SubtiList) | Protein (Swiss-Prot)[a] | MSD | Additional information[b] | Function | No. of peptide ratios 1[c] | No. of peptide ratios 2[d] | Mean ratio (stat. vs. exp.)[e] | Relative difference (%)[f] |
|---|---|---|---|---|---|---|---|---|
| **Proteins increased in level in stationary phase** | | | | | | | | |
| AppB | | 6 | rr | Oligopeptide ABC transporter (permease) | 1 | 1 | 4.2 | 12.6 |
| AppF | | | * | Oligopeptide ABC transporter (ATP-binding protein) | 1 | 1 | 2.0 | 12.3 |
| AtpD | ATPB | | * | ATP synthase (subunit beta) | 2 | 3 | 1.5 | 19.7 |
| CtaC | COX2 | 3 | lipo | Cytochrome caa3 oxidase (subunit II) | 2 | 2 | 1.5 | 1.5 |
| CtaE[g] | COX3 | 5 | | Cytochrome caa3 oxidase (subunit III) | 1 | 1 | 2.4 | 2.0 |
| DltD | | 1 | | D-Alanine transfer from undecaprenol-phosphate to the poly(glycerophosphate) chain | 1 | 1 | 2.0 | 3.8 |
| McpB | | 2 | | Methyl-accepting chemotaxis protein | 2 | 3 | 1.5 | 13.3 |
| PbpX | | 1 | rr | Penicillin-binding protein | 1 | 1 | 2.1 | 8.1 |
| QcrA | | 1 | rr | Menaquinol/cytochrome *c* oxidoreductase (iron-sulfur subunit) | 3 | 1 | 3.0 | 12.3 |
| QcrC | | 3 | | Menaquinol/cytochrome *c* oxidoreductase (cytochrome *b/c* subunit) | 2 | 2 | 2.2 | 5.3 |
| RapC | | | | Response regulator aspartate phosphatase | 2 | 1 | 3.2 | 16.2 |
| SdhA | DHSA | 1 | | Succinate dehydrogenase (flavoprotein subunit) | 4 | 6 | 2.4 | 4.9 |
| YbcS[g] | SKFC | 1 | | Unknown | 1 | 1 | 3.0 | 58.4 |
| YdhK | | 1 | | Unknown | 1 | 2 | 2.0 | 10.1 |
| YdjL | | | | Unknown; similar to L-iditol 2-dehydrogenase | 3 | 2 | 1.5 | 4.6 |
| YknW | | 5 | | Unknown | 1 | 1 | 2.1 | 6.2 |
| YknX[g] | | 1 | | Unknown | 15 | 9 | 1.6 | 0.2 |
| YknZ[g] | | 4 | | Unknown | 2 | 2 | 1.9 | 1.5 |
| YodF | | 12 | | Unknown; similar to proline permease | 1 | 1 | 1.5 | 16.1 |
| YoeA[g] | | 11 | | Unknown | 1 | 1 | 2.5 | 0.2 |
| YolF | | 1 | | Unknown | 3 | 3 | 1.8 | 11.2 |
| YqfA[g] | | 2 | | Unknown | 3 | 4 | 1.8 | 9.5 |
| YtrF[g] | | 4 | | Unknown | 4 | 4 | 1.8 | 8.2 |
| YtxG[h] | | 1 | | Unknown; similar to general stress protein | 3 | 2 | 3.1 | 20.3 |
| YueB[g,h] | | 6 | | Unknown | 10 | 11 | 3.8 | 26.5 |
| YukC | | 1 | | Unknown | 1 | 3 | 1.7 | 4.9 |
| YurO | | 1 | | Unknown; similar to multiple sugar-binding protein | 3 | 3 | 2.3 | 1.3 |
| YvgW[g] | | 7 | | Unknown; similar to heavy metal-transporting ATPase | 5 | 5 | 6.5 | 2.7 |
| **Proteins reduced in level in stationary phase** | | | | | | | | |
| ArgD | | | | *N*-acetylornithine aminotransferase | 1 | 1 | 0.7 | 12.6 |
| ArgG[h] | ASSY | | | Argininosuccinate synthase | 1 | 1 | 0.3 | 41.8 |
| ArgJ | | | | Ornithine acetyltransferase/amino-acid acetyltransferase | 3 | 1 | 0.4 | 15.7 |
| AroA | AROG | | | 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase/chorismate mutase-isozyme 3 | 3 | 5 | 0.3 | 4.6 |
| CarB | CARY | | | Carbamoyl-phosphate transferase-arginine (subunit B) | 1 | 1 | 0.4 | 13.5 |
| DhbF[h] | | | | Involved in 2,3-dihydroxybenzoate biosynthesis | 4 | 3 | 0.3 | 61.6 |
| FhuD | | 1 | | Ferrichrome ABC transporter (ferrichrome-binding protein) | 6 | 3 | 0.7 | 4.0 |
| GlyA | | | | Serine hydroxymethyltransferase | 4 | 3 | 0.6 | 2.1 |
| Hag[h] | FLA | | * | Flagellin protein | 4 | 2 | 0.3 | 26.5 |
| Hom | DHOM | | | Homoserine dehydrogenase | 6 | 2 | 0.6 | 2.5 |
| LeuA[h] | LEU1 | | | 2-Isopropylmalate synthase | 3 | 2 | 0.3 | 41.5 |
| LytB | CWBA | 1 | sec | Modifier protein of major autolysin LytC | 4 | 2 | 0.6 | 3.1 |
| MntA | | | * | Manganese ABC transporter (membrane protein) | 3 | 3 | 0.6 | 7.3 |
| OpuAA | | | * | Glycine betaine ABC transporter (ATP-binding protein) | 3 | 3 | 0.2 | 2.4 |

_____

**Table 2.** Continued

| Protein (SubtiList) | Protein (Swiss-Prot)[a] | MSD | Additional information[b] | Function | No. of peptide ratios 1[c] | No. of peptide ratios 2[d] | Mean ratio (stat. vs. exp.)[e] | Relative difference (%)[f] |
|---|---|---|---|---|---|---|---|---|
| PbpD | | 1 | sec | Penicillin-binding protein 4 | 2 | 3 | 0.6 | 10.1 |
| PdhC | ODP2 | | * | Pyruvate dehydrogenase (dihydrolipoamide acetyltransferase E2 subunit) | 2 | 3 | 0.7 | 4.5 |
| PonA | PBPA | 1 | | Penicillin-binding proteins 1A/1B | 14 | 17 | 0.6 | 6.8 |
| PrkC | | 2 | | Probable membrane-linked protein kinase | 2 | 2 | 0.6 | 11.7 |
| PurC | Pur7 | | | Phosphoribosylaminoimida-zole succinocarboxamide synthetase | 1 | 1 | 0.5 | 11.0 |
| PyrAB | CARB | | | Carbamoyl-phosphate synthetase (catalytic subunit) | 6 | 4 | 0.5 | 7.2 |
| SerA | | | | Phosphoglycerate dehydrogenase | 4 | 5 | 0.6 | 2.3 |
| SrfAA[g, h] | | | | Surfactin synthetase/competence | 9 | 17 | 0.2 | 27.2 |
| SrfAC[g] | | | | Surfactin synthetase/competence | 6 | 8 | 0.6 | 3.4 |
| ThrS | SYT1 | | | Threonyl-tRNA synthetase (major) | 1 | 3 | 0.6 | 9.1 |
| TufA | EFTU | | | Elongation factor Tu | 8 | 4 | 0.6 | 9.3 |
| YaaD | PDXS | | | Unknown; similar to superoxide-inducible protein | 1 | 1 | 0.6 | 0.9 |
| YclQ[g] | | 1 | | Unknown; similar to ferrichrome ABC transporter (binding protein) | 6 | 5 | 0.1 | 1.4 |
| YluC | RASP | 4 | | Unknown | 1 | 2 | 0.6 | 6.7 |
| YoaB | | 11 | | Unknown; similar to alpha-ketoglutarate permease | 1 | 1 | 0.3 | 12.9 |
| YpuA | | 1 | sec | Unknown | 4 | 3 | 0.6 | 11.3 |
| YqzC | | 1 | | Unknown | 2 | 2 | 0.5 | 18.9 |
| YueB[g] | | 6 | | Unknown | 10 | 10 | 0.1 | 8.2 |
| YurU | | | | Unknown | 3 | 1 | 0.6 | 19.9 |
| YvaW | SDPA | | | Unknown | 2 | 4 | 0.3 | 4.4 |
| YvaY | SDPC | 2 | | Unknown | 1 | 1 | 0.6 | 4.2 |
| YvcB | | | | Unknown | 4 | 4 | 0.3 | 2.9 |
| YvhJ | | 1 | | Unknown; similar to transcriptional regulator | 1 | 1 | 0.3 | 11.3 |
| YwtF | | 1 | sec | Unknown; similar to transcriptional regulator | 1 | 1 | 0.6 | 15.6 |
| YxeB[g] | | 1 | | Unknown; similar to ABC transporter (binding protein) | 12 | 9 | 0.6 | 1.8 |
| YxkC[g] | | 1 | | Unknown | 6 | 6 | 0.6 | 7.5 |

a) Swiss-Prot (release 54.6) accession is also displayed if it differs from the SubtiList accession to incorporate changes in functional annotation that have been introduced since the last release of SubtiList.
b) sec, secreted *via* Sec-system; rr, secreted *via* Tat-system; lipo, lipoprotein-motif; *, retained in the membrane (according to SecP-algorithm).
c) Number of quantified peptides in technical replicate 1, for proteins occurring in more than one gel slice the sum of different quantified peptides is displayed.
d) Number of quantified peptides in technical replicate 2, for proteins occurring in more than one gel slice the sum of different quantified peptides is displayed.
e) Mean value of the protein ratios from both technical replicates.
f) Difference of the protein ratios from both technical replicates divided by the mean value.
g) Protein occurred in more than one gel slice.
h) Relative differences larger than 20% but quantitative change bigger than three-fold.

cates was 9% for quantifications relying on at least two peptide ratios and 14% for single peptide quantifications. Thus, the reliability cut-off for the standard variation between technical replicates was set to 20%. Furthermore, proteins exhibiting an at least three-fold change were regarded as significantly quantified, even when variation was larger than 20%, thus generating a list of 281 proteins which fulfilled these criteria. Table 1 illustrates that the majority of 153 quantifiable proteins was based on two or more peptides in both replicates. Furthermore, 40 protein ratios were based on at least two peptides in one and one peptide in the other replicate and 88 additional quantifications relied on a single peptide ratio in each replicate. Integral membrane proteins add up to 60% of all quantified proteins. Of the 281 proteins subjected to rigorous quantification, 115 proteins (41%) were found to vary significantly in their level with growth phase (Table 1 and Supporting Information Table S6).

### 3.4 Comparison of SILAC- and $^{14}$N/$^{15}$N-metabolic labeling based quantification

The quantification of technical replicates of both the SILAC- and the $^{14}$N/$^{15}$N-metabolically labeled samples exhibited a very good reproducibility (Figs. 3A and B). Furthermore, the threshold of 1.5-fold that was chosen for the quantitative analysis was also supported by the overall distribution of the quantitative data in both experimental setups (Figs. 3C and D). Also the intra- and interexperimental variation of both datasets was highly consistent. Thus, the data exhibited a similar quality and allowed a direct comparison of the methods.

Comparing the data it was obvious that the analysis of the $^{14}$N/$^{15}$N-metabolically labeled samples resulted in 1.5 times more significant quantitative data compared to SILAC (192 *vs.* 281 proteins, Table 1). One reason for this phenomenon is the fact that all amino acids are labeled and thus all peptides, which were identified, delivered quantitative data. The SILAC approach, however, offers only quantitative data for

lysine containing peptides, which is theoretically the case for approximately 60% of all peptides. This disadvantage could be compensated for by constructing a *B. subtilis* strain comprising auxotrophies for both lysine and arginine and using for SILAC-based experiments arginine and lysine simultaneously.

Comparing the number of peptides which contributed to the quantification it is apparent from Fig. 4 that $^{14}$N/$^{15}$N-metabolic labeling and the SILAC method generated in general very similar results. However, there were two differences, the $^{14}$N/$^{15}$N-metabolic labeling produced more hits with reproducible one peptide quantifications and significantly more hits for which quantification was based on more than 10 peptide ratios. Both effects probably once again reflect the fact that all peptides could by definition be used for quantification in the $^{14}$N/$^{15}$N experiment. Looking at the data it is also clear that significant proportions of the quantitative data were based on single peptide quantifications (31% in $^{14}$N/$^{15}$N-metabolically labeled samples and 20% for the SILAC samples). Even if these ratios have been measured reproducibly in technical replicates there is a requirement for the development of more bioinformatics tools that rigorously evaluate single peptide identifications [38–40] which might help to increase the fraction of reliably assigned mass spectrometric data even further.

Summarizing the comparison of both methods, it becomes clear that SILAC as well as $^{14}$N/$^{15}$N-labeling provide very reliable quantification of the membrane fraction with the latter allowing better coverage. SILAC offers a more straightforward analysis but requires one or even better two auxotrophies allowing simultaneous labeling with lysine and arginine. The $^{14}$N/$^{15}$N-labeling approach is amendable to bacteria even if auxotrophic mutants are difficult to construct but analysis is more labor intensive. Here one would profit if programs would use the same database search results, gen-
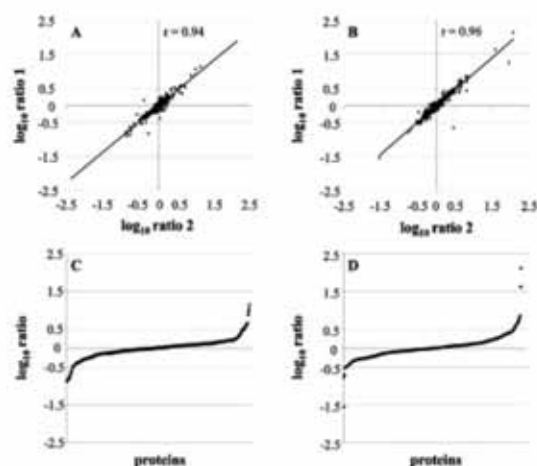


**Figure 3.** Reproducibility and distribution of the protein ratios. (A, B) Reproducibility of the data in the two technical replicates analyzed for amino acid limitation (A) and glucose limitation (B). Logarithms of protein ratios, which were determined by dividing the intensity values of exponential growth by those of stationary phase samples were plotted (A: SILAC – amino acid limitation; B: $^{14}$N/$^{15}$N-metabolic labeling – glucose limitation). The *y*-axis displays for each physiological condition the $\log_{10}$ protein ratios from the first technical replicate ($\log_{10}$ ratio 1) and the *x*-axis the $\log_{10}$ protein ratios from the second technical replicate ($\log_{10}$ ratio 2). The high correlation coefficients (A: *r* = 0.94 and B: *r* = 0.96) indicate the reproducibility of the experiments. (C, D) Logarithms of average protein ratios of technical replicates analyzed with SILAC (C – amino acid limitation) and $^{14}$N/$^{15}$N-metabolic labeling (D – glucose limitation) are plotted. In the *x*-axes all quantified proteins are ranked from the smallest to the largest average $\log_{10}$ ratio of both technical replicates to display the overall distribution of the ratios. The *y*-axis displays the average $\log_{10}$ protein ratios of both technical replicates.
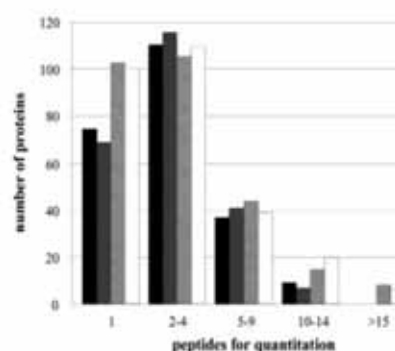


**Figure 4.** Comparison of the number of peptides available for quantification using SILAC and $^{14}$N/$^{15}$N-metabolic labeling. Protein quantifications are displayed as a function of peptide ratios (*x*-axis) that were used for the calculation of protein ratios (*y*-axis) for the technical replicates of the SILAC- (black and gray bars) and the $^{14}$N/$^{15}$N-metabolic labeling (shaded and white bars) approaches without filtering of outliers.

erated by the SEQUEST algorithm, both for the identification and the quantification avoiding time-consuming file conversions.

Furthermore, future analyses will certainly benefit from the application of more sophisticated software, *e.g.*, ASAP ratio [41], ProRata [42] or RelEx [43], for an accurate protein abundance ratio estimation.

### 3.5 Physiological consequences of lysine- and glucose limitation on the membrane proteome of *B. subtilis*

In this study, we wanted to compare the suitability of SILAC and a general uniform $^{14}N/^{15}N$-labeling for the analysis of the growth phase dependent changes in the membrane proteome of *B. subtilis*. For this reason we have chosen to analyze two important physiological scenarios. Lysine limitation in the *B. subtilis* strain IS58 which is a very specific limitation known to elicit the well-known stringent response and glucose limitation in the *B. subtilis* strain 168 which should also trigger the stringent response but on top of it a more general energy limitation.

When analyzing the observed changes in the membrane protein fraction one has probably to restrict interpretations to real membrane proteins. Changes in the still copurifying cytoplasmic proteins have to be viewed with caution because they might indeed reflect changes in level or merely a different distribution between the membrane fraction and the cytoplasmic fraction. In order to distinguish between effects on protein level and distribution, membrane fraction, cytoplasmic fraction and unfractionated crude protein extracts will have to be considered simultaneously in future studies.

The global consequences of glucose limitation on *B. subtilis* 168 physiology have already been investigated applying transcriptional profiling [44–46] or proteomics [2, 44]. However, the proteomic analyses of glucose starved cells have been restricted to cytoplasmic proteins and we now extend this analysis to the membrane proteome. The summary of the differentially regulated proteins, which is given in Table 3, documents the expected increase in the amount of *B. subtilis* transporters for alternative carbon-sources (DctP, GlpF, MsmX, MtlA, SacP, YvfH) like mannitol, lactate and glycerol. This broad spectrum is sort of surprising because the increase likely occurs in the absence of the specific carbon source – a phenomenon that has already been noted by Koburger *et al.* [46] in their transcriptomic study. Another example is the dipeptide transport system represented by the proteins DppB, DppC, and DppE. These proteins are encoded by genes belonging to the *dpp*-operon [47], which is repressed by glucose in growing cells and is expressed shortly after induction of stationary phase responses [48] in a Spo0A dependent manner. Additionally, the increased amount of the menaquinol cytochrome *c* reductase QcrABC after transition into stationary phase depends on Spo0A [49]. This result is supported by the early findings that Spo0A is activated during glucose limiting conditions [50].

The strong reduction of nonessential transport processes, like those for osmoprotective substances such as glycine betaine and choline (Opu transport systems) is probably caused by the need to cut down transport processes that consume energy. Such reduction in the level of membrane proteins might indicate increased degradation of proteins no longer needed during starvation in order to replenish the pool of building blocks.

The reduction of membrane associated proteins linked to chemotaxis and motility can also be interpreted as a general approach to cut down the energy-demanding process of motility in an attempt to save energy.

Furthermore, an increased amount of proteins of TCA cycle is observed (CitZ, Icd, SdhAB, SucC), which is in good agreement with previous reports [51–54]. However, we have to stress that of the TCA proteins just mentioned only SdhA can be argued with because only this protein contains an MSD.

As stated in the beginning of this chapter many of the proteins displaying significant changes in level have not been assigned a function yet and thus their role cannot be discussed. However, if the list of quantified proteins (Supporting Information Table S6) is inspected in more detail, it is possible to recognize a number of protein classes, the members of which display the same tendency in regulation. This additional class of regulated proteins encompasses for example the increase of the cytochrome aa3 oxidase (QoxAB) and the ATP-synthetase complex (AtpABDF). These results support the above-mentioned energy depletion of the cell. Furthermore, the slight increase of the subunits of Opp-transporter indicates the activation of Spo0A and thus the induction of early stationary phase responses [55].

Interpretation of the data of the lysine limitation is more difficult because the specific limitation is expected to trigger more subtle changes to the membrane proteome than glucose limitation. Furthermore, the impact of amino acid limitation on transcriptome and proteome has been studied by Eymann *et al.* [56], but this study utilized norvaline induced starvation for leucine and isoleucine in growing cells and covered only the initial phase up to 20 min after the application of norvaline. Thus, transcriptome or proteome analyses of amino acid starvation induced stationary phase to compare with are still lacking.

In this membrane fraction focused approach we observed during entry into stationary phase a significant increase in the level of the subunits B and F of the App peptide transport complex. This transporter was shown to be involved in sporulation and genetic competence before [57], but under the conditions examined in this study up-regulation of the transporter might just be destined to compensate for the lack of lysine. While the amount of the App-transporter increases, nonessential ATP-dependent transport processes, *e.g.*, for the import of cofactors, exhibited a reduction which could indicate an attempt to save resources. Simultaneously, proteins involved in membrane bioenergetic processes (QcrAB, CtaCE, AtpD) are increased. The induction of

**Table 3.** Summary of proteins displaying growth-phase dependent variation in their level in the $^{14}N/^{15}N$ experiment (*B. subtilis* 168, glucose limitation in minimal medium)

| Protein (SubtiList) | Protein (Swiss-Prot)[a] | MSD | Add. information[b] | Function | No. of peptide ratios 1[c] | No. of peptide ratios 2[d] | Mean ratio (stat. *vs.* exp.)[e] | Relative difference (%)[f] |
|---|---|---|---|---|---|---|---|---|
| **Proteins increased in stationary phase** | | | | | | | | |
| AcoL[g] | DLDH3 | | | Acetoin dehydrogenase E3 component (dihydro-lipoamide dehydrogenase) | 1 | 1 | 55.8 | 133.5 |
| Asd | DHAS | | | Aspartate-semialdehyde dehydrogenase | 1 | 1 | 3.7 | 2.6 |
| CitZ[h] | CISY2 | | | Citrate synthase II (major) | 4 | 4 | 2.3 | 8.5 |
| ClpC | | | | Class III stress response-related ATPase | 1 | 4 | 1.6 | 10.2 |
| DctP | DCTA | 8 | | C4-dicarboxylate transport protein | 1 | 1 | 41.6 | 16.2 |
| DnaK[h] | | | | Class I heat-shock protein (molecular chaperone) | 2 | 5 | 2.4 | 8.3 |
| DppB | | 6 | | Dipeptide ABC transporter (permease) (sporulation) | 1 | 1 | 2.1 | 11.9 |
| DppC | | 6 | | Dipeptide ABC transporter (permease) (sporulation) | 1 | 1 | 1.8 | 18.2 |
| DppE[h] | | 1 | lipo | Dipeptide ABC transporter (dipeptide-binding protein) (sporulation) | 9 | 7 | 1.9 | 1.6 |
| FolD | | | | Methylenetetrahydrofolate dehydrogenase/ methenyltetrahydrofolate cyclohydrolase | 1 | 1 | 2.3 | 8.8 |
| GapA[h] | G3P1 | | | Glyceraldehyde-3-phosphate dehydrogenase | 1 | 2 | 3.4 | 0.9 |
| GlnH[h] | | 1 | * | Glutamine ABC transporter (glutamine-binding protein) | 2 | 4 | 1.9 | 5.5 |
| GlpF | | 7 | | Glycerol uptake facilitator | 1 | 1 | 130.9 | 15.3 |
| GlyA[h] | | | | Serine hydroxymethyltransferase | 6 | 6 | 1.7 | 3.7 |
| HisS | SYH | | | Histidyl-tRNA synthetase | 2 | 1 | 2.0 | 10.4 |
| HtrA | | 1 | | Serine protease Do (heat-shock protein) | 1 | 2 | 1.6 | 6.1 |
| Icd | IDH | | | Isocitrate dehydrogenase | 5 | 6 | 7.0 | 2.1 |
| LuxS | | | | Probable autoinducer-2 production protein | 2 | 2 | 3.2 | 0.1 |
| MetE | | | | Cobalamin-independent methionine synthase | 4 | 3 | 2.5 | 10.4 |
| MmgE | | | | Function unknown | 1 | 1 | 1.6 | 1.2 |
| MsmX[h] | | | | Multiple sugar-binding transport ATP-binding protein | 1 | 1 | 2.7 | 3.3 |
| MtlA | PTM3C | 7 | | PTS mannitol-specific enzyme IICBA component | 1 | 1 | 6.4 | 4.4 |
| PurA | | | | Adenylosuccinate synthetase | 2 | 2 | 2.1 | 10.5 |
| PurB | PUR8 | | | Adenylosuccinate lyase | 1 | 1 | 4.1 | 11.9 |
| PycA[h] | PYC | | | Pyruvate carboxylase | 4 | 4 | 2.0 | 3.5 |
| QcrA | | 1 | rr | Menaquinol/cytochrome *c* oxidoreductase (iron-sulfur subunit) | 2 | 3 | 2.4 | 2.0 |
| QcrB | | 4 | | Menaquinol/cytochrome *c* oxidoreductase (cytochrome *b* subunit) | 4 | 3 | 2.5 | 6.4 |
| QcrC[h] | | 3 | | Menaquinol/cytochrome *c* oxidoreductase (cytochrome *b/c* subunit) | 1 | 2 | 2.4 | 8.6 |
| RapA | | | | Response regulator aspartate phosphatase | 1 | 1 | 6.7 | 15.2 |
| RocE[g] | | | | Amino acid permease | 1 | 1 | 3.2 | 50.3 |
| RpsC[g] | RS3 | | | Ribosomal protein S3 (BS3) | 1 | 1 | 5.8 | 26.9 |
| SdhB[h] | DHSB | | | Succinate dehydrogenase (iron-sulfur protein) | 5 | 5 | 1.6 | 0.7 |
| SucC[h] | | | | Succinyl-CoA synthetase (beta subunit) | 6 | 5 | 1.9 | 0.8 |
| YaaN | | | | Unknown; similar to toxic cation resistance | 2 | 1 | 1.6 | 9.9 |
| YcdH[h] | | 1 | lipo | Unknown; similar to ABC transporter (binding protein) | 4 | 4 | 2.1 | 8.2 |
| YcnJ | | 9 | | Unknown; similar to copper export protein | 1 | 2 | 2.6 | 0.2 |
| YhgC | | | | Unknown | 1 | 1 | 2.1 | 10.2 |
| YhxA[g] | | | | Unknown; similar to adenosylmethionine-8-amino-7-oxononanoate aminotransferase | 1 | 1 | 3.4 | 43.4 |
| YjhA | | 1 | * | Unknown | 4 | 4 | 2.6 | 3.7 |
| YkvW | | 6 | | Unknown; similar to heavy metal-transporting ATPase | 1 | 1 | 3.2 | 8.9 |
| YpmQ | SCO1 | 1 | lipo | Unknown | 1 | 1 | 1.6 | 12.1 |
| YqxM | | 1 | sec | Unknown | 1 | 1 | 1.7 | 19.2 |
| YrpE | | | lipo | Unknown | 3 | 3 | 2.5 | 1.3 |
| YtkA[h] | | 1 | lipo | Unknown | 1 | 1 | 2.1 | 5.1 |
| YtxH | | 1 | | Unknown; similar to general stress protein | 2 | 3 | 1.5 | 11.3 |

_____

**Table 3.** Continued

| Protein (SubtiList) | Protein (Swiss-Prot)[a] | MSD | Add. information[b] | Function | No. of peptide ratios 1[c] | No. of peptide ratios 2[d] | Mean ratio (stat. vs. exp.)[e] | Relative difference (%)[f] |
|---|---|---|---|---|---|---|---|---|
| YurY | V296 | | | Unknown; similar to ABC transporter (ATP-binding protein) | 1 | 1 | 4.5 | 6.1 |
| YvfH[N] | | 14 | | Unknown; similar to L-lactate permease | 3 | 3 | 4.2 | 0.2 |
| YvfV | | | | Unknown; similar to glycolate oxidase | 1 | 1 | 1.6 | 0.3 |
| YvgL | | 1 | lipo | Unknown; similar to molybdate-binding protein | 2 | 1 | 2.0 | 1.8 |
| YvrC | | 1 | lipo | Unknown; similar to iron-binding protein | 1 | 1 | 3.2 | 5.7 |
| YwbM | | 1 | * | Unknown | 4 | 3 | 1.8 | 12.7 |
| YwhK[g,h] | | | | Unknown | 1 | 1 | 4.7 | 32.0 |
| YwjH | TAL | | | Unknown; similar to transaldolase (pentose phosphate) | 2 | 2 | 1.5 | 2.7 |
| Yxal[N] | | 3 | | Unknown | 1 | 1 | 2.6 | 4.9 |
| **Proteins reduced in stationary phase** | | | | | | | | |
| AsnO | | | | Asparagine synthetase | 2 | 2 | 0.2 | 3.3 |
| ComEA[N] | | 1 | | Exogenous DNA-binding protein | 4 | 4 | 0.6 | 0.3 |
| ComER | | | | Nonessential gene for competence | 2 | 2 | 0.3 | 18.9 |
| DhbF | | | | Involved in 2,3-dihydroxybenzoate biosynthesis | 10 | 10 | 0.4 | 19.4 |
| FlgE | FLGG | | | Flagellar hook protein | 5 | 3 | 0.3 | 7.7 |
| FlhA | | 7 | | Flagella-associated protein | 3 | 3 | 0.3 | 19.5 |
| FliF[g,h] | | 2 | | Flagellar basal-body M-ring protein | 4 | 11 | 0.4 | 23.3 |
| FliY | | | | Flagellar motor switch protein | 1 | 1 | 0.3 | 0.6 |
| FtsH[h] | | 2 | | Cell-division protein/general stress protein (class III heat-shock) | 18 | 20 | 0.6 | 0.1 |
| GamP | PTW3C | 9 | | Probable PTS glucosamine-specific enzyme IICBA component | 1 | 1 | 0.6 | 0.8 |
| GoxB | GLOX | 1 | | Glycine oxidase | 2 | 2 | 0.5 | 1.0 |
| Hag | FLA | | | Flagellin protein | 4 | 4 | 0.03 | 6.1 |
| LeuA[N] | LEU1 | | | 2-Isopropylmalate synthase | 6 | 4 | 0.6 | 0.3 |
| LytA[N] | | 1 | lipo | Secretion of major autolysin LytC | 2 | 3 | 0.4 | 4.4 |
| McpA[N] | | 2 | | Methyl-accepting chemotaxis protein | 6 | 8 | 0.5 | 8.2 |
| McpB[h] | | 2 | | Methyl-accepting chemotaxis protein | 8 | 14 | 0.5 | 7.8 |
| McpC[N] | | 2 | | Methyl-accepting chemotaxis protein | 8 | 9 | 0.6 | 4.9 |
| MmsA | | | | Methylmalonate-semialdehyde dehydrogenase | 1 | 1 | 0.5 | 2.5 |
| MotA[N] | | 4 | | Motility protein (flagellar motor rotation) | 1 | 1 | 0.5 | 1.7 |
| MotB | | 1 | sec | Motility protein (flagellar motor rotation) | 1 | 1 | 0.3 | 12.1 |
| Nin | | | | Inhibitor of the DNA degrading activity of NucA | 1 | 2 | 0.6 | 3.0 |
| OpuAA[N] | | | | Glycine betaine ABC transporter (ATP-binding protein) | 2 | 2 | 0.4 | 1.7 |
| OpuAC | | 1 | lipo | Glycine betaine ABC transporter (glycine betaine-binding protein) | 8 | 3 | 0.6 | 6.8 |
| OpuBC | | | * | Choline ABC transporter (choline-binding protein) | 1 | 1 | 0.4 | 9.2 |
| OpuCA | | | | Glycine betaine/carnitine/choline ABC transporter (ATP-binding protein) | 3 | 4 | 0.2 | 0.8 |
| OpuCC | | 1 | lipo | Glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein) | 4 | 3 | 0.5 | 17.2 |
| OpuCD | | 5 | | Glycine betaine/carnitine/choline ABC transporter (membrane protein) | 1 | 1 | 0.4 | 0.3 |
| PspA | | | | Phage shock protein A homolog | 2 | 1 | 0.5 | 4.4 |
| SpoIIIAH[N] | SP3AH | 1 | | Mutants block sporulation after engulfment | 3 | 3 | 0.5 | 0.3 |
| SpoIVA[N] | SP4A | | | Required for proper spore cortex formation and coat assembly | 5 | 3 | 0.5 | 3.5 |
| SrfAA[N] | | | | Surfactin synthetase/competence | 36 | 35 | 0.5 | 11.1 |
| SrfAC | | | | Surfactin synthetase/competence | 1 | 7 | 0.5 | 12.9 |
| TlpB[N] | | 2 | | Methyl-accepting chemotaxis protein | 3 | 3 | 0.6 | 7.4 |
| YclF | | 13 | | Unknown; similar to di-tripeptide ABC transporter (membrane protein) | 1 | 1 | 0.5 | 4.4 |
| YclM | AK3 | | | Unknown; similar to aspartate kinase | 1 | 1 | 0.6 | 0.6 |
| YdcC[N] | | 1 | * | Unknown | 3 | 3 | 0.5 | 4.6 |

2074    A. Dreisbach *et al.*    *Proteomics* 2008, *8*, 2062–2076

**Table 3.** Continued

| Protein (SubtiList) | Protein (Swiss-Prot)[a] | MSD | Add. infor- mation[b] | Function | No. of peptide ratios 1[c] | No. of peptide ratios 2[d] | Mean ratio (stat. *vs.* exp.)[e] | Relative difference (%)[f] |
|---|---|---|---|---|---|---|---|---|
| YdhK | | 1 | * | Unknown | 2 | 3 | 0.5 | 6.9 |
| YfmT | | | | Unknown; similar to benzaldehyde dehydrogenase | 6 | 5 | 0.4 | 6.2 |
| YjcM | | 1 | sec | Unknown | 1 | 1 | 0.5 | 19.1 |
| YjcP | | 2 | | Unknown | 1 | 1 | 0.4 | 6.3 |
| YkrL | HPTX | 4 | | Unknown; similar to heat-shock protein | 4 | 4 | 0.6 | 2.0 |
| YkrT | MTNK | | | Unknown | 1 | 1 | 0.6 | 1.8 |
| YluC | RASP | 4 | | Unknown | 2 | 2 | 0.4 | 2.3 |
| YlxF | | 1 | sec | Unknown | 1 | 2 | 0.3 | 5.0 |
| YmdA[h] | | 1 | | Unknown | 11 | 12 | 0.6 | 1.1 |
| YobJ | | 3 | | Unknown | 4 | 4 | 0.6 | 2.0 |
| YonS[h] | | 1 | * | Unknown | 3 | 2 | 0.3 | 0.0 |
| YqeB | | 2 | | Unknown | 1 | 1 | 0.6 | 9.2 |
| YrkA | | 4 | | Unknown; similar to hemolysin-like | 1 | 1 | 0.6 | 0.6 |
| YrvD | | 2 | | Unknown | 1 | 1 | 0.6 | 18.1 |
| YtrF | | 4 | | Unknown | 2 | 1 | 0.6 | 11.1 |
| YukA | | | | Unknown | 4 | 4 | 0.5 | 12.8 |
| YukB | | 1 | | Unknown | 1 | 2 | 0.5 | 12.3 |
| YurP[h] | | | | Unknown; similar to glutamine-fructose-6-phosphate transaminase | 3 | 2 | 0.6 | 0.5 |
| YwjA | | 5 | | Unknown; similar to ABC transporter (ATP-binding protein) | 4 | 4 | 0.6 | 1.2 |
| YxkC[h] | | 1 | | Unknown | 9 | 8 | 0.3 | 2.1 |

a) Swiss-Prot (release 54.6) accession is also displayed if it differs from the SubtiList accession to incorporate changes in functional annotation that have been introduced since the last release of SubtiList.
b) sec, secreted *via* Sec-system; rr, secreted *via* Tat-system; lipo, lipoprotein-motif; *, retained in the membrane (according to SecP-algorithm).
c) Number of quantified peptides in technical replicate 1, for proteins occurring in more than one gel slice the sum of different quantified peptides is displayed.
d) Number of quantified peptides in technical replicate 2, for proteins occurring in more than one gel slice the sum of different quantified peptides is displayed.
e) Mean value of the protein ratios from both technical replicates.
f) Difference of the protein ratios from both technical replicates divided by the mean value.
g) Relative differences larger than 20% but quantitative change bigger than three-fold.
h) Protein occurred in more than one gel slice.

the *qcrABC* operon during transition into stationary phase and its dependence on Spo0A activation has already been demonstrated [49]. Expression of the SigH-dependent membrane proteins YtxG, YoeA, QcrA, and QcrC was also expected because SigH is activated upon amino acid limitation.

The membrane proteins YknX and YknZ were also observed with an increased quantity in the stationary phase cells. Both proteins are subunits of an uncharacterized ABC-transporter that is regulated by the extracytoplasmic sigma factor SigW [58]. It has already been shown that the maximum expression of *sigW* is reached during early stationary growth phase.

Additionally, in tendency an increase of ATP-synthetase in stationary phase (AtpABDFGH) in the lysine depleted medium was observed, which again might point to an energy depleted state of the starving cells.

## 4  Concluding remarks

In summary, we were able to show that both SILAC- and $^{14}$N/$^{15}$N-labeling provide similarly valuable data for the quantification of bacterial membrane proteins. Compared to *in vitro* labeling techniques these two methods offer the clear advantage that the label is introduced already during growth. Thus, the isolation of membrane proteins is carried out after the combination of differentially labeled cells in order to reduce the technical bias during the extended purification procedure.

Whereas the SILAC labeling is well suited for bacteria with amino acid auxotrophies the $^{14}$N/$^{15}$N-metabolic labeling can be applied to every bacterial strain growing in minimal medium containing a defined N-source. However, due to the lack of easily available convenient software for the evaluation of $^{14}$N/$^{15}$N-labeled samples, analysis of the data is con-

_____

siderably more time-consuming compared to the SILAC labeling. In general the $^{14}$N/$^{15}$N-labeling has some advantages such as labeling of all amino acids and thereby increasing the number of peptides for quantification. This clearly increases the reliability of the data. Ongoing physiological proteomic projects can now be complemented with the membrane proteome analysis, using the identical bacterial strains. The proof of principle scenarios, amino acid (lysine) and glucose starvation, which have been chosen in this study, have provided results that are compatible with and extend previous studies of the cytoplasmic proteome.

Both, SILAC as well as $^{14}$N/$^{15}$N labeling are compatible with 2-D PAGE, 2-D LC-MS/MS and GeLC-MS/MS and thus allow comprehensive simultaneous interrogation of cytoplasmic and enriched membrane proteomes.

## 5   References

[1] Hecker, M., Völker, U., Towards a comprehensive understanding of *Bacillus subtilis* cell physiology by physiological proteomics. *Proteomics* 2004, *4*, 3727–3750.

[2] Bernhardt, J., Weibezahn, J., Scharf, C., Hecker, M., *Bacillus subtilis* during feast and famine: Visualization of the overall regulation of protein synthesis during glucose starvation by proteome analysis. *Genome Res.* 2003, *13*, 224–237.

[3] Völker, U., Engelmann, S., Maul, B., Riethdorf, S. *et al.*, Analysis of the induction of general stress proteins of *Bacillus subtilis*. *Microbiology* 1994, *140*, 741–752.

[4] Tjalsma, H., Antelmann, H., Jongbloed, J. D., Braun, P. G. *et al.*, Proteomics of protein secretion by *Bacillus subtilis*: Separating the "secrets" of the secretome. *Microbiol. Mol. Biol. Rev.* 2004, *68*, 207–233.

[5] Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S. *et al.*, A proteomic view on genome-based signal peptide predictions. *Genome Res.* 2001, *11*, 1484–1502.

[6] Wilkins, M. R., Gasteiger, E., Sanchez, J. C., Bairoch, A., Hochstrasser, D. F., Two-dimensional gel electrophoresis for proteome projects: The effects of protein hydrophobicity and copy number. *Electrophoresis* 1998, *19*, 1501–1505.

[7] Adessi, C., Miege, C., Albrieux, C., Rabilloud, T., Two-dimensional electrophoresis of membrane proteins: A current challenge for immobilized pH gradients. *Electrophoresis* 1997, *18*, 127–135.

[8] Wall, D. B., Berger, S. J., Finch, J. W., Cohen, S. A. *et al.*, Continuous sample deposition from reversed-phase liquid chromatography to tracks on a matrix-assisted laser desorption/ionization precoated target for the analysis of protein digests. *Electrophoresis* 2002, *23*, 3193–3204.

[9] Washburn, M. P., Ulaszek, R., Deciu, C., Schieltz, D. M., Yates, J. R., III, Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal. Chem.* 2002, *74*, 1650–1657.

[10] Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., Gygi, S. P., Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. *J. Proteome Res.* 2003, *2*, 43–50.

[11] Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J. *et al.*, A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics* 2004, *4*, 2849–2876.

[12] Wolff, S., Otto, A., Albrecht, D., Zeng, J. S. *et al.*, Gel-free and gel-based proteomics in *Bacillus subtilis*: A comparative study. *Mol. Cell. Proteomics* 2006, *5*, 1183–1192.

[13] Bunai, K., Ariga, M., Inoue, T., Nozaki, M. *et al.*, Profiling and comprehensive expression analysis of ABC transporter solute-binding proteins of *Bacillus subtilis* membrane based on a proteomic approach. *Electrophoresis* 2004, *25*, 141–155.

[14] Bunai, K., Yamane, K., Effectiveness and limitation of two-dimensional gel electrophoresis in bacterial membrane protein proteomics and perspectives. *J. Chromatogr.* 2005, *815*, 227–236.

[15] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B. *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 2002, *1*, 376–386.

[16] Conrads, T. P., Alving, K., Veenstra, T. D., Belov, M. E. *et al.*, Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and 15N-metabolic labeling. *Anal. Chem.* 2001, *73*, 2132–2139.

[17] Wu, C. C., MacCoss, M. J., Howell, K. E., Matthews, D. E., Yates, J. R., III, Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* 2004, *76*, 4951–4959.

[18] Smith, I., Paress, P., Cabane, K., Dubnau, E., Genetics and physiology of the rel system of *Bacillus subtilis*. *Mol. Gen. Genet.* 1980, *178*, 271–279.

[19] Anagnostopoulos, C., Spizizen, J., Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* 1961, *81*, 741–746.

[20] Stülke, J., Hanschke, R., Hecker, M., Temporal activation of beta-glucanase synthesis in *Bacillus subtilis* is mediated by the GTP pool. *J. Gen. Microbiol.* 1993, *139*, 2041–2045.

[21] Möller, S., Croning, M. D., Apweiler, R., Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 2001, *17*, 646–653.

[22] Nakai, K., Kanehisa, M., A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 1992, *14*, 897–911.

[23] Sonnhammer, E. L. L., vonHeijne, G., Krogh, A., in: Glasgow, J., Littlejohn, T., Major, F., Lathrop, R. (Eds.), *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, Menlo Park, CA 1998, pp. 175–182.

[24] Tusnády, G. E., Simon, I., Principles governing amino acid composition of integral membrane proteins: Application to topology prediction. *J. Mol. Biol.* 1998, *283*, 489–506.

[25] Hirokawa, T., Boon-Chieng, S., Mitaku, S., SOSUI: Classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 1998, *14*, 378–379.

[26] Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S., van Dijl, J. M., Signal peptide-dependent protein transport in *Bacillus subtilis*: A genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* 2000, *64*, 515–547.

[27] Tjalsma, H., Feature-based reappraisal of the *Bacillus subtilis* exoproteome. *Proteomics* 2007, *7*, 73–81.

[28] Chang, S., Cohen, S. N., High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *Mol. Gen. Genet.* 1979, *168*, 111–115.

[29] Olsen, J. V., Ong, S. E., Mann, M., Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* 2004, *3*, 608–614.

[30] Schulze, W. X., Mann, M., A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* 2004, *279*, 10756–10764.

[31] Klein, C., Garcia-Rizo, C., Bisle, B., Scheffer, B. *et al.*, The membrane proteome of *Halobacterium salinarum*. *Proteomics* 2005, *5*, 180–197.

[32] Blonder, J., Goshe, M. B., Xiao, W., Camp, D. G., II *et al.*, Global analysis of the membrane subproteome of *Pseudomonas aeruginosa* using liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* 2004, *3*, 434–444.

[33] Fischer, F., Wolters, D., Rögner, M., Poetsch, A., Toward the complete membrane proteome: High coverage of integral membrane proteins through transmembrane peptide detection. *Mol. Cell. Proteomics* 2006, *5*, 444–453.

[34] Kyte, J., Doolittle, R. F., A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 1982, *157*, 105–132.

[35] Tjalsma, H., van Dijl, J. M., Proteomics-based consensus prediction of protein retention in a bacterial membrane. *Proteomics* 2005, *5*, 4472–4482.

[36] Molina, H., Parmigiani, G., Pandey, A., Assessing reproducibility of a protein dynamics study using *in vivo* labeling and liquid chromatography tandem mass spectrometry. *Anal. Chem.* 2005, *77*, 2739–2744.

[37] Sao-Jose, C., Lhuillier, S., Lurz, R., Melki, R. *et al.*, The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. *J. Biol. Chem.* 2006, *281*, 11464–11470.

[38] Rohrbough, J. G., Breci, L., Merchant, N., Miller, S., Haynes, P. A., Verification of single-peptide protein identifications by the application of complementary database search algorithms. *J. Biomol. Tech.* 2006, *17*, 327–332.

[39] Weatherly, D. B., Astwood, J. A., III, Minning, T. A., Cavola, C. *et al.*, A Heuristic method for assigning a false-discovery rate for protein identifications from Mascot database search results. *Mol. Cell. Proteomics* 2005, *4*, 762–772.

[40] Higdon, R., Kolker, E., A predictive model for identifying proteins by a single peptide match. *Bioinformatics* 2007, *23*, 277–280.

[41] Li, X. J., Zhang, H., Ranish, J. A., Aebersold, R., Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. *Anal. Chem.* 2003, *75*, 6648–6657.

[42] Pan, C., Kora, G., McDonald, W. H., Tabb, D. L. *et al.*, ProRata: A quantitative proteomics program for accurate protein abundance ratio estimation with confidence interval evaluation. *Anal. Chem.* 2006, *78*, 7121–7131.

[43] MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R., Yates, J. R., III, A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* 2003, *75*, 6912–6921.

[44] Yoshida, K., Kobayashi, K., Miwa, Y., Kang, C. M. *et al.*, Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res.* 2001, *29*, 683–692.

[45] Lorca, G. L., Chung, Y. J., Barabote, R. D., Weyler, W. *et al.*, Catabolite repression and activation in *Bacillus subtilis*: Dependency on CcpA, HPr, and HprK. *J. Bacteriol.* 2005, *187*, 7826–7839.

[46] Koburger, T., Weibezahn, J., Bernhardt, J., Homuth, G., Hecker, M., Genome-wide mRNA profiling in glucose starved *Bacillus subtilis* cells. *Mol. Genet. Genomics* 2005, *274*, 1–12.

[47] Remaut, H., Bompard-Gilles, C., Goffin, C., Frere, J. M., Van Beeumen, J., Structure of the *Bacillus subtilis* D-aminopeptidase DppA reveals a novel self-compartmentalizing protease. *Nat. Struct. Biol.* 2001, *8*, 674–678.

[48] Slack, F. J., Mueller, J. P., Sonenshein, A. L., Mutations that relieve nutritional repression of the Bacillus subtilis dipeptide permease operon. *J. Bacteriol.* 1993, *175*, 4605–4614.

[49] Yu, J., Hederstedt, L., Piggot, P. J., The cytochrome bccomplex (menaquinone:cytochrome *c* reductase) in *Bacillus subtilis* has a nontraditional subunit organization. *J. Bacteriol.* 1995, *177*, 6751–6760.

[50] Yamashita, S., Kawamura, F., Yoshikawa, H., Takahashi, H. *et al.*, Dissection of the expression signals of the spoA gene of *Bacillus subtilis*: Glucose represses sporulation-specific expression. *J. Gen. Microbiol.* 1989, *135*, 1335–1345.

[51] Hanson, R. S., Cox, D. P., Effect of different nutritional conditions on the synthesis of tricarboxylic acid cycle enzymes. *J. Bacteriol.* 1967, *93*, 1777–1787.

[52] Jin, S., Sonenshein, A. L., Transcriptional regulation of *Bacillus subtilis* citrate synthase genes. *J. Bacteriol.* 1994, *176*, 4680–4690.

[53] Kim, J. H., Voskuil, M. I., Chambliss, G. H., NADP, corepressor for the *Bacillus* catabolite control protein CcpA. *Proc. Natl. Acad. Sci. USA* 1998, *95*, 9590–9595.

[54] Tobisch, S., Zühlke, D., Bernhardt, J., Stülke, J., Hecker, M., Role of CcpA in regulation of the central pathways of carbon catabolism in *Bacillus subtilis*. *J. Bacteriol.* 1999, *181*, 6996–7004.

[55] Perego, M., Higgins, C. F., Pearce, S. R., Gallagher, M. P., Hoch, J. A., The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol. Microbiol.* 1991, *5*, 173–185.

[56] Eymann, C., Homuth, G., Scharf, C., Hecker, M., *Bacillus subtilis* functional genomics: Global characterization of the stringent response by proteome and transcriptome analysis. *J. Bacteriol.* 2002, *184*, 2500–2520.

[57] Koide, A., Hoch, J. A., Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol. Microbiol.* 1994, *13*, 417–426.

[58] Huang, X., Gaballa, A., Cao, M., Helmann, J. D., Identification of target promoters for the *Bacillus subtilis* extracytoplasmic function sigma factor, sigma W. *Mol. Microbiol.* 1999, *31*, 361–371.

_____

*Differential effect of YidC depletion on the membrane proteome of Escherichia coli under aerobic and anaerobic growth conditions*

RESEARCH ARTICLE

# Differential effect of YidC depletion on the membrane proteome of *Escherichia coli* under aerobic and anaerobic growth conditions

Claire E. Price[1], Andreas Otto[2], Fabrizia Fusetti[3], Dörte Becher[2], Michael Hecker[2] and Arnold J. M. Driessen[1]

[1] Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, Kluyver Centre for the Genomics of Industrial Fermentations and the Zernike Institute of Advanced Materials, University of Groningen, NN Haren, The Netherlands
[2] Institute for Microbiology, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany
[3] Membrane enzymology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh, AG Groningen, The Netherlands

YidC of *Escherichia coli* belongs to the evolutionarily conserved Oxa1/Alb3/YidC family. Members of the family have all been implicated in membrane protein biogenesis of respiratory and energy transducing proteins. The number of proteins identified thus far to require YidC for their membrane biogenesis remains limited and the identification of new substrates may allow the elucidation of properties that define the YidC specificity. To this end we investigated changes in the membrane proteome of *E. coli* upon YidC depletion using metabolic labeling of proteins with $^{15}N/^{14}N$ combined with a MS-centered proteomics approach and compared the effects of YidC depletion under aerobic and anaerobic growth conditions. We found that YidC depletion resulted in protein aggregation/misfolding in the cytoplasm as well as in the inner membrane of *E. coli*. A dramatic increase was observed in the chaperone-mediated stress response upon YidC depletion and this response was limited to aerobically grown cells. A number of transporter proteins were identified as possible candidates for the YidC-dependent insertion and/or folding pathway. These included the small metal ion transporter CorA, numerous ABC transporters, as well as the MFS transporters KgtP and ProP, providing a new subset of proteins potentially requiring YidC for membrane biogenesis.

## 1 Introduction

Approximately 30% of proteins encoded by *Escherichia coli* are destined for insertion into or translocation across the inner membrane. Membrane proteins perform many of the cell's essential functions such as energy transduction, solute transport and stimulus transduction. The general secretory (Sec) pathway is responsible for insertion of the majority of integral membrane proteins. The protein conducting channel of the Sec translocase consists of the heterotrimer SecYEG complex [1, 2] to which the ribosome associates to insert

**Correspondence:** Professor Arnold J. M. Driessen, Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, Kluyver Centre for the Genomics of Industrial Fermentations and the Zernike Institute of Advanced Materials, University of Groningen, 9751 NN Haren, The Netherlands
**E-mail:** a.j.m.driessen@rug.nl
**Fax:** +31-50-3632164

**Abbreviations: ABC,** ATP-binding cassette; **IMV,** inner membrane vesicles; **MFS,** major facilitator superfamily; **pfp,** percentage of false positives; **psp,** phage shock protein A; **Sec,** secretory; **TMS,** transmembrane sequence

newly synthesized membrane proteins into the lipid bilayer. In association with the motor protein SecA, the SecYEG channel mediates the translocation of large polar inter-membrane loops across the membrane [3]. Another protein found associated with the Sec translocase, YidC, can function together with the SecYEG channel as well as independently as a membrane insertase for small hydrophobic proteins.

YidC belongs to the evolutionarily conserved Oxa1/Alb3/YidC family which evolved before the divergence of the three major domains of life [4–6]. Members of the family have all been implicated in membrane protein biogenesis of respiratory and energy transducing proteins. To date, only two natural *E. coli* substrates have been identified as using the YidC-only, or Sec-independent, pathway for insertion, namely the c subunit of the $F_1F_o$ ATP synthase, $F_oc$ [7], and the channel of large conductance, MscL [8]. The a subunit of the $F_1F_o$ ATP synthase, $F_oa$ [9], subunit II of the cytochrome *o* oxidase, CyoA [10, 11], and subunit K of the NADH:ubiquinone oxidoreductase, NuoK [12], are inserted *via* the Sec pathway and also require YidC for their insertion. YidC has also been implicated in the folding of proteins following their insertion by the Sec translocase. The lactose permease, LacY [13], and MalF, of the maltose transport complex [14], have been shown to require YidC for correct folding and stability in the membrane. Finally, YidC mediates the membrane insertion of small phage coat proteins such as M13 procoat and Pf3 [15].

The number of proteins identified thus far to require YidC for their membrane biogenesis remains limited and the identification of new substrates may allow the elucidation of properties that confer YidC-dependence to this subset of membrane proteins. To this end we investigated changes in the membrane proteome of *E. coli* upon YidC depletion. YidC is essential for growth under both aerobic and anaerobic growth conditions [16, 17] likely owing to its role in the biogenesis of the energy transducing proteins that are essential under these conditions. We therefore investigated the effects of YidC depletion on the membrane proteome from cells grown aerobically and anaerobically. In the latter case fumarate was used as a terminal electron acceptor. The proteomic analysis of membrane proteins poses a number of challenges. Owing to their hydrophobic nature, membrane proteins are not suited to standard 2-D gel-based methods. Metabolic labeling of proteins with stable isotopes offers an alternative quantitation method. This can be combined with MS-centered proteomics approaches based on the separation of complex trypsin-digested samples by reverse phase chromatography in combination with MS analysis (LC-MS/MS).

We found that YidC depletion under aerobic conditions elicited a massive chaperone response to protein misfolding and/or aggregation while depletion under anaerobic conditions did not show this response. A number of transporter proteins were identified as possible candidates for the YidC-dependent membrane insertion and/or folding pathway.

## 2 Materials and methods

### 2.1 Bacterial strains and plasmids

The YidC depletion strain *E. coli* FTL10 [18] in which *yidC* is under the control of the *araBAD* promoter was a generous gift of Frank Sargent (University of East Anglia, Norwich, UK). *E. coli* FTL10 "YidC$^+$" and "YidC$^-$" strains were used as described previously [17]. Plasmids pEH1YidC [19] and pTrcYidC [20] were used for the overexpression of His-tagged and non-tagged versions of YidC, respectively.

### 2.2 Materials

Sodium fumarate was purchased from Sigma-Aldrich. $^{15}$N-ammonium sulphate was purchased from Cambridge Isotope Laboratories (USA) and trypsin from Promega (Germany). Antisera against DnaK, $F_oc$, GroEL, IbpA, NuoK, phage shock protein A (PspA) and LepB were generous gifts from Axel Mogk (Universität Heidelberg, Germany), Gabriele Deckers-Hebestreit (Universität Osnabrück, Germany), Saskia van der Vies (Vrije Universiteit Amsterdam), Ewa Laskowska (University of Gdansk, Poland), Takao Yagi (The Scripps Research Institute, USA), Jan Tommassen (Utrecht University, The Netherlands) and William Wickner (William Wickner, Dartmouth University, USA), respectively. Polyclonal SecA and YidC antibodies were from our laboratory collection. Alkaline phosphatase conjugated anti-chicken, anti-mouse and anti-rabbit IgG were purchased from Sigma-Aldrich.

### 2.3 Bacterial growth

For metabolic labeling experiments using $^{14}$N/$^{15}$N, YidC$^+$ and YidC$^-$ strains were grown aerobically or anaerobically at 37°C. For aerobic growth, cells were grown in parallel M63 minimal medium supplemented with 0.5% glucose [21]. Cells were depleted of YidC essentially as described previously [17, 22]. Briefly, cells were grown overnight in growth medium containing 0.5% arabinose after which cells were washed in fresh M63 minimal medium and resuspended in medium supplemented with glucose. The cells were grown to an OD$_{600}$ of 0.6 which corresponds to late exponential phase and then diluted two-fold with the same medium. The procedure was repeated until growth cessation of the YidC$^-$ strain. For anaerobic growth, cells were grown in basal anaerobic growth medium [23] containing glycerol and sodium fumarate added to 0.5% w/v and 10 mM, respectively, instead of glucose. Under these growth conditions, the cells were grown to an OD$_{660}$ of 0.3. Each strain was grown in duplicate in medium containing $^{14}$N ammonium sulphate as well as in medium supplemented with $^{15}$N ammonium sulphate. Cells grown in the presence of $^{15}$N ammonium sulphate were used as a reference pool as

described previously [24]. After cells were harvested, 225 OD-units of bacteria grown in the presence of $^{15}$N or $^{14}$N were combined and stored at $-80^{\circ}$C.

## 2.4 Sample preparation

Inner membrane vesicles (IMVs) were isolated as described previously [25]. The IMVs were subjected to sodium carbonate extraction [26] to further enrich for membrane proteins, and to reduce the amount of nonspecifically associated cytosolic proteins.

## 2.5 MS analysis

Inner membrane fractions were subjected to the GeLC MS-workflow including 1-D SDS gel electrophoresis followed by tryptic digestion as described in [27]. The resulting tryptic peptides were separated on a reverse phase chromatography column (Waters BEH 1.7 μm, 100-μm id × 100 mm) operated on a nanoACQUITY-UPLC (both Waters, Milford, MA, USA). Peptides were concentrated followed by desalting on a trapping-column (Waters nanoACQUITY UPLC column, Symmetry C18, 5 μm, 180 μm × 20 mm, Waters) for 3 min at a flow rate of 1 mL/min with 99% buffer A (90% acetonitrile, 0.1% acetic acid). The peptides were eluted and separated with a non-linear 80-min gradient from 5 to 60% ACN in 0.1% acetic acid at a constant flow rate of 400 nL/min. MS/MS data were acquired with the LTQ-Orbitrap mass spectrometer (Thermo Fisher, Bremen, Germany) equipped with a nanoelectrospray ion source. After a survey scan in the Orbitrap ($r = 30\,000$) MS$^2$ data was recorded for the five most intensive precursor ions in the linear ion trap.

The *.dta files were extracted from *.raw files using BioworksBrowser 3.3.1 SP1 (Thermo Fisher Scientific) with no charge state deconvolution and deisotoping performed on the data. The *.dta files were searched with SEQUEST version v28 (rev.12) (Thermo Fisher Scientific) against *E. coli* target-decoy protein sequence database (complete proteome set of *E. coli* K12 extracted from UniProtKB/Swiss-Prot [28] with a set of common laboratory contaminants) compiled using BioworksBrowser. The searches were implemented in two iterations: First, for the GeLCMS analyses the following search parameters were used: enzyme type, trypsin (KR); peptide tolerance, 10 ppm; tolerance for fragment ions, 1 amu; b- and y-ion series; methionine oxidation as a variable modification (15.99 Da); a maximum of three modifications *per* peptide was allowed. In the second iteration the mass shift of all amino acids completely labelled with $^{15}$N Nitrogen was taken into account in the search parameters.

Resulting *.dta and *.out files were assembled and filtered using DTASelect (version 2.0.25) (parameters GeLCMS: -y 2 -c 2 -C 4 --here --decoy Reverse_ -p 2 -t 2 -u --MC 2 -i 0.3 --fp 0.005). Data were revised using an in-house written java-script to ensure that each protein hit relied on at least two different peptides as judged by amino acid sequence.

Peptide ratios of heavy and light peptides according to the identification lists were determined by CenSus [29] and exported ($R^2$ values bigger than 0.7 and only unique peptides; proteins failing to be relatively quantified were checked manually in the graphical user interface for on/off proteins) to Excel. Proteins relatively quantified with at least two peptides were taken into account for the subsequent analysis.

The proteomics data were subsequently analyzed by the Rank Product analysis [30] using the software package R version 2.10.1 [31]. This generates a lists of proteins ranked according to log ratio and calculates a conservative estimate of the percentage of false positives (pfp). Proteins with pfp values smaller than 0.1 (10%) were regarded as differentially represented.

## 2.6 Isolation of protein aggregates

Protein aggregates were isolated from YidC$^+$ and YidC$^-$ cells according to the method described in [32]. The samples were separated by SDS-PAGE and stained with Bio-Safe Coomassie (Bio-Rad, USA). In-gel tryptic digestion was followed by peptide extraction and LC-MS/MS as described in [20]. MS analysis was performed using a MALDI-TOF/TOF 4800 Proteomics Analyzer instrument (Applied Biosystems, USA) in the *m/z* range 900–4000. Data acquisition was performed in positive ion mode. Peptides with signal-to-noise level above 100 were selected for MS/MS. Protein identification was performed with ProteinPilot$^{TM}$ 2.0 software using the Paragon$^{TM}$ Algorithm (Applied Biosystems/MDS Sciex), searching against the *E. coli* K12 UniProtKB/Swiss-Prot protein sequence database. Trypsin specificity and all default options were included in the search. The results were manually inspected and the identifications were accepted based on the independent identification of at least two unique peptides each with confidence of identification probabilities higher than 95%.

## 2.7 Identification of co-purifying proteins

For the overexpression of (His-tagged) YidC, Luria Bertani medium was used for aerobic growth and the anaerobic basal medium was supplemented with 0.1% yeast extract (Difco, USA). His-tagged YidC overexpressed from plasmid pEH1YidC was purified on a Ni$^{2+}$-NTA column in order to identify co-purifying proteins. The samples were separated by SDS-PAGE and identified as described for protein aggregates.

## 2.8 Protein determination and Western blotting

Protein concentrations of the membrane preparations and protein aggregates were determined with the DC Protein

_____

assay (Bio-Rad) using BSA as a standard. SDS-PAGE and immunoblot analyses were carried out according to methods that were described previously [33, 34]. Blotting efficiency was checked by immunoblot analysis using antisera directed against LepB. Signal capture and quantification were performed using the FUJIFILM LAS-4000 luminescent image analyzer.

## 3    Results

In total, 359 and 379 proteins were identified and quantified from the membrane preparations of aerobically and anaerobically grown *E. coli* cells, respectively. Each data set contained two biological repeats for each strain. Integral membrane and membrane-associated proteins were well represented (Supporting Information Fig. 1). In the aerobic data set 129 integral membrane proteins were quantified and in the anaerobic data set 116 were quantified. A total of 134 different integral membrane proteins were quantified.

The majority of the proteins in the data sets were unaffected by YidC depletion (Figs. 1A and B, and Supporting Information Table 1). Proteins exhibiting changes with an associated pfp value of less than 0.1 were considered to have changed significantly. Compared to the cytoplasmic proteins, a higher percentage of membrane or secreted proteins decreased following YidC depletion. In the aerobic data set only 5% of cytoplasmic proteins decreased compared to 20% of the integral membrane proteins upon YidC depletion. On the other hand, in the anaerobic data set none of the cytoplasmic proteins was decreased, whereas in the integral membrane proteins the levels were reduced to 6%. The integral membrane or membrane-associated proteins were subdivided into 13 subcategories according to their functions (Figs. 1C and D). In particular, proteins involved in energy transduction, transport and cell division were negatively affected by YidC depletion. The greatest fold decrease was observed for YidC, which was reduced in the aerobic and anaerobic data sets by 11- to 20-fold, respectively (Table 1). This demonstrates that YidC was extensively depleted during growth in the absence of arabinose.

### 3.1    Energy transducing proteins

In the current study subunits of 14 different energy transducing complexes were quantified. NDH-I consists of 13 subunits in *E. coli*, 7 of which were identified in this study (Supporting Information Table 1). While the membrane subunit NuoA decreased upon YidC depletion, subunits of the peripheral arm of complex I did not change significantly. Another membrane subunit, NuoK, has been shown to require YidC for insertion [12] but was not identified in this study. Immunoblot analysis showed that NuoK indeed decreased upon YidC depletion under both aerobic and anaerobic conditions (Supporting Information Fig. 2). Six of

the 13 subunits of the $F_1F_o$ ATP synthase were quantified but $F_o$a and $F_o$c, which have been shown previously to be YidC-dependent substrates, were not identified. Immunoblot analysis verified that the levels of $F_o$c decreased under both aerobic and anaerobic growth conditions (Supporting Information Fig. 2). In the aerobic data set, CydA of the cytochrome *d* terminal oxidase decreased. Also FrdA was observed to decrease upon YidC depletion, whereas in the anaerobic data set, CyoA and CyoB of the ubiquinol oxidase decreased while GlpB increased. In a recent study, the effect of YidC depletion on the transcriptome of *E. coli* was investigated [35]. It was found that the expression of subunits of the fumarate reductase decreased following YidC depletion, indicating that the effect may be caused by a decreased gene expression rather than impaired membrane biogenesis. However, transcription of subunits of complex I and the $F_1F_o$ ATP synthase increased upon YidC depletion. This may be a compensatory mechanism in response to impaired insertion of the YidC-dependent subunits in these complexes.

### 3.2    Transporter proteins

Following YidC depletion, subunits of 10 of the 24 ATP-binding cassette (ABC) transporters identified and quantified, decreased (Table 1). This included subunits of arginine, dipeptide, histidine, branched amino acid, lipoprotein-releasing, lipopolysaccharide and putrescine transporters. Also, levels of the subunits of four other ABC transporters were observed to increase, namely heme, methionine and phosphate import systems and the multidrug resistant-like ATP-binding protein, MdlB. A further nine transporter proteins were observed to decrease upon YidC depletion, including the small magnesium transport protein CorA and the major facilitator superfamily (MFS) transporters KgtP and ProP. The levels of the biopolymer transporter ExbB decreased by four-fold which was the greatest decrease observed after YidC. Protein levels of the phenylalanine-specific permease and a predicted transporter protein YbaL increased following YidC depletion.

### 3.3    Cell division and cell shape

It has previously been observed that YidC depleted cells grown aerobically are impaired in cell division [35]. A total of 12 proteins involved in cell division and cell shape were quantified and the levels of only two proteins, FtsX and NlpI, decreased upon YidC depletion (Table 1). FtsX is predicted to form an ABC transporter with FtsE [36], and a decrease in FtsX levels in the membrane may be responsible for the cell division impairment and filament formation observed. Levels of the shape determining proteins MreB and RodZ as well as WzzB increased upon YidC depletion. These changes were not pronounced when YidC depletion was carried out anaerobically.
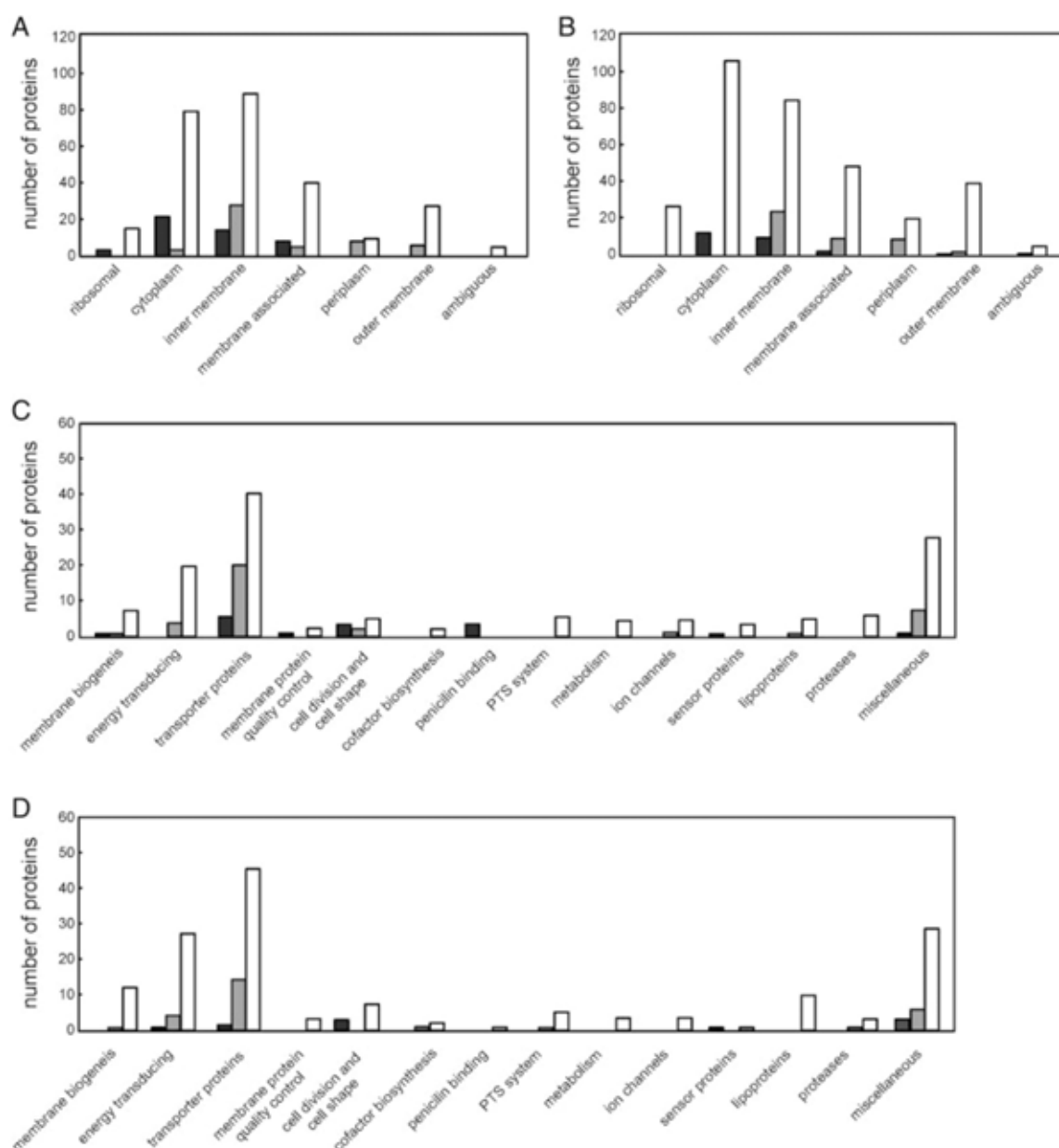
3239



**Figure 1.** Summary of localization and changes upon YidC depletion of proteins quantified in this study. Localizations were based on annotations in the UniProt KB/Swiss-Prot database for *E. coli* or where necessary Psort 2.0b [60]. (A and B) The number of proteins whose levels increase (black bars), decrease (grey) and remain unchanged (white bars) upon YidC depletion in the aerobic (A) and anaerobic (B) data sets. (C and D) Membrane proteins were divided into 13 categories based on function. The number of proteins whose levels increase (black bars), decrease (grey) and remain unchanged (white bars) upon YidC depletion in the aerobic (C) and anaerobic (D) data sets.

## 3.4 Miscellaneous

A further eight integral membrane proteins that do not fit into the categories shown in Fig. 1 were observed to decrease upon YidC depletion. These included the protoporphyrin oxidase, apolipoprotein N-acyltransferase, and

TolQ and TolR. Also, seven uncharacterized proteins decreased upon YidC depletion while one increased. Levels of the carboxypeptides DacA and DacC increased following YidC depletion.

YidC has been implicated in the quality control mechanisms that detect and degrade misfolded proteins in

Differential effect of YidC depletion on the membrane proteome of *Escherichia coli* under aerobic and anaerobic growth conditions

_____

C. E. Price *et al.*

*Proteomics* 2010, *10*, 3235–3247

**Table 1.** Membrane or membrane-associated proteins with significant changes following YidC depletion

| Gene name[a] | Protein name[b] | TMS/localization[c] | Log₂ (YidC⁺-YidC⁻)[d] | |
| --- | --- | --- | --- | --- |
| | | | Aerobic | Anaerobic |
| **Membrane biogenesis** | | | | |
| *yidC (oxaA)* | Inner membrane protein OxaA | 6 | −3.5 | −4.32 |
| *secA* | Protein translocase subunit SecA | c, im ass | 0.68 | −0.09 |
| **Energy transducing proteins** | | | | |
| *cydA* | Cytochrome *d* terminal oxidase, subunit I | 7 | −0.63 | −0.11 |
| *cyoA* | Ubiquinol oxidase subunit 2 | 2 | −0.36 | −0.36 |
| *cyoB* | Ubiquinol oxidase subunit 1 | 15 | 0.02 | −0.36 |
| *frdA* | Fumarate reductase flavoprotein subunit | c, im ass | −0.93 | 0 |
| *glpB* | Anaerobic glycerol-3-phosphate dehydrogenase subunit B | c, im ass | Not identified | 0.66 |
| *nuoA* | NADH-quinone oxidoreductase subunit A | 3 | −0.63 | −0.29 |
| *ynfG* | Probable anaerobic DMS reductase YnfG | c, im ass | Not identified | −0.4 |
| **ABC transporters** | | | | |
| *artI* | Arginine-binding periplasmic protein 1 | p | Not quantified[e] | −2.87 |
| *artP* | Arginine transporter ATP-binding protein ArtP | c, im ass | −0.25 | −0.29 |
| *cydC* | ATP-binding/permease protein CydC | 6 | 0.3 | 0.59 |
| *dppA* | Periplasmic dipeptide transport protein | p | −1.81 | −0.2 |
| *dppB* | Dipeptide transport system permease DppB | 7 | −0.61 | −0.43 |
| *dppD* | Dipeptide transport ATP-binding protein DppD | c, im ass | −0.51 | −0.22 |
| *hisP* | Histidine transport ATP-binding protein HisP | c, im ass | −0.71 | −0.16 |
| *livF* | High-affinty branched amino acid transport ATP-binding protein LivF | c, im ass | −0.45 | −0.65 |
| *livG* | High-affinty branched amino acid transport ATP-binding protein LivG | c, im ass | −0.32 | −0.63 |
| *livJ* | Leu/Ile/Val-binding protein | p | −1.37 | 0.03 |
| *lolD* | Lipoprotein-releasing system ATP-binding protein LolD | c, im ass | −0.62 | 0 |
| *lolE* | Lipoprotein-releasing system transmembrane protein LolE | 4 | −0.85 | −0.25 |
| *lptB* | Lipopolysaccharide export system ATP-binding protein LptB | c, im ass | −0.72 | −0.39 |
| *lptF* | Lipopolysaccharide export system permease protein LptF | 6 | −0.71 | −0.46 |
| *malK* | Maltose/maltodextrin import ATP-binding protein MalK | c, im ass | Not quantified[e] | −0.53 |
| *mdlB* | Multidrug resistance-like ATP-binding protein MdlB | 6 | 0.65 | Not identified |
| *metN* | Methionine import ATP-binding protein MetN | c, im ass | 0.7 | −0.18 |
| *metQ* | ᴅ-Methionine-binding lipoprotein MetQ | lp | 0.6 | 0.29 |
| *potA* | Spermidine/putrecine import ATP-binding protein PotA | c, im ass | 0.01 | −0.43 |
| *potF* | Putrescine-binding periplasmic protein | p | −1.15 | Not identified |
| *pstB* | Phosphate import ATP-binding protein PstB | c, im ass | 0.83 | 0.23 |
| *yecS* | Inner membrane amino-acid ABC transporter permease protein YecS | 3 | Not quantified[e] | −0.49 |
| *yjjK* | Uncharacterized ABC transporter ATP-binding protein YjjK | c, im ass | Not identified | −0.39 |
| **APC superfamily transporters** | | | | |
| *aroP* | Aromatic amino acid transporter protein AroP | 12 | −0.61 | −0.25 |
| *cycA* | ᴅ-Alanine/ᴅ-serine/glycine transporter | 12 | −1.01 | Not identified |

© 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

**www.proteomics-journal.com**

82

_____

3241

**Table 1.** Continued

| Gene name[a] | Protein name[b] | TMS/localization[c] | Log$_2$ (YidC$^+$-YidC$^-$)[d] | |
|---|---|---|---|---|
| | | | Aerobic | Anaerobic |
| *pheP* | Phenylalanine-specific permease | 12 | **1.46** | Not identified |
| **Other transport proteins** | | | | |
| *copA* | Copper-exporting P-type ATPase A | 8 | −0.71 | Not identified |
| *corA* | Magnesium transport protein CorA | 2 | −1.24 | −0.58 |
| *exbB* | Biopolymer transport protein ExbB | 3 | −2.18 | Not identified |
| *feoB* | Ferrous iron transporter protein B | 13 | −1.57 | Not identified |
| *kgtP* | Alpha-ketoglutarate permease | 12 | −0.73 | Not quantified[e] |
| *proP* | Proline/glycine betaine transporter | 12 | −0.62 | −0.23 |
| *ybaL* | Inner membrane protein YbaL | 13 | **1.39** | Not identified |
| *sstT* | Serine/threonine transporter SstT | 8 | −1.82 | **−0.44** |
| **Cell division and cell shape** | | | | |
| *ftsE* | Cell division ATP-binding protein FtsE | c, im ass | Not quantified[e] | **0.46** |
| *ftsX* | Cell division protein FtsX | 4 | −0.67 | −0.14 |
| *nlpI* | Lipoprotein NlpI | lp | −0.84 | −0.06 |
| *mreB* | Rod shape-determining protein MreB | c | **0.67** | **0.35** |
| *rodZ* | Cytoskeleton protein RodZ | 1 | **0.73** | 0.21 |
| *wzzB* | Chain length determinant protein | 2 | **0.67** | **0.14** |
| **Miscellaneous** | | | | |
| *dacA* | D-Alanyl-D-alanine carboxypeptidase DacA | 1 | **0.6** | 0.02 |
| *dacC* | D-Alanyl-D-alanine carboxypeptidase DacC | p, im ass | **0.89** | 0.09 |
| *hemG* | Protoporphyrin oxidase | amb | Not quantified[e] | **−0.74** |
| *hflK* | Protein HflK | 1 | **0.6** | 0.14 |
| *lnt* | Apolipoprotein *N*-acyltransferase | 6 | **0.59** | **0.3** |
| *mrdA* | Penicllin-binding protein 2 | p, im ass | **0.62** | Not identified |
| *mscS* | Small-conductance mechanosensitive channel | 3 | −0.63 | −0.01 |
| *phoQ* | Sensor protein PhoQ | 2 | **0.67** | **0.39** |
| *ptsG* | PTS system glucose-specific EIICB component | 10 | −0.37 | −0.33 |
| *rlpA* | Rare lipoprotein A | im lp | −0.88 | −0.06 |
| *sohB* | Probable protease SohB | 1 | −0.36 | −0.35 |
| *tolQ* | Protein TolQ | 3 | −0.79 | **−0.76** |
| *tolR* | Protein TolR | 1 | Not identified | **−0.94** |
| *ybbM* | UPF0014 inner membrane protein YbbM | 7 | −1.38 | Not identified |
| *ycjF* | UPF0283 membrane protein YcjF | 3 | **1.1** | **0.55** |
| *ydcL* | Uncharacterized lipoprotein YdcL | im lp | −0.88 | 0.18 |
| *yebE* | Inner membrane protien YebE | 1 | Not identified | **−0.94** |
| *yedE* | UPF0394 inner membrane protein YedE | 8 | Not quantified[e] | **−0.36** |
| *yegH* | UPF0053 protein YegH | 5 | −0.63 | 0.25 |
| *yfdH* | Bactoprenol glucosyl transferase homolog from prophage CPS-53 | 2 | −0.72 | −0.52 |
| *yjiG* | Inner membrane protein YjiG | 4 | Not identified | **−1.88** |
| *ypfJ* | Uncharacterized protein YpfJ | 1 | −0.73 | Not identified |
| *yqjE* | Inner membrane protein YqjE | 2 | **−1** | Not quantified[e] |

Numbers in bold indicate changes with associated pfp values of less than 0.1.
a) Gene designations in the UniProtKB/Swiss-Prot database for *E. coli*.
b) Protein designations in the UniProtKB/Swiss-Prot database for *E. coli*.
c) Localization based on Uniprot entry or where necessary Psort 2.0b. Abbreviations: c, cytoplasmc; im ass, inner membrane associated; im lp, inner membrane lipoprotein; p, periplasm; amb, ambiguous.
d) The log$_2$ ratio of each sample to the 15N reference was used to calculate the difference in protein levels between YidC+ and YidC− strains.
e) Not quantified: Protein was identified in at least one sample but insufficient data for quantitation.

_____

the inner membrane [37]. Upon YidC depletion, levels of HflK were observed to increase. HflK is, together with HflC, a regulator module of the ATP-dependent protease HflB (also known as FtsH) [38]. It has been shown previously that expression levels of *hflC* and *hflK* increase upon YidC depletion [35]. Therefore, the increased transcript and protein levels may be a response to the accumulation of misfolded membrane proteins under YidC depleting conditions.

## 3.5 Coimmunoprecipitation with His-YidC

Oxa1 has been isolated as a complex with *in vitro*-synthesized Atp9 (homologous to $F_oc$) as well as with the entire $F_1F_0$ ATP synthase suggesting a role for Oxa1 in the assembly of the protein complex [39]. In *E. coli*, *in vitro*-synthesized $F_oc$ has been shown to co-purify with YidC suggesting that it also contacts its substrates and remains briefly associated with them [40], whereas in *Bacillus subtilis*, the entire $F_1F_0$ ATP synthase can be found to be associated with the YidC homologs SpoIIIJ and YqjG [20]. Overexpression of YidC from plasmid pEH1YidC or pTrcYidC was induced and IMVs containing His-tagged YidC or wild-type YidC were purified on a $Ni^{2+}$-NTA column. Protein profiles were compared and protein bands specific for purification with His-YidC were excised, digested in gel with trypsin followed by LC-MS/MS. A total of 32 proteins were identified together with purified His-YidC (Supporting Information Table 2) of which 21 were integral membrane or membrane-associated proteins and 11 were cytoplasmic proteins. This included the transporter proteins CorA and MetN which were observed to decrease and increase, respectively, upon YidC depletion as well as PutP and YeeF which were not present in the metabolic labeling data sets. Several proteins involved in energy transduction were identified including the known YidC substrate CyoA, but also CyoB, HybA, Ndh and NuoC/D. Other membrane proteins identified included the sensor protein RstB, heme biosynthesis protein HemY and the cell shape determining proteins MraY and MreB. As previously described [37], the quality control proteins HflB, HflC and HflK were found associated with His-YidC. In addition, the HflC-related protein, QmcA and the chaperone DnaJ were identified reiterating a role for YidC in membrane protein folding and/or quality control.

## 3.6 Stress response proteins

Under aerobic growth conditions, protein levels of a large number of chaperones involved in the cell response to heat shock and protein misfolding and aggregation increased upon YidC depletion (Table 2). YidC depletion resulted in increased levels of the chaperones ClpB, DjlA, DnaJ, DnaK, GroL and HtpG (also known as Hsp90) between 1.7- and 4.2-fold. Levels

of the DeaD cold shock protein increased by nearly fivefold, whereas the small heat shock proteins IbpA and B increased by approximately 16-fold. The well-characterized PspA response was also observed upon YidC depletion with PspA, PspB and PspC increasing. This is also in agreement with recently described transcriptomic data in which the expression of the entire *pspABCDE* gene cluster increased by YidC depletion [35]. The response under anaerobic growth conditions was, however, entirely different. As previously characterized [17], there is no PspA response and although an increase was observed for DnaK and GroL, this was much reduced compared to that observed in the aerobic samples.

Many of the chaperones identified are cytoplasmic proteins and it was of interest whether the increase seen in the membranes represented increased levels in the cytoplasm or a sequestering of the chaperones to the membrane. Immunoblot analysis was therefore performed on cytoplasmic and membrane fractions (Fig. 2). Under aerobic growth conditions, levels of IbpA were found to increase in the membrane fraction with a smaller but reproducible increase in the cytoplasmic fraction indicating that there is a sequestering of the small chaperone to the membrane. Although IbpA was not identified in the metabolic labeling anaerobic data set, a similar but smaller response occurred under those conditions. Under aerobic conditions the chaperone GroEL showed a similar but not as pronounced sequestering to the membrane, while this was not evident for DnaK where levels increased equally in the cytoplasmic and membrane fractions under both aerobic and anaerobic conditions. Overall, the increase was more substantial in the aerobically grown cells consistent with the data from the metabolic labeling experiments. Immunoblot analysis of the inner membrane-associated protein PspA was consistent with the data from the metabolic labeling experiments.

## 3.7 Isolation of protein aggregates

The depletion of YidC under aerobic conditions resulted in the pronounced upregulation of cytoplasmic chaperones often associated with protein aggregation following heat shock, whereas the response was weak under anaerobic growth conditions. Therefore, we determined whether the different growth conditions affect the amount of protein aggregation or whether the cell responds differently depending on the availability of oxygen. Using selective solubilization in the detergent NP-40, aggregates were isolated from $YidC^+$ and $YidC^-$ strains grown under the same conditions as used for the proteomic analysis. Protein aggregates were quantified and expressed as a percentage of the total cellular protein (Supporting Information Fig. 3). Under aerobic conditions the amount of protein aggregates increased 1.6-fold, whereas under anaerobic conditions the amount increased by 1.25-fold. In total, 28 proteins were identified, none of which were integral membrane proteins (Supporting Information Table 3). Eight of these proteins were mislocalized outer membrane

**Table 2.** Proteins involved in stress responses with significant changes following YidC depletion

| Gene name[a] | Protein name[b] | TMS/localization[c] | Log$_2$ (YidC$^+$-YidC$^-$)[d] | |
|---|---|---|---|---|
| | | | Aerobic | Anaerobic |
| *clpB* | Chaperone protein ClpB | c | **2.05** | Not quantified[e] |
| *deaD* | Cold-shock DEAD box protein A | c | **2.28** | Not identified |
| *djlA* | DnaJ-like protein DjlA | 1 | **0.71** | −0.03 |
| *dnaJ* | Chaperone protein DnaJ | c | **2.11** | Not identified |
| *dnaK* | Chaperone protein DnaK | c, im ass | **1.67** | **0.58** |
| *dps* | DNA protection during starvation protein | c | Not identified | **0.68** |
| *groL* | 60 kDa chaperonin (GroEL protein) | c | **1.11** | **0.75** |
| *htpG* | Chaperone protein HtpG | c, im ass | **1.5** | Not identified |
| *ibpA* | Small heat shock protein IbpA | c | **4** | Not identified |
| *ibpB* | Small heat shock protein IbpB | c | **4.14** | Not identified |
| *kup* | Low affinity potassium transporter system protein Kup | 12 | **−1.24** | Not identified |
| *osmE* | Osmotically induced lipoprotein E | im lp | Not identified | **−0.58** |
| *pspA* | PspA | c, im ass | **1.91** | **−0.55** |
| *pspB* | PspB | 1 | **2.65** | **−0.48** |
| *pspC* | PspC | 1 ? | **2.61** | **−0.32** |
| *rpoE* | RNA polymerase sigma-E factor | c | Not quantified[e] | **0.52** |
| *rseP* | Regulator of sigma E protease | 3 | 0.13 | **−0.89** |
| *uspF* | Universal stress protein F | c | Not identified | **−0.28** |

Numbers in bold indicate changes with associated pfp values of less than 0.1.
a) Gene designations in the UniProtKB/Swiss-Prot database for *E. coli*.
b) Protein designations in the UniProtKB/Swiss-Prot database for *E. coli*.
c) Localization based on Uniprot entry or where necessary Psort 2.0b. Abbreviations: c, cytoplasm; im ass, inner membrane associated; im lp, inner membrane lipoprotein.
d) The log$_2$ ratio of each sample to the 15N reference was used to calculate the difference in protein levels between YidC+ and YidC- strains.
e) Not quantified: Protein was identified in at least one sample but insufficient data for quantitation.
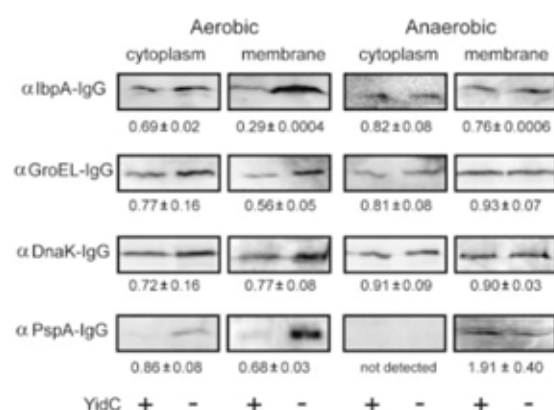


**Figure 2.** Chaperones IbpA and GroEL are sequestered to the membrane upon YidC depletion. Cytoplasmic and membrane fractions were analyzed by immunoblotting using antisera directed against IbpA, GroEL, DnaK and PspA. The ratio of the YidC$^+$/YidC$^-$ signals was calculated. Indicated are the means of three separate measurements ± the standard error of the mean.

and periplasmic proteins. Also proteins involved in the cellular stress response were identified, namely IbpA, DnaK, Dps and Spy. A large number of ribosome subunits were identified in the aggregates as well as prolyl-tRNA synthetase

and EF-Tu. MreB, which plays a role in the determination of cell shape in *E. coli* and which was observed to increase upon YidC depletion and co-purify with His-YidC, was also identified in the aggregates.

### 3.8 Periplasmic and outer membrane proteins

YidC depletion caused the accumulation of a number of mislocalized proteins in the cytoplasm some of which were also identified in the inner membrane preparation (Supporting Information Table 4). In the aerobic data set, 44 and 18% of the periplasmic proteins and outer membrane proteins quantified decreased upon YidC depletion (Fig. 1A, Supporting Information Table 3), whereas fewer changes were observed in the anaerobic data set (Fig. 1B). The observed decrease may be indicative of partial impaired translocation *via* the Sec translocase.

### 3.9 Ribosomal and cytoplasmic proteins

The levels of 21 cytoplasmic proteins increased following YidC depletion, including EF-Ts and SecA (Table 1, Supporting Information Table 5). Since SecA is localized in the cytoplasm as well as associated with the inner

membrane, the levels of SecA in these fractions were investigated using immunoblot analysis (Fig. 3). Under aerobic growth conditions, the level of SecA in the cytoplasm did not change significantly upon YidC depletion, while the amount associated with the membrane increased. Under anaerobic conditions no substantial change in the levels of SecA were observed. Also the levels of three ribosomal proteins increased upon YidC depletion in accordance with data which indicates that the transcription of a number of ribosomal subunits increases following YidC depletion [35].

## 4    Discussion

The aim of this study was to compare the impact of YidC depletion on the membrane proteome of aerobically and anaerobically grown *E. coli* cells, and to identify new candidate YidC substrates. The small metal ion transporter CorA was observed to decrease significantly under both aerobic and anaerobic conditions and was also found to co-purify with His-YidC. This 37 kDa protein has two transmembrane sequences (TMSs) with a large cytoplasmically located N-terminus. This helical hairpin might be a typical substrate for the YidC only pathway, but obviously its C-terminal location is rather unusual. In addition, a number of ABC transporters were found to decrease upon YidC depletion including the arginine, dipeptide, histidine, leucine-specific, lipoprotein-releasing, lipopolysaccharide and putrescine transporters. Levels of the maltose transporter MalFGK in membranes decreased upon YidC depletion as expected because of the role of YidC in the folding of MalF and its assembly into the MalFGK complex [14]. Thus, it seems that YidC fulfils a more general role in the biogenesis of ABC transporter subunits. Also the α-ketoglutarate permease KgtP and the proline transporter ProP decreased significantly upon YidC depletion and since these proteins share a similar topology to the lactose permease LacY, YidC may also be essential for their folding in the membrane. Not all MFS transporters were, however, negatively affected by YidC depletion indicating that the YidC-involvement needs to be evaluated case by case. Levels of the heptameric mechanosensitive ion channel MscS [41]



**Figure 3.** An increased amount of SecA is localized to the inner membrane under aerobic growth conditions. Cytoplasmic and membrane fractions were analyzed by immunoblotting using antisera directed against SecA. The ratio of the YidC⁺/YidC⁻ signals was calculated. Indicated are the means of three separate measurements ± the standard error of the mean.

decreased upon YidC depletion. Secondary effects, such as the reduction of the proton motive force as described previously [7], should always be taken into consideration when essential genes, such as YidC, are depleted in the cell. However, the validity of our approach is reinforced by observations that several known YidC substrates were identified.

Members of the Oxa1/Alb3/YidC family have been largely characterized in their role in the biogenesis of energy transducing complexes. For most of the subunits of energy transducing complexes identified in this study, no significant changes were observed. Interestingly, subunit I of the cytochrome *d* terminal oxidase decreased suggesting that YidC may also be involved in the biogenesis of this terminal oxidase. The smallest membrane embedded subunit of NDH-I, NuoK requires both YidC and the Sec translocase for insertion [12] and also its levels are down upon YidC depletion as demonstrated by immunodetection. Also, the membrane-embedded subunits NuoA and NuoL decreased upon YidC depletion. Complex I biogenesis has been proposed to include the formation of intermediate complexes resembling the three functional domains which in turn assemble into the haloenzyme. The NADH dehydrogenase subcomplex (NuoEFG) can form independently of the other subunits [42] and indeed changes in the levels of subunits F and G observed upon YidC depletion are very subtle. This suggests that the NADH dehydrogenase fragment assembles and is targeted to the membrane even when biogenesis of the membrane part of complex I is impaired.

A dramatic increase was observed in the chaperone-mediated stress response. Interestingly, this response was limited to aerobically grown cells, and appears to stem from the misfolding and aggregation of proteins in the cytoplasm. The PspA response is a well-studied response to YidC depletion [43]. The *psp* operon (*pspABCE*) is induced by many stress factors such as bacteriophage infection, ethanol, heat and osmotic shock, and has an unknown function in the maintenance of the proton motive force [44]. We have previously described that Psp response upon YidC depletion does not occur under anaerobic conditions [17]. The precise nature of the signals that induce *psp* expression remains to be determined but clearly depends on the nature of the stress as well as the availability of oxygen [45, 46]. It is unclear how *E. coli* deals with stresses under anaerobic conditions and whether other stress-related proteins replace the PspA response.

Since YidC depletion causes the aggregation of proteins in the cytoplasm the levels of the small chaperones IbpA and IbpB increased significantly under aerobic growth conditions. This response also occurs during heat shock and oxidative stress [47] as well as to overexpression of heterologous proteins [48]. The increase in IbpA levels was in particular pronounced in the membrane fraction which may indicate that IbpA acts on protein aggregation at the inner membrane surface. IbpA, together with IbpB, associates
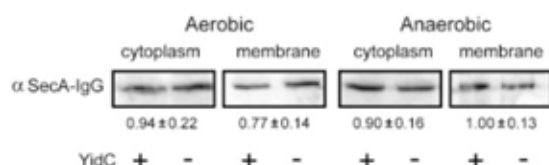
_____

with aggregated proteins to stabilize and protect them from irreversible denaturation and extensive proteolysis [49]. Aggregated proteins bound to the IbpAB complex are more efficiently refolded and reactivated by ATP-dependent chaperone systems such as ClpB, HtpG and DnaJ, DnaK and GlpE [50]. Under anaerobic conditions levels of IbpA, DnaK and GroEL were observed to increase but no significant sequestering to the membrane was observed. In both aerobic and anaerobic-grown cells, protein aggregation was observed upon YidC depletion. Proteins identified in the aggregates included mislocalized outer membrane and periplasmic proteins as well as chaperones. This most likely indicates a general defect in secretion in YidC-depleted cells. Remarkably, the difference in the extent of aggregation between aerobically and anaerobically grown cells was small. This suggests that *E. coli* may contain either a yet to be characterized chaperone systems that responds to protein aggregation under anaerobic conditions or that the cells are less able to respond to protein aggregation under anaerobic conditions. Alternatively, the slower growth rate of anaerobic cells (doubling time 8 h compared to 1 h for aerobically grown cells) may negate the need for upregulation of chaperone systems.

A similar chaperone response to that under aerobic conditions is observed when SecE is depleted from *E. coli* [51]. However in contrast to SecE depletion, YidC depletion causes an increase in the levels of the membrane protein quality control proteins HflBCK. This indicates that upon YidC depletion protein misfolding occurs not only in the cytoplasm but also in the membrane. YidC has been implicated in the correct folding of membrane proteins and their stability and incorporation into protein complexes. It may be in this capacity that YidC interacts with HflBCK to ensure correct folding of membrane proteins. The increase in the levels of ribosomal proteins present in the membrane fraction as well as SecA is also indicative of a defect in protein insertion and/or translocation *via* the Sec translocase. SecA translation has been shown to be upregulated in response to protein secretion defects [52, 53]. This response is modulated by SecM, whose gene is located upstream of *secA* [54]. Collectively, these data indicate that there is a certain degree of secretion hindrance when YidC is depleted from the cell but a complete blockage of the Sec translocase does not occur. SecA is not upregulated in anaerobically grown cells possibly indicative of a less severe secretion defect.

Although the current study was performed on inner membrane protein preparations there was some contamination of other proteins. Cytoplasmic proteins are very well-suited to trypsin digestion, therefore, the large number identified does not necessarily reflect a large quantity of cytoplasmic proteins present in the sample [55]. Also it is not clear whether the presence of outer membrane proteins in the inner membrane preparation is a contamination of outer membrane proteins because of the large quantity of porins such as OmpF and OmpC in the membrane [56] or a result of proposed contacts made between the inner and

outer membrane [57]. The proteomic analysis of membrane proteins remains challenging owing to their poor digestion by trypsin and unsuitability to conventional proteomic separation methods. The metabolic labeling and subsequent separation of trypsin fragments by LC-MS circumvents at least some of these problems. In addition, the use of an internal standard facilitates the comparison of biological replicates in a reliable manner [24]. In both the aerobic and anaerobic data sets approximately 120 integral membrane proteins containing between 1 and 13 TMSs were quantified. *E. coli* is predicted to contain 764 TMS-containing proteins [58] and in the present study 134 were quantitatively analyzed. Although this represents only 17.5% of the theoretical value, the number of proteins expressed under the conditions tested would be considerably less than 764. This indicates the suitability of this method for membrane proteome-wide studies in *E. coli*. On the other hand, small hydrophobic proteins were underrepresented in this study. For instance, $F_oc$ was not identified while the identified proteins with 1 or 2 TMSs generally contained large C- or N-terminal tails. Protein identification tools are based on trypsin generated fragments and TMSs are largely trypsin resistant owing either to a lack of charged residues and/or protection due to folding of the hydrophobic segments. Trypsin digestion combined with chemical cleavage may allow the identification of such small hydrophobic proteins.

In conclusion, YidC depletion results in protein aggregation/misfolding in the cytoplasm as well as misfolding in the inner membrane of *E. coli*. The chaperone response to this stress is much stronger under aerobic growth conditions compared to anaerobic conditions. The present study provides a set of potential new substrates for the design of *in vitro* insertion experiments such as those described previously [7, 59] to determine the extent of YidC involvement in their biogenesis.

## 5 References

[1] Breyton, C., Haase, W., Rapoport, T. A., Kuhlbrandt, W., Collinson, I., Three-dimensional structure of the bacterial

protein-translocation complex SecYEG. *Nature* 2002, *418*, 662–665.

[2] Van den Berg, B., Clemons Jr, W. M., Collinson, I., Modis, Y. et al., X-ray structure of a protein-conducting channel. *Nature* 2004, *427*, 36–44.

[3] Schiebel, E., Driessen, A. J. M., Hartl, F. U., Wickner, W., $\Delta\mu_{H}+$ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* 1991, *64*, 927–939.

[4] Pohlschroder, M., Hartmann, E., Hand, N. J., Dilks, K., Haddad, A., Diversity and evolution of protein translocation. *Annu. Rev. Microbiol* 2005, *59*, 91–111.

[5] Yen, M. R., Harley, K. T., Tseng, Y. H., SaierJr, M. H., Phylogenetic and structural analyses of the oxa1 family of protein translocases. *FEMS Microbiol Lett.* 2001, *204*, 223–231.

[6] Luirink, J., Samuelsson, T., de Gier, J. W., YidC/Oxa1p/Alb3: evolutionarily conserved mediators of membrane protein assembly. *FEBS Lett.* 2001, *501*, 1–5.

[7] van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., Driessen, A. J., $F_1F_O$ ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *J. Cell Biol.* 2004, *165*, 213–222.

[8] Facey, S. J., Neugebauer, S. A., Krauss, S., Kuhn, A., The mechanosensitive channel protein MscL is targeted by the SRP to the novel YidC membrane insertion pathway of *Escherichia coli*. *J. Mol. Biol.* 2007, *365*, 995–1004.

[9] Yi, L., Jiang, F., Chen, M., Cain, B. et al., YidC is strictly required for membrane insertion of subunits a and c of the $F_1F_O$ATP synthase and SecE of the SecYEG translocase. *Biochemistry* 2003, *42*, 10537–10544.

[10] du Plessis, D. J., Nouwen, N., Driessen, A. J., Subunit a of cytochrome *o* oxidase requires both YidC and SecYEG for membrane insertion. *J. Biol. Chem.* 2006, *281*, 12248–12252.

[11] van Bloois, E., Haan, G. J., de Gier, J. W., Oudega, B., Luirink, J., Distinct requirements for translocation of the N-tail and C-tail of the *Escherichia coli* inner membrane protein CyoA. *J. Biol. Chem.* 2006, *281*, 10002–10009.

[12] Price, C. E., Driessen, A. J., Conserved negative charges in the transmembrane segments of subunit K of the NADH:ubiquinone oxidoreductase determine its dependence on YidC for membrane insertion. *J. Biol. Chem.* 2010, *285*, 3575–3581.

[13] Nagamori, S., Smirnova, I. N., Kaback, H. R., Role of YidC in folding of polytopic membrane proteins. *J. Cell Biol.* 2004, *165*, 53–62.

[14] Wagner, S., Pop, O., Haan, G. J., Baars, L. et al., Biogenesis of MalF and the MalFGK$_2$ maltose transport complex in *Escherichia coli* requires YidC. *J. Biol. Chem.* 2008, *283*, 17881–17890.

[15] Kol, S., Nouwen, N., Driessen, A. J., Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes. *J. Biol. Chem.* 2008, *283*, 31269–31273.

[16] Samuelson, J. C., Chen, M., Jiang, F., Moller, I. et al., YidC mediates membrane protein insertion in bacteria. *Nature* 2000, *406*, 637–641.

[17] Price, C. E., Driessen, A. J., YidC is involved in the biogenesis of anaerobic respiratory complexes in the inner membrane of *Escherichia coli*. *J. Biol. Chem.* 2008, *283*, 26921–26927.

[18] Hatzixanthis, K., Palmer, T., Sargent, F., A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol. Microbiol.* 2003, *49*, 1377–1390.

[19] van der Laan, M., Houben, E. N., Nouwen, N., Luirink, J., Driessen, A. J., Reconstitution of Sec-dependent membrane protein insertion: nascent FtsQ interacts with YidC in a SecYEG-dependent manner. *EMBO Rep.* 2001, *2*, 519–523.

[20] Saller, M. J., Fusetti, F., Driessen, A. J., *Bacillus subtilis* SpoIIIJ and YqjG function in membrane protein biogenesis. *J. Bacteriol.* 2009, *191*, 6749–6757.

[21] Cohen, G. N., Rickenberg, H. V., Existence of specific acceptors for amino acids in *Escherichia coli*. *C.R. Hebd. Seances Acad. Sci.* 1955, *240*, 2086–2088.

[22] Nouwen, N., van der Laan, M., Driessen, A. J., SecDFyajC is not required for the maintenance of the proton motive force. *FEBS Lett.* 2001, *508*, 103–106.

[23] Lester, R. L., DeMoss, J. A., Effects of molybdate and selenite on formate and nitrate metabolism in *Escherichia coli*. *J. Bacteriol.* 1971, *105*, 1006–1014.

[24] MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R., Yates III, J. R., A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* 2003, *75*, 6912–6921.

[25] Kaufmann, A., Manting, E. H., Veenendaal, A. K., Driessen, A. J., van der Does, C., Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE. *Biochemistry* 1999, *38*, 9115–9125.

[26] Fujiki, Y., Hubbard, A. L., Fowler, S., Lazarow, P. B., Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J. Cell. Biol.* 1982, *93*, 97–102.

[27] Dreisbach, A., Otto, A., Becher, D., Hammer, E. et al., Monitoring of changes in the membrane proteome during stationary phase adaptation of *Bacillus subtilis* using *in vivo* labeling techniques. *Proteomics* 2008, *8*, 2062–2076.

[28] The UniProt Consortium., The Universal Protein Resource (UniProt) 2009. *Nucl. Acids Res.* 2009, *37*, D169–D174.

[29] Park, S. K., Venable, J. D., Xu, T., Yates III, J. R., A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* 2008, *5*, 319–322.

[30] Breitling, R., Armengaud, P., Amtmann, A., Herzyk, P., Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* 2004, *573*, 83–92.

[31] R Developement Core Team, R:A language and environment for statistical computing. 2006. Vienna, Austria.

[32] Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P., Bukau, B., Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. *Mol. Microbiol.* 2001, *40*, 397–413.

[33] Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970, *227*, 680–685.

[34] Towbin, H., Staehelin, T., Gordon, J., Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* 1979, *76*, 4350–4354.

[35] Wang, P., Kuhn, A., Dalbey, R. E., Global change of gene expression and cell physiology in YidC-depleted *Escherichia coli*. *J. Bacteriol.* 2010, *192*, 2193–2209.

[36] Schmidt, K. L., Peterson, N. D., Kustusch, R. J., Wissel, M. C. et al., A predicted ABC transporter, FtsEX, is needed for cell division in *Escherichia coli*. *J. Bacteriol.* 2004, *186*, 785–793.

[37] van Bloois, E., Dekker, H. L., Froderberg, L., Houben, E. N. et al., Detection of cross-links between FtsH, YidC, HflK/C suggests a linked role for these proteins in quality control upon insertion of bacterial inner membrane proteins. *FEBS Lett.* 2008, *582*, 1419–1424.

[38] Kihara, A., Akiyama, Y., Ito, K., A protease complex in the *Escherichia coli* plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. *EMBO J.* 1996, *15*, 6122–6131.

[39] Jia, L., Dienhart, M. K., Stuart, R. A., Oxa1 directly interacts with Atp9 and mediates its assembly into the mitochondrial $F_1F_O$-ATP synthase complex. *Mol. Biol. Cell* 2007, *18*, 1897–1908.

[40] Kol, S., Turrell, B. R., de Keyser, J., van der laan, M. et al., YidC-mediated membrane insertion of assembly mutants of subunit c of the $F_1F_O$ ATPase. *J. Biol. Chem.* 2006, *281*, 29762–29768.

[41] Bass, R. B., Strop, P., Barclay, M., Rees, D. C., Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel. *Science* 2002, *298*, 1582–1587.

[42] Schneider, D., Pohl, T., Walter, J., Dorner, K. et al., Assembly of the *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I). *Biochim. Biophys. Acta* 2008, *1777*, 735–739.

[43] van der Laan, M., Urbanus, M. L., Ten Hagen-Jongman, C. M., Nouwen, N. et al., A conserved function of YidC in the biogenesis of respiratory chain complexes. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 5801–5806.

[44] Kleerebezem, M., Crielaard, W., Tommassen, J., Involvement of stress protein PspA (phage shock protein A) of *Escherichia coli* in maintenance of the proton-motive force under stress conditions. *EMBO J.* 1996, *15*, 162–171.

[45] Weber, A., Kogl, S. A., Jung, K., Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in *Escherichia coli*. *J. Bacteriol.* 2006, *188*, 7165–7175.

[46] Jovanovic, G., Engl, C., Buck, M., Physical, functional and conditional interactions between ArcAB and phage shock proteins upon secretin-induced stress in *Escherichia coli*. *Mol. Microbiol.* 2009, *74*, 16–28.

[47] Kitagawa, M., Matsumura, Y., Tsuchido, T., Small heat shock proteins, IbpA and IbpB, are involved in resistances to heat and superoxide stresses in *Escherichia coli*. *FEMS Microbiol. Lett.* 2000, *184*, 165–171.

[48] Allen, S. P., Polazzi, J. O., Gierse, J. K., Easton, A. M., Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *J. Bacteriol.* 1992, *174*, 6938–6947.

[49] Kuczynska-Wisnik, D., Kedzierska, S., Matuszewska, E., Lund, P. et al., The *Escherichia coli* small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured *in vivo* during extreme heat shock. *Microbiology* 2002, *148*, 1757–1765.

[50] Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E., Bukau, B., Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol. Microbiol.* 2003, *50*, 585–595.

[51] Baars, L., Wagner, S., Wickstrom, D., Klepsch, M. et al., Effects of SecE depletion on the inner and outer membrane proteomes of *Escherichia coli*. *J. Bacteriol.* 2008, *190*, 3505–3525.

[52] McNicholas, P., Salavati, R., Oliver, D., Dual regulation of *Escherichia coli* secA translation by distinct upstream elements. *J. Mol. Biol.* 1997, *265*, 128–141.

[53] Schatz, P. J., Beckwith, J., Genetic analysis of protein export in *Escherichia coli*. *Annu. Rev. Genet.* 1990, *24*, 215–248.

[54] Nakatogawa, H., Ito, K., Secretion monitor, SecM, undergoes self-translation arrest in the cytosol. *Mol. Cell* 2001, *7*, 185–192.

[55] Graumann, J., Hubner, N. C., Kim, J. B., Ko, K. et al., Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell Proteomics.* 2008, *7*, 672–683.

[56] Molloy, M. P., Phadke, N. D., Maddock, J. R., Andrews, P. C., Two-dimensional electrophoresis and peptide mass fingerprinting of bacterial outer membrane proteins. *Electrophoresis* 2001, *22*, 1686–1696.

[57] Danese, P. N., Silhavy, T. J., Targeting and assembly of periplasmic and outer-membrane proteins in *Escherichia coli*. *Annu. Rev. Genet.* 1998, *32*, 59–94.

[58] Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. L., Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 2001, *305*, 567–580.

[59] van der Does, C., de Keyzer, J., van der Laan, M., Driessen, A. J., Reconstitution of purified bacterial preprotein translocase in liposomes. *Methods Enzymol.* 2003, *372*, 86–98.

[60] Gardy, J. L., Laird, M. R., Chen, F., Rey, S. et al., PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* 2005, *21*, 617–623.

_____

*Penicillin-binding protein folding is dependent on the PrsA peptidyl-prolyl cis-trans isomerase in Bacillus subtilis*

# Penicillin-binding protein folding is dependent on the PrsA peptidyl-prolyl *cis-trans* isomerase in *Bacillus subtilis*

Hanne-Leena Hyyryläinen,[1] Bogumila C. Marciniak,[2]
Kathleen Dahncke,[3†] Milla Pietiäinen,[1]
Pascal Courtin,[4] Marika Vitikainen,[1‡] Raili Seppala,[1]
Andreas Otto,[3] Dörte Becher,[3]
Marie-Pierre Chapot-Chartier,[4] Oscar P. Kuipers[2] and
Vesa P. Kontinen[1*]

[1]*Antimicrobial Resistance Unit, Department of Infectious
Disease Surveillance and Control, National Institute for
Health and Welfare (THL), P.O. Box 30, FI-00271
Helsinki, Finland.*
[2]*Molecular Genetics Group, Groningen Biomolecular
Sciences and Biotechnology Institute, University of
Groningen, Kerklaan 30, 9751 NN Haren, the
Netherlands.*
[3]*Institut für Mikrobiologie, Ernst-Moritz-Arndt Universität
Greifswald, Friedrich-Ludwig-Jahn-Str. 15, D-17489
Greifswald, Germany.*
[4]*INRA, UR477 Biochimie Bactérienne, F-78350
Jouy-en-Josas, France.*

## Summary

The PrsA protein is a membrane-anchored peptidyl-
prolyl *cis-trans* isomerase in *Bacillus subtilis* and
most other Gram-positive bacteria. It catalyses the
post-translocational folding of exported proteins and
is essential for normal growth of *B. subtilis*. We
studied the mechanism behind this indispensability.
We could construct a viable *prsA* null mutant in the
presence of a high concentration of magnesium.
Various changes in cell morphology in the absence of
PrsA suggested that PrsA is involved in the biosynthe-
sis of the cylindrical lateral wall. Consistently, four
penicillin-binding proteins (PBP2a, PBP2b, PBP3 and
PBP4) were unstable in the absence of PrsA, while
muropeptide analysis revealed a 2% decrease in the
peptidoglycan cross-linkage index. Misfolded PBP2a
was detected in PrsA-depleted cells, indicating that

PrsA is required for the folding of this PBP either
directly or indirectly. Furthermore, strongly increased
uniform staining of cell wall with a fluorescent vanco-
mycin was observed in the absence of PrsA. We also
demonstrated that PrsA is a dimeric or oligomeric
protein which is localized at distinct spots organized in
a helical pattern along the cell membrane. These
results suggest that PrsA is essential for normal
growth most probably as PBP folding is dependent on
this PPIase.

## Introduction

Intracellular folding of a protein into a native functional
structure is assisted by molecular chaperones and
foldase enzymes. A class of foldases ubiquitous in all
types of cells and cell compartments is formed by
peptidyl-prolyl *cis-trans* isomerases (PPIases), which
catalyse the isomerization of peptide bonds immediately
preceding proline residues (Schiene and Fischer, 2000;
Wang and Heitman, 2005; Lu and Zhou, 2007). Three
families of PPIases have been identified: cyclophilins,
FK506-binding proteins and parvulins (Rahfeld *et al.*,
1994). The archetype of the parvulin family of PPIases is
the *Escherichia coli* Par10, which consists of 92 amino
acid residues comprising the minimal catalytic domain
(Kühlewein *et al.*, 2004). In other parvulins, a catalytic
domain is flanked with N- and C-terminal regions of
various lengths and roles in substrate binding and/or
chaperone-like catalysis of folding (Lu *et al.*, 1996; Yaffe
*et al.*, 1997; Uchida *et al.*, 1999; Wu *et al.*, 2000; Behrens
*et al.*, 2001; Vitikainen *et al.*, 2004). Human Pin1 and its
counterpart parvulins in other eukaryotic cells specifically
recognize proline residues that are preceded by phospho-
rylated serine or threonine residues (Hani *et al.*, 1999; Lu
*et al.*, 1999; Yao *et al.*, 2001; Lu and Zhou, 2007). Other
parvulins have a wider substrate range, as their substrate
recognition is independent of phosphorylation (Hennecke
*et al.*, 2005; Stymest and Klappa, 2008).

PrsA is a lipoprotein bound to the outer face of the
cytoplasmic membrane in *Bacillus subtilis* and other Gram-
positive Firmicutes (Kontinen and Sarvas, 1993; Vitikainen
*et al.*, 2004). It consists of a diacylglycerol membrane
anchor, a large functionally unknown N-terminal domain
followed by a PPIase domain with similarity to the parvulin

family of PPIases and a small functionally unknown C-terminal domain (Vitikainen *et al.*, 2004). *Bacillus subtilis* PrsA exhibits PPIase activity but may also have a chaperone-like activity *in vivo* (Vitikainen *et al.*, 2004). The presence of a deletion in the PPIase domain of the *Lactococcus lactis* PrsA-like protein PmpA suggests that it may function only as a chaperone (Drouault *et al.*, 2002). The periplasmic SurA of *E. coli* is also a chaperone with a specialized role in the folding and assembly of outer-membrane proteins (Behrens *et al.*, 2001). Several extracellular proteins in various Gram-positive bacteria are secreted or matured in a PrsA-dependent manner (Kontinen and Sarvas, 1993; Hyyryläinen *et al.*, 2000; Vitikainen *et al.*, 2001, 2005; Drouault *et al.*, 2002; Lindholm *et al.*, 2006; Ma *et al.*, 2006; Alonzo *et al.*, 2009; Zemansky *et al.*, 2009). Overexpression of PrsA enhances α-amylase secretion from *Bacillus* and *L. lactis* cells (Kontinen and Sarvas, 1993; Vitikainen *et al.*, 2001, 2005; Lindholm *et al.*, 2006) including the biotechnically important thermoresistant AmyL α-amylase of *Bacillus licheniformis* (Kontinen and Sarvas, 1993). Some other extracellular proteins are also secreted at increased levels from PrsA-overexpressing cells (Wu *et al.*, 1998; Williams *et al.*, 2003). In *B. subtilis*, PrsA is an essential cell component in normal growth conditions indicating that it has an indispensable role in protein folding at the membrane-cell wall interface ('periplasm') (Vitikainen *et al.*, 2001). All three domains are essential for PrsA function (Vitikainen *et al.*, 2004). Inactivation of the D-alanylation system of the teichoic acids (Dlt) restores slight growth of *B. subtilis* cells lacking PrsA, suggesting that the increased net negative charge of the wall in the absence of Dlt partially suppresses the growth defect (Hyyryläinen *et al.*, 2000). In contrast to the rod-shaped *B. subtilis*, PrsA is a dispensable protein in several cocci, *L. lactis* (Drouault *et al.*, 2002), *Streptococcus pyogenes* (Ma *et al.*, 2006) and *Staphylococcus aureus* (Vitikainen *et al.*, unpublished).

In this study our purpose was to identify the indispensable cell component(s) which is (are) folded in a PrsA-dependent manner and to elucidate why PrsA is an essential protein in the rod-shaped bacterium *B. subtilis*, but non-essential in cocci. A hypothesis explaining this difference could be that PrsA catalyses the folding of a protein(s) involved in the biosynthesis of the cylindrical (lateral) cell wall and determination of the rod cell shape. The bacterial cell shape is maintained by a peptidoglycan cell wall (murein sacculus) and the actin-like proteins Mbl, MreB and MreBH, which form helical cables (cytoskeleton) that encircle the cell immediately beneath the cell membrane (Jones *et al.*, 2001; Carballido-Lopez and Errington, 2003; Soufo and Graumann, 2003; Defeu Soufo and Graumann, 2004; Stewart, 2005; Kawai *et al.*, 2009). The rod shape of *B. subtilis* is also dependent on other proteins including MreC and MreD, which are membrane proteins

and interact with each other and Mbl (Defeu Soufo and Graumann, 2006; van den Ent *et al.*, 2006). In the absence of any of these Mre proteins, cells are spherical or aberrant in shape or non-viable in normal growth conditions (Jones *et al.*, 2001; Leaver and Errington, 2005; Kawai *et al.*, 2009). Studies on the cell shape determination of *Caulobacter crescentus* and *B. subtilis* have also shown that MreB, MreC and MreD interact with penicillin-binding proteins (PBPs) (Figge *et al.*, 2004; Divakaruni *et al.*, 2005; van den Ent *et al.*, 2006; Kawai *et al.*, 2009). Peptidoglycan precursors are incorporated into the wall at distinct sites organized in a helical pattern along the lateral wall (Daniel and Errington, 2003; Tiyanont *et al.*, 2006; Divakaruni *et al.*, 2007). The Mre proteins and two PBP1-associated cell division proteins, EzrA and GpsB, are involved in the determination of the spatial organization and dynamics of the peptidoglycan synthesis (Claessen *et al.*, 2008; Kawai *et al.*, 2009). In *Corynebacterium glutamicum*, which does not have the *mreB* gene (Daniel and Errington, 2003), peptidoglycan precursors are incorporated for the whole lateral wall via the newly formed division poles. In *S. aureus* and probably in cocci more generally, peptidoglycan synthesis takes place only at the division septum and the hemispherical poles derived from it after cell division (Pinho and Errington, 2003).

High-molecular-weight PBPs are membrane-bound transglycosylase and transpeptidase enzymes which use peptidoglycan precursors to synthesize peptidoglycan chains and cross-link adjacent glycan chains to form a murein sacculus (Popham and Young, 2003; Sauvage *et al.*, 2008). Class A high-molecular-weight PBPs possess both transglycosylase and transpeptidase activities, whereas class B high-molecular-weight PBPs have only transpeptidase activity (Scheffers *et al.*, 2004). In addition to these two classes of high-molecular-weight PBPs, bacterial cells also contain low-molecular-weight PBPs (class C) which have neither transglycosylase nor transpeptidase activity, but function as carboxypeptidases or endopeptidases (Popham and Young, 2003; Sauvage *et al.*, 2008). The *B. subtilis* genome sequence has revealed 16 PBP-encoding genes, many of which are functionally redundant (Scheffers *et al.*, 2004). The PBPs have several distinct localization patterns in the cell suggesting dedicated functional roles for them in cell wall growth or cell division (Scheffers *et al.*, 2004). The PBP3 and PBP4a monofunctional transpeptidases and the PBP5 D-alanyl-D-alanine carboxypeptidase are localized in distinct spots or bands in the region of the lateral cell wall, suggesting their involvement in the elongation of the lateral wall. The bifunctional PBP1 is involved in the growth of both the lateral wall and the septum (Pedersen *et al.*, 1999; Scheffers *et al.*, 2004; Claessen *et al.*, 2008; Kawai *et al.*, 2009). The PBP2a and PbpH transpeptidases have a redundant, essential activity in the lateral wall synthesis and rod-shape determination

(Wei *et al.*, 2003), whereas the septal localization of PBP2b suggests a specific role for this transpeptidase in cell division (Scheffers *et al.*, 2004).

MreC and PBPs possess a number of proline residues (about 3% of amino acid residues) and the functional domains of these proteins are localized in the same cell compartment as PrsA suggesting that their folding could be dependent on PrsA. Misfolded proteins are rapidly degraded by the quality-control proteases such as HtrA (Hyyryläinen *et al.*, 2001). We studied the stability of MreC and PBPs in cells depleted of PrsA to find out whether misfolding of these proteins occurs in the absence of PrsA. Stabilities of three divisome proteins with a 'peri-plasmic' domain, FtsL, DivIB and DivIC (Daniel *et al.*, 1998; Katis and Wake, 1999) were also determined in a similar manner. Membrane proteomes were analysed to identify other possible PrsA-dependent membrane proteins. Various methods and approaches including electron microscopy, muropeptide analysis and labeling of PBPs and peptidoglycan precursors with fluorescent antibiotics were used to characterize the cell wall biosynthesis defect of PrsA-depleted cells and the functional role of PrsA in cell shape determination. Furthermore, we studied whether PrsA is evenly or non-evenly distributed along the cell membrane. The results showed that several PBPs are folded in a PrsA-dependent manner, suggesting that this is the likely cause for the growth arrest in the absence of PrsA.

## Results

*B. subtilis cells depleted of the PrsA protein are able to grow in the presence of a high concentration of magnesium*

In order to modulate cellular amount of PrsA, we have placed the *B. subtilis prsA* gene under the transcriptional control of the IPTG-inducible P*spac* promoter (Vitikainen *et al.*, 2001). Transmission electron microscope images of cells expressing P*spac-prsA* (IH7211) at low levels showed that PrsA depletion causes distinct changes in cell morphology. Severely PrsA-depleted cells which were still able to grow (P*spac-prsA* induced with 8 μM IPTG) were rod shaped but much larger than cells of the parental wild-type strain (compare Fig. 1A and B). At a PrsA level that was insufficient to support normal growth, large spherical cells were observed (Fig. 1C and D). Thick cell wall material was seen in residual division septa between the spherical cells, whereas in the periphery in many sites only a thin wall layer was left, suggesting that PrsA may be required for the biosynthesis of the lateral cell wall. These changes in cell morphology are similar to those seen in defects of the cell wall polymer biosynthesis and cell shape determination (Wei *et al.*, 2003; Leaver and Errington, 2005).
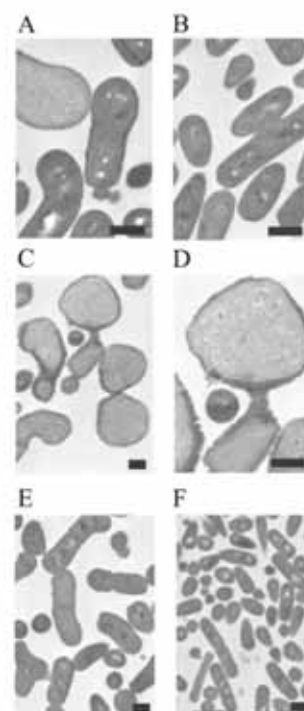


**Fig. 1.** Morphological changes of PrsA-depleted cells. Electron microscopy of PrsA-depleted cells was performed as described in *Experimental procedures*. The scale bars – 2 μm.
A. IH7211 (P*spac-prsA*) induced with 8 μM IPTG.
B. RH2111 wild-type strain.
C–D. IH7211 (P*spac-prsA*) in the absence of IPTG.
E. IH7211 (P*spac-prsA*) in the absence of IPTG and in the presence of 20 mM MgCl₂.
F. IH7211 (P*spac-prsA*) induced with 1 mM IPTG.

The MreC and MreD cell shape-determination proteins are required for lateral cell wall biosynthesis and normal growth of *B. subtilis*. However, a high concentration of magnesium (20 mM) in the growth medium restores some growth even in the complete absence of these proteins (Leaver and Errington, 2005). Cells depleted of MreC or MreD in the presence of 20 mM magnesium are spherical in shape. The mechanism of this suppression is currently unknown. Because magnesium also rescues growth of other mutants with defects in different aspects of cell wall biosynthesis (Δ*ponA* and Δ*mreB*), and which in contrast to *mreC* and *mreD* mutants maintain a wild type-like rod shape, it may affect indirectly peptidoglycan structure or turnover (Formstone and Errington, 2005). We studied whether the growth defect of PrsA-depleted cells can be suppressed by magnesium. The strain IH7211 (P*spac-prsA*) was grown in Antibiotic Medium 3 in the presence or absence of IPTG and the effect of 20 mM MgCl₂ on the growth was determined. Interestingly, magnesium restored the growth of PrsA-depleted IH7211 (Fig. 2). However, even though the PrsA-depleted cells grew in the
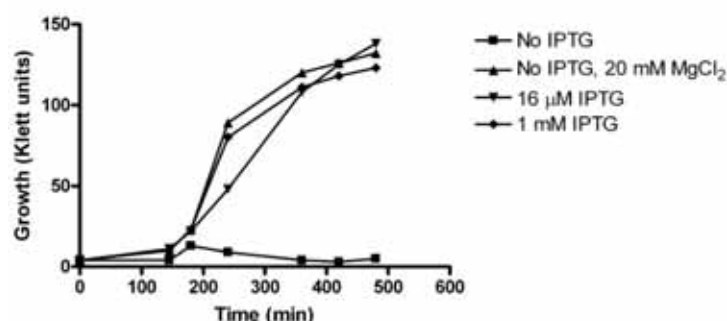
_____

**Fig. 2.** The suppression of the growth defect of PrsA-depleted cells by 20 mM MgCl₂. Cells of IH7211 (P*spac-prsA*) were grown in Antibiotic Medium 3 with the supplementations indicated.

presence of magnesium, it did not restore the 'normal' morphology of PrsA-expressing cells but they remained thick rods (Fig. 1E and F).

*PrsA is required for the stability of several penicillin-binding proteins*

The above results suggest that PrsA is involved in the folding of a membrane protein(s) which has (have) a large 'periplasmic' domain with an essential role(s) in the wall biosynthesis and/or cell shape determination. We used N-terminal green fluorescence protein (GFP) tags and specific antibodies against GFP, MreC and PrsA to determine effects of the PrsA depletion on the stability of potential PrsA-dependent membrane components. We started this study with GFP fusions, as most of these constructs were available and used previously in other studies (Scheffers *et al.*, 2004; Leaver and Errington, 2005). We were particularly interested in finding out whether the stability of MreC and PBPs is dependent on PrsA. The PBP 2b (PBP2b) interacts with the divisome proteins DivIB, DivIC and FtsL and affects their stability (Daniel *et al.*, 2006). Furthermore, an NMR study of DivIB suggested that *cis-trans* isomerization of a specific prolyl peptide bond causes a major structural change in this divisome protein and it was proposed that the *cis-trans* isomerization may function as a regulatory switch in cell division (Robson and King, 2006). Therefore, we were also interested in determining the stability of DivIB, DivIC and FtsL in PrsA-depleted cells.

The absence of PrsA did not decrease cellular levels of MreC and GFP-MreC as determined by immunoblotting with anti-MreC antibodies, suggesting that PrsA is not essential for the folding and stability of this cell shape-determination protein (Fig. S1). Hence, despite of the similar suppression of the growth defects of PrsA-depleted and MreC-depleted cells by magnesium, the growth inhibition in the absence of PrsA is most probably not caused by misfolding of MreC. In contrast, PBP2a, a high-molecular-weight PBP, which is involved in the lateral cell wall biosynthesis (Wei *et al.*, 2003), is most probably a PrsA-dependent protein, as indicated by a clear

decrease in the level of GFP-PBP2a in PrsA-depleted cells expressing this fusion protein (Fig. 3A) and in their isolated membranes (Fig. S1). The PrsA dependency of PBP1a/b, the largest of the PBPs of *B. subtilis*, was also studied with a GFP-PBP1a/b construct. GFP-PBP1a/b was stable in PrsA-depleted cells (Fig. S1), suggesting that PBP1a/b folding is independent of PrsA. The PrsA-dependency analysis of the DivIB, DivIC and FtsL divisome proteins revealed that none of them are PrsA dependent (Fig. S2).

The binding of β-lactam antibiotics to the native fold of PBPs can be used to visualize and determine cellular levels of enzymatically active PBPs. In order to elucidate the effect of the PrsA depletion on the levels of correctly folded PBPs, cell membranes were isolated from PrsA-depleted and non-depleted cells and PBPs were labeled with fluorescent penicillin, Bocillin-FL, and analysed with SDS-PAGE. There are 16 PBPs in *B. subtilis* and seven of them, PBP1a/b, PBP2a, PBP2b, PBP2c, PBP3, PBP4 and PBP5, could be identified with Bocillin-FL labeling (Fig. 3B). The identification of these PBPs is based on a similar labeling experiment with null mutants of the corresponding genes with the exception of PBP2b (Fig. S3) and their known migration in SDS-PAGE as determined in a previous study by another group (Popham and Setlow, 1996). The level of PBP2c encoded by the *pbpF* gene was clearly lower than the levels of the other PBPs. It was detected as a very weak band below the band of PBP2b (Fig. 3B and Fig. S3). The two forms of PBP1, a and b (Popham and Setlow, 1995), could not be separated in the mini-gels used in this analysis. The data demonstrate that four out of seven PBPs were affected by the PrsA depletion (Fig. 3B). The levels of active PBP2a, PBP2b, PBP3 and PBP4 were clearly lower in cells 'lacking' PrsA as compared with non-depleted cells, suggesting that their stability is dependent on the PrsA foldase, whereas the levels of PBP1a/b, PBP2c and PBP5 were not decreased, suggesting their independence of PrsA. These results are consistent with the similar effects of the PrsA depletion on GFP-PBP2a and GFP-PBP1a/b observed in the above immunoblotting experiments. The levels of the four PrsA-dependent PBPs were decreased in a similar manner also
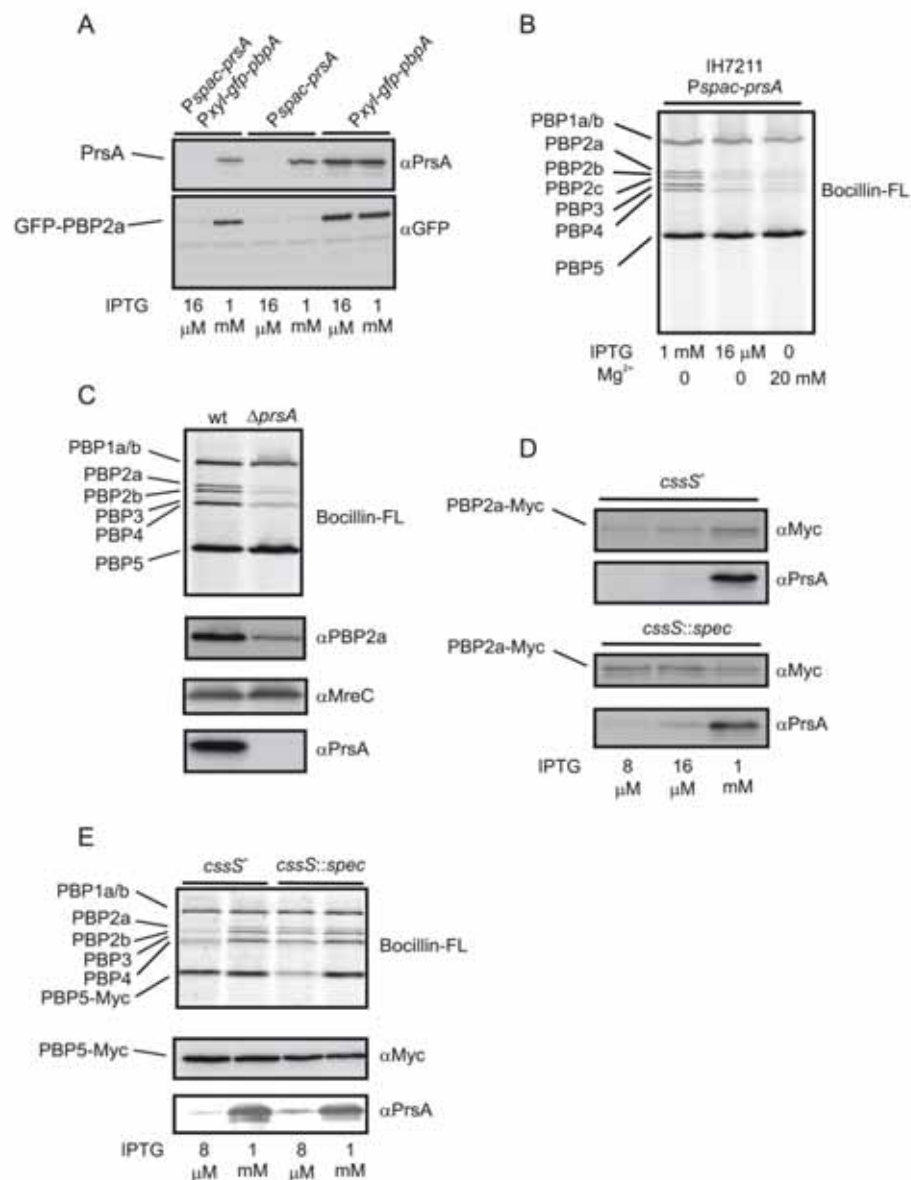
**Fig. 3.** PrsA depletion destabilizes penicillin-binding proteins.
A. Effect of PrsA depletion on the cellular level of GFP-PBP2a. Whole cell samples from *B. subtilis* IH8437 (P*xyl-gfp-pbpA* P*spac-prsA*) were prepared as described in *Experimental procedures* and levels of GFP-PBP2a and PrsA were determined by immunoblotting and chemiluminescence detection. P*xyl* was induced with 0.5% xylose.
B. Cytoplasmic membranes were isolated and levels of active PBPs were visualized by staining with Bocillin-FL (see *Experimental procedures*) and separation in SDS-PAGE. Effect of PrsA depletion and magnesium on the levels of active PBPs. The genes encoding PBP1a/b, PBP2a, PBP2b, PBP2c, PBP3, PBP4 and PBP5 are *ponA*, *pbpA*, *pbpB*, *pbpF*, *pbpC*, *pbpD* and *dacA* respectively.
C. Effect of the *prsA* null mutation (Δ*prsA*; *B. subtilis* IH9024) on levels of active PBPs (Bocillin FL staining) and total levels of PBP2a, MreC and PrsA (immunoblotting).
D. Levels of PBP2a-Myc in PrsA-depleted cells in the presence (*B. subtilis* IH9027; *pbpA-myc* P*spac-prsA*) and absence of CssS (*B. subtilis* IH9016; *pbpA-myc* P*spac-prsA cssS::spec*) as determined by immunoblotting.
E. Misfolding of PBP5-Myc in the absence of both PrsA and CssS. Bocillin-FL staining of active PBPs and immunoblotting of total PBP5-Myc in membranes of PrsA-depleted and non-depleted cells of IH9028 (*dacA-myc* P*spac-prsA*) and IH9029 (*dacA-myc* P*spac-prsA cssS::spec*).
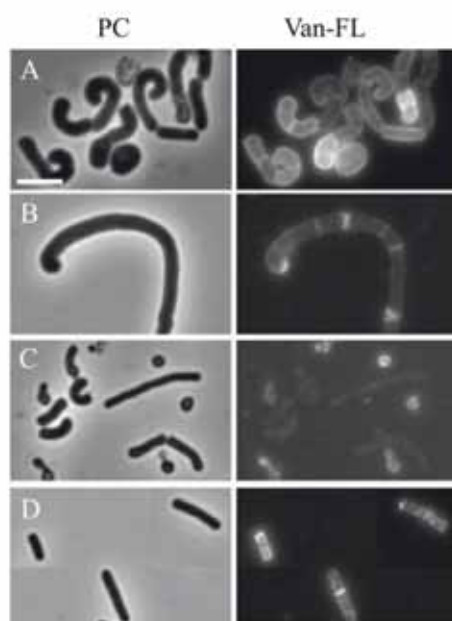
_____

PC    Van-FL



**Fig. 4.** Van-FL staining of *B. subtilis* Δ*prsA* mutant IH9024 (A–C) and 168 (D). Left panels – phase contrast images, right panels – Van-FL staining. Scale bar represents 6 µm (the same for all the images).
A. and D. Exponential growth phase.
B. Stationary phase.
C. Late stationary phase.

in the presence of 20 mM MgCl₂, suggesting that magnesium does not suppress their instability (Fig. 3B).

*The prsA gene can be deleted in the presence of a high concentration of magnesium*

The suppression of the growth defect of PrsA-depleted *B. subtilis* IH7211 cells in the presence of a high concentration of magnesium suggests that it might be possible to construct a *prsA* null mutant (gene replacement) on plates supplemented with magnesium. Indeed this was the case (see *Experimental procedures* for the mutant construction). The *prsA* null mutant grew on Antibiotic Medium 3 – plates supplemented with 20 mM MgCl₂ forming very tiny homogeneous colonies. Microscopic inspection of cells in the colonies showed that all bacteria were strongly deformed: either giant vesicular cells or thick twisted rods. These cells looked quite similar as PrsA-depleted cells of IH7211. In corresponding liquid cultures, the large deformed cells changed so that in overnight grown cultures small motile cocci-like cells, short bent rods, which were thinner than those in the exponential growth phase, and also fairly normal-looking rods were observed (see Fig. 4). The appearance of viable cocci-like *prsA* null mutant cells corroborates the evidence about the involvement of PrsA in lateral cell wall

biosynthesis and cell elongation. However, the presence of some rod-shaped bacteria in overnight cultures suggests that they were capable of synthesizing the cylindrical lateral wall in the absence of PrsA. We did not see on plates any fast-growing suppressors that would have overgrown the more slowly growing Δ*prsA* mutant.

Bocillin-FL labeling of *prsA* null mutant membranes showed that the level of activity of the same set of PBPs as in the case of PrsA-depleted IH7211 (PBP2a, PBP2b, PBP3 and PBP4) was decreased (Fig. 3C). Again, no effect was seen on PBP1a/b and PBP5. There was some experimental variation in the staining of PBP3 in membranes from PrsA-expressing strains (compare Fig. 3B and C), but not with the other PBPs. It may be more unstable than the other ones. We also determined protein levels of PBP2a and MreC in the null mutant and the wild-type parental strain by immunoblotting with anti-PBP2a and anti-MreC antibodies respectively. The PBP2a amount was clearly lower (> 50% reduction) in the absence of PrsA (Fig. 3C), whereas like in the case of PrsA-depleted IH7211 no effect of PrsA was seen on the MreC amount. This result suggests that in the absence of PrsA significant misfolding and degradation of several PBPs occurs. PrsA may not be involved in the folding of MreC.

*The CssRS two-component system is involved in the quality control of penicillin-binding proteins*

There are several 'periplasmic' proteases which could be involved in the quality control and degradation of PBPs including HtrA, HtrB, PrsW and WprA (Margot and Karamata, 1996; Stephenson and Harwood, 1998; Hyyryläinen *et al.*, 2001; Ellermeier and Losick, 2006; Heinrich *et al.*, 2008). Accumulation of misfolded proteins in the cell wall (misfolding or secretion stress) induces *htrA* and *htrB* gene expression in a manner dependent on the CssRS two-component system (Hyyryläinen *et al.*, 2001; Darmon *et al.*, 2002). CssRS is most probably dedicated to regulate the expression of only these two quality control protease genes (Hyyrylainen *et al.*, 2005). We used a knockout mutation of the *cssS* gene (*cssS::spec*) to study whether CssRS and the HtrA/B proteases it specifically regulates are involved in PBP degradation. We constructed strains which express PBP2a-Myc or PBP5-Myc proteins (PBPs modified with a C-terminal Myc-tag) and determined their levels in the presence and absence of *cssS::spec* and in PrsA-depleted and non-depleted cells by immunoblotting with anti-Myc antibodies. Bocillin-FL staining was also used to determine levels of active correctly folded PBPs in the *cssS::spec* mutant cells.

Immunoblotting revealed that the *cssS::spec* mutation decreased PBP2a-Myc degradation in PrsA-depleted cells (Fig. 3D), suggesting that HtrA and HtrB proteases are

involved in the degradation of misfolded PBP2a. The amount of PBP5-Myc as determined by immunoblotting was not significantly affected by *cssS::spec* (Fig. 3E). It was independent of CssRS and PrsA. However, Bocillin-FL labeling revealed that in the absence of CssS the level of active PBP5-Myc was clearly lower in PrsA-depleted cells than in non-depleted cells (Fig. 3E). The PrsA depletion and *cssS::spec* mutation alone had no effect on the level of active PBP5-Myc. These results show that CssRS, probably via HtrA/B proteases, and PrsA both are involved in the formation (folding) of active PBP5 (an overlapping function). The Bocillin-FL labeling (Fig. 3E) also suggested that absence of CssS (and HtrA/B proteases) increased the level of active PBP2a, PBP2b and PBP4 in PrsA-depleted cells. It seems that if PBP2a degradation is prevented (*cssS::spec*), at least some active correctly folded PBP2a is nevertheless formed with time despite of the absence or very low level of PrsA. PBP2a and PBP5 are quite different in this respect.

### Incorrectly folded PBP2a is formed in PrsA-depleted cells

In order to demonstrate that PrsA-depletion causes misfolding of PBP2a, we quantitated the levels of PBP2a-Myc in membranes of PrsA-depleted and non-depleted cells of *B. subtilis* IH9016 (P*spac-prsA pbpA-myc cssS::spec*) by immunoblotting with anti-Myc antibodies and the proportion of correctly folded PBP2a-Myc by Bocillin-FL labeling (Zhao *et al.*, 1999). Similar quantitation was also performed with the strain *B. subtilis* IH8266 (P*spac-prsA cssS::spec*) and anti-PBP2a antibodies. IH8266 expresses wild-type PBP2a instead of PBP2a-Myc.

The quantitation was performed with a series of twofold diluted membrane samples analysed by SDS-PAGE (Fig. S4). It was observed that the level of Bocillin-FL-binding, correctly folded full-length PBP2a-Myc and PBP2a were decreased by 27–40% in PrsA-depleted cells as compared with non-depleted cells. This indicates that PBP2a misfolding occurs in the absence of PrsA.

### Peptidoglycan cross-linkage degree is decreased and amount of muropeptides with pentapeptide chain is increased in cells depleted of PrsA

To elaborate the cell wall defect caused by the PrsA depletion, we performed a muropeptide analysis of cell walls isolated from PrsA-depleted and non-depleted cells of IH7211 (P*spac-prsA*) and the wild-type strain RH2111. Peptidoglycan was extracted from the cells in the late stationary phase as described in *Experimental procedures* and reduced muropeptides were analysed by RP-HPLC. The muropeptide profile consisted of 37 peaks, which were analysed by MALDI-TOF mass spectrometry (Table 1).

At a low level of PrsA (P*spac-prsA* induced with 8 µM IPTG) the amount of muropeptides with pentapeptide chain was significantly increased ($P < 0.0001$) as compared with that in RH2111 with the wild-type level of PrsA (Table 2). The induction of P*spac-prsA* expression in IH7211 with 1 mM IPTG restored a value close to that of the wild-type strain. The PrsA level obtained with the induction is about 60% of the wild-type PrsA level (Vitikainen *et al.*, 2001). The cross-linkage degree was significantly decreased ($P < 0.05$) by 2% in the PrsA-depleted cells compared with the wild-type strain (Table 2) and was restored to the wild-type level by the induction of P*spac-prsA* with 1 mM IPTG. Thus, we conclude that the observed differences for peptidoglycan structure are caused by the PrsA depletion.

### Van-FL imaging of the cell wall defect of prsA mutants

The cell wall defect was further characterized by imaging peptidoglycan biosynthesis in the *prsA* null mutant and PrsA-depleted cells of *B. subtilis* IH7211 with a fluorescent vancomycin (Van-FL). Van-FL binds to the terminal D-Ala-D-Ala moieties in non-cross-linked peptidoglycan precursors and growing glycan chains (Daniel and Errington, 2003; Tiyanont *et al.*, 2006). It has been shown by using Van-FL staining and fluorescence microscopy that lateral wall peptidoglycan polymers are synthesized in distinct spots organized in a spiral pattern (Daniel and Errington, 2003; Tiyanont *et al.*, 2006). PBPs that are located in the lateral wall in a similar spiral organization pattern are responsible for the synthesis of the lateral wall peptidoglycan. On the other hand, PBPs that normally synthesize the division septum are also capable of synthesizing lateral wall peptidoglycan (Daniel and Errington, 2003).

In cells of *B. subtilis* 168 and IH7211 (P*spac-prsA*) induced with 1 mM IPTG, fluorescence was mainly seen in the division septum (Figs 4D and 5C, right panels, respectively), but the spiral synthesis pattern of lateral wall peptidoglycan was also observed particularly in the case of exponential-phase cells of *B. subtilis* 168 (Fig. 4D). The *prsA* null mutant and PrsA-depleted cells of IH7211 (8 µM and 16 µM IPTG) were more intensively fluorescent than wild-type and non-depleted cells. In thick rods and spherical severely PrsA-depleted cells, fluorescence was strongly increased in the entire wall. As compared with the fairly moderate effect of severe PrsA depletion on the cross-linkage index, only 2%, the strong diffuse Van-FL fluorescence is surprising. This result may suggest that peptidoglycan (lipid II) precursors are more abundant in the membrane of PrsA-depleted than non-depleted cells and distributed evenly around whole deformed cells. Stationary-phase cells of the *prsA* null mutant, including the small cocci-like ones (Fig. 4B and

_____

**Table 1.** Muropeptide structures and quantification from peptidoglycan of *B. subtilis* RH2111 and conditional mutant IH7211 ($P_{spac}$-*prsA*) cultured with 8 μM or 1 mM IPTG.

| Peak number[a] | Proposed structures[b,c] | RH2111 | IH7211 8 μM IPTG | IH7211 1 mM IPTG |
|---|---|---|---|---|
| | | | Amount of muropeptides (%)[d] | |
| 1 | ds-tri | 2.65 | 1.26 | 2.63 |
| 2 | ds-tri (deAc) | 0.53 | 1.00 | 0.53 |
| 3 | ds-tri with 1 amidation | 10.58 | 13.39 | 10.84 |
| 4 | ds-tri (deAc) with 1 amidation | 2.07 | 1.70 | 1.70 |
| 5 | ds-tetra | 0.32 | 0.31 | 0.22 |
| 6 | ds-tetra with 1 amidation | 0.22 | 0.33 | 0.26 |
| 7 | ds-di | 1.39 | 1.73 | 1.62 |
| 8 | ds-tetra with 1 amidation | 0.75 | 1.02 | 0.71 |
| 9 | ds-tri-di with 2 amidations | 1.06 | 1.85 | 1.25 |
| 10 | ds-tetra with 2 amidations | 0.85 | 0.83 | 0.82 |
| 11 | ds-tri-tetra with 2 amidations | 0.44 | 0.22 | 0.53 |
| 12 | ds-penta with 1 amidation | 0.39 | 0.95 | 0.33 |
| 13 | ds-tri-ds-tetra with 1 amidation and missing GlcNac | 1.34 | 1.19 | 1.68 |
| 14 | ds-tri-ds-tetra | 1.44 | 0.80 | 1.16 |
| 15 | ds-tetra-tetra with 1 amidation | 0.32 | 0.27 | 0.28 |
| 16 | ds-tri-ds-tetra with 1 amidation | 8.78 | 6.63 | 10.51 |
| 17 | ds-tri-ds-tetra with 2 amidations and missing GlcNac | 0.66 | 1.43 | 0.86 |
| 18 | anhydro ds-tri with 1 amidation | 0.10 | 0.17 | 1.23 |
| | ds-tri-ds-tetra (deAc) with 1 amidation | 2.29 | 0.86 | 1.23 |
| 19 | ds-tri-ds-tetra (deAc) with 1 amidation | 1.15 | 0.69 | 1.06 |
| 20 | ds-tri-ds-tetra with 1 amidation | 1.89 | 1.31 | 2.12 |
| 21 | ds-tri-ds-tetra with 2 amidations | 26.34 | 30.79 | 26.32 |
| 22 | ds-tri-ds-tetra (deAc) with 2 amidations | 7.90 | 6.87 | 6.04 |
| 23 | ds-tri-ds-tetra (deAc) with 2 amidations | 5.26 | 4.81 | 4.37 |
| 24 | ds-tri-ds-tetra (deAc × 2) with 2 amidations | 1.69 | 1.22 | 1.15 |
| 25 | ds-tetra-ds-tetra with 2 amidations | 0.87 | 1.08 | 0.91 |
| 26 | ds-penta-ds-tetra with 1 amidation | 0.49 | 0.59 | 0.38 |
| 27 | ds-penta-ds-tetra with 2 amidations | 1.63 | 1.95 | 1.41 |
| 28 | ds-tetra-ds-tetra with 3 amidations | 0.18 | 0.12 | 0.19 |
| | ds-tri-ds-tetra-ds-tetra with 3 amidations | 0.18 | 0.12 | 0.19 |
| 29 | ds-penta-ds-tetra (deAc) with 2 amidations | 0.68 | 0.83 | 0.54 |
| 30 | ds-tri-ds-tetra +Ac with 2 amidations | 0.51 | 0.20 | 0.59 |
| | ds-tri-ds-tetra-ds-tetra with 2 amidations | 0.51 | 0.20 | 0.59 |
| 31 | ds-tri-ds-tetra-ds-tetra with 2 amidations | 2.20 | 1.38 | 2.95 |
| 32 | ds-tri-ds-tetra-ds-tetra with 3 amidations | 6.01 | 6.01 | 6.67 |
| 33 | ds-tri-ds-tetra-ds-tetra (deAc) with 3 amidations | 3.30 | 2.27 | 2.77 |
| 34 | ds-tri-ds-tetra-ds-tetra (deAc × 2) with 3 amidations | 1.18 | 0.66 | 1.05 |
| 35 | ds-penta-ds-tetra-ds-tetra with 3 amidations | 0.48 | 0.41 | 0.46 |
| 36 | anhydro ds-tri-ds-tetra with 2 amidations | 0.78 | 1.67 | 0.80 |
| 37 | anhydro ds-tri-ds-tetra with 2 amidations | 0.24 | 0.38 | 0.53 |
| | ds-tri-ds-tetra-ds-tetra-ds-tetra with 4 amidations | 0.37 | 0.46 | 0.53 |

**a.** Peak numbers refer to the different peaks separated by RP-HPLC from the muropeptide digest of *B. subtilis* peptidoglycan.
**b.** Proposed structures according to the masses determined by Maldi-Tof and to the identifications by Atrih *et al.* (1999).
**c.** ds, disaccharide (GlcNAc-MurNAc); tri, tripeptide (L-Ala-D-Glu-mDAP); tetra, tetrapeptide (L-Ala-D-Glu-mDAP-D-Ala); penta, pentapeptide (L-Ala-D-Glu-mDAP-D-Ala-D-Ala); mDAP, meso-diaminopimelic acid; deAc, deacetylated; +Ac, acetylated.
**d.** Percentage of each peak was calculated as the ratio of the peak area over the sum of areas of all the peaks identified in the table. The values presented are the mean values obtained on three independent experiments for each strain.

C), were less fluorescent than the deformed exponential-phase cells.

### PrsA-dependent membrane proteins as revealed by membrane proteome analysis

The instability of several PBPs in the absence of PrsA suggests that PrsA facilitates the folding of membrane proteins which have a large 'periplasmic' hydrophilic domain(s). In order to identify the PrsA-dependent mem-

brane proteins, membrane proteome analysis of IH7211 (P*spac-prsA*) was performed by using metabolic labeling of PrsA-depleted and non-depleted cells with $^{14}$N/$^{15}$N and mass spectrometric quantification of the changes in relative amounts of peptides derived from the labeled proteins (see *Experimental procedures* for methodological details). The proteome analysis consisted of two biological replicates (M1 and M2a) and a technical replicate of the experiment M2a (M2b). With the minimum criterion of two identified peptides, 192 different proteins could be quan-

**Table 2.** Relative amounts of monomers, dimers, trimers and tetramers and of the different peptide side-chains with a free carboxyl group (acceptor chain) in the peptidoglycan of the different strains.

| Muropeptides[a] | Amount of muropeptides (%)[b] | | |
| --- | --- | --- | --- |
| | RH2111 | IH7211 8 μM IPTG | IH7211 1 mM IPTG |
| Monomers | 19.8 | 22.7 | 20.9 |
| Dimers | 65.9 | 65.8 | 63.9 |
| Trimers | 13.9 | 11.1 | 14.7 |
| Tetramers | 0.4 | 0.5 | 0.5 |
| Cross-linkage degree | 42.5 (0.9) | 40.6 (0.7) | 42.1 (0.7) |
| Dipeptide | 1.4 (0.1) | 1.7 (0.2) | 1.6 (0.2) |
| Tripeptide | 51.4 (0.5) | 51.7 (0.4) | 51.9 (0.2) |
| Tetrapeptide | 2.8 (0.5) | 3.2 (0.1) | 2.7 (0.3) |
| Pentapeptide | 2.0 (0.1) | 2.8 (0.1) | 1.7 (0.1) |

a. The cross-linkage degree and the relative amounts of the different peptide side-chains were calculated according to Glauner *et al.* (1988).
b. Percentages presented are the mean values of three independent determinations for each strain according to muropeptide structure determination presented in Table 1. Values in brackets are standard deviations.

tified which were predicted to contain transmembrane helices or signal peptides. A majority of them were predicted to be membrane-localized proteins (157–171, depending on the algorithm used), either integral membrane proteins or lipoproteins, but also some established or likely exported/cell wall-associated proteins were identified (the proteins and their predicted subcellular localizations are listed in Table S1).

In PrsA-depleted cells, i.e. P*spac-prsA* induced with 2 μM IPTG, the level of PrsA protein was about 10% of that in non-depleted cells, i.e. P*spac-prsA* induced with 1 mM IPTG (Table 3). The levels of PBP2b, PBP3 and PBP4 were decreased (log2 ratio < −1) in PrsA-depleted

cell membranes (Table 3), consistent with the Bocillin-FL labelling and immunoblotting results above. Their amount was 30–40% of that in non-depleted cells and the decrease was observed in the both biological replicates as well as in the technical replicate. Furthermore, the level of PBP2a was decreased in M2a (Table 3). Peptides derived from PBP1a/b, PBP2c, PBP5 and PbpX were also quantified, but their levels were not significantly changed by PrsA depletion. Only a few other proteins were quantified decreased, most notably the bacteriophage SPP1 adsorption protein YueB, which was strongly decreased in all three replicates, being at the lowest only 15% of the level of the PrsA-expressing cells in M2a (Table 3). Transmembrane topology prediction of YueB (TMHMM) suggests that it has membrane-spanning segments close to the N- and C-termini and between them a large 'periplasmic' domain (about 900 amino acids). The reduced amount of YueB protein at a low level of PrsA suggests that this domain folds in a PrsA-dependent manner. The levels of OxaA2, ComE and YvrA proteins were slightly decreased in one or two of the replicates. The proteome analysis also suggested that some proteins were more abundant in PrsA-depleted membranes than in non-depleted ones. Among them are the WapA wall-associated protein and the LytA membrane protein, which were detected at about twofold higher levels in two of the replicates (Table S1).

*PrsA is localized in spots with a spiral-like pattern of organization along the cell membrane*

In order to find out whether the PrsA lipoprotein is distributed evenly around the cell membrane or in an uneven manner like MreC and several PBPs, we constructed the *B. subtilis* IH8478 strain which expresses PrsA modified with a C-terminal Myc-tag. This strain was subjected to a procedure (see *Experimental procedures*) in which PrsA-
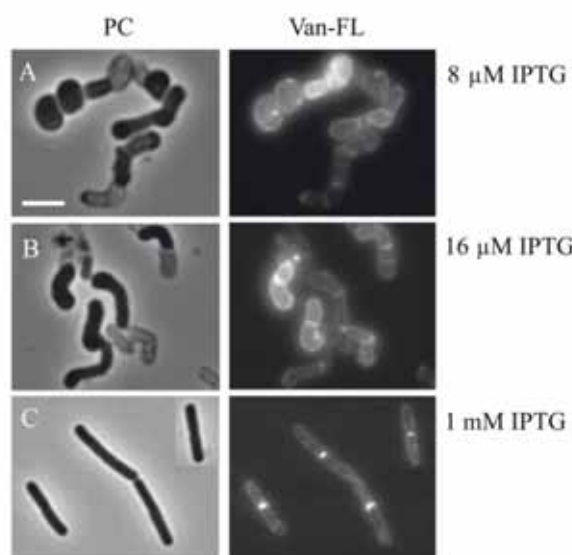


PC                    Van-FL

A                                          8 μM IPTG

B                                          16 μM IPTG

C                                          1 mM IPTG

**Fig. 5.** Van-FL staining of *B. subtilis* strain IH7211 (P*spac-prsA*). Left panels – phase contrast images, right panels – Van-FL staining. P*spac-prsA* expression was induced with IPTG as indicated. Scale bar represents 6 μm (the same for all the images).

Penicillin-binding protein folding is dependent on the PrsA peptidyl-prolyl *cis-trans* isomerase in *Bacillus subtilis*

_____

**Table 3.** Decreased levels of membrane proteins with a large 'periplasmic' domain in PrsA-depleted cells.

| Gene names | Protein Names | M1 | | | | M2a | | | | M2b | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Weighted average ratio (given by CENSUS) | Normalized Ratio | log2[a] RATIO | Number of quantified Peptides | Weighted average ratio (given by CENSUS) | Normalized Ratio | log2[a] RATIO | Number of quantified Peptides | Weighted average ratio (given by CENSUS) | Normalized Ratio | log2[a] RATIO | Number of quantified Peptides |
| prsA | Foldase protein PrsA | 0.12 | 0.12 | −3.01 | 12 | 0.16 | 0.13 | −2.90 | 11 | 0.15 | 0.12 | −3.11 | 6 |
| pbpB | Penicillin-binding protein 2b | 0.32 | 0.32 | −1.64 | 5 | 0.41 | 0.34 | −1.55 | 10 | 0.39 | 0.30 | −1.73 | 7 |
| pbpC | Penicillin-binding protein 3 | 0.40 | 0.41 | −1.30 | 8 | 0.41 | 0.34 | −1.55 | 17 | 0.41 | 0.32 | −1.66 | 10 |
| yueB | Bacteriophage SPP1 adsorption protein YueB | 0.44 | 0.44 | −1.18 | 2 | 0.18 | 0.15 | −2.73 | 12 | 0.26 | 0.20 | −2.31 | 10 |
| pbpD | Penicillin-binding protein 4 | 0.46 | 0.47 | −1.09 | 3 | 0.48 | 0.40 | −1.32 | 6 | 0.49 | 0.38 | −1.40 | 6 |
| comEA | ComE operon protein 1 | 0.83 | 0.84 | −0.25 | 3 | 0.53 | 0.44 | −1.18 | 2 | 0.94 | 0.73 | −0.46 | 2 |
| oxaA2 | Membrane protein OxaA 2 | 1.79 | 1.81 | 0.85 | 2 | 0.45 | 0.38 | −1.41 | 4 | 0.48 | 0.37 | −1.43 | 4 |
| yvrA | YvrA | 0.75 | 0.76 | −0.39 | 5 | 0.69 | 0.58 | −0.80 | 6 | 0.60 | 0.46 | −1.11 | 3 |
| pbpA | Penicillin-binding protein 2a | | | | | 0.41 | 0.34 | −1.55 | 2 | 0.75 | 0.58 | −0.78 | 3 |

a. The shading shows more than 2-fold decreased levels.

Myc was stained with anti-c-Myc antibodies, secondary antibodies conjugated with biotin and ExtrAvidin conjugated with Cy3. The stained PrsA-Myc was visualized by fluorescence microscopy. The localization pattern of PrsA-Myc was determined both in cells from the exponential and stationary phase of growth. To show specific binding of the antibodies, *B. subtilis* strain 168 (RH2111) was used simultaneously as a negative control. Indeed, no Cy3 signal was detected in the control strain (see Fig. S5).

The fluorescence images showed that PrsA is not distributed evenly in the membrane but it is localized in distinct spots that are lined up in spirals (Fig. 6). This pattern is stable throughout vegetative growth until stationary phase. However, the spiral structures are better resolved in exponentially growing cells than in stationary phase cells.

*PrsA is an oligomeric protein*

It has been shown that MreC is a dimeric protein and this structure may be important for its putative function as a scaffold for recruiting PBPs and spatially organizing the lateral cell wall synthesis (van den Ent *et al.*, 2006). We used formaldehyde cross-linking of PrsA and PrsA-Myc in whole cells and immunoblotting with anti-c-Myc and anti-PrsA antibodies to study whether PrsA is a monomeric or dimeric/multimeric protein. The non-random localization pattern suggests that PrsA may be associated with some other protein(s) which has the same or similar localization pattern, e.g. MreC or some of the PBPs. The cross-linking approach might also reveal such interactions.

The cross-linking of cells expressing either wild-type PrsA or PrsA-Myc revealed, in addition to the 33 kDa PrsA or PrsA-Myc monomers, two other PrsA-containing bands of higher molecular weights, one migrating (in SDS-PAGE) at approximately 65 kDa and the other one slightly above it (−68 kDa) (Fig. 7). The 65 kDa form was very heat resistant; it did not disappear at heating for 30 min at 95°C and analysis in SDS-PAGE in the presence or absence of dithiothreitol (Fig. 7 and data not shown). A small amount of this protein could also be detected in non-cross-linked cells. The molecular weight suggests that it might be a PrsA dimer. Consistently, *B. subtilis* PrsA expressed in and purified from *E. coli* contained this same form and its amount increased by cross-linking of the purified PrsA. The 68 kDa band was detected only in lanes containing cross-linked *B. subtilis* cell or PrsA protein (from *E. coli*) samples and it disappeared by heating at 95°C. It most probably contains oligomeric PrsA which migrated clearly faster than what is expected for PrsA trimers or tetramers. A few very weak bands of high-molecular-weight complexes (> 68 kDa) were also detected, but further studies are needed to elucidate
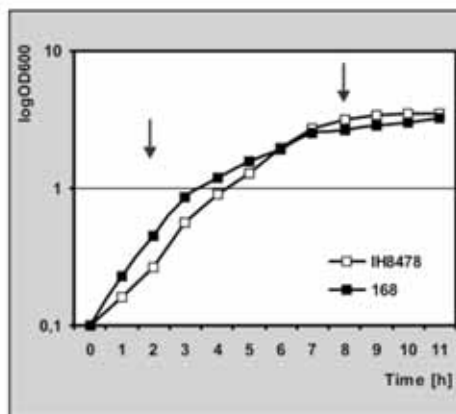
**Fig. 6.** Immunolocalization of PrsA-Myc in *Bacillus subtilis* IH8478 strain in exponential (A–C) and stationary phase (D–F). Immunostaining has been described in *Experimental procedures*. Scale bar – 2 μm. The upper panel shows growth curves of *B. subtilis* IH8478 and *B. subtilis* 168, the parental strain which does not express PrsA-Myc. Arrows indicate time points when samples were collected for immunofluorescence microscopy. A. and D. Phase contrast pictures. B. and E. Fluorescence pictures of Cy3-stained cells. C. and F. Fluorescence pictures after deconvolution.
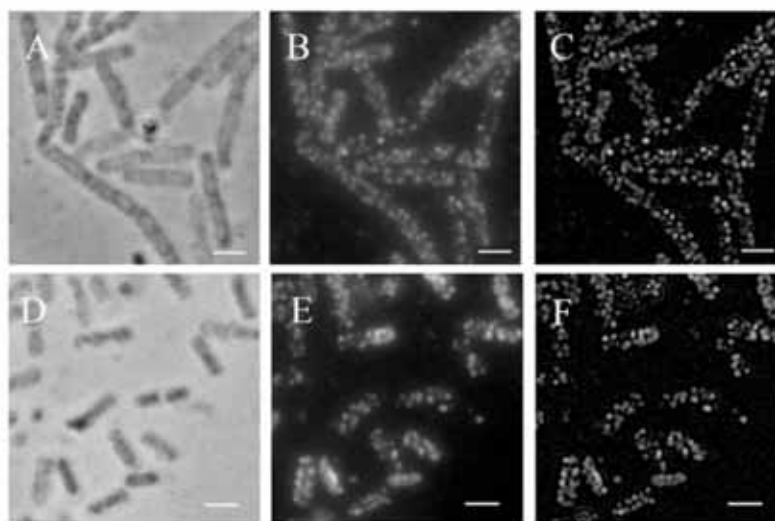
whether they contain PrsA associated with some other components of the membrane/wall.

Small-angle X-ray scattering (SAXS) was another method to elucidate the multimeric structure of PrsA (Rodgers *et al.*, 1996). *Escherichia coli*-produced PrsA protein was subjected to SAXS and the radius of gyration ($Rg$) was determined. PrsA had the Rg of 27–29 Å (Fig. S6), which further suggests that PrsA forms oligomers.

## Discussion

PrsA peptidyl-prolyl *cis-trans* isomerase has an essential role in extracytoplasmic protein folding in rod-shaped bacteria. The localization of the enzyme domain at the membrane-cell wall interface suggests that PrsA may assist the folding of membrane proteins which have large functional domains on the outer surface of the membrane. In this study we used several methodological approaches to identify membrane proteins which are dependent on

PrsA for folding with the emphasis on identifying the PrsA-dependent component(s) that is (are) involved in cell shape determination and/or cell wall synthesis.

PrsA is most probably involved in lateral cell wall biosynthesis of *B. subtilis* as suggested by the morphological changes in the absence of PrsA. Severely PrsA-depleted cells are spherical in shape and pearl necklace-like round-cell chains are formed, but a high concentration of magnesium probably stabilizes peptidoglycan enabling the bacterium to maintain its rod shape (thick) at very low PrsA levels. Consistently also small cocci-like cells were formed in stationary phase cultures of the *prsA* null mutant. In the wild type-like rods of the *prsA* null mutant, the lateral wall synthesis may have been restored by some compensation mechanism for instance a secondary suppressor mutation.

Our results suggest that PrsA is required for lateral cell wall biosynthesis as the folding and stability of those PBPs which are involved in the lateral wall synthesis are
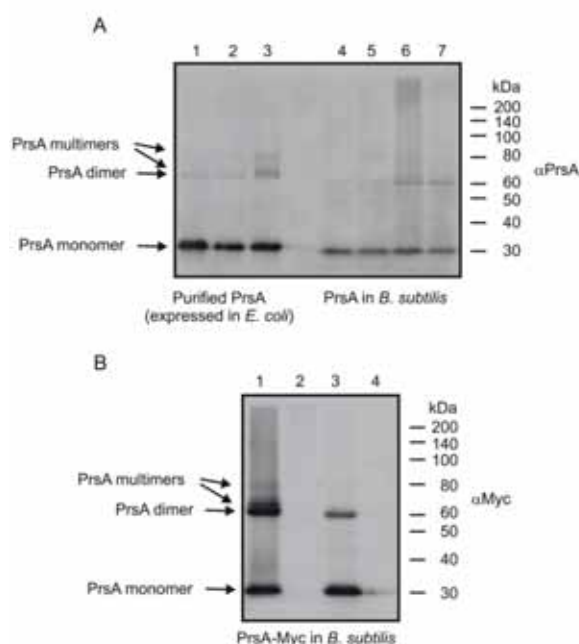
_____

**Fig. 7.** PrsA dimers and oligomers as revealed by formaldehyde cross-linking of whole *B. subtilis* cells and purified PrsA. Cross-linked PrsA (A) or PrsA-Myc (B) complexes were detected by immunoblotting.
A. Non-lipomodified PrsA expressed in the cytoplasm of *E. coli* and purified (lanes 1–3) and lipomodified PrsA in whole cells of *B. subtilis* RH2111 (lanes 4–7). Heating of non-cross-linked (lanes 1, 2, 4 and 5) or cross-linked (lanes 3, 6 and 7) samples in SDS-PAGE sample buffer at 37°C for 10 min (lanes 1, 3, 4 and 6) or at 95°C for 30 min in the presence of 10 mM dithiothreitol (lanes 2, 5 and 7).
B. Lipomodified PrsA-Myc in *B. subtilis* IH8478 cross-linked and heated at 37°C for 10 min (lane 1) or at 95°C for 30 min (lane 3). Corresponding samples from *B. subtilis* RH2111, the parental strain not expressing PrsA-Myc, were used as negative controls (lanes 2 and 4 respectively).

dependent on PrsA. Levels of several high-molecular-weight PBPs (PBP2a, PBP2b, PBP3 and PBP4) were significantly decreased in PrsA-depleted cells, in contrast to some other proteins (MreC and PBP1a/b) located in the same compartment. We could also demonstrate that misfolding of PBP2a occurred in the absence of PrsA and the misfolded PBP2a was degraded in a manner dependent on the CssRS two-component system. It seems that PrsA is required either directly or indirectly for PBP2a folding, and most probably also for the folding of the other PrsA-dependent PBPs, and the primary reason for the growth inhibition and cell wall synthesis defect is probably insufficient amount of active PBPs. Most likely HtrA and HtrB proteases are involved in the degradation of misfolded PBP2a, as CssRS is a dedicated regulator of their expression (Hyyrylainen *et al.*, 2005). HtrA/DegP-type proteases have in addition to the proteolytic activity a chaperone activity which catalyses protein folding (Spiess *et al.*,

1999; Antelmann *et al.*, 2003) and this may explain the effect of CssRS on the folding of PBP5. Our results suggest that PrsA and CssRS have an overlapping and redundant role in the folding of PBP5, but further work is needed to characterize the interplay of these components at the molecular level in detail.

Severe PrsA depletion decreased the cross-linkage degree but only by 2%. This is surprising as cells are spherical or large rods in severely PrsA-depleted cultures. PBPs that normally synthesize the division septum, particularly PBP1a/b, which was stable in PrsA-depleted cells, may have taken a larger role in the cross-linkage/synthesis of the whole cell wall peptidoglycan. This would explain the relatively moderate decrease in the cross-linkage index. The peptidoglycan structure is only moderately impaired, but the bacterium is spherical in shape. In contrast to wild-type cells, Van-FL stained strongly Δ*prsA* and PrsA-depleted exponential-phase cells and the fluorescence was fairly uniformly distributed around the whole cell membrane. The increased number of remaining pentapeptide side-chains and their even distribution in the wall might explain the increased Van-FL staining. However, the fairly moderate effect of PrsA-depletion on the peptidoglycan structure may not be consistent with this explanation. An alternative hypothesis might be that the level of membrane-bound peptidoglycan precursors was increased in these deformed cells and that the precursors either moved freely in the membrane or were translocated uniformly across the membrane.

Does PrsA assist directly PBP folding or is the effect indirect i.e. via catalysis of folding of a third component which influences PBP folding? We tried to demonstrate *in vitro* with purified PBP2a and PrsA a direct role for PrsA in the catalysis of PBP2a folding by measuring the folding kinetics of denatured PBP2a with the binding of Bocillin-FL, but it did not succeed. The result was very similar as previously with AmyQ α-amylase: the denatured protein folded rapidly to correct enzymatically active conformation and independently of PrsA. In the compartment at the membrane-wall interface ('periplasm'), AmyQ folding is strongly dependent on PrsA, but in protoplasts in the absence of the wall AmyQ folds independently of PrsA (Hyyryläinen *et al.*, 2001; Vitikainen *et al.*, 2001; Wahlstrom *et al.*, 2003). Thus, the folding assistance requirement is dependent on the cell wall environment. The situation with PBP2a and other PrsA-dependent PBPs is probably very similar. PrsA assistance is needed for folding in the 'periplasm' only. PBPs, and also AmyQ, are fairly large proteins and have numerous proline residues. Therefore, it is likely that direct assistance of chaperones and foldases is needed for their folding in the cell wall environment. Why is then PBP1a/b independent of PrsA? It is a larger protein than the other PBPs and AmyQ. One explanation could be that in addition to PrsA and CssRS

(HtrA/B), which both are involved in PBP5 folding, a third chaperone catalyses PBP1a/b folding.

A wider search for PrsA-dependent proteins was performed by analysing PrsA-dependent changes in membrane proteome by mass-spectrometry. The proteome analysis suggested that PPB2a, PBP2b, PBP3 and PBP4 are the main PrsA-dependent proteins in the membrane. In addition to them, only a few other membrane proteins exhibited decreased stability in PrsA-depleted cells, most significantly the SSP1 receptor YueB. This fairly small set of PrsA-dependent proteins supports a direct rather than an indirect effect of PrsA on PBP folding. One can envision an indirect effect if PrsA determines the isomeric state of a prolyl bond in a third component, e.g. MreC, the stability of both these conformers is similar and the position of the prolyl switch affects the folding and stability of PBPs.

Despite of the rod cell shape, *Corynebacteria* such as *C. glutamicum* and *Corynebacterium diphtheriae*, which belong to the Actinobacteria group of Gram-positive bacteria, do not possess PrsA (Kalinowski *et al.*, 2003). Obviously they do not need a PrsA-like foldase/chaperone to synthesize the lateral wall and maintain the rod cell shape. The PrsA-independency may be due to the different mode of lateral wall synthesis in these bacteria as compared with *B. subtilis* and most probably other rod-shaped Firmicutes; it has been shown that in *C. glutamicum* peptidoglycan is incorporated into the wall via cell poles in a manner dependent on the DivIVA protein (Daniel and Errington, 2003; Letek *et al.*, 2008). PrsA is also a dispensable protein in cocci (Drouault *et al.*, 2002; Ma *et al.*, 2006). Because in cocci peptidoglycan is assembled at the division septum and the hemispherical poles derived from it (Pinho and Errington, 2003), the same reason may explain why PrsA is dispensable in them.

In *B. subtilis*, the DivIB, DivIC and FtsL cell division proteins and the PBP 2b form a protein complex that assembles in an interdependent manner (Daniel *et al.*, 2006). It has also been shown that FtsL and DivIC are unstable in cells depleted of PBP2b (Daniel *et al.*, 2006). FtsL is a particularly unstable protein in the absence of the interacting proteins (Bramkamp *et al.*, 2006). PBP2b was one of those PBPs which were degraded in the absence of PrsA and this could affect the stability of the whole complex. Therefore, we studied the effect of PrsA depletion on the stability of the divisome proteins by using N-terminal GFP fusions of these proteins. Surprisingly, their stability was independent of the expression of PrsA in this assay. The membrane proteome analysis of PrsA-depleted cells also showed a clear decrease in the level of PBP2b but no effect on the level of DivIB; FtsL and DivIC could not be detected. An NMR study has shown that the DivIB protein of *Geobacillus stearothermophilus* is in two distinctly different conformations depending on the *cis-trans* isomerization of a Tyr-Pro peptide bond in the extra-

cytoplasmic β-domain and it was speculated that the *cis-trans* isomerization may modulate the assembly of the divisome complex (Robson and King, 2006). Our finding that PrsA depletion has no effect on the stability of DivIB, DivIC and FtsL does not support the hypothesis about a regulatory role of *cis-trans* isomerization in the divisome complex assembly in *B. subtilis*.

We also determined the localization of PrsA in the membrane by taking advantage of a *B. subtilis* strain that expresses Myc-tagged PrsA. Because the *prsA-myc* fusion gene is present as a single copy in the chromosome and under the control of the native *prsA* promoter, artefacts due to overproduction were avoided. The results showed that PrsA is not randomly distributed in the membrane. It (PrsA-Myc) was localized to the lateral cell membrane in which it formed distinct spots organized in a helical pattern. To our knowledge this is the first lipoprotein which has been shown to have the helical organization pattern. Still little is known about the mechanism underlying helical distribution of proteins along the membrane. The helical organization of PrsA raises the question whether it is associated with any of the other proteins with the similar organization pattern including the cytoskeleton proteins and PBPs. The helical pattern might be dependent on them. It has been shown that a cytoskeleton protein (MreBH) can determine the helical organization of a protein (LytE) on the extacytoplasmic side of the membrane (Carballido-Lopez *et al.*, 2006). We also demonstrated that PrsA forms dimers/oligomers, but the functional significance of the dimeric/oligomeric structure is still unclear.

PrsA foldase/chaperone catalyses post-translocation folding of exported proteins (Jacobs *et al.*, 1993; Hyyryläinen *et al.*, 2001; Vitikainen *et al.*, 2004). It has been shown that overexpression of PrsA over the wild-type level enhances secretion of some extracellular proteins particularly α-amylases of *Bacillus* sp. from industrial Gram-positive bacteria (Kontinen and Sarvas, 1993; Vitikainen *et al.*, 2001, 2005; Williams *et al.*, 2003; Lindholm *et al.*, 2006). Therefore, PrsA is an important tool for increasing yields in industrial protein production especially thermoresistant α-amylases for various applications needing degradation of starch to smaller sugars, including bioethanol production. Now we have shown that PrsA also has a housekeeping role in the cell. It is required directly or indirectly for PBP folding and lateral cell wall biosynthesis. A class of most important current antibiotics, β-lactams, exerts the antimicrobial effect by inhibiting PBPs. Our results suggest that inhibiting PrsA might be an alternative and additive way to inhibit cell wall biosynthesis of pathogenic rod-shaped bacteria and treat infectious diseases. PrsA-targeted antimicrobials could also inhibit bacterial pathogenesis by inhibiting secretion of critical virulence factors for instance listeriolysin O and broad-range phospholipase C in *Listeria monocytogenes* (Alonzo *et al.*,

Penicillin-binding protein folding is dependent on the PrsA peptidyl-prolyl *cis-trans* isomerase in *Bacillus subtilis*
_____

PrsA PPtase for penicillin-binding protein folding   121

**Table 4.** *B. subtilis* strains used in this study.

| Strain[a] | Relevant genotype | Reference |
|---|---|---|
| IH7211 | 168 prsA::pKTH3384 Pspac-prsA | Vitikainen et al., 2001 |
| IH8266 | 168 prsA::pKTH3384 Pspac-prsA cssS::spec | This study |
| IH8435 | 3417 prsA::pKTH3384 Pspac-prsA | This study |
| IH8437 | 3103 prsA::pKTH3384 Pspac-prsA | This study |
| IH8441 | 2083 prsA::pKTH3384 Pspac-prsA | This study |
| IH8445 | JAH66 prsA::pKTH3384 Pspac-prsA | This study |
| IH8456 | 168 ΔpbpF::ery | This study |
| IH8458 | 168 dacA::cat | This study |
| IH8460 | 168 ΔpbpD::ery | This study |
| IH8462 | 168 ΔponA::spec | This study |
| IH8464 | 168 pbpC::spec | This study |
| IH8471 | 168 divIC::pKTH3806 Pxyl-gfp-divIC | This study |
| IH8473 | 168 divIB::pKTH3814 Pxyl-gfp-divIB | This study |
| IH8478 | 168 prsA::pMUTIN-cMyc prsA-myc | This study |
| IH8480 | IH8471 prsA::pKTH3384 Pspac-prsA | This study |
| IH8483 | IH8473 prsA::pKTH3384 Pspac-prsA | This study |
| IH8999 | 168 pbpA::ery | This study |
| IH9003 | 168 pbpA::pKTH3828 pbpA-myc | This study |
| IH9013 | IH9003 cssS::spec | This study |
| IH9016 | IH9013 prsA::pKTH3384 Pspac-prsA | This study |
| IH9024 | 168 ΔprsA | This study |
| IH9025 | 168 dacA::pKTH3831 dac-myc | This study |
| IH9027 | IH9003 prsA::pKTH3384 Pspac-prsA | This study |
| IH9028 | IH9025 prsA::pKTH3384 Pspac-prsA | This study |
| IH9029 | IH9028 cssS::spec | This study |
| 2083 (RH2225) | 168 ponA::pSG1492 Pxyl-gfp-ponA | Scheffers et al., 2004 |
| 3103 (RH2227) | 168 pbpA::pSG5043 Pxyl-gfp-pbpA | Scheffers et al., 2004 |
| 3417 (RH2224) | 168 mreC::pSG5276 Pxyl-gfp-mreC | Leaver and Errington, 2005 |
| DPVB207 (RH2223) | PS832 ΔpbpH::spec pbpA::ery amyE::Pxyl-pbpH | Wei et al., 2003 |
| PS1869 (RH2234) | PS832 ΔpbpF::ery | McPherson et al., 2001 |
| PS1900 (RH2235) | PS832 dacA::cat | Popham et al., 1996 |
| PS2022 (RH2236) | PS832 ΔpbpD::ery | Popham and Setlow, 1996 |
| PS2062 (RH2237) | PS832 ΔponA::spec | McPherson et al., 2001 |
| PS2328 (RH2238) | PS832 pbpC::spec | Wei et al., 2003 |
| PS832 | Wild type, Trp⁺ revertant of 168 | Wei et al., 2003 |
| 168 (RH2111) | trpC2 | Kunst et al., 1997 |
| JAH66 (RH2229) | 1012 amyE::gfp-ftsL | Heinrich et al., 2008 |

**a.** In parenthesis are shown the strain codes in the collection of THL.

2009; Zemansky *et al.*, 2009). Because human and other eukaryotic cells also have functionally important parvulin-type PPIases, PrsA inhibitors should be PrsA specific. The designing of PrsA inhibitors could take advantage of structural and mechanistic differences between bacterial and eukaryotic parvulins (Heikkinen *et al.*, 2009). We can envision that such PrsA inhibitors could either inhibit specifically the PPIase activity or interfere with functions of the N- or C-terminal domains, which may be chaperone or interaction domains. Alternatively, eukaryotic cell membrane could be impermeable to PrsA inhibitors.

## Experimental procedures

### Bacterial strains, plasmids and growth conditions

*Bacillus subtilis* strains and plasmids used in this study are listed in Table 4 and Table S2 respectively. *Escherichia coli* DH5α was used for cloning. Bacteria were grown in LB or TY medium and on corresponding agar plates at 37°C, but also some other culture media were used for particular experiments. The metabolic labelling for membrane proteome analysis was performed with cells grown in a synthetic minimal medium (BMM) (Stulke *et al.*, 1993). Cultivations were performed in Antibiotic Medium 3 when the effect of magnesium on PrsA-depleted cells and Van-FL staining of peptidoglycan synthesis in *prsA* null mutant were studied. The growth media were supplemented with appropriate antibiotics when needed to maintain plasmids and chromosomal plasmid integrants in the cells. The antibiotics (and their concentrations) used in the cultivations were ampicillin (100 μg ml⁻¹), erythromycin (0.5, 1 or 100 μg ml⁻¹), lincomycin (12.5 μg ml⁻¹), chloramphenicol (5 μg ml⁻¹) and spectinomycin (100 μg ml⁻¹). The expression of Pspac-prsA was induced with 1 mM (full induction) or with 0, 2, 8, 16 or 24 μM (PrsA depletion) IPTG. For the PrsA depletion experiments, bacteria were taken from fresh agar plates, washed several times in LB medium and then used to inoculate the cultures. The expression from Pxyl was induced with 0.5% xylose.

## Plasmid constructions

To construct pKTH3805, a fragment of the coding region of *prsA* (bp 633–938) was PCR-amplified with the primers prsAmyc-fw (5′cacaggtaccatggacgaaacattcagcaaag) and prsAmyc-rv (5′cacacggccgtttagaattgcttgaagatgaagaagtg). The amplified fragment was inserted into the pGEM-T-Easy vector (Promega). The recombinant plasmid was digested with KpnI and EagI, the released 'prsA fragment was ligated with pMUTIN-cMyc (Kaltwasser *et al.*, 2002) cut with the same restriction enzymes and the ligation mixture was used to transform competent *E. coli* DH5α cells. The obtained plasmid carrying the 'prsA fragment (pKTH3805) was used to transform *B. subtilis* RH2111 (168). The Campbell-type integration of pKTH3805 into the *prsA* gene resulted in the formation of *prsA-myc*, which encodes PrsA bearing a C-terminal Myc-tag. The chromosomal integration of pKTH3805 replaced the wild-type *prsA* with *prsA-myc*. The viability of the resulted strain *B. subtilis* IH8478 indicates that PrsA-Myc is functional.

   *Bacillus subtilis* IH9003 expressing *pbpA-myc* was constructed by transforming 168 with plasmid pKTH3828. This plasmid was constructed by PCR-amplifying a 'pbpA-myc fragment with the primers pbpA-fw (5′cacaggatccatagcaaatg-gtggctaccgc) and pbpA-rv (cacagtcgacttacagatcttcttcgct-gatcagtttctgttcgttatcagaagacgttgtgtttctgc) and inserting the fragment between the BamHI and SalI sites of the pSG4902 plasmid (Scheffers *et al.*, 2004). The Campbell-type integration of pKTH3828 into the *pbpA* gene resulted in the formation of *pbpA-myc*.

   *Bacillus subtilis* IH9025 expressing *dacA-myc* was constructed in a similar manner as *B. subtilis* IH9003. The 'dacA-myc fragment in pKTH3831 was PCR-amplified with the primers dacA-fw (5′cacagtcgacggaacagctgaacgcaacg) and dacA-rv (5′cacagaattcttacagatcttcttcgctgatcagtttctgttcaaacca gccggtaccgtatc) and inserted between the SalI and EcoRI sites of pSG4902.

   To construct pKTH3814, a fragment of the *divIB* gene (bp 295–633) was PCR amplified with the primers divIB-fw (5′cacactcgaggtcatgaacccgggtcaagacc) and divIB-rv (5′cacagaattcaggaagcgatttgctgatctcc). The amplified fragment was inserted between the XhoI and EcoRI sites of the pSG4902 plasmid (Scheffers *et al.*, 2004). The resulted pKTH3814 plasmid was used to transform *B. subtilis* RH2111. pKTH3814, upon integration into the *B. subtilis* chromosome, disrupted the native *divIB* gene and created a *gfp-divIB* fusion (*B. subtilis* IH8473), which is expressed from the xylose-inducible P*xyl* promoter.

   *Bacillus subtilis* IH8471 expressing *gfp-divIC* was constructed in a similar manner as *B. subtilis* IH8473. A fragment of the *divIC* gene (bp 455–755) was PCR amplified with the primers divIC-fw (5′cacaggatcccctttgaattttccagggaacga) and divIC-rv (5′cacactcgaggacgtaatcctcatccttcaatttg). The 'divIC fragment was cloned in the pGEM-T-Easy vector. The fragment was released by digesting with BamHI and XhoI, and inserted into pSG4902. *Bacillus subtilis* RH2111 was transformed with the obtained plasmid, pKTH3806.

## Construction of the prsA null mutant

Chromosomal DNA was isolated from *B. subtilis* IH7075 (in the culture collection of THL) and used to transform *B. subtilis*

168 (RH2111). In IH7075, the *prsA* gene is deleted and replaced with the *cat* gene from pC194. The IH7075 strain also harbours plasmid pKTH3327, which carries *prsA* under the control of the P*spac* promoter and complements the chromosomal *prsA* deletion. The Δ*prsA* mutation deletes the coding sequence of *prsA*, and 153 bp and 75 bp of the up- and downstream regions respectively. The *prsA* null mutants of 168 were selected on Antibiotic Medium 3 –plates containing 20 mM MgCl$_2$ and chloramphenicol (5 μg ml$^{-1}$).

## SDS-PAGE and immunoblot analysis of PrsA, MreC, PBP2a and GFP-fusion proteins

Whole cell samples were prepared from cultures in the late exponential (Klett100 + 1 h) or early stationary (Klett100 + 3 h) phase of growth. Cells were harvested by centrifugation from 1 ml of culture, resuspended in 50 μl of protoplast buffer (20 mM potassium phosphate pH7.5, 15 mM MgCl$_2$, 20% sucrose and 1 mg ml$^{-1}$ lysozyme) and incubated at 37°C for 20 min. Then 50 μl of 2×Laemmli sample buffer was added and the samples were boiled for 10 min at 100°C. Proteins were separated in SDS-PAGE and blotted onto a PVDF nylon filter (Immobilon-P transfer membrane; Millipore). The immunodetection of PrsA, PrsA-Myc, MreC, PBP2a and GFP-fusion proteins was performed using specific antibodies against these proteins (or Myc-tag) and goat anti-rabbit IgG (H + L)-HRP conjugate (Bio-Rad). ECL reactions were performed using an Immun-Star WesternC kit according to the instructions of the manufacturer (Bio-Rad) and proteins were visualized and quantified with a Fluo-rChem HD2 imager and AlphaEase FC software (AlphaInnotech).

## Isolation of bacterial cell membranes and Bocillin-FL labelling of penicillin-binding proteins

Bacterial cell membranes were prepared from 500 ml of culture in LB medium. Cells were harvested at the cell density of Klett100 by centrifugation, washed once with potassium phosphate buffer (20 mM potassium phosphate pH 7.5, 140 mM NaCl) and resuspended in the washing buffer. The cells were disrupted by passing the suspension through a French press cell twice at 20 000 lb in$^{-2}$. The cell lysate was centrifuged at 8000 *g* for 10 min, followed by centrifugation of the supernatant at 217 500 *g* for 1 h. The membrane pellet was resuspended in the phosphate buffer (1 ml). Protein concentrations of the membrane preparations were determined using a 2-D Quant Kit (GE Healthcare) or Micro BCA (Pierce) according to the instructions of the manufacturers.

   In the Bocillin-FL labelling reaction, cell membranes (300 μg protein) were treated with 10 μM Bocillin-FL in a final volume of 100 μl. The reaction mixtures were incubated at 35°C for 30 min, followed by addition of 100 μl 2×Laemmli buffer and incubation at 100°C for 3 min. Samples corresponding 15 μg of protein were separated in 10% SDS-PAGE. Fluorescent PBPs were visualized using a fluoroimager Typhoon 9400 (Amersham Biosciences) with excitation at 488 nm and emission at 520 nm. Levels of active Bocillin-FL-labelled PBP2a were quantified with the ImageQuantTL software (Amersham Biosciences).

Penicillin-binding protein folding is dependent on the PrsA peptidyl-prolyl *cis-trans* isomerase in *Bacillus subtilis*
_____

PrsA PPIase for penicillin-binding protein folding   123

## Peptidoglycan extraction and muropeptide analysis

*Bacillus subtilis* cells were grown in LB medium, a volume corresponding $V \times OD = 300$ was taken in the late stationary phase (Klett100 + 6 h) and the cells were harvested by centrifugation. Peptidoglycan extraction and muropeptide analysis were performed as described previously for *L. lactis* (Courtin *et al.*, 2006). Purified peptidoglycan was digested with mutanolysin (muramidase from *Streptomyces globisporus*; Sigma-Aldrich). Soluble muropeptides were reduced with sodium borohydride as described (Atrih *et al.*, 1999) and separated on a Hypersil-100 column (C18, 250 by 4.6 mm, 5 µm; Thermo Finnigan) at 50°C by RP-HPLC. The column was eluted with 10 mM ammonium phosphate pH 4.6 and 0.18% sodium azide for 5 min, followed by a linear gradient of methanol (0–20%) for 270 min as described previously (Courtin *et al.*, 2006). Samples of eluted muropeptides were analysed by MALDI-TOF mass spectrometry (Voyager DE STR, Applied Biosystems, Framingham, MA) with α-cyano-4-hydrocinnamic acid matrix. For most fractions, 1 µl of sample was sufficient. For less abundant fractions, 100 µl was concentrated on a ZIP-Tip C18 pipette tip (Millipore) and eluted with 1 µl of solvent (50% acetonitrile and 0.15% trifluoroacetic acid) prior to analysis.

The cross-linkage degree was calculated according to Glauner (1988) with the formula: $(1/2 \Sigma \text{ dimers} + 2/3 \Sigma \text{ trimers} + 3/4 \Sigma \text{ tetramers})/ \Sigma$ all muropeptides.

The percentage of muropeptides with a certain side peptide chain (X = di, tri, tetra and penta) with free COOH (donor chain) was calculated according to Glauner (1988).

$$\text{percentage (X)} = \left[ \Sigma \text{ monomers(X)} + \frac{1}{2}\Sigma\text{dimers(X)} + \frac{1}{3}\Sigma\text{trimers(X)} + \frac{1}{4}\Sigma\text{tetramers(X)} \right] / \Sigma$$

all muropeptides.

Statistical analysis of the data was performed with Statgraphics Plus software (Manugistics, Rockville, MD, USA). The results obtained with two strains were compared with a two-sample comparison analysis and a *t*-test.

## Van-FL staining and fluorescence microscopy

*Bacillus subtilis* strains 168 and ΔprsA were grown over night in 10 ml of Antibiotic medium 3 (Difco) suplemented with 20 mM MgCl2 and 5 µg ml⁻¹ chloramphenicol when needed. Over night cultures were diluted in 20 ml of fresh medium to OD600 = 0.1 and grown 24 h. Cells were collected for Van-FL staining at exponential, stationary and late stationary phases. *Bacillus subtilis* strain IH7211 was grown overnight on a TY agar plate supplemented with 1 µg ml⁻¹ erythromycin. Material from plate was suspended in phosphate-buffered saline (PBS) to $OD_{600} = 1.0$ and washed three times with 1 ml of PBS. 20 µl of the suspension was used to inoculate 10 ml of TY containing 1 µg ml⁻¹ erythromycin and IPTG at three different concentrations: 8 µM, 16 µM and 1 mM. At $OD_{600} = 0.6$, samples were collected for Van-FL staining.

A culture of 0.5 ml was incubated 20 min with 1 µg ml⁻¹ fluorescently labelled vancomycin (BODIPY® FL vancomycin, Invitrogen) mixed in 1:1 ratio with unlabelled vancomycin

(Sigma). The cells were spotted on microscope slides (Knittel Gläser, Germany). The Van-FL-stained cells were viewed immediately under a fluorescence microscope (Olympus IX71) equipped with a Cool Snap HQ2 camera (Photometrics). Van-FL fluorescence was visualized with a bandpass 470/40 nm excitation filter and a bandpass 525/50 nm emission filter. Images were analysed using ImageJ (http://rsb.info.nih.gov/nih-image/) and Adobe Photoshop CS2 Version 9.0.

## Electron microscopy

For electron microscopy, cells were fixed with 2.5% glutaraldehyde overnight at 4°C as has been described (Leaver and Errington, 2005). The fixed cells were washed twice in phosphate buffer and then treated with 1% osmium tetroxide for 1 h. The samples were dehydrated with a series of treatments with ethanol and acetone, followed by embedding in Epon resin. The microscopy was performed in the Institute of Biotechnology, University of Helsinki, by using a JEOL 1200EX II electron microscope.

## Quantification of the effect of a minimal PrsA concentration on the membrane proteome of B. subtilis

For metabolic labelling, *B. subtilis* IH7211 (P_spac-prsA) was grown aerobically at 37°C in a synthetic minimal medium (BMM). The medium was either supplemented with ¹⁵N-ammonium sulphate, ¹⁵N-L-tryptophan and ¹⁵N-L-glutamate (Cambridge Isotope Laboratories, Andover, USA) or ¹⁴N-ammonium sulphate, ¹⁴N-L-tryptophan and ¹⁴N-L-glutamate. Cells from overnight cultures grown with 40 µM IPTG were washed with warm BMM and used to inoculate the cultures for the metabolic labelling. In order to realize two different biological states, 2 µM IPTG was used in the cultivation M1 with the light form of nitrogen source and 1 mM IPTG with the heavy one (see below). A biological replicate (M2a) was performed with identical cultures but the labels were switched. Cells were harvested in the exponential phase of growth and equivalent OD-units of bacteria either grown with ¹⁵N-BMM or ¹⁴N-BMM were combined in proportion 1:1 and stored at –20°C.

The combined cells were disrupted on ice using a French press cell, followed by separation of cell debris by centrifugation and extraction of cell membranes according to the method described by Eymann and collaborators (Eymann *et al.*, 2004). Membranes were centrifuged at 100 000 *g* for 60 min at 4°C, followed by purification steps with high salt buffer [20 mM Tris-HCl, pH7.5, 10 mM EDTA, 1 M NaCl, Complete Protease Inhibitor (Roche)], alkaline Na₂CO₃ buffer (100 mM Na₂CO₃-HCl, pH 11, 10 mM EDTA, 100 mM NaCl) and 50 mM TEAB. Finally the purified membranes were solubilized in SDS-PAGE sample buffer containing 20% SDS. Proteins were analysed using GeLC-MSMS analysing the resulting tryptic digests using an UPLC coupled to an LTQ-Orbitrap. In addition to the two biological replicates, a technical replicate, M2b, was performed. For this replicate, the proteins of M2a were separated and analysed with GeLC-MSMS for a second time. The data-analysis and relative quantification of the proteins was performed as described in Supporting Information (Table S1).

### Determination of the localization of PrsA

Single colonies of *B. subtilis* 168 and IH8478 (*prsA-myc*) grown on TY plates were used to inoculate 5 ml of TY liquid medium. Overnight cultures were diluted in 20 ml of TY medium to $OD_{600} = 0.1$ and grown till stationary phase. TY medium was supplemented with erythromycin to a final concentration of 1 µg ml⁻¹ when needed. Samples for immunofluorescence assay were collected 2 h before and 4 h after the transition from the exponential to the stationary growth phase (Fig. 6). Immunofluorescence staining was performed according to the method described by Harry and collaborators (Harry *et al.*, 1995) with the modifications described below.

Cells were fixed and made permeable as follows. 0.5 ml of bacterial culture in TY was mixed with an equal volume of 2×fixative solution containing 2.68% paraformaldehyde and 0.0050% glutaraldehyde and incubated for 15 min at room temperature (21–23°C) and 30 min on ice. After fixation the cells were washed three times in PBS and resuspended in GTE (50 mM glucose, 20 mM Tris-HCl pH 7.5, 10 mM EDTA). A fresh lysozyme solution in GTE was added to an aliquot of cells to a final concentration of 2 mg ml⁻¹ and cells were immediately spotted on multiwell slides (MP Biomedicals LLC) coated with 0.01% poly-L-lysine (Sigma-Aldrich). After 5 min incubation wells were washed with PBS and left to air dry.

For immunostaining, cells were blocked with PBS containing 2% BSA and 0.01% Tween (hereafter referred to as 'blocking solution') for 15 min at room temperature. Next cells were incubated with mouse anti-c-Myc antibodies (Gentaur) diluted 1:1000 in the blocking solution for 1 h at room temperature. After washing the cells 10 times with PBS, secondary anti-mouse biotin-conjugated antibody (Sigma-Aldrich) diluted 1:500 in the blocking solution was added, followed by incubation in the dark at room temperature for 1 h. Cells were washed again 10 times with PBS and incubated 1 h with 1:25 diluted ExtrAvidin Cy3 conjugate (Sigma-Aldrich) at room temperature in the dark. Samples were washed 10 times with PBS and mounted with Vectashield mounting medium (Vector Laboratories). Slides were stored at −20°C.

Sample imaging was performed using a wide-field Zeiss Axioscop50 fluorescence microscope (Carl Zeiss, Oberkochen, Germany) equipped with a Princeton Instruments 1300Y digital camera. Cy3 fluorescence was visualized with a bandpass (546 /12 nm) excitation filter, a 560 nm dichromatic mirror, and a bandpass (575–640 nm) emission filter. Images were analysed using ImageJ (http://rsb.info.nih.gov/nih-image/) and Adobe Photoshop CS2. Wide-field images were corrected for bleaching and unstable illumination using the Huygens Professional deconvolution software by Scientific Volume Imaging (http://www.svi.nl/).

### Cross-linking of PrsA

Protein cross-linking was performed with 1% formaldehyde as has been described (Jensen *et al.*, 2005). Bacteria were grown in LB medium at 37°C until culture density was 100 Klett units. Cells were harvested from 1 ml of the culture, washed once with 1 ml 0.1 M sodium phosphate buffer (pH 6.8) and cross-linked with formaldehyde for 10 min at room temperature. Cells were pelleted, washed once in phosphate buffer and solubilized in SDS-PAGE lysis buffer as has been described (Healy *et al.*, 1991). Cross-linked PrsA or PrsA-Myc proteins were analysed by SDS-PAGE and immunoblotting. Samples were heated either at 37°C for 10 min or at 95°C for 30 min (breaks formaldehyde cross-links) prior to SDS-PAGE.

### References

Alonzo, F. 3rd, Port, G.C., Cao, M., and Freitag, N.E. (2009) The posttranslocation chaperone PrsA2 contributes to multiple facets of *Listeria monocytogenes* pathogenesis. *Infect Immun* **77**: 2612–2623.

Antelmann, H., Darmon, E., Noone, D., Veening, J.W., Westers, H., Bron, S., *et al.* (2003) The extracellular proteome of *Bacillus subtilis* under secretion stress conditions. *Mol Microbiol* **49**: 143–156.

Atrih, A., Bacher, G., Allmaier, G., Williamson, M.P., and Foster, S.J. (1999) Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. *J Bacteriol* **181**: 3956–3966.

Behrens, S., Maier, R., de Cock, H., Schmid, F.X., and Gross, C.A. (2001) The SurA periplasmic PPIase lacking its parvulin domains functions in vivo and has chaperone activity. *EMBO J* **20**: 285–294.

Bramkamp, M., Weston, L., Daniel, R.A., and Errington, J. (2006) Regulated intramembrane proteolysis of FtsL protein and the control of cell division in *Bacillus subtilis*. *Mol Microbiol* **62**: 580–591.

Carballido-Lopez, R., and Errington, J. (2003) The bacterial cytoskeleton: *in vivo* dynamics of the actin-like protein Mbl of *Bacillus subtilis*. *Dev Cell* **4**: 19–28.

Carballido-Lopez, R., Formstone, A., Li, Y., Ehrlich, S.D., Noirot, P., and Errington, J. (2006) Actin homolog MreBH governs cell morphogenesis by localization of the cell wall hydrolase LytE. *Dev Cell* **11**: 399–409.

Claessen, D., Emmins, R., Hamoen, L.W., Daniel, R.A., Err-

ington, J., and Edwards, D.H. (2008) Control of the cell elongation-division cycle by shuttling of PBP1 protein in *Bacillus subtilis*. *Mol Microbiol* **68:** 1029–1046.

Courtin, P., Miranda, G., Guillot, A., Wessner, F., Mezange, C., Domakova, E., *et al.* (2006) Peptidoglycan structure analysis of *Lactococcus lactis* reveals the presence of an L,d-carboxypeptidase involved in peptidoglycan maturation. *J Bacteriol* **188:** 5293–5298.

Daniel, R.A., and Errington, J. (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell* **113:** 767–776.

Daniel, R.A., Harry, E.J., Katis, V.L., Wake, R.G., and Errington, J. (1998) Characterization of the essential cell division gene ftsL(yIID) of *Bacillus subtilis* and its role in the assembly of the division apparatus. *Mol Microbiol* **29:** 593–604.

Daniel, R.A., Noirot-Gros, M.F., Noirot, P., and Errington, J. (2006) Multiple interactions between the transmembrane division proteins of *Bacillus subtilis* and the role of FtsL instability in divisome assembly. *J Bacteriol* **188:** 7396–7404.

Darmon, E., Noone, D., Masson, A., Bron, S., Kuipers, O.P., Devine, K.M., and van Dijl, J.M. (2002) A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. *J Bacteriol* **184:** 5661–5671.

Defeu Soufo, H.J., and Graumann, P.L. (2004) Dynamic movement of actin-like proteins within bacterial cells. *EMBO Rep* **5:** 789–794.

Defeu Soufo, H.J., and Graumann, P.L. (2006) Dynamic localization and interaction with other *Bacillus subtilis* actin-like proteins are important for the function of MreB. *Mol Microbiol* **62:** 1340–1356.

Divakaruni, A.V., Loo, R.R., Xie, Y., Loo, J.A., and Gober, J.W. (2005) The cell-shape protein MreC interacts with extracytoplasmic proteins including cell wall assembly complexes in *Caulobacter crescentus*. *Proc Natl Acad Sci USA* **102:** 18602–18607.

Divakaruni, A.V., Baida, C., White, C.L., and Gober, J.W. (2007) The cell shape proteins MreB and MreC control cell morphogenesis by positioning cell wall synthetic complexes. *Mol Microbiol* **66:** 174–188.

Drouault, S., Anba, J., Bonneau, S., Bolotin, A., Ehrlich, S.D., and Renault, P. (2002) The peptidyl-prolyl isomerase motif is lacking in PmpA, the PrsA-like protein involved in the secretion machinery of *Lactococcus lactis*. *Appl Environ Microbiol* **68:** 3932–3942.

Ellermeier, C.D., and Losick, R. (2006) Evidence for a novel protease governing regulated intramembrane proteolysis and resistance to antimicrobial peptides in *Bacillus subtilis*. *Genes Dev* **20:** 1911–1922.

van den Ent, F., Leaver, M., Bendezu, F., Errington, J., de Boer, P., and Lowe, J. (2006) Dimeric structure of the cell shape protein MreC and its functional implications. *Mol Microbiol* **62:** 1631–1642.

Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., *et al.* (2004) A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics* **4:** 2849–2876.

Figge, R.M., Divakaruni, A.V., and Gober, J.W. (2004) MreB, the cell shape-determining bacterial actin homologue,

co-ordinates cell wall morphogenesis in *Caulobacter crescentus*. *Mol Microbiol* **51:** 1321–1332.

Formstone, A., and Errington, J. (2005) A magnesium-dependent mreB null mutant: implications for the role of mreB in *Bacillus subtilis*. *Mol Microbiol* **55:** 1646–1657.

Glauner, B. (1988) Separation and quantification of muropeptides with high-performance liquid chromatography. *Anal Biochem* **172:** 451–464.

Glauner, B., Holtje, J.V., and Schwarz, U. (1988) The composition of the murein of *Escherichia coli*. *J Biol Chem* **263:** 10088–10095.

Hani, J., Schelbert, B., Bernhardt, A., Domdey, H., Fischer, G., Wiebauer, K., and Rahfeld, J.U. (1999) Mutations in a peptidylprolyl-*cis*/*trans*-isomerase gene lead to a defect in 3′-end formation of a pre-mRNA in *Saccharomyces cerevisiae*. *J Biol Chem* **274:** 108–116.

Harry, E.J., Pogliano, K., and Losick, R. (1995) Use of immunofluorescence to visualize cell-specific gene expression during sporulation in *Bacillus subtilis*. *J Bacteriol* **177:** 3386–3393.

Healy, J., Weir, J., Smith, I., and Losick, R. (1991) Post-transcriptional control of a sporulation regulatory gene encoding transcription factor sigma H in *Bacillus subtilis*. *Mol Microbiol* **5:** 477–487.

Heikkinen, O., Seppala, R., Tossavainen, H., Heikkinen, S., Koskela, H., Permi, P., and Kilpelainen, I. (2009) Solution structure of the parvulin-type PPIase domain of *Staphylococcus aureus* PrsA – implications for the catalytic mechanism of parvulins. *BMC Struct Biol* **9:** 17.

Heinrich, J., Lunden, T., Kontinen, V.P., and Wiegert, T. (2008) The *Bacillus subtilis* ABC transporter EcsAB influences intramembrane proteolysis through RasP. *Microbiology* **154:** 1989–1997.

Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J., and Behrens, S. (2005) The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J Biol Chem* **280:** 23540–23548.

Hyyryläinen, H.-L., Vitikainen, M., Thwaite, J., Wu, H., Sarvas, M., Harwood, C., *et al.* (2000) d-alanine substitution of teichoic acids as a modulator of protein folding and stability at the cytoplasmic membrane/cell wall interface of *Bacillus subtilis*. *J Biol Chem* **275:** 26696–26703.

Hyyryläinen, H.L., Bolhuis, A., Darmon, E., Muukkonen, L., Koski, P., Vitikainen, M., *et al.* (2001) A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. *Mol Microbiol* **41:** 1159–1172.

Hyyrylainen, H.L., Sarvas, M., and Kontinen, V.P. (2005) Transcriptome analysis of the secretion stress response of *Bacillus subtilis*. *Appl Microbiol Biotechnol* **67:** 389–396.

Jacobs, M., Andersen, J.B., Kontinen, V., and Sarvas, M. (1993) *Bacillus subtilis* PrsA is required *in vivo* as an extra-cytoplasmic chaperone for secretion of active enzymes synthesized either with or without pro-sequences. *Mol Microbiol* **8:** 957–966.

Jensen, S.O., Thompson, L.S., and Harry, E.J. (2005) Cell division in *Bacillus subtilis*: FtsZ and FtsA association is Z-ring independent, and FtsA is required for efficient midcell Z-Ring assembly. *J Bacteriol* **187:** 6536–6544.

Jones, L.J., Carballido-Lopez, R., and Errington, J. (2001)

126  H.-L. Hyyryläinen et al.

Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell* **104**: 913–922.

Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., *et al.* (2003) The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol* **104**: 5–25.

Kaltwasser, M., Wiegert, T., and Schumann, W. (2002) Construction and application of epitope- and green fluorescent protein-tagging integration vectors for *Bacillus subtilis*. *Appl Environ Microbiol* **68**: 2624–2628.

Katis, V.L., and Wake, R.G. (1999) Membrane-bound division proteins DivIB and DivIC of *Bacillus subtilis* function solely through their external domains in both vegetative and sporulation division. *J Bacteriol* **181**: 2710–2718.

Kawai, Y., Daniel, R.A., and Errington, J. (2009) Regulation of cell wall morphogenesis in *Bacillus subtilis* by recruitment of PBP1 to the MreB helix. *Mol Microbiol* **71**: 1131–1144.

Kontinen, V., and Sarvas, M. (1993) The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Mol Microbiol* **8**: 727–737.

Kühlewein, A., Voll, G., Hernandez Alvarez, B., Kessler, H., Fischer, G., Rahfeld, J.U., and Gemmecker, G. (2004) Solution structure of *Escherichia coli* Par10: The prototypic member of the Parvulin family of peptidyl-prolyl *cis/trans* isomerases. *Protein Sci* **13**: 2378–2387.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**: 249–256.

Leaver, M., and Errington, J. (2005) Roles for MreC and MreD proteins in helical growth of the cylindrical cell wall in *Bacillus subtilis*. *Mol Microbiol* **57**: 1196–1209.

Letek, M., Ordonez, E., Vaquera, J., Margolin, W., Flardh, K., Mateos, L.M., and Gil, J.A. (2008) DivIVA is required for polar growth in the MreB-lacking rod-shaped actinomycete *Corynebacterium glutamicum*. *J Bacteriol* **190**: 3283–3292.

Lindholm, A., Ellmen, U., Tolonen-Martikainen, M., and Palva, A. (2006) Heterologous protein secretion in *Lactococcus lactis* is enhanced by the *Bacillus subtilis* chaperone-like protein PrsA. *Appl Microbiol Biotechnol* **73**: 904–914.

Lu, K.P., and Zhou, X.Z. (2007) The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat Rev Mol Cell Biol* **8**: 904–916.

Lu, K.P., Hanes, S.D., and Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **380**: 544–547.

Lu, P.J., Zhou, X.Z., Shen, M., and Lu, K.P. (1999) Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* **283**: 1325–1328.

Ma, Y., Bryant, A.E., Salmi, D.B., Hayes-Schroer, S.M., McIndoo, E., Aldape, M.J., and Stevens, D.L. (2006) Identification and characterization of bicistronic *speB* and *prsA* gene expression in the group A *Streptococcus*. *J Bacteriol* **188**: 7626–7634.

McPherson, D.C., Driks, A., and Popham, D.L. (2001) Two class A high-molecular-weight penicillin-binding proteins of *Bacillus subtilis* play redundant roles in sporulation. *J Bacteriol* **183**: 6046–6053.

Margot, P., and Karamata, D. (1996) The *wprA* gene of *Bacillus subtilis* 168, expressed during exponential growth, encodes a cell-wall-associated protease. *Microbiology* **142**: 3437–3444.

Pedersen, L.B., Angert, E.R., and Setlow, P. (1999) Septal localization of penicillin-binding protein 1 in *Bacillus subtilis*. *J Bacteriol* **181**: 3201–3211.

Pinho, M.G., and Errington, J. (2003) Dispersed mode of *Staphylococcus aureus* cell wall synthesis in the absence of the division machinery. *Mol Microbiol* **50**: 871–881.

Popham, D.L., and Setlow, P. (1995) Cloning, nucleotide sequence, and mutagenesis of the *Bacillus subtilis ponA* operon, which codes for penicillin-binding protein (PBP) 1 and a PBP-related factor. *J Bacteriol* **177**: 326–335.

Popham, D.L., and Setlow, P. (1996) Phenotypes of *Bacillus subtilis* mutants lacking multiple class a high-molecular-weight penicillin-binding proteins. *J Bacteriol* **178**: 2079–2085.

Popham, D.L., and Young, K.D. (2003) Role of penicillin-binding proteins in bacterial cell morphogenesis. *Curr Opin Microbiol* **6**: 594–599.

Popham, D.L., Helin, J., Costello, C.E., and Setlow, P. (1996) Analysis of the peptidoglycan structure of *Bacillus subtilis* endospores. *J Bacteriol* **178**: 6451–6458.

Rahfeld, J.-U., Rücknagel, K.P., Schelbert, B., Ludvig, B., Hacker, J., Mann, K., and Fischer, G. (1994) Confirmation of the existence of a third family among peptidyl-prolyl *cis/trans* isomerases. Amino acid sequence and recombinant production of parvulin. *FEBS Lett* **352**: 180–184.

Robson, S.A., and King, G.F. (2006) Domain architecture and structure of the bacterial cell division protein DivIB. *Proc Natl Acad Sci USA* **103**: 6700–6705.

Rodgers, K.K., Bu, Z., Fleming, K.G., Schatz, D.G., Engelman, D.M., and Coleman, J.E. (1996) A zinc-binding domain involved in the dimerization of RAG1. *J Mol Biol* **260**: 70–84.

Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., and Charlier, P. (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev* **32**: 234–258.

Scheffers, D.J., Jones, L.J., and Errington, J. (2004) Several distinct localization patterns for penicillin-binding proteins in *Bacillus subtilis*. *Mol Microbiol* **51**: 749–764.

Schiene, C., and Fischer, G. (2000) Enzymes that catalyse the restructuring of proteins. *Curr Opin Struct Biol* **10**: 40–45.

Soufo, H.J., and Graumann, P.L. (2003) Actin-like proteins MreB and Mbl from *Bacillus subtilis* are required for bipolar positioning of replication origins. *Curr Biol* **13**: 1916–1920.

Spiess, C., Beil, A., and Ehrmann, M. (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**: 339–347.

Stephenson, K., and Harwood, C.R. (1998) Influence of a cell-wall-associated protease on production of α-amylase by *Bacillus subtilis*. *Appl Environ Microbiol* **64**: 2875–2881.

Stewart, G.C. (2005) Taking shape: control of bacterial cell wall biosynthesis. *Mol Microbiol* **57**: 1177–1181.

Stulke, J., Hanschke, R., and Hecker, M. (1993) Temporal activation of beta-glucanase synthesis in *Bacillus subtilis* is mediated by the GTP pool. *J Gen Microbiol* **139**: 2041–2045.

_____

Stymest, K.H., and Klappa, P. (2008) The periplasmic peptidyl prolyl *cis-trans* isomerases PpiD and SurA have partially overlapping substrate specificities. *FEBS J* **275**: 3470–3479.

Tiyanont, K., Doan, T., Lazarus, M.B., Fang, X., Rudner, D.Z., and Walker, S. (2006) Imaging peptidoglycan biosynthesis in *Bacillus subtilis* with fluorescent antibiotics. *Proc Natl Acad Sci USA* **103**: 11033–11038.

Uchida, T., Fujimori, F., Tradler, T., Fischer, G., and Rahfeld, J.-U. (1999) Identification and characterization of a 14 kDa human protein as a novel parvulin like peptidyl-prolyl *cis/trans* isomerase. *FEBS Lett* **446**: 278–282.

Vitikainen, M., Pummi, T., Airaksinen, U., Wu, H., Sarvas, M., and Kontinen, V.P. (2001) Quantitation of the capacity of the secretion apparatus and requirement for PrsA in growth and secretion of α-amylase in *Bacillus subtilis*. *J Bacteriol* **183**: 1881–1890.

Vitikainen, M., Lappalainen, I., Seppala, R., Antelmann, H., Boer, H., Taira, S., *et al.* (2004) Structure-function analysis of PrsA reveals roles for the parvulin-like and flanking N- and C-terminal domains in protein folding and secretion in *Bacillus subtilis*. *J Biol Chem* **279**: 19302–19314.

Vitikainen, M., Hyyrylainen, H.L., Kivimaki, A., Kontinen, V.P., and Sarvas, M. (2005) Secretion of heterologous proteins in *Bacillus subtilis* can be improved by engineering cell components affecting post-translocational protein folding and degradation. *J Appl Microbiol* **99**: 363–375.

Wahlstrom, E., Vitikainen, M., Kontinen, V.P., and Sarvas, M. (2003) The extracytoplasmic folding factor PrsA is required for protein secretion only in the presence of the cell wall in *Bacillus subtilis*. *Microbiology* **149**: 569–577.

Wang, P., and Heitman, J. (2005) The cyclophilins. *Genome Biol* **6**: 226.

Wei, Y., Havasy, T., McPherson, D.C., and Popham, D.L. (2003) Rod shape determination by the *Bacillus subtilis* class B penicillin-binding proteins encoded by *pbpA* and *pbpH*. *J Bacteriol* **185**: 4717–4726.

Williams, R.C., Rees, M.L., Jacobs, M.F., Pragai, Z., Thwaite, J.E., Baillie, L.W., *et al.* (2003) Production of *Bacillus anthracis* protective antigen is dependent on the extracellular chaperone, PrsA. *J Biol Chem* **278**: 18056–18062.

Wu, S.C., Ye, R., Wu, X.C., Ng, S.C., and Wong, S.L. (1998) Enhanced secretory production of a single-chain antibody fragment from *Bacillus subtilis* by coproduction of molecular chaperones. *J Bacteriol* **180**: 2830–2835.

Wu, X., Wilcox, C.B., Devasahayam, G., Hackett, R.L., Arevalo-Rodriguez, M., Cardenas, M.E., *et al.* (2000) The Ess1 prolyl isomerase is linked to chromatin remodeling complexes and the general transcription machinery. *EMBO J* **19**: 3727–3738.

Yaffe, M.B., Schutkowski, M., Shen, M., Zhou, X.Z., Stukenberg, P.T., Rahfeld, J.U., *et al.* (1997) Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. *Science* **278**: 1957–1960.

Yao, J.L., Kops, O., Lu, P.J., and Lu, K.P. (2001) Functional conservation of phosphorylation-specific prolyl isomerases in plants. *J Biol Chem* **276**: 13517–13523.

Zemansky, J., Kline, B.C., Woodward, J.J., Leber, J.H., Marquis, H., and Portnoy, D.A. (2009) Development of a mariner-based transposon and identification of *Listeria monocytogenes* determinants, including the peptidyl-prolyl isomerase PrsA2, that contribute to its hemolytic phenotype. *J Bacteriol* **191**: 3950–3964.

Zhao, G., Meier, T.I., Kahl, S.D., Gee, K.R., and Blaszczak, L.C. (1999) BOCILLIN FL, a sensitive and commercially available reagent for detection of penicillin-binding proteins. *Antimicrob Agents Chemother* **43**: 1124–1128.

## Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

*Bacillus subtilis YqjG is required for genetic competence development*

# *Bacillus subtilis* YqjG is required

# for genetic competence development

Manfred J. Saller[1], Andreas Otto[2], Greetje A. Berrelkamp-Lahpor[1], Dörte Becher[2],

Michael Hecker[2] and Arnold J.M. Driessen[1*]

**Running title: *Bacillus subtilis* YidC homologs**

[1]Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology

Institute, Kluyver Centre for the Genomics of Industrial Fermentations and the

Zernike Institute of Advanced Materials, University of Groningen, 9751 NN Haren,

The Netherlands,

[2]Institute for Microbiology, Ernst-Moritz-Arndt-University Greifswald, F.-L.-Jahn

Str. 15, 17489 Greifswald, Germany

*Address correspondence to: Arnold J.M. Driessen, Tel. 31-50-3632164; Fax. 31-50-3632164; E-mail: a.j.m.driessen@rug.nl

1    Members of the evolutionary conserved Oxa1/Alb3/YidC family have been shown to

2    play an important role in membrane protein insertion, folding and/or assembly.

3    *Bacillus subtilis* contains two YidC-like proteins, denoted as SpoIIIJ and YqjG.

4    SpoIIIJ and YqjG are largely exchangeable, but SpoIIIJ is essential for spore

5    formation and YqjG cannot complement this activity. To elucidate the role of YqjG,

6    we determined the membrane proteome and functional aspects of *B. subtilis* cells

7    devoid of SpoIIIJ, YqjG or both. The data show that SpoIIIJ and YqjG have

8    complementary functions in membrane protein insertion and assembly. The reduced

9    levels of $F_1F_O$ ATP synthase in cells devoid of both SpoIIIJ and YqjG are due to a

10   defective assembly of the $F_1$-domain onto the $F_0$-domain. Importantly, for the first

11   time, a specific function is demonstrated for YqjG in genetic competence

12   development.

13

14

## 15    1.    Introduction

16

17   The YidC/Oxa1/Alb3 protein family, which comprises evolutionary conserved

18   integral membrane proteins, plays an important role in the insertion of proteins into

19   the bacterial, mitochondrial and thylakoid membrane. While these proteins share only

20   low primary sequence similarity, all members consist of a conserved hydrophobic

21   region of five transmembrane segments (TMS) connected by short hydrophilic loops.

22   The N- and C-terminal flanking regions are very diverse in length and sequence and

23   are even completely absent in some Oxa1p-like proteins [1]. With the *E. coli* YidC,

24   the conserved region suffices for the membrane protein insertion activity whereas the

25   flanking regions are dispensable [2]. So far no function has been proposed for the

1    large N-terminal periplasmic domain of YidC. Some YidC homologues bear a

2    positively charged, C-terminal extension that binds ribosome-nascent-chain (RNC)

3    complexes [3-5].

4          YidC facilitates the insertion of membrane proteins independently [6] or in

5    concert with the translocon [7, 8], the protein conducting pore formed by the

6    heterotrimeric complex SecYEG. Although only few membrane proteins have been

7    shown to be dependent on only YidC for their insertion, a preference for small

8    hydrophobic membrane protein subunits of large energy transducing complexes

9    appears evident [9]. The membrane insertion of subunit c of the $F_1F_O$ ATP synthase

10    ($F_Oc$) [6] and the mechanosensitive channel for large conductance (MscL) [10] is

11    dependent on YidC only. YidC also mediates the membrane insertion of the phage

12    coat proteins M13 and Pf3, whereas the insertion of subunit a of the $F_1F_O$ ATP

13    synthase ($F_Oa$) [6, 11] and the cytochrome o oxidase (CyoA) [7, 8] strictly depends on

14    both YidC and the translocon. Recent studies demonstrate a role of YidC in the

15    biogenesis of anaerobic respiratory chain complexes like the NADH dehydrogenase

16    complex I, and the fumarate and nitrate reductases [12, 13].

17          The number of Oxa1 homologues per cell or organelle varies between 1 (gram

18    negative bacteria) and 6 (*Arabidopsis thaliana*). In *B. subtilis* two homologues,

19    SpoIIIJ and YqjG, are present. Both proteins complement the function of YidC in *E.*

20    *coli* and promote membrane insertion of $F_Oc$ from *B. subtilis* and *E. coli* [14]. In

21    addition, SpoIIIJ and YqjG were found to form a complex with the entire $F_1F_O$ ATP

22    synthase reminiscent of the Oxa1/ATPase complex found in the yeast *Saccharomyces*

23    *cerevisiae* [15]. While SpoIIIJ and YqjG appear to be exchangeable in membrane

24    protein insertion and complex assembly [14], only SpoIIIJ is essential for spore

25    formation in *B. subtilis* and cannot be complemented by YqjG [16, 17]. During

1    vegetative growth expression of either of the homologues is sufficient for cell

2    viability, whereas the double deletion exhibits a lethal phenotype [16, 18, 19]. The

3    expression of YqjG is low but is induced upon initiation of sporulation and under

4    SpoIIIJ limiting conditions. A recent study provided evidence that translation of *yqjG*

5    is stalled until the cell encounters problems in membrane protein insertion [20]. These

6    data suggested that YqjG is a back-up protein for SpoIIIJ to ensure cell survival

7    during vegetative growth when SpoIIIJ is limiting. A function, specific for YqjG, has

8    not yet been demonstrated. In order to address this question, we chose to perform a

9    proteomic and functional study on *B. subtilis* strains that lack YqjG, SpoIIIJ or both

10    proteins. Our results indicate SpoIIIJ and YqjG fulfill complementary functions

11    within the cell, such as the biogenesis of energy-transducing complexes. Remarkably,

12    the protein levels of the cytoplasmic $F_1$-domain but not of the membrane-embedded

13    $F_O$-sector were decreased in a SpoIIIJ/YqjG double deletion strain indicating that

14    complex assembly, and not the membrane insertion of the $F_O$ subunits is responsible

15    for the defect within the $F_1F_O$ ATP synthase. Our studies demonstrate that YqjG plays

16    an exclusive and critical role in genetic competence development, which differentiates

17    this YidC protein from SpoIIIJ that has a function in sporulation.

18

19

20    **2.    Materials and methods**

21

22    **2.1    Antisera**

23

24    Antisera against various proteins from *B. subtilis* were generous gifts from T.

25    Mascher (Ludwig-Maximilians-Universität, München, Germany, LiaH), J.-M. van

1   Dijl (UMCG, Groningen, The Netherlands, BdbD and ComGC), V.P. Kontinen (THL,

2   Helsinki, Finnland, PrsA) and the late A. Kroeger (Frankfurt, Germany, QoxA). All

3   antibodies were polyclonal and raised in rabbits. After incubation with the secondary

4   antibody (anti-rabbit IgG conjugated with alkaline phosphatase; Sigma-Aldrich

5   Chemie GmbH, Steinheim, Germany) and CDP-*Star* (Roche Diagnostics GmbH,

6   Mannheim, Germany) chemiluminescent signals were detected by the Luminescent

7   Image Analyzer LAS-4000 (FujiFilm, Tokio, Japan).

8

9   **2.2.    Media and growth conditions**

10

11  *B. subtilis* was grown in Luria Bertani (LB), synthetic minimal medium (SMM) [21]

12  or chemically defined competence medium (CDCM) [22] at 37 ºC and shaken at 200

13  rpm. When appropriate, media were supplemented with spectinomycin (100 µg/ ml),

14  erythromycin (2 µg/ ml), tetracycline (12 µg/ ml), chloramphenicol (10 µg/ ml) or

15  kanamycin (10 µg/ ml).

16

17  **2.3     Strains and plasmids**

18

19  All primer sequences used in this study are shown in Table 1. All plasmids and strains

20  are listed in Table 2. *B. subtilis* was transformed as described before [23]. To delete

21  *spoIIIJ*, the 5'- and 3'- flanking regions of the gene were amplified by PCR from

22  genomic DNA of *B. subtilis* 168 as template and using primer pairs Spod1/Spod2 and

23  Spod3/Spod4, respectively. The tetracycline resistance cassette was obtained by PCR

24  using the primer pair Tet1/Tet2 and plasmid pDG1514 as template. Next, the

25  tetracycline cassette was fused between the *spoIIIJ* flanking regions by overlapping

1     PCR using primer set Spod1/Spod2. The resulting PCR product was transformed into

2     *B. subtilis* 168 and selected on LB plates containing tetracycline. Correct

3     integration/deletion was verified by PCR and the strain was named *B. subtilis* MS001

4     (*ΔspoIIIJ*). The same strategy was followed to delete *yqjG*. The flanking regions of

5     *yqjG* were amplified by PCR using primer pairs YqjGd1/YqjGd2-Spec and YqjGd3-

6     Spec/YqjG d4. Next, a PCR reaction with plasmid pDG1726 as template and primer

7     pair Spec1/Spec2 yielded the spectinomycin cassette. Again, the flanking regions and

8     the spectinomycin cassette were fused by PCR and transformed into *B. subtilis* 168.

9     After overnight incubation on LB plates supplemented with spectinomycin,

10    transformants were checked for correct integration by PCR (data not shown) and

11    denoted as *B. subtilis* MS003 (*ΔyqjG*).

12        By transforming either pAX-spoIIIJ or pAX-yqjG into *B. subtilis* MS001

13    (*ΔspoIIIJ*), a xylose-inducible *spoIIIJ* or *yqjG* gene was inserted into the genomic

14    *lacA* locus. Positive transformants were selected on LB plated containing

15    erythromycin and tetracycline. Correct double crossover insertion was verified by

16    PCR. Finally, in both strains the *yqjG* gene was deleted from the genome and replaced

17    by a spectinomycin resistance cassette. The last step was carried out essentially as

18    described above for the construction of *B. subtilis* MS003 (*ΔyqjG*). This resulted in *B.*

19    *subtilis* strains MS004 (*ΔyqjG*, *ΔspoIIIJ*, $P_{xylA}$ *spoIIIJ*) and MS005 (*ΔyqjG*, *ΔspoIIIJ*,

20    $P_{xylA}$ *yqjG*), both resistant against tetracycline, erythromycin and spectinomycin and

21    strictly dependent on xylose for growth.

22        To detect the expression of the ComG-operon of *B. subtilis* when grown in a

23    chemically defined medium, a *gfp*-reporter gene was fused downstream of the *comGA*

24    promoter region [22]. Herein, the plasmid pSGComGA1 (W.K. Smits unpublished), a

25    derivate of pSGComGA [24] carrying a kanamycin resistance cassette, was

1     transformed into *B. subtilis* 168, MS001 and MS003. The plasmid integrates by a

2     single crossover reaction into the *comGA* promoter region of the chromosome. The

3     resulting *B. subtilis* strains carrying the *gfp*-reporter gene were named MS006 (wild

4     type derivate), MS007 (*ΔspoIIIJ*) and MS008 (*ΔyqjG*), respectively.

5

6     **2.4     Cell growth for metabolic labeling**

7

8     Cells of *B. subtilis* strains 168, MS001 (*ΔspoIIIJ*), MS003 (*ΔyqjG*) and MS004

9     (*ΔyqjG, ΔspoIIIJ*, $P_{xylA}$ *spoIIIJ*) were grown under agitation (200 rpm) at 37 °C in

10    SMM [21] supplemented with either $^{15}$N-ammonium sulfate and $^{15}$N- L- tryptophane

11    (0.078 mM, 98 atom % excess, Cambridge Isotope Laboratories, Andover, USA) in

12    the labeled medium or with $^{14}$N- ammonium sulfate and $^{14}$N- L- tryptophane in the

13    unlabelled medium. Cultures were supplemented with the appropriate antibiotics and

14    either 0.2% (wt/vol) glucose (168, MS001 and MS003) or xylose (MS004) as the sole

15    carbon source. After overnight incubation, cultures were diluted 1:50 (168, MS001

16    and MS003) or 1:25 (MS004) in fresh medium, cell growth continued to an $OD_{600\,nm}$

17    of 1 and cells were harvested by centrifugation (10 min, 7500 x g). To obtain cells

18    devoid of both SpoIIIJ and YqjG, MS004 (*ΔyqjG, ΔspoIIIJ*, $P_{xylA}$ *spoIIIJ*) cells were

19    washed once in pre-warmed SMM without sugar, resuspended to an $OD_{600\,nm}$ of 0.5 in

20    SMM containing 0.2% (wt/vol) glucose, and cell growth continued to an $OD_{600\,nm}$ of

21    1. Next, cells were diluted two-fold in SMM containing 0.2% (wt/vol) glucose and

22    further incubated at 37 °C. This procedure was repeated until cells stopped growing,

23    which occurred mostly after the 5th generation. Cells were harvested, and equivalent

24    OD-units of bacteria grown on SMM supplemented with $^{14}$N or $^{15}$N were pooled,

25    quick-frozen in liquid nitrogen and stored at -80 °C.

1

2    **2.5    Membrane and sample preparation**

3

4    Membrane isolation was carried out essentially as described [25] excluding the

5    protoplasting step. Samples that were analyzed by the GeLCMS-workflow (See

6    below) were fractionated by SDS-PAGE following tryptic digestion [26].

7    *Proteomics measurements* - Tryptic digests of the SDS-PAGE gel pieces were loaded

8    on a reversed phase column (Waters BEH 1.7 µm, 100-µm i. D. x 100 mm) operated

9    on a nanoACQUITY-UPLC (Waters Corporation, Milford, MA, USA). Peptides were

10   first concentrated and desalted on a trap column (Waters nanoACQUITY UPLC

11   column, Symmetry C18, 5 µm, 180 µm x 20 mm) for 3 min at a flow rate of 1 ml/ min

12   with 99 % acetonitrile in 0.1 % acetic acid. Subsequently, the peptides were eluted

13   and separated with a non-linear 80-min gradient from 5 –60 % ACN in 0.1 % acetic

14   acid at a constant flow rate of 400 nl/min. MS/MS data were acquired with the LTQ-

15   Orbitrap mass spectrometer (Thermo Fisher, Bremen, Germany) equipped with a

16   nanoelectrospray ion source. After a survey scan in the Orbitrap (r= 30,000) MS data

17   were recorded for the five most intensive precursor ions in the linear ion trap. Singly

18   charged ions were not taken into account for MS analysis [21].

19

20   **2.6    Data analysis: protein identification**

21

22   The *.dta files were extracted from *.raw files using Bioworks Browser 3.3.1 SP1

23   (Thermo Fisher Scientific) without charge state deconvolution and deisotoping.

24   Subsequently, the data were searched with SEQUEST version v28 (rev.12) (Thermo

25   Fisher Scientific) against a *B. subtilis* target-decoy protein sequence database

1    (complete proteome set of *B. subtilis* extracted from UniprotKB release 12.7 (UniProt

2    Consortium, Nucleic acids research 2007, 35, D193-197) with a set of common

3    laboratory contaminants) compiled with BioworksBrowser. The searches were done

4    in two iterations: First, for the GeLCMS analyses the following search parameters

5    were used: enzyme type, trypsin (KR); peptide tolerance - 10 ppm; tolerance for

6    fragment ions - 1 amu; consideration of b- and y-ion series; variable modification -

7    methionine (15.99 Da); a maximum of three modifications per peptide was allowed.

8    In the second iteration the mass shift of all amino acids completely labeled with $^{15}$N

9    Nitrogen was taken into account in the search parameters. Resulting *.dta and *.out

10   files were assembled and filtered using DTASelect (version 2.0.25) (parameters

11   GeLCMS: -y 2 -c 2 -C 4 --here --decoy Reverse_ -p 2 -t 2 -u --MC 2 -i 0.3 --fp 0.005).

12   Protein identification data was cured with an in- house written java- script to ensure

13   that each protein hit relies on at least two different peptides as judged by amino acid

14   sequence.

15

16   **2.7**     **Data analysis: relative quantification**

17

18   The cured search results served as the base for the software CenSus [27] to obtain

19   quantitative data of $^{14}$N peaks (sample) and $^{15}$N peaks (pooled reference) [28]. Peptide

20   ratios were determined by CenSus and subsequently exported ($R^2$ values bigger than

21   0.7 and only unique peptides; proteins failing to be relatively quantified were checked

22   manually in the graphical user interface for on/off proteins). Proteins relatively

23   quantified with at least 2 peptides were taken into account for the subsequent analysis.

24

25   **2.8.**     **Data analysis: protein classification and localization**

1

2   The identified proteins were sorted into functional categories according to their

3   functional assessment at the protein database of NCBI. In order to predict the cellular

4   localization of all identified proteins searches against the LocateP Database

5   (http://www.cmbi.ru.nl/locatep-db/cgi-bin/locatepdb.py) [29] were performed.

6

7   **2.9.    ATPase activity assay**

8

9   The ATPase activity of isolated *B. subtilis* membranes was determined by means of

10  inorganic phosphate release and detected with the dye malachite green [30].

11  Membranes were diluted in assay buffer (50 mM HEPES/KOH pH 7.5, 50 mM KCl,

12  2.5 mM MgCl$_2$, and 2 mM DTT) to a concentration of 40 µg/ml and incubated at 37

13  ºC. The reaction was started by adding ATP (1 mM final concentration) and stopped

14  by the addition of EDTA (10 mM). In order to determine DCCD (N,N'-

15  Dicyclohexylcarbodiimide) sensitive ATPase activity, assays were performed in the

16  presence of 0.5 mM DCCD.

17

18  **2.10    Flow cytometry and transformability assay**

19

20  *B. subtilis* MS006, MS007 and MS008 were grown in CDCM medium supplemented

21  with the appropriate antibiotic. These strains were derived from *B. subtilis* 168, and

22  MS001 (ΔspoIIIJ) and MS003 (ΔyqjG) contained a *gfp* reporter gene fused to the

23  *comGA* promoter region. This enabled the fluorescent detection of comG expression.

24  Freshly grown overnight cultures were diluted 1:25 in CDCM and incubated until no

25  further increase of GFP fluorescence could be observed. At an hourly interval aliquots

1  were taken and 20000 cells directly measured on a Coulter Epics XL-MCL flow

2  cytometer (Beckman Coulter, Mijdrecht, the Netherlands). The setup of the flow

3  cytometric analysis was essentially as described [31]. In order to determine the

4  transformability, 10 µg plasmid DNA from pTRKH2 or pNZ8901 was transformed

5  into *B. subtilis* and cells were plated on LB agar plates both without (viable count)

6  and with antibiotic (transformant count). After overnight incubation at 37 ºC, the

7  colony forming units (CFUs) were counted to yield the transformation efficiency.

8

9

10

11

12

13

14  **3  Results**

15

16  **3.1  Membrane proteome of *B. subtilis* Δ*spoIIIJ* and Δ*yqjG* strains**

17

18  The membrane proteomes of the *B. subtilis* strains MS001 (*ΔspoIIIJ*) and MS003

19  (*ΔyqjG*) were compared with that of the wild type. For this purpose, *B. subtilis* 168,

20  MS001 and MS003 were grown in labeled and unlabeled SMM medium (see also 2.4)

21  supplemented with 0.2% (wt/vol) glucose and the appropriate antibiotics. After the

22  cultures reached an $OD_{600nm}$ of 1, cell growth was stopped and membranes were

23  isolated from the different strains (Figure 1). Proteins present in the membrane

24  fraction were subjected to mass spectrometric analysis by GeLCMS. For each

25  comparison four samples (two biological duplicates of each strain) were analyzed and

1    relatively quantified to the $^{15}$N-labeled reference. In total 572 (MS001/WT), 417

2    (MS003/WT) and 547 (MS004/MS003) proteins were identified and quantified (for

3    an overview see Supplementary data: Figure 1 and Table 1) in at least one biological

4    duplicate of each strain. Considering that 70 % of *B. subtilis* proteins are predicted to

5    be cytosolic by LocateP [29] and that membrane protein analysis by GeLCMS is

6    impaired by their physiochemical properties, membrane proteins are very well

7    represented in our data set. Between 45 and 55 % of all identified proteins are

8    predicted to be membrane proteins (Supplementary Figure 1) including multi-

9    spanning, C- or N-terminal anchored and lipid anchored proteins.

10       Upon analysis of the individual *ΔspoIIIJ* and *ΔyqjG* strains, it appeared that

11    the expression levels of most proteins are unchanged compared to the wild type

12    (Figure 2A and B). With a 2-fold in- or decrease in protein amount as a threshold, 19

13    proteins revealed a higher and only 2 showed a lower protein amount in the *ΔspoIIIJ*

14    strain, while in the *ΔyqjG* strain the protein amount of only 2 proteins was altered over

15    the threshold compared to the wild type (Table 3). As expected, SpoIIIJ was not

16    present in MS001 (*ΔspoIIIJ*) and its level was not altered in MS003 (*ΔyqjG*), while

17    YqjG was not present in MS003. In MS001, YqjG exhibited the greatest change in

18    amount of all quantified proteins (20-fold increase), which is in line with previous

19    studies that showed increased levels of *yqjG* transcripts in the absence of SpoIIIJ [18,

20    20, 32]. Most of the proteins with an increased quantity in MS001 (*ΔspoIIIJ*) have a

21    function in the late stages of spore development, e.g., SpoIIIAH, SpoIVA, SpoIVFA,

22    SpoVID, YdcC and GerM (Table 3). Furthermore, the penicillin-binding protein

23    PbpF, the ribosomal protein RpsP, the putative NADH dehydrogenases YumB, the

24    major UDP-glucose dehydrogenase YwqF and four proteins with yet unknown

25    function (ComER, YqeZ, YuzA, Ytta) were significantly increased in protein amount.

1 Another three proteins involved in transport exhibited higher levels, whereas the

2 uncharacterized ABC-transporter subunit YwjA and the malate kinase sensor MalK

3 were present at lower levels. The alkaline serine protease AprX is absent in MS001

4 (*ΔspoIIIJ*) and over-represented in MS003 (*ΔyqjG*). Also the ribosomal subunit RplA

5 was found to be overrepresented in MS003 while the level of the aerotactic transducer

6 HemAT and YwpH, a single strand DNA-binding protein, was diminished (Table 3).

7 These data indicate that there are only few changes in the membrane proteome upon

8 the deletion of the SpoIIIJ or YqjG function, except for the increased amounts of

9 proteins that function in the late stages of spore development in the *ΔspoIIIJ* strain.

10

11 **3.2    Membrane proteome of cells depleted both for SpoIIIJ and YqjG**

12

13 Since SpoIIIJ and YqjG likely have overlapping functions, relative quantities of

14 protein amounts in strain MS004, containing both a deletion of *yqjG* and having the

15 expression of *spoIIIJ* under control of the xylose promoter, were examined. For

16 comparison strain MS003 (Δ*yqjG*) was used. The effect of the depletion of both

17 proteins from the cell was by far more drastic than the effect observed in the single

18 deletion strains (Figure 2C). SpoIIIJ was present in MS003 but absent in MS004

19 indicating that the cells were completely depleted of SpoIIIJ during cell growth in

20 absence of xylose.

21      As expected, major changes in protein expression levels could be observed

22 (Supplementary data, Table 1). Out of the 547 quantified proteins, 160 exhibited a

23 significant (> 2-fold) overexpression and 142 were found at reduced amounts (Figure

24 2C). Remarkably, the majority (85 %) of proteins with increased amount were

25 cytoplasmic proteins (Figure 2C, inset, yellow lines). On the other hand, the majority

1     of the proteins with diminished amount are membrane proteins (61 %) (Figure 2C,

2     inset, blue lines). This demonstrates that the depletion of the SpoIIIJ/YqjG function

3     has a marked effect on the membrane proteome.

4        Whilst no obvious stress response was induced in the Δ*spoIIIJ* and Δ*yqjG*

5     strains, depletion of SpoIIIJ in a Δ*yqjG* strain (MS004) resulted in increased levels of

6     specific stress response proteins (Table 4). The highest induction (more than 110-fold)

7     is observed for LiaH, which is normally induced under cell wall stress. LiaH shows

8     sequence similarity with the phage shock protein A (PspA) from *E. coli*. Its

9     expression is dependent on the two-component system LiaRS and is strongly

10    repressed by LiaF [33]. Protein amounts of LiaS and LiaF were about 10- and 17-fold

11    higher than in strain MS003 (Δ*yqjG*). The pronounced LiaH response in MS004 was

12    verified by immunoblotting (Figure 1). It should be noted, that the amount of LiaH in

13    membranes derived from MS005 (YqjG-depletion in a Δ*spoIIIJ* strain) was less than

14    in MS004 (SpoIIIJ-depletion in a Δ*yqjG* strain). A possible reason therefore is that

15    MS004 cells were more strongly depleted than MS005. On the other hand, LiaH could

16    neither be detected in membranes of the wild type nor in the single deletion strains

17    MS001 and MS003. In addition, another five stress response proteins were found in

18    increased amounts in the double depletion strain. The relative amounts of membrane

19    associated HtpX, YdjF and YokG was about 4.5-, 4.5- and 2.5-fold increased,

20    respectively (Table 4). Finally, the Clp protease system was induced upon

21    SpoIIIJ/YqjG depletion. Whereas ClpC did not show an altered protein level, ClpP

22    and ClpX were found to be increased 5.1- and 4.8-fold in amount. Other stress

23    response proteins like the well-characterized stress response chaperones GroELS or

24    DnaKJ, did not exhibit any changes.

1    In addition, there are several classes of membrane proteins whose expression

2    is affected by the SpoIIIJ/YqjG depletion. Only the most notable are discussed below.

3    The protein amount of PrsA, a lipid-anchored membrane protein with an essential role

4    in protein secretion [34], was found unaffected in both the proteomic and the western

5    blot analysis (Figure 1) and was therefore used as an internal standard. A list of all

6    identified and quantified proteins and their localization (predicted by LocateP [29]) is

7    presented in the Supplementary data (Table S1).

8

9    **3.3    Functional overlap of SpoIIIJ and YqjG in membrane protein insertion**

10   **and assembly of energy-transducing complexes**

11

12   The proteomic analysis is consistent with a complementary function of SpoIIIJ and

13   YqjG in the biogenesis of energy transducing protein complexes. In both MS001 and

14   MS003, none of the energy transduction subunits/complexes revealed an altered

15   protein expression level beside the putative NADH dehydrogenase YumB, which

16   increased 3.1-fold in amount upon deletion of *spoIIIJ* (Supplementary data, Table S1).

17   Despite of sharing sequence similarity to other NADH dehydrogenases, YumB was

18   recently reported to be possibly wrongly annotated [35], which is supported by the

19   wild type growth behavior of a Δ*yumB* strain [36]. YumB might however be a minor

20   NADH dehydrogenase that is functionally masked by other dehydrogenases, e.g.

21   YjlD. In the MS004 strain (SpoIIIJ-depletion in a Δ*yqjG* strain), CtaC was unchanged

22   and YjlD, the major NADH dehydrogenase [36], was reduced 3.1-fold in protein

23   amount as compared to strain MS003 (Δ*yqjG*). The amount of the cytochrome oxidase

24   subunit QoxA in MS004 membrane was almost 2-fold reduced (Supplementary data,

1    Table 1). Immunoblot analysis, however, indicated a greater loss of QoxA upon

2    SpoIIIJ- and YqjG-depletion, as no QoxA could be detected (Figure 1).

3        The proteomic approach did not identify all integral membrane subunits of the

4    $F_1F_O$ ATP synthase in all compared strains (Table 5). Therefore, the DCCD-sensitive

5    ATPase activity of membranes derived from the various strains was determined.

6    Membranes derived from the wild type and the single *yqjG* and *spoIIIJ* deletion

7    mutants showed essentially equal levels of functional $F_1F_O$ ATP synthases (Figure 3).

8    However, the ATPase activity of MS004 (SpoIIIJ-depletion in a $\Delta yqjG$ strain) and

9    MS005 (YqjG-depletion in a $\Delta spoIIIJ$ strain) dropped to about 40 % of that found for

10    the wild type indicating a loss of functional ATP synthase complexes in cells devoid

11    of both SpoIIIJ and YqjG. Remarkably, the proteomic data revealed that the amounts

12    of the membrane-embedded $F_O a$, $F_O b$ and $F_O c$ did not drastically change, but rather all

13    subunits of the cytoplasmic $F_1$-domain ($\alpha - \varepsilon$) were 4-fold (Table 5) less abundant in

14    the membrane fraction derived from strain MS004 (SpoIIIJ-depletion in a $\Delta yqjG$

15    strain) as compared to MS003 ($\Delta yqjG$). This data strongly suggests that the main

16    defect with the $F_1F_O$ ATP synthase is not at the level of the membrane insertion of $F_O$

17    subunits, but rather in the functional assembly of the $F_1$-domain onto the $F_O$-domain.

18

19    **3.4.**    **Differential effect of *spoIIIJ* and *yqjG* deletion on genetic competence**

20    **development**

21

22    The protein levels in the single deletion strains MS001 and MS003 were hardly

23    affected and nearly similar to wild type levels. When a less stringent ± 1.5-fold

24    change was considered, a greater number of proteins of a functional category followed

25    the same direction. This is particularly evident for proteins involved in competence

_____

1    development (Table 6), some of which are essential for DNA binding and uptake. The

2    relative amounts of this group of proteins showed the same trend and were increased

3    in MS001 (Δ*spoIIIJ*) and decreased in MS003 (Δ*yqjG*) as compared to the wild type.

4    Next, the DNA uptake ability was tested by the transformability of the wild type, and

5    the two deletion strains MS001 and MS003. Cells were transformed with plasmid

6    (pTRKH2 or pNZ8901) or genomic DNA and transferred to LB plates with

7    (transformants) or without antibiotics (viable count). As shown in Figure 4A the

8    transformation frequencies are independent of the DNA source used. For the wild type

9    the transformation efficiencies (number of transformants/number of viable cells) were

10   $0.38 \times 10^{-3}$ (plasmid DNA) and $0.42 \times 10^{-3}$ (genomic DNA). These values increased

11   for the MS001 strain in which *spoIIIJ* was deleted yielding frequencies of $1.21 \times 10^{-3}$

12   and $1.18 \times 10^{-3}$ while in the MS003 strain in which *yqjG* was deleted the numbers

13   decreased to $0.18 \times 10^{-3}$ and $0.25 \times 10^{-3}$, respectively. Thus, it appears that the

14   transformability is diminished in the Δ*yqjG* strain to about 50 % of the wild type,

15   whereas, a marked 3-fold increase in the transformability was noted in the *spoIIIJ*

16   deletion strain. Results derived from transformation experiments were shown to fit the

17   standards for statistical significance by one-way analysis of variance (ANOVA). All

18   tests resulted in p-values lower than 0.01 (P<0.01).

19          Next, the expression levels of the ComG-operon were monitored by flow

20   cytometry as a measure of competence development [22]. In *B. subtilis* MS006 (wild

21   type), MS007 (Δ*spoIIIJ*) and MS008 (Δ*yqjG*) a *gfp* reporter gene was fused to the

22   *comGA* promoter. Expression of the ComG-operon should induce GFP expression,

23   thus resulting in increased levels of fluorescence [31]. This construct has been shown

24   to be an excellent reporter for identification of competent cells after initiation of

25   competence development [24]. To exclude that the insertion of the *comG-gfp* reporter

1    has itself an effect on genetic competence development, transformation frequencies

2    were determined for *B. subtilis* MS006, MS007 and MS008 (Figure 4A). Again,

3    similar values ($0.41 \times 10^{-3}$, $1.17 \times 10^{-3}$ and $0.19 \times 10^{-3}$) were obtained for the uptake of

4    plasmid DNA and therefore a direct effect of the reporter can be excluded. For the

5    flow cytometry detection, cells were grown in CDCM and the fluorescence was

6    measured every hour. During the lag and early exponential phase only few fluorescent

7    cells could be detected (data not shown). With all strains, at late exponential growth

8    phase, the GFP fluorescence accumulates and peaked at the early stationary phase

9    (Figure 4B). However, a clear difference in the number of GFP fluorescent or ComG-

10    operon expressing cells could be observed among the different strains. While with the

11    wild type, about 18% of the total cell counts showed fluorescence, this was

12    substantially higher with the *ΔspoIIIJ* strain MS007, i.e. around 50 %. In contrast,

13    only 11 % of the MS008 cells (*ΔyqjG*) were GFP-fluorescent. Furthermore, the

14    protein levels of ComGC and BdbD in *B. subtilis* 168, MS001 (*ΔspoIIIJ*) and MS003

15    (*ΔyqjG*) were monitored by immunoblotting of membranes derived from cells that

16    showed the maximal competence in the late exponential phase (Figure 4C). The

17    amount of BdbD, a thiol-disulfide oxidoreductase required for genetic competence

18    [37], remained unaltered. On the other hand, the level of ComGC, the major

19    pseudopilus forming subunit [38], was dramatically affected. As expected from the

20    proteomic and flow cytometric data, ComGC levels were increased in the *spoIIIJ*

21    deletion MS001 strain compared to the wild type, while they were markedly reduced

22    in the *yqjG* deletion MS003 strain. These data demonstrate that the ComG-operon

23    expression and transformability correlate and that natural competence is reduced by

24    deletion of *yqjG*. On the other hand natural competence is increased upon *spoIIIJ*

25    deletion. Also, in the *spoIIIJ* deletion strain the amount of YqjG is increased

1  approximately 20-fold suggesting that YqjG activity is a major controlling factor

2  contributing to the development of natural competence.

3

4

5  **4  Discussion**

6

7  In *Bacillus subtilis*, the general exchangeability of the two YidC homologues [14, 39-

8  41] and the induction of *yqjG* expression upon SpoIIIJ depletion [18, 20, 32] suggests

9  that YqjG serves as a back-up protein at limited SpoIIIJ activity to sustain viability of

10  cells. However, in contrast to other organisms that contain multiple YidC paralogs,

11  SpoIIIJ and YqjG exhibit a highly similar topology and organization. Often, one of

12  the paralogs contains an elongated C-terminus that has been implicated in ribosome-

13  binding and membrane targeting [4, 5]. Since this is not the case with the two YidC

14  homologues of *B. subtilis*, a different functional specialization is expected. To reveal a

15  possible YqjG-specific function, we performed proteomic studies on *yqjG* and *spoIIIJ*

16  single deletion strains as well as on *B. subtilis* cells that could be depleted for both

17  paralogous proteins.

18      In general, our data are consistent with a largely complementary function for

19  SpoIIIJ and YqjG. The growth rates of the single deletion strains (MS001 and

20  MS003) were essentially identical to that of the wild type (data not shown).

21  Furthermore, the expression levels of most quantified proteins did not show

22  significant changes (Figure 2A and B). Also the protein levels of various known YidC

23  substrates were found unchanged as well as the $F_1F_O$-ATPase activity. Only when

24  both SpoIIIJ and YqjG were depleted from the cells, major changes in protein

25  composition could be observed. In such cells, the majority of the identified membrane

1    proteins were found at lower levels in line with the proposed function of the YidC

2    homologues as membrane protein insertases [14]. The use of depletion strains of

3    essential proteins has a general disadvantage, as it might be difficult to distinguish

4    between primary and secondary effects. However, the remarkable correlation of our

5    results with previous findings favors the conclusion that the observed effect on the

6    membrane proteome is a direct result of a YidC deficiency. In addition, we noticed

7    that depletion of the YidC homologues from *B. subtilis* cells led to the similar stress

8    response as YidC depletion in *E. coli*. SpoIIIJ/YqjG double depletion resulted in an

9    immense (more than 110-fold) increase of LiaH. LiaH is homologous to the phage

10   shock protein (Psp) A, which is overexpressed in YidC depleted *E. coli* cells [9]. The

11   enhanced expression of PspA is a response to defects in the cell envelope integrity

12   [33].

13        In contrast to a previous study [18], the level of QoxA, a homologue of the *E.*

14   *coli* YidC substrate CyoA, was found to be reduced in cells lacking both SpoIIIJ and

15   YqjG. The other subunits of the cytochrome oxidase were found unchanged. In *E.*

16   *coli*, YidC is essential for membrane insertion of subunits a and c of the $F_1F_O$ ATP

17   synthase [6, 11] and both SpoIIIJ and YqjG were previously shown to mediate the

18   membrane insertion of *B. subtilis* and *E. coli* $F_Oc$ proteins [14]. Therefore, it was

19   rather surprising that upon SpoIIIJ and YqjG depletion the levels of the membrane

20   embedded $F_O$ domain remained essentially unaltered whereas there was a marked

21   reduction in the amount of membrane-associated $F_1$ domain. The lack of a clear effect

22   on the $F_O$ domain could be due to a transcriptional up-regulation response of *atpB* and

23   *atpE* ($F_Oa$ and $F_Oc$) upon SpoIIIJ and YqjG depletion as reported in *E. coli* [42]. The

24   loss of $F_1$ subunits and consequently, the DCCD-sensitive $F_1F_O$-ATPase activity

1    suggests that there is indeed a functional defect in the assembly further supporting the

2    notion that SpoIIIJ and YqjG fulfill a key role in ATP synthase biogenesis [14].

3    The overlapping functions of SpoIIIJ and YqjG support cell viability under

4    vegetative growth, since expression of one of the two proteins is sufficient for growth

5    of *B. subtilis* [16, 18]. However, during sporulation SpoIIIJ mediates a crucial

6    function in spore formation, while YqjG is dispensable in this process [17]. Several

7    studies have been performed in order to elucidate the specific requirements for

8    SpoIIIJ during sporulation [43-45]. It is believed that SpoIIIJ plays an exclusive role

9    in the assembly of a novel type of secretion tunnel, which is composed of proteins

10   encoded by the *spoIIIA* operon and that spans the space between inner and outer pre-

11   spore membrane. As shown recently, SpoIIIJ physically interacts with the membrane

12   protein SpoIIIAE [45], which is speculated to be a SpoIIIJ specific substrate. Our data

13   indicate a specific role of YqjG in the development of natural competence and

14   transformability. Although the effects of YqjG on the membrane protein amounts

15   appear subtle, there is an obvious and significant correlation between YqjG

16   expression and genetic competence. It should be noted that the quantitative proteomic

17   analysis were carried out under conditions which are unfavorable for competence

18   development (minimal medium, $OD_{600}$ of 1), while transformability, Western blotting

19   and ComG-GFP expression (Figure 4) were determined under optimal competence

20   conditions (CDCM at late stationary phase). This difference explains the smaller

21   changes in protein levels as observed in the MS analysis as opposed to the substantial

22   changes in the individual protein detection assays. Deletion of *yqjG* resulted in a 50 %

23   reduction of transformability while YqjG overexpression (*spoIIIJ* deletion) caused a

24   3-fold increase of transformation efficiency. Although YqjG is not essential for

1    competence development, its expression clearly enhances the cells' ability for

2    transformation and stable integration.

3         How does YqjG affect genetic competence? The first major process in protein

4    biogenesis is transcription, which in prokaryotes involves binding of a sigma factor

5    and subsequent targeting of RNA polymerase to the gene locus. It has been shown

6    that ComK is the key transcription factor for competence genes in *B. subtilis* [46]. It

7    could well be that YqjG is involved in ComK activation. A similar function has been

8    suggested for SpoIIIJ during sporulation, where the presence of SpoIIIJ is essential for

9    the activation of the pre-spore specific sigma factor G [44]. Amongst the genes under

10   control of ComK are the components of the ComG operon. They belong to the group

11   of late competence genes and are essential for exogenous DNA binding and transport

12   to the membrane. Our flow cytometric analysis demonstrates that the transcription of

13   the *comG* operon correlates with the YqjG expression level, which provides further

14   evidence that YqjG is involved in a regulatory function although this might be an

15   indirect response due to the competence defect. Enhanced gene transcription does not

16   necessarily result in increased protein levels, since other processes might be limiting

17   in protein biogenesis, e.g. translation, targeting and posttranslational modifications.

18   The ComG proteins together with the oxidoreductase BdbCD and the prepilin

19   peptidase ComC are needed for the formation of the pilin-like structure involved in

20   DNA uptake. ComGC is the major pilin subunit that shows similarity to type 2

21   secretion system pseudopilins and type 4 pilins. The prepilins are first inserted into

22   the membrane, and upon the removal of the charged N-terminus by ComC, BdbCD

23   introduce disulphide bonds into the maturating pilin subunits whereupon these are

24   assembled to the pilus [38]. Even though the level of BdbD remained unaltered upon

25   *yqjG* deletion, only minute traces of ComGC could be detected in the membrane.

_____

1     YqjG might be required for the ComGC membrane insertion, an activity that

2     apparently cannot be complemented by SpoIIIJ. Another mechanism by which YqjG

3     may affect competence development follows from findings with *Enterococcus*

4     *faecalis* [47]. The N-terminus of the *E. faecalis* YidC, designated as CcfA, is

5     predicted to be lipid modified just like the N-termini of SpoIIIJ and YqjG. After N-

6     terminal modifications including processing by signal peptidase II, a small peptide

7     with the amino acid sequence LVTLVFV, is released from the cell. This peptide acts

8     as a pheromone and causes the induction of conjugative plasmid transfer between *E.*

9     *faecalis* cells. The amino acid sequence of the N-termini of SpoIIIJ and YqjG differ

10     from this exact pheromone peptide sequence. Moreover, SpoIIIJ does not seem to

11     encode for a pheromone peptide as no effect on sporulation in *B. subtilis* nor on the

12     aggregation of *E. faecalis* cells could be observed [44, 47]. Since the *yqjG* expression

13     in these SpoIIIJ expressing cells is below detection levels [32], a possible pheromone

14     activity associated with YqjG cannot be excluded. It would be of interest to test this

15     hypothesis with a strain overexpressing YqjG, such as the *spoIIIJ* deletion strain.

16         Our data demonstrate that SpoIIIJ and YqjG fulfill complementary and

17     essential functions during vegetative growth. Whereas SpoIIIJ is known to fulfill a

18     specific function in a late stage of sporulation, we now show for the first time that

19     YqjG is specifically involved in genetic competence development. Therefore, we

20     conclude that the two YidC homologs are functionally specialized in distinct cellular

21     differentiation processes in *B. subtilis*.

22

3

4  *The authors have declared no conflict of interest.*

5

6

7

8

9

10

11  **Legends to the Figures**

12

13  FIG. 1. Absence of SpoIIIJ and YqjG induces LiaH response and reduces QoaX

14  levels. The protein compositions of membranes isolated from *B. subtilis* 168 (lane 1),

15  MS001 (lane 2), MS003 (lane 3), MS004 (lane 4) and MS005 (lane 5) were visualized

16  by SDS-PAGE and Coomassie Brilliant Blue staining. The membrane levels of LiaH,

17  PrsA and QoaX were detected by immunoblotting using specific antisera.

18

19  FIG. 2. *B. subtilis* cells depleted of SpoIIIJ and YqjG show accumulation of

20  cytoplasmic and a decrease of membrane proteins. Changes of protein expression

21  levels were plotted in log2 scale. Proteins with $\geq$ 2- fold expression increase are

22  shown in green, with $\leq$ 2-fold decrease in red. A: MS001($\Delta spoIIIJ$)/WT; B:

23  MS003($\Delta yqjG$)/WT; C: MS004($\Delta yqjG$, $\Delta spoIIIJ$, $P_{xylA}$ $spoIIIJ$)/MS003($\Delta yqjG$). The

24  predicted cellular localization of the proteins is depicted in the insets of each sub-

25  figure. Yellow: cytoplasmic proteins, blue: membrane proteins, orange: secreted

26  proteins.

27

28  FIG. 3. Effect of SpoIIIJ and YqjG depletion on the $F_1F_O$-ATPase activity in

29  isolated membranes. The activity of the $F_1F_O$-ATPase was determined by the release

30  of inorganic phosphate in the presence (white bars) or absence (black bars) of 0.5 mM

31  DCCD.

32

1    FIG. 4. Differential effect of *spoIIIJ*- and *yqjG*-deletion on genetic competence

2    development. A: Transformation frequencies of *B. subtilis* 168 (wild type), MS001

3    (Δ*spoIIIJ*) and MS003 (Δ*yqjG*); and in *B. subtilis* MS006, MS007 and MS008 bearing

4    the *comG*-*gfp* insertion. Cells were transformed with either plasmid or genomic DNA

5    and the transformation frequency was calculated. B: Expression of the *comG*-operon

6    in *B. subtilis* MS006 (WT, *comG*-*gfp*), MS007 (Δ*spoIIIJ*, *comG*-*gfp*) and MS008

7    (Δ*yqjG*, *comG*-*gfp*) was monitored by fluorescence measurements in a flow

8    cytometer. The diagram shows the maximum fluorescence value reached at late

9    stationary growth phase. C: Immunoblot detection of ComCG and BdbD with known

10   functions in genetic competence in membranes isolated from *B. subtilis* 168 (lane 1),

11   MS001 (lane 2) and MS003 (lane 3).

12

1    **Table 1**. Primers.

2

| Primer name | Sequence 5'→3' |
|---|---|
| Spod1 | GGTTGCAAGAAATAAAAGCA |
| Spod2 | GCGTCGACGTCATATGGATCCACTCTTTTATTGCAAAGCA |
| Spod3 | CTAGAATTCGAGCTCCCGGGTTTCCTCATTAAAGGACCTG |
| Spod4 | ACAGAAATAGAAATTCGTCC |
| YqjGd1 | TAAGATCGTAATTCCAAGCG |
| YqjGd2-Spec | CGAAAATCGCCATTCGCCAGCCTCCTTTTTATAAATGCGG |
| YqjGd3-Spec | AACCCTTGCATAGGGGGATCTTGACAGTGCAAAACATCGT |
| YqjGd4 | TTTAGCATAATGAAGCAAGG |
| FoSpoBamHI | GGAGGATCCATGTTGTTGAAAAGGAGAATAGG |
| ReSpoBamHI | GCGGATCCTCACTTTTTCTTTCCTCCGGCTTTTTGC |
| FoYqjGBamHI | GGAGGATCCATGTTAAAAACATATCAAAAACTTTTGGC |
| ReYqjGBamHI | GCGGATCCTTATTTCACCGACTCAGTAAGAGC |
| Tet1 | GATCCATATGACGTCGACGC |
| Tet2 | CCCGGGAGCTCGAATTCTAG |
| Spec1 | CTGGCGAATGGCGATTTTCG |
| Spec2 | GATCCCCCTATGCAAGGGTT |

3
4
5    **Table 2**. Strains and plasmids.
6

| Plasmids | Relevant properties | Reference |
|---|---|---|
| pAX01 | integrative vector (pBR322 derivate) for xylose inducible expression of genes (P$_{XylA}$ promoter) in *B. subtilis*, Em$^r$ | [48] |
| pAX-spoIIIJ | pAX01 carrying *B. subtilis* *spoIIIJ* | This study |
| pAX-yqjG | pAX01 carrying *B. subtilis* *yqjG* | This study |
| pDG1514 | contains tetracycline resistance gene cassette | [49] |
| pDG1726 | contains spectinomycin resistance gene cassette | [49] |
| pSGComGA1 | P$_{comG}$-*gfp*, Kan$^r$ | W.K. Smits unpublished |
| pTRKH2 | *E. coli* gram-positive cloning vector, Em$^r$ | [50] |
| pNZ8901 | SURE expression vector, P$_{spaSmut}$, Cm$^r$ | [51] |
| **Strains** | | |
| *B. subtilis* | | |
| 168 | trpC2 | [52] |
| MS001 | 168, *spoIIIJ*::Tet$^r$ | This study |
| MS003 | 168, *yqjG*::Spec$^r$ | This study |
| MS004 | MS001, *yqjG*::Spec$^r$, *lacA*:: P$_{xylA}$-*spoIIIJ*, Em$^r$ | This study |
| MS005 | MS001, *yqjG*:: Spec$^r$, *lacA*:: P$_{xylA}$-*yqjG*, Em$^r$ | This study |
| MS006 | 168, P$_{comG}$-*gfp*, Kan$^r$ | This study |
| MS007 | MS001, P$_{comG}$-*gfp*, Kan$^r$ | This study |
| MS008 | MS003, P$_{comG}$-*gfp*, Kan$^r$ | This study |

7
8

1    **Table 3**. Identified proteins with expression change in *B. subtilis* MS001 (*ΔspoIIIJ*)
2    and MS003 (*ΔyqjG*).
3
4

| Protein ID\|gene | Protein | Localization | Fold change |
|---|---|---|---|
| **MS001** | | | **MS001/WT** |
| P54544\|oxaA2 | YqjG | Multi-TM's | 19.97 |
| P49785\|spoIIIAH | Stage III sporulation protein AH | N-term. anchor | 6.32 |
| P39696\|comER | Unknown | Cytoplasmic | 5.82 |
| P26936\|spoIVFA | Stage IV sporulation protein FA | Cytoplasmic | 5.46 |
| P39072\|gerM | Spore germination protein | Lipid anchor | 4.20 |
| O05267\|yumB | NADH dehydrogenase-like protein | Cytoplasmic | 3.12 |
| P96718\|ywqF | Major UDP-glucose dehydrogenase | Cytoplasmic | 3.10 |
| P35149\|spoIVA | Stage IV sporulation protein A | Cytoplasmic | 3.03 |
| P37963\|spoVID | Stage VI sporulation protein D | Cytoplasmic | 2.79 |
| P54465\|yqeZ | Unknown | Multi-TM's | 2.77 |
| O34431\|yloB | Calcium-transporting ATPase | Multi-TM's | 2.69 |
| P96619\|ydcC | Sporulation protein. sigma E promotor controlled | Lipid anchor | 2.55 |
| P46338\|pstS | Phosphate ABC transporter (binding protein) | Lipid anchor | 2.55 |
| O32087\|yuzA | Unknown | Multi-TM's | 2.39 |
| Q795Q5\|yttA | Unknown | Multi-TM's | 2.36 |
| P54598\|yhcN | Inner spore membrane protein | Lipid anchor | 2.16 |
| P38050\|pbpF | Penicillin-binding protein 2C | N-term. anchor | 2.10 |
| P21474\|rpsP | 30S ribosomal protein S16 | Cytoplasmic | 2.08 |
| O31707\|yknU | Uncharacterized ABC transporter (ATP-binding protein) | Multi-TM's | 2.01 |
| O05250\|malK | Malate kinase sensor.YufL | Multi-TM's | -2.08 |
| P45861\|ywjA | Uncharacterized ABC transporter (ATP-binding protein) | Multi-TM's | -2.36 |
| | | | |
| ON/OFF proteins | | | |
| Q01625\|oxaA1 | SpoIIIJ | Multi-TM's | OFF in MS001 |
| O31788\|aprX | Alkaline serine protease | Cytoplasmic | OFF in MS001 |
| O05272\|asnO | Asparagine synthetase | Cytoplasmic | OFF in WT |
| | | | |
| **MS003** | | | **MS003/WT** |
| O31788\|aprX | Alkaline serine protease | Cytoplasmic | 2.60 |
| Q06797\|rplA | 50S ribosomal protein L1 | Cytoplasmic | 2.53 |
| O07621\|hemAT | Haem-based aerotactic transducer | Cytoplasmic | -2.92 |
| P94590\|ywpH | Single-strand DNA-binding protein | Cytoplasmic | -2.51 |

5
6
7
8
9

1 **Table 4.** Effect of SpoIIIJ and YqjG-depletion on the stress response protein levels at
2 the membrane.
3
4

| Protein ID \|gene | Protein | Fold change | | |
|---|---|---|---|---|
| | | MS001 / 168 | MS003 / 168 | MS004 / MS003 |
| O32201\|liaH | Cell wall stress response protein | no ID | no ID | 112.21 |
| O32199\|liaF | Cell wall stress response protein | no ID | no ID | 16.68 |
| O32198\|liaS | Sensor histidine kinase | no ID | no ID | 9.58 |
| P80244\|clpP | ATP-dependent Clp protease proteolytic subunit | 1.09 | 1.02 | 5.13 |
| P50866\|clpX | ATP-dependent Clp protease ATP-binding subunit | -1.01 | 1.24 | 4.76 |
| O31657\|htpX | Heat-shock protein | 1.06 | 1.01 | 4.56 |
| P54617\|ydjF | Phage shock protein A homolog | -1.13 | -1.41 | 4.29 |
| O32000\|yokG | Delta-endotoxin | no ID | no ID | 2.51 |
| P94520\|ysdB | Sigma-w pathway protein. cell envelope stress | -1.03 | no ID | 1.65 |
| P37571\|clpC | Class III stress response-related ATPase | -1.05 | 1.20 | 1.58 |
| P17820\|dnaK | Class I heat-shock protein (molecular chaperone) | 1.03 | -1.03 | 1.27 |
| Q45539\|csbB | Stress response protein | -1.41 | -1.19 | 1.18 |
| P40780\|ytxH | General stress protein | 1.15 | 1.06 | -1.09 |
| P37476\|ftsH | General stress protein (class III heat-shock) | 1.06 | 1.02 | -1.15 |
| P80875\|yceD | General stress protein 16U | no ID | no ID | -1.16 |
| P40779\|ytxG | General stress protein | 1.13 | 1.15 | -1.33 |
| P17631\|dnaJ | Heat-shock protein | no ID | no ID | -1.33 |
| P28599\|groS | 10 kDa chaperonin | no ID | no ID | -1.46 |
| P28598\|groL | Class I heat-shock protein (chaperonin) | 1.08 | -1.14 | -1.87 |
| P71051\|yveL | Putative tyrosine-protein kinase | no ID | -1.28 | no ID |
| P39778\|hslU | ATP-dependent protease ATP-binding subunit | 1.13 | no ID | no ID |
| P94362\|yxkI | Heat-shock protein | -1.13 | no ID | no ID |

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

1 **Table 5.** Effect of SpoIIIJ and YqjG-depletion on the levels of the subunits of the
2 $F_1F_0$-ATPase at the membrane.
3
4

| Protein ID \|gene | Protein | Localization | Fold change | | |
|---|---|---|---|---|---|
| | | | MS001 / 168 | MS003 / 168 | MS004/ MS003 |
| P37808\|atpA | ATP synthase $F_1$ subunit α | Cytoplasmic | 1.20 | 1.27 | **-3.78** |
| P37809\|atpD | ATP synthase $F_1$ subunit β | Cytoplasmic | 1.17 | 1.27 | **-3.89** |
| P37810\|atpG | ATP synthase $F_1$ subunit γ | Cytoplasmic | 1.18 | 1.64 | **-2.93** |
| P37811\|atpH | ATP synthase $F_1$ subunit δ | Cytoplasmic | 1.16 | 1.35 | **-3.18** |
| P37812\|atpC | ATP synthase $F_1$ subunit ε | Cytoplasmic | 1.14 | 1.17 | **-4.72** |
| P37813\|atpB | ATP synthase $F_0$ subunit a | Multi-TM's | 1.19 | no ID | -1.31 |
| P37814\|atpF | ATP synthase $F_0$ subunit b | N-term. anchor | 1.12 | 1.20 | -1.40 |
| P37815\|atpE | ATP synthase $F_0$ subunit c | Multi-TM's | no ID | no ID | 1.66 |

5
6
7 **Table 6.** Effect of SpoIIIJ and YqjG-depletion on the level of proteins involved in
8 competence development at the membrane.
9
10

| Protein ID \|gene | Protein | Fold change | |
|---|---|---|---|
| | | MS001 / 168 | MS003 / 168 |
| Q99027\|comP | Two-component sensor histidine kinase | no ID | no ID |
| P39694\|comEA | Unspecific high-affinity DNA-binding protein | 1.59 | -1.80 |
| P39145\|comFA | ATP dependent helicase | 1.47 | -1.38 |
| P25953\|comGA | Traffic NTPase | 1.09 | -1.42 |
| P25955\|comGC | Major pseudopilin | 1.36 | -1.68 |
| P25959\|comGG | Minor pseudopilin | 1.81 | no ID |
| P68569\|bdbA | Thiol-disulfide oxidoreductase | 1.23 | -1.32 |
| O32218\|bdbD | Thiol-disulfide oxidoreductase | 1.12 | no ID |
| P12667\|nucA | Membrane-associated nuclease | 1.30 | -1.54 |
| P12669\|nin | Inhibitor of the DNA degrading activity of NucA | 1.29 | -1.16 |
| P16971\|recA | Multifunctional SOS repair regulator | 1.34 | -1.41 |
| P94590\|ywpH | Single-strand DNA-binding protein | 1.69 | -2.52 |
| O06973\|yvcJ | Nucleotide-binding protein | no ID | no ID |

11
12
13

14

15

16

17
18

1    **5    References**
2
3

4

5    [1]    Yen, M. R., Harley, K. T., Tseng, Y. H., and Saier, M. H., Jr.,  Phylogenetic and
6           structural analyses of the oxa1 family of protein translocases. *FEMS Microbiol.*
7           *Lett.*2001, 204, 223-231.

8    [2]    Jiang, F., Chen, M., Yi, L., de Gier, J. W. *et al.*,  Defining the regions of
9           *Escherichia coli* YidC that contribute to activity. *J. Biol. Chem.*2003, 278,
10          48965-48972.

11   [3]    Jia, L., Dienhart, M., Schramp, M., McCauley, M. *et al.*,  Yeast Oxa1 interacts
12          with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1.
13          *EMBO J.*2003, 22, 6438-6447.

14   [4]    Falk, S., Ravaud, S., Koch, J., and Sinning, I.,  The C terminus of the Alb3
15          membrane insertase recruits cpSRP43 to the thylakoid membrane. *J. Biol.*
16          *Chem.*2010, 285, 5954-5962.

17   [5]    Kohler, R., Boehringer, D., Greber, B., Bingel-Erlenmeyer, R. *et al.*,  YidC and
18          Oxa1 form dimeric insertion pores on the translating ribosome. *Mol. Cell*2009,
19          34, 344-353.

20   [6]    van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N. *et al.*,  $F_1F_O$ ATP synthase
21          subunit c is a substrate of the novel YidC pathway for membrane protein
22          biogenesis. *J. Cell Biol.*2004, 165, 213-222.

1    [7]    du Plessis, D. J., Nouwen, N., and Driessen, A. J. M.,  Subunit a of cytochrome

2          o oxidase requires both YidC and SecYEG for membrane insertion. *J. Biol.*

3          *Chem.*2006, 281, 12248-12252.


4    [8]    van Bloois, E., Haan, G. J., de Gier, J. W., Oudega, B. *et al.*,  Distinct

5          requirements for translocation of the N-tail and C-tail of the *Escherichia coli*

6          inner membrane protein CyoA. *J. Biol. Chem.*2006, 281, 10002-10009.


7    [9]    van der Laan, M., Urbanus, M. L., ten Hagen-Jongman, C. M., Nouwen, N. *et*

8          *al.*,  A conserved function of YidC in the biogenesis of respiratory chain

9          complexes. *Proc. Natl. Acad. Sci. U. S. A.*2003, 100, 5801-5806.


10   [10]   Facey, S. J., Neugebauer, S. A., Krauss, S., and Kuhn, A.,  The

11         mechanosensitive channel protein MscL is targeted by the SRP to the novel

12         YidC membrane insertion pathway of *Escherichia coli. J. Mol. Biol.*2007, 365,

13         995-1004.


14   [11]   Kol, S., Majczak, W., Heerlien, R., van der Berg, J. P. *et al.*,  Subunit a of the

15         $F_1F_o$ ATP synthase requires YidC and SecYEG for membrane insertion. *J. Mol.*

16         *Biol.*2009, 390, 893-901.


17   [12]   Price, C. E. and Driessen, A. J.,  YidC is involved in the biogenesis of anaerobic

18         respiratory complexes in the inner membrane of *Escherichia coli. J. Biol.*

19         *Chem.*2008, 283, 26921-26927.

1    [13] Price, C. E. and Driessen, A. J., Conserved negative charges in the

2         transmembrane segments of subunit K of the NADH:ubiquinone oxidoreductase

3         determine its dependence on YidC for membrane insertion. *J. Biol. Chem.*2010,

4         285, 3575-3581.

5    [14] Saller, M. J., Fusetti, F., and Driessen, A. J., *Bacillus subtilis* SpoIIIJ and YqjG

6         function in membrane protein biogenesis. *J. Bacteriol.*2009, 191, 6749-6757.

7    [15] Jia, L., Dienhart, M. K., and Stuart, R. A., Oxa1 directly interacts with Atp9 and

8         mediates its assembly into the mitochondrial $F_1F_O$ ATP synthase complex. *Mol.*

9         *Biol. Cell*2007, 18, 1897-1908.

10   [16] Murakami, T., Haga, K., Takeuchi, M., and Sato, T., Analysis of the *Bacillus*

11        *subtilis spoIIIJ* gene and its paralogue gene, *yqjG. J. Bacteriol.*2002, 184, 1998-

12        2004.

13   [17] Errington, J., Appleby, L., Daniel, R. A., Goodfellow, H. *et al.*, Structure and

14        function of the *spoIIIJ* gene of *Bacillus subtilis*: a vegetatively expressed gene

15        that is essential for sigma G activity at an intermediate stage of sporulation. *J.*

16        *Gen. Microbiol.*1992, 138, 2609-2618.

17   [18] Tjalsma, H., Bron, S., and van Dijl, J. M., Complementary impact of paralogous

18        Oxa1-like proteins of *Bacillus subtilis* on post-translocational stages in protein

19        secretion. *J. Biol. Chem.*2003, 278, 15622-15632.

1    [19]  Thomaides, H. B., Davison, E. J., Burston, L., Johnson, H. *et al.*,  Essential

2          bacterial functions encoded by gene pairs. *J. Bacteriol.*2007, 189, 591-602.

3    [20]  Chiba, S., Lamsa, A., and Pogliano, K.,  A ribosome-nascent chain sensor of

4          membrane protein biogenesis in *Bacillus subtilis. EMBO J.*2009, 28, 3461-3475.

5    [21]  Stulke, J., Hanschke, R., and Hecker, M.,  Temporal activation of beta-glucanase

6          synthesis in *Bacillus subtilis* is mediated by the GTP pool. *J. Gen.*

7          *Microbiol.*1993, 139, 2041-2045.

8    [22]  Smits, W. K., Dubois, J. Y., Bron, S., van Dijl, J. M. *et al.*,  Tricksy business:

9          transcriptome analysis reveals the involvement of thioredoxin A in redox

10         homeostasis, oxidative stress, sulfur metabolism, and cellular differentiation in

11         *Bacillus subtilis. J. Bacteriol.*2005, 187, 3921-3930.

12   [23]  Bron, S. and Venema, G.,  Ultraviolet inactivation and excision-repair in

13         *Bacillus subtilis*. I. Construction and characterization of a transformable

14         eightfold auxotrophic strain and two ultraviolet-sensitive derivatives. *Mutat.*

15         *Res.*1972, 15, 1-10.

16   [24]  Smits, W. K., Eschevins, C. C., Susanna, K. A., Bron, S. *et al.*,  Stripping

17         *Bacillus*: ComK auto-stimulation is responsible for the bistable response in

18         competence development. *Mol. Microbiol.*2005, 56, 604-614.

1    [25] Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J. *et al.*,  A comprehensive

2        proteome map of growing *Bacillus subtilis* cells. *Proteomics.*2004, 4, 2849-

3        2876.


4    [26] Dreisbach, A., Otto, A., Becher, D., Hammer, E. *et al.*,  Monitoring of changes

5        in the membrane proteome during stationary phase adaptation of *Bacillus*

6        *subtilis* using in vivo labeling techniques. *Proteomics.*2008, 8, 2062-2076.


7    [27] Park, S. K., Venable, J. D., Xu, T., and Yates, J. R., III,  A quantitative analysis

8        software tool for mass spectrometry-based proteomics. *Nat. Methods*2008, 5,

9        319-322.


10   [28] MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R. *et al.*,  A correlation algorithm

11       for the automated quantitative analysis of shotgun proteomics data. *Anal.*

12       *Chem.*2003, 75, 6912-6921.


13   [29] Zhou, M., Boekhorst, J., Francke, C., and Siezen, R. J.,  LocateP: genome-scale

14       subcellular-location predictor for bacterial proteins. *BMC. Bioinformatics.*2008,

15       9, 173.


16   [30] Lanzetta, P. A., Alvarez, L. J., Reinach, P. S., and Candia, O. A.,  An improved

17       assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.*1979, 100,

18       95-97.

1 [31] Veening, J. W., Hamoen, L. W., and Kuipers, O. P., Phosphatases modulate the

2 bistable sporulation gene expression pattern in *Bacillus subtilis. Mol.*

3 *Microbiol.*2005, 56, 1481-1494.

4 [32] Rubio, A., Jiang, X., and Pogliano, K., Localization of translocation complex

5 components in *Bacillus subtilis*: enrichment of the signal recognition particle

6 receptor at early sporulation septa. *J. Bacteriol.*2005, 187, 5000-5002.

7 [33] Jordan, S., Junker, A., Helmann, J. D., and Mascher, T., Regulation of LiaRS-

8 dependent gene expression in *Bacillus subtilis*: identification of inhibitor

9 proteins, regulator binding sites, and target genes of a conserved cell envelope

10 stress-sensing two-component system. *J. Bacteriol.*2006, 188, 5153-5166.

11 [34] Kontinen, V. P. and Sarvas, M., The PrsA lipoprotein is essential for protein

12 secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Mol.*

13 *Microbiol.*1993, 8, 727-737.

14 [35] Hsiao, T. L., Revelles, O., Chen, L., Sauer, U. *et al.*, Automatic policing of

15 biochemical annotations using genomic correlations. *Nat. Chem. Biol.*2010, 6,

16 34-40.

17 [36] Gyan, S., Shiohira, Y., Sato, I., Takeuchi, M. *et al.*, Regulatory loop between

18 redox sensing of the NADH/NAD(+) ratio by Rex (YdiH) and oxidation of

19 NADH by NADH dehydrogenase Ndh in *Bacillus subtilis. J. Bacteriol.*2006,

20 188, 7062-7071.

1    [37]  Meima, R., Eschevins, C., Fillinger, S., Bolhuis, A. *et al.*,  The *bdbDC* operon of

2         *Bacillus subtilis* encodes thiol-disulfide oxidoreductases required for

3         competence development. *J. Biol. Chem.*2002, 277, 6994-7001.

4    [38]  Chen, I., Provvedi, R., and Dubnau, D.,  A macromolecular complex formed by

5         a pilin-like protein in competent *Bacillus subtilis. J. Biol. Chem.*2006, 281,

6         21720-21727.

7    [39]  Jiang, F., Yi, L., Moore, M., Chen, M. *et al.*,  Chloroplast YidC homolog

8         Albino3 can functionally complement the bacterial YidC depletion strain and

9         promote membrane insertion of both bacterial and chloroplast thylakoid

10        proteins. *J. Biol. Chem.*2002, 277, 19281-19288.

11   [40]  Preuss, M., Ott, M., Funes, S., Luirink, J. *et al.*,  Evolution of mitochondrial oxa

12        proteins from bacterial YidC. Inherited and acquired functions of a conserved

13        protein insertion machinery. *J. Biol. Chem.*2005, 280, 13004-13011.

14   [41]  van Bloois, E., Koningstein, G., Bauerschmitt, H., Herrmann, J. M. *et al.*,

15        *Saccharomyces cerevisiae* Cox18 complements the essential Sec-independent

16        function of *Escherichia coli* YidC. *FEBS J.*2007, 274, 5704-5713.

17   [42]  Wang, P., Kuhn, A., and Dalbey, R. E.,  Global change of gene expression and

18        cell physiology in YidC-depleted *Escherichia coli. J. Bacteriol.*2010, 192, 2193-

19        2209.

1    [43]   Camp, A. H. and Losick, R., A novel pathway of intercellular signalling in

2          *Bacillus subtilis* involves a protein with similarity to a component of type III

3          secretion channels. *Mol. Microbiol.*2008, 69, 402-417.

4    [44]   Serrano, M., Corte, L., Opdyke, J., Moran, C. P., Jr. *et al.*, Expression of *spoIIIJ*

5          in the prespore is sufficient for activation of sigma G and for sporulation in

6          *Bacillus subtilis*. *J. Bacteriol.*2003, 185, 3905-3917.

7    [45]   Serrano, M., Vieira, F., Moran, C. P., Jr., and Henriques, A. O., Processing of a

8          membrane protein required for cell-to-cell signaling during endospore formation

9          in *Bacillus subtilis*. *J. Bacteriol.*2008, 190, 7786-96.

10    [46]   van Sinderen D., Luttinger, A., Kong, L., Dubnau, D. *et al.*, *comK* encodes the

11          competence transcription factor, the key regulatory protein for competence

12          development in *Bacillus subtilis*. *Mol. Microbiol.*1995, 15, 455-462.

13    [47]   Antiporta, M. H. and Dunny, G. M., *ccfA*, the genetic determinant for the

14          cCF10 peptide pheromone in *Enterococcus faecalis* OG1RF. *J. Bacteriol.*2002,

15          184, 1155-1162.

16    [48]   Hartl, B., Wehrl, W., Wiegert, T., Homuth, G. *et al.*, Development of a new

17          integration site within the *Bacillus subtilis* chromosome and construction of

18          compatible expression cassettes. *J. Bacteriol.*2001, 183, 2696-2699.

19    [49]   Guerout-Fleury, A. M., Shazand, K., Frandsen, N., and Stragier, P., Antibiotic-

20          resistance cassettes for *Bacillus subtilis*. *Gene*1995, 167, 335-336.

1    [50]  O'Sullivan, D. J. and Klaenhammer, T. R.,  High- and low-copy-number

2          *Lactococcus* shuttle cloning vectors with features for clone screening.

3          *Gene*1993, 137, 227-231.


4    [51]  Bongers, R. S., Veening, J. W., Van Wieringen, M., Kuipers, O. P. *et al.*,

5          Development and characterization of a subtilin-regulated expression system in

6          *Bacillus subtilis*: strict control of gene expression by addition of subtilin. *Appl.*

7          *Environ. Microbiol.*2005, 71, 8818-8824.


8    [52]  Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M. *et al.*,  The complete

9          genome sequence of the gram-positive bacterium *Bacillus subtilis. Nature*1997,

10         390, 249-256.

11

12

13

Figure 1

_____

Figure 2

Figure 3

A

B

Relative cell count

GFP

MS008 (ΔyqjG)

MS007 (ΔspoIIIJ)

MS006 (wild type)

Fluorescence intensity [a.u.]

C

BdbD

ComGC

| Lane | 1 | 2 | 3 |
| --- | --- | --- | --- |
|  | WT | MS001 | MS003 |
|  |  | (ΔspoIIIJ) | (ΔyqjG) |
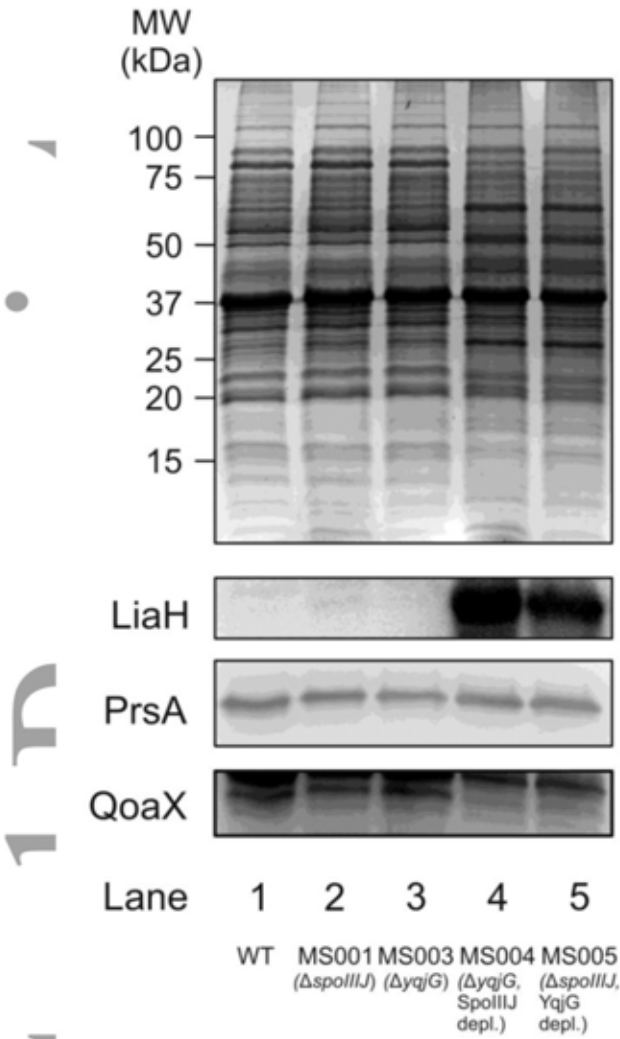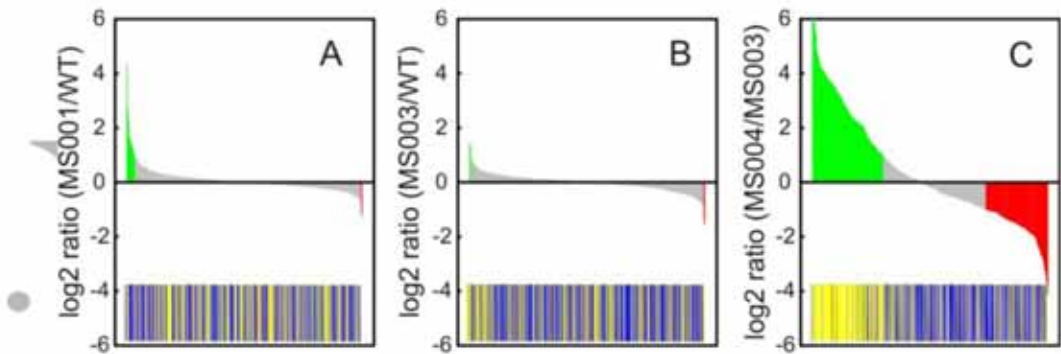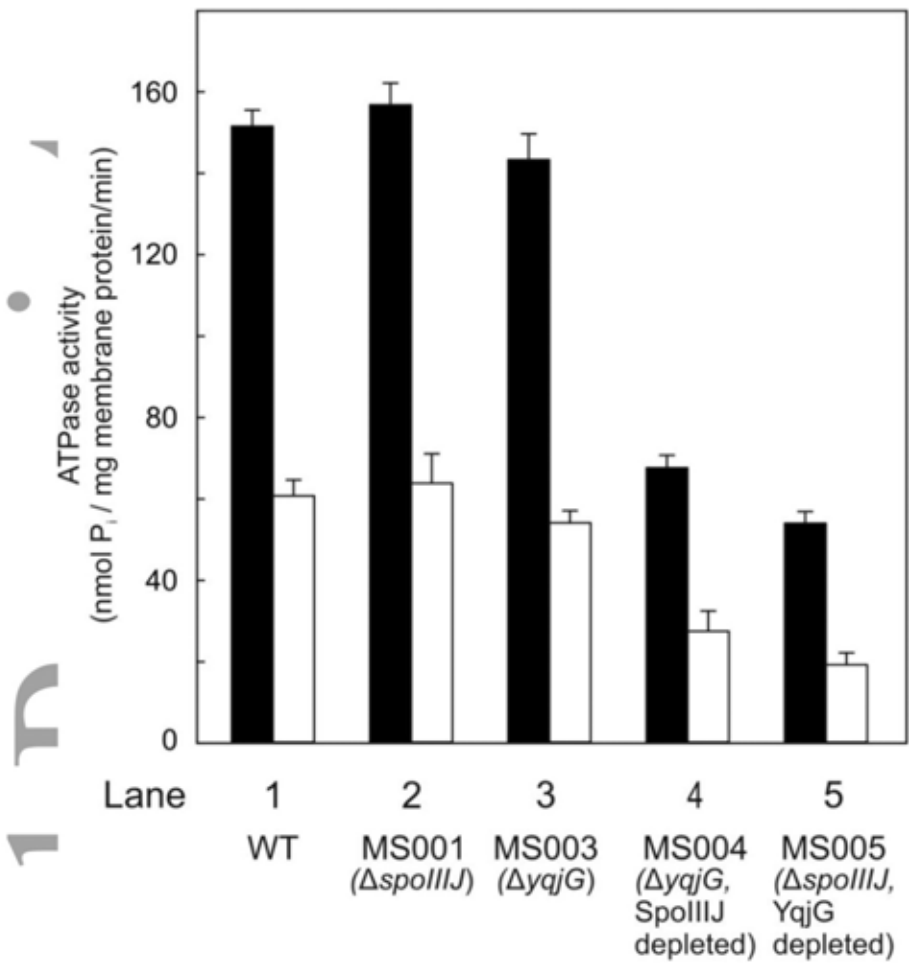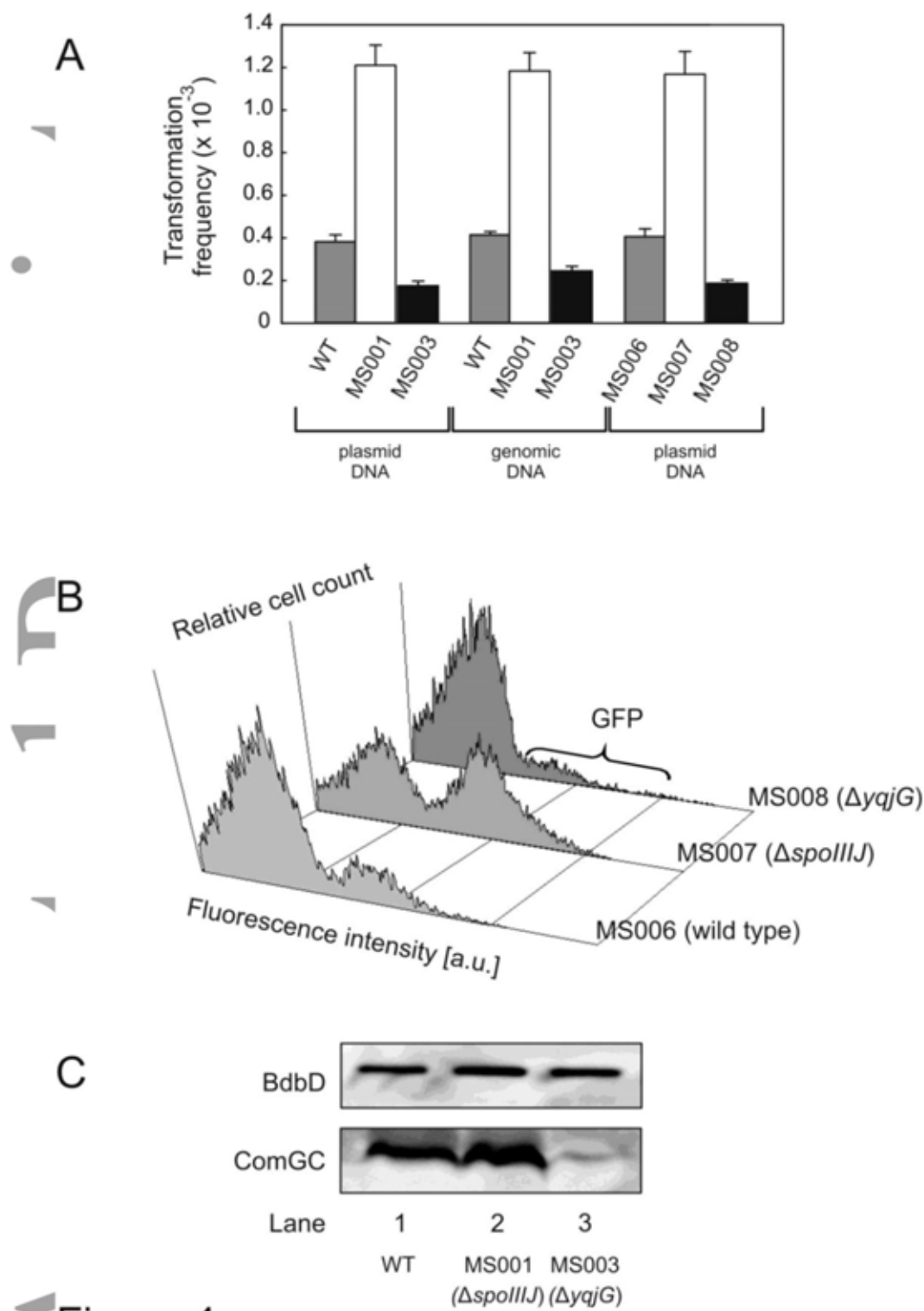
Figure 4

1

2

_____

*Quantitative Cell Surface Profiling for SigB-Dependent Protein Expression in the Human Pathogen Staphylococcus aureus via Biotinylation approach*

# Quantitative Cell Surface Proteome Profiling for SigB-Dependent Protein Expression in the Human Pathogen *Staphylococcus aureus* via Biotinylation Approach

Kristina Hempel, Jan Pané-Farré, Andreas Otto, Susanne Sievers, Michael Hecker, and Dörte Becher*

*Institute for Microbiology, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany*

Most of the *Staphylococcus aureus* virulence factors are either cell surface exposed or secreted. Here we report a global and quantitative analysis of staphylococcal cell surface-associated proteins using a combination of $^{14}N^{15}N$ metabolic labeling, biotinylation, and GeLC−MS/MS. To address the important question of *S. aureus* pathophysiology, we analyzed the influence of the alternative sigma factor $\sigma^B$ on the expression of cell surface-associated proteins. Therefore, we compared the methicillin-resistant *S. aureus* wild-type strain COL with its *sigB* mutant, because $\sigma^B$ might play a crucial role in the pattern of the surface proteome. A total of 296 proteins from growing and nongrowing cells could be quantified. One third of these proteins are known as cell surface-associated, including 3 sortase substrates, 3 cell wall-associated proteins, 35 lipo-, 23 membrane-, and 34 signal peptide-containing proteins comparing wild-type and *sigB* mutant. Fourty nine surface-associated proteins were modulated by $\sigma^B$, including 21 proteins already known to be SigB-dependent or SigB-influenced. More proteins were down- (31 proteins) than up-regulated (18 proteins) in the *sigB* mutant. Our approach revealed 28 surface-associated proteins not previously reported as SigB-dependent or -influenced, expanding the group of surface-associated proteins and virulence factors modulated by SigB.

**Keywords:** biotinylation • cell surface proteins • mass spectrometry • $^{14}N^{15}N$ metabolic labeling • SigmaB • *Staphylococcus aureus*

## Introduction

The interaction of a bacterium with its host is specifically based on the protein set presented on the surface of the bacterial cell. The human pathogenic bacterium *Staphylococcus aureus* produces an arsenal of virulence factors, particularly surface-exposed or secreted ones that cause a broad range of infections ranging from skin lesions to critical diseases such as endocarditis, osteomyelitis, toxic shock syndrome, and sepsis.[1-5] Antibiotic therapies have become more challenging due to increasing multidrug resistance, such as vancomycin/methicillin-resistant *S. aureus* strains,[6,7] and requiring the development of new antistaphylococcal therapeutic strategies. Analyzing the outer subset of proteins will lead to a deeper understanding of infection-related cell physiology and particularly of host-pathogen interaction.

The surface-associated subproteomic fraction comprises membrane proteins with exterior loops, lipoproteins, proteins covalently and noncovalently attached to the cell wall, as well as proteins containing signal peptides that are usually secreted but possibly also adhere to the cell surface.[8] Lipoproteins are characterized by the consensus sequence of lipobox in a signal sequence, a signature that has been used effectively to predict 66 lipoproteins in *S. aureus* COL.[9,10] Cell wall proteins are covalently attached to peptidoglycan by sortase (SrtA), a membrane-bound transpeptidase, via a special mechanism requiring a C-terminal sorting signal consisting of a conserved LPXTG-motif.[11,12] The alternative NPQTN motif is recognized by SrtB.[13] Genome analyses revealed 22 sortase substrates in several *S. aureus* strains, with 20 present in *S. aureus* strain COL,[14,15] the model organism in this study. Seven additional surface proteins that are not covalently attached to the cell wall were predicted from genome-wide analyses.[14] Those cell wall proteins and surface components that adhere to extracellular host tissue and plasma proteins for initializing infection process are known as MSCRAMMs (microbial surface components recognizing adhesive matrix molecules).[16]

Rendering this essential subproteomic fraction accessible to qualitative and quantitative proteome analyses is one of the challenging objectives in understanding pathophysiology and bacterial-host interactions. Different methods for isolating surface proteins of gram-positive bacteria have been established, including (i) serial extraction of proteins by different, highly concentrated salts,[17] (ii) enzymatic extraction, and (iii) biotinylation with subsequent enrichment of surface-associated proteins.

Enzymatic extraction of surface proteins was performed with lysostaphin (which lyses staphylococcal peptidoglycan) and

* To whom correspondence should be addressed. Institut für Mikrobiologie, Ernst-Moritz-Arndt-Universität Greifswald, Friedrich-Ludwig-Jahn-Str. 15, D-17489 Greifswald, Germany. E-mail: dbecher@uni-greifswald.de. Phone: +49-3834-864230. Fax: +49-3834-864202.

subsequent 2D-PAGE of the cell wall digest supernatant.[18–20] In another approach of enzymatic extraction intact bacteria were incubated with trypsin and cleaved surface proteins were identified via LC–MS/MS.[21,22]

Biotinylation can be used as a selective labeling method for surface exposed proteins. Numerous biotinylation reagents for the study of surface proteins are available (reviewed in ref 23). The biotinylation approach was initially established for eukaryotic cells[24] and adapted and advanced for gram-negative bacteria.[25–28] It was applied to gram-positive bacteria as *S. aureus* by Gatlin et al.[19] In this first comprehensive cell envelope proteome analysis of *S. aureus*, biotinylation was used in addition to cell wall digestion by lysostaphin. Gatlin and co-workers identified 48 surface proteins, including membrane-, lipo-, cell wall- and export signal-containing proteins as well as cytosolic contaminants, with the majority of surface proteins identified by the combination of lysostaphin digestion and 2D-PAGE. However, a quantitative analysis of the staphylococcal surface proteome is still lacking.

Protein quantification is crucial for physiological studies to monitor changes of protein abundance in response to different physiological stimuli or conditions. Different methods for quantitative mass spectrometry are reviewed by Bantscheff et al.[29] Label-free quantification has already been performed to monitor changes in the cell surface glyco-proteome of *D. melanogaster*.[30] Here we applied a metabolic labeling approach using $^{15}N$. Bacteria were cultured in $^{15}N$-enriched cell culture medium to introduce a stable isotope signature into proteins during growth.[31] Relative quantification was performed by comparing the intensities of isotope clusters of the unlabeled peptide (comprised of naturally abundant stable isotopes) and the $^{15}N$-labeled peptide.

To combine this quantitative study with a crucial physiological question, we analyzed the impact of the alternative sigma factor $\sigma^B$ on cell surface-associated protein expression in *S. aureus*.[32] SigmaB controls a large regulon in *S. aureus*, comprising about 300 genes, among them many with a function in virulence.[33–36] The impact of SigB on *S. aureus* pathogenicity was demonstrated in a number of different laboratories. For instance, active SigB was shown to be required for full virulence in a central venous catheter-related infection as well as a septic arthritis mice model and during the early stage of infection in an endocarditis rat model.[37–39] The exact molecular mechanism for the contribution of SigB to *S. aureus* related infections is not clear at present, but it has been suggested that the positive effect of SigB on the expression of MSCRAMMs (e.g., ClfA, FnbA) may contribute to the adhesion to host tissue, thereby promoting colonization and persistence.[38] In line with this notion is the observation that strains expressing SigB have an advantage over their *sigB* mutants during internalization into human osteoblasts, a process facilitated by surface-associated factors such as FnbA that is under partial control by SigB.[40] In addition, a large number of secreted virulence factors are negatively regulated by SigB. Many of those are expressed at high cell densities upon entry into stationary phase.[35] Since this group of virulence factors includes proteases, hemolysins, lipases, and different toxins that harm the host, it has been speculated that suppression of these factors may further contribute to intracellular persistence of *S. aureus*. To date about 60 potential surface-associated proteins have been described as SigB-dependent, although mostly in microarray-based studies.[36,41] More specifically, 27 genes (encoding 4 lipoproteins, 19 signal peptide-containing proteins and 4

sortase substrates) were shown to be under positive control and 29 genes (encoding 4 lipoproteins, 22 signal peptide-containing proteins and 3 sortase substrates) under negative control by SigB. In addition, genes encoding integral membrane proteins were down- (56 proteins) and up- (18 proteins) regulated in a *sigB* mutant.[36,41] Thus, the investigation of the impact of SigB on this subproteomic fraction is an ideal model system for our quantitative proteomic approach.

In this study, we present a GeLC–MS based proteomic analysis combining a biotinylation approach with $^{14}N/^{15}N$ metabolic labeling, which enables comprehensive quantitative analysis of enriched surface-exposed proteins. $^{14}N/^{15}N$ metabolic labeling was performed during cultivation prior to biotinylation of staphylococcal surface proteins. For biotinylation we used Sulfo-*N*-hydroxysulfosuccinimide-disulfide-Biotin (Sulfo-NHS-SS-Biotin), a membrane-impermeable reagent that reacts specifically with the $\varepsilon$-amino-group of lysine residues of surface-associated proteins. Isolation and purification of labeled proteins was carried out by affinity-chromatography. Proteins were eluted by cleaving the disulfide bond through reduction. We focused on alterations of surface protein composition of *S. aureus* COL wild type compared to a *sigB* mutant because the expression of a number of important virulence factors is modulated by SigB. Our results provide new insights in the regulation of surface-associated virulence factors of *S. aureus*.

## Experimental Procedures

*S. aureus* **Strains and Growth Conditions for** $^{14}N/^{15}N$ **Metabolic Labeling.** *Staphylococcus aureus* COL (Mec, high-Mc' clinical isolate, Tc'),[42] further indicated as wild type (WT) and *S. aureus* COL Δ*sigB* (deletion of *sigB* operon, Em'),[33] were used in this study. *S. aureus* COL WT and Δ*sigB* were grown under agitation at 37 °C in Bioexpress 1000 medium (Cambridge Isotope Laboratories, Andover, MA). For comparison of the protein expression of *S. aureus* COL WT and Δ*sigB* both strains were grown either in unlabeled or $^{15}N$-enriched medium. The experiment was repeated four times: twice by combining $^{15}N$-labeled WT cells with $^{14}N$-grown Δ*sigB* cells and a further two times by switching to $^{15}N$-labeled Δ*sigB* cells combined with $^{14}N$-grown WT cells. Label switch was performed to compensate for contingent differences between labeled and unlabeled BioExpress 1000 cell growth media. Exponential phase cells were harvested at $A_{600}$ of 0.5 and populations entering stationary phase 2 h later.

**Preparation of the Cell Surface Proteome by Biotinylation.** For preparation of the cell surface protein fraction, cells were harvested at 4000× *g* for 5 min at 4 °C, and 1 g of cells (wet cell weight) were resuspended in 5 mL ice-cold PBS (pH 8.0) with 1 mM PMSF on ice. Equivalent volumes of WT grown in unlabeled medium and Δ*sigB* grown in $^{15}N$-enriched medium were combined. In total 4 biological replicates were performed, thereby allowing for label switching when samples were combined vice versa. The biotinylation reaction was performed by adding 100 μL fresh Sulfo-NHS-SS-Biotin (Pierce, Rockford, IL) solution to 1 mL of intact cells, a final concentration of 1.5 mM Sulfo-NHS-SS-Biotin. A 1% (w/v) solution of Sulfo-NHS-SS-Biotin was prepared by adding 5 mg to 500 μL PBS (pH 8.0) immediately before use. Cells were incubated by gentle shaking for 2 h on ice. To stop the reaction and to remove nonreacted biotinylation reagent cells were centrifuged at 4000× *g* for 5 min at 4 °C and washed three times with ice-cold PBS (pH 8.0)/ 500 mM glycine. A pellet of 1 mL reaction

*Surface Proteome Profiling in S. aureus*

volume was resuspended in 500 μL PBS (pH 8.0) with 1 mM PMSF on ice and transferred to a 1.5 mL tube containing glass beads (Sartorius). No lysostaphin was added. Disruption of cells was performed mechanically in a Ribolyzer (Hybaid, Ashford, UK) at 6 m/s² twice for 20 s. The lysate was centrifuged (20 000× g for 30 min at 4 °C) and the biotinylated proteins were isolated and purified by NeutrAvidin agarose (Pierce, Rockford, IL) affinity-purification. For a reaction volume of 500 μL. protein mixture 50 μL of NeutrAvidin agarose resin were washed twice with PBS (pH 8.0)/ 1% NP-40 and centrifuged at 1000× g for 1 min at 4 °C. The resin was mixed with the cell lysate for 90 min by gently shaking on ice. The supernatant was removed and the resin-bound complex washed 6 times with PBS (pH 8.0)/ 1% NP-40. Biotinylated proteins were eluted by adding 20 μL of reductive SDS sample buffer containing 62.5 mM Tris/ HCl, pH 6.8, 2% SDS, 20% glycerol, 50 mM DTT and 5% β-mercaptoethanol. Additionally, we carried in each case a parallel control, which was treated equally except for the biotinylation reagent Sulfo-NHS-SS-Biotin, which was not added. This control verified successful Sulfo-NHS-SS-Biotin-tagging of proteins in the sample and showed that only a small number of proteins bind nonspecifically to avidin resin in the control, even under harsh washing conditions.

**Analysis of Protein Extracts by GeLC−MS.** Eluted protein samples were separated by 1D-SDS-PAGE.[43] After staining with Coomassie Brilliant Blue R-250, the complete protein separation lane was cut into 10 equal gel slices. In-gel digestion was performed with trypsin (Promega, Madison, WI) as described by Eymann et al.[44] Peptides obtained from in-gel digestion were separated by liquid chromatography and measured online by ESI-mass spectrometry. LC−MS/MS analyses were performed using a nanoACQUITY UPLC system (Waters, Milford, MA) coupled to an LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific,Waltham, MA) creating an electro spray by the application of 1.5 kV between Picotip Emitter (SilicaTip, FS360−20−10 Coating P200P, New Objective) and transfer capillary. Peptides were loaded onto a trap column (Symmetry C18, 5 μm, 180 μm inner diameter × 20 mm, Waters) and washed for 3 min with 99% buffer A (0.1% (v/v) acetic acid) at a flow rate of 10 μL/ min. Elution was performed onto an analytical column (BEH130 C18, 1.7 μm, 100 μm inner diameter × 100 mm, Waters) by a binary gradient of buffer A and B (100% (v/v) acetonitrile, 0,1% (v/v) acetic acid) over a period of 80 min with a flow rate of 1000 nl/ min. For MS/MS analysis a full survey scan in the Orbitrap (m/z 300−2000) with a resolution of 30 000 was followed by MS/MS experiments of the five most abundant precursor ions acquired in the LTQ via CID. Precursors were dynamically excluded for 30 s, and unassigned charge states as well as singly charged ions were rejected. Because of the high technical reproducibility of the LC−MS/MS analyses, the experiment was focused on 4 biological replicates as the biological variance is higher than the analytical one.

**Data Analysis.** For protein identification *.dta files were generated from *.raw files using Bioworks Browser 3.3.1 SP1 (Thermo Fisher Scientific). Charge state deconvolution and deisotoping were not performed. Identification of peptides was carried out by database search using SEQUEST version 28 (rev.12) (Thermo Fisher Scientific) against a target-decoy database (5318 entries). This database was composed of all protein sequences of *S. aureus* COL extracted from the National Center for Biotechnology Information (NCBI) bacteria genomes (http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genome&Cmd=

Retrieve&dopt=Protein+Table&list_uids=610). A set of the reversed sequences created by BioworksBrowser 3.2 EF2 as well as common contaminants such as keratin was appended. The search parameters specified trypsin (KR) cutting fully enzymatically, allowing two missed cleavage sites. Peptide tolerance for precursor ions was set to 10 ppm and fragment ion tolerance to 1 amu. The highest mass deviation that was observed was about 2 ppm. Only b- and y-ion series were included, considering methionine oxidation as variable modification with a maximum of three modifications per peptide. Each *.raw file was searched twice, once without any static modification and a second time by allowing for the substitution of all $^{14}$N with $^{15}$N atoms. Resulting *.out files of each sample were combined using DTASelect 2.0.[45] Overall $^{14}$N and $^{15}$N DTASelect files were merged with Contrast and filtered with DTASelect to obtain protein hits based on at least two different peptides and a false discovery rate on peptide level lower than 1% (filter used: DTASelect -y 2 -c 2 -C 4 --here --decoy Reverse_ -p 2 -t 2 -u --MC 2 -i 0.3 --fp 0.005). The final DTASelect-filter file was adjusted to remove peptide hits of the same sequence (differently charged and post-translationally modified peptides count to be the same) employing a user-written script. For the generation of spectral count values to illustrate the abundance of identified peptides, all spectra detected for each sequence were retained by DTASelect filtering with $t = 0$.[46]

For the prediction of transmembrane domains, the TMHMM 2.0 algorithm was used.[47] Signal sequences were estimated by the software tool SignalP 3.0.[48,49] When both the presence of a signal sequence and a cleavage site were presumed, and the signal peptide overlapped with a TMD to more than half of the latter's length, the TMD was considered to be a false-prediction, and subtracted. Proteins were considered to be lipid-anchored if projected by DOLOP[9] or Augur.[10] If proteins exhibited an LPXTG- or NPQTN motif, they were regarded as sortase substrates covalently bound to the cell wall.[14] Further cell wall-associated proteins were defined according to previous findings[14,50] and PSORT.[51] For protein quantification *.raw files had to be transferred to MS1 files (RawExtractor 1.9.3 by Tao Xu, Yates Laboratory at The Scripps Research Institute). The software tool Census version 1.31 uses MS1 files and the adjusted final DTASelect-filter file for protein quantification.[52] Quantification results were exported to Excel by implementing a determination factor of 0.7. A protein was considered reliably quantified when its ratio was determined in 2 biological replicates with not less than two peptides in at least one of the replicates. Normalization of protein ratios was carried out over the median of all log2 ratios. Significant changes in protein abundance were approved by a normalized log2 ratio above or lower than +0.8 and −0.8, respectively. Proteins that have been identified, but failed to give reliable quantitative data, were manually inspected for poor quality or presence in only one of the samples, $^{14}$N or $^{15}$N. If their presence in only one of the samples was based on more than one peptide in both biological replicates, they were added to the list of quantified proteins as "on"/"off" proteins. In this case, the correspondent $^{14}$N- or $^{15}$N-peptide was missing for quantification.

## Results

**Protein Identification by Biotinylation.** Biotinylation was used to capture the surface proteins of *S. aureus* COL. GeLC−MS/MS of the purified biotinylated proteins (Figure 1) led to the identification of 143, 124, 259, and 228 proteins from the four exponential phase culture samples (see Supplemental
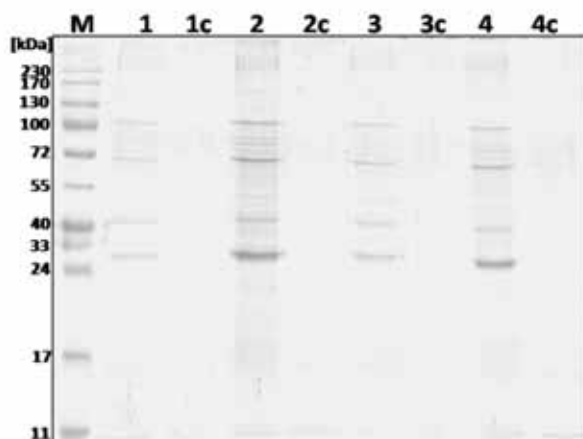
research articles

**Figure 1.** SDS-PAGE of the biotinylated, enriched cell surface protein fraction. Surface proteins from a combined preparation of ¹⁵N-labeled *S. aureus* COL WT and ¹⁴N-Δ*sigB* cells were separated. Shown are two biological replicates harvested each in exponential and stationary phase of growth. Each surface protein preparation was accompanied by a control reaction (c) performed without biotinylation reagent. Gel was stained with Coomassie Brilliant Blue R-250. **M** Protein Marker. **1** Exponential growth. **1c** Corresponding control. **2** Stationary phase. **2c** Control. **3** Second biological replicate exponential growth. **3c** Control. **4** Second biological replicate stationary phase. **4c** Control. Two further biological replicates were analyzed with label switch combining ¹⁴N-*S. aureus* COL WT and ¹⁵N-labeled Δ*sigB* cells resulting in a comparable picture (data not shown).

Table 1, Supporting Information) and 353, 299, 289, and 264 proteins from stationary phase cells (see Supplemental Table 2, Supporting Information). In total, 309 proteins were identified from exponentially growing cells among the four biological replicates. These proteins include 4 sortase substrates, 3 cell wall-associated proteins, 32 lipoproteins, 30 membrane proteins, 31 proteins containing a signal peptide, and 209 proteins predicted as cytoplasmic subcellular fraction (Figure 2A). For stationary phase of growth 441 proteins could be identified overall, including 4 sortase substrates, 4 cell wall-associated proteins, 40 lipoproteins, 39 membrane proteins, 40 proteins containing a signal peptide and 314 cytoplasmic proteins (Figure 2A). Addition of lysostaphin to our biotinylation approach could not improve the identification results in terms of getting a higher number of sortase substrates.

A control reaction accomplished simultaneously without biotinylation did not yield clear protein bands, although few proteins including pyruvate carboxylase (Pyc), acetyl-CoA carboxylase (AccB) and translation elongation factor G (FusA) were identified (Figure 1). However, Pyc and AccB are biotin carboxyl carrier proteins that bind naturally to avidin columns.[53] FusA is a high-abundance protein and thus difficult to remove completely from the NeutrAvidin agarose. The group of identified cytoplasmic proteins in the biotinylation sample is therefore considered to be biotinylated. In fact, a viable count determination before and after the biotinylation process showed a reduced number of CFUs, indicating cell death through the biotinylation of the cell surface (see Supplementary Figure 1, Supporting Information). Therefore, cell lysis is expected for part of these cells during the biotinylation experiment.

The abundance of the identified peptides of the subproteomic fractions analyzed by biotinylation differed (Figure 2B).

The number of spectra per peptide (spectral counts) gives an estimate of peptide abundance. The proportion of identified cell surface-associated proteins was about 30%. Considering the spectral counts as a method for evaluating the protein abundance of the identified proteins, the proportion of peptides belonging to cell surface-associated proteins was about 54% for growing and 56% for nongrowing cells (Figure 2B), indicating an effective enrichment. Additionally most of the cytosolic proteins were found in only one or two biological replicates (Figure 3), whereas most of the cell surface-associated proteins were identified in 3 or even in all 4 experiments, serving as evidence for the reproducibility of the enrichment procedure.

**Quantification of *S. aureus* Δ*sigB* Cell Surface Proteome.** In growing cells 150 proteins, which fulfilled the required criteria for quantification, were quantified comparing *sigB* mutant to its respective wild type (WT). Sixteen additional "on/off" proteins were exclusively present in either WT or Δ*sigB* (Supplemental Table 3, Supporting Information). Together, these 166 proteins comprise 3 sortase substrates, 1 cell wall-associated protein, 23 lipoproteins, 14 membrane proteins, 21 proteins containing a signal peptide, and 104 cytoplasmic proteins (Figure 4A). In stationary phase, 271 proteins including 27 on/off proteins were quantified, including 2 sortase substrates, 3 cell wall-associated proteins, 34 lipoproteins, 20 membrane proteins, 30 proteins containing a signal peptide, and 182 cytoplasmic proteins (Figure 4A, Supplemental Table 4, Supporting Information). In summary, 98 surface-associated proteins could be quantified for growing and nongrowing cells (Table 1). A comparison of the ratios for two biological replicates in a scatter plot demonstrated low variation and therefore high reproducibility of the presented method (Figure 5).

Alterations in protein amount were shown for a total of 53 proteins in growing *S. aureus* Δ*sigB* cells compared to the wild type. One protein was only present in the *sigB* mutant whereas 15 proteins were only identified in the wild type. Twelve proteins were increasing and 25 proteins decreasing in their amount compared to the WT. In nongrowing cells 92 proteins changed their amount when $\sigma^B$ was deleted including 8 "on"-proteins, 31 increasing, 34 decreasing, and 19 "off"-proteins.

In summary we found 117 proteins in altered amounts for both growth stages comparing wild type and *sigB* mutant. Of these, 49 proteins were assigned to the cell surface and extracellular compartments (Figure 4B), the focus of the present work, including 3 sortase substrates, 2 cell wall-associated proteins, 18 lipoproteins, 8 membrane proteins, and 18 proteins with a secretion signal (Table 1).

For SigB-regulation, we expect a down-regulation or even an off-protein in the *sigB* mutant in case the gene is solely (i, positive direct) or partially (ii, positive indirect) under SigB-control and therefore SigB-dependent. Up-regulation of proteins in the *sigB* mutant would indicate a (iii) negative indirect influence of SigB.

During at least one of the two time points analyzed, 31 surface-associated proteins showed a decreased amount in the *sigB* mutant, including off-proteins (Table 1). Fourteen of these had already been described as being SigB-dependent.[35,36,41,54] Among them were the fibrinogen-binding protein ClfA, the two putative lytic transglycosylases SceD and IsaA, the signal peptide-containing and putative surface protein SACOL2197, and the uncharacterized lipoprotein SACOL0444. A SigB-dependent promoter was experimentally confirmed for the *clfA* and *SACOL0444* genes and predicted for *sceD*, providing an
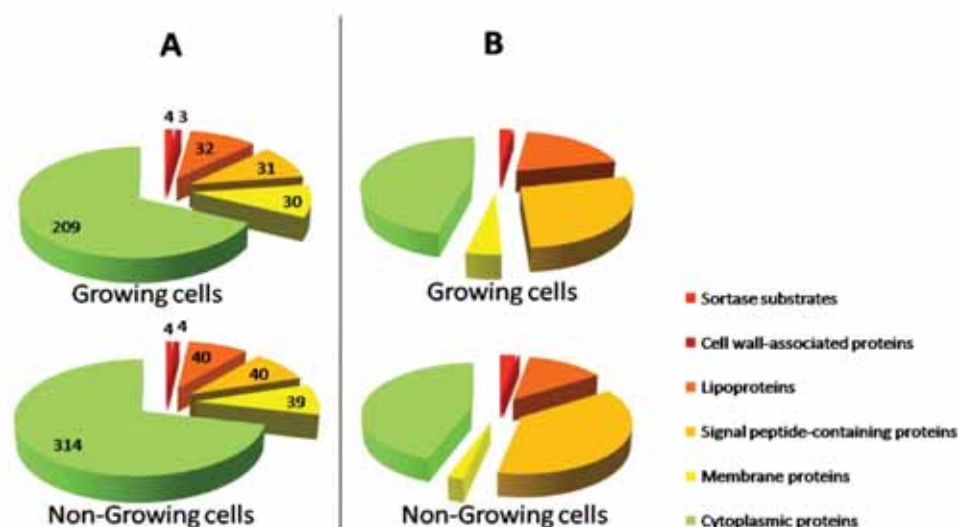
_____

**Figure 2.** Identification of proteins. (A) Number of proteins. In summary the biotinylation of *S. aureus* COL WT and ΔsigB cells for surface protein enrichment led to the identification of 309 proteins in growing and 441 proteins in nongrowing cells in four biological replicates. Proteins were identified with at least 2 significant peptides via GeLC−MS/MS. Protein classification according to subproteomes reveals that about one-third of the identified proteins are part of cell envelope-associated proteins. (B) Number of spectral counts normalized by Ø molecular mass. To estimate the abundance of the identified proteins the number of spectra per identified peptide (spectral counts) were normalized by the average molecular mass of the corresponding subproteomic fractions to eliminate the bias caused by protein size. Considering the spectral counts the abundance of cell envelope-associated peptides is higher.
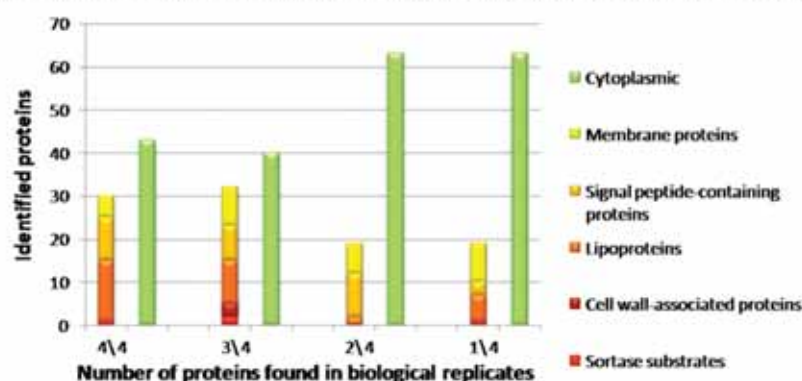


**Figure 3.** Number of proteins found in biological replicates. The number of proteins identified in all (4/4), three (3/4), two (2/4) or only one (1/4) biological replicate is shown for exponential growth phase experiment. Most of the cytoplasmic proteins (right column in green) were found in one or two of the four biological replicates. A higher proportion of cell surface-associated proteins (left column in yellowish colors) were identified in all four or at least in three of four experiments.

explanation for the strongly (i) positive, direct SigB-dependent accumulation of these proteins.[35,55] Interestingly, SACOL0444 and SACOL2197 were characterized as off-proteins at both time points indicating that these proteins are completely absent in the *sigB* mutant. This is in agreement with previous analyses that have demonstrated that transcription of the respective genes occurs only in strains with active SigB (refs 36 and 41 and our unpublished results). Surprisingly, in contrast to *SACOL0444*, no obvious SigB-binding site can be identified upstream of *SACOL2197*,[36,41] indicating a (ii) positive, indirect regulation pattern. Both proteins display a domain organization that suggests a function in virulence. SACOL2197 consists of a single MAP (MHC class II analogous protein) domain. The MAP domain is found in five copies in Eap (extracellular adherence protein) for which multiple functions in host-pathogen interaction, including adhesion to host tissue, immune modulation,

and anti-inflammatory functions have been demonstrated.[56] The lipoprotein SACOL0444 consists of two in tandem PepSY domains that are usually found at the N-terminus of proteases that are synthesized as chaperones where they display protease inhibitor activity.[57]

Additional proteins that were down-regulated in the mutant, and that were previously described as SigB-dependent include LytM, LysM, CapA, and the lipoprotein SACOL2365. Furthermore, we quantified 17 additional surface-associated proteins with decreasing amounts in the mutant that were not characterized as SigB-dependent so far. For example ClfB and 10 lipoproteins that include members of the periplasmatic binding protein family were also down-regulated in the *sigB* mutant. Interestingly, the signal peptide-containing protein YycH, whose gene is cotranscribed with the genes encoding the essential WalK/WalR (YycF/YycG) two-component system, was
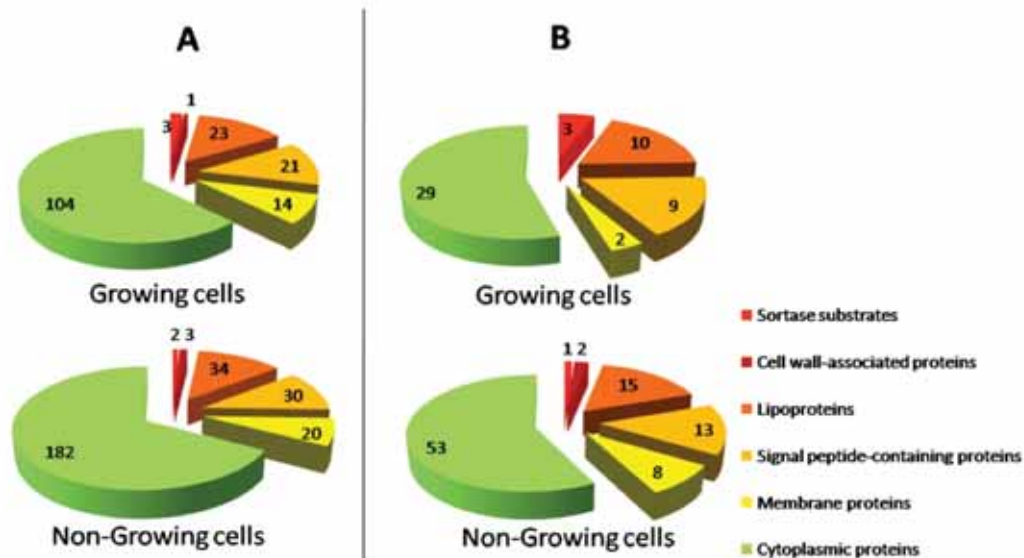
159

**Figure 4.** Proportion of quantified and regulated proteins. (A) Quantification. Quantitative data could be obtained for 166 and 271 proteins in exponential and stationary growth phase, respectively. Thus the amount of 62 and 89 cell envelope-associated proteins could be determined. (B) Regulation. Considering solely the proteins with altered amounts in ΔsigB 45% of all quantified proteins have cell envelope-associated functions. Fifty-three and 92 proteins, including 16 and 27 on/off proteins, in exponential and stationary phase of growth, respectively, changed their amount significantly.

down-regulated in the *sigB*-mutant. WalK/WalR is known to control the expression of cell wall related functions.[58]

Compared to the wild-type situation, we observed a different tendency for the amount of surface-associated proteins. Proteins down-regulated at both time points in the *sigB* mutant either remained unchanged or increased in stationary phase cells of the wild type. In contrast, proteins down-regulated only at the exponential phase of growth in the SigB-depleted strain did not usually accumulate during stationary phase.[59]

Together, 18 surface-associated proteins were found in increasing amounts in the *sigB* mutant strain. With the exception of SACOL2302, the (iii) negative, indirect effect of SigB on the regulated surface-associated protein was restricted to the stationary phase (Figure 6). Seven out of these 17 proteins quantified in our analysis were already characterized as suppressed by SigB in earlier studies.[35,36] Of these, Nuc, GlpQ, SspB1, SspB2, Aur, and Hlb belong to the RNAIII quorum sensing system regulated set of secreted virulence factors.[35] In summary, we found 11 additional proteins with increasing amounts that were not described to be under negative SigB-influence before. Interestingly, we found the SecDF component of the sec-dependent secretion machinery to be expressed at a higher level in the *sigB*-deleted strain. Five thus far uncharacterized lipoproteins accumulated at higher amounts in the *sigB* mutant, of which 3 were predicted to function as periplasmatic binding proteins.

Finally, SACOL1522 that was characterized as positively SigB-dependent and SACOL0303 and SACOL2476 that were characterized as negatively SigB-influenced in earlier DNA-microarray-based studies[36,41] were also quantified by our approach but not considered as regulated according to our cutoff criteria.

## Discussion

**Quantitative Biotinylation Approach for Capturing the Surface Proteome of *S. aureus*.** Exploring the surface proteome of pathogenic bacteria is one of the current goals in proteomics.

A gold-standard method for the selective and gelfree capturing of surface-exposed proteins of the bacterial cell has not been established to date. Tagging of cell surface-associated proteins with Sulfo-NHS-SS-Biotin followed by NeutrAvidin affinity-chromatography resulted in clear enrichment for a specific protein fraction (Figure 1). The biotinylation approach led to the identification of 100 proteins in growing and 127 proteins in nongrowing cells that are known as cell envelope-associated (Figure 2A).

Nine cell wall proteins besides an uncharacterized one, ClfA, ClfB, EbpS, Pls, SasH, Atl, IsaB, and SdrE including 5 proteins with LPXTG motive were found in growing and nongrowing cells (see Supplemental Table 1 and 2, Supporting Information). These proteins represent about one-third of the surface proteins of *S. aureus* COL predicted by Gill et al.[14] and Clarke and Foster.[50] Further studies on infection-associated conditions (e.g., iron starvation) may increase the number of identifiable proteins, as additive virulence gene expression should be achieved. For *S. aureus*, COL 66 surface-exposed lipoproteins are predicted.[9,10] In this study, a total of 40 lipoproteins could be found representing 61% of the theoretical predicted lipoproteome. Interestingly, 43 signal peptide-containing proteins representing 30% of all staphylococcal extracellular proteins were found in the biotinylated surface protein fraction, indicating that these proteins are loosely associated with the cell surface. Among these 43 proteins, 29 signal peptide-containing proteins were found in addition to a previous study of the extracellular proteome via 2D-PAGE.[60] Altogether 135 cell envelope-associated proteins including membrane-, lipo-, cell wall-, and export signal-containing proteins could be identified (see Supplemental Table 1 and 2, Supporting Information). This is the highest coverage of the *S. aureus* surface proteome to date, particularly because of the advantages of GeLC−MS/MS-based proteome analyses compared to 2D-PAGE. The applied approach adds new proteins on the way toward the identification of the entire proteome of *S. aureus*. It allows the detection

*Surface Proteome Profiling in S. aureus*

**Table 1.** Quantified Surface-Associated Proteins Comparing *S. aureus* COL Δ*sigB* and WT

| protein accession number | protein | φ log2 ratio in exponential growth[a] | | φ log2 ratio in stationary phase[a] | | localization[b] |
|---|---|---|---|---|---|---|
| SACOL0995 | – | off | ↓ | −2.81 | ↓ | Lipoprotein |
| **SACOL0444**[e] | – | off | ↓ | off | ↓ | Lipoprotein |
| **SACOL2197** | – | off | ↓ | off | ↓ | Signal peptide-containing |
| SACOL0506 | – | −5.77 | ↓ | −2.88 | ↓ | Lipoprotein |
| SACOL0884 | – | −3.26 | ↓ | −1.38 | ↓ | Lipoprotein |
| **SACOL0856** | ClfA | −2.80 | ↓ | off | ↓ | Sortase substrate |
| **SACOL2584** | IsaA | −1.45 | ↓ | −1.00 | ↓ | Signal peptide-containing |
| SACOL2252 | – | −0.96 | ↓ | −1.07 | ↓ | Membrane |
| SACOL0050 | Pls | −1.16 | ↓ | −0.70 | ↓[c] | Sortase substrate |
| **SACOL2365** | – | −0.74 | ↓[c] | −0.81 | ↓ | Lipoprotein |
| **SACOL2088** | SceD | −0.68 | ↓[c] | −1.26 | ↓ | Signal peptide-containing |
| **SACOL0723** | LysM | off | ↓ | | | Signal peptide-containing |
| **SACOL2295** | – | off | ↓ | | | Signal peptide-containing |
| **SACOL0263** | LytM | −2.75 | ↓ | | | Signal peptide-containing |
| SACOL2652 | ClfB | −2.66 | ↓ | | | Sortase substrate |
| SACOL0799 | – | −2.14 | ↓ | −0.32 | ↔ | Lipoprotein |
| SACOL2277 | – | −1.38 | ↓ | 0.30 | ↔ | Lipoprotein |
| **SACOL0099** | SirA | −1.16 | ↓ | 1.31 | ↔[d] | Lipoprotein |
| SACOL0021 | YycH | −1.13 | ↓ | | | Signal peptide-containing |
| SACOL0033 | MecA | −0.98 | ↓ | −0.21 | ↔ | Signal peptide-containing |
| **SACOL0272** | – | −1.05 | ↓ | 1.83 | ↑ | Membrane |
| SACOL2167 | – | −0.94 | ↓ | 1.80 | ↑ | Lipoprotein |
| **SACOL0136** | Cap5A | | | off | ↓ | Membrane |
| SACOL0888 | – | | | off | ↓ | Lipoprotein |
| SACOL0187 | – | | | off | ↓ | Lipoprotein |
| **SACOL0907** | Seb | | | off | ↓ | Signal peptide-containing |
| SACOL2660 | IsaB | | | −2.09 | ↓ | Cell wall-associated |
| SACOL1847 | – | | | −1.76 | ↓ | Cell wall-associated |
| SACOL2451 | – | | | −1.23 | ↓ | Lipoprotein |
| SACOL1225 | – | | | −0.84 | ↓ | Lipoprotein |
| **SACOL0270** | – | −0.55 | ↔ | −1.02 | ↓ | Signal peptide-containing |
| **SACOL2476** | – | −0.16 | ↔ | 1.29 | ↔[d] | Lipoprotein |
| SACOL0938 | DltD | −0.76 | ↔ | 0.40 | ↔ | Signal peptide-containing |
| SACOL0486 | – | −0.72 | ↔ | −0.02 | ↔ | Lipoprotein |
| SACOL0778 | – | −0.68 | ↔ | −0.46 | ↔ | Membrane |
| SACOL0217 | – | −0.57 | ↔ | 0.31 | ↔ | Lipoprotein |
| SACOL2403 | – | −0.50 | ↔ | −0.14 | ↔ | Lipoprotein |
| SACOL1609 | Pbp3 | −0.48 | ↔ | −0.18 | ↔ | Membrane |
| SACOL1964 | CamS | −0.34 | ↔ | −0.35 | ↔ | Lipoprotein |
| SACOL1630 | – | −0.26 | ↔ | 0.14 | ↔ | Membrane |
| SACOL2272 | ModA | −0.23 | ↔ | 0.42 | ↔ | Lipoprotein |
| SACOL1490 | Pbp2 | −0.22 | ↔ | 0.11 | ↔ | Signal peptide-containing |
| SACOL2467 | – | −0.08 | ↔ | 0.16 | ↔ | Lipoprotein |
| SACOL1398 | – | −0.03 | ↔ | 0.53 | ↔ | Signal peptide-containing |
| SACOL0969 | SpsB | 0.02 | ↔ | 0.24 | ↔ | Membrane |
| **SACOL1522** | – | 0.04 | ↔ | −0.26 | ↔ | Cell wall-associated |
| SACOL1777 | – | 0.08 | ↔ | 0.11 | ↔ | Membrane |
| SACOL1455 | – | 0.12 | ↔ | 0.15 | ↔ | Membrane |
| SACOL2348 | – | 0.26 | ↔ | 0.20 | ↔ | Signal peptide-containing |
| SACOL2383 | – | 0.27 | ↔ | 0.54 | ↔ | Signal peptide-containing |
| SACOL1066 | – | 0.28 | ↔ | 0.31 | ↔ | Signal peptide-containing |
| SACOL0449 | – | 0.36 | ↔ | 0.21 | ↔ | Lipoprotein |
| SACOL1194 | Pbp1 | 0.36 | ↔ | 0.55 | ↔ | Signal peptide-containing |
| SACOL1704 | MreC | 0.39 | ↔ | 0.04 | ↔ | Signal peptide-containing |
| SACOL0555 | – | 0.42 | ↔ | 0.30 | ↔ | Membrane |
| SACOL1065 | – | 0.54 | ↔ | 0.34 | ↔ | Signal peptide-containing |
| SACOL1101 | – | 0.63 | ↔ | −0.01 | ↔ | Lipoprotein |
| SACOL1897 | – | 0.75 | ↔ | 0.38 | ↔ | Lipoprotein |
| SACOL2152 | – | −0.37 | ↔ | | | Signal peptide-containing |
| SACOL2582 | – | −0.20 | ↔ | | | Membrane |
| SACOL0699 | Pbp4 | −0.17 | ↔ | | | Membrane |
| SACOL1589 | – | −0.14 | ↔ | | | Lipoprotein |
| SACOL1281 | – | 0.51 | ↔ | | | Membrane |
| SACOL0851 | – | | | −0.56 | ↔ | Lipoprotein |
| SACOL0669 | – | | | −0.39 | ↔ | Signal peptide-containing |
| SACOL1514 | GpsA | | | −0.26 | ↔ | Signal peptide-containing |

**Table 1** Continued

| protein accession number | protein | φ log2 ratio in exponential growth[a] | | φ log2 ratio in stationary phase[a] | | localization[b] |
|---|---|---|---|---|---|---|
| SACOL2407 | – | | | −0.16 | ** | Lipoprotein |
| SACOL1574 | – | | | 0.08 | ** | Lipoprotein |
| SACOL1159 | SdhA | | | 0.18 | ** | Signal peptide-containing |
| SACOL2552 | – | | | 0.22 | ** | Membrane |
| SACOL1991 | – | | | 0.28 | ** | Membrane |
| **SACOL0303** | – | | | 0.30 | ** | Signal peptide-containing |
| SACOL0665 | – | | | 0.33 | ** | Lipoprotein |
| SACOL0803 | – | | | 0.40 | ** | Lipoprotein |
| SACOL2539 | SrtA | | | 0.41 | ** | Membrane |
| SACOL2179 | – | | | 0.42 | ** | Signal peptide-containing |
| SACOL2436 | – | | | 0.52 | ** | Membrane |
| SACOL1989 | – | | | 0.53 | ** | Signal peptide-containing |
| SACOL0128 | – | | | 0.56 | ** | Lipoprotein |
| SACOL2381 | – | | | 0.72 | ** | Membrane |
| SACOL2302 | – | 0.83 | ↑ | 0.51 | ** | Signal peptide-containing |
| SACOL1045 | – | | | 1.01 | ↑ | Lipoprotein |
| **SACOL0985** | – | | | 1.02 | ↑ | Signal peptide-containing |
| SACOL1932 | – | | | 1.09 | ↑ | Membrane |
| **SACOL0962** | GlpQ | | | 1.31 | ↑ | Signal peptide-containing |
| SACOL0193 | – | | | 1.35 | ↑ | Lipoprotein |
| SACOL0275 | – | | | 1.48 | ↑ | Membrane |
| **SACOL2003** | Hlb | | | 2.52 | ↑ | Signal peptide-containing |
| **SACOL0860** | Nuc | | | 2.80 | ↑ | Membrane |
| **SACOL1056** | SspB1 | | | on | ↑ | Signal peptide-containing |
| **SACOL1970** | SspB2 | | | on | ↑ | Signal peptide-containing |
| **SACOL2659** | Aur | | | on | ↑ | Signal peptide-containing |
| SACOL1069 | QoxA | −0.31 | ** | 0.91 | ↑ | Membrane |
| SACOL2412 | – | −0.31 | ** | 0.90 | ↑ | Lipoprotein |
| SACOL1692 | SecDF | 0.22 | ** | 0.93 | ↑ | Membrane |
| SACOL1825 | – | 0.45 | ** | 0.82 | ↑ | Signal peptide-containing |
| SACOL0768 | – | 0.04 | ** | 0.79 | ↑[c] | Lipoprotein |
| SACOL1070 | QoxB | 0.38 | ** | 0.78 | ↑[c] | Lipoprotein |

[a] Average log2 fold changes from 4 biological replicates of *S. aureus* COL Δ*sigB* compared to WT. ↑, Higher amount (log 2 ratio >0.8 in at least 2 biological replicates); **, no change in amount; ↓, lower amount (log 2 ratio < −0.8 in at least 2 biological replicates); on, exclusivly present in *sigB* mutant; off, exclusivly present in WT. [b] For the prediction of transmembrane domains the TMHMM 2.0 algorithm was used.[47] Proteins were considered to be lipo-anchored if projected by DOLOP[8] or Augur.[10] If proteins exhibited an LPXTG- or NPQTN motif they were regarded as sortase substrates.[14,15] Cell wall-associated proteins were defined according to previous findings.[14,50] Signal sequences were estimated by the software tool SignalP 3.0.[48,49] [c] Reliably quantified in more than 2 biological replicates, with average fold change in 2 biological replicates above and lower than 0.8 and −0.8, respectively. [d] Reliably quantified in more than 2 biological replicates, with fold changes above or lower threshold in only 1 biological replicate. [e] Proteins in bold have been described already.[35,36,41,54] Detailed quantitative data see in Supplemental Tables 3 and 4, Supporting Information.

of surface-exposed proteins and of proteins whose full function and localization is unknown so far. Identified extracellular proteins such as the extracellular autolysin SACOL1825 can have both enzymatic and adhesive functions[61] and are therefore found on the surface. The fraction of cytoplasmic proteins is assumed to contaminate the surface-associated fraction for two reasons. First, cell lysis during cultivation and experimental processes (see Supplementary Figure 1, Supporting Information) frees intracellular proteins which are then also tagged with Sulfo-NHS-SS-Biotin and therefore enriched and identified. Enrichment of cytoplasmic proteins due to unspecific binding on NeutrAvidin agarose can be excluded as only a few biotin carboxyl carrier proteins were found in the control reaction. Second, the identification of intracellular proteins could indicate an additional unknown function of some proteins at the cell surface. External presence of enolase or GapA2, also found in this study, was already shown previously.[62-64] Any unexpected permeability of the biotinylation reagent would be nullified as the incorporated disulfide bridge would be cleaved in the reductive milieu of the cytosol. Cleavage of the biotinylation reagent would make protein labeling impossible.

$^{14}$N$^{15}$N metabolic labeling in combination with biotinylation of the labeled surface proteins enables the investigation of

differences in surface protein abundance. On the basis of the above-mentioned criteria 54 and 62% of the identified proteins could be quantified in growing and nongrowing cells, respectively. Exemplified by the analysis of *S. aureus* COL WT and Δ*sigB* in exponential and stationary phase, more than one-third of the quantified proteins are cell surface-associated (Figure 4A). The presented method is therefore suitable for quantitative proteomic studies of bacterial surface proteins.

**SigB-Dependent Changes in the Surface-Associated Proteome.** The expression of the cell surface-associated and secreted proteome is under complex control by overlapping regulatory circuits that involve a large number of different transcription factors.[65] In this study, we quantified a total of 49 surface-associated proteins with altered amounts in the *sigB*-mutant (Table 1). In support of the reliability of our approach, we found that for 14 of these proteins SigB-dependence and for 7 proteins negative SigB-influence have previously been reported, albeit mostly at the transcriptional level.[35,36,41,54]

The majority of the proteins for which we obtained quantitative data could still be identified in the *sigB*-deleted strain. These observations suggest that expression of these proteins is not solely SigB-dependent and that the temporal impact of SigB on expression is modulated by additional regulators that
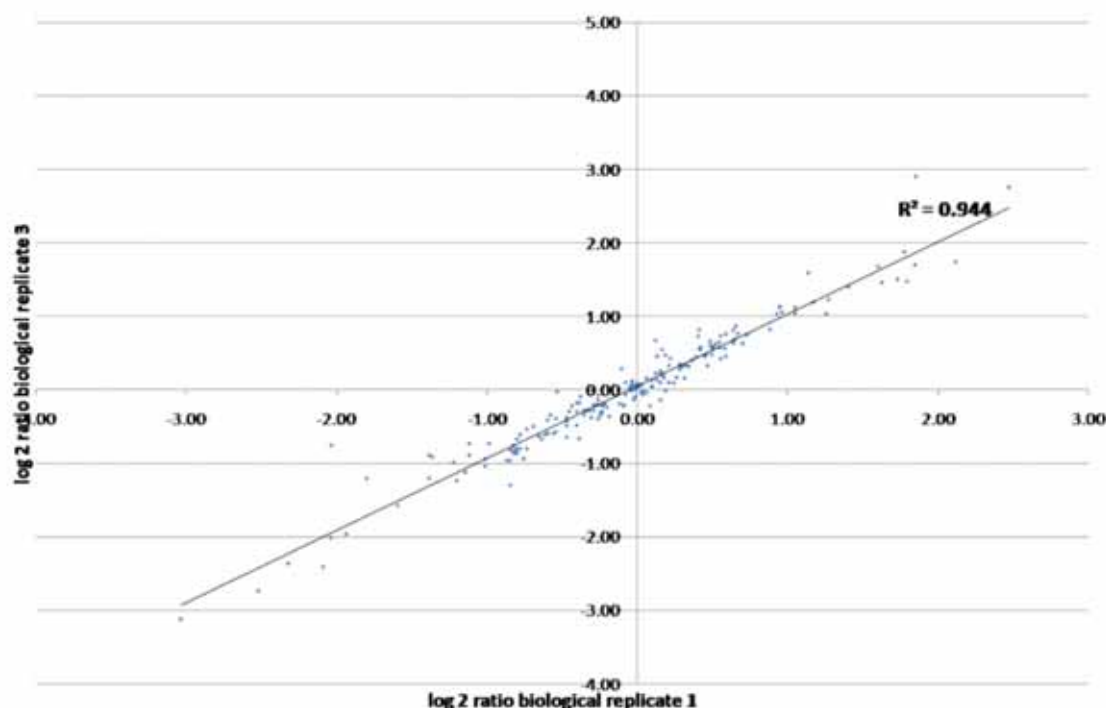
**Figure 5.** Exemplified comparison of two biological replicates. Normalized log2 ratios of the quantified proteins in two biological experiments demonstrates the high reproducibility of the quantification results.
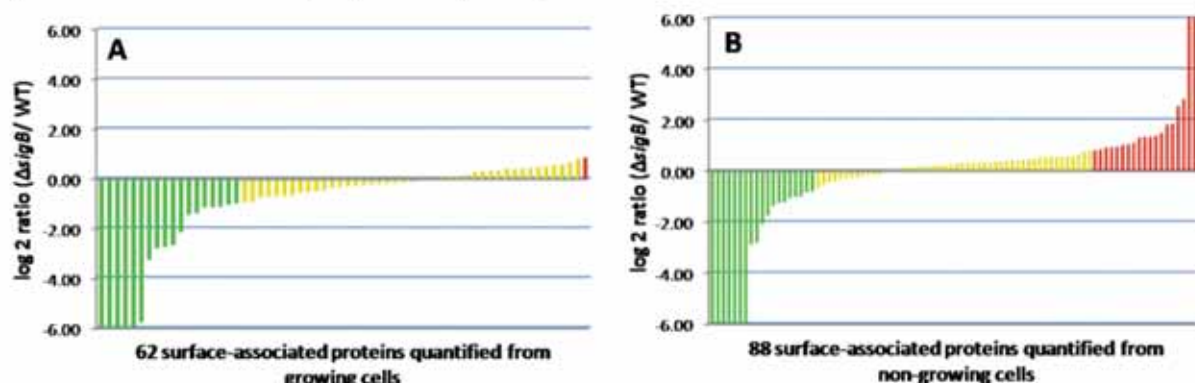


**Figure 6.** Up- and down-regulation of cell envelope-associated proteins. Differential protein expression profiling of surface proteins including membrane proteins, lipoproteins, sortase substrates, cell wall-associated proteins and signal peptide-containing proteins in (A) growing and (B) nongrowing cells comparing *S. aureus* sigB mutant and wild type. Log2 ratio above and lower than 0.8 and −0.8 signifies increasing or decreasing amounts, respectively. For illustrating "on/ off-proteins" a log2 ratio of −6 represents a protein "off" in the *sigB* mutant and a log2 ratio of 6 a protein "on" in the *sigB* mutant.

act up- or downstream of SigB. For most of the proteins found at lower levels in the *sigB*-mutant, a consensus sequence-directed search failed to identify potential SigB-binding sites upstream of the respective genes (data not shown), supporting the presence of additional regulators that likely act in concert with SigB. Further studies are clearly required to establish the mechanistic link between the expression of these proteins and SigB. Transcription of the two DNA binding proteins SarA and SarS is under partial control by SigB.[66,67] Both proteins act as global regulators and are important components of the complex network that controls virulence gene expression in *S. aureus*.[68] For many of the proteins quantified in our approach, neither SarA nor SarS dependence has been reported to date. However, we have found that the down-regulated proteins IsaA, SceD,

LytM, LysM, and SACOL2295, for which DNA-array data indicate a SigB-dependence, were recently demonstrated to be under positive transcriptional control by the essential WalK/ WalR (YycG/YycF) two-component system.[58] WalKR was shown to positively control autolytic activity in *Staphylococcus aureus*.[58]

SigB modulates the capacity of *S. aureus* to adhere and invade host tissue.[40] Consistent with this notion we observed a strong decrease in the amount of the two fibrinogen binding clumping factors ClfA and ClfB in the *sigB* mutant. This situation is somewhat reminiscent of what has been described for *Listeria monocytogenes* where a role for SigB via expression of the crucial adhesions and internalization molecules InlA and InlB during host cell invasion has been clearly demonstrated.[69]

## research articles

ClfB of *S. aureus* may be of crucial importance during persistent nasal-colonization in humans.[70] Furthermore, SceD, a putative lytic transglycosylase, was shown to be essential for nasal colonization in a cotton rat model.[71] Interestingly, expression of SceD was greatly upregulated in response to salt stress,[71] a stimulus that also activates SigB.[41,72] A SigB-dependent promoter was predicted in front of the *sceD* gene[54] and it is tempting to speculate that active SigB may be important during persistence in the nose where changes in osmolarity may be frequent. In addition to SceD, further proteins with a potential role in cell wall dynamics were identified in our study to be under positive (IsaA, LytM, SACOL0270, SACOL2295) as well as negative (SACOL1932, SACOL1825) regulatory influence by SigB. In this context it is also interesting to note that overexpression of SigB resulted in cell wall thickening and increased resistance to cell wall-acting antibiotics.[73] Indeed, several mutations that led to an increased resistance to cell wall-active antibiotics were demonstrated to increase SigB-activity.[74] However, the levels of penicillin-binding proteins (PBPs), important in the final steps of peptidoglycan biosynthesis and involved in the resistance of *S. aureus* to beta-lactams, were not altered in our study (Table 1). These results are in accordance with a study by Morikawa and colleagues[73] were no differences in the transcription of PBPs were detected between an *S. aureus* WT strain and its *sigB* mutant. Only overexpression of plasmid encoded SigB led to an increased transcription of PBP genes and it was suggested that this is likely to occur via an indirect pathway.[73]

An increased accumulation of surface-associated proteins in the *sigB* mutant was almost exclusively observed at stationary phase. It is conceivable that this negative influence of SigB is indirect under these conditions since SigB can only act as a positive transcription factor. Increased expression in a *sigB* mutant during stationary phase was already shown for many extracellular enzymes.[35,54] In this case, $\sigma^B$ seems to delay the synthesis of extracellular enzymes in the stationary phase. For Nuc, GlpQ, SspB1, SspB2, Aur, and Hlb, which were quantified by our approach, this observation could be explained by their dependence on the cell density activated Agr two-component system. It has been shown that the level of RNAIII, the regulatory molecule of the Agr two-component system, was higher in strains with a deletion in *sigB*.[75] However, in a different study, only a minor effect of SigB on the level of RNAIII was reported, leaving the link between SigB and the Agr controlled protein expression an open question.[35]

In summary, we provided quantitative data for 98 staphylococcal surface-associated proteins including 49 proteins with altered amounts in a *sigB* mutant. A total of 28 newly characterized proteins have not previously been described as SigB-dependent or SigB-influenced. The SigB-regulated surface proteome supports a role for SigB in the modulation of the cell wall metabolism and the expression of important adhesion molecules. Both aspects are likely to contribute to the fitness of *S. aureus* during persistence in its host niche. To our knowledge we present the first comprehensive quantitative analysis of surface proteins for *S. aureus* COL. The comparison of WT and *sigB* mutant showed the potential of the combination of $^{14}N^{15}N$ metabolic labeling, biotinylation and GeLC-MS/MS to analyze even proteins with unknown functions. Our quantitative biotinylation approach is now an established mass spectrometry-based method for exploring surface protein pattern, thereby adding a next layer to the full understanding of *S. aureus* pathophysiology.

## References

(1) Mertz, P. M.; Cardenas, T. C.; Snyder, R. V.; Kinney, M. A.; Davis, S. C.; Plano, L. R. Staphylococcus aureus virulence factors associated with infected skin lesions: influence on the local immune response. *Arch Dermatol.* **2007**, *143*, 1259–63.

(2) Millar, B. C.; Prendergast, B. D.; Moore, J. E. Community-associated MRSA (CA-MRSA): an emerging pathogen in infective endocarditis. *J. Antimicrob. Chemother.* **2008**, *61*, 1–7.

(3) Saavedra-Lozano, J.; Mejias, A.; Ahmad, N.; Peromingo, E.; Ardura, M. I.; Guillen, S.; Syed, A.; Cavuoti, D.; Ramilo, O. Changing trends in acute osteomyelitis in children: impact of methicillin-resistant Staphylococcus aureus infections. *J. Pediatr. Orthop.* **2008**, *28*, 569–75.

(4) Schlievert, P. M.; Shands, K. N.; Dan, B. B.; Schmid, G. P.; Nishimura, R. D. Identification and characterization of an exotoxin from Staphylococcus aureus associated with toxic-shock syndrome. *J. Infect. Dis.* **1981**, *143*, 509–16.

(5) Castaldo, E. T.; Yang, E. Y. Severe sepsis attributable to community-associated methicillin-resistant Staphylococcus aureus: an emerging fatal problem. *Am Surg.* **2007**, *73*, 684–7, and discussion 687–8.

(6) Weigel, L. M.; et al. Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus. *Science* **2003**, *302*, 1569–71.

(7) Moran, G. J.; Krishnadasan, A.; Gorwitz, R. J.; Fosheim, G. E.; McDougal, L. K.; Carey, R. B.; Talan, D. A. Methicillin-resistant S. aureus infections among patients in the emergency department. *N. Engl. J. Med.* **2006**, *355*, 666–74.

(8) Tjalsma, H.; van Dijl, J. M. Proteomics-based consensus prediction of protein retention in a bacterial membrane. *Proteomics* **2005**, *5*, 4472–82.

(9) Madan Babu, M.; Sankaran, K. DOLOP--database of bacterial lipoproteins. *Bioinformatics* **2002**, *18*, 641–3.

(10) Billion, A.; Ghai, R.; Chakraborty, T.; Hain, T. Augur--a computational pipeline for whole genome microbial surface protein prediction and classification. *Bioinformatics* **2006**, *22*, 2819–20.

(11) Navarre, W. W.; Schneewind, O. Proteolytic cleavage and cell wall anchoring at the LPXTG motif of surface proteins in gram-positive bacteria. *Mol. Microbiol.* **1994**, *14*, 115–21.

(12) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **1999**, *285*, 760–3.

(13) Mazmanian, S. K.; Ton-That, H.; Su, K.; Schneewind, O. An iron-regulated sortase anchors a class of surface protein during Staphylococcus aureus pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2293–8.

(14) Gill, S. R.; et al. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a biofilm-producing methicillin-resistant Staphylococcus epidermidis strain. *J. Bacteriol.* **2005**, *187*, 2426–38.

(15) Sibbald, M. J.; et al. Mapping the pathways to staphylococcal pathogenesis by comparative secretomics. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 755–88.

(16) Foster, T. J.; Hook, M. Surface protein adhesins of Staphylococcus aureus. *Trends Microbiol.* **1998**, *6*, 484–8.

research articles

(17) Schaumburg, J.; Diekmann, O.; Hagendorff, P.; Bergmann, S.; Rohde, M.; Hammerschmidt, S.; Jansch, L.; Wehland, J.; Karst, U. The cell wall subproteome of Listeria monocytogenes. *Proteomics* **2004**, *4*, 2991–3006.

(18) Nandakumar, R.; Nandakumar, M. P.; Marten, M. R.; Ross, J. M. Proteome analysis of membrane and cell wall associated proteins from Staphylococcus aureus. *J. Proteome Res.* **2005**, *4*, 250–7.

(19) Gatlin, C. L.; et al. Proteomic profiling of cell envelope-associated proteins from Staphylococcus aureus. *Proteomics* **2006**, *6*, 1530–49.

(20) Planchon, S.; Chambon, C.; Desvaux, M.; Chafsey, I.; Leroy, S.; Talon, R.; Hebraud, M. Proteomic analysis of cell envelope from Staphylococcus xylosus C2a, a coagulase-negative staphylococcus. *J. Proteome Res.* **2007**, *6*, 3566–80.

(21) Tjalsma, H.; Lambooy, L.; Hermans, P. W.; Swinkels, D. W. Shedding & shaving: disclosure of proteomic expressions on a bacterial face. *Proteomics* **2008**, *8*, 1415–28.

(22) Severin, A.; Nickbarg, E.; Wooters, J.; Quazi, S. A.; Matsuka, Y. V.; Murphy, E.; Moutsatsos, I. K.; Zagursky, R. J.; Olmsted, S. B. Proteomic analysis and identification of Streptococcus pyogenes surface-associated proteins. *J. Bacteriol.* **2007**, *189*, 1514–22.

(23) Elia, G. Biotinylation reagents for the study of cell surface proteins. *Proteomics* **2008**, *8*, 4012–24.

(24) Daniels, G. M.; Amara, S. G. Selective labeling of neurotransmitter transporters at the cell surface. *Methods Enzymol.* **1998**, *296*, 307–18.

(25) Sabarth, N.; Lamer, S.; Zimny-Arndt, U.; Jungblut, P. R.; Meyer, T. F.; Bumann, D. Identification of surface proteins of Helicobacter pylori by selective biotinylation, affinity purification, and two-dimensional gel electrophoresis. *J. Biol. Chem.* **2002**, *277*, 27896–902.

(26) Myers-Morales, T.; Cowan, C.; Gray, M. E.; Wulff, C. R.; Parker, C. E.; Borchers, C. H.; Straley, S. C. A surface-focused biotinylation procedure identifies the Yersinia pestis catalase KatY as a membrane-associated but non-surface-located protein. *Appl. Environ. Microbiol.* **2007**, *73*, 5750–9.

(27) Ge, Y.; Rikihisa, Y. Surface-exposed proteins of Ehrlichia chaffeensis. *Infect. Immun.* **2007**, *75*, 3833–41.

(28) Harding, S. V.; Sarkar-Tyson, M.; Smither, S. J.; Atkins, T. P.; Oyston, P. C.; Brown, K. A.; Liu, Y.; Wait, R.; Titball, R. W. The identification of surface proteins of Burkholderia pseudomallei. *Vaccine* **2007**, *25*, 2664–72.

(29) Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J.; Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **2007**, *389*, 1017–31.

(30) Schiess, R.; Mueller, L. N.; Schmidt, A.; Mueller, M.; Wollscheid, B.; Aebersold, R. Analysis of cell surface proteome changes via label-free, quantitative mass spectrometry. *Mol. Cell. Proteomics.* **2009**, *8*, 624–38.

(31) Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6591–6.

(32) Wu, S.; de Lencastre, H.; Tomasz, A. Sigma-B, a putative operon encoding alternate sigma factor of Staphylococcus aureus RNA polymerase: molecular cloning and DNA sequencing. *J. Bacteriol.* **1996**, *178*, 6036–42.

(33) Kullik, I.; Giachino, P.; Fuchs, T. Deletion of the alternative sigma factor sigmaB in Staphylococcus aureus reveals its function as a global regulator of virulence genes. *J. Bacteriol.* **1998**, *180*, 4814–20.

(34) Gertz, S.; Engelmann, S.; Schmid, R.; Ziebandt, A. K.; Tischer, K.; Scharf, C.; Hacker, J.; Hecker, M. Characterization of the sigma(B) regulon in Staphylococcus aureus. *J. Bacteriol.* **2000**, *182*, 6983–91.

(35) Ziebandt, A. K.; Becher, D.; Ohlsen, K.; Hacker, J.; Hecker, M.; Engelmann, S. The influence of agr and sigmaB in growth phase dependent regulation of virulence factors in Staphylococcus aureus. *Proteomics* **2004**, *4*, 3034–47.

(36) Bischoff, M.; Dunman, P.; Kormanec, J.; Macapagal, D.; Murphy, E.; Mounts, W.; Berger-Bachi, B.; Projan, S. Microarray-based analysis of the Staphylococcus aureus sigmaB regulon. *J. Bacteriol.* **2004**, *186*, 4085–99.

(37) Jonsson, I. M.; Arvidson, S.; Foster, S.; Tarkowski, A. Sigma factor B and RsbU are required for virulence in Staphylococcus aureus-induced arthritis and sepsis. *Infect. Immun.* **2004**, *72*, 6106–11.

(38) Entenza, J. M.; Moreillon, P.; Senn, M. M.; Kormanec, J.; Dunman, P. M.; Berger-Bachi, B.; Projan, S.; Bischoff, M. Role of sigmaB in the expression of Staphylococcus aureus cell wall adhesins ClfA and FnbA and contribution to infectivity in a rat model of experimental endocarditis. *Infect. Immun.* **2005**, *73*, 990–8.

(39) Lorenz, U.; Huttinger, C.; Schafer, T.; Ziebuhr, W.; Thiede, A.; Hacker, J.; Engelmann, S.; Hecker, M.; Ohlsen, K. The alternative sigma factor sigma B of Staphylococcus aureus modulates virulence in experimental central venous catheter-related infections. *Microbes Infect.* **2008**, *10*, 217–23.

(40) Nair, S. P.; Bischoff, M.; Senn, M. M.; Berger-Bachi, B. The sigma B regulon influences internalization of Staphylococcus aureus by osteoblasts. *Infect. Immun.* **2003**, *71*, 4167–70.

(41) Pane-Farre, J.; Jonas, B.; Forstner, K.; Engelmann, S.; Hecker, M. The sigmaB regulon in Staphylococcus aureus and its regulation. *Int. J. Med. Microbiol.* **2006**, *296*, 237–58.

(42) Shafer, W. M.; Iandolo, J. J. Genetics of staphylococcal enterotoxin B in methicillin-resistant isolates of Staphylococcus aureus. *Infect. Immun.* **1979**, *25*, 902–11.

(43) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680–5.

(44) Eymann, C.; et al. A comprehensive proteome map of growing Bacillus subtilis cells. *Proteomics* **2004**, *4*, 2849–76.

(45) Tabb, D. L.; McDonald, W. H.; Yates, J. R., 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **2002**, *1*, 21–6.

(46) Xia, Q.; Wang, T.; Taub, F.; Park, Y.; Capestany, C. A.; Lamont, R. J.; Hackett, M. Quantitative proteomics of intracellular Porphyromonas gingivalis. *Proteomics* **2007**, *7*, 4323–37.

(47) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–80.

(48) Nielsen, H.; Krogh, A. Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 122–30.

(49) Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **2004**, *340*, 783–95.

(50) Clarke, S. R.; Foster, S. J. Surface adhesins of Staphylococcus aureus. *Adv Microb Physiol.* **2006**, *51*, 187–224.

(51) Nakai, K.; Horton, P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* **1999**, *24*, 34–6.

(52) Park, S. K.; Venable, J. D.; Xu, T.; Yates, J. R., 3rd. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* **2008**, *5*, 319–22.

(53) Choi-Rhee, E.; Cronan, J. E. The biotin carboxylase-biotin carboxyl carrier protein complex of Escherichia coli acetyl-CoA carboxylase. *J. Biol. Chem.* **2003**, *278*, 30806–12.

(54) Ziebandt, A. K.; Weber, H.; Rudolph, J.; Schmid, R.; Hoper, D.; Engelmann, S.; Hecker, M. Extracellular proteins of Staphylococcus aureus and the role of SarA and sigma B. *Proteomics* **2001**, *1*, 480–93.

(55) Homerova, D.; Surdova, K.; Kormanec, J. Optimization of a two-plasmid system for the identification of promoters recognized by RNA polymerase containing Mycobacterium tuberculosis stress response sigma factor, sigmaF. *Folia Microbiol. (Praha)* **2004**, *49*, 685–91.

(56) Chavakis, T.; Wiechmann, K.; Preissner, K. T.; Herrmann, M. Staphylococcus aureus interactions with the endothelium: the role of bacterial "secretable expanded repertoire adhesive molecules" (SERAM) in disturbing host defense systems. *Thromb. Haemost.* **2005**, *94*, 278–85.

(57) Yeats, C.; Rawlings, N. D.; Bateman, A. The PepSY domain: a regulator of peptidase activity in the microbial environment. *Trends Biochem. Sci.* **2004**, *29*, 169–72.

(58) Dubrac, S.; Boneca, I. G.; Poupel, O.; Msadek, T. New insights into the WalK/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in Staphylococcus aureus. *J. Bacteriol.* **2007**, *189*, 8257–69.

(59) Becher, D.; Hempel, K.; Sievers, S.; Zühlke, D.; Pané-Faré, J.; Otto, A.; Fuchs, S.; Albrecht, D.; Engelmann, S.; Völker, U.; van Dijl, J. M.; Hecker, M. A proteomic view of an important human pathogen - towards the quantification of the entire *Staphylococcus aureus* proteome. *PLoS ONE* **2009**, *4* (12), e8176,. doi:10.1371/journal.pone.0008176.

(60) Rogasch, K.; et al. Influence of the two-component system SaeRS on global gene expression in two different Staphylococcus aureus strains. *J. Bacteriol.* **2006**, *188*, 7742–58.

(61) Otto, M. Virulence factors of the coagulase-negative staphylococci. *Front. Biosci.* **2004**, *9*, 841–63.

(62) Bergmann, S.; Rohde, M.; Chhatwal, G. S.; Hammerschmidt, S. alpha-Enolase of Streptococcus pneumoniae is a plasmin(ogen)-binding protein displayed on the bacterial cell surface. *Mol. Microbiol.* **2001**, *40*, 1273–87.

_____

## research articles

Hempel et al.

(63) Carneiro, C. R.; Postol, E.; Nomizo, R.; Reis, L. F.; Brentani, R. R. Identification of enolase as a laminin-binding protein on the surface of Staphylococcus aureus. *Microbes Infect.* **2004**, *6*, 604–8.

(64) Jin, H.; Song, Y. P.; Boel, G.; Kochar, J.; Pancholi, V. Group A streptococcal surface GAPDH, SDH, recognizes uPAR/CD87 as its receptor on the human pharyngeal cell and mediates bacterial adherence to host cells. *J. Mol. Biol.* **2005**, *350*, 27–41.

(65) Novick, R. P.; Geisinger, E. Quorum sensing in staphylococci. *Annu. Rev. Genet.* **2008**, *42*, 541–64.

(66) Deora, R.; Tseng, T.; Misra, T. K. Alternative transcription factor sigmaSB of Staphylococcus aureus: characterization and role in transcription of the global regulatory locus sar. *J. Bacteriol.* **1997**, *179*, 6355–9.

(67) Tegmark, K.; Karlsson, A.; Arvidson, S. Identification and characterization of SarH1, a new global regulator of virulence gene expression in Staphylococcus aureus. *Mol. Microbiol.* **2000**, *37*, 398–409.

(68) Dunman, P. M.; et al. Transcription profiling-based identification of Staphylococcus aureus genes regulated by the agr and/or sarA loci. *J. Bacteriol.* **2001**, *183*, 7341–53.

(69) Kim, H.; Marquis, H.; Boor, K. J. SigmaB contributes to Listeria monocytogenes invasion by controlling expression of inlA and inlB. *Microbiology* **2005**, *151*, 3215–22.

(70) Wertheim, H. F.; Walsh, E.; Choudhurry, R.; Melles, D. C.; Boelens, H. A.; Miajlovic, H.; Verbrugh, H. A.; Foster, T.; van Belkum, A. Key role for clumping factor B in Staphylococcus aureus nasal colonization of humans. *PLoS Med.* **2008**, *5*, e17.

(71) Stapleton, M. R.; Horsburgh, M. J.; Hayhurst, E. J.; Wright, L.; Jonsson, I. M.; Tarkowski, A.; Kokai-Kun, J. F.; Mond, J. J.; Foster, S. J. Characterization of IsaA and SceD, two putative lytic transglycosylases of Staphylococcus aureus. *J. Bacteriol.* **2007**, *189*, 7316–25.

(72) Senn, M. M.; Giachino, P.; Homerova, D.; Steinhuber, A.; Strassner, J.; Kormanec, J.; Fluckiger, U.; Berger-Bachi, B.; Bischoff, M. Molecular analysis and organization of the sigmaB operon in Staphylococcus aureus. *J. Bacteriol.* **2005**, *187*, 8006–19.

(73) Morikawa, K.; Maruyama, A.; Inose, Y.; Higashide, M.; Hayashi, H.; Ohta, T. Overexpression of sigma factor, sigma(B), urges Staphylococcus aureus to thicken the cell wall and to resist beta-lactams. *Biochem. Biophys. Res. Commun.* **2001**, *288*, 385–9.

(74) Bischoff, M.; Berger-Bachi, B. Teicoplanin stress-selected mutations increasing sigma(B) activity in Staphylococcus aureus. *Antimicrob. Agents Chemother.* **2001**, *45*, 1714–20.

(75) Bischoff, M.; Entenza, J. M.; Giachino, P. Influence of a functional sigB operon on the global regulators sar and agr in Staphylococcus aureus. *J. Bacteriol.* **2001**, *183*, 5171–9.

PR901143A

_____

*A Comprehensive Proteomics and Transcriptomics Analysis*

*of Bacillus subtilis Salt Stress adaptation*

# A Comprehensive Proteomics and Transcriptomics Analysis of *Bacillus subtilis* Salt Stress Adaptation[∇][†]

Hannes Hahne,[1][‡] Ulrike Mäder,[2] Andreas Otto,[1] Florian Bonn,[1] Leif Steil,[2] Erhard Bremer,[3] Michael Hecker,[1] and Dörte Becher[1][*]

*Institut für Mikrobiologie, Ernst-Moritz-Arndt Universität, Greifswald, Germany[1]; Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, Ernst-Moritz-Arndt Universität, Greifswald, Germany[2]; and Laboratory for Microbiology, Department of Biology, Philipps-Universität, Marburg, Germany[3]*

In its natural habitats, *Bacillus subtilis* is exposed to changing osmolarity, necessitating adaptive stress responses. Transcriptomic and proteomic approaches can provide a picture of the dynamic changes occurring in salt-stressed *B. subtilis* cultures because these studies provide an unbiased view of cells coping with high salinity. We applied whole-genome microarray technology and metabolic labeling, combined with state-of-the-art proteomic techniques, to provide a global and time-resolved picture of the physiological response of *B. subtilis* cells exposed to a severe and sudden osmotic upshift. This combined experimental approach provided quantitative data for 3,961 mRNA transcription profiles, 590 expression profiles of proteins detected in the cytosol, and 383 expression profiles of proteins detected in the membrane fraction. Our study uncovered a well-coordinated induction of gene expression subsequent to an osmotic upshift that involves large parts of the SigB, SigW, SigM, and SigX regulons. Additionally osmotic upregulation of a large number of genes that do not belong to these regulons was observed. In total, osmotic upregulation of about 500 *B. subtilis* genes was detected. Our data provide an unprecedented rich basis for further in-depth investigation of the physiological and genetic responses of *B. subtilis* to hyperosmotic stress.

*Bacillus subtilis*, a metabolically versatile Gram-positive bacterium, is able to adapt efficiently to a wide variety of stress conditions such as heat stress or phosphate deprivation. It is frequently exposed to nutrient limitations, and in response to such growth-restricting conditions, *B. subtilis* mounts a complex developmental program that results in the formation of a highly stress-resistant spore (14). In its natural habitat, the upper layers of soil, *B. subtilis* is often challenged with sudden and often long-lasting changes in osmolality due to flooding and drying of the soil. These extreme conditions threaten the cell with rupture or dehydration and necessitate active countermeasures to ensure survival and growth (7).

After sudden osmotic upshift, *B. subtilis* rapidly accumulates $K^+$ from the environment via two different $K^+$ uptake systems (KtrAB and KtrCD) (26). This initial phase of osmotic adaptation is followed by the accumulation of compatible solutes thereby permitting a reduction in the cellular $K^+$ level (65). Compatible solutes are either endogenously synthesized by *B. subtilis* or accumulated from exogenous sources (7). They counteract the deleterious effects of a hypertonic environment on cellular water content and cell physiology and allow *B. subtilis* to proliferate under a wide range of external osmolalities. Proline, the major compatible solute produced by *B.*

*subtilis* (65), is synthesized via a dedicated osmostress-responsive pathway that comprises at least two paralogue enzymes of the vegetative proline biosynthetic pathway (4, 28). These paralogue proline biosynthetic enzymes are ProH, a pyrroline-5-carboxylate reductase, and ProJ, a glutamate 5-kinase. Additionally, YerD, a putative ferredoxin-dependent glutamate synthase, might be involved in providing the precursor for proline biosynthesis subsequent to a severe osmotic upshift (28). *B. subtilis* possesses five osmotically regulated transport systems (OpuA to OpuE) for the acquisition of a broad spectrum of compatible solutes from environmental sources, and the uptake of these compounds provides a considerable degree of osmostress resistance (7).

Sudden osmotic down-shocks trigger a rapid influx of water, driving up turgor, which threatens the integrity of the cell. Like many other microorganisms, *B. subtilis* expels compatible solutes and ions via mechanosensitive channels (MscL and YkuT) to reduce the osmotic potential of its cytoplasm and thereby counteracts the influx of water (24, 64).

Under severe hyperosmotic conditions that no longer allow growth and cell division, *B. subtilis* activates the $\sigma^B$-controlled general stress response which provides a nonspecific and preemptive multiple stress resistance to the cells (21). Environmental and cellular stress conditions are detected by a complex set of sensor and regulatory proteins that are organized in the stressosome (43). Signal output from the stressosome activates the alternative transcription factor $\sigma^B$, which in turn triggers the coordinated transcription of about 150 genes comprising a general stress regulon (51, 52). Thirty-seven general stress proteins have an important function in cellular protection against a severe salt shock since the disruption of their structural genes causes a salt-sensitive phenotype. Such a phenotype is also

exhibited by a *sigB* mutant and by the disruption of *yerD*. Among the $\sigma^B$-controlled genes are those coding for the import systems for the compatible solutes glycine betaine and proline, OpuD and OpuE (28). Transcription of both the *opuE* and *opuD* genes is under simultaneous control of $\sigma^B$ and the vegetative sigma factor $\sigma^A$, thereby linking the nonspecific general stress response with the salt-specific adaptive response of *B. subtilis* (7, 21).

Two extracytoplasmic function (ECF) sigma factor regulons, the $\sigma^W$ and $\sigma^M$ regulons, are also induced subsequent to a hyperosmotic shock. Both the $\sigma^W$ and $\sigma^M$ regulons are typically induced by cell envelope stress, and $\sigma^M$ is essential for growth and cellular survival in high-salt concentrations (29, 51). Severe salt stress has pleiotropic effects on the physiology of *B. subtilis*: the composition of the cytoplasmic membrane is altered (40, 41), cell wall properties are adjusted (42), and swarming capability of the cells is largely impaired (57).

The physiological response of *B. subtilis* to changing osmolality has been analyzed in substantial detail at the level of single genes or proteins (7), but the potential of large-scale transcriptomics and proteomics approaches has not yet been fully explored. Such studies are well suited to provide a global view on the cellular responses to a particular environmental challenge. Hoffmann et al. used classical two-dimensional gel electrophoresis (2-DE) to compare protein expression patterns of *B. subtilis* cultures growing under low- and high-salinity conditions (25), and Höper et al. performed a detailed analysis of the protein synthesis pattern along the growth curve of severely salt-stressed *B. subtilis* using 2-DE in combination with dual-channel imaging and gel warping (27). A transcriptomics approach analyzing mRNA changes on a time-resolved scale following a hyperosmotic shock was performed by Steil et al., who focused on transcriptional changes that are taking place in cells of a Δ*sigB* mutant strain that were either continuously cultured at high salinity or subjected to an osmotic upshift (57).

Many membrane proteins, especially transporters of compatible solutes and ions, play a crucial role in adaptation of *B. subtilis* to salt stress. However, no study has yet addressed the changes in the membrane proteome of salt-stressed *B. subtilis* cells in a quantitative manner. Furthermore, quantitative studies investigating salt-induced changes in a prokaryotic membrane proteome are generally rare (6, 22, 31, 67, 68). For these reasons, we designed an experimental strategy that combined quantitative shotgun proteomics for the analysis of both the cytosolic and the cytoplasmic membrane proteome with whole-genome microarray data to profile in a time-dependent fashion the physiological adaptation of *B. subtilis* to a sudden and strong osmotic upshift.

## MATERIALS AND METHODS

**Growth conditions and preparation of the crude membrane fraction.** The *B. subtilis* 168 (*trpC2*) wild-type strain (2) was grown aerobically at 37°C and 180 rpm in Belitsky minimal medium (58) without glutamate to avoid its usage as precursor for biosynthesis of the compatible solute proline. Exponentially growing cells (optical density at 500 nm [$OD_{500}$] of 0.4) were challenged with 6% (wt/vol) NaCl, and samples were taken before and 10, 30, 60, and 120 min after the onset of stress. Cells were harvested by centrifugation (8,000 × g for 10 min at 4°C), and cell pellets were washed twice with TBS (50 mM Tris, 150 mM NaCl, pH 7.5). Cell lysis was performed in a French press (minicell; SLM Aminco, Rochester, NY). Cell debris was removed by centrifugation (20,000 × g for 20

min at 4°C), and the protein concentration of the supernatant was determined (NanoRotiquant; Roth, Germany).

The above-described growth experiment was performed three times in $^{14}$N-labeled medium and three times in $^{15}$N-labeled medium for relative quantification of proteins on a proteome-wide level using either normal or $^{15}$N-labeled ammonium sulfate as a supplement. $^{15}$N reagents were purchased from Euriso-tope (Saint Aubin Cedex, France). Equal amounts of protein from each $^{14}$N-labeled sample time point were mixed with the respective pooled reference proteins containing equal amounts of $^{15}$N-labeled samples from all five time points. The mixed $^{14}$N/$^{15}$N-labeled samples were subjected to ultracentrifugation (100,000 × g for 60 min at 4°C). The resulting pellet was designated as the crude membrane fraction, and an aliquot with a protein content of 50 mg was used as the starting material for subsequent membrane preparations. The supernatant was designated as the cytosolic fraction and analyzed further. A schematic representation of the workflow for the proteomics experiment is given in Fig. S1 in file S1 in the supplemental material.

**Membrane purification and SDS-PAGE of membrane and cytosolic proteins.** The membrane preparation was performed exactly as described previously (19). Briefly, to remove contaminating cytosolic proteins, the crude membrane fraction was subjected to a two-step purification procedure using a strongly alkaline carbonate buffer followed by a high-salt buffer. The purified membrane fraction was solubilized using SDS-PAGE sample buffer, and proteins from both subcellular fractions were separated *via* SDS-PAGE using a 10% (wt/vol) acrylamide minigel for membrane proteins and a 12% (wt/vol) minigel for cytosolic proteins. A gel lane was cut into 12 equidistant pieces, and the tryptic in-gel digestion and the following peptide elution for liquid chromatography-tandem mass spectrometry (LC-MS/MS) were carried out as described previously (15).

**Nano-LC-MS/MS analysis.** The nano-LC-MS/MS analysis of peptides derived from tryptic in-gel digestion was performed on a linear trap quadrupole (LTQ) Oribtrap (Thermo Fisher Scientific, Waltham, MA) equipped with a nano-ACQUITY UPLC (Waters, Milford, MA). Peptides were loaded onto a trapping column (nanoAcquity Symmetry UPLC column, $C_{18}$, 5 μm, 180 μm by 20 mm; Waters) at a flow rate of 10 μl/min and washed for 3 min with 99% buffer A. Peptides were then eluted and separated *via* an analytical column (nanoAcquity BEH130 UPLC column, $C_{18}$, 1.7 μm, 100 μm by 100 mm; Waters) with a decreasing buffer gradient (from 99% buffer A to 60% buffer B [0.1% acetic acid, 90% acetonitrile, distilled water [d$H_2$O]]) in a time frame of 80 min. The mass spectrometric analysis started with a full survey scan in the Orbitrap (*m/z* 300 to 2,000, resolution of 60,000) followed by collision-induced dissociation and acquisition of MS/MS spectra of the four most abundant precursor ions in the LTQ. Precursors were dynamically excluded for 30 s, and unassigned charge states as well as singly charged ions were rejected.

**Database search and filtering of data sets.** Sequest peak list files (*.dta) were generated from the raw instrument data (*.raw) using Extract_msn with default settings implemented in BioworksBrowser 3.3.1 SP1. The database search was performed with SEQUEST (version 27, rev. 12) (13), and the result files (*.out) were combined and filtered with DTASelect 1.9 and Contrast (59). The searched database contained the target sequences, which include the complete proteome set of *B. subtilis* extracted from UniProtKB (release 12.7) (63) and a set of common laboratory contaminants as well as a decoy database. The precursor tolerance was set to 10 ppm, and tolerance for fragment ions was set to 1 atomic mass unit (amu). In these searches, we only allowed fully tryptic peptides, two missed cleavage sites, and methionine oxidation (+15.9949) as differential modification. Only b- and y-ion series were included in the database search. All searches were performed for either light or heavy (with $^{15}$N as fixed modification) peptides using the same set of parameters. Peptide filter criteria were adjusted by keeping the empirically determined false-positive rate (FPR) (50) for each set of combined $^{14}$N and $^{15}$N search results on a protein level below 3% and on a peptide level below 1%. The following filter criteria were sufficiently stringent: minimum XCorr scores were set to 2.2, 2.8 and 3.5 for +2, +3 and +4 charged ions, respectively, and a ΔCN value of 0.1 was demanded.

**Quantification, normalization, and filtering of proteomic data.** The quantification of metabolically labeled peptides was performed using the quantification program Census (49). MS1 files were generated from the raw instrument data with the RawExtractor program (version 1.9.3) and SEQUEST search results of heavy and light peptides were combined with Contrast and filtered with DTA-Select 1.9, applying the above-mentioned filter criteria. Both MS1 files and the combined search results were passed to Census, and quantification was performed using the default setting for high-resolution mass spectrometry data. Quantified peptides were extracted from Census, using a default filter and a stringent determinant score of 0.8. Singleton peptides (i.e., peptides with one isotopomer below the detection threshold) as well as proteins with only one quantified unique peptide were discarded. Protein ratios were calculated from

peptide ratios using the weighted average approach implemented in Census. Weighted average protein ratios were $\log_2$ transformed and median centered. The final data set retained only those proteins that were quantified with at least two $^{14}$N- or $^{15}$N-labeled peptides per sample and that had at least two expression values from separate biological replicates at the initial reference time point. Expression profiles from different replicates were averaged (see Tables S1 and S2 in file S2 in the supplemental material).

**Transcriptome analysis.** Cell harvesting and preparation of total RNA were performed as described previously (16). RNA samples were DNase treated and purified using the RNeasy minikit (Qiagen, Hilden, Germany). The quality of RNA preparations was controlled using the Agilent 2100 Bioanalyzer according to the manufacturer's instructions (Agilent Technologies, Waldbronn, Germany). Generation of the Cy3/Cy5-labeled cDNAs and hybridization to whole-genome DNA microarrays (Eurogentec, Cologne, Germany) containing DNA fragments originating from 4,022 *B. subtilis* genes as duplicate spots were performed as described by Jürgen et al. (35). The slides were scanned with a ScanArray Express scanner (PerkinElmer Life and Analytical Sciences, Monza, Italy), and signal intensities of the individual spots were quantified with the ScanArray Express image analysis software (PerkinElmer Life and Analytical Sciences). Three hybridizations with mRNA from three biological replicates were performed with the samples labeled with Cy5 and a pooled reference labeled with Cy3, resulting in six measurements per gene. Genes were considered significantly expressed if 70% of the feature pixels had intensities more than 2 standard deviations above the background pixel intensity in four of the six measurements. Intensity-dependent (Lowess) normalization was performed using the GeneSpring software (Agilent Technologies), and the ratios of duplicate spots were averaged for further analyses, resulting in three biologically independent expression values per time point. The expression profiles from three biological replicates were averaged, and incomplete expression profiles discarded. The mRNA expression data are available as supplemental material (see Table S3 in file S2 in the supplemental material).

**Statistical, clustering, and functional annotation analyses.** The TIGR Multiexperiment Viewer belonging to the TM4 software suite (53) was used for one-way analysis of variance (ANOVA), figure of merit (FOM) analyses, and K-medians clustering (KMC) of both the transcriptomic and the proteomic data sets. The confidence level for the ANOVA results was set to 99.9% ($P < 10^{-3}$) for transcriptomic data and to 95% ($P < 0.05$) for proteomic data. The KMC was only performed on statistically significant mRNA and protein expression profiles. The number of clusters for the KMC approach was derived using the figure of merit implementation of the TM4 software suite (53). Briefly, a figure of merit is an estimate of the predictive power of a clustering algorithm. Since the FOM and KMC calculations were both performed in the Euclidean space, the average expression profiles were standardized in order to have a mean expression value of zero and standard deviation of 1. This simple standardization ensures that similar expression profiles are close in Euclidean space, thereby clustering together. The averaged and standardized mRNA and protein expression profiles were subjected to KMC using 11 clusters for transcriptomic data and nine clusters for proteomic data. Euclidean distance was used as metric, and a maximum of 50 iterations was allowed. The Gene Functional Classification tool provided by DAVID (Database for Annotation, Visualization and Integrated Discovery) bioinformatics resources was used to derive the first biologically meaningful interpretations of gene or protein clusters (11, 30), and for all DAVID tools, default parameters were employed. Lists of regulon members were compiled for the $\sigma^B$, $\sigma^M$, $\sigma^W$, and $\sigma^X$ regulons using the DBTBS database (55) and additional data from genome-wide studies (8, 12, 33, 51, 52).

**Membrane protein and signal peptide prediction.** Four algorithms were used to predict integral membrane proteins (IMPs) and transmembrane helices (TMH): TMHMM2.0 (39, 56), SOSUI1.1 (23, 45, 46), HMMTOP2.0 (61, 62), and Phobius (37, 38). Signal peptides were predicted using LipoP1.0 (34), SignalP3.0 (5, 47), and Phobius. Prediction results were combined in the following way. (i) Integral membrane proteins were considered if at least two of all four algorithms (TMHMM2.0, SOSUI, HMMTOP, and Phobius) predicted transmembrane helices and (in order to discriminate between N-terminal TMHs and signal peptides) if Phobius predicted at least one TMH or no signal peptide sequence. (ii) Lipoproteins were considered if LipoP1.0 predicts a signal peptidase II cleavage site and Phobius a signal peptide. (iii) Finally, secreted proteins were considered if SignalP3.0 as well as Phobius predicted a signal peptide and if the protein was not assigned as a lipoprotein using the programs mentioned in parameter ii. This cellular localization prediction of proteins resulted in the assignment of 1,027 integral membrane proteins, 99 lipoproteins, and 274 Sec-type secreted proteins in the complete proteome set of *B. subtilis*.

**Microarray data accession number.** The microarray data are accessible through GEO Series accession no. GSE18345 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE18345).

## RESULTS AND DISCUSSION

Aiming for a comprehensive picture of the transcriptional response of *B. subtilis* to a sudden and strong increase in medium osmolality, we monitored the mRNA expression profile using whole-genome microarrays. We complemented this data set with a $^{14}$N/$^{15}$N metabolic labeling strategy in order to track changes in the relative protein levels of the cytosol and the cytoplasmic membrane proteome.

Bacterial cultures were propagated in a minimal medium to early exponential phase ($OD_{500}$ of 0.4) and were then exposed to a strong osmotic upshift by the addition of 6% (wt/vol) NaCl. This sudden increase in salinity resulted in a growth arrest for 60 to 90 min, after which growth was resumed at a strongly reduced rate (27, 60). Samples for the proteomic as well as for the transcriptomic experiments were taken along the growth curve before and 10, 30, 60, and 120 min after the addition of NaCl.

**Global changes on transcriptome and proteome level.** Overall, in our proteome study we could identify 2,169 proteins with at least one unique peptide (FPR, 11.2% on protein level) and 1,710 proteins with at least two unique peptides (FPR, 0.6%) along the growth curve. The proteins identified in our study represent 53% and 42% of the theoretical protein complement of *B. subtilis*, respectively. Moreover, the protein identifications are based on 14,334 and 13,875 unique peptide identifications (FPR, 1.9% and 0.1%, respectively, on the peptide level). With an average of 6.6 unique peptides or about 26% average sequence coverage per protein identification, this large data set represents a rich resource for further analyses (targeted analyses and/or reanalyses) of changes in the *B. subtilis* proteome under salt stress conditions, but this data set (for details, see files S1 and S3 in the supplemental material) might prove useful also for studies not connected with the cellular osmotic stress response.

To monitor global changes in gene transcription subsequent to an osmotic upshift, microarrays of the entire gene inventory of *B. subtilis* were used to screen the transcription profile of the osmotically challenged cells at various times. The results of these experiments were then compared to changes observed in the proteome profile of these cells (Fig. 1). Immediately after the hyperosmotic shock with 6% (wt/vol) NaCl, the *B. subtilis* cells ceased growth, and concomitantly the number of significantly expressed mRNA species decreased dramatically from 3,447 in exponentially growing cells down to 2,007 mRNA species 30 min after the cells were exposed to salt stress. In contrast, the total number of identified proteins remained nearly unaffected by the salt treatment of the *B. subtilis* cultures. Only the proteins identified in the membrane fraction showed a substantial increase (20%) within the first 10 min after salt stress. Of the 218 proteins identified, only a subset (30%) can be considered as truly associated with the cytoplasmic membrane either as integral membrane proteins, lipoprotein, or secreted or peripheral membrane protein. The remaining 70% of the proteins detected in the membrane fraction are of cytosolic origin. A functional annotation analysis of these
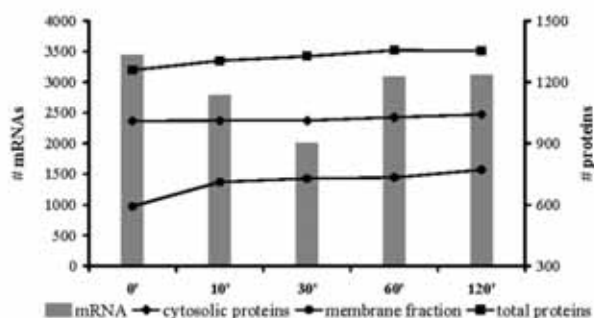
170

FIG. 1. Numbers of significantly expressed transcripts and reliably identified proteins along the growth curve. Microarrays were screened for significantly expressed genes using pixel statistics as described in Materials and Methods, and the corresponding numbers of protein identifications along the growth curve are based on at least two unique peptide identifications from all three biological replicates.

proteins using the DAVID Gene Functional Classification tool revealed five overrepresented groups among these cytosolic proteins: ribosomal proteins, proteins involved in histidine and arginine biosynthesis, aminoacyl-tRNA synthetases, and glycolytic enzymes. It is certainly possible that these proteins end up in the membrane fraction due to their physiological roles during salt stress adaptation. However, one should keep in mind that denatured proteins or proteins subjected to degradation might appear in the membrane fraction as well.

In contrast, more than 70% of the 98 proteins identified in growing cells—but not in early salt-exposed cells—are associated with the cytoplasmic membrane. This finding may point to either a general protection mechanism against hyperosmotic shock or the vulnerability of membrane proteins to high salt concentrations.

***K*-medians clustering of significantly regulated mRNA and proteins.** The stringent filtering procedures of the transcriptome and proteome data set resulted in a comprehensive data set of 3,961 mRNA expression profiles as well as expression profiles for 590 proteins identified in the cytosolic fraction and 383 proteins identified in the membrane fraction. Among these, 165 proteins could be quantified in either subcellular fraction. Within these data sets, the one-way ANOVA identifies 949 mRNAs with significant changes at $P < 10^{-3}$, as well as 29 proteins in the cytosol and 59 proteins in the membrane fraction with significant changes at $P < 0.05$. Among the 59 significant proteins of the membrane fraction were, with overlapping predictions, 17 integral membrane proteins, 6 lipoproteins, 5 secreted proteins, and 33 proteins that were either peripheral membrane proteins or cytosolic contaminants. The difference between significantly regulated mRNAs and proteins illustrates nicely the different induction behavior and turnover rates of bacterial mRNAs (with half-lives between seconds and minutes) (20) and proteins (half-lives of minutes to hours).

KMC was applied to both the statistically significant transcriptome and proteome data sets in order to group mRNAs and proteins with a similar temporal expression pattern in response to the imposed osmotic upshift. The results are displayed in Fig. 2. The corresponding functionally enriched gene and protein groups are summarized in Table 1, and detailed information is available as supplemental material (see files S4A and S4B in the supplemental material). Overall, mRNA and protein clusters are in good accordance. In particular, the temporary downregulation of most vegetative functions due to the growth arrest of the osmotically challenged *B. subtilis* cells is well reflected by both the mRNA and the protein complement. However, most changes in response to the osmotic upshift at the mRNA level were observed within the first 30 to 60
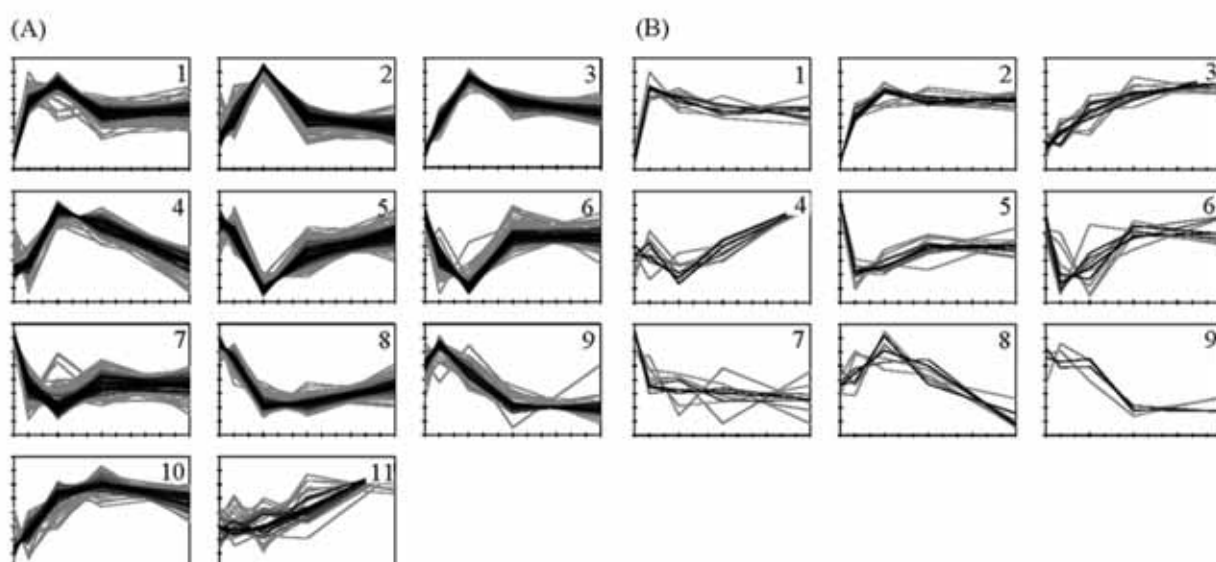
(A)                                                (B)



FIG. 2. Results of the *K*-medians clustering of the transcriptomic (A) and proteomic (B) expression profiles. The cytosolic and the membrane proteomes were clustered together. The expression profiles are color-coded, corresponding to their similarity (correlation coefficient) to the cluster median. The KMC results as well as the respective functional analysis data are available in files S4A and S4B in the supplemental material.

171

TABLE 1. Summary of functional annotation enrichment analyses for the KMC results of the transcriptomic and proteomic data

| Cluster | No. of members | Enriched functional group(s) | DAVID functional annotation clustering result(s) (P value)[a] | Quantile (%) | Fold change (vs. 0 min) at: | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | 10 min | 30 min | 60 min | 120 min |
| **Transcriptome** | | | | | | | | |
| 1 | 127 | This cluster contains 54 genes previously assigned to the σ$^B$-dependent general stress regulon. Of these, 46 show a minimum 8-fold induction 30 min after osmotic upshift. | Stress response (0.035), CBS domain (0.041) | 75 | 10.9 | 19.2 | 6.0 | 7.2 |
| | | | | 50 | 5.2 | 10.5 | 3.3 | 3.9 |
| | | | | 25 | 3.0 | 4.3 | 2.2 | 2.2 |
| 2 | 104 | This cluster contains 83 genes of unknown function. The remaining 21 genes are functionally very divergent. 47 genes encode IMPs, lipoproteins, or secreted proteins. | Membrane protein (0.026), DNA binding (0.031) | 75 | 2.0 | 6.1 | 1.7 | 1.5 |
| | | | | 50 | 1.5 | 3.6 | 1.4 | 1.2 |
| | | | | 25 | 1.1 | 2.5 | 1.1 | −1.1 |
| 3 | 109 | This cluster contains 23 σ$^B$-dependent genes and 27 genes involved in transport processes (e.g., *glnT, gutP, mntACDH, mrpBCDFG*, and *opu* transporters). | Transporter (5.9E−4), lipid-anchored proteins/ABC transporter (0.006), amino/organic acid transport (0.009), tellurium resistance protein (0.013), stress response (0.014) | 75 | 3.2 | 10.5 | 5.0 | 3.8 |
| | | | | 50 | 1.9 | 6.1 | 3.3 | 2.5 |
| | | | | 25 | 1.6 | 3.9 | 2.3 | 2.0 |
| 4 | 74 | This cluster contains 53 genes of unknown function. The remaining 21 genes are functionally very divergent. This cluster also includes 26 genes encoding IMPs. | Membrane protein (0.046) | 75 | 1.2 | 3.7 | 2.9 | 1.4 |
| | | | | 50 | 1.1 | 2.4 | 2.0 | 1.1 |
| | | | | 25 | −1.1 | 1.8 | 1.4 | −1.2 |
| 5 | 114 | This cluster shows strong enrichment of genes of primary metabolism (TCA cycle, metabolism of amino acids, nucleotides, lipids, and coenzymes). | Amino acid biosynthesis (1.5E−5), primary metabolism/biosynthesis (1.6E−5), ligase (i.e., aminoacyl-tRNA synthtases), lysine biosynthesis (0.012), arginine biosynthesis (0.013), coenzyme/prosthetic group biosynthesis (0.029), TCA cycle (0.034), nucleotide-binding (0.040) | 75 | 1.0 | −2.2 | −1.3 | 1.0 |
| | | | | 50 | −1.2 | −3.0 | −1.7 | −1.2 |
| | | | | 25 | −1.5 | −4.6 | −2.2 | −1.6 |
| 6 | 99 | This cluster shows enrichment of genes of the primary metabolism (oxidative phosphorylation, glycolysis, amino acid, nucleotide, lipid, and coenzyme metabolism). | SAM and 4Fe-4S enzymes (0.029), metal binding (0.033), (phospho)transferase (0.048) | 75 | −1.7 | −2.6 | −1.1 | −1.2 |
| | | | | 50 | −2.1 | −3.5 | −1.4 | −1.4 |
| | | | | 25 | −3.2 | −5.5 | −1.8 | −1.6 |
| 7 | 99 | This cluster comprises 61 genes of unknown function and several small groups with divergent functional annotations. | 2-component system (0.060)* | 75 | −2.1 | −2.6 | −1.9 | −2.0 |
| | | | | 50 | −2.6 | −3.5 | −2.5 | −2.4 |
| | | | | 25 | −3.4 | −5.2 | −3.1 | −3.1 |
| 8 | 105 | Chemotaxis and motility genes are strongly enriched in this cluster. Moreover, this cluster contains genes involved in a broad range of transport processes and amino acid metabolism. | Motility (7.6E−7), chemotaxis (0.003), membrane protein (0.007), methyl-accepting transducer (0.012), amino acid biosynthesis (0.017) | 75 | −1.1 | −2.9 | −2.8 | −1.9 |
| | | | | 50 | −1.4 | −3.7 | −3.9 | −2.5 |
| | | | | 25 | −1.8 | −5.8 | −6.2 | −3.5 |
| 9 | 32 | This cluster comprises 18 genes coding for the phage-like element PBSX. | LysM repeat-containing proteins/cell wall (0.016) | 75 | 1.4 | 1.1 | −1.4 | −1.5 |
| | | | | 50 | 1.3 | −1.1 | −1.6 | −1.7 |
| | | | | 25 | 1.2 | −1.4 | −2.1 | −2.0 |
| 10 | 58 | This cluster contains no prevalent functional gene group(s), but many small groups. Among these are genes for proline biosynthesis and Opu transporters. | Proline/glutamine family biosynthesis (0.055)* | 75 | 1.6 | 3.2 | 4.8 | 3.0 |
| | | | | 50 | 1.3 | 2.3 | 3.1 | 2.1 |
| | | | | 25 | 1.1 | 1.8 | 2.3 | 1.6 |
| 11 | 28 | This cluster contains no prevalent functional gene group(s). One small group is related to sporulation and germination (5 genes). | Sporulation/germination (0.083)* | 75 | 1.3 | 1.5 | 1.9 | 4.0 |
| | | | | 50 | 1.1 | 0.0 | 1.5 | 2.5 |
| | | | | 25 | −1.1 | −1.2 | 1.3 | 2.3 |
| **Proteome** | | | | | | | | |
| 1 | 10 | All proteins are non-membrane proteins of the primary metabolism detected in the membrane fraction. | Biosynthetic process (0.182)†, ATP binding (0.184)† | 75 | 2.3 | 2.2 | 1.9 | 1.9 |
| | | | | 50 | 2.0 | 1.9 | 1.6 | 1.6 |
| | | | | 25 | 1.9 | 1.7 | 1.5 | 1.5 |
| 2 | 15 | Peripheral membrane proteins as well as ribosomal proteins dominate this cluster. In addition, three (putative) detoxification proteins (SodA, YaaN, YceC) grouped in this cluster were detected in the membrane fraction. | Ribosomal proteins (0.080) | 75 | 2.1 | 2.6 | 2.9 | 2.7 |
| | | | | 50 | 1.9 | 2.3 | 2.3 | 2.1 |
| | | | | 25 | 1.6 | 2.2 | 1.9 | 1.9 |
| 3 | 12 | Membrane proteins involved in osmostress resistance (Opu proteins, YtxH, and YuaG) are enriched in this cluster. | Glycine betaine/proline transport (0.036) | 75 | 1.2 | 1.9 | 2.8 | 3.1 |
| | | | | 50 | 1.1 | 1.6 | 2.1 | 2.4 |
| | | | | 25 | −1.0 | 1.3 | 1.7 | 2.0 |

A Comprehensive Proteomics and Transcriptomics Analysis of *Bacillus subtilis* Salt Stress adaptation
_____

Vol. 192, 2010                    ANALYSIS OF *B. SUBTILIS* SALT STRESS ADAPTATION    875

TABLE 1—*Continued*

| Cluster | No. of members | Enriched functional group(s) | DAVID functional annotation clustering result(s) (*P* value)[a] | Quantile (%) | Fold change (vs. 0 min) at: | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | 10 min | 30 min | 60 min | 120 min |
| 4 | 8 | Cytosolic proteins with only very weak significant changes are comprised in this cluster. The assignment of any biological significance to this cluster is therefore difficult or questionable. | Oxidoreductase (0.020) | 75 | 1.1 | −1.0 | 1.1 | 1.4 |
| | | | | 50 | 0.0 | −1.1 | 1.1 | 1.3 |
| | | | | 25 | −1.0 | −1.1 | −1.0 | 1.2 |
| 5 | 13 | 12 proteins of this cluster are known/predicted to be IMPs, lipoproteins, or secreted proteins. All proteins of this cluster were detected in the membrane fraction. | Membrane protein (0.038) | 75 | −1.4 | −1.3 | −1.2 | −1.2 |
| | | | | 50 | −1.6 | −1.5 | −1.2 | −1.3 |
| | | | | 25 | −1.7 | −1.6 | −1.4 | −1.4 |
| 6 | 10 | 7 proteins of this cluster are known to be IMPs or lipoproteins. These proteins were detected in the membrane fraction. | Membrane protein (0.402)† | 75 | −1.3 | −1.3 | 1.0 | 0.5 |
| | | | | 50 | −1.9 | −1.6 | −1.1 | −1.1 |
| | | | | 25 | −2.7 | −3.3 | −1.1 | −1.3 |
| 7 | 9 | 8 proteins of this cluster are IMPs, lipoproteins, or known to be cell surface associated, mostly detected in the membrane fraction. Two cell wall proteins, PonA and PbpC, also fall within this group. | Cell envelope biogenesis/developmental process (0.192)† | 75 | −1.3 | −1.3 | −1.4 | −1.3 |
| | | | | 50 | −1.6 | −1.5 | −1.8 | −1.8 |
| | | | | 25 | −1.6 | −1.9 | −2.2 | −2.0 |
| 8 | 7 | All proteins are of cytosolic origin and show only very weak significant changes. The inference of biological significance is therefore difficult or questionable. | Primary metabolic process (0.206)† | 75 | 1.1 | 1.2 | 1.1 | −1.1 |
| | | | | 50 | 1.0 | 1.1 | 1.0 | −1.1 |
| | | | | 25 | 0.0 | 1.1 | 1.0 | −1.1 |
| 9 | 4 | 2 proteins of this very small cluster are involved in motility (Hag and FlgG), show a ≥2-fold downregulation, and were detected in the membrane fraction. | No terms clustered | 75 | 1.2 | −0.4 | −1.5 | −1.2 |
| | | | | 50 | 0.0 | −1.1 | −1.8 | −1.6 |
| | | | | 25 | −1.1 | −1.2 | −2.1 | −2.1 |

[a] Different significance thresholds for the DAVID Functional Annotation Clustering approach were applied for the transcriptomic data and the proteomic data. Since the number of genes per cluster is reasonably high for the transcriptomic data, a *P* value of ≤0.05 was applied. Because the number of proteins per cluster is, with a few exceptions, rather low, a less stringent *P* value (≤0.1) was chosen as a significance threshold. The *P* value is given in parentheses. *, no significant (*P* ≤ 0.05) functional annotation cluster (the next best hit is displayed); †, no significant (*P* ≤ 0.1) functional annotation cluster (the next best hit is displayed).

min, whereas the corresponding changes in the protein pattern occur at a far slower rate. The changes in the proteome of the salt-stressed cells were probably not complete at the last time point (120 min) analyzed in our experiments.

**General and cell envelope stress response.** The initial response of *B. subtilis* to a sudden and strong osmotic upshift is characterized by a strong and immediate activation of the σ$^B$ regulon, leading to synthesis and accumulation of proteins conferring a broad stress resistance to *B. subtilis* cells (21). The σ$^B$ regulon comprises at least 37 proteins that are critical for providing osmostress resistance, as judged by the salt-sensitive phenotype of their structural genes (28). Almost all of these σ$^B$-dependent salt stress protective proteins are significantly upregulated either early at the mRNA level or in later adaptation stages of the osmotically challenged cells on the protein level (Fig. 3A). Only a small number of these proteins are of known function. However, their immediate and strong induction subsequent to the osmotic up-shock and, at later times, their accumulation in osmotically stressed cells strongly suggest a physiologically important function for this group of proteins within the acclimatization process of *B. subtilis* to high-salinity surroundings.

In addition to the σ$^B$ regulon, three large ECF regulons (σ$^M$, σ$^W$, and σ$^X$), which specifically govern the physiological response to cell envelope stress, are affected by a suddenly increase in the external osmolality. The σ$^M$ and σ$^W$ regulons are known to be involved in salt stress adaptation (29, 51). A compiled set of 132 known or probable members of the σ$^M$, σ$^W$, and σ$^X$ regulons was sorted according to expression patterns and regulon memberships (Fig. 3B; see Table S1 in file S5 in the supplemental material). This classification revealed four groups of genes with similar expression profiles, each of which was associated with one prevalent ECF transcription factor. The color-coded expression pattern of these genes indicates a sequential activation of the general and cell-surface stress regulons. The σ$^B$ regulon is activated immediately after the shift to high salinity, and the onset of σ$^B$-controlled gene expression is followed by the induction of the σ$^W$ regulon, which shows maximum activation after 30 min. The σ$^M$ regulon is the last to be induced, with maximum induction 60 min after the salt shock. We also observed partial repression of the σ$^X$ regulon. Our data clearly show that there is a sequential activation of the transcription of the σ$^B$, σ$^W$, and σ$^M$ regulons, but one should keep in mind that there is evidence for functional and regulatory redundancy of the ECFs (44).

We observed that the differential activation of the above-discussed stress regulons at the mRNA level by salt stress was not reflected one-to-one at the protein level. Fifty-four expression profiles were acquired for 43 proteins whose structural genes are under the (partially overlapping) control of σ$^M$, σ$^W$, and σ$^X$ (see Table S2 in file S5 in the supplemental material). Remarkably, among this set of proteins, only six proteins showed significant changes in response to salt stress. The penicillin-binding protein PonA and the membrane-localized protease FtsH are under the control of σ$^M$, and both proteins exhibited a significant downregulation in the membrane fraction. In contrast, three "Y" proteins (YaaN, YuaG, and YuaI) under the control of σ$^W$ show significant upregulation. In ad-
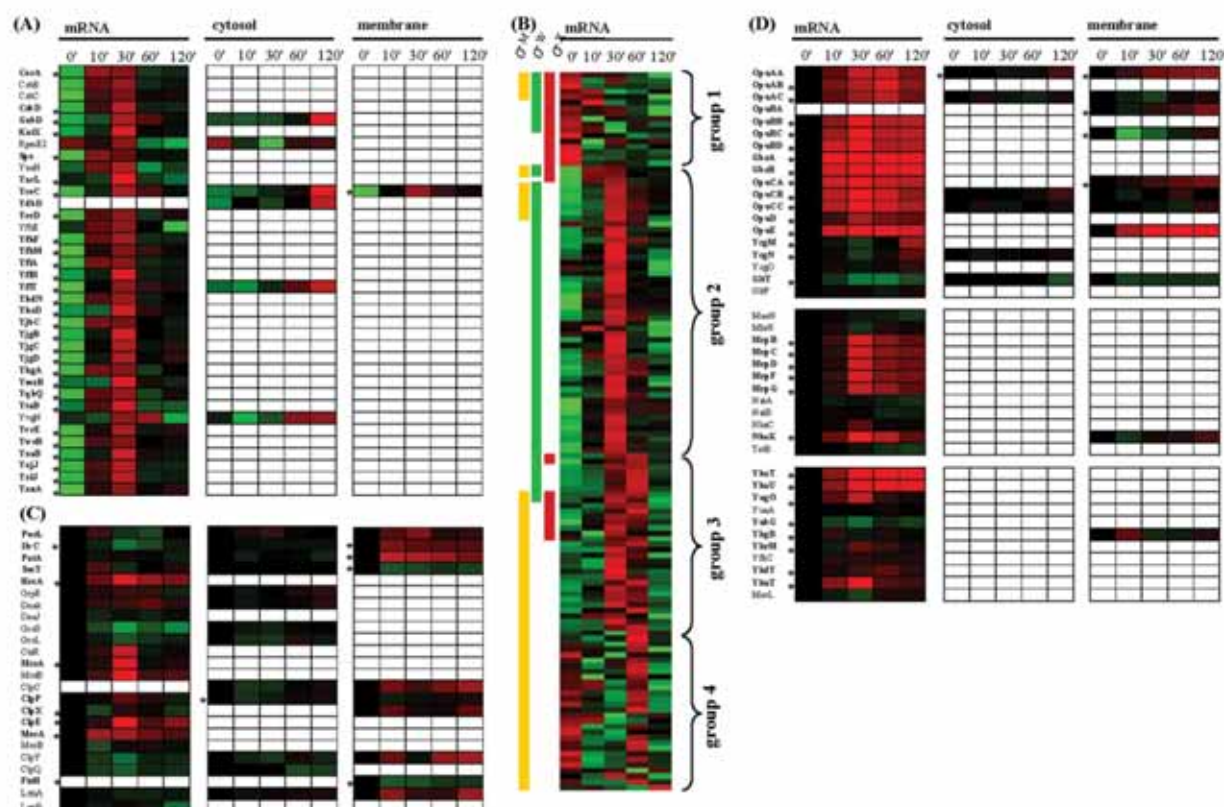
173

FIG. 3. General, cell envelope, and specific osmostress responses. Expression profiles shown in panels A and B are based on standardized log₂ ratio values, and expression profiles in panels C and D are log₂ ratio based. The minimum and maximum displayed log₂ ratios are $\pm 3$ for the transcriptomics data and $\pm 2$ for the proteomics data. Genes/proteins exhibiting significant changes are asterisked. (A) The 35 proteins given here belong to the $\sigma^B$-dependent general stress regulon and were already shown to be indispensable for the adaptation to high salinity (28). (B) Displayed are 131 mRNAs which are presumed or shown to be under the control of either $\sigma^M$, $\sigma^W$, or $\sigma^X$. In order to reveal overall expression patterns, the mRNA expression profiles were sorted according to their ECF membership and their expression profile. (C) Expression profiles of chaperones and ATP-dependent proteases as well as regulator and adapter proteins involved in protein quality control are shown. (D) Specific osmostress response proteins. The upper group contains proteins involved in uptake and synthesis of compatible solutes, whereas both lower groups comprise systems (presumably) involved in Na⁺ extrusion (middle group) and K⁺ extrusion (lower group).

dition, YceC, which is a member of the detoxification operon *yceCDEFGH* and under the control of $\sigma^A$, $\sigma^M$, and $\sigma^W$ is significantly upregulated in the membrane fraction. Proteins under the control of $\sigma^X$ (DltD, LytR, PbpX, YabP, YwbL, and YwbM) are mostly integral membrane proteins, lipoproteins, or secreted proteins and were detected in our experiments in the membrane fraction. Each of these six proteins exhibits a nonsignificant downregulation in the salt-stressed cells.

Overall, expression profiles of proteins and mRNAs correspond reasonably well, but some proteins exhibit regulation patterns that are opposed to their observed profile at the mRNA level. Several explanations for this unexpected behavior are conceivable. Since many ECF-controlled genes encode for proteins that are known to be secreted or localized at the cellular surface, a plausible explanation is that the respective proteins are subject to a localization shift upon stimulation (e.g., secretion). However, targeted protein degradation and mRNA-related processes are possible explanations. In any case, further experimental evidence is required to elucidate the

physiological significance of the observed counterintuitive changes in the context of hyperosmotic stress.

**Protein quality control.** In *B. subtilis*, the functional integrity of proteins is ensured through protein quality control mechanisms that include the action of molecular chaperones, protein-folding catalysts, and ATP-dependent proteases. As mentioned above, shortly after the shift to hyperosmotic conditions numerous cytosolic proteins with predicted vegetative functions were observed to be enriched in the membrane fraction. A representative set of these types of proteins (PurL, phosphoribosylformylglycinamidine synthase II involved in purine biosynthesis; IlvC, ketol-acid reductoisomerase involved in branched-chain amino acid synthesis; and PatA, an aminotransferase involved in lysine biosynthesis) is shown in Fig. 3C. We also found that numerous integral membrane proteins or peripheral membrane proteins (e.g., ATP-binding proteins of ABC transporters) showed a clear downregulation in the membrane fraction after hyperosmotic shock. These proteins are represented in Fig. 3C by SecY, which possesses 10 transmem-

brane helices and belongs to the Sec protein secretion system. Both facts indicate the involvement of protein quality control mechanisms in response to the osmotic upshift, either in the refolding or degradation of unfolded vegetative proteins or in the (targeted) degradation of membrane proteins. Figure 3C shows expression profiles of some quality control proteins. Significant induction on mRNA level could be observed for *hrcA*, *mcsA*, *clpP*, *clpX*, *clpE*, and *mecA*. On the protein level, it is interesting to note that an enrichment of chaperones (ClpC, ClpX, and ClpY) is observed in the membrane fraction concomitantly with their depletion in the cytosolic fraction. In addition, the ATP-dependent LonA protease was upregulated in the membrane fraction and hence could carry out degradation of misfolded proteins, while the membrane-anchored protease FtsH, which degrades both cytosolic and membrane proteins (54), was itself subject to degradation. The issue of whether the observed enrichment of vegetative proteins in the membrane fraction corresponds to a real physiologically relevant change in protein localization during salt stress needs further careful studies. However, our data hint that the sudden imposition of hyperosmotic conditions might induce denaturation and misfolding of both cytosolic and membrane proteins in *B. subtilis*.

**Osmospecific response.** The stress response of *B. subtilis* to sudden increases in osmolality comprises not only $\sigma^B$-controlled general stress proteins, but also many salt-specific stress proteins. The transcription of the genes encoding this latter group of proteins appears to be mostly independently from alternative sigma factors. Many of the salt-specific proteins are of known function. These proteins display a significant upregulation at either the mRNA or protein level (Fig. 3D). In particular, all Opu transporters (OpuA to OpuE) for compatible solute acquisition are immediately upregulated at the mRNA level, with maximum induction after 30 min, and the corresponding transport proteins accumulate to significant levels in the cytoplasmic membrane after 60 or 120 min. It has to be noted in this context that the specific substrates of these osmoprotectant uptake systems (7) were not present in the growth medium. Consequently, their induction reflects their osmotic control. When *B. subtilis* uses proline as a nutrient rather than as an osmoprotectant, this amino acid is imported *via* the high-affinity proline uptake system YcgO and degraded by the YcgM and YcgN enzymes to glutamate (7). Expression of the *ycgMNO* operon is induced by low levels of extracellular proline (S. Moses and E. Bremer, unpublished data). We observed a significant upregulation of the transcription of the *ycgMNO* operon in the salt-stressed cells in the late adaptation phase; this apparent osmotic induction of the *ycgMNO* operon probably reflects the leakage of proline from the osmotically stressed and hence proline-producing cells and the subsequent induction of *ycgMNO* expression (S. Moses and E. Bremer; unpublished data). The H$^+$/glutamate symporter GltP did not show a significant change in mRNA level, whereas the H$^+$ and Na$^+$/glutamate symporter GltT was downregulated at both the mRNA and the protein level, consistent with the cellular need for a tightly controlled low intracellular Na$^+$ level. Twenty-nine possible candidates for the active Na$^+$ extrusion systems were extracted from UniProtKB using the GO annotation "sodium ion transport" (GO:0006814). Among these, two Na$^+$/H$^+$ antiporters were found to be significantly upregulated

in the osmotically challenged *B. subtilis* cells: (i) the *mrpABCEDEFG* operon, involved in pH homeostasis (32), and the Mrp complex, which probably constitutes the major Na$^+$ extrusion system operating in *B. subtilis*; and (ii) NhaK, an Na$^+$, K$^+$, Li$^+$, and Rb$^+$/H$^+$ antiporter (18). Disruption of the *mrpABCEDEFG* operon causes a strong sensitivity toward Na$^+$ (36), and upregulation of *nhaK* expression in response to either NaCl or KCl has already been reported (18).

When *B. subtilis* is exposed to a sudden shift from hyper- to hypo-osmotic conditions, the cells have to jettison very rapidly those ions and organic solutes that were accumulated under high-osmolality growth conditions in order to reduce water entry (7). Expulsion of these solutes by *B. subtilis* is accomplished *via* the transient opening of mechanosensitive channels (24). A single channel of large conductivity (MscL type) and three channel proteins of small conductivity (MscS types YhdY, YkfC, and YkuT) are present in *B. subtilis* (24, 64). We observed a minor reduction in the mRNA level for the *mscL* gene, but found a significant induction on the mRNA level for the *yhdY* and *ykuT* genes in our transcriptome study. An upregulation of the *yhdY* and *ykuT* genes at high salinity has been observed previously (64). The transient opening of mechanosensitive channels in *B. subtilis* serves as a safety valve for osmotically down-shocked cells. The observed induction of *yhdY* and *ykuT* might thus be interpreted as a preemptive measure of high-salinity-challenged *B. subtilis* cells for the next osmotic down-shock. It also should be noted that the *ykuT* gene is part of the $\sigma^B$ regulon.

The initial response of *B. subtilis* to an osmotic upshift is the massive accumulation of K$^+$ ions (26, 65). The K$^+$ content of the cells is reduced again when compatible solutes are accumulated in the second adaptation phase to the high-salinity challenge (65). To search for *B. subtilis* proteins potentially involved in the extrusion of K$^+$, a list of all proteins with the GO annotation "potassium ion transport" (GO:0006813) was extracted from the UniProtKB database. Among the nine proteins retrieved by this search, the K$^+$ export system YhaTU (17) and the putative K$^+$ channel protein YugO were significantly induced at the mRNA level. These observations suggest that both the YhaTU and the YugO systems might function in a concerted fashion in the export of K$^+$ during the prolonged adaptation process of *B. subtilis* to high-salinity surroundings.

After the initial uptake of K$^+$, *B. subtilis* starts to synthesize large amounts of the compatible solute proline, which is accumulated to maintain turgor and general cellular functionality (26, 65). Two routes for proline biosynthesis from its precursor glutamate operate in *B. subtilis* (4): ProB, ProG, and ProI catalyze anabolic proline biosynthesis, and at least two paralogue proteins (ProH and ProJ) are presumed to be responsible for high-level proline synthesis under hyperosmotic conditions (4, 7, 28). The ProA enzyme is used by both proline biosynthetic routes (4, 7). Consistent with the cellular role for these two proline biosynthetic routes, we found that the genes involved in the anabolic proline biosynthesis were repressed, presumably due to the growth arrest, but those used for the osmoadaptive proline production (*proHJ*) were induced (Fig. 4).

The large amounts of proline produced under osmotic stress conditions (65) require a continuous supply of the precursor molecules 2-oxoglutarate and glutamate. Recently, Höper et
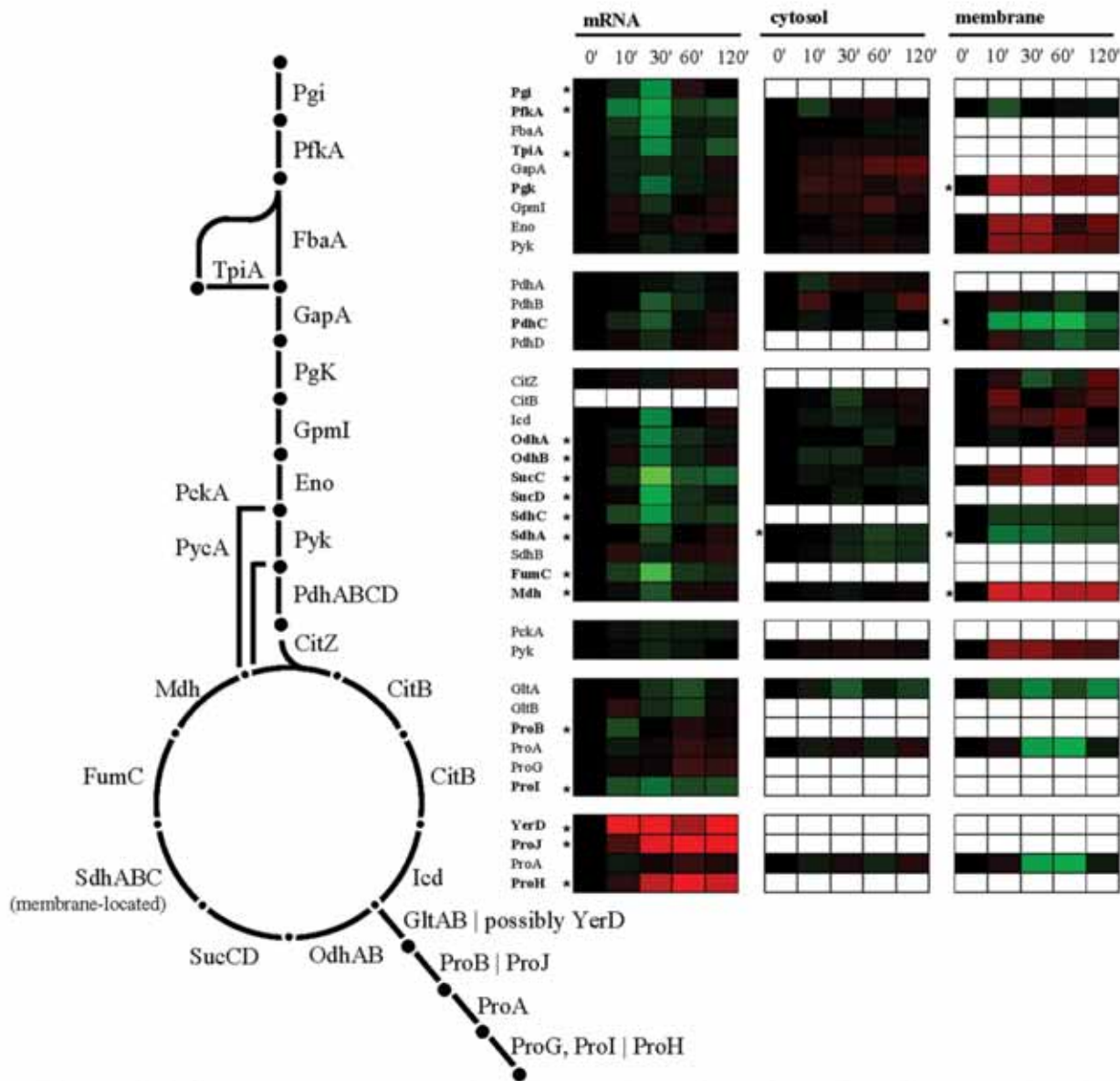
FIG. 4. Glycolysis, TCA cycle, and vegetative as well as salt-specific proline biosynthesis. Expression profiles are displayed on the basis of log$_2$ ratios. The minimum and maximum displayed log$_2$ ratios are ±3 for the transcriptomics data and ±2 for the proteomics data. Genes/proteins exhibiting significant changes are asterisked.

al. (27) observed in a comprehensive proteome study of salt-shocked *B. subtilis* cells that the enzymes of the tricarboxylic acid (TCA) cycle leading to the synthesis of 2-oxoglutarate remained present in the salt-stressed cells, whereas the synthesis and amounts of those enzymes that participate in the TCA cycle reactions from 2-oxoglutarate to oxaloacetate were reduced. In this way, the osmotically challenged *B. subtilis* cell ensures an adequate supply of 2-oxoglutarate that then can be converted via the GOGAT (glutamine-2-oxoglutarate aminotransferase) enzyme into glutamate, the immediate precursor

for the biosynthesis of proline. Our transcriptome and proteome data support the findings of Höper et al. (27).

The adaptation of *B. subtilis* to high salinity is also accompanied by rearrangements in either the composition or structure of the cell envelope. In particular, the lipid and fatty acid composition of the cytoplasmic membrane is affected (42). Lopez et al. found an increase in the content of cardiolipin at the expense of phosphatidylglycerol, an increase in saturated straight-chain fatty acids, and a decrease in the level of iso-branched fatty acids, as well as an increase in unsaturated fatty
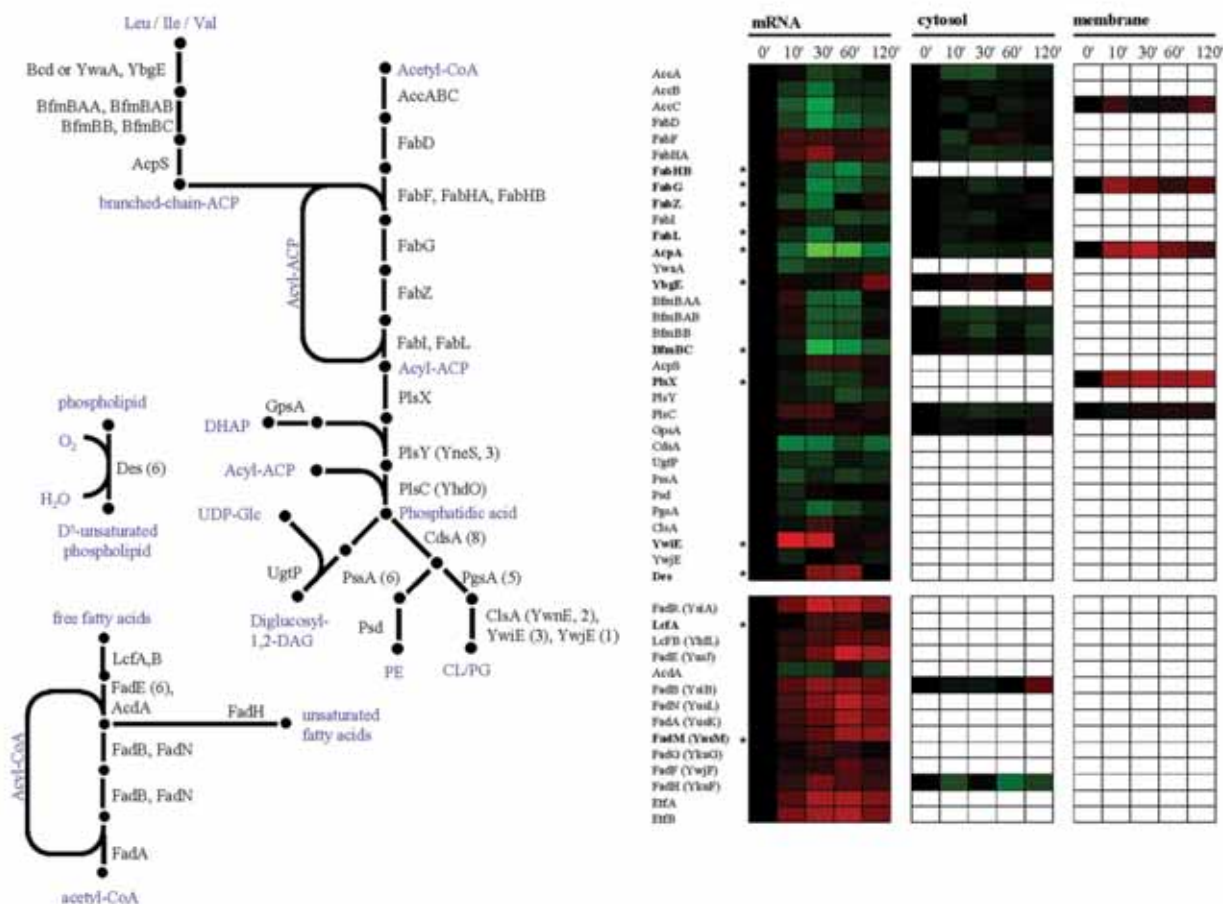
FIG. 5. Fatty acid and lipid biosynthesis and fatty acid degradation. Expression profiles are displayed on the basis of $\log_2$ ratios. The minimum and maximum displayed $\log_2$ ratios are $\pm 3$ for the transcriptomics data and $\pm 2$ for the proteomics data. Genes/proteins exhibiting significant changes are asterisked. PE, phosphatidylethanolamine; CL/PG, cardiolipin/phosphatidylglycerol.

acids. These changes are thought to lead to a higher salt resistance of the *B. subtilis* cell (41). Due to the salinity-induced growth arrest observed in our experiments, most enzymes involved in fatty acid and lipid synthesis are repressed on the mRNA level and are downregulated on the protein level (Fig. 5). In particular, the significant repression of the genes encoding for the fatty acid synthesis proteins AcpA, FabG, FabZ, and FabL strongly indicates a reduced need for membrane building blocks in these almost nongrowing cells.

In contrast to the above-described findings, we observed an induction for a few genes and proteins involved in fatty acid synthesis. The initial condensation step of the fatty acid synthesis in *B. subtilis* is carried out by two β-ketoacyl acyl carrier protein (ACP) synthases, FabHA and FabHB, which can use either acetyl coenzyme A (CoA) or branched acyl-CoA primers (9). A third β-ketoacyl-ACP synthase, FabF, carries out chain elongation. Interestingly, in the course of cellular adaptation of the *B. subtilis* cells to high salinity, we found that the transcription of the *fabHA* and *fabHB* genes was reciprocally regulated. The *fabHB* gene was repressed, and, interestingly, the encoded enzyme has the higher specific activity for

branched-chain primers (9). This result is in good agreement with the decreased need for isobranched fatty acids under salt stress conditions (41). More detailed studies on the apparent different roles for the FabHA and FabHB enzymes during salt stress adaptation might thus be rewarding.

*B. subtilis* has three (putative) cardiolipin synthase genes, *clsA* (formerly known as *ywnE*), *ywiE*, and *ywjE*. In our data, transcription of the *clsA* gene as well as the *ywjE* gene is only marginally affected by salt stress, but the σ$^B$-controlled *ywiE* gene is induced more than 8-fold. Nevertheless, *ywiE* and *ywjE* are not essential during hyperosmotic conditions, whereas *clsA* is required for increased osmotolerance in *B. subtilis* (28, 40). A third distinct biochemical change in the membrane composition is the increase in unsaturated fatty acids, which is catalyzed by a membrane-located Δ$^5$-desaturase encoded by the cold-inducible *des* gene (1). The expression of the *des* gene is regulated via the DesKR two-component regulatory system responsive to membrane lipid fluidity (10). After a 60-min imposition of hyperosmotic stress, we observed a 3-fold induction of the *des* gene. This is not only in good agreement with the observation made by Lopez et al. (40) on membrane fatty
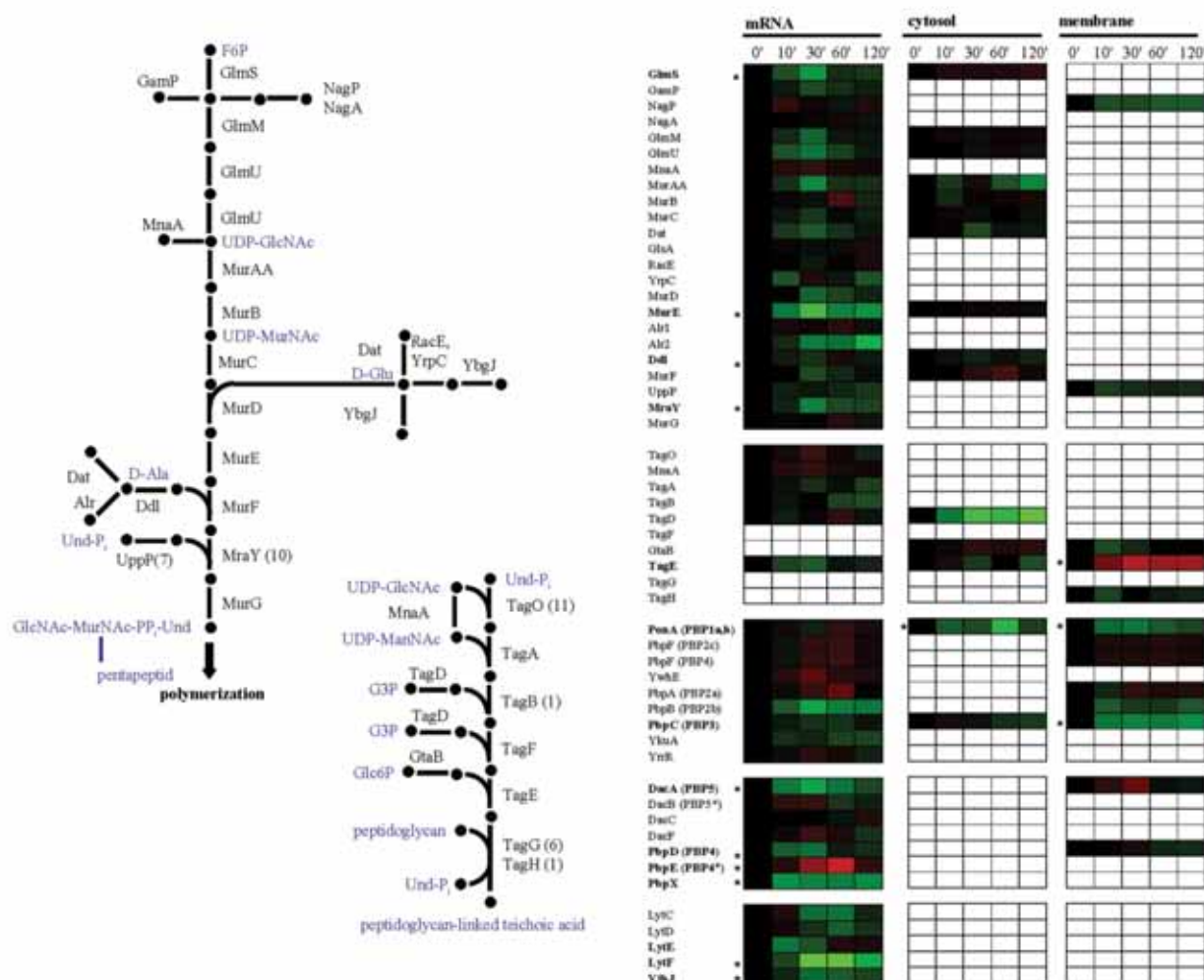
FIG. 6. Biosynthesis of peptidoglycan and teichoic acids, different groups of penicillin-binding proteins, and vegetative autolysins. Expression profiles are displayed on the basis of log$_2$ ratios. The minimum and maximum displayed log$_2$ ratios are $\pm 3$ for the transcriptomics data and $\pm 2$ for the proteomics data. Genes/proteins exhibiting significant changes are asterisked.

acid and lipid composition under hyperosmotic conditions, but it also suggests that the DesKR two-component system senses changes in membrane fluidity after the exposure of *B. subtilis* cells to NaCl in the growth medium.

We found that the anabolic reactions of the fatty acid and lipid metabolism are downregulated, likely a consequence of the growth arrest. However, the expression of genes coding for proteins involved in degradation of free fatty acids *via* β-oxidation show global upregulation (Fig. 5). These genes are mostly controlled by the FadR repressor, which is negatively regulated by long-chain acyl-CoAs. It is worth notice that both end products from straight-chain as well as branched-chain acyl-CoA degradation end up in the TCA cycle either as acetyl-CoA or succinyl-CoA and could in principle complement the anaplerotic reactions catalyzed by the PckA and PycA enzymes.

Along with the composition of the cytoplasmic membrane, exposure of *B. subtilis* to high salinity also influences the properties of the cell wall (42). Given the growth arrest that the *B.*

*subtilis* cells experienced in our salt stress experiment, nearly all enzymes involved in the biosynthesis of the cell wall and its precursors are downregulated on both the mRNA and protein levels (Fig. 6). Few genes and enzymes show the opposite regulation. Among this group of salt-stress-induced genes is penicillin-binding protein Pbp4* (*pbpE*); upregulation of *pbpE* expression has already been reported, and the disruption of *pbpE* results in a salt-sensitive phenotype (48). Expression of *pbpE* is controlled by the alternative sigma factor σ$^W$, and Pbp4* functions as an endopeptidase that is involved in the cleavage of peptidoglycan cross-links. Hence, Pbp4* could play an important role in the cell wall rearrangements that take place during osmoadaptation of *B. subtilis* (48).

**Concluding remarks.** The strategy we employed for the study of severely salt-stressed *B. subtilis* cells examines in parallel and in a time-resolved manner the quantitative changes of the transcriptome, the cytosolic proteome, and the membrane proteome. This genome-wide approach covers not only the

whole mRNA complement of *B. subtilis* cells, but also the two largest subproteomes. A higher proteome coverage of salt-stressed *B. subtilis* can be achievable in future studies by employing tools that specifically target the most hydrophobic complement of the proteome (66), the secretome (3), and the cell wall proteome of *B. subtilis* (Hempel et al., unpublished data).

The data presented here provide a comprehensive picture of the dynamic changes in gene expression and the concomitant alterations in the level of proteins that occur in *B. subtilis* cells exposed to a severe salt shock. Our study complements previous genetic and physiological studies on the osmotic stress response of *B. subtilis*. In combination with other genome-wide transcriptional profiling (57) and proteomic studies (27) of salt-stressed *B. subtilis* cells, our data provide an unprecedented rich basis for further in-depth investigation of the physiological and genetic responses of *B. subtilis* to hyperosmotic stress.

## REFERENCES

1. **Aguilar, P. S., J. E. Cronan, Jr., and D. de Mendoza.** 1998. A *Bacillus subtilis* gene induced by cold shock encodes a membrane phospholipid desaturase. J. Bacteriol. **180**:2194–2200.
2. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81**:741–746.
3. **Antelmann, H., H. Tjalsma, B. Voigt, S. Ohlmeier, S. Bron, J. M. van Dijl, and M. Hecker.** 2001. A proteomic view on genome-based signal peptide predictions. Genome Res. **11**:1484–1502.
4. **Belitsky, B. R., J. Brill, E. Bremer, and A. L. Sonenshein.** 2001. Multiple genes for the last step of proline biosynthesis in *Bacillus subtilis*. J. Bacteriol. **183**:4389–4392.
5. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340**:783–795.
6. **Bidle, K. A., P. A. Kirkland, J. L. Nannen, and J. A. Maupin-Furlow.** 2008. Proteomic analysis of *Haloferax volcanii* reveals salinity-mediated regulation of the stress response protein PspA. Microbiology **154**:1436–1443.
7. **Bremer, E.** 2002. Adaptation to changing osmolarity, p. 385–391. *In* A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and its closest relatives: from genes to cells. ASM Press, Washington, DC.
8. **Cao, M., and J. D. Helmann.** 2004. The *Bacillus subtilis* extracytoplasmic-function sigmaX factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides. J. Bacteriol. **186**:1136–1146.
9. **Choi, K. H., R. J. Heath, and C. O. Rock.** 2000. beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. J. Bacteriol. **182**:365–370.
10. **Cybulski, L. E., D. Albanesi, M. C. Mansilla, S. Altabe, P. S. Aguilar, and D. de Mendoza.** 2002. Mechanism of membrane fluidity optimization: isothermal control of the *Bacillus subtilis* acyl-lipid desaturase. Mol. Microbiol. **45**:1379–1388.
11. **Dennis, G., Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, and R. A. Lempicki.** 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. **4**:P3.
12. **Eiamphungporn, W., and J. D. Helmann.** 2008. The *Bacillus subtilis* sigmaM regulon and its contribution to cell envelope stress responses. Mol. Microbiol. **67**:830–848.
13. **Eng, J. K., A. L. McCormack, and J. R. Yates.** 1994. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. J. Am. Soc. Mass Spectrom. **5**:976.
14. **Errington, J.** 2003. Regulation of endospore formation in *Bacillus subtilis*. Nat. Rev. Microbiol. **1**:117–126.
15. **Eymann, C., A. Dreisbach, D. Albrecht, J. Bernhardt, D. Becher, S. Gentner, T. Tam le, K. Büttner, G. Buurman, C. Scharf, S. Venz, U. Völker, and M. Hecker.** 2004. A comprehensive proteome map of growing *Bacillus subtilis* cells. Proteomics **4**:2849–2876.
16. **Eymann, C., G. Homuth, C. Scharf, and M. Hecker.** 2002. *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. J. Bacteriol. **184**:2500–2520.
17. **Fujisawa, M., M. Ito, and T. A. Krulwich.** 2007. Three two-component transporters with channel-like properties have monovalent cation/proton antiport activity. Proc. Natl. Acad. Sci. U. S. A. **104**:13289–13294.
18. **Fujisawa, M., A. Kusumoto, Y. Wada, T. Tsuchiya, and M. Ito.** 2005. NhaK, a novel monovalent cation/H⁺ antiporter of *Bacillus subtilis*. Arch. Microbiol. **183**:411–420.
19. **Hahne, H., S. Wolff, M. Hecker, and D. Becher.** 2008. From complementarity to comprehensiveness—targeting the membrane proteome of growing *Bacillus subtilis* by divergent approaches. Proteomics **8**:4123–4136.
20. **Hambraeus, G., C. von Wachenfeldt, and L. Hederstedt.** 2003. Genome-wide survey of mRNA half-lives in *Bacillus subtilis* identifies extremely stable mRNAs. Mol. Genet. Genomics **269**:706–714.
21. **Hecker, M., and U. Völker.** 2001. General stress response of *Bacillus subtilis* and other bacteria. Adv. Microb. Physiol. **44**:35–91.
22. **Herranen, M., N. Battchikova, P. Zhang, A. Graf, S. Sirpio, V. Paakkarinen, and E. M. Aro.** 2004. Towards functional proteomics of membrane protein complexes in *Synechocystis sp.* PCC 6803. Plant Physiol. **134**:470–481.
23. **Hirokawa, T., S. Boon-Chieng, and S. Mitaku.** 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics **14**:378–379.
24. **Hoffmann, T., C. Boiangiu, S. Moses, and E. Bremer.** 2008. Responses of *Bacillus subtilis* to hypotonic challenges: physiological contributions of mechanosensitive channels to cellular survival. Appl. Environ. Microbiol. **74**:2454–2460.
25. **Hoffmann, T., A. Schutz, M. Brosius, A. Völker, U. Völker, and E. Bremer.** 2002. High-salinity-induced iron limitation in *Bacillus subtilis*. J. Bacteriol. **184**:718–727.
26. **Holtmann, G., E. P. Bakker, N. Uozumi, and E. Bremer.** 2003. KtrAB and KtrCD: two K⁺ uptake systems in *Bacillus subtilis* and their role in adaptation to hypertonicity. J. Bacteriol. **185**:1289–1298.
27. **Höper, D., J. Bernhardt, and M. Hecker.** 2006. Salt stress adaptation of *Bacillus subtilis*: a physiological proteomics approach. Proteomics **6**:1550–1562.
28. **Höper, D., U. Völker, and M. Hecker.** 2005. Comprehensive characterization of the contribution of individual SigB-dependent general stress genes to stress resistance of *Bacillus subtilis*. J. Bacteriol. **187**:2810–2826.
29. **Horsburgh, M. J., and A. Moir.** 1999. Sigma M, an ECF RNA polymerase sigma factor of *Bacillus subtilis* 168, is essential for growth and survival in high concentrations of salt. Mol. Microbiol. **32**:41–50.
30. **Huang, D. W., B. T. Sherman, Q. Tan, J. Kir, D. Liu, D. Bryant, Y. Guo, R. Stephens, M. W. Baseler, H. C. Lane, and R. A. Lempicki.** 2007. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. **35**:W169–W175.
31. **Huang, F., S. Fulda, M. Hagemann, and B. Norling.** 2006. Proteomic screening of salt-stress-induced changes in plasma membranes of *Synechocystis sp.* strain PCC 6803. Proteomics **6**:910–920.
32. **Ito, M., A. A. Guffanti, B. Oudega, and T. A. Krulwich.** 1999. mrp, a multigene, multifunctional locus in *Bacillus subtilis* with roles in resistance to cholate and to Na⁺ and in pH homeostasis. J. Bacteriol. **181**:2394–2402.
33. **Jervis, A. J., P. D. Thackray, C. W. Houston, M. J. Horsburgh, and A. Moir.** 2007. SigM-responsive genes of *Bacillus subtilis* and their promoters. J. Bacteriol. **189**:4534–4538.
34. **Juncker, A. S., H. Willenbrock, G. Von Heijne, S. Brunak, H. Nielsen, and A. Krogh.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci. **12**:1652–1662.
35. **Jürgen, B., S. Tobisch, M. Wumpelmann, D. Gordes, A. Koch, K. Thurow, D. Albrecht, M. Hecker, and T. Schweder.** 2005. Global expression profiling of *Bacillus subtilis* cells during industrial-close fed-batch fermentations with different nitrogen sources. Biotechnol. Bioeng. **92**:277–298.
36. **Kajiyama, Y., M. Otagiri, J. Sekiguchi, S. Kosono, and T. Kudo.** 2007. Complex formation by the mrpABCDEFG gene products, which constitute a principal Na+/H+ antiporter in *Bacillus subtilis*. J. Bacteriol. **189**:7511–7514.
37. **Kall, L., A. Krogh, and E. L. Sonnhammer.** 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res. **35**:W429–W432.
38. **Kall, L., A. Krogh, and E. L. Sonnhammer.** 2004. A combined transmembrane topology and signal peptide prediction method. J. Mol. Biol. **338**:1027–1036.
39. **Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305**:567–580.
40. **Lopez, C. S., A. F. Alice, H. Heras, E. A. Rivas, and C. Sanchez-Rivas.** 2006. Role of anionic phospholipids in the adaptation of *Bacillus subtilis* to high salinity. Microbiology **152**:605–616.
41. **Lopez, C. S., H. Heras, H. Garda, S. Ruzal, C. Sanchez-Rivas, and E. Rivas.** 2000. Biochemical and biophysical studies of *Bacillus subtilis* envelopes under hyperosmotic stress. Int. J. Food Microbiol. **55**:137–142.
42. **Lopez, C. S., H. Heras, S. M. Ruzal, C. Sanchez-Rivas, and E. A. Rivas.** 1998. Variations of the envelope composition of *Bacillus subtilis* during growth in hyperosmotic medium. Curr. Microbiol. **36**:55–61.
43. **Marles-Wright, J., T. Grant, O. Delumeau, G. van Duinen, S. J. Firbank,**

882    HAHNE ET AL.    J. BACTERIOL.

P. J. Lewis, J. W. Murray, J. A. Newman, M. B. Quin, P. R. Race, A. Rohou, W. Tichelaar, M. van Heel, and R. J. Lewis. 2008. Molecular architecture of the "stressosome," a signal integration and transduction hub. Science **322:** 92–96.

44. **Mascher, T., A. B. Hachmann, and J. D. Helmann.** 2007. Regulatory overlap and functional redundancy among *Bacillus subtilis* extracytoplasmic function sigma factors. J. Bacteriol. **189:**6919–6927.

45. **Mitaku, S., and T. Hirokawa.** 1999. Physicochemical factors for discriminating between soluble and membrane proteins: hydrophobicity of helical segments and protein length. Protein Eng. **12:**953–957.

46. **Mitaku, S., T. Hirokawa, and T. Tsuji.** 2002. Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. Bioinformatics **18:**608–616.

47. **Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne.** 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10:**1–6.

48. **Palomino, M. M., C. Sanchez-Rivas, and S. M. Ruzal.** 2009. High salt stress in *Bacillus subtilis*: involvement of PBP4* as a peptidoglycan hydrolase. Res. Microbiol. **160:**117–124.

49. **Park, S. K., J. D. Venable, T. Xu, and J. R. Yates III.** 2008. A quantitative analysis software tool for mass spectrometry-based proteomics. Nat. Methods **5:**319–322.

50. **Peng, J., J. E. Elias, C. C. Thoreen, L. J. Licklider, and S. P. Gygi.** 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J. Proteome Res. **2:**43–50.

51. **Petersohn, A., M. Brigulla, S. Haas, J. D. Hoheisel, U. Völker, and M. Hecker.** 2001. Global analysis of the general stress response of *Bacillus subtilis*. J. Bacteriol. **183:**5617–5631.

52. **Price, C. W., P. Fawcett, H. Ceremonie, N. Su, C. K. Murphy, and P. Youngman.** 2001. Genome-wide analysis of the general stress response in *Bacillus subtilis*. Mol. Microbiol. **41:**757–774.

53. **Saeed, A. I., V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, J. Braisted, M. Klapa, T. Currier, M. Thiagarajan, A. Sturn, M. Snuffin, A. Rezantsev, D. Popov, A. Ryltsov, E. Kostukovich, I. Borisovsky, Z. Liu, A. Vinsavich, V. Trush, and J. Quackenbush.** 2003. TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:**374–378.

54. **Schumann, W.** 1999. FtsH—a single-chain charonin? FEMS Microbiol. Rev. **23:**1–11.

55. **Sierro, N., Y. Makita, M. de Hoon, and K. Nakai.** 2008. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. Nucleic Acids Res. **36:**D93–D96.

56. **Sonnhammer, E. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. **6:**175–182.

57. **Steil, L., T. Hoffmann, I. Budde, U. Völker, and E. Bremer.** 2003. Genome-wide transcriptional profiling analysis of adaptation of *Bacillus subtilis* to high salinity. J. Bacteriol. **185:**6358–6370.

58. **Stülke, J., R. Hanschke, and M. Hecker.** 1993. Temporal activation of beta-glucanase synthesis in *Bacillus subtilis* is mediated by the GTP pool. J. Gen. Microbiol. **139:**2041–2045.

59. **Tabb, D. L., W. H. McDonald, and J. R. Yates III.** 2002. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. **1:**21–26.

60. **Tam, L. T., H. Antelmann, C. Eymann, D. Albrecht, J. Bernhardt, and M. Hecker.** 2006. Proteome signatures for stress and starvation in *Bacillus subtilis* as revealed by a 2-D gel image color coding approach. Proteomics **6:**4565–4585.

61. **Tusnady, G. E., and I. Simon.** 1998. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J. Mol. Biol. **283:**489–506.

62. **Tusnady, G. E., and I. Simon.** 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics **17:**849–850.

63. **UniProt Consortium.** 2007. The universal protein resource (UniProt). Nucleic Acids Res. **35:**D193–D197.

64. **Wahome, P. G., and P. Setlow.** 2008. Growth, osmotic downshock resistance and differentiation of *Bacillus subtilis* strains lacking mechanosensitive channels. Arch. Microbiol. **189:**49–58.

65. **Whatmore, A. M., J. A. Chudek, and R. H. Reed.** 1990. The effects of osmotic upshock on the intracellular solute pools of *Bacillus subtilis*. J. Gen. Microbiol. **136:**2527–2535.

66. **Wolff, S., H. Hahne, M. Hecker, and D. Becher.** 2008. Complementary analysis of the vegetative membrane proteome of the human pathogen *Staphylococcus aureus*. Mol. Cell. Proteomics **7:**1460–1468.

67. **Xu, C., H. Ren, S. Wang, and X. Peng.** 2004. Proteomic analysis of salt-sensitive outer membrane proteins of *Vibrio parahaemolyticus*. Res. Microbiol. **155:**835–842.

68. **Xu, C., S. Wang, H. Ren, X. Lin, L. Wu, and X. Peng.** 2005. Proteomic analysis on the expression of outer membrane proteins of *Vibrio alginolyticus* at different sodium concentrations. Proteomics **5:**3142–3152.

_____

*A Proteomic View of an Important Human Pathogen-Towards the Quantification of the Entire Staphylococcus aureus Proteome*

PLoS one

# A Proteomic View of an Important Human Pathogen – Towards the Quantification of the Entire *Staphylococcus aureus* Proteome

**Dörte Becher[1], Kristina Hempel[1], Susanne Sievers[1], Daniela Zühlke[1], Jan Pané-Farré[1], Andreas Otto[1], Stephan Fuchs[1], Dirk Albrecht[1], Jörg Bernhardt[1], Susanne Engelmann[1], Uwe Völker[2], Jan Maarten van Dijl[3], Michael Hecker[1]***

1 Institute for Microbiology, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany, 2 Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany, 3 Department of Medical Microbiology, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands

## Abstract

The genome sequence is the "blue-print of life," but proteomics provides the link to the actual physiology of living cells. Because of their low complexity bacteria are excellent model systems to identify the entire protein assembly of a living organism. Here we show that the majority of proteins expressed in growing and non-growing cells of the human pathogen *Staphylococcus aureus* can be identified and even quantified by a metabolic labeling proteomic approach. *S. aureus* has been selected as model for this proteomic study, because it poses a major risk to our health care system by combining high pathogenicity with an increasing frequency of multiple antibiotic resistance, thus requiring the development of new anti-staphylococcal therapy strategies. Since such strategies will likely have to target extracellular and surface-exposed virulence factors as well as staphylococcal survival and adaptation capabilities, we decided to combine four subproteomic fractions: cytosolic proteins, membrane-bound proteins, cell surface-associated and extracellular proteins, to comprehensively cover the entire proteome of *S. aureus*. This quantitative proteomics approach integrating data ranging from gene expression to subcellular localization in growing and non-growing cells is a proof of principle for whole-cell physiological proteomics that can now be extended to address physiological questions in infection-relevant settings. Importantly, with more than 1700 identified proteins (and 1450 quantified proteins) corresponding to a coverage of about three-quarters of the expressed proteins, our model study represents the most comprehensive quantification of a bacterial proteome reported to date. It thus paves the way towards a new level in understanding of cell physiology and pathophysiology of *S. aureus* and related pathogenic bacteria, opening new avenues for infection-related research on this crucial pathogen.

## Introduction

*Staphylococcus aureus* is a Gram-positive human pathogen of increasing significance, mainly due to its high incidence and the increasing spread of antibiotic resistances. Multiple virulence factors allow *S. aureus* to cause a broad spectrum of infectious diseases, ranging from superficial abscesses of the skin to severe diseases such as endocarditis, osteomyelitis, toxic shock syndrome or sepsis. Furthermore, *S. aureus* is particularly important in healthcare settings, where it is causing up to 40% of nosocomial infections. Vancomycin and related antibiotics form the drugs of last resort against resistant strains. Therefore, the recent emergence of vancomycin/methicillin-resistant *S. aureus* strains represents a major threat for the health care system, requiring the development of new therapeutic options against *S. aureus* infections [1]. Since new therapeutic strategies could involve neutralization of virulence factors or direct elimination of *S. aureus* a better understanding of both cell physiology and pathophysiology is urgently required to successfully combat *S. aureus* infections. This crucial goal can be reached by a functional genomics approach which opens a new perspective in *S. aureus* research because cell physiology and pathogenicity can be analyzed at a genome-wide scale.

The release of the first complete bacterial genome sequence in 1995 opened the era of functional genomics. In the same year, the term proteome was coined to describe the complete set of proteins synthesized under specific physiological circumstances [2]. Whereas the genome sequence only provides the "blue-print of life", proteomics is required to bring this "blue print of life" to cell function, because proteins can be regarded as the main players of life. On the other hand the genome sequence is necessary to identify all cellular proteins expressed at genome-wide scale by mass spectrometry. Clearly, the route from the genome sequence via transcriptomics, proteomics and metabolomics towards a systems biology perspective leads to a new quality in understanding life in general. Because of their low complexity - only a few

A Proteomic View of an Important Human Pathogen- Towards the Quantification of the Entire *Staphylococcus aureus* Proteome

Quantitation *S.aureus* Proteome

hundred different proteins make a cell viable - bacteria are extremely attractive model systems to bring the genome sequence to real life.

First proteomics studies date back to 1975 when O'Farrell published the two-dimensional gel electrophoresis technique that allows the separation of thousands of protein spots in an area of 20×20 cm in size [3]. After the initial enthusiasm for gel-based proteomics up to the early nineties, it became evident that only parts of a bacterial proteome can be visualized by this method. Many proteins escape detection by gel-based proteomics, hydrophobic membrane proteins and low abundance proteins being the most prominent classes. The implementation of mass spectrometry of peptides with mild ionization methods and online LC-MS/MS provides a suitable alternative to gel-based proteomics that has seen spectacular advances over the past 10 years, providing a completely new quality in proteomic research [4–6].

Comprehensive proteomic coverage has recently been reported for the non-pathogenic eukaryotic model organism *Saccharomyces cerevisiae* [7]. Mann and co-workers were the first to show that a combination of state of the art proteomics consisting of highly sensitive mass spectrometry, *in vivo* labeling with stable isotopes and appropriate bioinformatic analysis, allows the identification and even comparative quantification of almost the entire proteome of exponentially growing haploid and diploid *S. cerevisiae* cells.

Here we show that almost the entire proteome of *S. aureus* COL, whose genome sequence became available in 2005 [8], can be identified and even quantified by a combination of 2-D gel-based and LC-MS/MS-based proteomic techniques. Since therapeutic strategies directed against *S. aureus* could either target extracellular and secreted proteins or the pathogens metabolism, we integrated the analysis of the different subproteomes to provide evidence that an integrated proteomic view of cell physiology and virulence will help to come to a new quality in the understanding of general cellular processes of *S. aureus*. This approach would thus generate the prerequisites for a deeper understanding of the behaviour of this pathogen in its human host. The proteomic view of metabolism and virulence of *S. aureus* presented here constitutes to our knowledge the most comprehensive quantitative study of a bacterial proteome to date and opens a new dimension in infection-related research of this crucial pathogen.

## Results and Discussion

In order to visualize the entire proteome of *S. aureus*, we started with the hypothesis that roughly 2000 proteins are being synthesized from the genome that contains about 2600 genes [8]. This assumption is supported by our transcriptional profiling studies. Using DNA microarray techniques from about 2400 genes spotted on the DNA array about 1975 (82%) genes were found to be expressed above background level in growing and 1111 (46%) genes in stationary-phase cells (see Text S1, Table S1). Thus, approximately two thousand genes are expressed in total in both conditions corresponding to an expression level of 83% (see Table S1). In order to visualize the entire proteome we intentionally aimed at a combination of different subproteomic fractions (Figure 1) instead of a comprehensive analysis of one single fraction. Most of the virulence factors that are in the focus of infection-related proteomic studies of *S. aureus* are either cell surface-exposed or even secreted into the extracellular space. Combined analysis of different subproteomes would allow simultaneous coverage of infection-relevant immunogenic, surface-exposed and extracellular proteins besides the intracellular
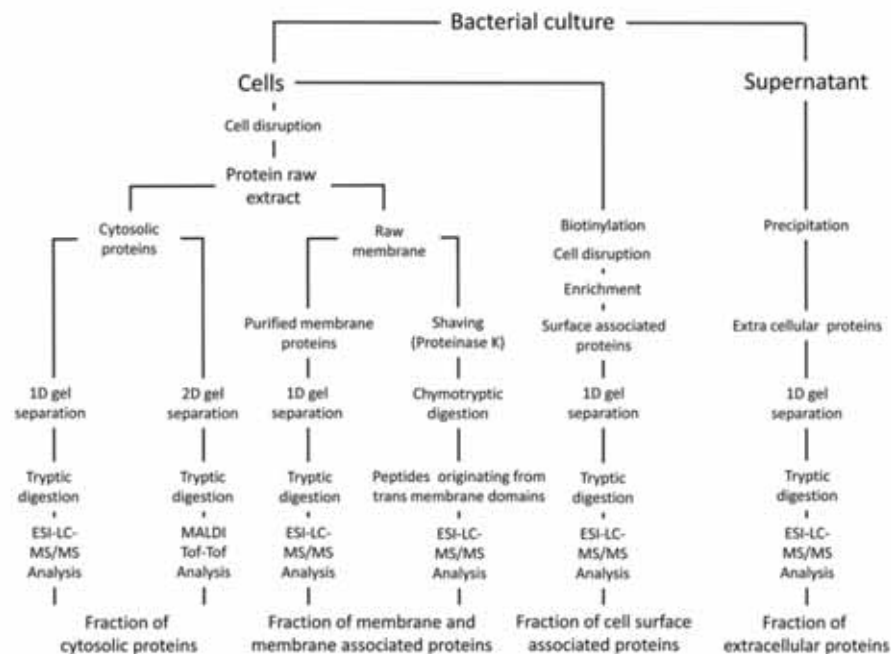


**Figure 1. Analysis strategy for investigation of all sub-proteomic fractions of *S. aureus*.** The mixture of the differentially labeled cells has been carried out on different levels. For analysis of the cytosolic proteins and membrane proteins by LC-MS/MS the samples were mixed on protein level using aliquots with the same protein amount. For 2-D gel analysis identical amounts of protein were separated on individual gel. For the fraction of cell surface-associated proteins the cells were mixed prior to the biotinylation step in equal amounts. Extracellular proteins were mixed after the precipitation step.
doi:10.1371/journal.pone.0008176.g001

A Proteomic View of an Important Human Pathogen- Towards the Quantification of the
Entire *Staphylococcus aureus* Proteome
_____

Quantitation *S.aureus* Proteome

proteins that are required for basic metabolism and other cellular processes. Furthermore, such a pre-fractionation would also reveal detailed information about the subproteomic distribution of proteins. Therefore, four particular proteomic fractions, namely the cytosolic, membrane-bound, cell surface-associated and extracellular subproteomes were analyzed in order to define and assemble the entire proteome.

Even if the visualization of the protein inventory is the main topic of this project physiological studies require protein expression profiling and quantification to follow the up- and down-regulation of single proteins in response to physiological stimuli. In this study exponentially growing cells were compared to non-growing cells of the stationary phase as a proof of principle, because a shift from growing to non-growing conditions is probably a typical situation in most of the natural microhabitats in the human host.

## Cytosolic Proteins

To initiate the analysis of cytosolic proteins we first employed gel-based proteomics. Gel-based proteomics techniques have been extensively used to characterize the regulation of gene expression and metabolism of *S. aureus* in response to infection-related stimuli such as oxygen starvation [9], different kinds of stress such as oxidative stress [10], and were also used to assign the involved proteins to regulons [11], the building blocks of the genetic adaptation network. Two-dimensional gel electrophoresis (2-D PAGE) is still a powerful tool to address physiological issues, in particular because most of the metabolic pathways and the most obvious stress/starvation responses are visualized by gel-based proteomics in a convenient, eye-catching and intuitive way ("metabolism at a glance" [12,13]).

For this study we analyzed the proteome pattern in growing and non-growing cells cultivated in Bioexpress medium in the analytical windows of pI 4–7 and for the first time also of pI 6–11. In total 683 proteins were identified (Figure 2, Figure S1) covering most of the metabolic pathways (Figure S2). Different patterns of cytosolic proteins that respond to the transition from growth to stationary phase can be visualized and easily be quantified by image analysis.

The data indicate a dramatic reprogramming of protein synthesis in the non-growing state with many vegetative proteins repressed (e.g. by the stringent response) and others strongly induced (red labeled in Figure 2A) forming a complex adaptational protein network in order to face starvation and stress. A comprehensive stress/starvation proteomics signature library [14] is a useful toolbox for prediction of the physiological state of cells [15,16], which will be important for physiological studies as for defining the main nutrients (metabolic proteome signatures) and the predominant stress/starvation stimuli of *S. aureus* cells under infection-relevant conditions [for related studies see 17, 18].

Because 2-D gels cover only about 25% of the proteome, 1-D gel-based protein prefractionation in combination with LC-MS/MS (GeLC-MS/MS) was employed to detect the missing proteins, and to identify and quantify the majority of expressed proteins. Using this technique, more than 1100 proteins were identified in the cytosolic fraction of growing and non-growing cells, most of them having a predicted pI in the range of 3.5 to 7.5 (Figure 2B).

Besides the proteins already covered by 2-D PAGE, about 500 new proteins were identified by the GeLC-MS/MS approach, mostly low abundant proteins, but also very acidic or alkaline, as well as very large or small proteins (Figure 3). In total, a proteomic coverage of about 67% was achieved from 1796 predicted cytosolic proteins. Notably, the proteomic coverage of highly expressed genes is up to 90% while for low-expressed genes the coverage is lower (Figure 4).

Importantly, the dynamics of protein synthesis in growing and non-growing cells can be followed and compared by integrating 2-D gel-based and GeLC-MS/MS-based approaches. In Figure 2B, the quantitative GeLC-MS/MS data are combined with the experimental 2-D gel data (Figure 2A) in a virtual 2-D gel, using the same color code for up- and down-regulated proteins. Importantly, the two approaches revealed similar levels of up- or down-regulation of proteins in non-growing cells (Figure 3, Table S2).

The color code used in Figure 2 visualizes three major classes of proteins:

I.  yellow labeled proteins probably no longer synthesized in non-growing cells but still present and more or less stable

II. green labeled proteins no longer synthesized and even degraded in non-growing cells and

III. red labeled proteins enriched in non-growing cells due to increased synthesis

Transcription of class I and class II proteins have been analyzed by Northern blotting for a few typical candidate genes (Table S3, Figure S3). In non-growing cells the transcription of the genes encoding "yellow and green proteins" was repressed at similar levels at the onset of stationary phase, leading to a shut off of the new synthesis of the proteins. These proteins, however, already synthesized in growing cells and still present in non-growing cells, have two different fates: Either they are stable and kept at the same level as in growing cells or they are present in lower levels than in growing cells strongly indicating degradation. The reason for this different behavior- either stable or degraded- might be a reflection of the activity of the proteins. Those proteins that are still active might be stable (yellow proteins) whereas the proteins no longer active and thus not required might be degraded (green proteins), thus at least partially relieving nutrient starvation. We suggest that such unemployed proteins synthesized and required in growing cells are no longer active and integrated in functional complexes in non-growing cells, which makes them vulnerable to proteolytic attack, for example by Clp proteases [unpublished data, see also 19]. Ribosomal proteins, translational factors and some key enzymes for amino acid biosynthesis are the most obvious members of this group of "green proteins" (Figure 2). The molecular mechanisms that make the unemployed proteins accessible to the Clp machine have to be studied in future experiments.

This recycling of inactive proteins is probably a major nutrient resource for non-growing cells starved for carbon and energy sources. Particularly the ribosomes synthesized in high amounts in growing and partly inactive in non-growing cells form a major endogenous nutrient source. This degradation of ribosomes is observed at the RNA and protein level (Figure 2, Figure S4). The recycling of ribosomes is probably a crucial component of the physiology of starving cell [20].

The proteins particularly synthesized in the stationary phase (class III, red labeled proteins in Figure 2) will likely have adaptive functions in the non-growing state, among them specific starvation proteins that may help to make new substrates available (e.g. carbon overflow products) and even specific and general stress proteins to make the non-growing cells more stress resistant than growing cells [21]. Among the "red proteins" we found e.g. the PEP carboxykinase (PckA), a gluconeogenetic enzyme indicating the absence of glucose, some TCA cycle enzymes indicating carbon source limitation in non-growing cells and proteins

A Proteomic View of an Important Human Pathogen- Towards the Quantification of the Entire *Staphylococcus aureus* Proteome
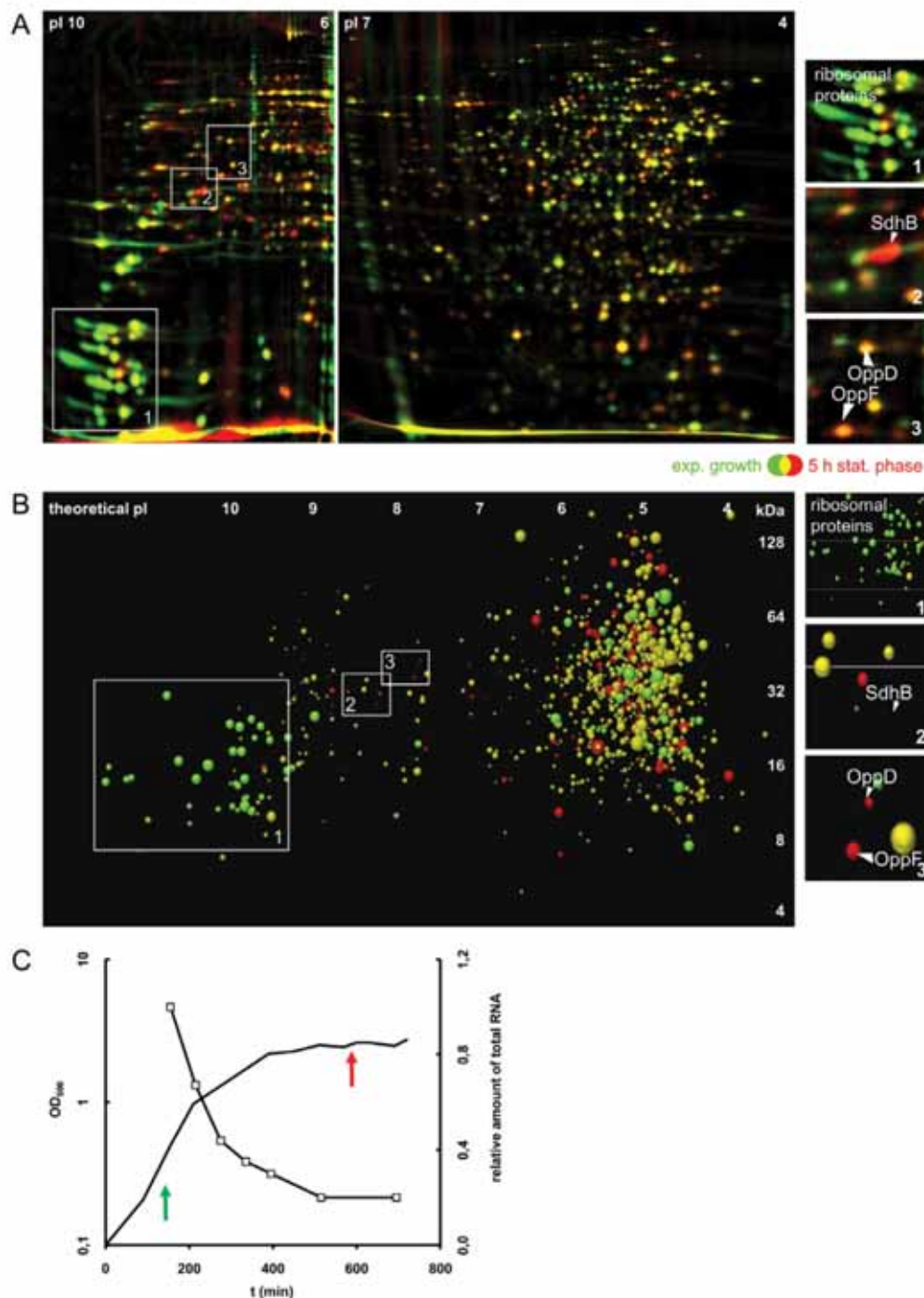
_____

Quantitation *S.aureus* Proteome



**Figure 2. Changes of protein expression pattern.** Framed sections are presented separately on right side of the figures. **A 2-D gel-based expression analysis by Dual Channel Imaging:** The overlay of the false color image of growing cells (green) and non-growing cells (red) is shown. This overlay results in the following color code: I. proteins probably no longer synthesized in non-growing cells but still present and more or less stable (yellow color), II. proteins no longer synthesized and even degraded in non-growing cells (green color) and III. proteins enriched in non-growing cells (red labelled). **B GeLC-MS/MS-based expression analysis:** Spot position derived from theoretical pI and theoretical molecular weight, spot size derived from spectral counts determined by Census software, spot color derived from log2 ratio of stat/exp, (green<−0.8, red>+0.8, yellow≥−0.8, ≤+0.8), Figure has been created in Microsoft Excel. **C Growth curve and RNA content:** Sampling points (arrows) are indicated in the growth curve (line graph). The relative decrease of total RNA content with decreasing growth rate is shown in the same graph (line with squares).
doi:10.1371/journal.pone.0008176.g002
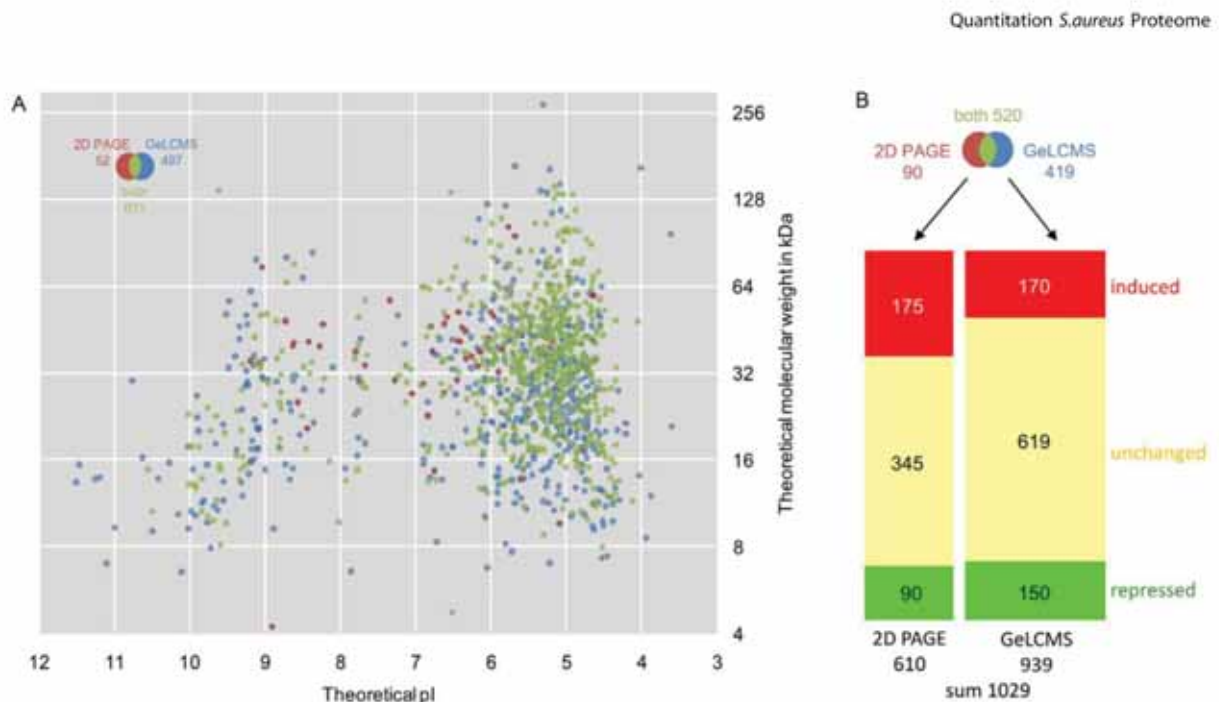
**Figure 3. Comparative analysis of the cytosolic protein fraction. A Qualitative analysis:** In summary 1180 proteins were identified, including 1095 proteins predicted as cytosolic proteins (but also proteins predicted to be cell wall-associated - 13, membrane proteins – 29, extracellular proteins - 23 and lipoproteins -20). Thus, 61% of the predicted cytosolic proteins were identified. **B Quantitative analysis:** 1029 proteins have been quantified in summary in both approaches with an overlap of 520 proteins. With the 2-D gel-based approach 610 proteins could be quantified and 265 proteins were found altered in amount (175 proteins - labeled in red - in higher amount and 90 proteins - labeled in green - in lower amount in the stationary phase!). By GeLC-MS 939 proteins could be quantified and 320 proteins were changed in quantity (170 proteins - labeled in red - in higher and 150 proteins - labeled in green - in lower amount in the stationary phase). Therefore, about one third of the cytosolic proteome has to be considered as regulated.
doi:10.1371/journal.pone.0008176.g003

belonging to the phosphate regulon (e.g. PhoP, SACOL1424) which suggests phosphate starvation. These proteomic signatures are useful tools for a general evaluation of cell physiology and metabolism (detailed information see Table S2, Figure S4).

A Voronoi treemap-based layout of KEGG's *S. aureus* COL gene orthology was used to intuitively visualize the changes in the protein expression pattern in non-growing cells compared to growing cells. The up- (red color) and down-regulation (green color) of most of the proteins assigned to functional categories can be visualized which is illustrated for the already mentioned down-regulation of the translational machines such as ribosomal proteins or aminoacyl-tRNA-synthetases in non-growing cells. Additional pronounced changes in different branches of metabolism are highlighted in Figure 5.

### Membrane Proteins

The membrane proteome was analyzed by a combination of a GeLC-MS/MS [21] and a shaving approach [22]. Whereas the first approach also identified a large fraction of cytosolic proteins - either membrane-associated or cytosolic contaminations - membrane shaving only revealed proteins with 1 to 21 trans-membrane domains. It is interesting to note that only by the shaving approach the most hydrophobic species of membrane-spanning peptides, usually lacking lysine- and arginine residues as targets for tryptic digestion, were identified. Employing a combination of both approaches, 56% of all predicted membrane proteins could be detected (Figure 6). The coverage is even higher if we consider that not all membrane proteins were expressed under the physiological conditions tested in this study. In addition to integral membrane

proteins, the membrane-targeted analysis resulted in the identification of 125 proteins with a cytoplasmic location, but temporary or permanent membrane association. It should be noted that this number represents only an estimate of peripheral membrane proteins. Nevertheless, this group of membrane-associated proteins is important, due to their essential roles in processes such as membrane and cell wall metabolism, signal transduction and respiration.

In this study the membrane proteome of *S. aureus* was quantified for the first time. Several integral- and membrane-associated proteins, such as the fructose-specific enzyme II (EII), four further EII components of unknown substrate specificity of the phosphotransferase system (PTS) and the glycerol uptake facilitator GlpF were up-regulated in non-growing cells due to carbon source limitation (Figure 7). Besides up-regulation of GlpF, two further transporters importing $C_3$ carbon sources, namely, phosphoglycerate and glycerol-3-phosphate and four cytosolic enzymes acting downstream of $C_3$ catabolite uptake were detected in increased quantity in the membrane fraction. This finding indicates that the glycerol consuming pathway forms enzyme complexes associated with the membrane bound uptake facilitator. Glycerol consumption was tightly connected to virulence in two close relatives of *S. aureus*, the intracellular pathogens *Listeria monocytogenes* [23,24] and *Mycoplasma pneumonia* [25]. Such observations justify further research to obtain a better understanding of glycerol utilization in *S. aureus* and its potential involvement in pathogenicity. Furthermore, 11 transporters, mostly of the ATP binding cassette (ABC) type, involved in the uptake of amino acids and oligopeptides were also up-regulated. The strong upregulation of
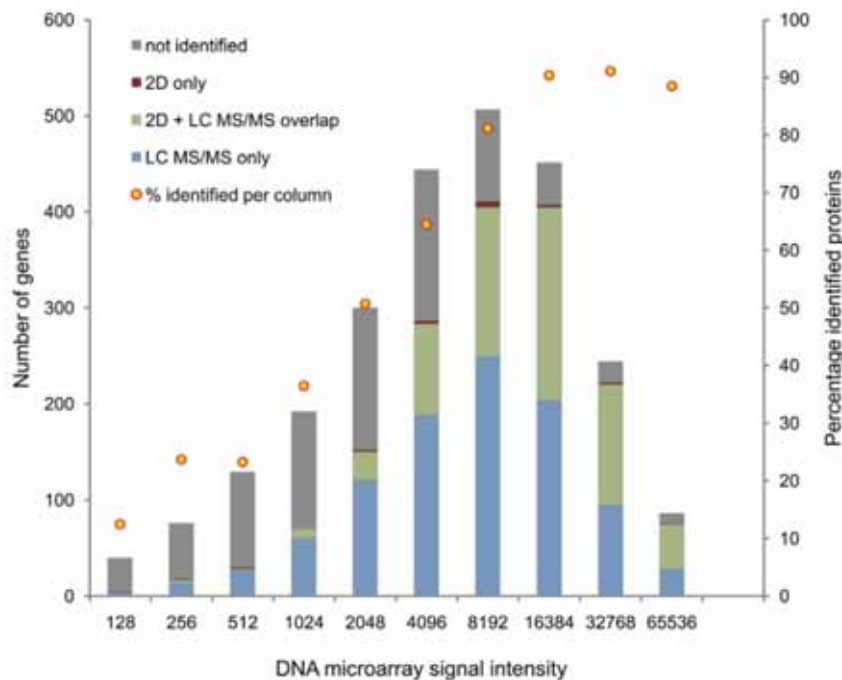
_____

**Figure 4. Comparison of identified proteins depending on the expression strength on RNA level.** Up to 90% of the gene products of strongly expressed genes (definition of expressed genes – cf. Text S1) are identified, whereas low abundance proteins are clearly underrepresented among the identified proteins.
doi:10.1371/journal.pone.0008176.g004

many transport proteins is a result of glucose starvation in non-growing cells.

Among the proteins showing decreased quantities, the high affinity iron compound ABC transporters were most striking (Figure 7). Five of the six down-regulated proteins represent lipoproteins that are responsible for binding of iron substrates. However, the physiological significance of this finding remains to be elucidated [26].

In general the transition from exponential growth to stationary phase is associated with broad quantitative changes in the membrane proteome involving both protein induction and degradation in non-growing cells (30% of the proteins covered being regulated). Nevertheless, there are proteome facets which seem to remain constant in quantity throughout different phases of growth. These include the Sec machinery for protein translocation, which serves in membrane protein biogenesis and protein secretion during all growth stages [27]. Summarizing this part, the present study documents the highest quantitative coverage of a bacterial membrane proteome to date.

## Cell Surface-Associated and Extracellular Proteins

From an infection-related point of view, secreted proteins still bound to the cell surface or secreted into the extracellular space form the most crucial class of proteins of *S. aureus*, because most of the virulence factors can be found in these two subproteomic fractions.

The group of cell surface-exposed proteins contains, besides membrane proteins with extended extracellular loops, also lipid-modified proteins, proteins covalently coupled to peptidoglycan by sortase, and cell wall-associated proteins not covalently bound to cell wall components [28]. To define this class of proteins, a biotinylation approach was employed with a reagent that is unable to enter the cell, thus strongly favouring the biotinylation of surface-exposed protein domains. After affinity purification on NeutrAvidin columns biotinylated proteins were enriched, identified and quantified. By this technique 146 cell surface-associated proteins were identified, among them 48 membrane proteins, 4 proteins covalently attached to peptidoglycan, 37 lipoproteins and 57 cell wall-associated proteins containing a signal peptide, and various cytosolic contaminants. Some of these cell wall-associated proteins were also found in the extracellular space, indicating that this subgroup of secreted proteins is only loosely associated with the cell surface. Lipoproteins and membrane proteins with peripheral loops were found in the membrane fraction as well as in the biotinylated fraction as expected (Figure S5). Surface adhesins such as the fibronectin-binding proteins (Fnbps) [29] and the fibrinogen-binding protein ClfB [30] are preferentially synthesized during growth, whereas ClfA is higher expressed in non-growing cells [31]. Judged from the quantitative biotinylation approach ClfB showed a six times decreased level whereas the ClfA level increased more than three times in non-growing cells compared to growing cells. Other components like the immuno-dominant protein IsaB [32,33] were present in more than tenfold increased amount on the surface of cells in stationary phase (Figure S6). Furthermore, there are also many changes in the level and pattern of membrane and lipoproteins covered by this biotinylation approach. Again, this is the highest coverage of identified and quantified cell surface-exposed proteins published so far.

Following virulence protein production along the growth curve, we observed a shift from the synthesis of certain cell surface-anchored proteins in growing cells, which may facilitate host cell adhesion, biofilm formation or even cell invasion [1] to the
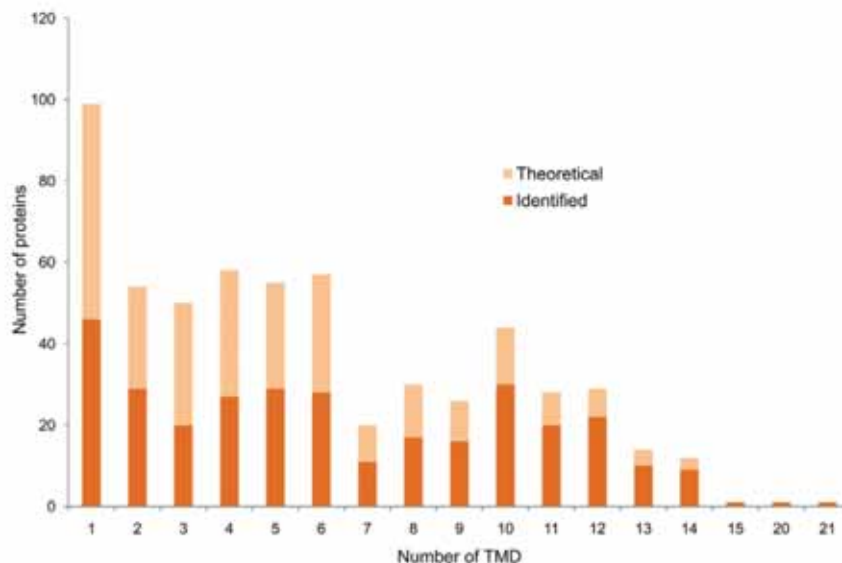
**Figure 5. Coverage of gene functional categories.** The coverage of gene functional categories was displayed by using the KEGG orthology of *S. aureus* COL which was mapped in a Voronoi treemap layout [38]. Treemaps display hierarchically organized information by using a space filling approach according to their hierarchy level (A). Smallest blue mosaic tiles correspond to proteins belonging to metabolic pathways (glycolysis, TCA etc.) which depict the metabolic brunches (e.g. carbohydrates, nucleotides) and which belong to metabolism at all. Changes in protein amount caused by starvation are indicated in shade of green (reduced level) and red (increased level). Proteins labeled in grey were not quantified. (B). Homogeneously regulated clusters have been indicated.
doi:10.1371/journal.pone.0008176.g005

**Figure 6. Numbers of theoretical and identified integral membrane proteins.** Comparison of identified integral membrane proteins with the numbers of theoretical membrane proteins considering predicted trans-membrane domains (TMDs). This comparison indicates that there was no bias against proteins with a higher number of TMDs.
doi:10.1371/journal.pone.0008176.g006

synthesis of extracellular proteins in the stationary phase. These extracellular proteins are virulence factors, such as toxins and superantigens, but also enzymes which are required for cell spreading. The analysis of the secretome is the method of choice for covering the entity of extracellular virulence factors which are extremely variable in different *S. aureus* strains. Usually, gel-based approaches guarantee a high coverage of this fraction [34], but gel-free approaches yield protein identification that overlap with those from gel-based approaches and also allow detection of additional extracellular proteins. In this study we used a quantitative GeLC-MS/MS approach and identified 57 extracellular proteins particularly in the stationary phase (Figure S5). It is relevant to note that not all proteins characterized by an N-terminal signal sequence and signal peptidase I cleavage site will be secreted to the extracellular space. The existence of so called Sec-attached proteins which are retained at the cytoplasmic membrane by still unknown mechanisms was discovered recently [35]. Unfortunately, it is still difficult to predict the proteins that belong to this class. Importantly, the present panorama view of the *S. aureus* secretome can now be used to address crucial issues, such as the assignment of secreted proteins to virulence regulons, or assessment of the virulence potential of well-defined clinical isolates.

## A Proteomics and Transcriptomic Comparison

Using the genes that were detected as expressed above background level in our transcriptional profiling study as a baseline, the products of about 80% of the expressed genes were identified at the proteome level (Figure 8). Almost all proteins were detected for highly expressed genes, whereas this proportion covered by proteomics decreases with decreasing gene expression level (see Figure 4). These results indicate that some proteins, in most cases probably low-abundance proteins, are still missing in our protein inventory of cells because of the low expression rate of the corresponding genes. Furthermore, we found that gene products present at very low level were exclusively detected by

LC/MS-MS-techniques compared to gel-based proteomics reinforcing the statement of their higher sensitivity. In growing cells there is a good correlation between transcriptomics and proteomics. This is no longer true for non-growing cells mainly because many proteins synthesized in growing cells are still present but no longer synthesized.

## Conclusion and Outlook

Our progress on the way towards the entire proteome of *S. aureus* becomes evident from the combined subproteomic data (Figure 8). In total a coverage of about 80% of the expressed genes was obtained. Notably, proteins were identified with an average sequence coverage of 35%, although a much lower coverage would have been sufficient to identify a protein with high confidence.

Besides detailed information on the expression of proteins, reliable information about the localization of almost all proteins in the cell was gained. Several proteins were exclusively identified in one of the four subproteomic fractions indicating their localization in the cytosol or targeting to the cytosolic membrane, cell wall or even to the extracellular milieu. For proteins of still unknown function this information is of special interest as it provides clues for the identification of their cellular role. An interesting example is SACOL 1373, a protein of unknown function that remains membrane-associated during growth, but is secreted into the medium during stationary phase.

The current study does not only provide the protein inventory of *S. aureus* cells, but also quantitative data on the majority of proteins expressed in growing compared to non-growing cells in all four subproteomic fractions. A list containing all data on up- and down-regulation of the entire protein inventory, including information on subcellular localization of proteins, is presented in Table S2. The clustering of major protein groups is indicated by the color code used in Figure 2. In many cases proteins are either down- (green) or up-regulated (red). On the other hand there are also many proteins probably no longer synthesized but stable in
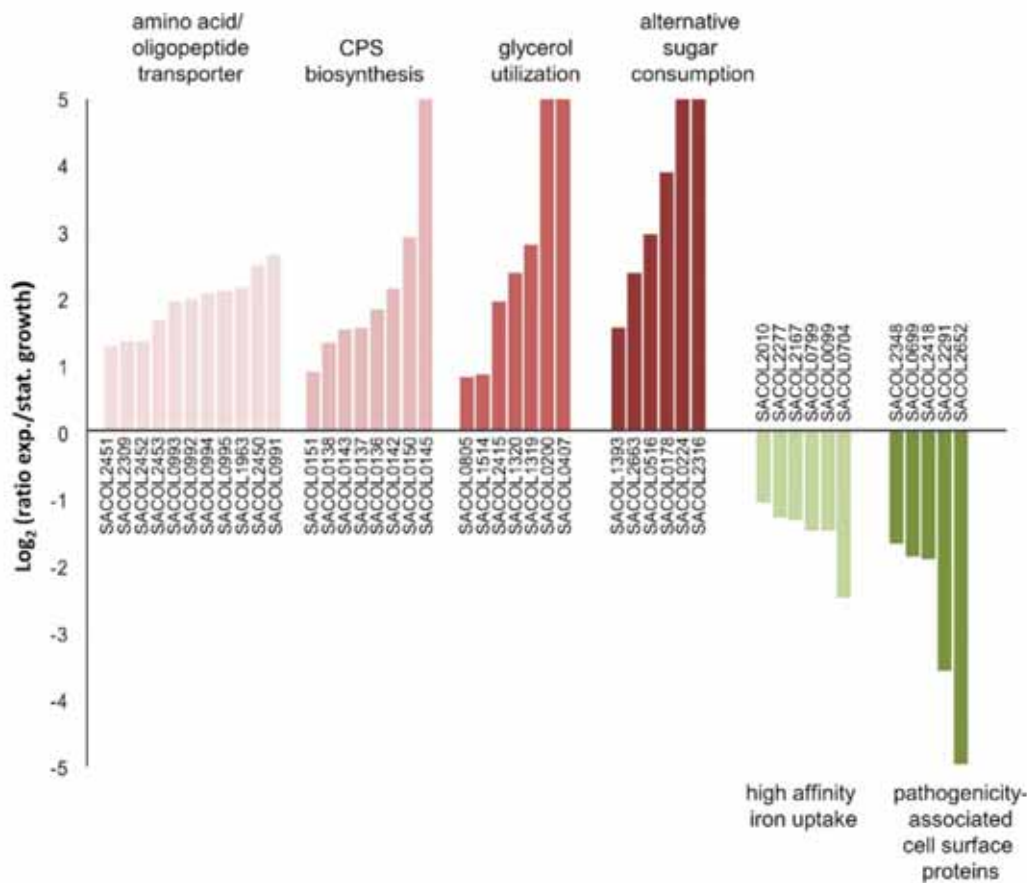
**Figure 7. Membrane proteins which were found in altered amount.** Several of the membrane proteins increased in amount in stationary phase are involved in the transport of amino acids or oligopeptides [from left to right: SACOL2451, SACOL2309, SACOL2452, SACOL2453, SACOL0993 (OppD), SACOL0992 (OppC), SACOL0994 (OppF), SACOL0995, SACOL1963, SACOL2450, SACOL0991 (OppB)], CPS (capsular polysaccharide) biosynthesis [SACOL0151 (Cap5P), SACOL0138 (Cap5C), SACOL0143 (Cap5H), SACOL0137 (Cap5B), SACOL0136 (Cap5A), SACOL0142 (Cap5G), SACOL0150 (Cap5O), SACOL0145 (Cap5J), glycerol utilization [SACOL0805, SACOL1514 (GpsA), SACOL2415 (Gpm), SACOL1320 (GlpK), SACOL1319 (GlpF), SACOL0200, SACOL0407 (GlpT)], or consumption of alternative sugars [SACOL1393 (GlcT), SACOL2663, SACOL0516, SACOL0178, SACOL0224, SACOL2316]. Of the proteins down-regulated in stationary phase numerous could be assigned to the category of iron uptake facilitation [SACOL2010, SACOL2277, SACOL2167, SACOL0799, SACOL0099 (SirA), SACOL0704] or to the class of cell surface proteins which are associated to pathogenicity of *S. aureus* [SACOL2348, SACOL0699 (Pbp4), SACOL2418 (Sbi), SACOL2291 (SsaA2), SACOL2652 (ClfB)]. Proteins indicated with a log2 ratio of 5 and -5 represent "on"/"off" proteins, respectively.
doi:10.1371/journal.pone.0008176.g007

non-growing cells (yellow). This quantitative proteomics approach comparing growing *vs.* non-growing cells is only a starting point, and the huge amount of proteome data requires many follow-up studies in order to understand the protein dynamics in a physiological context. Furthermore, this approach can now be extended to many other physiological issues, infection-related ones included. The future challenge will now be to transfer this huge amount of proteome data to new hypotheses which will finally generate new physiological knowledge.

This progress in bacterial proteomics will have a great impact on future studies on cell physiology of *S. aureus* and related Gram-positive bacteria, and it will thus dramatically improve our knowledge about the behavior of these pathogens in the human host. First of all protein expression profiling in response to environmental stimuli can now be accomplished with a high proteomic coverage to identify almost all proteins that make the cell viable. Furthermore, the virulence protein inventory of a cell can be captured, analyzed, and quantified. Protein expression

profiling, however, is only the first step towards the understanding of life processes. Using a combination of gel-based and gel-free proteomics the final destination and fate of each single protein can be followed at a proteome-wide scale [13]. Detailed studies on the dynamics of the membrane proteome, on the phosphoproteome and its role in signal transduction, on the damage, repair and degradation of proteins or on the protein interaction network will provide a new quality in understanding cell physiology and pathophysiology of *S. aureus*, thereby opening a new era in infection-related research.

## Materials and Methods

*S. aureus* COL was grown in $^{15}$N-labeled or unlabeled BioExpress® 1000 medium (Cambridge Isotope Laboratories, Andova, MA, USA.) under vigorous agitation at 37°C. Cells were harvested at an $OD_{600}$ of 0.5 to sample exponentially growing cells or five hours after the cells completely reached stationary phase
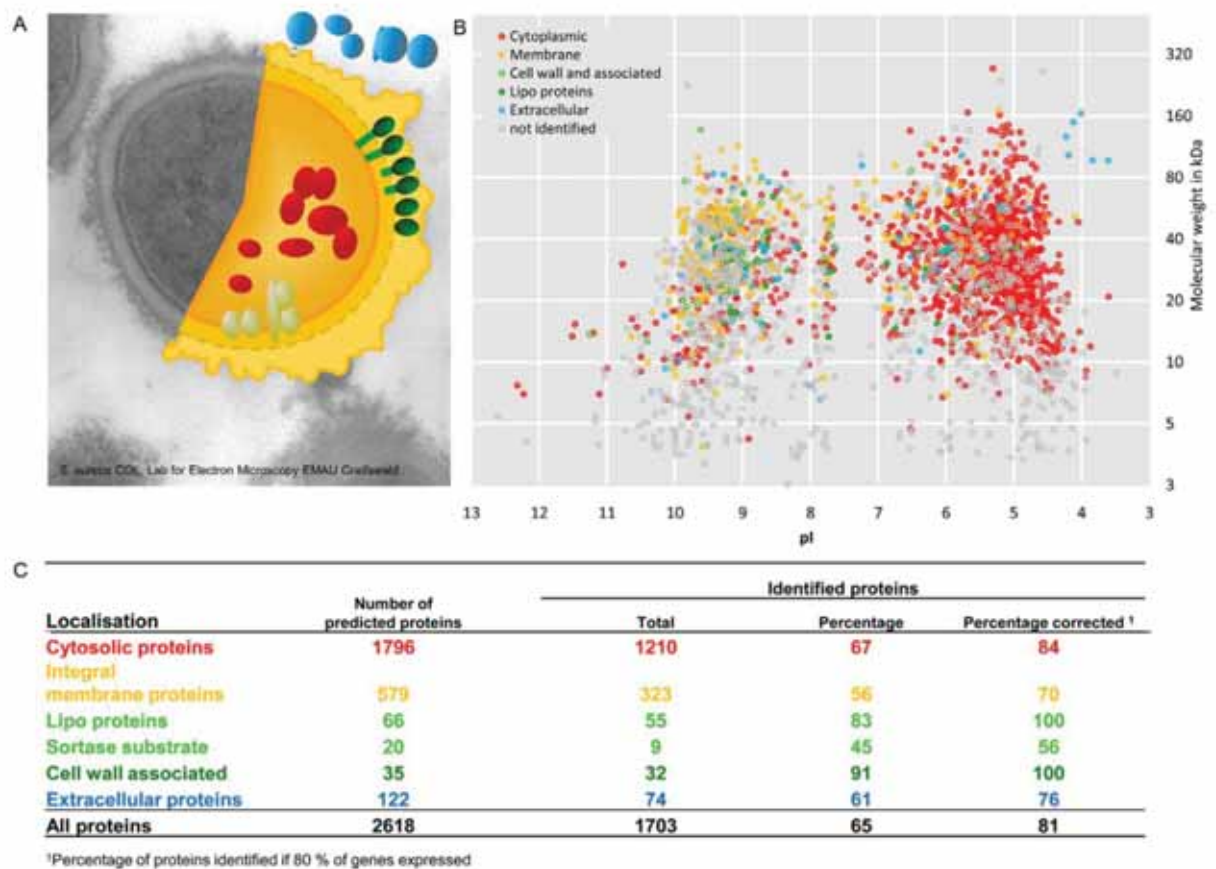
**Figure 8. Summary of identified proteins.** This summary considers the predicted localization in the cell considering the expected gene expression. **A:** Electron microscopy picture of *S. aureus* COL with illustrated subcellular protein fractions. **B:** Identified proteins colored according to the predicted localization in the cell displayed in a theoretical 2-D gel. **C:** Summary of identified proteins in relation to expected gene expression. Software tools used for the prediction of protein localization are given in the full methods section (Text S1).
doi:10.1371/journal.pone.0008176.g008

($OD_{600}$ about 4). Cells were disrupted and protein raw extracts of cytosolic and membrane proteins were prepared as described [9,22]. Extracts of extracellular proteins were prepared by TCA precipitation at 4°C over night. For preparation of extracts of surface-associated proteins, cells were treated with Sulfo-NHS-SS-Biotin prior to disruption. Biotinylated proteins were enriched by NeutrAvidin agarose affinity purification. Protein raw extracts of labeled and unlabeled cells were either analyzed separately by 2-D gel image analysis or 1:1 mixed and further separated via 1-D PAGE (Figure 1). Proteins separated in 2-D PAGE were identified via MALDI-ToF/MS after in-gel digestion [36] using a Proteome Analyzer 4800 (Applied Biosystems, Foster City, CA) and subsequent MASCOT search. Peptides originated from in-gel digestion of 1-D gels were separated by RP chromatography using a nanoACQUITY™ UPLC™ System (Waters, Milford, MA). Separated peptides were online subjected to MS/MS analysis in a LTQ-Orbitrap (Thermo Fisher Scientific, Waltham, MA, USA), which was run in data-dependent mode. Resulting spectra were searched against a target-decoy database using SEQUEST (version 27 rev. 12). The database was composed of all protein sequences of *S. aureus* COL extracted from the National Center for Biotechnology Information (NCBI) bacteria genomes plus com-

mon contaminants and an appended set of the reversed sequences created by Bioworks Browser. Quantitative data were obtained implementing the Census software tool [37].

**Full Methods** and any associated references are available as Text S1.

## Supporting Information

**Text S1** Supplementary materials and methods.
Found at: doi:10.1371/journal.pone.0008176.s001 (0.09 MB DOC)

**Figure S1** Reference 2-D map of cytoplasmic proteins of S. aureus COL grown in Bioexpress 1000 (Cambridge Isotope Laboratories, Andova, MA, USA) medium in the pI range 4–7. Cells were harvested at OD600 of 0.5 to sample exponentially growing cells. In this reference map 553 proteins were sketched, represented in 732 spots. It has to be considered that 428 proteins appear in single spots whereas 125 proteins are distributed in at least two spots, indicating post-translational modifications.
Found at: doi:10.1371/journal.pone.0008176.s002 (3.87 MB TIF)

**Figure S2** Identified proteins in the 2-D map (pI 4–7 and 6–10) involved in the main metabolic pathways. Green labeled proteins

A Proteomic View of an Important Human Pathogen- Towards the Quantification of the
Entire *Staphylococcus aureus* Proteome

_____

Quantitation *S.aureus* Proteome

were identified whereas yellow labeled proteins were not identified in the 2-D gel based approach.
Found at: doi:10.1371/journal.pone.0008176.s003 (4.69 MB TIF)

**Figure S3** Northern blot analysis of selected genes. Total RNA was extracted from S. aureus COL cells cultured in BioExpress® medium at different time points during growth and stationary phase. Equal amounts of RNA were separated by denaturing RNA gel electrophoresis and blotted onto positively charged nylon membranes. The membranes were hybridized with digoxigenin-labeled RNA probes for the respective genes. The diagrams show the quantified Northern blot signal intensities bevor (open squars) and after scaling (closed squars) for the decrease in total RNA content. The schematic representations of the gene loci are based on the sequence of S. aureus COL. Major transcripts representing the predicted transcriptional organization of the operons based on the Northern blot analyses are shown as arrows. Dual channel false color images show the results from 2-D gel-based expression analysis for the genes analyzed by Northern blot: growing cells (green) and non-growing cells (red).
Found at: doi:10.1371/journal.pone.0008176.s004 (1.26 MB TIF)

**Figure S4** Correlation between the protein ratios of biological replicates used to obtain the quantitative data. The calculated correlation was 0.84. Selected regulated functional groups are highlighted.
Found at: doi:10.1371/journal.pone.0008176.s005 (0.26 MB TIF)

**Figure S5** Regulation of proteins identified in the fraction of cell surface-associated and extracellular proteins.
Found at: doi:10.1371/journal.pone.0008176.s006 (1.46 MB TIF)

**Figure S6** Regulated cell surface and cell wall-associated proteins. Class I proteins are colored in yellow, class II proteins in green and class III proteins in red. 1-SACOL2652 (ClfB), 2-SACOL1940, 3-SACOL2348, 4-SACOL1725 (RplT), 5-SACOL1066 (Fmt), 6-SACOL1825, 7-SACOL0938 (DltD), 8-SACOL0712, 9-SA-

COL2383, 10-SACOL1687, 11-SACOL0551, 12-SACOL1111, 13-SACOL1895, 14-SACOL1028 (HtrA), 15-SACOL0021 (YycH), 16-SACOL0539 (PurR), 17-SACOL0189, 18-SACOL1522 (EbpS), 19-SACOL1989, 20-SACOL0022 (YycI), 21-SACOL1140 (IsdA), 22-SACOL0610 (SdrE), 23-SACOL1788, 24-SACOL1836, 25-SACOL0968 (SpsA), 26-SACOL1514 (GpsA), 27-SACOL1168 (Efb), 28-SACOL2549, 29-SACOL0050 (Pls), 30-SACOL2019 (SdrH), 31-SACOL0856 (ClfA), 32-SACOL1062 (Atl), 33-SA-COL2002 (Map), 34-SACOL0024, 35-SACOL1847, 36-SA-COL2660 (IsaB).
Found at: doi:10.1371/journal.pone.0008176.s007 (0.65 MB TIF)

**Table S1** Signal intensities and calculated ratios from DNA microarray experiment.
Found at: doi:10.1371/journal.pone.0008176.s008 (0.32 MB PDF)

**Table S2** Identified proteins with corresponding quantitative value.
Found at: doi:10.1371/journal.pone.0008176.s009 (0.29 MB PDF)

**Table S3** Oligonucleotides used in this study.
Found at: doi:10.1371/journal.pone.0008176.s010 (0.01 MB PDF)

## Author Contributions

Conceived and designed the experiments: DB MH. Performed the experiments: KH SS DZ JPF. Analyzed the data: DB KH SS DZ JPF AO SF DA JB SE UV JMvD. Contributed reagents/materials/analysis tools: AO. Wrote the paper: DB UV MH.

## References

1. Boucher HW, Corey GR (2008) Epidemiology of methicillin-resistant *Staphylococcus aureus*. Clin Infect Dis 46: 344–9.
2. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, et al. (1995) Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. Electrophoresis 7: 1090–4.
3. O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 10: 4007–21.
4. Cravatt BF, Simon GM, Yates JR 3rd (2007) The biological impact of mass-spectrometry-based proteomics. Nature 7172: 991–1000.
5. Loevenich SN, Brunner E, King NL, Deutsch EW, Stein SE, et al. (2009) The *Drosophila melanogaster* PeptideAtlas facilitates the use of peptide data for improved fly proteomics and genome annotation. BMC Bioinformatics 10: 59.
6. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 12: 1367–72.
7. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, et al. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature 7217: 1251–4.
8. Gill SR, Fouts DE, Archer GL, Mongodin EF, Deboy RT, et al. (2005) Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J Bacteriol 7: 2426–38.
9. Fuchs S, Pané-Farré J, Kohler C, Hecker M, Engelmann S (2007) Anaerobic gene expression in *Staphylococcus aureus*. J Bacteriol 11: 4275–89.
10. Hochgräfe F, Wolf C, Fuchs S, Liebeke M, Lalk M, et al. (2008) Nitric oxide stress induces different responses but mediates comparable protein thiol protection in *Bacillus subtilis* and *Staphylococcus aureus*. J Bacteriol 14: 4997–5008.
11. Rogasch K, Rühmling V, Pané-Farré J, Höper D, Weinberg C, et al. (2006) Influence of the two-component system SaeRS on global gene expression in two different *Staphylococcus aureus* strains. J Bacteriol 22: 7742–58.
12. Markert S, Arndt C, Felbeck H, Becher D, Sievert SM, et al. (2007) Physiological proteomics of the uncultured endosymbiont of *Riftia pachyptila*. Science 5809: 247–50.

13. Hecker M, Antelmann H, Büttner K, Bernhardt J (2008) Gel-based proteomics of Gram-positive bacteria: a powerful tool to address physiological questions. Proteomics 23-24: 4958–75.
14. Bandow JE, Hecker M (2007) Proteomic profiling of cellular stresses in *Bacillus subtilis* reveals cellular networks and assists in elucidating antibiotic mechanisms of action. Prog Drug Res 79: 81–101.
15. Batzilla CF, Rachid S, Engelmann S, Hecker M, Hacker J, et al. (2006) Impact of the accessory gene regulatory system (Agr) on extracellular proteins, codYexpression and amino acid metabolism in *Staphylococcus epidermidis*. Proteomics 12: 3602–13.
16. Resch A, Leicht S, Saric M, Pásztor L, Jakob A, et al. (2006) Comparative proteome analysis of *Staphylococcus aureus* biofilm and planktonic cells and correlation with transcriptome profiling. Proteomics 6: 1867–77.
17. Rollenhagen C, Sörensen M, Rizos K, Hurvitz R, Bumann D (2004) Antigen selection based on expression levels during infection facilitates vaccine development for an intracellular pathogen. Proc Natl Acad Sci U S A 23: 8739–44.
18. Becker D, Selbach M, Rollenhagen C, Ballmaier M, Meyer TF, et al. (2006) Robust Salmonella metabolism limits possibilities for new antimicrobials. Nature 440: 303–7.
19. Gerth U, Kock H, Kusters I, Michalik S, Switzer RL, et al. (2008) Clp-dependent proteolysis down-regulates central metabolic pathways in glucose-starved *Bacillus subtilis*. J Bacteriol 1: 321–31.
20. Kuroda A, Nomura K, Ohtomo R, Kato J, Ikeda T, et al. (2001) Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in *E. coli*. Science 293: 705–8.
21. Dreisbach A, Otto A, Becher D, Hammer E, Teumer A, et al. (2008) Monitoring of changes in the membrane proteome during stationary phase adaptation of *Bacillus subtilis* using in vivo labeling techniques. Proteomics 10: 2062–76.
22. Wolff S, Hahne H, Hecker M, Becher D (2008) Complementary analysis of the vegetative membrane proteome of the human pathogen *Staphylococcus aureus*. Mol Cell Proteomics 8: 1460–8.
23. Abram F, Su WL, Wiedmann M, Boor KJ, Coote P, et al. (2008) Proteomic analyses of a *Listeria monocytogenes* mutant lacking sigmaB identify new

A Proteomic View of an Important Human Pathogen- Towards the Quantification of the Entire *Staphylococcus aureus* Proteome

_____

Quantitation *S.aureus* Proteome

components of the sigmaB regulon and highlight a role for sigmaB in the utilization of glycerol. Appl Environ Microbiol 3: 594–604.

24. Joseph B, Mertins S, Stoll R, Schär J, Umesha KR, et al. (2008) Glycerol metabolism and PrfA activity in *Listeria monocytogenes*. J Bacteriol 15: 5412–30.

25. Hames C, Halbedel S, Hoppert M, Frey J, Stülke J (2009) Glycerol metabolism is important for cytotoxicity of *Mycoplasma pneumoniae*. J Bacteriol 3: 747–53.

26. Maresso AW, Schneewind O (2006) Iron acquisition and transport in *Staphylococcus aureus*. Biometals 2: 193–203.

27. Sibbald MJ, Ziebandt AK, Engelmann S, Hecker M, de Jong A, et al. (2006) Mapping the pathways to staphylococcal pathogenesis by comparative secretomics. Microbiol Mol Biol Rev 3: 755–88.

28. Gatlin CL, Pieper R, Huang ST, Mongodin E, Gebregeorgis E, et al. (2006) Proteomic profiling of cell envelope-associated proteins from *Staphylococcus aureus*. Proteomics 5: 1530–49.

29. Saravia-Otten P, Müller HP, Arvidson S (1997) Transcription of *Staphylococcus aureus* fibronectin binding protein genes is negatively regulated by agr and an agr-independent mechanism. J Bacteriol 17: 5259–63.

30. McAleese FM, Walsh EJ, Sieprawska M, Potempa J, Foster TJ (2001) Loss of clumping factor B fibrinogen binding activity by *Staphylococcus aureus* involves cessation of transcription, shedding and cleavage by metalloprotease. J Biol Chem 32: 29969–78.

31. Loughman A, Fitzgerald JR, Brennan MP, Higgins J, Downer R, et al. (2005) Roles for fibrinogen, immunoglobulin and complement in platelet activation promoted by *Staphylococcus aureus* clumping factor A. Mol Microbiol 3: 804–18.

32. Lorenz U, Ohlsen K, Karch H, Thiede A, Hacker J (2000) Immunodominant proteins in human sepsis caused by methicillin resistant *Staphylococcus aureus*. Adv Exp Med Biol 485: 273–8.

33. Clarke SR, Foster SJ (2006) Surface adhesins of *Staphylococcus aureus*. Adv Microb Physiol 51: 187–224.

34. Ziebandt AK, Becher D, Ohlsen K, Hacker J, Hecker M, et al. (2004) The influence of agr and sigmaB in growth phase dependent regulation of virulence factors in *Staphylococcus aureus*. Proteomics 10: 3034–47.

35. Tjalsma H, van Dijl JM (2005) Proteomics-based consensus prediction of protein retention in a bacterial membrane. Proteomics 17: 4472–82.

36. Eymann C, Dreisbach A, Albrecht D, Bernhardt J, Becher D, et al. (2004) A comprehensive proteome map of growing *Bacillus subtilis* cells. Proteomics 4: 2849–76.

37. Park SK, Venable JD, Xu T, Yates JR 3rd (2008) A quantitative analysis software tool for mass spectrometry-based proteomics. Nat Methods 4: 319–22.

38. Balzer M, Deussen O (2005) Voronoi treemaps. In Proc. IEEE Symposium on Information Visualization, InfoVis 05: 49–56.

_____

*Systems- wide temporal proteomic profiling in glucose-starved Bacillus subtilis*

_____

## ARTICLE

# Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*

Andreas Otto[1], Jörg Bernhardt[1], Hanna Meyer[2], Marc Schaffer[1], Florian-A. Herbst[1], Juliane Siebourg[3], Ulrike Mäder[1,4], Michael Lalk[2], Michael Hecker[1] & Dörte Becher[1]

Functional genomics of the Gram-positive model organism *Bacillus subtilis* reveals valuable insights into basic concepts of cell physiology. In this study, we monitor temporal changes in the proteome, transcriptome and extracellular metabolome of *B. subtilis* caused by glucose starvation. For proteomic profiling, a combination of *in vivo* metabolic labelling and shotgun mass spectrometric analysis was carried out for five different proteomic subfractions (cytosolic, integral membrane, membrane, surface and extracellular proteome fraction), leading to the identification of ~52% of the predicted proteome of *B. subtilis*. Quantitative proteomic and corresponding transcriptomic data were analysed with Voronoi treemaps linking functional classification and relative expression changes of gene products according to their fate in the stationary phase. The obtained data comprise the first comprehensive profiling of changes in the membrane subfraction and allow in-depth analysis of major physiological processes, including monitoring of protein degradation.

[1] Ernst-Moritz-Arndt-Universität Greifswald, Institute for Microbiology, Greifswald 17487, Germany. [2] Ernst-Moritz-Arndt-Universität Greifswald, Institute of Pharmacy, Greifswald 17487, Germany. [3] ETH Zürich, Computational Biology Group, Mattenstrasse 26, Basel CH-4058, Switzerland. [4] Ernst-Moritz-Arndt-Universität Greifswald, Interfaculty Institute for Genetics and Functional Genomics, Greifswald 17487, Germany. Correspondence and requests for materials should be addressed to D.B. (email: dbecher@uni-greifswald.de).

# ARTICLE

For a complete understanding of physiological processes in a cell, the integration of tailored approaches addressing different levels of regulation is needed. Therefore *Bacillus subtilis*, a Gram-positive prokaryote serving as a model organism[1], was subject to a wide range of transcriptomic[2,3], proteomic[4,5] and metabolomic studies[6,7]. Unlike transcriptome analysis, proteomic studies do not necessarily cover all possible gene products of the cell. For a long time, the entirety of all proteins represented a far too complex mixture to be analysed at once, due to the large ranges of protein abundance, pI (iso-electric point), molecular weight, solubility or localization[8]. A recent example demonstrates that these limitations in analytics of large-scale proteomics can be overcome by application of different digest enzymes and gas-phase fractionation during mass spectrometry analysis[9].

Soluble proteins are the most intensively studied species of proteins due to their favourable accessibility. In two dimensional gel-based studies, the majority of the proteins localized in the cytosol, the extracellular space or the cell wall have been differentially accessed[10]. More effort has to be spent on proteins with multiple membrane-spanning domains, extreme pIs or molecular weights or with low abundance. In contrast to highly abundant soluble proteins[11,12] they need to be enriched and specifically analysed.

Elaborate methodologies are known for subfractionation of cell extracts other than cytosolic proteins. Extracellular proteins secreted into the growth medium can be acid precipitated and subsequently analysed[13]. Cell wall bound and surface localized proteins are addressed by LiCl extraction and subsequent proteomic analysis via two dimensional gel electrophoresis[14]. Proteins from the cell surface of intact cells can be biotinylated and subsequently affinity purified[15].

Owing to the hydropathy and low abundance of the proteins[16], the membrane proteome still represents one of the least accessible subproteomes. Proteomic analyses of membrane proteins of *B. subtilis* have been developed over a longer period of time[5,17,18], Wolff *et al.*[19] and Hahne *et al.*[18] have proven that a combination of highly complementary enrichment and analytical techniques allows a very comprehensive analysis of the membrane proteome in Gram-positive bacteria.

The aim of this study was the systems-wide profiling of compartment-specific changes on the proteome level in *B. subtilis* caused by glucose starvation. State-of-the-art complete metabolic labelling with stable nitrogen isotopes as described by MacCoss *et al.*[20] and high mass accuracy mass spectrometry were applied to retrieve a maximum level of proteomic data. We investigated cytosolic proteins, two different membrane fractions consisting of transmembrane and membrane-attached proteins (membrane-shaving fraction (MSF) and enriched membrane fraction (EMF)), proteins attached to the cell surface (biotinylation-enrichment fraction (BEF)) and secreted extracellular proteins (extracellular fraction) of *B. subtilis*. To substantiate and complement the proteomic profiling, transcriptome data were acquired together with the quantification of extracellular metabolites. The proteomic and transcriptomic data were visualized by using Voronoi treemaps to enable the intuitional accessibility and immediate comparison of the systems-wide data.

In this study, we demonstrate temporal dynamic processes at the protein level that are so far unrivalled. Particularly, the temporal survey of membrane proteins in combination with the other more easily accessible subproteomes allows a broad monitoring of protein dynamics. The data reveal a comprehensive and unbiased overview on the physiological processes in *B. subtilis* as a model enabling in-depth analyses, including post-translational regulation of biosynthetic pathways, and an overview on time-dependent processes in the bacterial membrane as protein synthesis and degradation.

## Results

**Qualitative and quantitative analysis of the proteome.** Metabolically labelled *B. subtilis* proteins from growing and non-growing cells were fractionated to obtain five different subproteomes for a comprehensive view on regulatory and physiological changes (Fig. 1). Entry into

stationary phase was provoked by glucose deprivation at an optical density (500 nm) of 1 (Supplementary Fig. S1). Based on our workflow, allowing highly accurate (average mass deviation of <2 p.p.m. in our analysis) and reliable data (0.11% false positives on average on protein level), we were able to considerably cover the proteome of *B. subtilis*. In total, 2,142 proteins were identified with at least two
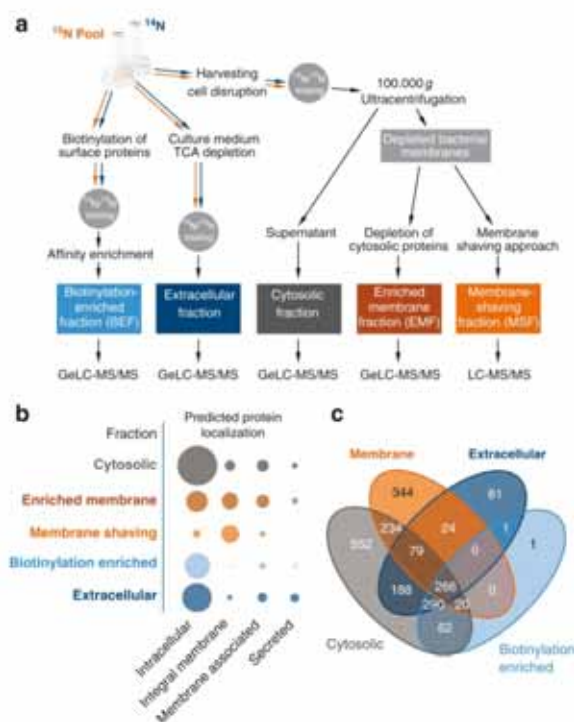


**Figure 1 | Distribution of proteins identified according to the subfractionation workflow. (a)** Workflow of the proteomic experiment. The preparation of the five different subcellular fractions is exemplified with the most crucial steps. The starting point for all proteomic samples is the bacterial batch cell culture. The first crossroads is the separation between complete cells for the BEF (biotinylation-enriched fraction) aiming at the surface-exposed proteins, the cell culture medium for the extracellular proteins and cells for the cytosolic/membrane fractions. ¹⁴N/¹⁵N mixing depicts the stage in sample preparation, where mixing of the labelled pool and the unlabelled samples occurred. The earliest time point in experimental setup is chosen to minimize introduced effects by the labelling procedure to the quantitative data. Proteomic analyses (GeLC-MS and LC-MS) were performed with high-resolution and high mass accuracy mass spectrometry (2,142 total identified proteins with an FPR (false-positive rate) of 0.11% and 1.86 p.p.m. mass accuracy). **(b)** Distribution of proteins identified with a predicted subcellular localization in the different subfractions analysed. The area of the circles is proportional to number of identified proteins. **(c)** VENN plot of protein identifications in the different designated subcellular fractions. Numbers for the membrane fraction are summarized from EMF and MSF. Prediction of subcellular localization is according to Zhou *et al.*[52] Loci of localization assigned were as follows: intracellular; integral membrane proteins (multitransmembrane, multitransmembrane (lipid-modified N termini)); membrane-associated proteins (lipid anchored, LPxTG cell wall anchored, N-terminally anchored (no cleavage site), N-terminally anchored (with cleavage site), C-terminally anchored (with cleavage site), intracellular/TMH start after 60); secreted (secreted via minor pathways (bacteriocin) (no cleavage site), secretory (released) (with cleavage site)).

196

# ARTICLE

**Table 1 | Proteins identified according their predicted subcellular location.**

| Predicted localization | Cytosol | EMF | MSF | BEF | Extracellular | Combined | Theoretical proteome | Coverage |
|---|---|---|---|---|---|---|---|---|
| Intracellular | 1,446 | 411 | 52 | 584 | 767 | 1,511 | 2,797 | 54.0% |
| Integral membrane | 94 | 251 | 287 | 2 | 18 | 361 | 832 | 43.4% |
| Membrane associated | 122 | 155 | 28 | 47 | 75 | 187 | 343 | 54.5% |
| Secreted | 26 | 32 | 3 | 7 | 71 | 83 | 133 | 62.4% |
| All proteins | 1,688 | 849 | 370 | 640 | 931 | 2142 | 4,105 | 52.2% |

Number of proteins identified for every proteomic subfraction analysed and assigned to their theoretical (predicted) localization. The coverage depicts the percentage of identified proteins in all subfractions combined for a specific predicted localization compared with the theoretical proteome.

peptides, which represents 52.2% of the entire predicted proteome (Table 1; Supplementary Data 1). We were able to assign 552 proteins solely found in the cytosolic fraction (Fig. 1), while 344 proteins were exclusively found to be associated with or completely inserted into the bacterial membrane. Furthermore, 81 proteins were identified only in the extracellular space. A substantial number of proteins were found in multiple subcellular fractions: 509 proteins were found in two, 389 in three and 266 proteins in all subproteomes.

Subfractionation of the bacterial cell extract led to an enrichment of proteins with varying physical properties and different localizations. The proteins predicted as intracellular were found mainly in the cytosolic fraction (Fig. 1). As expected, most integral membrane proteins were detected by using the membrane enrichment and the membrane-shaving approach. Membrane-associated proteins can be found primarily in the membrane-enrichment fraction and in the cytosolic fraction. The predicted secreted proteins could be identified to the highest extent in the extracellular fraction.

Based on the identification data, we carried out a relative quantification with the software Census[21]. This software compares the ^{14}N mass traces of the respective samples with the ^{15}N mass traces of pooled labelled protein extracts used as a standard. Altogether, we were able to obtain quantitative data for 2,048 proteins. The quantitative results for every subfraction are provided in Supplementary Data S2–S6.

**Correlation of proteome and transcriptome data.** According to Becher *et al.*[22], three groups of proteins can be distinguished during the time course in cells entering the stationary phase: proteins no longer synthesized in non-growing cells but still present and relatively stable, proteins no longer synthesized and diminishing in amount, probably as a result of degradation, and proteins enriched in non-growing cells due to increased synthesis. In non-growing cells, many proteins are still present and probably (at least partially) active but no longer synthesized. The most significant advantage of a global proteomic study is that it yields information on the relative protein abundance, whereas transcriptomic studies monitor mRNA abundances as an indicator of changes in transcriptional activity and hence are unsuitable for capturing the protein inventory of cells. We combined both approaches and also performed an array-based transcriptome analysis under the same cultivation conditions to correlate our data on proteomic abundance with the corresponding transcriptional gene expression profiles.

**Visualization via Voronoi treemaps.** To support the analysis of our gene expression data from transcriptome and proteome levels, we used the functional gene categorization from KEGG orthology[23,24] and a structured representation of gene regulatory information derived from SubtiWiki[25,26]. Both schemes classify genes/proteins in an acyclic multihierarchical tree graph according to their function or regulation, respectively. For a well-arranged visualization, we adapted treemaps for an intuitive display of large omics data sets with their relative expression data and functional/regulatory classification (Figs 2,3).
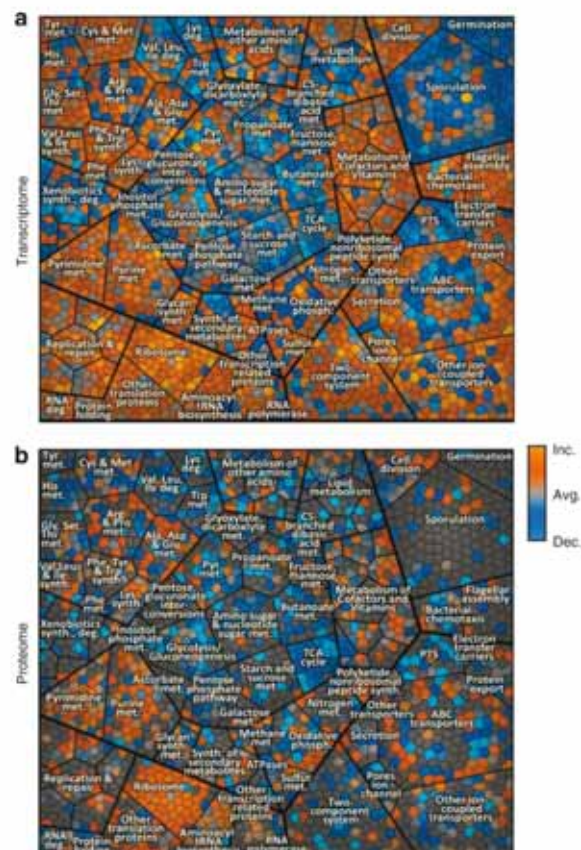
**Figure 2 | KEGG-orthology treemap of growing *B. subtilis*.** (**a**) Gene expression of growing *B. subtilis* compared with average expression during the time course in cells entering the stationary phase. (**b**) Relative protein amount determined in cytosolic fraction of growing *B. subtilis* compared with the average protein amount during the investigated time course. Each cell in the graph displays a single gene locus that belongs to other functionally related elements in parent convex-shaped categories. These are again summarized in higher-level categories. Functionally related elements seem in close neighbourhood to each other. Gene functional data are based on KEGG-orthology (for example, main level (metabolism)/first sublevel (carbohydrate metabolism)/second sublevel/(glycolysis)). To visualize differences in expression level/protein amount compared with the average level colour coding was applied as following: blue—decreased level (dec.), grey—same level as average (avg.), orange—increased level (inc.). These figures are part of the time course analysis (Supplementary Movies S1 and S3), monitoring the changes from exponential growth to late stationary phase.
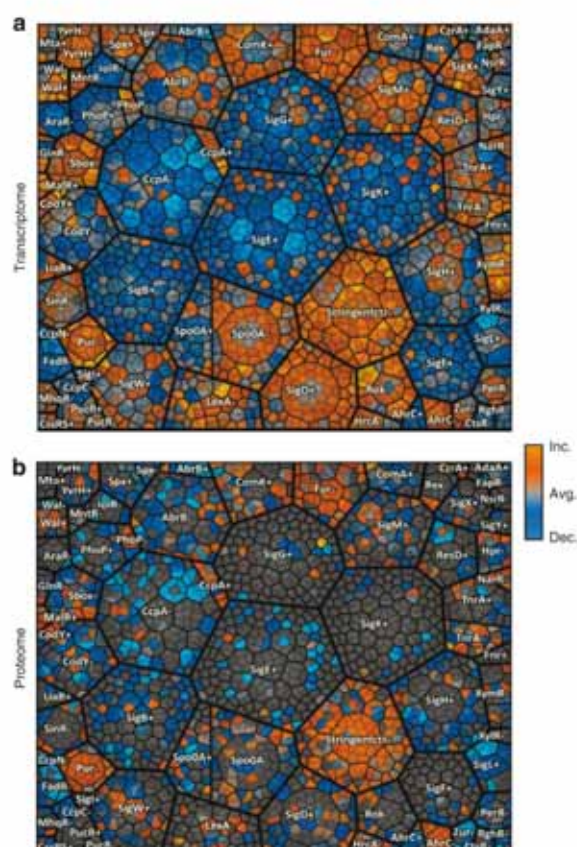
**Figure 3 | Regulon treemap of growing B. subtilis.** (a) Gene expression of growing *B. subtilis* compared with average expression during the time course in cells entering the stationary phase. (b) Relative protein amount determined in cytosolic fraction of growing *B. subtilis* compared with the average protein amount during the investigated time course. Each cell in the graph displays a single gene locus that belongs to other hierarchically/structurally related elements in parent convex-shaped categories. These are again summarized in higher-level regulatory categories. Functionally related elements seem in close neighbourhood to each other. Treemap design is based on hierarchically structured regulatory data (black borders: regulon/thin black borders within the regulons: operon/smallest cells: gene). +/− depict regulons being induced (+) or repressed (−) depending on the regulator assigned to the area. To visualize differences in expression level/protein amount compared with the average level colour coding was applied as following: blue—decreased level (dec.), grey—same level as average (avg.), orange—increased level (inc.). These figures are part of the time course analysis (Supplementary Movies S2 and S4), monitoring the changes from exponential growth to late stationary phase.

The treemap concept was established by Shneiderman[27] and originally served for the display of hierarchical structured data from software modules or file systems. The visualization is not based on a tree graph but on the planar combination of rectangular elements. Balzer and Deussen[28] improved this concept and developed Voronoi Treemaps, which use irregular, convex elements for treemap construction. This concept is very powerful and allows a visual perception of hierarchies that may contain several thousand elements. We adapted this concept[29] to link hierarchically structured regulatory (regulon/operon/gene) or gene functional data (for example, main level (metabolism)/first sublevel (carbohydrate metabolism/second sublevel/(glycolysis)) within Voronoi treemaps with data from gene expression measurements. Our treemaps combine areas representing functionally related elements in parent convex-shaped categories, which again become summarized in higher-level categories. After treemap construction and optimization, functionally related elements seem in close neighbourhood to each other.

Expression data were visualized in the treemaps using a colour gradient: For illustration of protein expression level, colours of the range blue (lower than average) grey (equal to average) and orange (higher than average) were applied to Voronoi cells (Figs 2,3).

Most prominent examples of negative changes in protein amount belong to the nucleotide metabolism, the translational machinery and the biosynthesis of amino acids, as can be deduced from the Supplementary movie S1 illustrating the proteomic cytosolic fraction. Proteins related to the degradation of amino acids and to the metabolism of carbohydrates increase in amount.

The Supplementary movie S2 displays the alteration of protein amounts with a treemap based on the transcriptional regulatory units. Regulons with many genes marking the cessation of exponential growth and the transition towards the stationary phase caused by glucose starvation are: the stringent control (consisting mostly of ribosomal proteins), PurR, Fur, FadR, SigL and CcpA.

For a comparison and a comprehensive representation of processes taking place in the starving cell, the same visualization by Voronoi treemaps is provided for the transcriptome data in Supplementary Movies S3 and S4.

**In-depth analysis of changes on proteome/transcriptome level.** In addition to the general overview on processes in the cell as provided by Voronoi treemaps, it is possible to monitor biological processes on the proteome level through in-depth analysis of the data. To follow the relative change in abundance throughout the complete time course for a specific protein, data for each protein were normalized to the first time point. The quantitative data were subjected to a significance analysis over the time course using analysis of variance. About a quarter of all identified proteins could be assigned to be significantly ($P$-value $< 0.01$) altered in amount (Supplementary data S2–S6).

As expected[7], most obvious changes occurred for proteins of the carbohydrate metabolism: the key enzymes for gluconeogenesis, GapB and PckA, strongly increase in amount, whereas glycolytic enzymes of the glycolysis remain relatively unaffected. Enzymes of the tricarboxylic acid cycle (TCA) increase more than twofold in protein amount (Fig. 4).

The proteomic profiles of several biosynthetic enzymes suggest that they undergo targeted degradation: besides the already known protease substrates[30], the data reveal PyrB, PurF, ArgG, ArgJ, LeuA, Sat and DhbC as new substrate candidates, according to the steeper slopes of decrease in protein amount compared with other members of functionally related biosynthetic pathways (Fig. 5). The decrease in protein amount is in agreement with the decrease at the transcript level.

The analysis of soluble proteins of the major metabolic pathways is well established and a wealth of information is known on *B. subtilis* coping with nutrient stresses. On the contrary, the membrane proteome covering the interface of the cell to the environment has been underrepresented in proteomic studies for a long time. For a broad understanding of processes in the model organism *B. subtilis*, this subgroup was included in the present study.

Proteins involved in glycerol, ribose, lactate, nucleoside, succinate and fumarate uptake and zinc uptake exhibit an increased amount in the stationary phase. Conversely, a decrease was found for transporters of malate, $Fe^{3+}$ citrate, $Fe^{3+}$ hydroxamate, hydroxymethylthiamin (HMP)/thiamine transport and unknown substrates. ABC transporter systems found to be increasing in protein amount are specific for glutamate (GlnH/GlnQ), arginine (RocC, RocE) and for oligopeptides (YhjB, *dpp* and *app* operons).
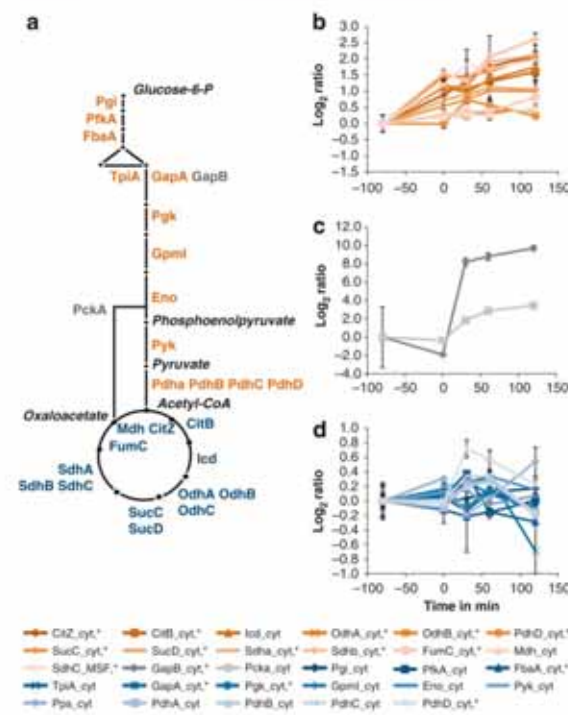
ARTICLE



**Figure 4 | Quantitative changes in proteins of the central carbon metabolism.** (a) Graphical representation of the central carbon metabolism. The main metabolic intermediates are italicized. (b) Change of protein amount for glycolytic enzymes. (c) Change of protein amount for gluconeogenic key enzymes. (d) Change of protein amount for TCA enzymes. Error bars indicate s.d. of the biological replicates (n = 3).
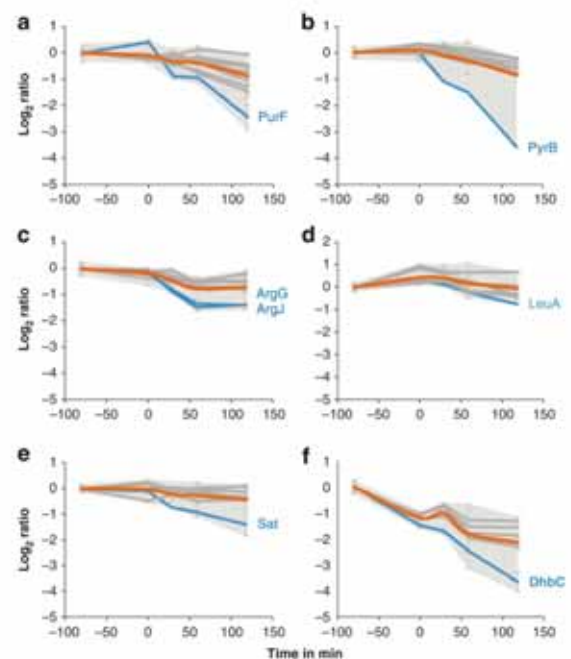
**Figure 5 | Changes in protein amount for proteins of selected biosynthetic pathways.** (a) Members of the purine biosynthetic pathway (PurR; PurB; PurK; PurN; PurH*; PurM;PurD*; PurL*; PurQ*; PurC*; PurF*). (b) Members of the pyrimidine biosynthetic pathway (PyrK; Pyre; PyrC*; PyrF;PyrAA*; PyrD; PyrAB; PyrB). (c) Members of the arginine biosynthetic pathway (ArgC; ArgD; ArgF*; ArgH*;ArgB; CarB; ArgG*; ArgJ*). (d) Members of the leucine/isoleucine biosynthetic pathway (LeuC; LeuD; IlvC; LeuB*;IlvB*; IlvH*; LeuA*). (e) Members of the cysteine biosynthetic pathway (YtiB; YvgQ; CysH; YvgR; CysC; Sat*). (f) Members of the bacillibactin biosynthetic pathway (DhbA*;DhbB*; DhbE*; DhbF*; DhbC*). Log₂ ratios are corrected for the first time point. Asterisks indicate proteins that are significantly altered as determined by analysis of variance (P-value < 0.01). Error bars indicate s.d. of the biological replicates (n = 3). Grey shading: area of maximal s.d.s. Red: centroid of all proteins displayed. Blue: possible protease targets.

## Discussion

The major focus of this work is directed towards a global view on the processes on the proteome level in *B. subtilis* cells faced with glucose starvation. The combination of metabolic labelling with the existing toolbox of methods to access different proteomic subfractions opens up the possibility for systems-wide studies, previously only possible at the transcriptome level.

Consequently, the analysis of 2,142 proteins regarding their dynamic change during a biological important process—the shift from growing to non-growing state—in a single study is a great step forward compared with previous studies of *B. subtilis*, such as Hahne et al.[31], as reviewed by Wolff et al.[12] or as comprehensive studies most recently published by Soufi et al.[32] However, it has to be considered that gene products of lowly expressed genes are still underrepresented in our study (Fig. 6).

Despite our extensive fractionation scheme before mass spectrometric measurement, numerous proteins were additionally identified in other than the predicted subcellular fraction (Table 1). For cytosolic proteins this can be explained by cell lysis after transition to the stationary phase and the still existing potential for improvement regarding the fractionation. In addition, overrepresentation of anticipated cytosolic proteins throughout the whole analysis may be explained by their increased analytical accessibility in our GeLC-MS approach and the high proportion of cytosolic proteins in the most abundantly transcribed groups. Furthermore, after imposition of starvation, cytosolic 'unemployed' proteins are prone to aggregation and degradation[30,33]. Proteins without substrates can form insoluble aggregates in the cell, which are spun down via ultracentrifugation

after cell disruption in our preparation process, as was also observed by Hahne et al.[31] (Supplementary Figs S2–S4).

In general, enrichment of proteins of a specific localization has worked out well. Proteins belonging to a predicted subcellular location were found to the most extent in the corresponding subfraction of our workflow (Fig. 1b). A total of 1,511 cytosolic proteins were identified resulting in a considerable coverage of the 2,797 cytosolic proteins. The large set of quantitative data obtained for the cytosolic proteins enables a display with Voronoi treemaps, representing the main metabolic functional clusters in a single, intuitively accessible graph. This avoids inherent shortfalls of shotgun proteomics yielding only tables with numerical values or representations; for example, by heatmaps lacking functional relatedness in visualization. Both orthology-based and regulon-based Voronoi treemaps consistently display changes on two complementary and physiologically meaningful levels of physiological regulation (Figs 2 and 3).

Necessary physiological regulation at the transition from exponential growth to the late stationary phase implies marked changes at the transcript and proteome level. Concomitant sampling of large-scale proteomics and transcriptomics data opens up the possibility for comparing both on a global scale. As expected, a correlation
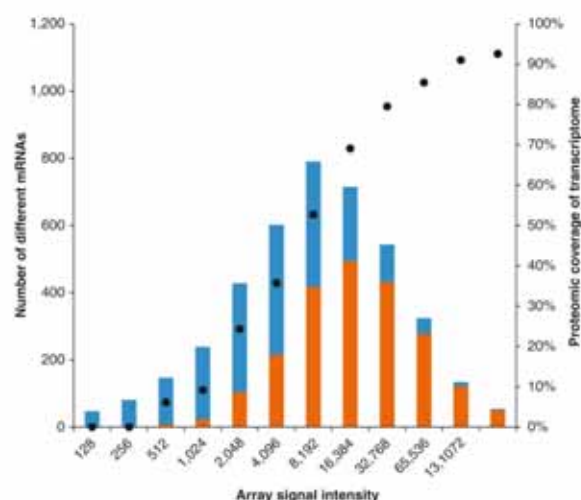
**Figure 6 | Comparison of identified proteins depending on the maximum transcript intensities.** The transcripts are binned according to their maximum array signal intensity over the whole time course. The number of genes belonging to the respective resembling group of transcript intensity is indicated in the bar chart in blue. Displayed in red is the proportion of transcripts with accordingly identified translation products (proteins). The proteomic coverage of highly expressed genes is >90%, whereas for lowly expressed genes the coverage is significantly lower, as displayed with the circles on the y axis on the right side.
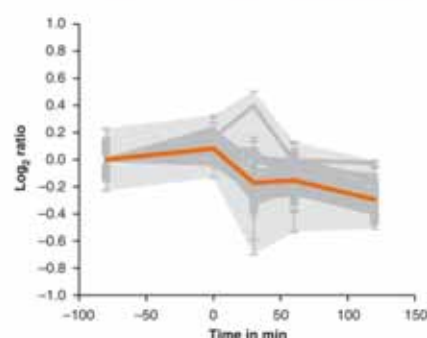


**Figure 7 | Change of protein amount for ribosomal proteins from the cytosolic fraction.** Relative quantitative changes for proteins belonging to the ribosome (YpfD; RpsL; RpsO; RpsJ*; RpmB; RpsP;*RpsD; Rpsl; RpsH; RplD; RpsM*; RplB*; RpsE*; RpsF*; RpsR; RpsS; RpsG*; RpsT*; RplC*; RpmA; RplU*; RplR*; RplQ*; RpsQ*; RplE*; RplK; RpsB; Rpml; RplS*; RplF; Rpll; RplX; RplW*; RpsK*; RplA*; RplP*; RplO*; RpsC; RplV*; RplT*; RplJ*; RplN*; RplL*; RplM*). Log, ratios are corrected for the first time point. Asterisks indicate proteins that are significantly altered as determined by analysis of variance (P-value<0.01). Error bars indicate s.d. of the biological replicates (n=3). Grey shading: area of maximal s.d.s. Red: centroid of all proteins displayed.

analysis of the two levels of regulation according to Fournier *et al.*[34] shows faster changes at the transcript level and a delayed cellular reaction on the proteome level (Supplementary Fig. S5).

The temporal offset of early mRNA and later protein changes could be related to the immediate cellular response on the transcript level and the time needed to accumulate the translated proteins. Accumulated proteins will then remain longer at their cellular localization contrary to their corresponding short-lived transcripts. The same situation holds true for mRNA/protein degradation: immediate responses towards decreased levels of mRNA are followed by slower degradation or even persistence of the protein species.

The data obtained in this study reflect how glucose-starved cells modulate the metabolism by synthesis of new metabolic activity, degradation of specific components under starvation conditions and search for alternative nutrients. As a direct consequence of glucose depletion, the cells switch from glycolytic to gluconeogenic metabolism. This is based on the existing glycolytic pathway and additionally requires new enzymatic activity of glycerine aldehyde-3-phosphate-dehydrogenase (GapB) and phosphoenolpyruvate carboxykinase (PckA)[35] (Fig. 4c). Due to the importance of the TCA in stationary phase metabolism for energy generation, proteins of the TCA are induced significantly[7]. Regulation of expression of genes for the TCA in *B. subtilis* is complex[36,37]. Based on the proteomic data, regulation of the enzymes is in agreement with the knowledge so far[7]. Altogether, information about changes on proteome and transcriptome level for the main carbohydrate metabolism agree with already published data, underpinning the reliability of the data generated in the present work.

Upon transition to the stationary phase, extensive reprogramming of gene expression takes place[3], shutting down gene expression needed for exponential growth, including amino-acid biosynthesis, purine and pyrimidine synthesis and the translational machinery (Supplementary movie S3). Confirming the findings of Eymann *et al.*[38] and as displayed in Supplementary movie S4, genes of the stringent response are immediately downregulated by nutrient starvation. Ribosomal

proteins as part of the negative stringent response regulon decrease markedly in protein amount (Fig. 7), leading to a reduction <70% of the original amount of the ribosomal machinery inside the cell 2h after nutrient depletion. The ribosomes contain a substantial fraction of intracellular protein content[39], and consequently the degradation and recycling of amino acids represent a source of nutrients in times of starvation.

As growth ceases, biosynthesis of amino acids is no longer needed. Besides the downregulation of biosynthetic genes, the protein amount of a large fraction of enzymes decreases. In concordance with transcriptomic data, enhanced degradation is seen for key enzymes of biosynthetic pathways. Noticeably, degradation rates for enzymes catalysing the first committed steps of biosynthetic pathways were higher than those of other constituents of the respective pathways (Fig. 5). While such elevated protease-dependent degradation has been reported previously for PurF and PyrB[30], our large-scale study adds new potential substrates for specific degradation (ArgG, ArgJ, LeuA, Sat and DhbC) that remain to be further investigated in future studies.

The general stress response mediated by the alternative RNA polymerase sigma factor SigB provides the *B. subtilis* cell with multiple stress resistance strategies in anticipation of future stresses. The entire SigB regulon is induced directly after the transition to the starvation phase (Supplementary Movie S4). Changes on transcript level for the sigB regulon are only reflected by six SigB-dependent general stress proteins on protein level. Besides the general stress marker proteins YvyD and GsiB, the proteases ClpC and ClpP increase in protein amount at the later time points 60 and 120 min after the transition. This is in good correlation with the increased protein aggregation and retrieval of aggregated cytosolic proteins in the EMF (Supplementary Figs S2–S4).

Previous work revealed that a substantial number of membrane proteins is affected by amino-acid depletion or glucose downshift in the membrane of *B. subtilis*[40]. Indeed, we could detect major changes in the membrane proteome, so far only described in transcriptomic studies.

Among the transporters that were diminished in protein amount, the major iron uptake systems revealed highly congruent changes (see Supplementary Fig. S6). The marked decrease imposed on the

ARTICLE

Fur[36] regulon suggests non-iron-limiting conditions. It remains to be elucidated if this is caused by the decreased biosynthetic activity of the cell compared with exponential growth. Remarkably, this notion is in accordance with data obtained for the closely related *Staphylococcus aureus*[22].

The largest group of induced transporter systems belongs to the CcpA regulon that is derepressed on cessation of growth due to the glucose downshift. The Supplementary movies S4 and S2 of the regulon-based Voronoi treemap gives an impression of the complexity of derepression and/or substrate induction of CcpA-controlled operons and exemplifies changes of transporter proteins. Operons within the CcpA regulon are turned on in a sequential manner, resulting in a time-dependent insertion of new transporter proteins into the membrane (for details see Supplementary Figs S7–S14). Accordingly, most operons for catabolism of alternative carbon sources are under the transcriptional control of multiple layers of a regulating network that allows for fine-tuning by temporal dependence and strict order of preference.

The high slopes of increase in protein amount result from the absence of most of the transporter proteins in the exponential growth phase. Taking the normalized spectral abundance factors (NSAFs) for these proteins into account, the final concentration of the respective proteins seems to be low in comparison with other transporters identified in the MSF (illustrated in Supplementary Fig. S15). It becomes clear that *B. subtilis* deploys new transport capacities for alternative carbon and energy sources at the onset of stationary phase, following glucose depletion to allow sensing and utilization of new nutrients.

Interestingly, determination of extracellular metabolites reveals that most of the substrates of the newly installed transporter systems are absent or at least not detectable in our study (Supplementary Fig. S1). A particular example is the induction of the citrate transporter CitM[41]. In contrast to the two other homologous uptake systems for different C4 TCA cycle intermediates, DctP[42] and MaeA[43], no substrate is present in the medium during the entire cultivation and therefore substrate induction of CitM is not likely to take place. On the contrary, the malate transporter MaeA decreases in protein amount as a consequence of malate exhaustion in the growth medium. Furthermore, DctP increases markedly in amount in the stationary phase as its substrates, succinate and fumarate, accumulate in the medium during cultivation and are used as nutrients and building blocks in the starved cell. It remains to be elucidated if the substrates found in the growth medium cause antitermination/induction (degradation products of lysed bacteria or exported metabolites) or if an increase in protein amount is a consequence solely of derepression of the CcpA-mediated carbon catabolite repression.

Similar to the temporal derepression of CcpA-dependent operons, parts of the Spo0A regulon were induced for a short time at the beginning of the stationary phase (Supplementary movie S3). Low-threshold induction of Spo0A genes leads to expression of the cannibalism operons *skf* and *sdp* (Supplementary Fig. S16)[44,45]. The transient expression of genes and accumulation of these two systems—encoding among others, proteins being located at the cell membrane—to either protect the cell and to enhance/generate nutrient availability adds another important aspect of glucose starvation.

Taken together, this work is the first global study that monitors time-dependent changes of integral membrane proteins of *B. subtilis*. Changes in the transport systems of the bacterial cell encountering limited nutrient availability are crucial switches for the adaptation of metabolic pathways and consequently for survival. Despite the short time of derepression/induction of the CcpA-regulated genes, it could be shown that new sets of nutrient uptake systems are inserted into the membrane. Preventive or substrate-induced provision of additional transporter specificity seems to be a major event after the onset of stationary phase following nutrient depletion in *B. subtilis*.

To summarize, this work represents the first comprehensive study on concomitant transcriptome and large-scale proteome dynamics to be followed over five time points for *B. subtilis* coping with physiological perturbations. We monitored temporal and spatial changes on the proteome level covering all subcellular localizations in our global approach on glucose-starved *B. subtilis* to an unparalleled extent. So far, the analysis of soluble proteins of the major metabolic pathways has been well established and a wealth of information is known on *B. subtilis* coping with nutrient stresses.

We were able to add qualitative and quantitative information gathered on the proteome level for the membrane proteome covering the interface of the cell to the environment that was underrepresented in proteomic studies for a long time. In this regard, time dependent insertion of additional substrate transporters into the membrane after glucose starvation could be shown.

The combination of subfractionation techniques with state-of-the-art metabolic labelling and highly accurate mass spectrometry allows a more comprehensive overview of highly dynamic temporal and spatial processes in the model organism *B. subtilis* than was possible before. Most importantly, by concurrent sampling of large-scale proteomic and transcriptomic data, it was possible to reveal rapid alterations at the transcript level reflected at a slower rate of the proteome level, introducing new functionalities needed in stationary phase and degrading specific enzymes in biosynthetic pathways.

Classification of proteins according to changes or stability was carried out and the implications for different physiological situations were shown. Elucidation of protein stability of specific pathways and the underlying mechanism seem to be one of the most promising topics to follow-up. It could be shown that degradative processes involve different subcellular locations that should be involved in future studies on targeted degradation in *B. subtilis* and its relatives.

Systems-wide quantitative proteomic studies are valuable tools to display the actual situation of the "players of life", the proteins, in a biological sample. We assume that due to the experimental setup and the used strain that is widely spread among the Bacillus community, this work with its spectral and raw data available in open repositories will be the base for future targeted approaches in the Bacillus community and beyond.

## Methods

**Bacterial growth and metabolic labelling.** *B. subtilis* wild-type strain 168 (*trpC2*)[46] was grown aerobically at 37 °C in a synthetic minimal medium[47] supplemented with either 15N-ammonium sulphate/15N-L-tryptophan (0.078 mM, 98 atom % excess, Cambridge Isotope Laboratories) or 14N-ammonium sulphate and 14N-L-tryptophan. Glucose and L-malate were added to a final concentration of 0.05% each to induce glucose starvation after growth to an OD$_{500}$ of 1.0. Cells were collected by centrifugation at different time points along the growth curve (exponential growth, transition phase, transition phase + 30, + 60 and + 120 min). Relative quantification based on metabolic labelling of cell extracts was carried out as published by MacCoss *et al.*[20] 15N-labelled cells were prepared in triplicate to obtain three biological replicates (*n* = 3).

**Preparation of mixed metabolically labelled cell extracts.** The washed cells were disrupted mechanically in a French press (Simaminco SLM) and the cell debris was removed via centrifugation (14000 *g* at 4 °C for 20 min).

Subsequently, the protein concentration of the unlabelled 14N-samples and the pooled 15N-samples was determined. The same protein amount of the 14N-unlabelled and labelled 15N-pooled protein extracts were mixed to obtain a metabolically labelled cell extract for further fractionation procedures.

**Fractionation of metabolically labelled cell extracts.** Membrane proteins were enriched according to the protocol published in Eymann *et al.*[5], leaving out the extraction of proteins by *n*-dodecyl-β-D-maltoside treatment. In brief, sedimented cell membranes are subjected to subsequent ultracentrifugation/washing steps, including washing of the pellets with high-salt and carbonate buffer resulting in the EMF that is depleted of cytosolic proteins. The cytosolic fraction represents the supernatant after the first ultracentrifugation step in the membrane-enrichment protocol.

The preparation of the integral membrane peptides is described by Wolff *et al.*[19] Briefly, the bacterial membrane is spun down via ultracentrifugation, subjected to a carbonate washing step and digested with ProteinaseK in urea to deplete all soluble

201

# ARTICLE

loops from the membrane. Transmembrane domains resting in the cell membrane are subsequently digested with Chymotrypsin in a buffer containing a detergent (RapiGest, Waters Corporation) suitable in mass spectrometry-based workflows, resulting in the MSF.

**Preparation of extracellular proteins.** To analyse the extracellular proteome, proteins in the cell medium were acid precipitated according to Antelmann et al.[48] In brief, proteins in the supernatant were precipitated on ice over night by adding 10% TCA to the supernatant that resulted from cell gathering. The precipitate was washed thoroughly with ethanol and the resulting protein pellet was dissolved in urea/thiurea. After determination of protein concentration and pooling of the ^{15}N-labelled extracellular proteins, equal protein amounts of the ^{15}N-samples and ^{15}N-pooled samples were mixed to obtain the metabolically labelled extracellular fractions.

**Preparation of biotinylation-enriched proteins samples.** The biotinylation approach and enrichment of surface proteins were performed following the protocol published by Hempel et al.[49] In brief, labelled (^{15}N) and unlabelled (^{14}N) cells were collected by centrifugation and subsequently biotinylated by Sulpho-NHS-SS-Biotin on ice. After quenching of the reaction and washing, cells were mixed before cell lysis to obtain the metabolically labelled extract after the following cell disruption. Biotinylated proteins were affinity purified with NeutrAvidin agarose. Elution of the purified and washed biotinlyated proteins was carried out by adding reductive SDS sample buffer.

**Sample preparation and proteomics measurement.** Preparation of protein fractions and explanation of the workflow leading to mass spectrometric results is described in more detail in the Supplementary Methods. In brief, samples that were analysed by the GeLC-MS workflow (EMF, cytosolic fraction, extracellular fraction, BEF) were fractionated by 1D SDS Gel electrophoresis followed by tryptic digestion, as described in Dreisbach et al.[46] The membrane-shaving (MSF) samples were prepared according to Wolff et al.[19] The tryptic-digested proteins were subjected to a reversed phase C18 column chromatography operated on a nanoACQUITY-UPLC (Waters Corporation). Peptides were first concentrated and desalted on a trapping column. Mass spectrometry (MS) and MS/MS data were acquired with an online coupled LTQ-Orbitrap mass spectrometer (Thermo Fisher). The digests resulting from the membrane-shaving fraction were subjected to reversed phase C18 column chromatography operated on a nanoACQUITY-UPLC (Waters Corporation) with a one column setup online coupled with an LTQ-Orbitrap mass spectrometer (Thermo Fisher).

Subsequent to sample preparation, samples were subjected to LC-MS/MS analysis as described in Supplementary Methods.

All mass spectrometry raw data are deposited into the PRIDE database with accession numbers of 11382–11458 (http://www.ebi.ac.uk/pride/).

**Data analysis.** The *.dta files extracted from *.raw files using BioworksBrowser 3.3.1 SP1 (Thermo Fisher Scientific) with no charge state deconvolution and deisotoping performed on the data were searched with SEQUEST version v28 (rev.12) (Thermo Fisher Scientific) against a B. subtilis target-decoy protein sequence database (complete proteome set of B. subtilis extracted from UniprotKB release 12.7 (UniProt Consortium, Nucleic acids research 2007, 35, D193-197) with a set of common laboratory contaminants) compiled with BioworksBrowser. The searches were performed in two iterations: First, for the GeLC-MS analyses the following search parameters were used: enzyme type, trypsin (KR); peptide tolerance, 10 p.p.m.; tolerance for fragment ions, 1 a.m.u.; b- and y-ion series; variable modification, methionine (15.99 Da) a maximum of three modifications per peptide was allowed. For the membrane shaving approach, the following search parameters were applied: enzyme type, none; peptide tolerance, 10 p.p.m.; tolerance for fragment ions, 1 a.m.u.; b- and y-ion series; an oxidation of methionine (15.99 Da) and a carboxyamidomethylation (57.02 Da) of cysteine were considered as variable modifications with a maximum of three modifications per peptide. In the second iteration, the mass shift of all amino acids completely labelled with ^{15}N-nitrogen was taken into account in the search parameters.

Resulting *.dta and *.out files were assembled and filtered using DTASelect (version 2.0.25) (parameters GeLC-MS: -y 2 -c 2 -C 4 --here --decoy Reverse_ -p 2 -t 2 -u --MC 2 -i 0.3 --fp 0.005; parameters membrane shaving: --nostats -1 1.9 -2 2.2 -3 3.3 -4 3.75 -i 0.299 -a false -p 1 -y 0 -d 0.1 -t 2 -u --here). The protein false-positive rate was calculated for each analysis according to Peng et al.[50]

**Data analysis relative quantification.** The workflow for relative quantification was carried out according to MacCoss et al.[20] The cured search results served as the base for the further analysis using the software Census[21] to obtain quantitative data of ^{14}N-peaks (sample) and ^{15}N-peaks (pooled reference[20]). The software extracts the respective mass traces of the two isotopologues within a defined scan range centred around the fragment scan, leading to successful identification of the respective peptide. The ratio of the peak intensities is subsequently calculated for all overview scans contained in the chosen peak boundaries. Peptide ratios and combined protein ratios are finally exported ($R^2$-values > 0.7 and only unique peptides; proteins failing to be relatively quantified were checked manually in the graphical

user interface for on/off proteins). Proteins relatively quantified with at least two peptides were taken into account for the subsequent analysis.

All quantification results of a complete GeLC-MS run were median centred and ratios were log₂-transformed and averaged over the biological replicates. Time course data for each protein were adjusted to the first time point as reference point; proteins without quantitative information for the first time point were not taken into account for a detailed discussion.

For a more detailed description on data analysis, including combination of quantification results, normalization and analysis of variance, see Supplementary Methods.

**Normalized spectral abundance factors.** To estimate the relative proportion of transporter proteins within the MSF, the NSAFs were calculated for these proteins according to Zybailov et al.[51] In brief, the NSAF for a protein is the number of spectral counts (SpC; the total number of MS/MS spectra accounting for this specific protein) divided by the length $L$ of the protein and divided by the sum of SpC/L for all proteins in the proteomic experiment. The NSAF is a measure for the relative portion of a protein in a proteomic analysis and proportional to the protein amount in the sample.

**Prediction of protein localization.** Predictions of protein localization in the cell were assigned according to the LocateP algorithm[52].

**Voronoi treemaps.** Algorithms for treemap calculations have been established in JAVA1.6 by using the corresponding JDK and have been formerly described in detail[53].

**Extracellular metabolite sampling and measurement by ^{1}H-NMR.** Detection and determination of concentrations for extracellular metabolites is described in detail in Supplementary Methods.

**Transcriptome analysis.** Generation of transcriptome data is described in detail in Supplementary methods. The complete data set is accessible through GEO Series accession no. GSE24058.

## References

1. Sonenshein, A. L. & Hoch, J.A.R.L. (eds). *Bacillus subtilis and Its Closest Relatives: From Genes to Cells* (American Society for Microbiology Press, 2002).
2. Eiamphungporn, W. & Helmann, J. D. The *Bacillus subtilis* sigma(M) regulon and its contribution to cell envelope stress responses. *Mol. Microbiol.* **67**, 830–848 (2008).
3. Koburger, T., Weibezahn, J., Bernhardt, J., Homuth, G. & Hecker, M. Genome-wide mRNA profiling in glucose starved *Bacillus subtilis* cells. *Mol. Genet. Genomics* **274**, 1–12 (2005).
4. Tjalsma, H. *et al.* Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. *Microbiol. Mol. Biol. Rev.* **68**, 207–233 (2004).
5. Eymann, C. *et al.* A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics* **4**, 2849–2876 (2004).
6. Dauner, M., Storni, T. & Sauer, U. *Bacillus subtilis* metabolism and energetics in carbon-limited and excess-carbon chemostat culture. *J. Bacteriol.* **183**, 7308–7317 (2001).
7. Kleijn, R. J. *et al.* Metabolic fluxes during strong carbon catabolite repression by malate in *Bacillus subtilis*. *J. Biol. Chem.* **285**, 1587–1596 (2010).
8. Anderson, N. L. & Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell Proteomics* **1**, 845–867 (2002).
9. de Godoy, L. M. *et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251–1254 (2008).
10. Hecker, M. & Volker, U. Towards a comprehensive understanding of *Bacillus subtilis* cell physiology by physiological proteomics. *Proteomics* **4**, 3727–3750 (2004).
11. Rabilloud, T. Membrane proteins and proteomics: love is possible, but so difficult. *Electrophoresis* **30**(Suppl 1), S174–S180 (2009).
12. Wolff, S. *et al.* Towards the entire proteome of the model bacterium *Bacillus subtilis* by gel-based and gel-free approaches. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **849**, 129–140 (2007).
13. Hirose, I. *et al.* Proteome analysis of *Bacillus subtilis* extracellular proteins: a two-dimensional protein electrophoretic study. *Microbiology* **146** (Part 1), 65–75 (2000).
14. Antelmann, H., Yamamoto, H., Sekiguchi, J. & Hecker, M. Stabilization of cell wall proteins in *Bacillus subtilis*: a proteomic approach. *Proteomics* **2**, 591–602 (2002).
15. Gatlin, C. L. *et al.* Proteomic profiling of cell envelope-associated proteins from *Staphylococcus aureus*. *Proteomics* **6**, 1530–1549 (2006).
16. Santoni, V., Molloy, M. & Rabilloud, T. Membrane proteins and proteomics: un amour impossible? *Electrophoresis* **21**, 1054–1070 (2000).
17. Bunai, K. *et al.* Proteomic analysis of acrylamide gel separated proteins immobilized on polyvinylidene difluoride membranes following proteolytic digestion in the presence of 80% acetonitrile. *Proteomics* **3**, 1738–1749 (2003).
18. Hahne, H., Wolff, S., Hecker, M. & Becher, D. From complementarity to comprehensiveness—targeting the membrane proteome of growing *Bacillus subtilis* by divergent approaches. *Proteomics* **8**, 4123–4136 (2008).

19. Wolff, S., Hahne, H., Hecker, M. & Becher, D. Complementary analysis of the vegetative membrane proteome of the human pathogen *Staphylococcus aureus*. *Mol. Cell Proteomics* 7, 1460–1468 (2008).

20. MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R. & Yates, J. R. III A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal. Chem.* 75, 6912–6921 (2003).

21. Park, S. K., Venable, J. D., Xu, T. & Yates, J. R. III A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* 5, 319–322 (2008).

22. Becher, D. et al. A proteomic view of an important human pathogen—towards the quantification of the entire *Staphylococcus aureus* proteome. *PLoS One* 4, e8176 (2009).

23. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30 (2000).

24. Kanehisa, M. A database for post-genome analysis. *Trends Genet.* 13, 375–376 (1997).

25. Lammers, C. R. et al. Connecting parts with processes: SubtiWiki and SubtiPathways integrate gene and pathway annotation for *Bacillus subtilis*. *Microbiology* 156, 849–859 (2010).

26. Florez, L. A., Roppel, S. F., Schmeisky, A. G., Lammers, C. R. & Stulke, J. A community-curated consensual annotation that is continuously updated: the *Bacillus subtilis* centred wiki SubtiWiki. *Database (Oxford)* 2009, bap012 (2009).

27. Shneiderman, B. Tree visualization with tree-maps: 2-d space-filling approach. *Transactions on Graphics* 11, 92–99 (1992).

28. Balzer, M. & Deussen, O. Voronoi Treemaps in *IEEE Symposium on Information Visualization (InfoVis 2005)* 7 (2005).

29. Bernhardt, J., Funke, S., Hecker, M. & Siebourg, J. Visualizing gene expression data via Voronoi Treemaps in *Sixth International Symposium on Voronoi Diagrams* 233–241 (2009).

30. Gerth, U. et al. Clp-dependent proteolysis down-regulates central metabolic pathways in glucose-starved *Bacillus subtilis*. *J. Bacteriol.* 190, 321–331 (2008).

31. Hahne, H. et al. A comprehensive proteomics and transcriptomics analysis of *Bacillus subtilis* salt stress adaptation. *J. Bacteriol.* (2009).

32. Soufi, B. et al. Stable isotope labeling by amino acids in cell culture (SILAC) applied to quantitative proteomics of *Bacillus subtilis*. *J. Proteome Res.* 9, 3638–3646 (2010).

33. Michalik, S. et al. Proteolysis during long-term glucose starvation in *Staphylococcus aureus* COL. *Proteomics* 9, 4468–4477 (2009).

34. Fournier, M. L. et al. Delayed correlation of mRNA and protein expression in Rapamycin-treated cells and a role for Ggc1 in cellular sensitivity to Rapamycin. *Mol. Cell. Proteomics* 9, 271–284 (2010).

35. Bernhardt, J., Weibezahn, J., Scharf, C. & Hecker, M. *Bacillus subtilis* during feast and famine: visualization of the overall regulation of protein synthesis during glucose starvation by proteome analysis. *Genome Res.* 13, 224–237 (2003).

36. Sonenshein, A. L. Control of key metabolic intersections in *Bacillus subtilis*. *Nat. Rev. Microbiol.* 5, 917–927 (2007).

37. Jin, S. & Sonenshein, A. L. Transcriptional regulation of *Bacillus subtilis* citrate synthase genes. *J. Bacteriol.* 176, 4680–4690 (1994).

38. Eymann, C., Homuth, G., Scharf, C. & Hecker, M. *Bacillus subtilis* functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. *J. Bacteriol.* 184, 2500–2520 (2002).

39. Tissieres, A. & Watson, J. D. Ribonucleoprotein particles from *Escherichia coli*. *Nature* 182, 778–780 (1958).

40. Dreisbach, A. et al. Monitoring of changes in the membrane proteome during stationary phase adaptation of *Bacillus subtilis* using *in vivo* labeling techniques. *Proteomics* 8, 2062–2076 (2008).

41. Yamamoto, H., Murata, M. & Sekiguchi, J. The CitST two-component system regulates the expression of the Mg-citrate transporter in *Bacillus subtilis*. *Mol. Microbiol.* 37, 898–912 (2000).

42. Asai, K., Baik, S. H., Kasahara, Y., Moriya, S. & Ogasawara, N. Regulation of the transport system for C4-dicarboxylic acids in *Bacillus subtilis*. *Microbiology* 146 (Pt 2), 263–271 (2000).

43. Doan, T. et al. The *Bacillus subtilis* ywkA gene encodes a malic enzyme and its transcription is activated by the YufL/YufM two-component system in response to malate. *Microbiology* 149, 2331–2343 (2003).

44. Fujita, M., Gonzalez-Pastor, J. E. & Losick, R. High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. *J. Bacteriol.* 187, 1357–1368 (2005).

45. Gonzalez-Pastor, J. E., Hobbs, E. C. & Losick, R. Cannibalism by sporulating bacteria. *Science* 301, 510–513 (2003)(AAAS, 2003).

46. Anagnostopoulos, C. & Spizizen, J. Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* 81, 741–746 (1961).

47. Stulke, J., Hanschke, R. & Hecker, M. Temporal activation of beta-glucanase synthesis in *Bacillus subtilis* is mediated by the GTP pool. *J. Gen. Microbiol.* 139, 2041–2045 (1993).

48. Antelmann, H. et al. A proteomic view on genome-based signal peptide predictions. *Genome Res.* 11, 1484–1502 (2001).

49. Hempel, K. et al. Quantitative cell surface proteome profiling for SigB-dependent protein expression in the human pathogen *Staphylococcus aureus* via biotinylation approach. *J. Proteome Res.* 9, 1579–1590 (2010).

50. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. & Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* 2, 43–50 (2003).

51. Zybailov, B. et al. Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J. Proteome Res.* 5, 2339–2347 (2006).

52. Zhou, M., Boekhorst, J., Francke, C. & Siezen, R. J. LocateP: genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinformatics* 9, 173 (2008).

## Acknowledgments

## Author contributions

A.O., D.B., M.H. conceived and designed the experiments; A.O., F.H., M.S. performed the experiments; A.O., D.B., J.B., H.M., U.M. and M.L. analysed the data; J.S. contributed analysis tools; A.O., M.H. and D.B wrote the paper.

## Additional information

_____

# *Curriculum Vitae*

| **Personal Information:** | Andreas Willy Otto |
|---|---|
| | Born: 21.06.1979 in Kiel |

**Education:**

| 07/86 - 06/89 | Grundschule am Sonderburger Platz, Kiel |
|---|---|
| 08/89 - 07/98 | Ernst Barlach Gymnasium, Kiel |
| | Abitur |

**Military service:**

| 10/98 – 10/99 | German Naval Forces |
|---|---|

**Academic studies:**

| 10/99- 09/04 | Study of Biochemistry |
|---|---|
| | Ernst- Moritz- Arndt- Universität Greifswald |
| 10/03 – 09/04 | Diploma thesis: „Untersuchungen zum Interaktom von *Bacillus subtilis*" |
| | Institut für Mikrobenphysiologie und Molekularbiologie, Prof. D. M. Hecker |

**Ph.D. thesis:**

| from 10/04 | Scientific employee at the E.- M.- Arndt- Universität Greifswald |
|---|---|
| | Ph.D. thesis at the Institut für Mikrobiologie und Molekularbiologie |

---

_____

## *Danksagung*

Meinem Doktorvater, Herrn Prof. Hecker sei herzlich gedankt für die Möglichkeit, diese Arbeit anzufertigen. Von ihm wurde bei mir das Interesse an Physiologie und zukunftsgewandten Technologien der Lebenswissenschaften in den Vorlesungen geweckt, im Studium und in der Diplomarbeit Vertrauen gegeben, Ehrgeiz geweckt und mir die Chance gegeben als Doktorand unter seiner Anleitung zu arbeiten.

Die Betreuung durch Dr. Dörte Becher sei nicht minder zu würdigen – es war ein langer Weg hin zu dieser Arbeit, gesäumt von langen, intensiven Diskussionen und abendfüllenden Sitzungen über dem Rechner mit schwerwiegenden Entscheidungen - aber auch Freude über exzellente Ergebnisse und Erfolge, die maßgeblich zur Motivation beigetragen haben. Ich schätze (inzwischen) alle Zwischenberichte zu Quantpro, alle Vorträge auf den Projektsitzungen wie z.B. Euroscope mit den dazugehörigen Dienstreisen und kleinen Aufgaben (Andreas, kannst Du mal kurz gucken…) als Teil der Zeit, die mit in den letzten Jahren viel Freude bereitet haben.

Allen Mitarbeitern der Massenspektrometrie möchte ich danken für die Zeit, die von erfolgsorientiertem Arbeiten bis hin zu kreativer Freizeitgestaltung mit einzigartigen Charakteren alles geboten haben.

Naturgemäß kann man nicht alle und jeden in einer solchen Danksagung nennen – gerade im Bezug dessen, dass diese Arbeit ohne die bestehenden Kollaborationen nie möglich gewesen wäre. Dennoch danke ich allen Koautoren, Helfern und Mitarbeitern des AK Hecker und aller anderen Arbeitsgruppen für die Zeit in Greifswald.

Meiner Familie gebührt Dank für die Unterstützung während meiner Ausbildung – sei es im Studium oder auch in der Promotionszeit. Meinen Eltern möchte ich dementsprechend diese Arbeit widmen.

Nicht zu vergessen sei Patricia –Sie hat mich in einer schwierigen Zeit der Promotion kennengelernt und alle Höhen und Tiefen – sehr komprimiert- mitbekommen.

_____

*Eidesstattliche Erklärung*

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Ernst-Moritz-Arndt-Universität Greifswald noch an einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die darin angegebenen Hilfsmittel verwendet habe.