

Research article

Open Access

Phylogenetic distribution and membrane topology of the LytR-CpsA-Psr protein family

Judith Hübscher, Lucas Lüthy, Brigitte Berger-Bächi and Patricia Stutzmann Meier*

Address: Institute of Medical Microbiology, University of Zürich, Zürich, Switzerland

Email: Judith Hübscher - jhuebsch@imm.uzh.ch; Lucas Lüthy - Lucas.Luethy@tecan.com; Brigitte Berger-Bächi - bberger@imm.uzh.ch; Patricia Stutzmann Meier* - pstutzmann@imm.uzh.ch

* Corresponding author

Published: 19 December 2008

Received: 25 September 2008

BMC Genomics 2008, 9:617 doi:10.1186/1471-2164-9-617

Accepted: 19 December 2008

This article is available from: <http://www.biomedcentral.com/1471-2164/9/617>

© 2008 Hübscher et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The bacterial cell wall is the target of many antibiotics and cell envelope constituents are critical to host-pathogen interactions. To combat resistance development and virulence, a detailed knowledge of the individual factors involved is essential. Members of the LytR-CpsA-Psr family of cell envelope-associated attenuators are relevant for β -lactam resistance, biofilm formation, and stress tolerance, and they are suggested to play a role in cell wall maintenance. However, their precise function is still unknown. This study addresses the occurrence as well as sequence-based characteristics of the LytR-CpsA-Psr proteins.

Results: A comprehensive list of LytR-CpsA-Psr proteins was established, and their phylogenetic distribution and clustering into subgroups was determined. LytR-CpsA-Psr proteins were present in all Gram-positive organisms, except for the cell wall-deficient *Mollicutes* and one strain of the *Clostridiales*. In contrast, the majority of Gram-negatives did not contain LytR-CpsA-Psr family members. Despite high sequence divergence, the LytR-CpsA-Psr domains of different subclusters shared a highly similar, predicted mixed α/β -structure, and conserved charged residues. PhoA fusion experiments, using MsrR of *Staphylococcus aureus*, confirmed membrane topology predictions and extracellular location of its LytR-CpsA-Psr domain.

Conclusion: The LytR-CpsA-Psr domain is unique to bacteria. The presence of diverse subgroups within the LytR-CpsA-Psr family might indicate functional differences, and could explain variations in phenotypes of respective mutants reported. The identified conserved structural elements and amino acids are likely to be important for the function of the domain and will help to guide future studies of the LytR-CpsA-Psr proteins.

Background

The cell envelope forms a protective shield around bacteria and is also the site of primary host-pathogen interactions. Its composition and surface characteristics are

therefore important in pathogenesis, and may, in view of increasing resistance against all commonly used cell wall-directed antibiotics, present novel potential antibacterial targets.

The LytR-CpsA-Psr family of cell envelope-associated transcriptional attenuators has been brought into focus of scientific interest upon the discovery that members of this family influence various virulence factors as well as antibiotic resistance of important human pathogens and, interestingly, seem to play a role in bacterial cell envelope maintenance [1-5]. Therefore, this protein family represents a promising target to gain more insight into virulence and antibiotic resistance development.

The LytR-CpsA-Psr family members are putative transmembrane proteins carrying a so-called LytR-CpsA-Psr domain, which is predicted to be extracellular. LytR was first described in *Bacillus subtilis*, where it acts as an attenuator of the expression of both itself and the divergently transcribed *lytABC* operon, which encodes a putative lipoprotein (LytA), an *N*-acetylmuramoyl-L-alanine amidase (LytC) and its modifier LytB [3]. CpsA is suggested to play a role in transcription activation of the capsular polysaccharide synthesis operon of *Streptococcus agalactiae* [2]. Psr was initially proposed to be a repressor of penicillin-binding protein 5 (PBP5) synthesis in *Enterococcus hirae* [6], however, Sapunari et al. could neither confirm an effect on PBP5 synthesis nor autolysis nor β -lactam resistance [7]. In contrast, the LytR homolog BrpA affects autolytic activity of *Streptococcus mutans* [1] and positively influences biofilm formation ability as well as acidic and oxidative stress tolerance [5]. Furthermore, BrpA inactivation alters phagocytosis by human polymorphonuclear leukocytes and the outcome of bacteremia in a rat model [8]. In *Staphylococcus aureus*, the LytR-CpsA-Psr member MsrR contributes to β -lactam resistance [4]. MsrR of *S. aureus* as well as Psr of *Enterococcus faecalis* were both shown to increase virulence in the model host *Caenorhabditis elegans* [9,10]. Furthermore, expression of SA0908, another member of the LytR-CpsA-Psr family in *S. aureus*, is increased during infection in a murine renal abscess model [11].

The function of the LytR-CpsA-Psr domain, however, is still unknown and so far information about these proteins is based on phenotypic characterizations. A more comprehensive knowledge about the occurrence of LytR-CpsA-Psr proteins and analyses of their sequence will provide a basis for experimental determination of interactions or structure/function relationships.

In the present study, we investigated the phylogenetic distribution of the LytR-CpsA-Psr proteins and analyzed secondary structure predictions. In addition, the staphylococcal LytR-CpsA-Psr protein MsrR was used as a model to confirm membrane topology.

Results and discussion

Sequence collection of LytR-CpsA-Psr proteins

A comprehensive sequence collection was obtained by searching the InterPro database [12] for LytR-CpsA-Psr

members (InterPro entry IPR004474) and using PSI-BLAST [13] to identify further homologous sequences as described in the Methods section. Sequences representing fragments of the LytR-CpsA-Psr domain were excluded from the studies. The revised dataset comprised 1'079 sequences.

Occurrence of the LytR-CpsA-Psr domain

Members of the LytR-CpsA-Psr family were found in the eight bacterial phyla *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Cyanobacteria*, *Deinococcus-Thermus*, *Firmicutes*, *Spirochaetes*, and *Thermotogae* (Figure 1). In addition, a LytR-CpsA-Psr fragment was identified in *Plesiocystis pacifica* strain SIR-1, which is a member of the δ -*Proteobacteria*.

The LytR-CpsA-Psr domain is restricted to the kingdom of bacteria with two exceptions occurring in the moss *Physcomitrella patens* subsp. *patens* (*Bryophyta*, Moss Superclass V) and the freshwater amoeba *Paulinella chromatophora*. *Physcomitrella* is widely used as a model organism in plant genetics as it undergoes integrative homologous recombination with a high efficiency [14], while *Paulinella* contains photosynthetic inclusions (chromatophores) affiliated with the cyanobacteria *Prochlorococcus* and *Synechococcus* spp. [15]. BLAST analysis of the LytR-CpsA-Psr sequences found in *Paulinella* and *Physcomitrella* revealed high similarities to those found in *Synechococcus* sp. WH5701 and *Paenibacillus* sp. JDR-2, respectively (data not shown). It is therefore conceivable that the LytR-CpsA-Psr proteins found in *Physcomitrella* and *Paulinella* are of bacterial origin.

Analysis of the genomes available at the National Center for Biotechnology Information (NCBI) [16] revealed that all of the completely sequenced *Firmicutes* contain at least one member of the LytR-CpsA-Psr family, except for *Clostridium kluyveri* strain DSM 555 and the whole order of *Mollicutes*. In the phylum of *Actinobacteria*, LytR-CpsA-Psr proteins were found in every completely sequenced genome and the same was true for *Deinococcus-Thermus*, *Spirochaetes*, and *Thermotogae*. In contrast, only two strains of the *Bacteroidetes*, i.e. *Pedobacter* sp. BAL39 (*Shingobacteriales*) and *Bacteroides capillosus* ATCC 29799 (*Bacteroidales*) harbor LytR-CpsA-Psr proteins, but none of the completely sequenced strains. Among the *Chloroflexi*, the *Dehalococcoides* do not possess LytR-CpsA-Psr members, whereas they were found in the orders *Chloroflexales* and *Herpetosiphonales*. Most of the *Cyanobacteria* genomes contain LytR-CpsA-Psr members, but only two out of twelve *Prochlorococcus marinus* genomes.

Characteristics of LytR-CpsA-Psr carriers

The organisms having LytR-CpsA-Psr proteins represent the immense diversity of bacteria with regard to morphology, ultrastructure, motility, metabolic characteristics, and habitat preferences. LytR-CpsA-Psr family members were

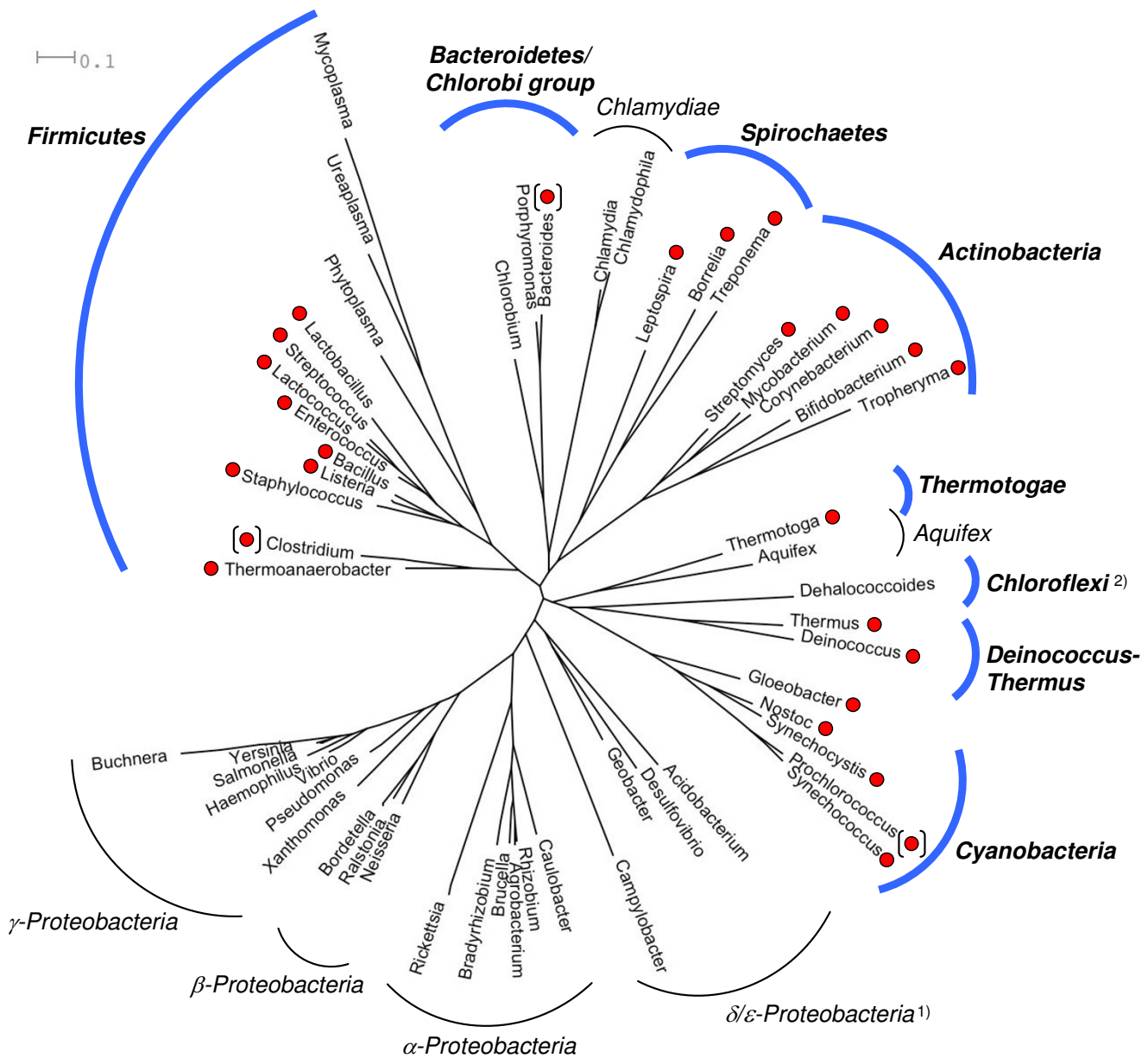


Figure 1
Overview of the occurrence of LytR-CpsA-Psr family proteins in bacteria. The phylogenetic tree is based on the interactive tree of life (iTOL) [65], which was pruned to a selection of genera representing the different phyla. The eight phyla, where LytR-CpsA-Psr family proteins were found, are printed in bold, and red circles indicate their presence in the respective genera. In case not all members of a certain genus harbor LytR-CpsA-Psr proteins, the circle was put in brackets. For a comprehensive description of the occurrence of LytR-CpsA-Psr domain refer to the text. ¹⁾A single fragment of a LytR-CpsA-Psr protein was detected in the *delta-Proteobacterium Plesiocystis pacifica* strain SIR-I. ²⁾Except for the *Dehalococcoides*, which are representing the phylum of the *Chloroflexi* in this tree, all members of the *Chloroflexi* contain LytR-CpsA-Psr family members.

found in pathogens of human, animals, or plants as well as in non-pathogenic organisms, in generalists and specialists such as hyperthermophilic bacteria. This heterogeneity raises the question: Is there a common trait to the LytR-CpsA-Psr carriers? One of the most striking observations was that the LytR-CpsA-Psr proteins are absent in the

large Gram-negative phylum of *Proteobacteria* and many other Gram-negative groups, while all of the Gram-positive bacteria contain at least one copy. Given the proposed role of the LytR-CpsA-Psr proteins in cell envelope maintenance and the lack of the respective domain in the cell wall-deficient order *Mollicutes*, this finding might indicate

a correlation with the Gram-positive cell wall in particular. However, the list of LytR-CpsA-Psr carriers also comprised *Spirochaetes* and numerous organisms classified as Gram-negatives, such as members of the *Cyanobacteria*, *Thermus*, *Thermotogae*, *Bacteroidetes*, and *Chloroflexi*.

Nevertheless, many of the Gram-negative LytR-CpsA-Psr carriers show features of a typical Gram-positive cell wall, such as a thick, multilayered peptidoglycan, a high cross-linking extent, or polysaccharides cross-linked to the peptidoglycan. Intermediate types between a typical Gram-negative and Gram-positive cell wall are for example observed in various *Cyanobacteria* [17], *Chloroflexus aurantiacus* [18], and *Deinococcus radiodurans* [19]. In conclusion, however, a common feature related to peptidoglycan composition or amount was not apparent.

Loss and gain of LytR-CpsA-Psr proteins in the course of evolution

The LytR-CpsA-Psr proteins described so far are not essential, and viable knock out mutants were obtained in *S. aureus*, *S. mutans*, and *E. hirae* [1,4,7]. These particular organisms, however, contain more than one member of the LytR-CpsA-Psr family. It would be interesting to know if the inactivation of all LytR-CpsA-Psr proteins affects viability. On the other hand, our sequence collection showed that this protein family was lost in certain organisms in the course of evolution. As mentioned above, only two *Prochlorococcus* strains carry LytR-CpsA-Psr members. The respective strains, MIT9303 and MIT9313, are distinct from other *Prochlorococcus* isolates in that the size and GC content of their genomes are more similar to those of the genus *Synechococcus* [20]. The differences in genome size reflect gene gain and loss during the evolution of *Prochlorococcus*, which lost the LytR-CpsA-Psr proteins after divergence from the MIT9303/MIT9313 clade. Interestingly, besides the LytR-CpsA-Psr members, numerous genes that are present in MIT9303/MIT9313 and most *Synechococcus*, but absent in the other completely sequenced *Prochlorococcus*, are involved in cell envelope biogenesis [20]. Apart from *Prochlorococcus*, *C. kluveri* is likely to be another example for the loss of LytR-CpsA-Psr proteins. *C. kluveri* differs from other members of the genus *Clostridium* by unique metabolic features and its genome shows only very low synteny on protein level to other *Clostridia* [21].

The occurrence of the LytR-CpsA-Psr family in the *Thermotogae*, which are considered an evolutionary early branch of bacteria [22], might indicate that this family already appeared in a common procaryotic ancestor and was lost by some lineages during the process of diversification, e.g. in the *Proteobacteria*. From this point of view, the incidence in *P. pacifica* SIR-1 could be the result of lateral gene transfer. However, the GC content of the LytR-CpsA-Psr protein encoding gene and the GC content of *P. pacifica* SIR-1 do not differ and a more thorough investigation

would be required to address this question. Yet, lateral gene transfer plays a significant role in evolution [23], and it is possible that LytR-CpsA-Psr proteins were (re-) acquired in the time course of procaryotic diversification.

Cluster analysis

Our sequence collection revealed that from one, e.g. in *C. tetani*, to eleven, e.g. in *Streptomyces coelicolor*, any number of LytR-CpsA-Psr members could be found in a single species. In order to investigate if there existed sub-families and a particular distribution, we performed a cluster analysis based on pair-wise sequence similarities resulting from a BLAST all against all search using CLANS [24] with a P-value cut off of 10^{-35} .

The results mainly reflected taxonomic relationships. Accordingly, the clusters were termed using the first character(s) of the respective phylum followed by consecutive numbers in case of multiple clusters. The majority of sequences grouped in six clusters, i.e. A1 (*Actinobacteria* cluster 1), Ch (*Chloroflexi* cluster), F1–F3 (*Firmicutes* clusters 1 to 3), and M (taxonomically mixed cluster) (Figure 2 and Additional file 1). Besides these main clusters, seven 'peripheric' clusters containing four to 18 proteins were observed: A2, A3, Cy (*Cyanobacteria* cluster), F4, F5, S (*Spirochaetes* cluster), and T (*Thermotogae* cluster). In addition, five sequences of the *Deinococcus-Thermus* group that were found in close proximity to each other were considered an additional cluster D. A small number of proteins did not cluster with any other of the LytR-CpsA-Psr members.

A step-wise increase in the P-value cut off caused the clusters A1, Ch, F1, F2, and M to merge in a supercluster (data not shown), suggesting a close relationship of the enclosed sequences. Decreasing the P-value cut off resulted in disconnection of the clusters into subclusters mainly following phylogenetic affiliations (data not shown). To verify whether unequal family sizes or redundant sequences distorted cluster analysis, CLANS was subsequently applied to a reduced dataset of 576 sequences with a maximum similarity of 90%. The use of this reduced dataset had no significant effect on the outcome of the clustering (data not shown) and the following discussion is based on the results obtained using the unfiltered list of LytR-CpsA-Psr proteins.

Cluster M – taxonomically mixed central cluster

Cluster M was the only one containing proteins of organisms belonging to various phyla, i.e. *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, *Deinococcus-Thermus*, and *Cyanobacteria*, and might represent those sequences that have diverged the least from a presumed ancestral LytR-CpsA-Psr protein. A clearly separated subcluster (M_a) contained exclusively sequences of *Prochlorococcus* and *Synechococcus* species as well as of the amoeba *Paulinella*.

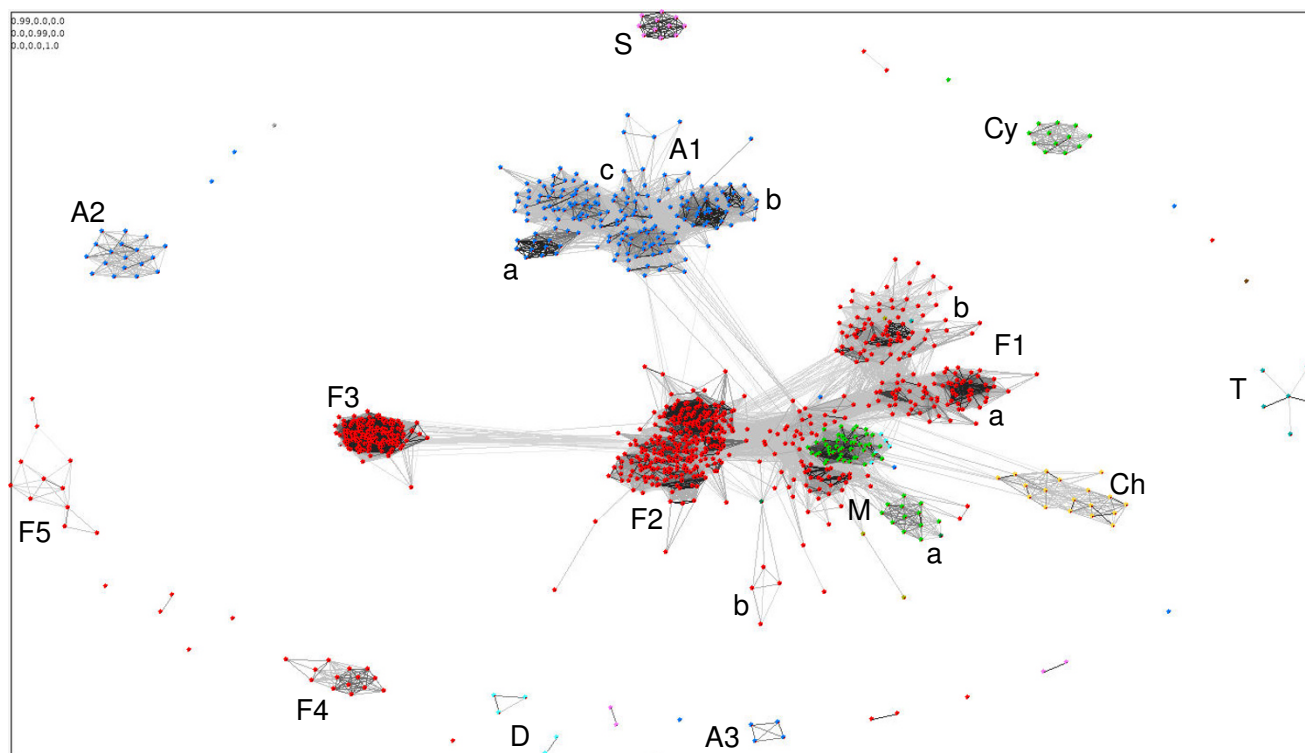


Figure 2

Two-dimensional CLANS clustering of the LytR-CpsA-Psr proteins. Proteins are represented by dots and colored according to their taxonomic relationship: blue, *Actinobacteria*; olive, *Bacteroidetes*; green, *Cyanobacteria*; yellow, *Chloroflexi*; turquoise, *Deinococcus-Thermus*; red, *Firmicutes*; dark green, *Physcomitrella*; brown, *Plesiocystis*; pink, *Spirochaetes*; petrol, *Thermotoga*. BLAST sequence similarities are indicated by lines shaded from light grey (P-values < 10^{-35}) to black (P-values < 10^{-200}). Clusters are indicated using the first character(s) of the corresponding phylum and are numbered in case there is more than one cluster of the same phylum. M denotes the taxonomically mixed 'central' cluster. Subclusters within A1, F1, and M are marked with small letters.

Cluster A1 – common *Actinobacteria* sequences

Most LytR-CpsA-Psr proteins of the high GC Gram-positive organisms grouped in cluster A1. A1-type sequences were identified in each member of the *Actinobacteria* and therefore represent the sequences characteristic for this phylum. The proteins of this cluster were predominantly predicted to have a single transmembrane segment with an extracellular LytR-CpsA-Psr domain, a trait attributed to the majority of LytR-CpsA-Psr proteins.

Within cluster A1, three separated subclusters were distinguished. Subcluster A1_a enclosed solely sequences of different *Mycobacterium* species, whereas A1_b and A1_c both covered sequences of various *Actinobacteria*.

Cluster A1 also harboured FrnA [UniProt:O68907], which contains an SBP_bac5 domain (bacterial extracellular solute-binding protein family 5 domain; PFAM:PF00496). The family 5 solute-binding proteins comprise peptide-binding proteins, such as the oligopeptide transporter

subunit OppA of Gram-negative bacteria. FrnA was exclusively found in *Streptomyces roseofulvus*, and is encoded by the first gene of a 25 kb-region, containing genes involved in biosynthesis of the polyketide antibiotic frenolicin [25]. In *Saccharopolyspora erythraea*, two A1-type proteins, SACE_6482 and SACE_6483, are encoded adjacent to *aveBII* (SACE_6480), a dTDP-glucose 4,6-dehydratase involved in the biosynthesis of a polyketide sugar unit precursor and hence of the macrolide antibiotic erythromycin A [26]. However, any involvement of LytR-CpsA-Psr proteins in biosynthesis of secondary metabolites by organisms of the order *Actinomycetales* remains open.

Cluster A2 – *Actinobacteria* proteins with multiple transmembrane segments

A few LytR-CpsA-Psr family members of the *Actinobacteria* clustered in A2. In contrast to the A1-type proteins and the LytR-CpsA-Psr family members in general, those proteins found in A2 were calculated to cross the membrane two to four times. In accordance with predictions for LytR-CpsA-

Psr members with one transmembrane helix and independently of the number of predicted transmembrane helices, the LytR-CpsA-Psr domains of A2-type proteins were mainly suggested to be extracellular by the various programs used. A2-type proteins were only present in some organisms of the *Propionibacterineae*, *Streptosporangineae*, and *Micrococccineae*.

Cluster A3 – *Streptomyces*-specific proteins

This cluster contained four sequences, which are rich in glycine and the polar amino acid glutamine predominantly located within a stretched, uncharged segment close to the N-terminus. The A3-type proteins also differ from other LytR-CpsA-Psr members by their exceptional length ranging from 559 to 617 aa and were only found in *Streptomyces*.

Cluster F1 – *Psr/MsrR*-like proteins of *Firmicutes*

Apart from those sequences comprised in cluster M, the *Firmicutes* sequences divided into three main clusters F1–F3 and two peripheric clusters F4 and F5. Cluster F1 contained proteins of the orders *Bacillales* and *Lactobacillales* (subcluster F1_a) as well as of *Clostridiales* and the two non-*Firmicutes* *Pedobacter* sp. (phylum *Bacteroidetes*) and *Fervidobacterium nodosum* (phylum *Thermotogae*) (subcluster F1_b).

F1-type proteins included Psr of enterococci, which is suggested to be an activator of its own transcription, especially in the presence of ampicillin [27]. In the food-borne pathogen *Listeria monocytogenes*, transcription of *psr* is increased in response to growth at 10°C [28]. Cluster F1 also included MsrR of *Staphylococcus aureus*, which is involved in β -lactam resistance [4]. *msrR* belongs to the cell wall stress stimulon, a set of genes collectively induced upon exposure to antibiotics damaging the cell wall [29,30].

Cluster F2 – *LytR/BrpA*-like proteins of *Firmicutes*

Cluster F2 mainly contained sequences of the orders *Bacillales* and *Lactobacillales* and included LytR of *B. subtilis* and BrpA of *S. mutans*. BrpA has an impact on autolysis, biofilm formation, stress tolerance, and virulence [1,5], and LytR acts as an attenuator of the *lytABC* operon encoding the major autolytic amidase of *B. subtilis* [3]. LytR belongs to the σ^X regulon, which controls genes participating in cell envelope metabolism and modulation [31,32]. YwtF, a further F2-type protein in *B. subtilis*, is a member of the σ^M regulon. The σ^M regulon plays a role in response to cell envelope stress, and in contrast to *lytR*, *ywtF* gene expression is upregulated by bacitracin and vancomycin [33–35]. A similar situation as in *B. subtilis* was observed for transcription of *lytR* and *ywtF* in *B. licheniformis* [36]. The two F2-type LytR-CpsA-Psr members of *S. aureus*, SA0908 and

SA2103 in strain N315, both belong to the cell wall stress stimulon as does *msrR* of cluster F1 [30,37,38].

In *Bacillus thuringiensis* subsp. *thuringiensis* strain 407-1, the F2-type LytR-CpsA-Psr protein EcfY is thought to be encoded in an operon with *sigW* (similar to extracytoplasmic function σ factor) and *ecfX* (putative anti- σ factor). *sigW* and *ecfX* contribute to β -exotoxin production, and *ecfY* was proposed to be involved in negative control of *sigW* expression [39].

In cluster F2, a LytR-CpsA-Psr protein with a type 2 phosphatidic acid phosphatase superfamily domain (PAP2, PFAM:PF01569) was contained, which is encoded by *E. faecalis* EF_3245. Certain PAP2 family members have undecaprenyl pyrophosphate phosphatase activity, and thus may play a role in peptidoglycan biosynthesis by providing the lipid carrier undecaprenyl phosphate [40,41]. Moreover, the LytR-CpsA-Psr protein of the eukaryote *Physcomitrella* also clustered in F2.

The staphylococcal protein LytR (SA0251 in strain N315), the response regulator of the two-component system LytSR [42], does not belong to the LytR-CpsA-Psr protein family despite its name.

Cluster F3 – *CpsA*-like *Firmicutes* proteins with multiple transmembrane segments

The third *Firmicutes* cluster was formed by the CpsA-like proteins, which are predominantly found in the genus *Streptococcus*. In streptococci, *cpsA* (also *cpsX*, *epsA*, *wzg*) is the first gene of the capsular biosynthesis locus, and CpsA is assumed to be a transcriptional regulator of capsule production in *S. agalactiae* [2]. In *S. pneumoniae*, however, although encapsulation is reduced in a *cpsA* mutant, no evidence for an effect of CpsA on transcription of the *cps* operon was found [43]. Yet, the mutant exhibited reduced tyrosine phosphorylation of CpsD, which is thought to be required for production of elevated amounts of capsule [43].

All F3-type proteins were predicted to have three transmembrane helices with an intracellular N-terminus. Additionally, CpsA and related proteins of streptococci possess a DNA polymerase processivity factor domain (DNA-PPF, PFAM:PF02916), a subdomain of the replisome sliding clamp unit. However, the subunits of the sliding clamp ring are composed of two homologous subdomains, the DNA-PPF domain and a gp45-slide C subdomain (PFAM:PF09116) [44], which is absent in CpsA. Moreover, in CpsA the DNA-PPF domain was predicted to span the third putative membrane segment with the major portion of the domain located extracellularly. These findings imply that the DNA-PPF domain found in CpsA proteins fulfils a different function. The extracellular portion of the

CpsA DNA-PPF domain overlaps with a structural periplasmic binding protein-like II domain (SCOP superfamily 53850) as identified using SCOP (Structural Classification Of Proteins) [45]. This superfamily includes various binding proteins such as phosphate-binding proteins, glutamine-binding proteins, or ferric-binding proteins. Interestingly, SCOP superfamily 53850 also covers the family 5 solute-binding proteins discussed with respect to the A1-type LytR-CpsA-Psr protein FrnA.

A number of proteins of the *Clostridiales* also clustered in F3. Although these proteins did not produce a significant hit to the DNA-PPF domain in a PFAM search, they were of similar length as CpsA (456–559 amino acids), and four of them [UniProt:A5KKW2, A8RE01, A8R906, and B1C5F5] were also predicted to carry a periplasmic binding protein-like II domain.

Clusters F4 and F5 – small 'peripheral' Firmicutes subclusters

Cluster F4 only contained sequences of *Lactococcus lactis* and *Streptococcus thermophilus*, whereas F5 only contained sequences of certain *Clostridiales*. The F4- and F5-type sequences showed the same transmembrane organization, and family domain architecture as those of the *Firmicutes* main clusters F1 and F2.

Clusters Ch, Cy, D, S and T – other taxonomy-specific clusters

The sequences of the *Chloroflexi* grouped in one single cluster (Ch) closely related to M, and all LytR-CpsA-Psr members of *Thermotoga* and one of the two members found in *F. nodosum* assembled in the *Thermotogae* cluster T. Cluster D comprised those *Deinococcus-Thermus* sequences not included in cluster M. Cluster Cy enclosed a set of *Synechococcus* and *Prochlorococcus* sequences, and finally the LytR-CpsA-Psr members of the *Spirochaetes Borrelia* were grouped in cluster S. The sequences of the other *Spirochaetes*, i.e. *Leptospira* and *Treponema*, were not included in any cluster.

All sequences enclosed in these clusters showed for the LytR-CpsA-Psr proteins a characteristic organization with a short cytoplasmic domain, one transmembrane segment, and a putative extracellular domain carrying the LytR-CpsA-Psr element.

Phylogenetic distribution

For a more detailed analysis of the phylogenetic distribution of the LytR-CpsA-Psr proteins, 30 fully sequenced strains representing the six phyla *Actinobacteria*, *Cyanobacteria*, *Chloroflexi*, *Deinococcus-Thermus*, *Firmicutes*, and *Thermotogae* were selected (Additional file 2). Figure 3 summarizes the number of LytR-CpsA-Psr proteins detected in these strains and the results of the CLANS clustering related to the species tree, which was constructed based on the 16S rRNA gene sequences.

Comparison of the genome sizes and the number of LytR-CpsA-Psr members in the respective organisms revealed a positive correlation (Figure 4). There were, however, exceptions like *Nostoc* sp. PCC 7120 with a genome size of 7.2 Mb but only two LytR-CpsA-Psr proteins.

The accumulation of multiple LytR-CpsA-Psr proteins most probably often happened through duplication from an ancestral species- or group-specific sequence, leading to as many as 11 members in *S. coelicolor*. Ten of the *S. coelicolor* LytR-CpsA-Psr proteins clustered in A1, while the eleventh [UniProt:Q9RDK8] clustered in A3. In contrast to Q9RDK8, all genes of the ten A1-type proteins lie within the central core region of the *S. coelicolor* genome, and seven of the LytR-CpsA-Psr genes are consecutively encoded. The core region of *S. coelicolor* comprises genes likely to be essential, such as genes involved in cell division, DNA replication, transcription, translation, and amino acid biosynthesis [46].

Gene duplication might also have happened in an ancestor of the *Deinococcus-Thermus* group, which is likely to have carried two members of the LytR-CpsA-Psr family: one of the M cluster-type and a more differentiated, phylum-specific one. Both of these LytR-CpsA-Psr proteins have been retained in all of the fully sequenced organisms of this taxon.

An identical distribution of LytR-CpsA-Psr proteins was found in the genus *Staphylococcus* and *Listeria*, which share one member in the MsrR/Psr-like cluster F1 and two members in the LytR/BrpA-like cluster F2. Only *Staphylococcus haemolyticus* encodes a fourth LytR-CpsA-Psr protein, CapM, which clusters in F2 and is the last gene of a putative capsular polysaccharide biosynthesis locus *capA-capM*. *capA-capG* show $\geq 76\%$ amino acid identity to the *cap5[8]A-cap5[8]G* genes of *S. aureus*, while *capH-capM* are unique to *S. haemolyticus* and might be of exogenous origin [47]. The three LytR-CpsA-Psr proteins common to staphylococci and listeria most likely were present in an evolutionary older organism before their phylogenetic separation, however, the cluster analysis leaves the phylogeny of the LytR-CpsA-Psr proteins in general open, and a hypothesis that cluster M represents the most original group remains speculative.

In order to infer evolutionary relatedness of the LytR-CpsA-Psr proteins represented by the 30 strains selected, phylogenetic trees were constructed. However, the LytR-CpsA-Psr proteins show only low amino acid identity and the domain characterizing this family is only about 150 amino acids in length (average value of the respective 93 sequences), rendering construction of phylogenetic trees problematic. Reliable bootstrap support was only obtained close to the terminal nodes, though the results of

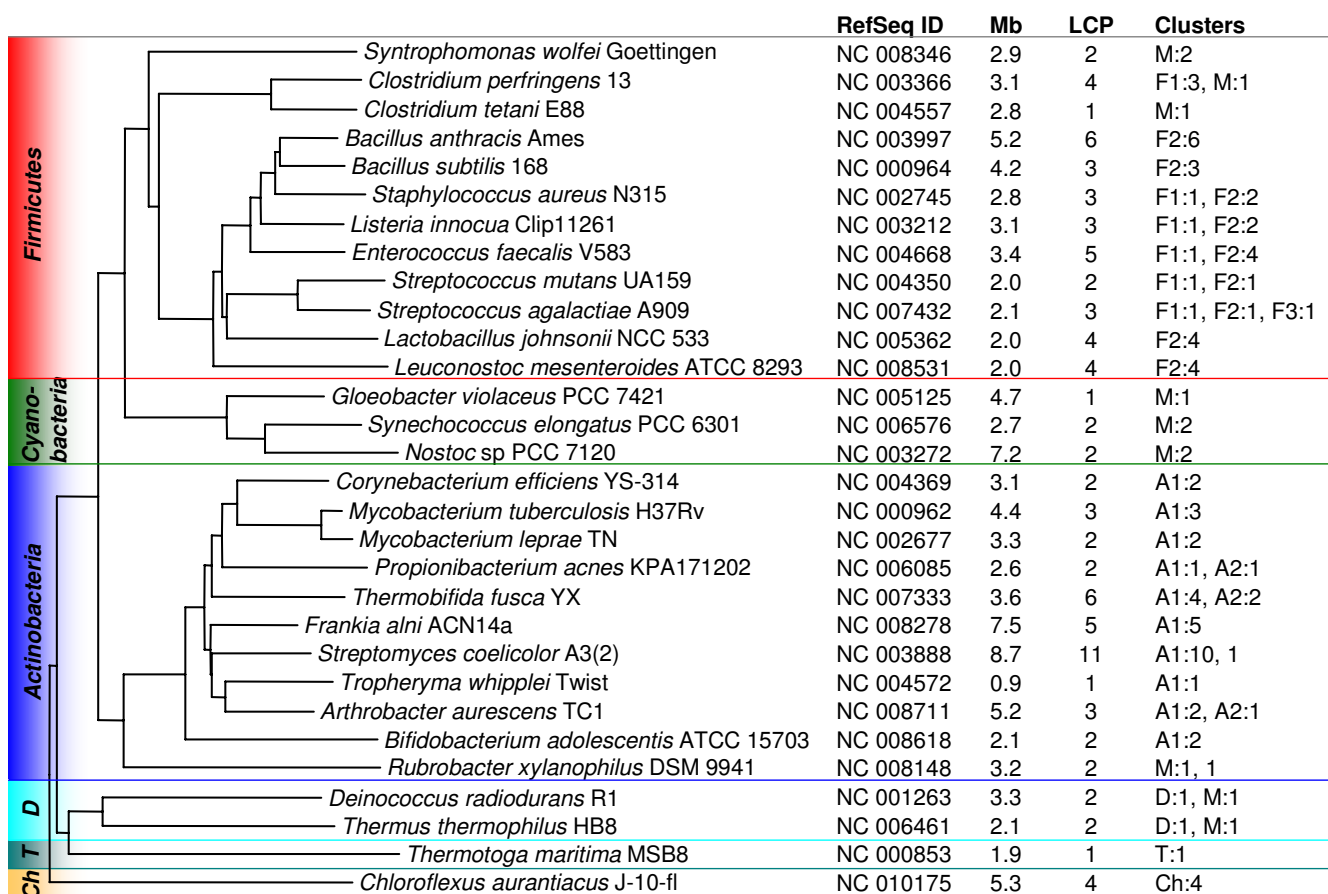


Figure 3
Distribution of LytR-CpsA-Psr proteins in selected genomes. The distribution of the LytR-CpsA-Psr family members was investigated in 30 fully sequenced bacterial strains. Phylogenetic relationships are shown by means of a species tree based on the 16S rRNA gene sequences. The NCBI RefSeq database accession numbers and the sizes of the genomes (total of all DNA molecules) in mega bases (Mb) are indicated. The number of LytR-CpsA-Psr family members (LCP) identified in the corresponding genome are given and the last column summarizes in which CLANS cluster they were found as follows: 'denotation of the cluster': 'number of proteins'. A number without an indication of a cluster means that this sequence did not group with any other sequences.

the cluster analyses were supported (Additional file 3), reflecting taxonomic relationships as shown in the 16S rRNA gene tree in Figure 3.

Secondary structure prediction

A subset of twelve sequences representing the LytR-CpsA-Psr proteins of clusters M, A1, F1, F2, F3, Ch, and T was subjected to secondary structure prediction. The sequences analyzed included the three eponymous proteins LytR, CpsA, and Psr, as well as other family members that have been described in literature, i.e. BrpA, MsrR, and SA0908. The sequences varied in length from 293 to 485 amino acids and they showed very different contents of charged amino acids and theoretical isoelectric points (pI) ranging from 5.17 to 9.66 (Additional file 4). All

sequences contained one consensus transmembrane region predicted by all of the four servers used, except for CpsA, which was predicted to have three transmembrane helices. The majority of the membrane topology predictions suggested an intracellular N-terminus, implicating an extracellular location of the LytR-CpsA-Psr domain.

The putative cytoplasmic tails of eight sequences were very short ranging from only six to 19 amino acids while CE2746 of *C. efficiens* [UniProt:Q8FLW3] showed a long cytoplasmic domain of 121–123 amino acids. A graphical overview of the respective architectures, illustrating also variations in the length of the extracellular domains, is given in Figure 5. The overall identities of the amino acid

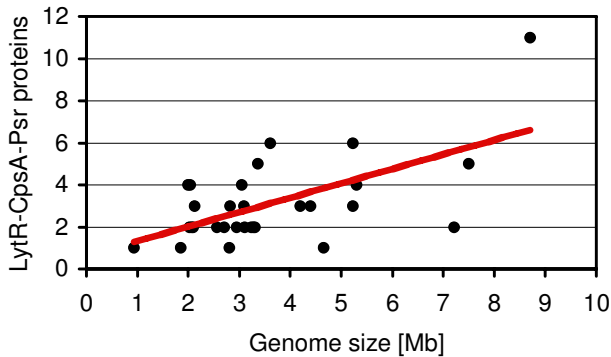


Figure 4
Relationship of the genome size to the number of LytR-CpsA-Psr proteins. With an increasing genome size generally an increased number of LytR-CpsA-Psr family members was found in the 30 strains investigated. The coefficient of correlation was 0.621, indicating an intermediate positive correlation, which is illustrated by the linear regression line (red).

sequences were only between 17 and 39% (Additional file 5).

There was a good overlap of the predicted structural elements within the LytR-CpsA-Psr domain and its flanking regions (extended LytR-CpsA-Psr domain), revealing a mixed α/β -structure (Figure 6). Secondary structure predictions have to be interpreted with caution and only experimental resolution will allow accurate interpretation of structure-function relationships. However, preliminary circular dichroism (CD) spectra of the purified and refolded extracellular domain of MsrR of *S. aureus* support predicted α -helix and β -sheet contents (L. Lüthy, unpublished).

The amino acid identities of the extended LytR-CpsA-Psr domains ranged from 21 to 44% (Additional file 5) and were hence only little above those of the full length proteins. Interestingly, the few fully conserved residues included nine charged amino acids: aspartic acid, arginine, and lysine. Charged residues are often involved in the active site of proteins, where they interact with charged ligands or bind substrates. Moreover, they may play a role in formation of protein-protein complexes and are important for folding, stability, and solubilization of proteins [48]. It is therefore very likely that at least some of the conserved residues are indispensable for the function of the LytR-CpsA-Psr domain in these proteins.

High sequence divergence was observed with regard to some of the peripheric clusters, e.g. the *Spirochaetes* cluster S. Although the secondary structure predictions mainly

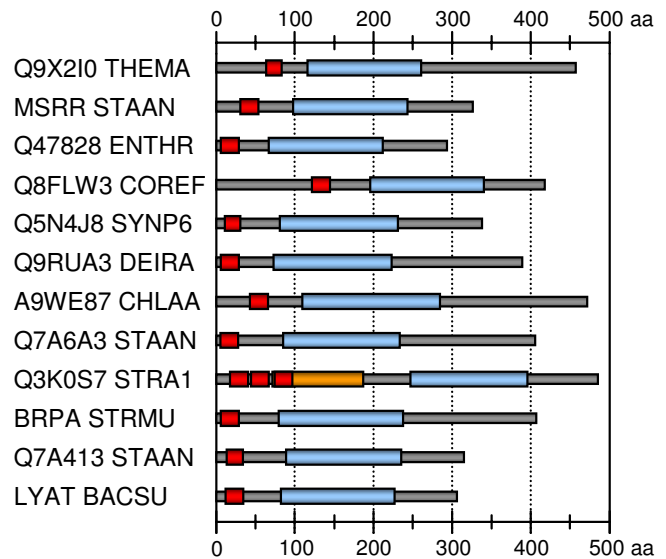


Figure 5
Architectures of the LytR-CpsA-Psr proteins used for secondary structure prediction. Most of the sequences showed a for LytR-CpsA-Psr proteins typical domain organization with a short N-terminal tail, a single transmembrane domain (red), and a long C-terminal part carrying the LytR-CpsA-Psr domain (light blue). An exception of this architecture is found in the group of CpsX-like proteins, represented by CpsX of *S. agalactiae* (Q3K0S7_STRA1), which contain three predicted transmembrane helices and a DNA polymerase processivity factor domain (orange) in addition to the LytR-CpsA-Psr domain. The sequence identifiers correspond to the UniProt entry names (see Additional file 4).

agreed with those of the other LytR-CpsA-Psr members, there was a very weak conservation of amino acids (data not shown), possibly indicating a different function.

Confirmation of membrane topology

The predicted membrane topology of the LytR-CpsA-Psr proteins was verified by PhoA fusions using the staphylococcal protein MsrR as a representative of the protein family. The amino acid sequence of MsrR is identical in twelve of the 14 sequenced *S. aureus* strains available on the NCBI website [16] and there is only one amino acid difference in strains Mu50 and Mu3, which both have lysine at position 146 instead of glutamic acid (Additional files 5 and 6). For the following analyses, the sequence represented by *S. aureus* strain N315 was employed (see also Additional file 4).

PhoA fusions and activity assay

Five different PhoA fusion proteins were constructed, with MsrR lacking either 0, 80, 147, 246, or 295 amino acids from the C-terminus (full length, $\Delta 80$, $\Delta 147$, $\Delta 246$, and $\Delta 295$) (Figure 7A), and were expressed in the *phoA*-*E. coli*

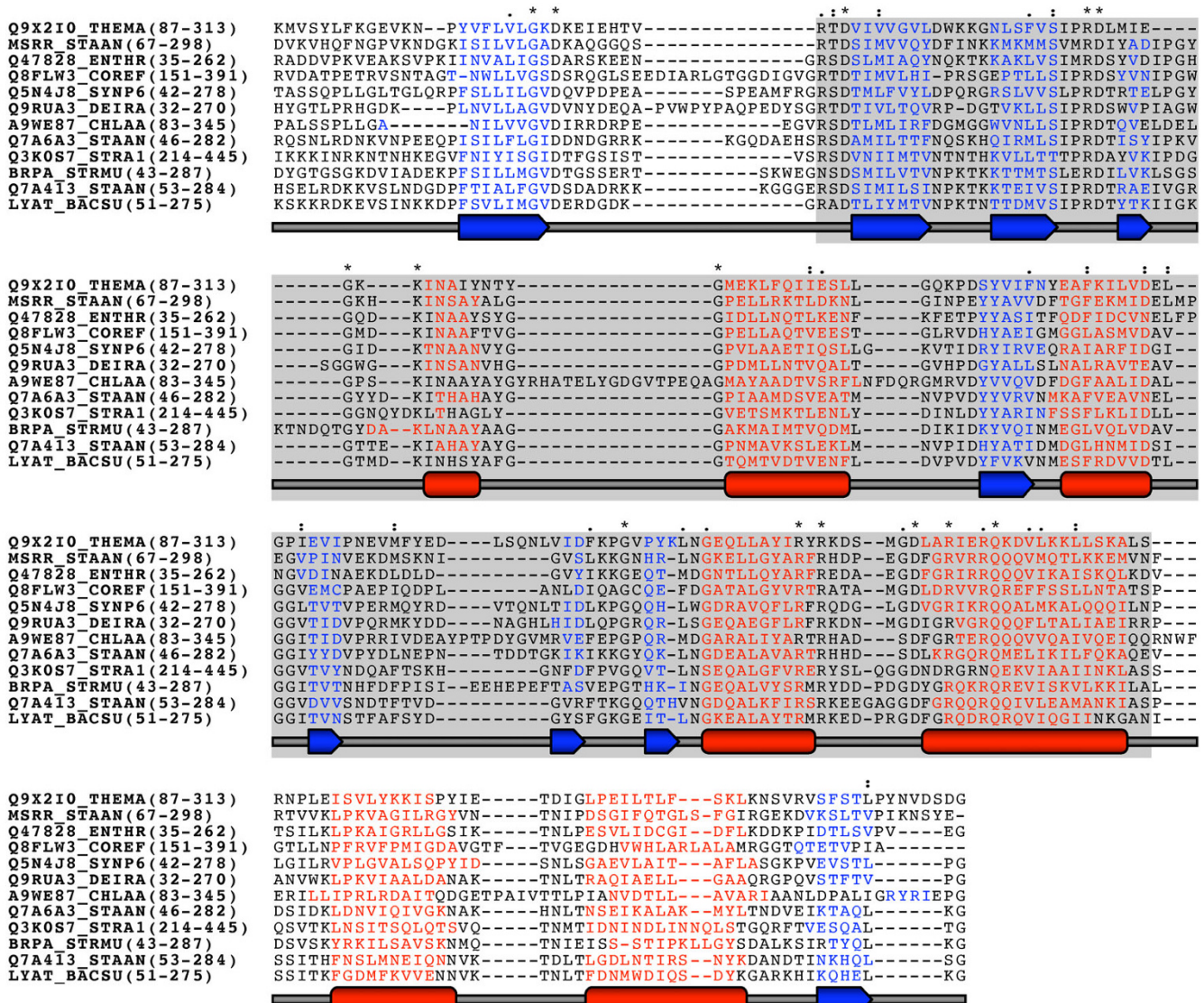


Figure 6
Secondary structure prediction of the extended LytR-CpsA-Psr domain. Full length sequences were aligned using MUSCLE and secondary structures were predicted using the prediction servers SsPro, PSIPRED, and Jpred. The consensus α -helices (red) and β -sheets (blue) based on the results of the three servers are shown for the extended LytR-CpsA-Psr domain. A schematic drawing underneath the aligned sequences further illustrates the sequential order of predicted α -helices (barrels) and β -sheets (arrows). The LytR-CpsA-Psr domain as defined in the PFAM database (PF03816) is shaded in grey. Fully conserved residues are marked by an asterisk (*) above the aligned sequences. A dash (-) represents a gap in the alignment. A colon (:) indicates that one of the following groups is conserved: STA, MILV, MILF, or NDEQ; a dot (.) indicates that one of the following groups is conserved: FVLIM, NEQHRK, SNDEQK, or SAG. The sequence identifiers correspond to the UniProt entry names (see Additional file 4), and the amino acid range covering the extended LytR-CpsA-Psr domain of each sequence is given in parentheses.

strain CC118. Correct folding of PhoA relies on the formation of disulfide bonds and occurs in the periplasm only. Therefore, PhoA activity can only be measured if the fusion protein spans the membrane in a suitable orientation. The fusions of PhoA to the full length protein and to

MsrR truncated by 80, 147, and 246 amino acids from the C-terminus led to significant PhoA activity, indicating that PhoA had been translocated to the periplasmic space (Figure 7B), however, there were large variations in activity between the different fusion constructs. The highest activ-

ity was reached when PhoA was fused to MsrR shortly after the transmembrane domain ($\Delta 246$), whereas the fusions towards the C-terminus ($\Delta 147$, $\Delta 80$, and full length) showed lower activities. Western blot hybridization using α -PhoA antibodies revealed partial degradation of the $\Delta 147$, $\Delta 80$, and full length fusion constructs, which likely accounts for the observed differences in PhoA activity (Figure 7C). In contrast, one strong, single band of the expected size was detected for the $\Delta 246$ as well as for the $\Delta 295$ fusion constructs. The $\Delta 295$ fusion lacks the transmembrane domain and did not yield PhoA activity above the threshold value, indicating that the fusion construct remained in the cytoplasm.

These findings confirmed the predicted arrangement of the LytR-CpsA-Psr protein MsrR in the membrane: The short N-terminal tail is located inside of the cell and is followed by one transmembrane helix and an extracellular domain carrying the LytR-CpsA-Psr domain. Most members of the LytR-CpsA-Psr protein family show very similar predicted architectures as MsrR.

Conclusion

Members of the LytR-CpsA-Psr family are known to influence clinically important attributes of various pathogens [2,4,5]. The LytR-CpsA-Psr domain was found to be restricted to the kingdom of bacteria. Being present in hyperthermophilic organisms of evolutionary early branches, the LytR-CpsA-Psr proteins appear to constitute an ancient protein family. All Gram-positive organisms were found to possess LytR-CpsA-Psr proteins, except for the *Mollicutes* that seem to have lost this protein family in the course of diversification. In contrast, the LytR-CpsA-Psr domain was only observed in a minority of the Gram-negatives, and it is conceivable that the LytR-CpsA-Psr proteins were lost in *Proteobacteria* and *Chlamydiae* as well as in most of the *Bacteroidetes* during evolution. LytR-CpsA-Psr proteins are shared by the most diverse prokaryotic phyla, including the thermophiles of the first bacterial branches, and the genes encoding LytR-CpsA-Psr proteins often belong to the core genome. Although these proteins do not appear to be essential, they presumably confer a selective advantage to their carriers, as becomes apparent in vitro when mutant cells are challenged with antibiotic, acidic, or oxidative stress [4,5].

The finding that various members of this protein family are part of the cell wall stress stimulon in different organisms [30,33-38] might further suggest a role in ensuring cell envelope integrity. Nevertheless, the specific functions of the LytR-CpsA-Psr proteins are probably more diverse than originally anticipated, as indicated by the results of the cluster analysis, explaining discrepancies in phenotypes of knock out mutants [1-4,7].

Our analyses of the occurrence and distribution of the LytR-CpsA-Psr proteins as well as the determination of putative structural elements and conserved amino acids of the family-specific domain will provide a guideline for the design of future experiments to elucidate structure-function relationships.

Methods

Sequence collection

As of August 20, 2008, accession numbers of LytR-CpsA-Psr protein family members included in the InterPro database (InterPro entry IPR004474), a consortium of the member databases PROSITE, Pfam, Prints, ProDom, SMART, and TIGRFAMs [12], were obtained. Homologues sequences not present in InterPro were identified by PSI-BLAST [13] against the NCBI non-redundant protein sequence database (nr) using MsrR of *S. aureus* strain N315 [UniProt:Q99Q02] as a query and an E-value threshold of 10^{-4} . Amino acid sequences were retrieved from UniProt [49] or from NCBI [16].

Cluster analysis

To identify closely related sequences within the LytR-CpsA-Psr members, CLANS (CLuster ANalysis of Sequences) was used [24]. For two-dimensional visualization of pairwise sequence similarities, based on P-values of high-scoring segment pairs (HSPs) resulting from a BLAST all against all search, the P-value cut off was set to 10^{-35} and the attraction (attract) and repulsion (repuls) exponents were set to 2. Network-based clustering was applied to assist designation of clusters.

Structure analysis

Sequence compositions and theoretical pI were analyzed using ProtParam [50]. Transmembrane regions were predicted using TMHMM [51], the "DAS"-Transmembrane Prediction server [52], TMPred [53], and HMMTOP [54]. Secondary structures were predicted using SsPro [55], PSIPRED [56], and Jpred [57].

Amino acid sequence alignments

Amino acid sequences were aligned using MUSCLE version 3.6 [58] and protein identity matrices of the MUSCLE alignments were calculated using ClustalX 2.1 [59].

Phylogenetic trees

Phylogenetic trees were constructed using programs of the PHYLIP 3.67 package [60]. For construction of a species tree, aligned 16S rRNA genes were downloaded from the ribosomal RNA database RDP-II [61]. Misaligned characters were manually edited using SeaView version 2.2 [62]. The C- and N-termini were trimmed and gap-only sites were deleted from the alignment. A distance matrix, computed with DNADIST using the F84 model of nucleotide substitution, was used for construction of a distance tree

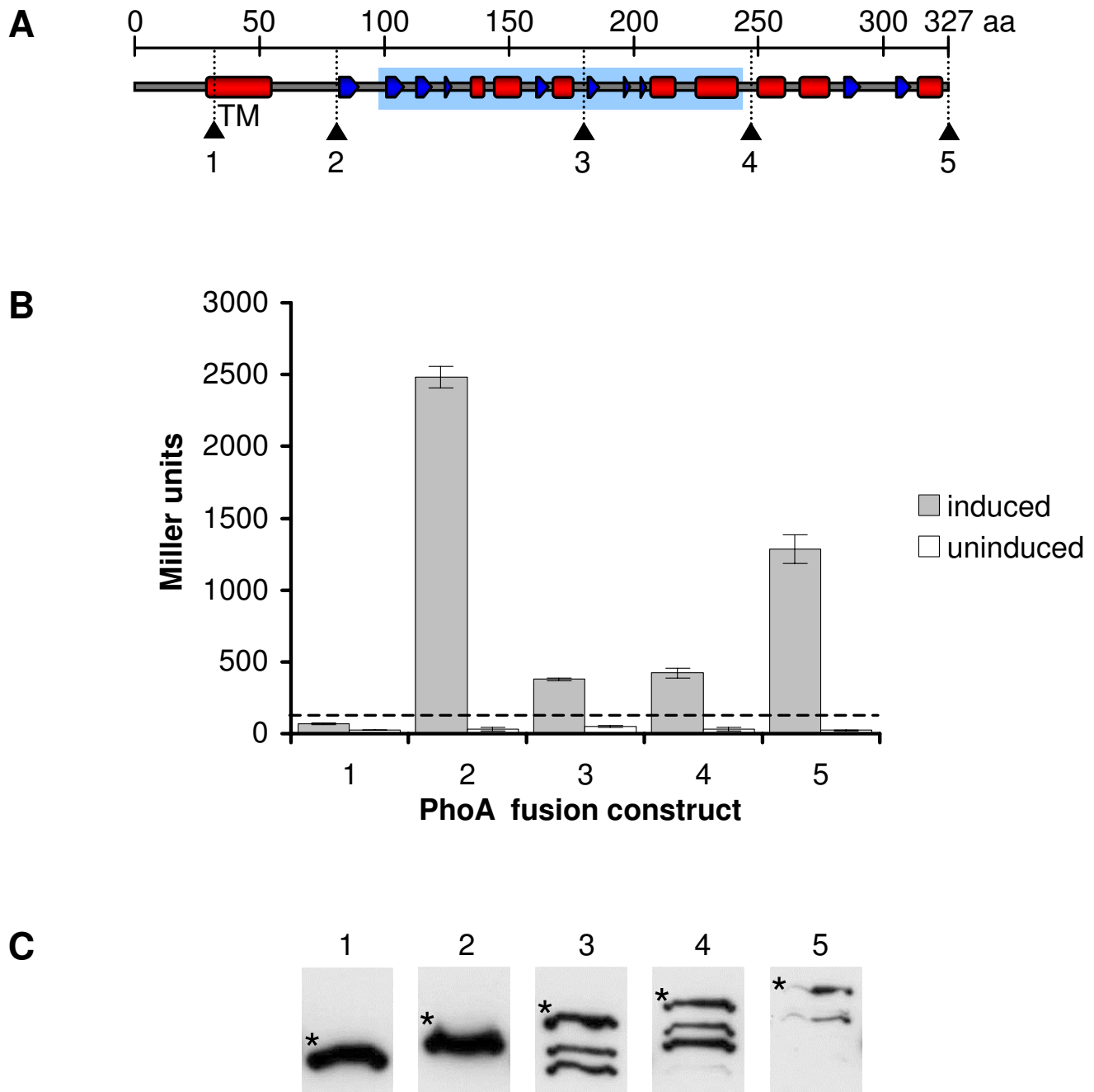


Figure 7
Confirmation of membrane topology of MsrR. (A) Schematic drawing of MsrR showing the predicted secondary structure (red barrel, α -helix; blue arrow, β -sheet) and transmembrane helix (TM). Black arrows indicate the positions where MsrR was fused to PhoA yielding fusion constructs number 1 ($\Delta 295$, deletion of 295 aa from the C-terminus), 2 ($\Delta 246$), 3 ($\Delta 147$), 4 ($\Delta 80$), and 5 (full length). (B) PhoA activity of fusion constructs after induction with arabinose as well as of uninduced control samples. The average and standard deviations of the results obtained using three different clones of each fusion construct are shown. The dashed line indicates a threshold of 120 Miller Units, which equals twice the highest background activity measured. (C) Detection of PhoA fusion constructs by Western blot hybridization using α -PhoA antibodies. For all fusion constructs, bands corresponding to the expected size were detected (marked by an asterisk): 52.6 kDa (1), 58.0 kDa (2), 69.3 kDa (3), 76.9 kDa (4), and 85.5 kDa (5). In the lanes containing constructs number 3–5, several smaller bands were present, which most probably represent degradation products.

Table 1: Strains and plasmids

Strain or plasmid	Relevant genotype or phenotype ^{a)}	Reference or source
Strains		
<i>S. aureus</i> Newman	Clinical isolate (ATCC 25904)	[66]
<i>E. coli</i> DH5α CC118	<i>supE44 ΔlacU169(φ 80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i> <i>Δ(ara-leu)7679 ΔlacX74 ΔphoA20 galE galK thi rpsE rpoB argE(am) recA1</i>	Invitrogen [67]
Plasmids		
pHA-1	<i>E. coli</i> plasmid containing an arabinose inducible promoter 5' of a signal sequence-less <i>phoA</i> reporter gene, <i>Am^r</i>	[68]

^{a)}*Am^r*, Ampicillin resistance marker

with FITCH (Fitch-Margoliash criterion) allowing global rearrangements and randomizing the input order of the sequences 10 times. For construction of protein sequence trees, bootstrap resampling was performed using SEQBOOT. Consensus distance trees were generated using PROTDIST with the Jones-Taylor-Thornton substitution matrix, NEIGHBOR, and CONSENSE. Phylogenetic trees were drawn with Dendroscope version 1.2.4 [63].

Bacterial strains and culture conditions

Strains and plasmids used in this study are listed in Table 1. Bacteria were maintained at 37 °C on Luria Bertani (LB) (Difco Laboratories, Detroit, MI, USA) agar plates supplemented with 100 µg/ml ampicillin or on sheep-blood agar plates.

PhoA fusion constructs

PCR products covering different lengths of *msrR* of *S. aureus* strain Newman were generated using the forward primer 5'-GCGCCTCGAGATGGATAAAGAACTAATGA-3' and the following reverse primers: 5'-GCGCGGTACCTCTTCATCTAAAAAGTCTTT-3' (full length), 5'-GCGCGGTACCGTTCTAAAAATTAACCATTTTC-3' (Deletion of 80 aa from the C-terminus, Δ80), 5'-GCGCGGTACCTCAGGCATTAATTCATCAA-3 (Δ147), 5'-GCGCGGTACCCCATCATTTTTACTGGTCC-3' (Δ246), and 5'-GCGCGGTACCAATTTCTAATTTTCTTCTT-3' (Δ295). Recognition sites of restriction endonucleases are underlined. The amplicons were digested using XhoI and KpnI and inserted into pHA-1 downstream of the *araB* arabinose-inducible promoter and upstream of the *phoA* gene lacking both the 5' segment coding for the signal sequence and the first five residues of the mature protein. The resulting plasmids were transferred into the *phoA* *E. coli* strain CC118.

Expression of PhoA fusion proteins and PhoA activity assay

Overnight cultures of *E. coli* strain CC118 harboring the *msrR-phoA* fusion constructs of interest were diluted 1:100 in LB broth containing 100 µg/ml ampicillin and grown to an optical density at 600 nm of 0.5. Bacterial cultures were then divided into halves: in one half, protein expression was induced with 0.2% arabinose, while the other half was left untreated. After growth for an additional hour, 1 ml of induced and non-induced cells, respectively, were harvested for determination of PhoA activity by a *p*-nitrophenyl phosphate (*p*NPP) (Sigma-Aldrich) cleavage assay [64]. In addition, 200 µl of each cell suspension were collected for verification of PhoA fusion protein expression by Western blot using an α-PhoA antibody conjugated to horseradish peroxidase (HRP) (Abcam Ltd., UK).

Western blot analysis

Whole cells were heated for 5 min at 65 °C in 5× SDS loading buffer supplemented with 0.05% *N*-lauroylsarcosyl and subjected to SDS-12%-PAGE. Separated proteins were transferred onto nitrocellulose membranes (Hybond-ECL; Amersham Biosciences), and for detection of PhoA, a HRP-conjugated α-PhoA antibody (Abcam) diluted 1:100'000 and the SuperSignal West Pico Chemiluminescent detection kit (Pierce) were used.

Authors' contributions

JH carried out sequence-based analyses, performed PhoA assays and Western blots, and drafted the manuscript. LL constructed PhoA fusion plasmids. BBB participated in the design and coordination of the study and contributed to the writing of the manuscript. PSM participated in the design of the study and contributed to the interpretation of the results and writing of the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Results of CLANS clustering analysis. CLANS clusters of the *LytR-CpsA-Psr* proteins based on pair-wise sequence similarities resulting from a BLAST all against all search using a P-value cut off of 10^{-35} . The clusters were named using the first character(s) of the respective bacterial phylum and were numbered in case there were multiple clusters of the same phylum. The NCBI GI numbers and UniProt/SwissProt accessions are given. A, Actinobacteria; Ch, Chloroflexi; Cy, Cyanobacteria; D, Deinococcus-Thermus; F, Firmicutes; S, Spirochaetes; T, Thermotogae.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-617-S1.xls>]

Additional file 2

***LytR-CpsA-Psr* proteins used in phylogenetic analyses.** Table showing the *LytR-CpsA-Psr* family members identified in the 30 selected, completely sequenced bacterial strains.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-617-S2.xls>]

Additional file 3

Neighbour-joining tree of *LytR-CpsA-Psr* proteins. A phylogenetic tree was constructed based on a MUSCLE alignment of the full length sequences. Bootstrap values only support the branching order towards the terminal nodes and deeper branches have to be considered unresolved (red values). Using solely the *LytR-CpsA-Psr* domain for the initial alignment, changing parameters, e.g. for calculation of the distance matrix, or applying another tree construction algorithm did not improve the results (data not shown). The UniProt entry names are given, and the subtrees are colored using the same color scheme as in Figures 2 and 3 (blue, Actinobacteria; green, Cyanobacteria; yellow, Chloroflexi; turquoise, Deinococcus-Thermus; red, Firmicutes; petrol, Thermotogae). Brackets indicate the corresponding CLANS clusters.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-617-S3.pdf>]

Additional file 4

Primary sequence analyses of *LytR-CpsA-Psr* proteins and results of transmembrane region prediction. The primary sequences of the full length proteins as well as of the *LytR-CpsA-Psr* domain as defined in the PFAM database (PF03816) were analyzed using ProtParam. Transmembrane segments were predicted using the web-based servers TMHMM, Tmpred, Das, and HMMTOP. The first and the last residue of the predicted transmembrane domains are indicated.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-617-S4.pdf>]

Additional file 5

Percent identity matrices of *LytR-CpsA-Psr* protein sequences. The sequence identifiers correspond to the UniProt entry names (see also Additional file 4). (A) Percent identity matrix of the full length sequences used for secondary structure prediction. The clusters as identified using CLANS are given. (B) Percent identity matrix of the extended *LytR-CpsA-Psr* domains of the sequences used for secondary structure prediction. For each sequence, the amino acid range is given in parentheses. (C) Percent identity matrix of the *MsrR* protein sequences in staphylococci. STAS1, *S. saprophyticus* ATCC15305; STAAAN, *S. aureus* N315; STAAM, *S. aureus* Mu50; STAES, *S. epidermidis* ATCC12228; STAHJ, *S. haemolyticus* JCSC1435.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-617-S5.pdf>]

Additional file 6

Sequence comparison of *MsrR* in staphylococci. The lysine (K) residue at position 146 of *MsrR* of *S. aureus* Mu50 and Mu3 is indicated by a red box. The *LytR-CpsA-Psr* domain as defined in the PFAM database (PF03816) is shaded in grey, and conserved residues as determined in Figure 6 are highlighted in blue. Residues of the predicted transmembrane regions are printed in red. Sequences are labelled using UniProt entry names. STAS1, *S. saprophyticus* ATCC15305; STAAAN, *S. aureus* N315; STAAM, *S. aureus* Mu50; STAES, *S. epidermidis* ATCC12228; STAHJ, *S. haemolyticus* JCSC1435.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-617-S6.pdf>]

Acknowledgements

We are grateful to D. Sjöstrand and T. Urbig (Department of Biochemistry and Biophysics, Stockholm University, Sweden) for providing plasmid pHA-1.

This study was supported by European Commission grant LSHM-CT-2003-503335 (SBF03.0098), and Swiss National Science Foundation grants 31-117707 and PMPDB-114323 to BBB and PSM, respectively.

References

1. Chatfield CH, Koo H, Quivey RG Jr: **The putative autolysin regulator *LytR* in *Streptococcus mutans* plays a role in cell division and is growth-phase regulated.** *Microbiology* 2005, **151**(Pt 2):625-631.
2. Cieslewicz MJ, Kasper DL, Wang Y, Wessels MR: **Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci.** *J Biol Chem* 2001, **276**(1):139-146.
3. Lazarevic V, Margot P, Soldo B, Karamata D: **Sequencing and analysis of the *Bacillus subtilis* *lytRABC* divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier.** *J Gen Microbiol* 1992, **138**(Pt 9):1949-1961.
4. Rossi J, Bischoff M, Wada A, Berger-Bächi B: ***MsrR*, a putative cell envelope-associated element involved in *Staphylococcus aureus* *sarA* attenuation.** *Antimicrob Agents Chemother* 2003, **47**(8):2558-2564.
5. Wen ZT, Baker HV, Burne RA: **Influence of *BrpA* on critical virulence attributes of *Streptococcus mutans*.** *J Bacteriol* 2006, **188**(8):2983-2992.
6. Ligozzi M, Pittaluga F, Fontana R: **Identification of a genetic element (*psr*) which negatively controls expression of *Enterococcus faecalis* *psr*.** *J Biol Chem* 2004, **279**(12):12000-12006.

- coccus hirae* penicillin-binding protein 5. *J Bacteriol* 1993, **175(7)**:2046-2051.
7. Sapunovic F, Franssen C, Stefanic P, Amoroso A, Dardenne O, Coyette J: **Redefining the role of *psr* in β -lactam resistance and cell autolysis of *Enterococcus hirae*.** *J Bacteriol* 2003, **185(20)**:5925-5935.
 8. Nakano K, Fujita K, Nishimura K, Nomura R, Ooshima T: **Contribution of biofilm regulatory protein A of *Streptococcus mutans*, to systemic virulence.** *Microbes Infect* 2005, **7(11-12)**:1246-1255.
 9. Bae T, Banger AK, Wallace A, Glass EM, Aslund F, Schneewind O, Missiakas DM: ***Staphylococcus aureus* virulence genes identified by *bursa aurealis* mutagenesis and nematode killing.** *Proc Natl Acad Sci USA* 2004, **101(33)**:12312-12317.
 10. Maadani A, Fox KA, Mylonakis E, Garsin DA: ***Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host.** *Infect Immun* 2007, **75(5)**:2634-2637.
 11. Lowe AM, Beattie DT, Deresiewicz RL: **Identification of novel staphylococcal virulence genes by in vivo expression technology.** *Mol Microbiol* 1998, **27(5)**:967-976.
 12. Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MD, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJ, Zdobnov EM: **The InterPro database, an integrated documentation resource for protein families, domains and functional sites.** *Nucleic Acids Res* 2001, **29(137-40)** [<http://www.ebi.ac.uk/interpro/>].
 13. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25(17)**:3389-3402.
 14. Cove D: **The moss *Physcomitrella patens*.** *Annu Rev Genet* 2005, **39**:339-358.
 15. Marin B, Nowack ECM, Melkonian M: **A plastid in the making: evidence for a second primary endosymbiosis.** *Protist* 2005, **156**:425-432.
 16. **National Center for Biotechnology Information (NCBI)** [<http://www.ncbi.nlm.nih.gov/>]
 17. Hoiczky E, Hansel A: **Cyanobacterial cell walls: news from an unusual prokaryotic envelope.** *J Bacteriol* 2000, **182(5)**:1191-1199.
 18. Jürgens UJ, Meissner J, Fischer U, König WA, Weckesser J: **Ornithine as a constituent of the peptidoglycan of *Chloroflexus aurantiacus*, diamminopimelic acid in that of *Chlorobium vibriforme* f. *thiosulfatophilum*.** *Arch Microbiol* 1987, **148(1)**:72-76.
 19. Quintela JC, Garcia-del Portillo F, Pittenauer E, Allmaier G, de Pedro MA: **Peptidoglycan fine structure of the radiotolerant bacterium *Deinococcus radiodurans* Sark.** *J Bacteriol* 1999, **181(1)**:334-337.
 20. Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S, Chen F, Lapidus A, Ferriera S, Johnson J, Steglich C, Church GM, Richardson P, Chisholm SW: **Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*.** *PLoS Genet* 2007, **3(12)**:e231.
 21. Seedorf H, Fricke WF, Veith B, Bruggemann H, Liesegang H, Strittmatter A, Miethke M, Buckel W, Hinderberger J, Li F, Hagemeyer C, Thauer RK, Gottschalk G: **The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features.** *Proc Natl Acad Sci USA* 2008, **105(6)**:2128-2133.
 22. Olsen GJ, Woese CR, Overbeek R: **The winds of (evolutionary) change: breathing new life into microbiology.** *J Bacteriol* 1994, **176(1)**:1-6.
 23. Doolittle WF: **Phylogenetic classification and the universal tree.** *Science* 1999, **284(5423)**:2124-2129.
 24. Frickey T, Lupas A: **CLANS: a Java application for visualizing protein families based on pairwise similarity.** *Bioinformatics* 2004, **20(18)**:3702-3704.
 25. Bibb MJ, Sherman DH, Omura S, Hopwood DA: **Cloning, sequencing and deduced functions of a cluster of *Streptomyces* genes probably encoding biosynthesis of the polyketide antibiotic frenolicin.** *Gene* 1994, **142(1)**:31-39.
 26. Linton KJ, Jarvis BW, Hutchinson CR: **Cloning of the genes encoding thymidine diphosphoglucose 4,6-dehydratase and thymidine diphospho-4-keto-6-deoxyglucose 3,5-epimerase from the erythromycin-producing *Saccharopolyspora erythraea*.** *Gene* 1995, **153(1)**:33-40.
 27. Rice LB, Carías LL, Hutton-Thomas R, Sifaoui F, Gutmann L, Rudin SD: **Penicillin-binding protein 5 and expression of ampicillin resistance in *Enterococcus faecium*.** *Antimicrob Agents Chemother* 2001, **45(5)**:1480-1486.
 28. Liu S, Graham JE, Bigelow L, Morse PD II, Wilkinson BJ: **Identification of *Listeria monocytogenes* genes expressed in response to growth at low temperature.** *Appl Environ Microbiol* 2002, **68(4)**:1697-1705.
 29. McAleese F, Wu SW, Sieradzki K, Dunman P, Murphy E, Projan S, Tomasz A: **Overexpression of genes of the cell wall stimulon in clinical isolates of *Staphylococcus aureus* exhibiting vancomycin-intermediate-*S. aureus*-type resistance to vancomycin.** *J Bacteriol* 2006, **188(3)**:1120-1133.
 30. Utaida S, Dunman PM, Macapagal D, Murphy E, Projan SJ, Singh VK, Jayaswal RK, Wilkinson BJ: **Genome-wide transcriptional profiling of the response of *Staphylococcus aureus* to cell-wall-active antibiotics reveals a cell-wall-stress stimulon.** *Microbiology* 2003, **149(Pt 10)**:2719-2732.
 31. Cao M, Helmann JD: **The *Bacillus subtilis* extracytoplasmic-function σ^X factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides.** *J Bacteriol* 2004, **186(4)**:1136-1146.
 32. Huang X, Helmann JD: **Identification of target promoters for the *Bacillus subtilis* σ^X factor using a consensus-directed search.** *J Mol Biol* 1998, **279(1)**:165-173.
 33. Cao M, Wang T, Ye R, Helmann JD: **Antibiotics that inhibit cell wall biosynthesis induce expression of the *Bacillus subtilis* σ^W and σ^M regulons.** *Mol Microbiol* 2002, **45(5)**:1267-1276.
 34. Eiamphungporn W, Helmann JD: **The *Bacillus subtilis* σ^M regulon and its contribution to cell envelope stress responses.** *Mol Microbiol* 2008, **67(4)**:830-848.
 35. Mascher T, Margulis NG, Wang T, Ye RW, Helmann JD: **Cell wall stress responses in *Bacillus subtilis*: the regulatory network of the bacitracin stimulon.** *Mol Microbiol* 2003, **50(5)**:1591-1604.
 36. Wecke T, Veith B, Ehrenreich A, Mascher T: **Cell envelope stress response in *Bacillus licheniformis*: integrating comparative genomics, transcriptional profiling, and regulon mining to decipher a complex regulatory network.** *J Bacteriol* 2006, **188(21)**:7500-7511.
 37. Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K: **Two-component system *VraSR* positively modulates the regulation of cell-wall biosynthesis pathway in *Staphylococcus aureus*.** *Mol Microbiol* 2003, **49(3)**:807-821.
 38. McCallum N, Spehar G, Bischoff M, Berger-Bächi B: **Strain dependence of the cell wall-damage induced stimulon in *Staphylococcus aureus*.** *Biochim Biophys Acta* 2006, **1760(10)**:1475-1481.
 39. Espinasse S, Gohar M, Lereclus D, Sanchis V: **An extracytoplasmic-function sigma factor is involved in a pathway controlling β -exotoxin I production in *Bacillus thuringiensis* subsp. *thuringiensis* strain 407-1.** *J Bacteriol* 2004, **186(10)**:3108-3116.
 40. Bernard R, El Ghachi M, Mengin-Lecreux D, Chippaux M, Denizot F: **BcrC from *Bacillus subtilis* acts as an undecaprenyl pyrophosphate phosphatase in bacitracin resistance.** *J Biol Chem* 2005, **280(32)**:28852-28857.
 41. El Ghachi M, Derbise A, Bouhss A, Mengin-Lecreux D: **Identification of multiple genes encoding membrane proteins with undecaprenyl pyrophosphate phosphatase (UppP) activity in *Escherichia coli*.** *J Biol Chem* 2005, **280(19)**:18689-18695.
 42. Brunskill EW, Bayles KW: **Identification and molecular characterization of a putative regulatory locus that affects autolysis in *Staphylococcus aureus*.** *J Bacteriol* 1996, **178(3)**:611-618.
 43. Bender MH, Cartee RT, Yother J: **Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*.** *J Bacteriol* 2003, **185(20)**:6057-6066.
 44. Shamoo Y, Steitz TA: **Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex.** *Cell* 1999, **99(2)**:155-166.
 45. Andreeva A, Howorth D, Brenner SE, Hubbard TJ, Chothia C, Murzin AG: **SCOP database in 2004: refinements integrate structure and sequence family data.** *Nucleic Acids Res* 2004:D226-229.
 46. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman

- A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabinowitsch E, Rajadream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrrell BG, Parkhill J, Hopwood DA: **Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)**. *Nature* 2002, **417(6885)**:141-147.
47. Takeuchi F, Watanabe S, Baba T, Yuzawa H, Ito T, Morimoto Y, Kuroda M, Cui L, Takahashi M, Ankai A, Baba S, Fukui S, Lee JC, Hiramatsu K: **Whole-genome sequencing of *Staphylococcus haemolyticus* uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species**. *J Bacteriol* 2005, **187(21)**:7292-7308.
 48. Gitlin I, Carbeck JD, Whitesides GM: **Why are proteins charged? Networks of charge-charge interactions in proteins measured by charge ladders and capillary electrophoresis**. *Angew Chem Int Ed Engl* 2006, **45(19)**:3022-3060.
 49. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS: **The universal protein resource (UniProt)**. *Nucleic Acids Res* 2005:D154-159.
 50. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A: **Protein identification and analysis tools on the ExPASy server**. *Humana Press* 2005.
 51. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL: **Predicting transmembrane protein topology with a hidden markov model: application to complete genomes**. *Journal of Molecular Biology* 2001, **305(3)**:567-580.
 52. Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A: **Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method**. *Protein Eng* 1997, **10(6)**:673-676.
 53. Hofmann K, Stoffel W: **TMbase - A database of membrane spanning protein segments**. *Biol Chem* 1993, **374**:166.
 54. Tusnady GE, Simon I: **The HMMTOP transmembrane topology prediction server**. *Bioinformatics* 2001, **17(9)**:849-850.
 55. Cheng J, Randall AZ, Sweredoski MJ, Baldi P: **SCRATCH: a protein structure and structural feature prediction server**. *Nucleic Acids Res* 2005:W72-76.
 56. Jones DT: **Protein secondary structure prediction based on position-specific scoring matrices**. *J Mol Biol* 1999, **292(2)**:195-202.
 57. Cuff J, Clamp M, Siddiqui A, Finlay M, Barton G: **JPred: a consensus secondary structure prediction server**. *Bioinformatics* 1998, **14(10)**:892-893.
 58. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucl Acids Res* 2004, **32(5)**:1792-1797.
 59. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: **Clustal W and Clustal X version 2.0**. *Bioinformatics* 2007, **23(21)**:2947-2948.
 60. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6**. In *Distributed by the author* University of Washington, Department of Genome Sciences. Seattle; 2005.
 61. Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM, Bandela AM, Cardenas E, Garrity GM, Tiedje JM: **The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data**. *Nucl Acids Res* 2007, **35(suppl_1)**:D169-172.
 62. Galtier N, Gouy M, Gautier C: **SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny**. *Comput Appl Biosci* 1996, **12(6)**:543-548.
 63. Huson D, Richter D, Rausch C, Dezulian T, Franz M, Rupp R: **Dendroscope: An interactive viewer for large phylogenetic trees**. *BMC Bioinformatics* 2007, **8(1)**:460.
 64. Michaelis S, Inouye H, Oliver D, Beckwith J: **Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli***. *J Bacteriol* 1983, **154(1)**:366-374.
 65. Letunic I, Bork P: **Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation**. *Bioinformatics* 2007, **23(1)**:127-128.
 66. Duthie ES, Lorenz LL: **Staphylococcal coagulase; mode of action and antigenicity**. *J Gen Microbiol* 1952, **6(1-2)**:95-107.
 67. Lee E, Manoil C: **Mutations eliminating the protein export function of a membrane-spanning sequence**. *J Biol Chem* 1994, **269(46)**:28822-28828.
 68. Drew D, Sjöstrand D, Nilsson J, Urbig T, Chin CN, de Gier JW, von Heijne G: **Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis**. *Proc Natl Acad Sci USA* 2002, **99(5)**:2690-2695.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

