Cell surface proteome of the marine planctomycete Rhodopirellula baltica

ARTICLE *in* PROTEOMICS · JUNE 2012

Impact Factor: 3.81 · DOI: 10.1002/pmic.201100512 · Source: PubMed

RESEARCH ARTICLE

Cell surface proteome of the marine planctomycete *Rhodopirellula baltica*

Birgit Voigt[1,2], Cao Xuan Hieu[1*,**], Kristina Hempel[1], Dörte Becher[1], Rabea Schlüter[1], Hanno Teeling[3], Frank Oliver Glöckner[3], Rudolf Amann[3], Michael Hecker[1,2] and Thomas Schweder[2,4]*

[1] Institute for Microbiology, Ernst-Moritz-Arndt-University, Greifswald, Germany
[2] Institute of Marine Biotechnology, Greifswald, Germany
[3] Max Planck Institute for Marine Microbiology, Bremen, Germany
[4] Institute for Pharmacy, Ernst-Moritz-Arndt-University, Greifswald, Germany

The surface proteome (surfaceome) of the marine planctomycete *Rhodopirellula baltica* SH1^T was studied using a biotinylation and a proteinase K approach combined with SDS-PAGE and mass spectrometry. 52 of the proteins identified in both approaches could be assigned to the group of potential surface proteins. Among them are some high molecular weight proteins, potentially involved in cell–cell attachment, that contain domains shown before to be typical for surface proteins like cadherin/dockerin domains, a bacterial adhesion domain or the fasciclin domain. The identification of proteins with enzymatic functions in the *R. baltica* surfaceome provides further clues for the suggestion that some degradative enzymes may be anchored onto the cell surface. YTV proteins, which have been earlier supposed to be components of the proteinaceous cell wall of *R. baltica*, were detected in the surface proteome. Additionally, 8 proteins with a novel protein structure combining a conserved type IV pilin/N-methylation domain and a planctomycete-typical DUF1559 domain were identified.

Received: September 27, 2011
Revised: January 24, 2012
Accepted: February 15, 2012

Keywords:
Cell wall proteins / Microbiology / *Rhodopirellula baltica* / Surface proteome

1 Introduction

Surface proteins represent "identity cards" of bacterial cells and are of special importance for the interaction of the bacteria with their environment [1–4]. Cell surface proteins can have diverse roles such as mediating inter- and intra-species cell–cell interactions and cell-to-cell communication [5]. They participate in the usage of nutrients from the environment (e.g., cellulosomes) [6] and in binding and transport of different substances across the cell wall [7–9]. Furthermore, such proteins monitor environmental changes and are involved in chemical sensing [10]. Porin proteins allow both the passage of nutrients into the cells and the removal of toxic substances from the cells [11, 12]. Adherence, invasion, and colonization of hosts organs or tissues by pathogens often require presentation of crucial proteins on the surface of cells [13–16]. Since the surface proteome is an important cellular protein fraction, different methods have been developed to study this subproteome [17–20].

Planctomycetes possess a rigid cell wall made up mainly of proteins without any peptidoglycan or liposaccharides [21–23]. During their complex yeast-like lifestyle, processes like rosette-aggregation of cells or budding cell division involve cell–cell interactions [24]. Furthermore,

Correspondence: Professor Thomas Schweder, Pharmaceutical Biotechnologie, Ernst-Moritz-Arndt-University, F.-L.-Jahn-Straße 17, D-17487 Greifswald, Germany
E-mail: schweder@uni-greifswald.de
Fax: +49-3834-864238

*These authors contributed equally to this work.
**Current Address: Cao Xuan Hieu, Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

holdfast and appendage-like structures, possibly for adherence in the environment, extend from the bacterial cell surface of adult planctomycete cells [23]. Hence, the analysis of surface and cell-wall components may result in new insights into the development and physiology of planctomycetes.

Previous studies have revealed the protein sorting mechanisms of both Gram-positive and Gram-negative model bacteria [25–27]. The advanced cell compartmentalization of planctomycetes [28, 29] suggests that some novel protein sorting mechanisms may have evolved in these bacteria. Their cytoplasm is separated by a single intracellular membrane into two main compartments, the paryphoplasm and the pirellulosome. The pirellulosome, surrounded by this intracellular membrane, contains the DNA nucleoid. Since all ribosome-like particles are localized in the pirellulosome as well, sometimes linearly arrayed along invaginations of the intracytoplasmic membrane [30], translation likely occurs exclusively in this compartment. Thus, there must be a mechanism to target newly synthesized proteins through both the intracellular membrane and the cytoplasmic membrane to the cell surface.

Rhodopirellula baltica SH1^T is one of the best-studied planctomycetes on the proteomic level. This marine bacterium was isolated from the water column of the Kieler Bight, a bay in the southwestern Baltic Sea [23]. *R. baltica* is a marine aerobic, heterotrophic bacterium [28], which shows at least two different life styles, either as free-swimming or as sessile, attached-living cells [31]. Continuing our proteomic study [32] on *R. baltica* SH1^T, we used a "protein shaving" approach applying proteinase K and a biotinylation method to address the surface proteome of intact cells. We identified 52 potential surface proteins in both approaches among them nine proteins possibly involved in cell–cell interaction and eight proteins with a novel protein domain structure.

2 Materials and methods

2.1 Surface protein shaving by Proteinase K treatment

R. baltica SH1^T (DSM 10527) was grown aerobically in defined mineral M13 medium [33] with ammonium chloride and glucose as nitrogen and carbon sources, respectively. Bacterial cells in the exponential growth phase (OD_{600} 0.7) were harvested by centrifugation. The cells (0.1 g wet weight) were resuspended in PBS buffer and incubated with 10 μg of Proteinase K (Sigma Aldrich, St. Louis, MO, USA) on ice for 30 min. Cells were removed by centrifugation (6000 rpm, 4°C, 10 min). The supernatant, containing the shaved peptides, was mixed with 2 × SDS-PAGE sample buffer (0.15 M Tris (pH 6.8), 1.2% SDS, 30% glycerol, 15% β-mercaptoethanol, a trace bromophenol blue). 1D SDS-PAGE containing 12.5% acrylamide were done in a vertical gel apparatus (Bio-Rad Laboratories, Hercules, CA, USA or Hoefer Scientific Instruments, Holliston, MA, USA).

2.2 Biotinylation of surface protein and cell wall fraction

Biotinylation of putative surface proteins was done according to Hempel et al. [19]. Intact growing cells (0.1 g wet weight) of *R. baltica* were incubated with 0.7 μg Sulfo-NHS-Biotin (Pierce, Rockford, IL, USA) in PBS buffer for 1 h at 4°C. Residual biotin molecules were subsequently saturated by glycine PBS buffer. Bacterial cells were disrupted using the grinding kit (Sigma Aldrich) and sonication at 54 W for 5 min in order to achieve complete cell disruption. With the cell pellet, a cell wall preparation was carried out as described previously [32]. The biotinylated proteins in the supernatant were purified using NeutrAvidin agarose beads (Thermo Scientific, Waltham, MA, USA). The beads were collected by centrifugation (2000 rpm, 4°C, 2 min). After washing the beads seven times in PBS, surface proteins were released from the beads by adding 2 × SDS-PAGE sample buffer and incubation at 80°C for 5 min. The proteins were then separated by 1D SDS-PAGE.

2.3 Nano-HPLC-ESI MS/MS

Bands from the 1D SDS-PAGE gels were excised and the proteins digested with Trypsin (Sigma Aldrich). Peptides originated from in-gel digestion of 1D gels were separated by RP chromatography using a nano-ACQUITYTM UPLCTM System (Waters, Milford, MA, USA). 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 [20, 34]. Proteins were identified by searching all MS/MS spectra in ".dta" format against a *R. baltica* database (BX119912 version 4, 24-MAR-2007; http://www.uniprot.org/uniprot/?query=Rhodopirellula±baltica&sort=score) using Sorcerer™-SEQUEST® (Sequest v. 2.7 rev. 11, Thermo Scientific) including Scaffold_3.1.2 (Proteome Software, Portland, OR, USA) [35]. SEQUEST® was searched with a parent ion tolerance of 10 ppm and a fragment ion mass tolerance of 1.00 Da. Up to two missed tryptic cleavages were allowed. Methionine oxidation (+15.99492 Da) and cysteine carbamidomethylation (+57.021465 Da) were set as variable modifications. For protein identification, a stringent SEQUEST® filter for peptides was used (Xcorr versus charge state: 1.90 for singly, 2.5 for doubly, and 3.8 for triply charged ions and deltaCn value greater than 0.10) and at least two peptides per proteins were required. With this method all identified proteins had a protein probability score of 100%. Protein probabilities were assigned by the Protein Prophet algorithm.

2.4 YTV domain protein detection by Western blot

A 15-mer YTV domain-specific peptide (C-TYTVNVP YTETVEQT) was chemically synthesized (purity 80%) by the JPT Peptide Technologies (Berlin, Germany). After conjugation with hemocyanin from *Limulus polyphemus* (LPH), 2 mg of the YTV domain-specific peptide was used for immunization of mice (BioGenes, Berlin, Germany). Mice polyclonal antisera raised against YTV proteins were diluted 1:10 000 and used in Western blot. The secondary antibody (goat anti-mouse immunoglobulin G) conjugated with alkaline phosphatase was used at a 1:100 000 dilution (Sigma). Immunoblots were developed with nitroblue tetrazolium-5-bromo-4-chloro-3-indolylphosphate as a color substrate [36].

2.5 Genome annotation databases and sequence prediction servers

Pfam [37] and InterPro [38] databases were used for identification of protein domain families and motifs. The evolutionary analysis of 845 full length DUF1559 proteins (Pfam:PF07596) was inferred using the Maximum Likelihood method and the Jones–Taylor–Thornton (JTT) matrix-based model [39] that is provided in the software MEGA5 [40]. Aligned sequences of whole proteins from the Pfam database were used for this analysis. The bootstrap consensus tree was inferred from 500 replicates [41]. Branches corresponding to one species or to a taxonomic group that reproduced in less than 50% bootstrap replicates were collapsed. Initial tree(s) for the heuristic search were obtained automatically as follows. When the number of common sites was <100 or less than one-fourth of the total number of sites, the maximum parsimony method was used; otherwise neighbor-joining (BIONJ) method with Markov Cluster (MCL) distance matrix was used. The tree (Fig. 4 and Supporting Information Fig. S1) is drawn in scale, with branch lengths measured in the number of substitutions per site. All ambiguous positions were eliminated for each sequence pair. A total of 1887 positions were in the final dataset. The consensus sequence of the identified DUF1559 signatures was created by ClustalX sequence alignments and predicted using the Pratt' Pattern Matching server (http://www.ebi.ac.uk/Tools/pratt/). The HMM logo from the Pfam database of predicted signatures was used for Fig. 3. The proportion of identical residues between 8 identified DUF1559 sequences in the Pfam alignment was computed using the Sequence Identity Matrix function of BioEdit 7.1.3.0 (http://www.mbio.ncsu.edu/bioedit/bioedit.html).

Signal peptide and transmembrane helix prediction was done with the SignalP 3.0 software (http://www.cbs.dtu.dk/services/SignalP/) [42] and PSORTb v.2.0 Precomputed Genome software (http://www.psort.org/genomes/genomes.pl) for both classical and non-classical signal peptide types and a new motif (Pfam:PF07595) from Studholme et al. [43].

3 Results and discussion

3.1 Identification of cell surface proteins by proteomic approaches

In order to investigate the surfaceome of *R. baltica* SH1^T, we used two enrichment approaches, biotinylation and proteinase K treatment. In the first method, primary amine groups of surface-exposed proteins were labeled with Sulfo-NHS-biotin followed by an affinity purification step based on the strong noncovalent linkage between biotin and avidin agarose beads. The second strategy, shaving of the external peptides of cell wall-associated proteins by protease treatment, provided a complementary approach. MS resulted in the identification of 67 and 79 proteins in the biotinylation and proteinase K approach, respectively (Supporting Information Table S1). In both approaches, proteins containing signal peptides are enriched (30% and 32% in the biotinylation and proteinase K approach, respectively) compared to a protein analysis of a soluble whole cell extract (21% [32]). Fifty-two of the proteins identified in both approaches could be assigned as potential surface proteins because they either are predicted to contain signal peptides or membrane spanning domains or contain domains which are associated with cell surface or extracellular functions (Table 1, Fig. 1). Of these 52 proteins only six were found in both approaches suggesting that different subproteomes are targeted by the two methods (Fig. 2). Only proteins with large surface exposed domains are favored subjects of the proteinase K method. The average size of proteins identified with this method is 147 kDa, in comparison with 79 kDa in the biotinylation assay. Some very large proteins having molecular masses of more than 600 kDa were identified only by the proteinase K method (Table 1). On the other hand, it is possible that surface structures inhibited biotinylation of certain proteins thereby preventing their detection. Furthermore, it should be noted that the planctomycete cell wall structure is strong enough to resist high concentrations of detergents (e.g., SDS [21, 22]). Thus, in spite of being biotinylated, some proteins might have remained in the cell wall because of covalent linkage to the cell wall or tight attachment to it and thereby escaping detection.

Classical signal peptide sequences were predicted in 35 (67%) of the putative surface proteins (Table 1). Another ten proteins (19%), exclusively identified in the proteinase K experiment, contain an N-terminal domain, which represents a novel class of signal peptide motifs [43]. This new planctomycete-typical motif with the consensus sequence RRLxxExLExRxLLA has limited similarity to the twin arginine transporter (TAT) motif [44] as it contains two neighboring arginines. Identified proteins which contain this new signal peptide are quite large with the smallest protein having a molecular mass of 73 kDa and the other 9 proteins having masses between 155 and 768 kDa. Bioinformatics searches in sequenced prokaryotic genomes revealed that very large proteins in bacteria are often located on the cell surface [45].

Table 1. Potential surface proteins of *R. baltica*

Function	Accession number	Mr (kDa)	Biotin	Prot K	SP	New SP	TMHMM	pSortb	Gel
YTV proteins									
Putative uncharacterized protein	RB850	66	x		Y			<SIGNALP>	CW
Putative uncharacterized protein	RB2247	65	x		Y			<SIGNALP>	CW
Putative uncharacterized protein	RB5788	64	x		Y			<SIGNALP>	
Putative uncharacterized protein	RB7455	92	x		N			<Cytoplasm>	CW
Surface, cell attachment									
Probable surface-associated protein CshA	RB844	159		x	N	Yp		<SECRETOMEP>	
Fat protein-possibly involved in cell–cell attachment	RB886	207		x	N	Yp		<SECRETOMEP>	
Putative uncharacterized protein	RB5524	768		x	N	Yp		<Cytoplasm>	
Putative uncharacterized protein	RB6428	17		x	Y			<SIGNALP>	2D
Probable aggregation factor core protein MAFp3	RB6459	215		x	N	Yp		<Cytoplasm>	
Putative uncharacterized protein	RB7321	651		x	N	Yp		<Cytoplasm>	
Similar to surface-associated protein CshA	RB7341	797		x	Y	Yp		<SIGNALP>	
Probable membrane protein	RB9053	155		x	N	Yp		<Cytoplasm>	
Putative uncharacterized protein	RB12697	178		x	N	Yp		<Cytoplasm>	2D
Fimbria/prepilinDUF1559 proteins									
Putative uncharacterized protein	RB778	44		x	Y		1	<Fimbria>	
Putative uncharacterized protein	RB870	46	x	x	N		1	<SECRETOMEP>	
Putative uncharacterized protein	RB2782	64	x	x	N		1	<TMHMM>	2D
Putative uncharacterized protein	RB8194	44		x	Y		1	<Fimbria>	
Putative uncharacterized protein	RB8223	45		x	N		1	<Cytoplasm>	
Putative uncharacterized protein	RB10581	41	x	x	Y		1	<Fimbria>	CW
Putative uncharacterized protein	RB11041	38		x	Y		1	<SIGNALP>	
Putative uncharacterized protein	RB13319	45	x	x	N		1	<Cytoplasm>	
Enzymes									
Probable N-acetylglucosamine-6-phosphate deacetylase (NgaA)	RB977	131	x		Y			<SIGNALP>	2D
Probable NADH-dependent dehydrogenase	RB1555	52	x	x	Y			<TATP>	2D/CW
Putative hydrolase	RB3405	86	x	x	Y			<SIGNALP>	2D
Probable NADH-dependent dehydrogenase	RB5365	52	x		Y			<TATP>	2D
Probable cyclophilin type peptidylprolyl isomerase	RB6278	160		x	N	Yp		<Cytoplasm>	
Similar to serine/threonine protein kinase-related protein	RB7541	75	x		Y		1	<LIPOP>	
Similar to cycloartenol synthase	RB11008	43		x	Y			<SIGNALP>	2D
Molybdopterin oxidoreductase, iron–sulfur binding subunit	RB12666	131	x		Y			<SIGNALP>	2D
Transport/membrane proteins									
Similar to multidrug resistance protein (MexB)	RB2383	125	x		Y		14	<TMHMM>	
Protein-export membrane protein (SecD)	RB4633	119		x	N		13	<TMHMM>	
Hypothetical									
Probable polymorphic membrane protein B/C family	RB577	73		x	N	Yp		<Cytoplasm>	
Putative uncharacterized protein	RB2012	55	x		Y			<SIGNALP>	
Probable outer membrane lipoprotein (IbeB)	RB2381	68	x		Y			<SIGNALP>	
Putative uncharacterized protein	RB2901	43		x	Y		1	<SIGNALP>	2D
Putative uncharacterized protein	RB3077	629		x	N	Yp		<Cytoplasm>	
Putative uncharacterized protein	RB3899	70	x		Y			<SIGNALP>	
Putative uncharacterized protein	RB4256	37	x		Y			<SIGNALP>	2D/CW

Table 1. Continued

Function	Accession number	Mr (kDa)	Biotin	Prot K	SP	New SP	TMHMM	pSortb	Gel
Putative uncharacterized protein	RB4703	78	x		Y		1	<SIGNALP>	2D
Putative large multifunctional protein	RB5608	50	x		Y			<SIGNALP>	2D
Putative uncharacterized protein	RB5657	15		x	Y			<SIGNALP>	CW
Putative uncharacterized protein	RB6221	113	x		Y			<SIGNALP>	2D
Putative uncharacterized protein	RB8039	177	x		Y			<SIGNALP>	
Putative uncharacterized protein	RB8055	44	x		Y			<SIGNALP>	
Putative uncharacterized protein	RB9101	20		x	Y			<SIGNALP>	2D
Putative uncharacterized protein	RB9201	229	x		Y		1	<SIGNALP>	
Putative uncharacterized protein	RB9261	46	x		Y			<SIGNALP>	2D
Putative uncharacterized protein	RB10300	116	x		Y			<SIGNALP>	
Putative uncharacterized protein	RB10666	205		x	N	Yp		<Cytoplasm>	
Putative uncharacterized protein	RB12461	61		x	Y		1	<SIGNALP>	2D
Putative uncharacterized protein	RB12634	18		x	Y			<SIGNALP>	2D
Putative uncharacterized protein	RB12700	49	x		Y			<SIGNALP>	

Biotin, protein found in biotinylation experiment; Prot K, protein found in proteinase K experiment; SP, signal peptide prediction; new SP, newly proposed SP [43]; TMHMM, number of predicted transmembrane helices; pSortb, predicted locations by PSORTb v.2.0; Gel, protein was identified in earlier proteome analysis [31], 2D, protein found in 2D gel analysis; CW, protein found in cell wall protein analysis.

3.2 Cell adhesion proteins and a novel protein structure combining a type IV pilin domain and a planctomycete typical domain DUF1559

Analysis of the proteinase K-treatment-enriched *R. baltica* cell surface protein fraction allowed the identification of several candidate proteins for cell–cell interactions. Some of these proteins exhibited similarities with surface proteins from other organisms; all of them contain domains that have been shown to be typical for surface proteins like cadherin/dockerin domains (Pfam: PF00404, PF00028), a bacterial adhesion domain (Pfam: PF05738) or the fasciclin domain (Pfam: PF02469). Except for the fasciclin domain containing protein (RB6428), the proteins are more than 100 kDa in size (Table 1). All proteins in this group, aside from the fasciclin, contain nonclassical planctomycete specific signal peptides [43]. Fasciclin domains are present in proteins known to mediate cell adhesion [46]. Proteins containing fasciclin domains have also been found in several bacteria among them a symbiotically living cyanobacterium where it was associated with symbiosis [47]. One protein (RB9053) contains a "bacterial adhesion" domain found in adhesive pili [48]. In *Staphylococcus epidermidis* a protein containing such domains mediates binding to human fibrinogen [49]. Proteins containing cadherin and/or dockerin domains were originally implicated in formation of cellulosomes [6]. Cellulosomes are complex nanomachines assembled of enzymes involved in degradation of plant cell walls. In recent years, such domains have been more and more found in proteins that have no glycoside hydrolase function [50]. For some of these proteins, involvement in cell–cell adhesion has been supposed [51]. le Duc et al. [52] demonstrated that E-cadherin in human cell lines functions as a mechanosensor at cell–cell junctions. Many proteins with cadherin/dockerin modules contain multiple domains. The *R. baltica* cadherin/dockerin proteins, however, contain mostly only one repeat of the domains.

Further characterization and comparative profiling may provide new insights into the functions of these proteins in surface attachment, cell aggregation, or phage

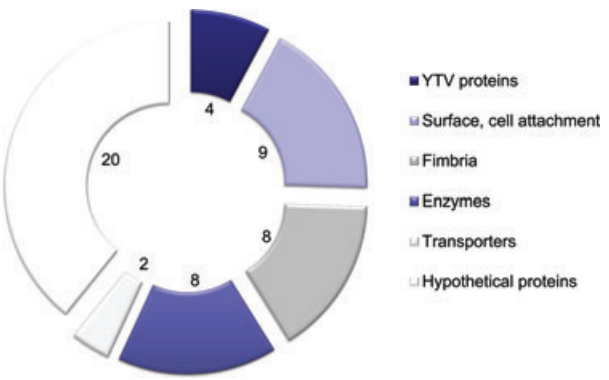

Figure 1. Functional classification of identified potential cell surface proteins based on the genome annotation.

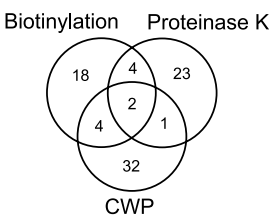

Figure 2. Comparison of the number of potential surface proteins identified by biotinylation and proteinase K treatment with the previously published cell wall proteome (CWP) [32].

Table 2. The occurrence of 845 proteins containing DUF1559 domain in 12 species.

Species tree[a)]	1 × DUF1559			2 × DUF1559		3 × DUF1559
	1	2	3	4	5	6
+−Planctomycetes//Planctomycetales (575)	1					
| |	1					
| +−Candidatus *Kuenenia stuttgartiensis* (1)						
| |						
| +−Uncultured planctomycete 8FN (1)						
| |						
| +−Planctomycetaceae (573)						
| |						
| +−*Pirellula staleyi ATCC 27377* (70)	52	17			1	
| |						
| +−*Rhodopirellula baltica* SH 1 (74)	49[b)]	21	1	1	2	
| |						
| +−*Planctomyces limnophilus* DSM 3776 (80)	60	17			3	
| |						
| +−*Planctomyces maris* DSM 8797 (142)	101	33	1		5	2
| |						
| +−*Blastopirellula marina* DSM 3645 (207)	152	52		1	1	1
|						
+−Lentisphaerae (168)						
| |						
| +−*Lentisphaera araneosa* HTCC2155 (49)	33	15	1			
| |						
| +−*Victivallis vadensis* ATCC BAA-548 (119)	74	45				
|						
+−Verrucomicrobia (101)						
| |						
| +−*Opitutaceae bacterium* TAV2 (58)	55	3				
| |						
| +−*Pedosphaera parvula* Ellin514 (43)	38	5				
|						
+−Firmicutes//Bacillaceae (1)						
| |						
| +−*Bacillus thuringiensis* T13001 (1)		1				

All proteins were categorized into the following domain structures: (1) prepilin domain DUF1559; (2) DUF1559; (3) enzymatic domain DUF1559; (4) prepilin domain 2× DUF1559; (5) 2× DUF1559; (6) 3 × DUF1559.
a) The number of DUF1559 proteins identified in the genome are shown in brackets.
b) All eight mass spectrometrically identified DUF1559 proteins in this study have the "prepilin domain-DUF1559" structure.

resistance of planctomycetes. When the population density rises above a certain level, planctomycete cells aggregate to form rosettes [53]. A switch between swarmer and attached cells was also found to be caused by changing environmental conditions (ammonium limitation [31]). In native habitats, adult *R. baltica* cells attach firmly onto snow particles by a unique hold stalk structure [54]. Furthermore, transcriptomic analysis of the life cycle of *R. baltica* indicated that the cell wall is modified when the physiological state of the cells changes [55]. Genes involved in reorganization of the cell envelope were found to be regulated during life cycle.

Studholme et al. [43] found in a study of planctomycete-specific protein domains in 74 proteins a novel domain, DUF1559 (Pfam: PF07596), that showed a weak (13%) similarity to the *Chlamydophila caviae* solute binding protein CCA00262 (Uniprot:Q823Z1). Interestingly, all eight DUF1559 proteins mass spectrometrically identified in our study (Table 1) also exhibit a prepilin domain (Pfam: PF07963, InterPro: IPR012902). This domain is a class III signal peptide especially for bacterial type IV pilins [56] and archaeal flagellins [57]. Several proteins of the Gram-negative bacteria-like pilin and flagella systems such as CpaB, PilB, PilT, PilM, FliC, and FliD were identified in a previous proteomic investigation of *R. baltica* [32]. The DUF1559 proteins might contribute to such structures as pilins or pseudopillins or even be components of various other structures on the cell surface of *R. baltica* such as the holdfast or crateriform.

According to the Pfam database (date access June 2011), the combination of a DUF1559 and a Prepilin domain is the most common structure of proteins containing the DUF1559

Figure 3. Five consensus protein motifs of identified Prepilin-DUF1559 domain proteins based on the RB879 sequence as an example. P, the prepilin/N-methyl domain (IPR012902, PF07963). I–IV, four motifs of the domain DUF1559 (InterPro:IPR011453, Pfam:PF07596).

domain in *Planctomycete* species for which genome sequences are available (Table 2). This domain organization is also found in the genomes of *Lentisphaerae* and *Verrucomicrobia* species, which is consistent with their close phylogenetic relationship to the *Planctomycetes* [58]. Based on the sequence alignment of the eight identified Prepilin-DUF 1559 proteins, we propose five conserved motifs in these proteins, one for the Prepilin domain and four other for the DUF1559 domains (Fig. 3). Our phylogenetic analysis of all known proteins containing the DUF1559 domain suggests that this protein family is an interesting case study for adaptive functional radiation with high rates of mutation and duplication events (Fig. 4 and Supporting Information Fig. S1). The eight proteins identified in the cell surface proteome grouped into different clades with orthologous proteins of other *Planctomycete* species in the gene tree created by the Maximum Likelihood method. Most of the nonplanctomycete DUF1559 proteins belong to one main evolutionary group. That group also contains the only representative DUF1559 protein of the "anammox planctomycete" candidatus *Kuenenia stuttgartiensis*. This supports the idea that this species branched off early in the planctomycete phylum, and might even be recognized as an independent phylum due to its unique and specialized physiological characteristics [59]. The protein sequences of RB11041 (Q7UJV0) and RB2782 (Q7UV98) that belong to this "ancient" clade differ drastically from other identified Prepillin-DUF1559 proteins (12.2–26.3% identity). Meanwhile, the other six of eight identified Prepillin-DUF1559 proteins clustered into a large clade including 27 DUF1559 containing proteins (Supporting Information Fig. S1). This clade shows nonmonophyletic supported subgroups as indicated by high bootstrap values. It would be interesting to further investigate the biological functions of those closely related DUF1559 proteins.

3.3 Enzymes are components of the cell surface proteome

Eight of the identified putative surface proteins are annotated as enzymes and some more exhibit enzymatic protein domains (Fig. 1). With one exception all enzymes contain classical signal peptide sequences (Table 1). The N-acetylglucosamine-6-phosphate deacetylase NagA_3 (RB977) and the diverse other enzymes may contribute to conversion of various polymer compounds. The finding of enzymes perhaps attached to the cell surface, supports our suggestion that *R. baltica* may anchor some degradative enzymes onto their cell surface, because no such enzymes were detected in the culture medium [32]. Additional degradative enzymes may be localized in the paryphoplasm or even in specialized vesicle-like membrane-surrounded compartments in *R. baltica* [23, 60].

Furthermore, some molecular chaperones and protein processing enzymes were detected in the cell surface proteome (Supporting Information Table S1). Although none of these proteins contains a signal peptide, they might still be secreted and attached to the cell surface as has been described for various species [61–63]. These enzymes might play a key role in the surface protein translocation and folding [36, 64]. Additionally, preprotein translocases (SecA, SecD; Supporting Information Table S1) and a protein with similarities to outer membrane lipoproteins (IbeB; Table 1) were found.

Several enzymes identified in the surface proteome are also found abundantly in the soluble intracellular fraction [32], such as TufA (Supporting Information Table S1). These proteins could occur in the surface protein fraction due to the cell lysis. However, the presence of supposedly intracellular enzymes, including GapA, Eno, and elongation factors, is quite common on cell surfaces of microorganisms [61–63]. Such anchorless surface proteins, lacking signal peptides or hydrophobic membrane spanning domains, were indicated to facilitate bacterial colonization or cell adherence [65–67].

3.4 The presence of YTV proteins in the SDS-resistant protein cell wall fraction

In our previous study on the cell wall-associated protein fraction, YTV proteins were highlighted by their unique sequence and amino acid composition that is consistent with the

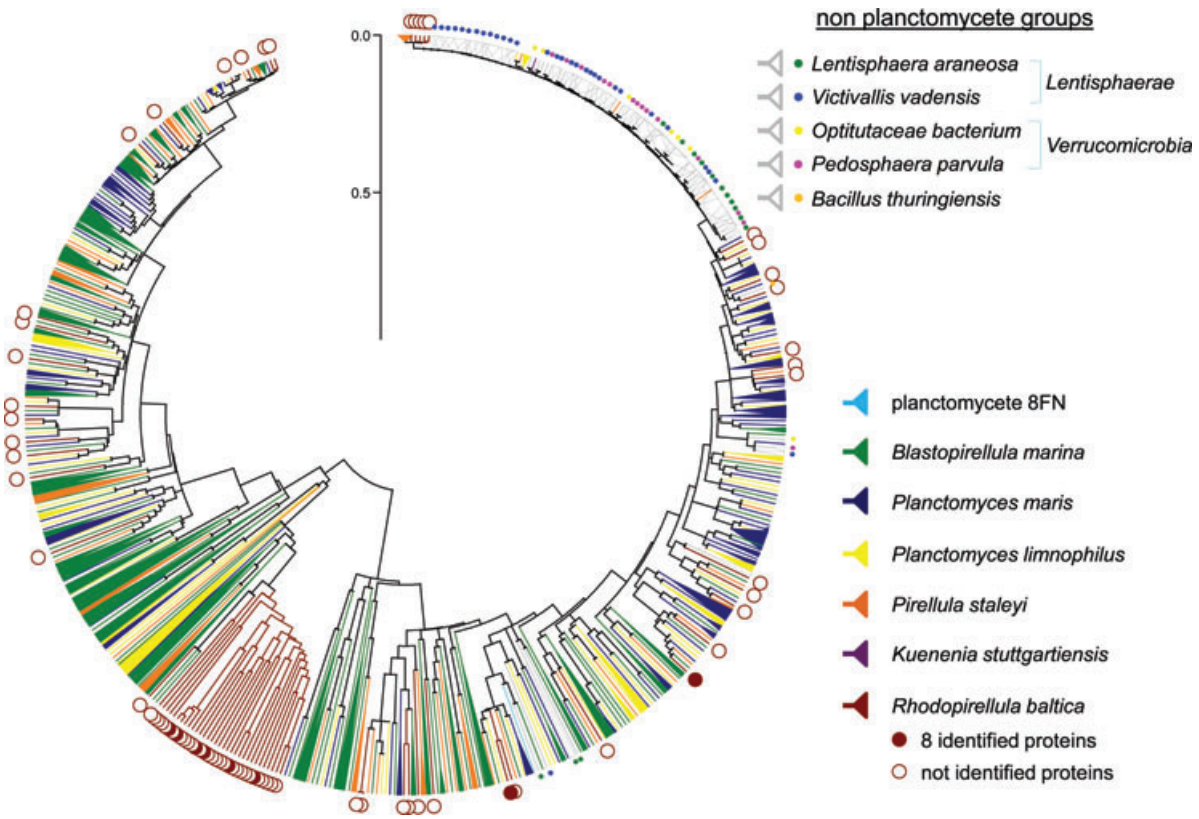

Figure 4. Evolutionary relationship of 845 full-length DUF1559 domain containing proteins (Pfam:PF07596) constructed by the Maximum Likelihood method and the Jones–Taylor–Thornton (JTT) matrix-based model (see Section 2). The consensus tree was inferred from 500 bootstrap replicates. Branches corresponding to one species or to a taxonomic group (except for *Rhodopirellula baltica*), which reproduced in less than 50% bootstrap replicates, are collapsed and colored as indicated in the legend. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The main clade of 27 *R. baltica* proteins (including their corresponding UniProt accession numbers) with bootstrap values above 50% is separately shown in Supporting Information Fig. S2. Further detailed information about bootstrap values and sequence annotations can be found in the Supporting Information Fig. S2.

chemical composition of the cell wall sacculi [32]. In this investigation, all three earlier identified and one additional YTV protein were detected in the biotinylated protein fraction in the most prominent 1-D gel band with more than 170 kDa in size (Fig. 5). None of the YTV proteins, however, has a molecular weight higher than 100 kDa. Likely, the YTV proteins are forming a complex that does not depolymerize during the SDS-sample buffer treatment and SDS-PAGE separation. A monoclonal antibody against a consensus 15mer peptide specific to YTV protein signatures (Pfam: PF07639) from the available complete genomes of planctomycetes and metagenomic sequences was raised. The Western blot results confirmed the presence of YTV domain containing proteins in the SDS resistant protein band (Fig. 5).

4 Concluding remarks

In this study, we present proteomic data from two cell surface protein enrichment experiments. Cell surface proteins in this regard mean either proteins anchored onto the cell wall or the cytoplasmic membrane and exposed to the environment or proteins that are components of the cell wall. Ultimately, we added 29 newly identified proteins to our previous extensive proteome analysis [32] and indicated the cellular location of a total of 52 proteins. However, further detailed proteome studies and/or other complement verification methods, such as electron or confocal microscope imaging, are needed to verify localization. Most of the identified potential surface proteins were predicted to possess signal peptide sequences and/or transmembrane spanning domains. Among them, nine proteins were found to be similar to known cell surface structure proteins or containing cell adhesion domains.

Since the DUF1559 domain is the top eighth largest InterPro domain of a hypothetical *R. baltica* proteome, the redundancy of this protein family has likely contributed to enlarge the planctomycete genomes. Our assumption that proteins with both a Prepilin and a DUF1559 domain form the planctomycete fimbriae structure needs further experimental data to prove.

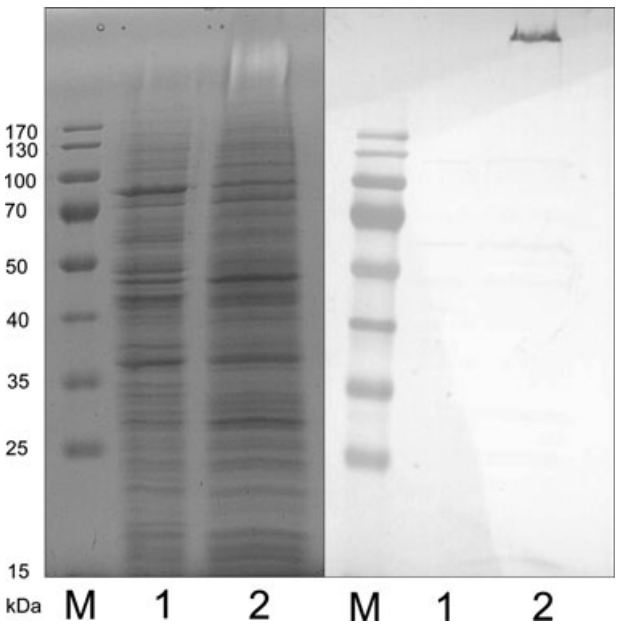

Figure 5. SDS-PAGE (left) and Western blot with YTV-antiserum (Pfam: PF07639) (right) of the intracellular (1) and biotinylated protein fraction (2).

C.X.H. was supported by the Vietnamese Ministry of Education and Training (MOET) and the Deutscher Akademischer Austausch Dienst (DAAD). This work was supported by the Bundesministerium für Bildung und Forschung (BMBF, grant number 03F0364B), the TBI of Mecklenburg-Vorpommern (AZ 230-630.8-TIFA-356), and the Max Planck Society.

The authors have declared no conflict of interest.

5 References

[1] Marraffini, L. A., Dedent, A. C., Schneewind, O., Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* 2006, *70*, 192–221.

[2] Scott, J. R., Barnett, T. C., Surface proteins of gram-positive bacteria and how they get there. *Annu. Rev. Microbiol.* 2006, *60*, 397–423.

[3] Albers, S. V., Szabo, Z., Driessen, A. J., Protein secretion in the Archaea: multiple paths towards a unique cell surface. *Nat. Rev. Microbiol.* 2006, *4*, 537–547.

[4] Engelhardt, H., Are S-layers exoskeletons? The basic function of protein surface layers revisited. *J. Struct. Biol.* 2007, *160*, 115–124.

[5] Senadheera, D., Cvitkovitch, D. G., Quorum sensing and biofilm formation by *Streptococcus mutans*. *Adv. Exp. Med. Biol.* 2008, *631*, 178–188.

[6] Gilbert, H. J., Cellulosomes: microbial nanomachines that display plasticity in quaternary structure. *Mol. Microbiol.* 2007, *63*, 1568–1576.

[7] Dramsi, S., Magnet, S., Davison, S., Arthur, M., Covalent attachment of proteins to peptidoglycan. *FEMS Microbiol. Rev.* 2008, *32*, 307–320.

[8] Zolghadr, B., Weber, S., Szabo, Z., Driessen, A. J., Albers, S. V., Identification of a system required for the functional surface localization of sugar binding proteins with class III signal peptides in *Sulfolobus solfataricus*. *Mol. Microbiol.* 2007, *64*, 795–806.

[9] Fraiberg, M., Borovok, I., Weiner, R. M., Lamed, R., Discovery and characterization of cadherin domains in *Saccharophagus degradans* 2–40. *J. Bacteriol.* 2010, *192*, 1066–1074.

[10] Lopez, D., Kolter, R., Extracellular signals that define distinct and coexisting cell fates in *Bacillus subtilis*. *FEMS Microbiol. Rev.* 2010, *34*, 134–149.

[11] Gibson, K., Kumagai, Y., Rikihisa, Y., Proteomic analysis of *Neorickettsia sennetsu* surface-exposed proteins and porin activity of the major surface protein P51. *J. Bacteriol.* 2010, *192*, 5898–5905.

[12] Koskenniemi, K., Laakso, K., Koponen, J., Kankainen, M. et al., Proteomics and transcriptomics characterization of bile stress response in probiotic *Lactobacillus rhamnosus* GG. *Mol. Cell Proteomics* 2011, *10*, M110 002741.

[13] Hauck, C. R., Ohlsen, K., Sticky connections: extracellular matrix protein recognition and integrin-mediated cellular invasion by *Staphylococcus aureus*. *Curr. Opin. Microbiol.* 2006, *9*, 5–11.

[14] Noh, S. M., Brayton, K. A., Brown, W. C., Norimine, J. et al., Composition of the surface proteome of *Anaplasma marginale* and its role in protective immunity induced by outer membrane immunization. *Infect. Immun.* 2008, *76*, 2219–2226.

[15] Ventura, C. L., Malachowa, N., Hammer, C. H., Nardone, G. A. et al., Identification of a novel *Staphylococcus aureus* two-component leukotoxin using cell surface proteomics. *PLoS One* 2010, *5*, e11634.

[16] Bohle, L. A., Riaz, T., Egge-Jacobsen, W., Skaugen, M. et al., Identification of surface proteins in *Enterococcus faecalis* V583. *BMC Genomics* 2011, *12*, 135.

[17] Cordwell, S. J., Technologies for bacterial surface proteomics. *Curr. Opin. Microbiol.* 2006, *9*, 320–329.

[18] Rodriguez-Ortega, M. J., Norais, N., Bensi, G., Liberatori, S. et al., Characterization and identification of vaccine candidate proteins through analysis of the group A streptococcus surface proteome. *Nat. Biotechnol.* 2006, *24*, 191–197.

[19] Hempel, K., Pane-Farre, J., Otto, A., Sievers, S. 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, *9*, 1579–1590.

[20] Otto, A., Bernhardt, J., Meyer, H., Schaffer, M. et al., Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat. Commun.* 2010, *1*, 137.

[21] König, H., Schlesner, H., Hirsch, P., Cell wall studies on budding bacteria of the *Planctomyces* αιρυετσαII group and on a *Prosthecomicrobium* sp. *Arch. Microbiol.* 1984, *138*, 200–205.

[22] Liesack, W., Koenig, E., Schlesner, H., Hirsch, P., Chemical composition of the peptidoglycan-free cell envelopes of budding bacteria of the *Pirella*/*Planctomyces* group. *Arch. Microbiol.* 1986, *145*, 361–366.

[23] Schlesner, H., Rensmann, C., Tindall, B. J., Gade, D. et al., Taxonomic heterogeneity within the *Planctomycetales* as derived by DNA–DNA hybridization, description of *Rhodopirellula baltica* gen. nov., sp. nov., transfer of *Pirellula marina* to the genus *Blastopirellula* gen. nov. as *Blastopirellula marina* comb. nov. and emended description of the genus *Pirellula*. *Int. J. Syst. Evol. Microbiol.* 2004, *54*, 1567–1580.

[24] Ward, N., Staley, J. T., Fuerst, J. A., Giovannoni, S. et al., in: Dworkin, M. (Ed.), *The Prokaryotes: A Handbook on the Biology of Bacteria*, 3rd ed., Springer, New York, NY 2006, pp. 757–793.

[25] 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.: MMBR* 2004, *68*, 207–233.

[26] Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P. et al., Protein secretion and outer membrane assembly in *Alphaproteobacteria*. *FEMS Microbiol. Rev.* 2008, *32*, 995–1009.

[27] Weiner, E. M., Robson, S., Marohn, M., Clubb, R. T., The Sortase A enzyme that attaches proteins to the cell wall of *Bacillus anthracis* contains an unusual active site architecture. *J. Biol. Chem.* 2010, *285*, 23433–23443.

[28] Glockner, F. O., Kube, M., Bauer, M., Teeling, H. et al., Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 8298–8303.

[29] Fuerst, J. A., Intracellular compartmentation in planctomycetes. *Annu. Rev. Microbiol.* 2005, *59*, 299–328.

[30] Lindsay, M. R., Webb, R. I., Strous, M., Jetten, M. S. et al., Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. *Arch. Microbiol.* 2001, *175*, 413–429.

[31] Frank, C. S., Langhammer, P., Fuchs, B. M., Harder, J., Ammonium and attachment of *Rhodopirellula baltica*. *Arch. Microbiol.* 2011, *193*, 365–372.

[32] Hieu, C. X., Voigt, B., Albrecht, D., Becher, D. et al., Detailed proteome analysis of growing cells of the planctomycete *Rhodopirellula baltica* SH1T. *Proteomics* 2008, *8*, 1608–1623.

[33] Schlesner, H., The development of media suitable for the microorganisms morphologically resembling *Planctomyces* spp., *Pirellula* spp., and other *Planctomycetales* from various aquatic habitats using dilute media. *Syst. Appl. Microbiol.* 1994, *17*, 135–145.

[34] 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.

[35] Nesvizhskii, A. I., Keller, A., Kolker, E., Aebersold, R., A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 2003, *75*, 4646–4658.

[36] Gerth, U., Kirstein, J., Mostertz, J., Waldminghaus, T. et al., Fine-tuning in regulation of Clp protein content in *Bacillus subtilis*. *J. Bacteriol.* 2004, *186*, 179–191.

[37] Finn, R. D., Tate, J., Mistry, J., Coggill, P. C. et al., The Pfam protein families database. *Nucleic Acids Res.* 2008, *36*, D281–D288.

[38] Mulder, N. J., Kersey, P., Pruess, M., Apweiler, R., In silico characterization of proteins: UniProt, InterPro and Integr8. *Mol. Biotechnol.* 2008, *38*, 165–177.

[39] Jones, D. T., Taylor, W. R., Thornton, J. M., The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 1992, *8*, 275–282.

[40] Tamura, K., Peterson, D., Peterson, N., Stecher, G. et al., MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 2011, *28*, 2731–2739.

[41] Felsenstein, J., Confidence-limits on phylogenies—an approach using the bootstrap. *Evolution* 1985, *39*, 783–791.

[42] Bendtsen, J. D., Nielsen, H., von Heijne, G., Brunak, S., Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 2004, *340*, 783–795.

[43] Studholme, D. J., Fuerst, J. A., Bateman, A., Novel protein domains and motifs in the marine planctomycete *Rhodopirellula baltica*. *FEMS Microbiol. Lett.* 2004, *236*, 333–340.

[44] Berks, B. C., Palmer, T., Sargent, F., The Tat protein translocation pathway and its role in microbial physiology. *Adv. Microb. Physiol.* 2003, *47*, 187–254.

[45] Reva, O., Tummler, B., Think big–giant genes in bacteria. *Environ. Microbiol.* 2008, *10*, 768–777.

[46] Kohsaka, H., Takasu, E., Nose, A., In vivo induction of postsynaptic molecular assembly by the cell adhesion molecule Fasciclin2. *J. Cell Biol.* 2007, *179*, 1289–1300.

[47] Paulsrud, P., Lindblad, P., Fasciclin domain proteins are present in nostoc symbionts of lichens. *Appl. Environ. Microbiol.* 2002, *68*, 2036–2039.

[48] Buts, L., Bouckaert, J., De Genst, E., Loris, R. et al., The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds N-acetylglucosamine. *Mol. Microbiol.* 2003, *49*, 705–715.

[49] Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S. et al., A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell* 2003, *115*, 217–228.

[50] Peer, A., Smith, S. P., Bayer, E. A., Lamed, R., Borovok, I., Noncellulosomal cohesin- and dockerin-like modules in the three domains of life. *FEMS Microbiol. Lett.* 2009, *291*, 1–16.

[51] Conrady, D. G., Brescia, C. C., Horii, K., Weiss, A. A. et al., A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. *Proc. Natl. Acad. Sci. USA* 2008, *105*, 19456–19461.

[52] le Duc, Q., Shi, Q., Blonk, I., Sonnenberg, A. et al., Vinculin potentiates E-cadherin mechanosensing and is recruited to actin-anchored sites within adherens junctions in a myosin II-dependent manner. *J. Cell Biol.* 2010, *189*, 1107–1115.

[53] Neef, A., Amann, R., Schlesner, H., Schleifer, K. H., Monitoring a widespread bacterial group: in situ detection of planctomycetes with 16S rRNA-targeted probes. *Microbiology* 1998, *144*(Pt 12), 3257–3266.

[54] Gade, D., Stuhrmann, T., Reinhardt, R., Rabus, R., Growth phase dependent regulation of protein composition in *Rhodopirellula baltica*. *Environ. Microbiol.* 2005, *7*, 1074–1084.

[55] Wecker, P., Klockow, C., Schuler, M., Dabin, J. et al., Life cycle analysis of the model organism *Rhodopirellula baltica* SH 1(T) by transcriptome studies. *Microb. Biotechnol.* 2010, *3*, 583–594.

[56] Hansen, J. K., Forest, K. T., Type IV pilin structures: insights on shared architecture, fiber assembly, receptor binding and type II secretion. *J. Mol. Microbiol. Biotechnol.* 2006, *11*, 192–207.

[57] Ng, S. Y., Chaban, B., Jarrell, K. F., Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J. Mol. Microbiol. Biotechnol.* 2006, *11*, 167–191.

[58] Wagner, M., Horn, M., The *Planctomycetes*, *Verrucomicrobia*, *Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 2006, *17*, 241–249.

[59] Fuerst, J. A., Sagulenko, E., Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. *Nat. Rev. Microbiol.* 2011, *9*, 403–413.

[60] van Niftrik, L., Geerts, W. J., van Donselaar, E. G., Humbel, B. M. et al., Linking ultrastructure and function in four genera of anaerobic ammonium-oxidizing bacteria: cell plan, glycogen storage, and localization of cytochrome C proteins. *J. Bacteriol.* 2008, *190*, 708–717.

[61] Tjalsma, H., Lambooy, L., Hermans, P. W., Swinkels, D. W., Shedding and shaving: disclosure of proteomic expressions on a bacterial face. *Proteomics* 2008, *8*, 1415–1428.

[62] Singh, V. K., Jayaswal, R. K., Wilkinson, B. J., Cell wall-active antibiotic induced proteins of *Staphylococcus aureus* identified using a proteomic approach. *FEMS Microbiol. Lett.* 2001, *199*, 79–84.

[63] Gozalbo, D., Gil-Navarro, I., Azorin, I., Renau-Piqueras, J. et al., The cell wall-associated glyceraldehyde-3-phosphate dehydrogenase of *Candida albicans* is also a fibronectin and laminin binding protein. *Infect. Immun.* 1998, *66*, 2052–2059.

[64] Liu, Y., Fu, X., Shen, J., Zhang, H. et al., Periplasmic proteins of *Escherichia coli* are highly resistant to aggregation: reappraisal for roles of molecular chaperones in periplasm. *Biochem. Biophys. Res. Commun.* 2004, *316*, 795–801.

[65] Chhatwal, G. S., Anchorless adhesins and invasins of Gram-positive bacteria: a new class of virulence factors. *Trends Microbiol.* 2002, *10*, 205–208.

[66] 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*, 1273–1287.

[67] 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.

