Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 1934–1951
Carbohydrate RESEARCH

Review

S-layer nanoglycobiology of bacteria

Paul Messner,* Kerstin Steiner, Kristof Zarschler and Christina Schäffer*

Universität für Bodenkultur Wien, Zentrum für NanoBiotechnologie A-1180 Wien, Gregor-Mendel-Strasse 33, Austria

Received 2 October 2007; received in revised form 5 December 2007; accepted 20 December 2007
Available online 16 January 2008

Presented at Eurocarb 14th Lübeck, Germany, September 2007

Abstract—Cell surface layers (S-layers) are common structures of the bacterial cell envelope with a lattice-like appearance that are formed by a self-assembly process. Frequently, the constituting S-layer proteins are modified with covalently linked glycan chains facing the extracellular environment. S-layer glycoproteins from organisms of the *Bacillaceae* family possess long, O-glycosidically linked glycans that are composed of a great variety of sugar constituents. The observed variations already exceed the display found in eukaryotic glycoproteins. Recent investigations of the S-layer protein glycosylation process at the molecular level, which has lagged behind the structural studies due to the lack of suitable molecular tools, indicated that the S-layer glycoprotein glycan biosynthesis pathway utilizes different modules of the well-known biosynthesis routes of lipopolysaccharide O-antigens. The genetic information for S-layer glycan biosynthesis is usually present in S-layer glycosylation (*slg*) gene clusters acting in concert with housekeeping genes. To account for the nanometer-scale cell surface display feature of bacterial S-layer glycosylation, we have coined the neologism 'nanoglycobiology'. It includes structural and biochemical aspects of S-layer glycans as well as molecular data on the machinery underlying the glycosylation event. A key aspect for the full potency of S-layer nanoglycobiology is the unique self-assembly feature of the S-layer protein matrix. Being aware that in many cases the glycan structures associated with a protein are the key to protein function, S-layer protein glycosylation will add a new and valuable component to an 'S-layer based molecular construction kit'. In our long-term research strategy, S-layer nanoglycobiology shall converge with other functional glycosylation systems to produce 'functional' S-layer *neo*glycoproteins for diverse applications in the fields of nanobiotechnology and vaccine technology. Recent advances in the field of S-layer nanoglycobiology have made our overall strategy a tangible aim of the near future.

Keywords: S-layer glycoprotein; Glycosylation enzymes; Glycosylation gene cluster; Nanobiotechnology; Carbohydrate engineering; Self-assembly

Contents

* Corresponding authors. Tel.: +43 1 47654 2202; fax: +43 1 478 9112 (P.M.); e-mail addresses: paul.messner@boku.ac.at; christina.schaeffer@boku.ac.at

0008-6215/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2007.12.025

1. Introduction

1.1. The S-layer protein self-assembly system

Prokaryotic surface layers (S-layers) have been described for the first time time in 1952. We now approach the sixth decade of S-layer research.[1,2] S-layers, in general, are two-dimensional crystalline arrays formed of individual subunits by an entropy-driven self-assembly event, which completely cover bacterial cells.[3] Over time, we faced a considerable change in the focus of S-layer research, accounting for novel areas and demands opening up in conjunction with applied research. At the very beginning of S-layer research, identification and ultrastructural characterization of S-layers from prokaryotes of different habitats and sources were the main tasks (for reviews see Refs. 3–6), since the mid 1970s chemical characterization of archaeal as well as of bacterial S-layer proteins became the central research focus (for reviews, see Refs. 7–15). The mid-1970s were also the time when glycosylation of S-layers had been discovered.[16,17] Almost 30 years later, it is now established knowledge that glycosylation represents the major modification of S-layer proteins, both in archaea and in bacteria.[18] (Please note that the term archaebacteria was replaced by archaea and the term eubacteria was replaced by bacteria[19]).

In 1986, the first S-layer gene sequences encoding the two S-layer proteins of *Brevibacillus brevis* 47 have been published.[20] Now more than 50 sequences of S-layer genes are available in public data bases, with ten of them encoding S-layer proteins with confirmed glycosylation. These are the S-layer proteins of the archaea *Haloarcula japonica* TR-1 (GenBank D87290), *Halobacterium salinarum* R_1M_1 (GenBank J02767), *Haloferax volcanii* DS2 (GenBank M62816), *Methanothermus fervidus* DSM 2088 (GenBank X58297), *Methanothermus sociabilis* DSM 3496 (GenBank X58296), and of the bacteria *Geobacillus stearothermophilus* NRS 2004/3a (GenBank AF328862), *Aneurinibacillus thermoaerophilus* DSM 10155/G^+ (GenBank AY395579), *A. thermoaerophilus* L420-91^T (GenBank AY395578), *Geobacillus tepidamans* GS5-97^T(GenBank AY883421), and *Thermoanaerobacter kivui* DSM 2030 (GenBank M31069). With the demonstration of the feasibility of recombinant S-layer protein production in heterologous expression systems, new avenues for S-layer protein research have opened up, putting forward the use of S-layer protein self-assembly systems for a wide spectrum of applications (for reviews, see Refs. 21–24). Regardless of the source of the S-layer, either after isolation from the bacterial cell wall by treatment with chaotropic agents or after heterologous expression in a suitable host, S-layer subunits characteristically self-assemble into monomolecular, two-dimensional arrays with oblique (p1, p2), square (p4) or hexagonal (p3, p6) symmetry, which are identical with those observed on the respective native bacterial cell.[25] These so-called self-assembly structures can have flat, cylindrical or even vesicular appearance, be either mono- or multilayered, and have variable dimensions (up to several μm^2). Usually, there is an anisotropical charge distribution and considerable differences in the surface topography of the self-assembly structures. These are formed either in suspension, on solid supports (e.g., gold chips, silicon wafers, plastic materials), on air-liquid interfaces, on liposomes or on lipid layers.[24,26] Due to the spacing between the subunits within the two-dimensional S-layer lattices ranging between 3 and 35 nm, S-layers are regarded unique patterning elements for basic and applied research in the field of nanobiotechnology, including both life and non-life sciences.[23,24,27]

1.2. Prokaryotic glycosylation

The increasing evidence that prokaryotes can glycosylate proteins, especially the finding that several human pathogens contain glycoproteins in their surface appendages,[28] has put an end to the doctrine that restricted glycoproteins to eukaryotes. About a decade ago, Sandercock and colleagues[29] tried to sort the accumulated literature on prokaryotic glycoproteins and they discriminated between non-S-layer glycoproteins and S-layer glycoproteins. Among non-S-layer glycoproteins are glycosylated enzymes, such as cellulases and xylanases, membrane-associated glycoproteins, surface-associated glycoproteins, and glycosylated antigens that are shed into the surrounding environment by the organisms.[18] The best investigated glycoproteins from this group are the flagella and pili of archaea and bacteria,[30] and the N-glycoprotein species of *Campylobacter jejuni*.[31,32] A common feature of this group of

glycoproteins is that the glycans are usually relatively short (about 1–20 glycoses) in comparison to the long-chain glycans from bacterial S-layer glycoproteins (with up to approximately 150 glycoses).[33] In contrast, glycans from archaeal S-layer glycoproteins are usually also relatively short.[34]

With S-layer glycoproteins constituting a major class of prokaryotic glycoproteins, much of the current knowledge about prokaryotic glycosylation has been derived from S-layer glycoprotein research. Initial analyses originate from the S-layer glycoprotein of the halophilic archaeon *H. salinarum*.[16] This work was expanded in great detail by Sumper, Wieland, and co-workers in the 1980s. The authors reported not only on the structures but also on the biosynthesis of halobacterial S-layer glycoproteins; furthermore, they provided the first sequence data of the gene encoding a glycosylated archaeal S-layer protein.[12,35] Other archaea that were investigated in greater detail were the haloarchaeon *H. volcanii*,[15,34,36] and the methanogens *Methanothermus fervidus*[13] and *Methanococcus voltae*.[37] Recently, *M. voltae* flagellins have been shown to contain a novel N-linked trisaccharide. Analysis of trypsin-generated peptides derived from the *M. voltae* S-layer glycoprotein revealed a modification by the same trisaccharide, suggesting a common glycosylation process for the two proteins.[38]

As mentioned before, at about the same time when glycosylation on haloarchaea was reported for the first time, Sleytr and Thorne[17] have discovered glycosylation of S-layer proteins from the bacteria *Thermoanaerobacter thermohydrosulfuricus* and *Thermoanaerobacterium thermosaccharolyticum*. Since then, S-layer glycoproteins from several other bacteria have been extensively studied,[39] leading to the awareness of the wide distribution of S-layer glycoproteins among bacteria. Based on a considerable body of S-layer glycan structures from organisms of the *Bacillaceae* family, investigation of the S-layer protein glycosylation process at the molecular level has been initiated about 10 years ago. This endeavor has been lagging behind the structural work due to the lack of suitable molecular tools. When eventually in 2002 the first S-layer gene sequence of the S-layer glycoprotein carrying bacterium *G. stearothermophilus* NRS 2004/3a became available,[40] and only two years later it became evident that S-layer protein glycosylation of the investigated bacterium is encoded by an S-layer glycosylation (*slg*) gene cluster,[41] a novel research direction has emerged, for which we have coined the neologism S-layer 'nanoglycobiology'.[39]

1.3. S-layer nanoglycobiology

S-layer nanoglycobiology accounts for the nanometer-scale cell surface display feature of bacterial S-layer glycosylation and includes structural and biochemical aspects of S-layer glycans as well as molecular data on the machinery underlying the glycosylation event. A key aspect for the full potency of S-layer nanoglycobiology is the unique self-assembly feature of the S-layer protein matrix.[39] Being aware that in many cases the glycan structures associated with a protein are the key to protein function,[42,43] S-layer protein glycosylation will add a new and very valuable component to an 'S-layer based molecular construction kit'.[23] Methods for organizing functional materials at the nanometer level are essential for the development of novel fabrication techniques.[23,24] In particular, molecular self-assembly systems that exploit the molecular scale manufactory precision of biological systems are the prime candidates in nanobiotechnology.

Considering that, at the current state of knowledge, S-layer glycoprotein glycans represent 'nonsense structures' (no discrete function could be attributed to any bacterial S-layer glycan so far), the recent demonstration of the functional transfer of protein glycosylation pathways into the experimental model organism *Escherichia coli*[44,45] opens new avenues for functionalization of S-layer proteins by glycoengineering. It is conceivable that this technology, as was the case for DNA and protein engineering, will become an important tool both in basic and applied S-layer glycoprotein research to analyze the S-layer protein glycosylation process and to artificially equip S-layer proteins with 'functional' glycosylation motifs for (nano)biotechnological or biomedical purposes.

2. Perspectives: S-layer *neo*glycoprotein production

Engineering of tailor-made self-assembly S-layer glycoproteins, the so-called S-layer *neo*glycoproteins, will decisively change our capabilities in influencing and controlling complex biological systems and to conceptuate novel self-assembly nanomaterials.

In our long-term research strategy, the detailed knowledge of the S-layer polypeptide and of the glycosylation feature shall converge (Fig. 1). In principle, alteration of the native S-layer glycan or assembly of completely new glycans on permissive sites of the S-layer protein portion can be envisaged. In any case, the detailed and molecular understanding of the native S-layer protein glycosylation process is a prerequisite. Due to the complexity of S-layer glycan biosynthesis, involving a large number of enzymes for nucleotide sugar biosynthesis, glycosyl transfer reactions, polymerization, membrane transfer of the oligosaccharide chain, and its ligation to distinct sites on the target protein, S-layer glycoproteins have escaped (nano)biotechnological applications so far.

To utilize an S-layer protein as a target for engineered glycosylation, based on the knowledge of the amino acid

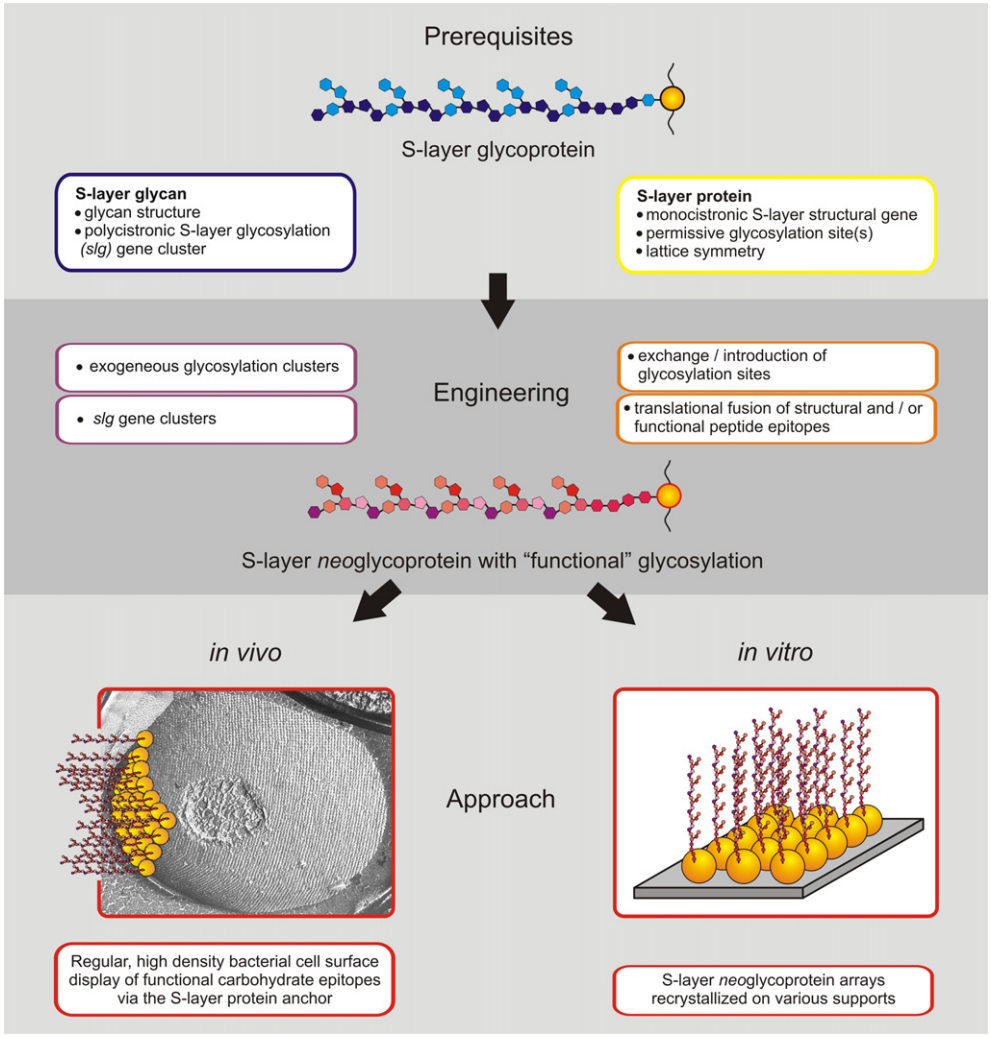

Figure 1. Illustration of prerequisites and strategies for the production of S-layer *neo*glycoproteins.

sequence, the native glycosylation sites, and potentially additional permissive sites for glycosylation have to be determined that allow introduction of exogenous glycosylation sequences into the S-layer protein. In addition, incorporation of structural or functional domains into the S-layer protein by protein engineering techniques will allow tuning S-layer *neo*glycoprotein properties for specific purposes.

Aiming at the nanobiotechnological utilization of the S-layer protein self-assembly feature, in combination with 'functional' glycosylation, it will be important to investigate for any given glycan structure, whether a change in the self-assembly behavior of the S-layer protein takes place upon recombinant glycosylation. It was important to learn from comparative self-assembly studies of native S-layer glycoproteins and the corresponding recombinant, and, consequently, non-glycosylated proteins, that glycosylation does not affect S-layer lattice formation (A. Scheberl, C. Schäffer, P. Messner, unpublished observation).

Our strategy follows two principal lines of development. The first one is the in vivo display of 'functional' glycoproteins on the surface of bacteria enabled by means of recombinant DNA technology. This has become an increasingly used strategy in various applications in microbiology, nanobiotechnology, and vaccinology.[46] Besides outer membrane proteins, lipoproteins, autotransporters, or subunits of surface appendages, which are being evaluated for that kind of applications, the use of the S-layer (glyco)protein cell surface anchor is a very attractive and promising alternative. An impressing example related to this line of development was stated by Paton and co-workers,[47] who demonstrated that a recombinant *E. coli* that displayed a Shiga toxin receptor mimic on its cell surface was capable of adsorbing and neutralizing Shiga toxins with very high efficiency. The in vitro line of development utilizes the recrystallization capability of the S-layer portion on a broad spectrum of supports. In either line, the S-layer 'anchor' offers the unique

advantage of providing a crystalline, regular 'immobilization matrix' that should eventually allow the controlled and periodic surface display of 'functional' glycosylation motifs (Fig. 1).

The conceptuation of S-layer *neo*glycoproteins clearly benefits from the data accumulated in the course of the transfer of the *C. jejuni* N-glycosylation machinery (*pgl* system) together with the *C. jejuni* target protein AcrA into *E. coli*, enabling heterologous expression of a recombinant N-glycoprotein.[44,45] These studies furthermore revealed differences between the prokaryotic and eukaryotic N-glycosylation systems. Bacteria obviously require a negatively charged amino acid at the −2 position to the glycosylated Asn in the sequon for N-glycosylation to occur, resulting in the stringent acceptor sequence D/E-YN-X-S/T (Y, XaP).[48] Concerning the PglB oligosaccharyl::protein transferase of *C. jejuni*, which is responsible for the transfer of the glycan to the AcrA protein as the final step of the glycosylation event, relaxed oligosaccharide substrate specificity has been reported, allowing the transfer of different glycans from the lipid carrier undecaprenyl pyrophosphate to an acceptor protein.[45,49] The corresponding enzyme from the S-layer protein glycosylation pathway has been identified as one of the key modules of S-layer *neo*glycoprotein production. Elucidation of the mechanism of action and the substrate specificity of this enzyme (named WsaB) from *G. stearothermophilus* NRS 2004/3a is currently being investigated in our laboratory. The obvious differences in the protein glycosylation process between bacteria and eukaryotes will clearly have to be accounted for, when aiming at the 'humanization' of S-layer proteins through 'functional' glycosylation.

As S-layer *neo*glycoprotein production represents a fresh area of research, the benefits of S-layer *neo*glycoproteins for potential nanobiotechnology applications may currently be only deduced form the successful cell surface display of foreign peptide epitopes. Among many examples in the literature for peptide epitope display via the S-layer protein anchor (for summary, see Ref. 24), connecting to the field of nanoglycobiology, we have recently constructed a chimeric S-layer protein displaying the fully active RmlA enzyme that is involved in the biosynthesis of nucleotide-activated L-rhamnose.[50] Recent advances in the field of S-layer nanoglycobiology that will be discussed in this article have made our overall strategy a tangible aim of the near future.

3. Diversity of bacterial S-layer protein glycosylation

3.1. Organisms possessing a glycosylated S-layer

Since the early days of S-layer glycoprotein research, it was evident that these cell surface components occur on archaea as well as on bacteria. S-layer glycoproteins have been known for their occurrence among the major lineages of archaea,[15,51] with most of the data concerning *H. salinarum*,[12] *M. fervidus*,[52] and *H. volcanii*.[36] Among bacteria, for a long time only Gram-positive members of the *Bacillaceae* family have been known to possess S-layer glycoproteins, including the species *G. stearothermophilus*, *G. tepidamans*, *Paenibacillus alvei*, *A. thermoaerophilus*, *Thermoanaerobacterium thermosaccharolyticum, Thermoanaerobacter thermohydrosulfuricus*, *T. kivui*, and *Desulfotomaculum nigrificans*.[18,39] Only very recently there were the first reports on the occurrence of glycosylated S-layer proteins in the Gram-negative species *Tannerella forsythia*[53] and *Bacteroides distasonis*.[54] Evidence was obtained from biochemical analyses and so far nothing is known about either glycan structure or linkage of the glycans to the S-layer protein portion. However, in contrast to the known S-layer glycoproteins from *Bacillaceae* investigated by our group in the past,[18,39] these glycosylated S-layer proteins originate from potential pathogens and, therefore, might be of medical relevance.[55] The relation of S-layer protein glycosylation to pathogenicity might implicate a first, discrete function of S-layer glycans, and, therefore, this research direction is currently being established in our laboratory. Concerning an overall function of S-layer glycoproteins in non-pathogenic bacteria, it is conceivable that, by representing the outermost cell surface structure of a bacterium, they participate in diverse cell surface phenomena and, simultaneously, contribute to a high diversification potential of the bacterial cell surface, which may be advantageous in the competitive natural habitat.

3.2. Biochemical and structural insights into S-layer protein glycosylation

The glycosylation degree of S-layer proteins generally varies between 1% and 10% (w/w), but is also subject to change, depending on the laboratory cultivation conditions of the bacteria. Presently, about 15 different S-layer glycoprotein glycan structures have been fully or at least partially elucidated and there are currently more than 25 further indications for glycan modifications of S-layer proteins according to biochemical evidence.[39,56] Figure 2 gives a comprehensive overview on S-layer glycoprotein glycan structures from bacteria. A common feature of almost all bacterial S-layer glycoproteins is the presence of long glycans made of repeats. The observed structures and glycosidic linkage types exceed by far the display found in eukaryotic glycoproteins. Commonly, bacterial S-layer glycan chains are long linear or branched homo- or heterosaccharides with 50–150 glycoses that constitute about 15–50 repeating units.[18,39] The monosaccharide constituents of bacterial S-layer glycan chains include a wide range

of neutral hexoses, 6-deoxyhexoses, and amino sugars. In addition, this spectrum is further extended by rare sugars, such as Qui*p*3NAc, Fuc*p*3NAc, D-Rha*p*, D-Fuc*p* or D-*glycero*-D-*manno*-heptose, which are otherwise

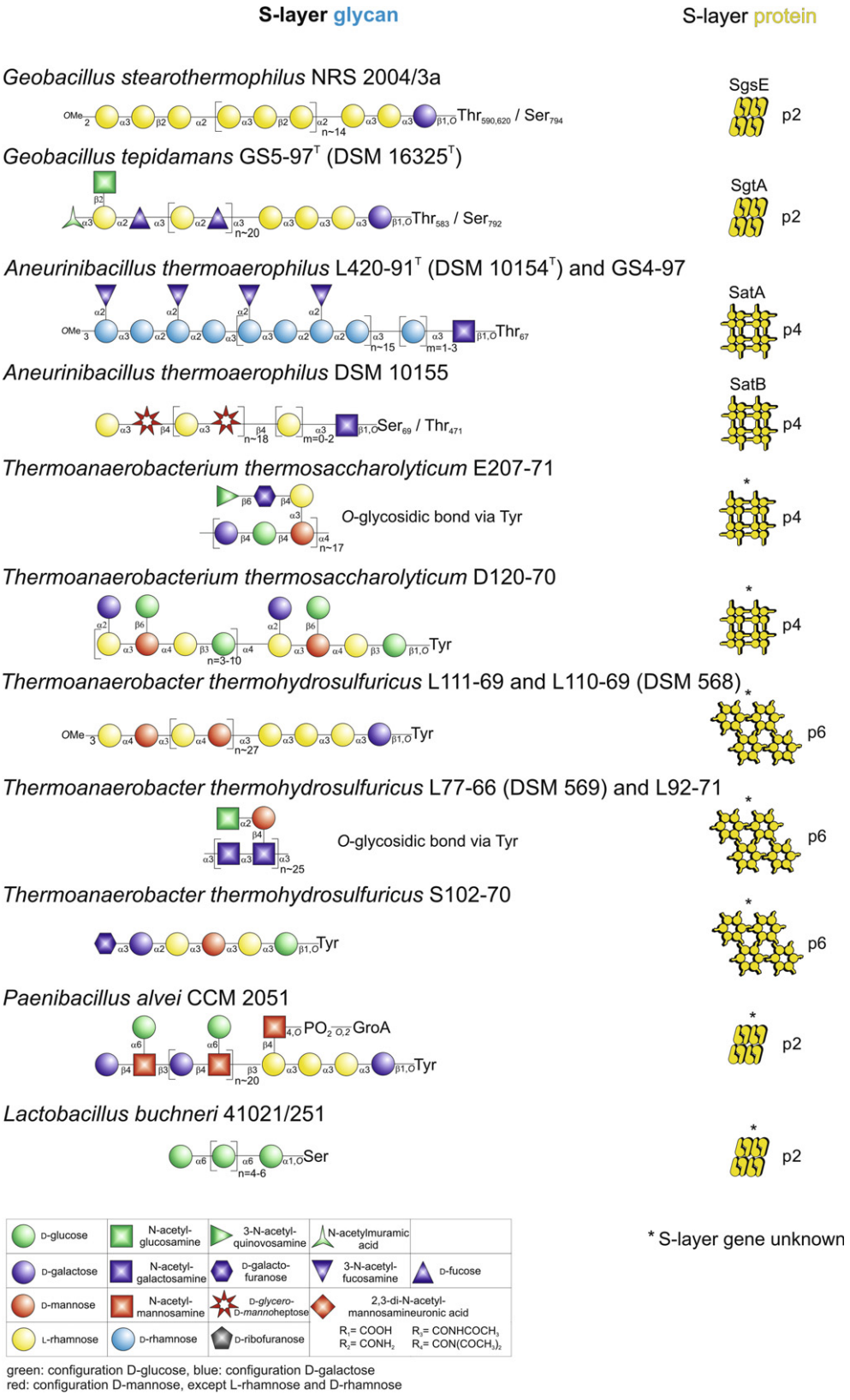

Figure 2. Bacterial S-layer glycan structures and attachment sites of S-layer proteins.

typical constituents of lipopolysaccharide (LPS) O-antigens of Gram-negative bacteria.[57] The typical linkages of these S-layer glycans to the protein portion are O-glycosidic linkages to serine, threonine, and tyrosine; so far, *N*-glycans have been found only in archaea.[39]

Bacterial S-layer glycoproteins possess a tripartite structure, which compares to lipopolysaccharide (LPS) O-antigens. This scheme comprises a glycan chain built of a variable number of repeating units that are linked via a variable core oligosaccharide to the S-layer protein backbone.[39,58,59] In the case where the oligosaccharide core is missing, the carbohydrate chain is directly bound to the glycosylated amino acid via the first repeating unit; the anomeric configuration of the linkage sugar may remain unchanged (no core) or be inverted (pseudocore). In some S-layer glycoproteins, capping of the terminal sugar residue at the non-reducing end of the glycan chain with non-carbohydrate constituents such as *O*-methyl-groups (2-*O*-Me, 3-*O*-Me) is present.[40,60] In this regard, a very interesting modification has been observed just recently. The S-layer glycoprotein glycan of *G. tepidamans* GS5-97^T is capped by α-(*R*)-*N*-acetylmuramic acid, the key component of bacterial peptidoglycan, at carbon 3, and by β-*N*-acetylglucosamine at carbon 2 of the terminal rhamnose residue.[61] A possible explanation for the presence or absence of a capping motif at the non-reducing end of the glycan chains might be derived, when relating S-layer glycoprotein glycan biosynthesis to LPS O-antigen biosynthesis.[57] There, depending on the route of biosynthesis, methylation was reported to function as termination signal for chain elongation.[57,62]

Detailed structural and genetic analyses have shown that bacterial S-layer glycoproteins, for example, from *G. stearothermophilus* NRS 2004/3a, *G. tepidamans* GS5-97^T, and *A. thermoaerophilus* strains, possess only a small number of potential glycosylation sites.[40,59] For instance, on SgsE, the S-layer protein of *G. stearothermophilus* NRS 2004/3a, the amino acids Thr-590, Thr-620, and Ser-794 have been identified as potential glycosylation sites. For the S-layer protein SgtA of *G. tepidamans*, also three distinct glycosylation sites are obvious, in *A. thermoaerophilus* DSM 10155/G$^+$ the amino acids Ser-69 and Thr-471 of SatB represent putative glycosylation sites, and for *A. thermoaerophilus* L420-91^T position Thr-67 of SatA has so far been identified as a glycosylation site. However, it is not known to which extent each of these sites is occupied in the mature glycoprotein. Consequently, an S-layer glycoprotein array is a mixture of variably glycosylated S-layer protein species. In summary, the glycosylation event adds an enormous variation potential to the respective S-layer protein. All features of bacterial S-layer glycoproteins known to date, including potential modifications and glycosylation sites, are summarized in Figure 2.

3.3. The S-layer protein portion

One of the prerequisites for the production of self-assembly S-layer *neo*glycoproteins is the detailed knowledge of the S-layer protein matrix.

Independent of the presence of glycosylation, S-layer proteins *per se* are water-insoluble proteins. Results of amino acid analysis indicated large amounts of glutamic acid and aspartic acid (~15 mol %), a high lysine content (~10 mol %), and large amounts of hydrophobic amino acids (~40–60 mol %). Hydrophilic and hydrophobic amino acids do not form extended clusters. Most S-layer subunits are weakly acidic proteins with isoelectric points in the range 4–6, with the exception of the S-layer proteins of lactobacilli.[24] Secondary structure predictions derived from comparison of protein sequence data and circular dichroism measurements revealed that S-layer proteins have an average α-helix content of ~20% and a β-sheet content of ~40%; aperiodic folding and the β-turn content may vary between 5% and 45%.[23]

As no tertiary structure of a full-size S-layer protein is known so far, our approach for 'functional' S-layer protein glycosylation is the utilization of the native glycosylation sites on the protein, because these sites can be expected to be located in surface-exposed loops within the bulk of the S-layer protein, allowing the display of attached glycans. The identification of the native S-layer glycosylation sites is based on the availability of the primary sequence of the protein, with which the sequences of proteolytically derived S-layer glycopeptides are aligned. The amino acid sequence of an S-layer protein can in principle be obtained either by a genomic approach or a proteomic approach, with the former being more frequently applied. From a recent combined approach, the primary sequence of the S-layer protein SgtA of *G. tepidamans* GS5-97^T was obtained. The purified S-layer protein was applied to proteolytic digests and the obtained peptides were analyzed by nano-ESI-QTOF tandem mass spectrometry.[63] Alignment of the sequenced peptides of SgtA showed high similarity to the N-terminus of the S-layer protein SgsE of *G. stearothermophilus* NRS 2004/3a. The identified conserved amino acid sequences were used for the design of degenerate primers for PCR amplification reactions to enable further sequencing, which finally resulted in the complete sequence of the S-layer structural gene *sgtA*.[64]

Studies on S-layer proteins from different bacteria revealed that S-layer proteins are multidomain proteins, comprising a cell-wall targeting region, which is involved in anchoring of the protein to the peptidoglycan of the bacterial cell wall by interaction with a species specific secondary cell wall polymer,[65,66] and a self-assembly domain, which makes up the larger part of the protein. Based on amino acid sequence alignment and on experimental data, for S-layer glycoproteins

from the *Bacillaceae*, which are serving as a base for S-layer *neo*glycoprotein production, two structural organization principles exist.[23,24] For the species *G. stearothermophilus*, the self-assembly domain is located at the C-terminal region of the S-layer protein,[67] while S-layer proteins of the species *A. thermoaerophilus* possess an N-terminal self-assembly domain. The finding that all glycosylation sites that have been identified on S-layer proteins so far are located within the proposed self-assembly region of the S-layer protein is of high relevance for the conceptuation of S-layer *neo*glycoproteins. Aiming at a highly efficient S-layer *neo*glycoprotein production, deletion of the cell wall binding domain of the S-layer protein without affecting the protein's self-assembly capability or its glycosylation potential should be possible in principle. Indeed, it was shown that deleting 130 or 330 amino acids from the N-terminus of SgsE of *G. stearothermophilus* NRS 2004/3a apparently does not affect S-layer self-assembly.[50]

3.4. Understanding S-layer protein glycosylation patterns

S-layer glycoprotein arrays are usually composed of individual, high-molecular mass glycoprotein subunits.[18] A first indication for the presence of a glycosylated S-layer protein in a given organism can be easily inferred from a positive carbohydrate staining reaction of a highly abundant protein band on an SDS-PA gel of intact bacterial cells, appearing in an estimated molecular mass range between 45 and 200 kDa.[24] Figure 3 gives an overview of the diversity of S-layer protein glycosylation patterns according to the migration behavior of glycoproteins on SDS-PA gels, visualized by Coomassie Blue and periodic acid-Schiff staining (PAS) reaction. It is evident from SDS-PAGE analyses that S-layer glycoproteins migrate as single or multiple bands, or even have a ladder-like appearance, an effect known from LPS O-antigens and referred to as nanoheterogeneity in the context of S-layer nanoglycobiology. In any case, also the corresponding, non-glycosylated S-layer protein is present, which is in accordance with the finding that glycosylation is lagging behind S-layer protein synthesis (R. Novotny, A. Scheberl, and C. Schäffer, unpublished data).

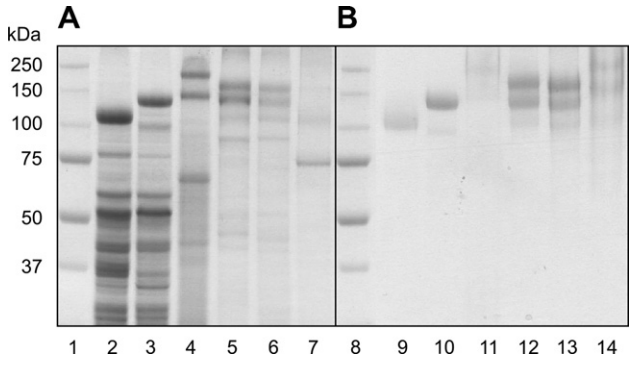

Figure 3. SDS-PAGE analysis of S-layer protein glycosylation patterns on intact bacterial cells by (A) Coomassie Blue staining and by (B) periodic acid-Schiff (PAS) staining reaction. Lanes 1, 8, molecular mass standard; lanes 2, 9, *A. thermoaerophilus* L420-91^T; lanes 3, 10, *A. thermoaerophilus* DSM 10155/G$^+$; lanes 4, 11, *P. alvei* CCM 2051^T; lanes 5, 12, *G. stearothermophilus* NRS 2004/3a; lanes 6, 13, *G. tepidamans* GS5-97^T, lanes 7, 14, *T. thermosaccharolyticum* E207-71. Amounts of 5 μg (A) and 10 μg (B) were loaded onto the gel.

SDS-PAGE analysis of the S-layer glycoprotein of *T. thermosaccharolyticum* E207-71 reveals an apparent molecular mass of 75 kDa for the non-glycosylated protein and nanoheterogeneity of the glycoprotein migrating as 14 different bands in the mass range of 103–213 kDa (Fig. 3, lanes 7, 14).[68] The glycosylated subunits of *A. thermoaerophilus* DSM 10155/G$^+$ possess an apparent molecular mass of 143 kDa, whereas the non-glycosylated band is downshifted to 76 kDa (Fig. 3, lanes 3, 10), which is in accordance with the calculated molecular mass derived from the amino acid sequence of the mature structural protein SatB (78.3 kDa). *A. thermoaerophilus* L420-91^T shows only one glycosylated band at an apparent molecular mass of 109 kDa on SDS-PAGE and a non-glycosylated band at 76 kDa (Fig. 3, lanes 2, 9), which corresponds to the calculated molecular mass of SatA (81.4 kDa). *P. alvei* CCM 2051^T displays three bands on SDS-PAGE with apparent molecular masses of 105, 155, and 240 kDa, of which the two high-mass bands are glycosylated (Fig. 3, lanes 4, 11). On an SDS-PA gel, the mature S-layer glycoproteins of *G. stearothermophilus* NRS 2004/3a (Fig. 3, lanes 5, 12) and of *G. tepidamans* GS5-97^T (Fig. 3, lanes 6, 13) are separated into four bands with apparent molecular masses of 93, 119, 147, and 170 kDa, and 93, 119, 140, and 166 kDa, respectively.[59] In either case, the three high-molecular-mass bands give a positive PAS staining reaction for carbohydrates. The 93-kDa bands are non-glycosylated, and the estimated molecular mass concurs with the calculated masses for the respective S-layer protein after cleavage of the 30-amino acid signal peptide from the precursor protein, that is, 93.7 kDa for SgsE and 92.3 kDa for SgtA.

As the detailed knowledge of the glycosylation pattern is another prerequisite for S-layer *neo*glycoprotein production, the interpretation of SDS-PAGE profiles of S-layer glycosylation was recently refined by mass spectrometry approaches. This was necessary, because it is known that glycoproteins have a retarded electrophoretic mobility due to the attached glycan portion, resulting in too high molecular mass estimates, and that the resolution of individual glycoprotein bands on a PA gel is not good enough to display glycan chain length distribution in detail. To address these questions, mass spectrometry (MS) is a more reliable alternative.[69–71]

Already in 1995, it was shown for the S-layer glycoprotein of *T. thermosaccharolyticum* E207-71 that the

mass of 75.62 kDa obtained by matrix-assisted laser desorption/ionization (MALDI)-MS (mass accuracy, 0.2%) was in accordance with the value of ~75 kDa derived from SDS-PAGE analysis.[68,72] A second broad, only partially resolved signal centered around 87.62 kDa corresponded presumably to the glycoprotein with carbohydrate-related heterogeneity. With the S-layer glycoprotein SgsE of *G. stearothermophilus* NRS 2004/3a, it has been demonstrated, how MS approaches can elucidate S-layer glycosylation patterns based on the known S-layer glycan structure.[59] As described above, on SDS-PA gels, the S-layer glycoprotein preparation of this bacterium is separated into three glycoprotein species. Previous NMR analyses have revealed that the S-layer glycans consist of trisaccharide repeats with the structure →[2)-α-L-Rha*p*-(1→3)-β-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→]$_n$, with the terminal rhamnose residue at the non-reducing end modified on carbon-2 by O-methylation, and of a short core saccharide consisting of α-(1→3)-linked L-Rha*p* residues, attached to carbon 3 of a β-D-galactose residue that serves as the linkage sugar to distinct sites on the S-layer polypeptide backbone.[40] As these data could not fully explain the SDS-PAGE banding pattern, infrared MALDI orthogonal time-of-flight MS (IR-MALDI-oTOFMS), nanoelectrospray ionization quadrupol time-of-flight mass spectrometry (nano-ESI-QTOFMS), and Fourier transform ion cyclotron resonance infrared multiphoton dissociation mass spectrometry (FTICR-IRMPDMS) were adapted for analysis of this high-molecular-mass S-layer glycoprotein and glycopeptides thereof.[59,73] Optimization and adaptation of IR-MALDI-oTOF MS to the water-insoluble S-layer glycoprotein allowed the determination of the average masses of the three inherently heterogenic glycoprotein species of SgsE to be 101.66 kDa, 108.68 kDa, and 115.73 kDa, corresponding to SgsE with up to three attached glycan chains, because the average mass differences between two neighboring peaks of the singly charged ions were 7.09 kDa, which is the average mass of a glycan chain with 15 repeating units. However, nanoheterogeneity of the glycan chain cannot be detected by this method due to peak broadening resulting from reduced resolution in the high-mass range and adduct formation of the matrix. Analysis of isolated glycopeptides by nanoESI-QTOFMS and FTICR-IRMPDMS enabled clear assignment of nanoheterogeneity to each glycan chain, with each of them revealing the most prevalent variation between 12 and 18 trisaccharide repeating units and the possibility of extension of the already known di-L-rhamnose core region by one additional L-Rha residue, and unambiguous identification of a third glycosylation site on the SgsE S-layer protein, namely, at position Thr-590, in addition to the known sites of Thr-620 and Ser-794. These data have led to the current interpretation that in the 101.66-kDa glycoprotein species only one glycosylation site is occupied, in the 108.68-kDa glycoprotein species two glycosylation sites are occupied, and in the 115.73-kDa glycoprotein species three glycosylation sites are occupied, whereas the 94.46-kDa band represents non-glycosylated S-layer protein. A similar situation was observed for the S-layer glycoprotein SgtA of *G. tepidamans* GS5-97^T, where four peaks corresponding to the four bands in SDS-PAGE could be determined to be 94.1, 101.3, 108.5, and 115.7 kDa.[64]

4. The molecular machinery behind S-layer protein glycosylation

4.1. Model organisms

Currently, *G. stearothermophilus* NRS 2004/3a is the best investigated model organism for addressing questions relevant for S-layer nanoglycobiology. However, to obtain a more general insight into the S-layer glycosylation process there are five more organisms under investigation in our laboratory; these are *G. tepidamans* GS5-97^T, *A. thermoaerophilus* DSM 10155/G$^+$ and L420-91^T, *T. thermosaccharolyticum* E207-71, and *P. alvei* CCM 2051^T. The selection of these organisms is based on the availability of the S-layer glycan structure; the current status of research on these organisms is quite different. It is important to note that *P. alvei* CCM 2051^T is so far the only one of the selected organisms that can uptake foreign DNA, making it a prime candidate for in vivo display of S-layer *neo*glycoproteins (K. Zarschler, C. Schäffer, P. Messner, unpublished data).

All model organisms are Gram-positive, moderately thermophilic or mesophilic, spore-forming bacteria that originate either from soil samples (*G. stearothermophilus* NRS 2004/3a), from extraction plants of Austrian beet sugar factories (*T. thermosaccharolyticum* E207-71, *G. tepidamans* GS5-97^T, *A. thermoaerophilus* DSM 10155/G$^+$ and L420-91^T), or from foul brood of bees (*P. alvei* CCM 2051^T).

Both *Geobacillus* strains, the well characterized *G. stearothermophilus* NRS 2004/3a and the just recently classified *G. tepidamans* GS5-97^T, possess S-layer glycoproteins assembling into arrays with oblique symmetry.[74,75] While the S-layer glycan of *G. stearothermophilus* NRS 2004/3a is a poly-L-rhamnan chain consisting of trisaccharide repeats with the structure →[2)-α-L-Rha*p*-(1→3)-β-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→]$_n$, the more complex glycan of *G. tepidamans* GS5-97^T is composed of disaccharide repeats with the structure →[3)-α-L-Rha*p*-(1→2)-α-D-Fuc(1→]$_n$.[40,61] The S-layer glycoprotein SatA of *A. thermoaerophilus* L420-91^T constitutes a square S-layer lattice and its glycan chain is composed of identical hexasaccharide repeats containing D-rhamnose and 3-acetamido-3,6-dideoxy-D-galact-

ose (D-Fuc*p*3NAc) in the molar ratio of 2:1.[60,76] Another member of the species *Aneurinibacillus*, namely *A. thermoaerophilus* DSM 10155/G$^+$, possesses the only heptose-containing S-layer glycan known so far. It is composed of disaccharide repeats of D-rhamnose and D-*glycero*-D-*manno*-heptose residues.[58,77] *T. thermosaccharolyticum* E207-71 possesses the most complex repeating unit structure of our model organisms with a branched S-layer hexasaccharide repeat composed of D-Gal*f*, L-Rha, D-Gal, D-Glc, D-Man, and the rare glycose 3-*N*-acetylquinovosamine (D-Qui*p*3NAc).[68]

Due to the ease of its transformability, we are recently focusing on the mesophilic bacterium *P. alvei* CCM 2051^T. More than 40 years ago, a granular, macromolecular pattern in the cell wall of this bacterium was observed and several years later, the structure of the S-layer and its glycan chain have been investigated.[78–81] Cells of *P. alvei* CCM 2051^T are covered with an oblique S-layer glycoprotein lattice. The S-layer glycan is composed of branched trisaccharide repeats with the constituents D-Gal, D-Glc, and D-ManNAc; in addition, a 2-phosphoglyceric acid residue is present in the core of this S-layer glycan.[81]

4.2. S-layer glycosylation gene clusters

An important milestone toward understanding S-layer protein glycosylation was the identification and sequencing of several S-layer glycosylation (*slg*) gene clusters. Analysis of S-layer glycosylation on the molecular level clearly benefited from the detailed molecular knowledge of the biosynthesis routes of LPS O-antigens and capsular polysaccharides.[82] Currently, most data are available from the organisms *G. stearothermophilus* NRS 2004/3a (GenBank AF328862),[83] *G. tepidamans* GS5-97^T (GenBank AY883421)[64] and *A. thermoaerophilus* strains L420-91^T (GenBank AY442352),[41] and DSM 10155/G$^+$ (GenBank AF324836).[41] In addition, a partial *slg* gene cluster sequence is available from *T. thermosaccharolyticum* E207-71 (GenBank AY422724)[41] (Table 1, Fig. 4).

Based on the common principle that sugars are incorporated into growing glycan chains from the respective nucleotide-activated precursor, we surveyed the literature for the genes involved in the biosynthesis of nucleotide-activated L-rhamnose (dTDP-β-L-rhamnose),[84] D-fucose (dTDP-α-D-fucose),[64] D-rhamnose (GDP-α-D-rhamnose),[85] 3-*N*-acetylfucosamine (dTDP-α-D-Fuc*p*3NAc),[86] D-*glycero*-α-D-*manno*-heptose (GDP-D-*glycero*-α-D-*manno*-heptose),[87,88] and 3-*N*-acetylquinovosamine (dTDP-α-D-Qui*p*3NAc),[89] all of which are constituents of the S-layer glycans of our model organisms. For entry into the *slg* gene clusters, degenerate primers were designed based on these gene sequences. Further up- and downstream sequencing eventually revealed the presence of whole gene clusters. The current picture of the *slg* gene clusters is the following. Depending on the complexity of the encoded S-layer glycan, the clusters are ~16 to ~25 kb in size and transcribed as polycistronic units.[41] They include nucleotide sugar pathway genes that are arranged consecutively, glycosyl transferase genes, glycan processing genes, and transporter genes. The presence of insertion sequences and the decrease of the G + C content at the *slg* locus in comparison to the respective bacterial genome suggest that the investigated organisms have acquired their specific S-layer glycosylation potential by lateral gene transfer. From the assigned genes, it is evident that none of the *slg* gene clusters encodes the biosynthesis of the nucleotide-activated linkage sugar of the S-layer glycan (UDP-Gal and UDP-GalNAc, respectively). Thus, S-layer protein glycosylation additionally requires the participation of housekeeping genes that map outside the cluster. The gene encoding the respective S-layer target protein is transcribed monocistronically and independently of the *slg* cluster genes (compare with Fig. 4). Its chromosomal location is not necessarily in close vicinity to the *slg* gene cluster.

For deducing a common organization principle of *slg* gene clusters, as known from the clusters encoding the biosynthesis of other bacterial polysaccharides, such as the LPS O-antigens of Gram-negative bacteria,[57,90] or the exopolysaccharides of lactic acid bacteria,[91] the number of *slg* clusters sequenced so far is too low. Furthermore, the current sequence information does not allow the identification of specific genes on the chromosome of the *Bacillaceae*, such as the *galF* and *gnd* genes in *E. coli* and *Salmonella enterica*,[90] or the *hemH* and *gsk* genes in *Yersinia enterocolitica*,[92] between which the *slg* locus is preferentially located. Nevertheless, recently, the comparison of the *slg* gene clusters of *G. stearothermophilus* NRS 2004/3a and *G. tepidamans* GS5-97^T revealed that the clusters are organized in a similar way.[64] The first three ORFs downstream of the S-layer gene include highly homologous genes coding for a putative ligase and a putative rhamnosyltransferase and are followed by the ABC-2 transporter encoding genes *wzm* and *wzt*. The subsequent sequence segments are different. In *G. tepidamans* GS5-97^T, there are five ORFs including four putative glycosyltransferase genes and the dTDP-4-dehydro-6-deoxyglucose reductase gene *fcd*. This segment is terminated by three transposases and one small ORF with unknown function. In *G. stearothermophilus* NRS 2004/3a, only two ORFs in that region of the gene cluster are found, a putative methyltransferase gene and a putative glycosyltransferase gene. Downstream of these variable parts of the *slg* gene clusters, the *rmlACBD* genes, a putative rhamnosyltransferase gene and the putative UDP-galactose lipid carrier transferase gene, are located. These parts of both gene clusters have the highest homology (up to 96% similarity). The

Table 1. Data base homologies of the genes contained in the *slg* gene clusters and S-layer structural genes

Nomenclature[a]		aa	MW (kDa)	G + C%	Assigned protein function (database link)
New	Old				
(a) Description of the slg gene cluster of Geobacillus stearothermophilus NRS 2004/3a (GenBank accession number AF328862)					
wsaA	ORFG101	169	19.9	43	TPR-repeat containing protein (ABM68315)
wsaB	ORFG102	526	59.9	37.4	Lipid A core::surface polymer ligase (ABM68316)
wsaC	ORFG103	324	38.2	33.0	Rhamnosyltransferase (ABM68317)
wzm	*wzm*	268	30.4	33.0	ABC transporter integral membrane protein (ZP_01501790)
wzt	*wzt*	409	46.2	33.7	ABC transporter nucleotide-binding protein (NP_643907)
wsaE	ORFG106	1127	132.6	31.8	Methyltransferase (YP_984868)
wsaF	ORFG107	413	48.3	30.0	Rhamnosyltransferase (ZP_01037878)
rmlA	*rmlA*	299	33.2	39.6	Glucose-1-phosphate thymidylyltransferase (ABM68329)
rmlC	*rmlC*	183	21.3	38.4	dTDP-dehydrorhamnose 3,5-epimerase (ABM68330)
rmlB	*rmlB*	342	38.8	38.4	dTDP-D-glucose 4,6-dehydratase (ABM68331)
rmlD	*rmlD*	282	31.9	38.5	dTDP-dehydrorhamnose reductase (ABM68332)
wsaD	ORFG112	289	33.4	37.0	Rhamnosyltransferase (ABM68333)
wsaP	ORFG113	471	54.5	37.9	UDP-galactose-lipid carrier transferase (ABM68334)
transposase	*istA*	184	21.7	42.2	Transposase (NP_634718)
(b) Description of the slg gene cluster of Geobacillus tepidamans GS5-97^T (GenBank accession number AY883421)					
wsbA		160	19.2	41.7	TPR-domain containing protein (AAR99603)
wsbB		526	59.7	37.1	Lipid A core::O-antigen ligase (AAR99604)
wsbC		325	38.5	32.1	Rhamnosyltransferase (AAR99605)
wzm		269	32.0	27.1	ABC transporter integral membrane protein (AAS49124)
wzt		395	45.5	29.1	ABC transporter nucleotide-binding protein (BAA82537)
wsbG		283	33.9	32.3	Glycosyltransferase (EDN15657)
wsbH		434	51.2	27.9	Glycosyltransferase (YP_323850)
wsbI		332	39.3	29.8	Rhamnosyltransferase (BAA19639)
wsbJ		414	48.6	31.2	Glycosyltransferase (YP_061583)
fcd		308	35.5	26.4	dTDP-4-dehydro-6-deoxy-glucose reductase (ZP_01501805)
transposase		181	21.4	41.3	Transposase (CAA79750)
transposase		122	14.5	38.8	Transposase (BAD18133)
transposase		59	7.1	46.2	Transposase (NP_243386)
wsbL		62	7.1	43.0	Unknown function
rmlA		299	33.3	41.6	Glucose-1-phosphate thymidylyltransferase (AAR99610)
rmlC		183	21.3	37.2	dTDP-4-dehydrorhamnose 3,5-epimerase (AAR99611)
rmlB		339	38.7	37.7	dTDP-D-glucose 4,6-dehydratase (AAR99612)
rmlD		282	32.0	37.8	dTDP-4-dehydrorhamnose reductase (AAR99613)
wsbD		250	33.3	37.1	Rhamnosyltransferase (AAR99614)
wsbP (incomplete ORF)		441	54.0	39.2	UDP-galactose-lipid carrier transferase (AAR99615)
(c) Description of the slg gene cluster of Aneurinibacillus thermoaerophilus L420-91^T (GenBank accession number AY442352)					
wscA	ORFA101	883	97.0	48.4	S-layer associated protein (YP_076994)
wscC	ORFA102	188	21.2	44.0	Transcriptional regulator (YP_753396)
gmd	*gmd*	341	39.0	37.0	GDP-mannose-4,6-dehydratase (YP_685009)
rmd	*rmd*	309	34.5	35.9	GDP-4-dehydro-6-deoxy-D-mannose-reductase (ZP_00768807)
wzm	*wzm*	261	30.5	33.2	ABC transporter integral membrane protein (YP_727348)
wzt	*wzt*	408	45.7	35.4	ABC transporter nucleotide-binding protein (BAA28325)
wscD	ORFA107	467	55.4	32.0	Methyltransferase (NP_616119)
wscE	ORFA108	353	41.2	31.0	Glycosyltransferase (YP_303938)
wscF	ORFA109	393	46.4	32.2	Hexosyltransferase (YP_303939)
transposase	ORFA110	103	11.94	41.1	Transposase (YP_001124563)
wscG	ORFA111	746	86.2	29.4	Integral membrane protein (AAK53474)
fdtA	*fdtA*	139	16.0	34.8	dTDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase (EDN36731)
fdtC	*fdtC*	192	21.0	35.1	dTDP-D-Fuc*p*3N acetylase (ZP_01774133)
fdtB	*fdtB*	363	41.0	35.4	dTDP-3-keto-6-deoxy-D-galactose aminase (ABF93267)
wscH	ORFA115	310	35.8	34.4	Glycosyltransferase (YP_001420868)
wscI	ORFA116	119	13.7	25.8	Integral membrane (GtrA-like) protein (YP_365446)
rmlA	*rmlA*	305	33.8	36.0	Glucose-1-phosphate thymidylyltransferase (ABM68329)
rmlB	*rmlB*	343	39.7	33.7	dTDP-D-glucose 4,6-dehydratase (ZP_01872234)
transposase	ORFA119	132	15.7	47.0	Transposase (NP_243386)
transposase	ORFA120	172	19.7	47.5	Transposase (BAD18254)
manC	*manC*	464	52.8	30.1	Mannose-1-phosphate guanylyltransferase (YP_149157)

Table 1 (*continued*)

Nomenclature[a]		aa	MW (kDa)	G + C%	Assigned protein function (database link)
New	Old				
wscJ	ORFA122	389	44.9	30.4	Mannosyltransferase (CAJ74004)
wscK	ORFA123	377	43.1	27.1	Glycosyltransferase (ZP_01425292)
wscL	ORFA124 (fragment)	111	12.6	48.3	Mannosyltransferase (NP_285363)
(d) Description of the slg gene cluster of Aneurinibacillus thermoaerophilus DSM 10155/G$^+$ (GenBank accession number AF324836)					
wsdA	ORFA201	883	97.0	48.2	S-layer associated protein (YP_076994)
wsdC	ORFA202	188	21.1	44.0	Transcription regulator (YP_753396)
wzm	*wzm*	267	31.4	32.3	ABC transporter integral membrane protein (YP_001232541)
wzt	*wzt*	435	49.4	33.6	ABC transporter nucleotide-binding protein (YP_001311813)
wsdD	ORFA205	303	34.7	33.8	Methyltransferase (YP_001311812)
wsdE	ORFA206	618	72.4	32.4	Glycosyltransferase (ZP_01624791)
wsdF	ORFA207	438	50.9	34.8	Glycosyltransferase (YP_294944)
hddA	*hddA*	341	38.1	36.5	D-*Glycero*-D-*manno*-heptose 7-phosphate kinase (ZP_00592568)
gmhA	*gmhA*	198	21.5	36.0	Sedoheptulose 7-phosphate isomerase (YP_179590)
hddC	*hddC*	230	26.6	35.1	D-*Glycero*-D-*manno*-heptose 1-phosphate guanylyltransferase (NP_809389)
gmhB	*gmhB*	179	20.8	31.1	D-*Glycero*-D-*manno*-heptose 1,7-bisphosphate phosphatase (YP_001060251)
transposase	ORFA212	114	13.7	35.7	Transposase fragment (YP_001127525)
transposase	ORFA213	78	8.9	38.9	Transposase fragment (YP_001127525)
rmlA	*rmlA*	296	32.9	43.0	Glucose-1-phosphate thymidylyltransferase (ABM68329)
rmlC	*rmlC*	182	20.7	44.0	dTDP-6-deoxy-D-glucose-3,5-epimerase (AAQ23680)
rmlB	*rmlB*	341	39.1	41.4	dTDP-D-glucose 4,6-dehydratase (ABM68331)
rmlD	*rmlD*	282	31.7	45.2	dTDP-4-dehydrorhamnose reductase (AAQ23682)
wsdG	ORFA218 (partial)	295	23.2	41.0	Sugar transferase (YP_385128)
(e) Description of the incomplete slg gene cluster of Thermoanaerobacterium thermosaccharolyticum E207-71 (*GenBank accession number* AY422724)					
glf	*glf*	372	44.6	27.1	UDP-galactopyranose mutase (YP_149169)
wse1	ORFT102	321	37.5	26.6	Rhamnosyltransferase (NP_275486)
wse2	ORFT103	278	33.6	20.7	Glycosyltransferase (ZP_01372717)
qdtC	*qdtC*	265	29.7	29.2	dTDP-D-Qui*p*3N acetylase (ZP_01065559)
qdtA	*qdtA*	136	16.0	26.5	dTDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase (AAS55720)
qdtB	*qdtB*	365	41.5	29.4	dTDP-3-keto-6-deoxy-D-glucose aminase (YP_001309709)
wzx	*wzx*	491	55.8	25.3	Polysaccharide transporter (YP_001392372)
rmlB	*rmlB*	351	40.7	32.2	dTDP-D-glucose 4,6-dehydratase (ZP_00778370)
rmlA	*rmlA*	302	33.6	35.0	Glucose-1-phosphate thymidylyltransferase (ZP_00778369)
rmlD	*rmlD*	294	33.5	32.8	dTDP-dehydrorhamnose reductase (ZP_00778368)
rmlC	*rmlC* (partial)	(83)	(9.7)	(32.0)	dTDP-4-dehydrorhamnose 3,5-epimerase (ZP_00778367)

(f) Description of S-layer structural genes encoding target proteins for glycosylation

Species	Strain	Gene	G + C%	Precursor/signal peptide (aa)	MW (kDa)	pI	Glycosylation sites on protein precursor	Protein sequence accession number
Geobacillus stearothermophilus	NRS 2004/3a	*sgsE*	47.6	903/30	93.7	6.01	Thr-590, Thr-620, Ser-794	AAL46630
Geobacillus tepidamans	GS5-97^T	*sgtA*	38.0	901/30	92.3	5.38	Ser-792, Thr-583	AAX46285
Aneurinibacillus thermoaerophilus	L420-91^T	*satA*	43.6	759/30	78.5	5.72	Thr-67	AAS44591
Aneurinibacillus thermoaerophilus	DSM 10155/G$^+$	*satB*	44.2	738/30	75.4	5.15	Ser-69, Thr-794	AAS44592

[a] The identical color code for the components involved in S-layer glycoprotein glycan biosynthesis is used in Figures 4, 6 and in Table 1: light blue, monosaccharide biosynthesis; red, glycan assembly; orange, glycan transfer; green, export; gray, transposase; black, unknown function; dark blue, elongated glycan chain; pink, linkage glycose; yellow, S-layer protein.

glycoproteins of both organisms possess an extended tripartite structure.[39] It seems that the variable part in the center of the *slg* gene clusters is responsible for the biosynthesis of the individual repeating units and terminating elements, whereas the region with higher homology codes for proteins involved in assembling the core region, transport of the glycan to the cell surface, and its ligation to the S-layer protein. This resembles the organization of

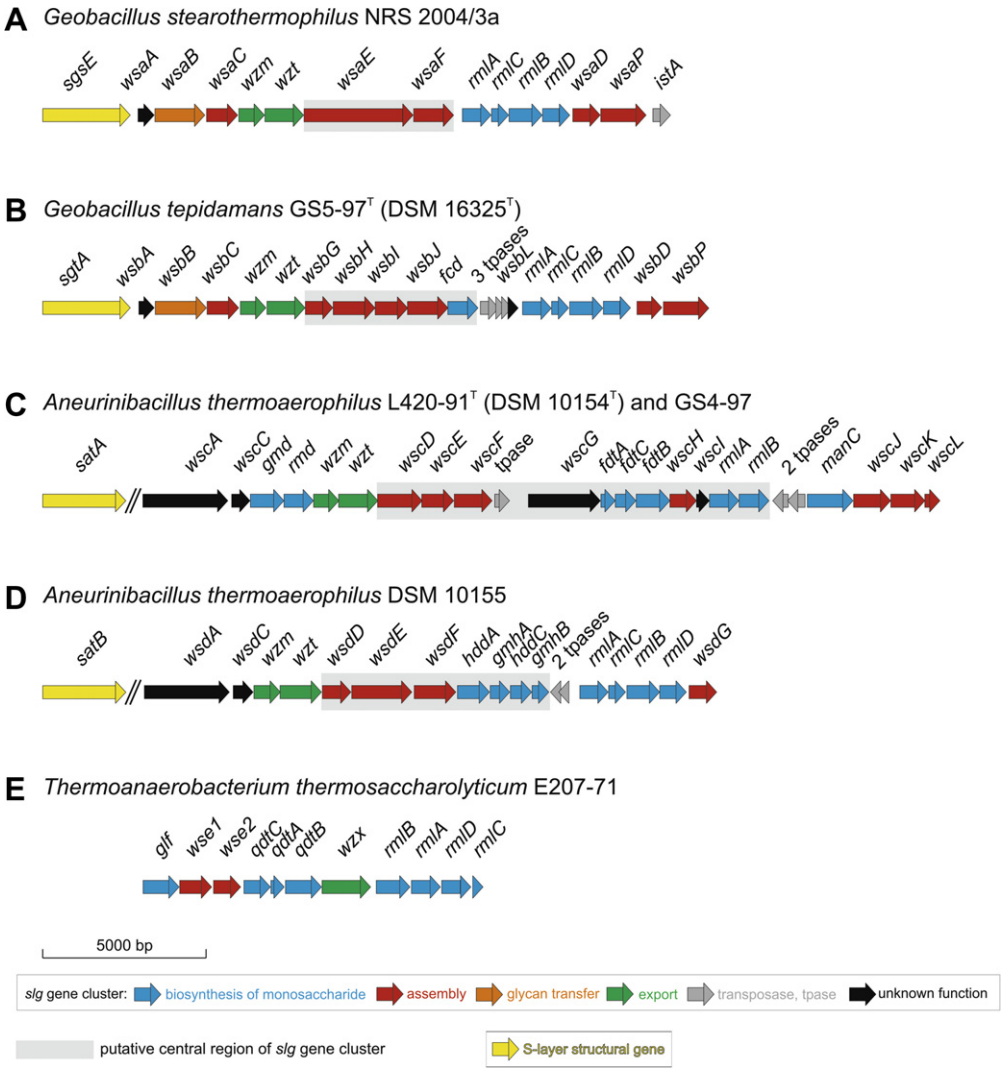

Figure 4. Genetic organization of *slg* gene clusters of (A) *G. stearothermophilus* NRS 2004/3a; (B) *G. tepidamans* GS5-97^T; (C) *A. thermoaerophilus* L420-91^T; (D) *A. thermoaerophilus* DSM 10155/G$^+$; and (E) *T. thermosaccharolyticum* E207-71. The identical color code for the components involved in S-layer glycoprotein glycan biosynthesis is used in Figures 4, 6 and in Table 1: light blue, monosaccharide biosynthesis; red, glycan assembly; orange, glycan transfer; green, export; gray, transposase; black, unknown function; dark blue, elongated glycan chain; pink, linkage glycose; yellow, S-layer protein.

O-antigen gene clusters in Gram-negative organisms, where the variability of O-antigens is considered as a result of recombination events in the central region of the O-antigen gene clusters.[93,94]

To conform the designation to the bacterial polysaccharide gene nomenclature[95] (BPGD database, http://www.microbio.usyd.edu.au/BPGD), we named the genes of the *slg* gene clusters *wsx***Y**, where *w* stands for glycan biosynthesis, *s* for S-layer-associated, *x**, starting from *a*, indicates the bacterial organism, in which a given *slg* gene cluster occurs, and *Y**, starting from *A*, reflects the consecutive order of the genes within a cluster. In case of an incomplete gene cluster, the capital letters are replaced by consecutive numbers, starting from 1. There are two exceptions, *B*, independent on the position within the cluster, is always reserved for the gene encoding the oligosaccharyl::protein transferase and *P* is reserved for the initiation enzyme.

4.3. Enzymes from S-layer protein glycosylation pathways

Most of the protein functions encoded by the different genes contained in the *slg* gene clusters have been preliminary assigned according to data base alignments. Functional characterization has so far been performed only with different proteins that are involved in the biosynthesis of nucleotide sugars and for the initiation enzyme of S-layer glycan biosynthesis WsaP.

4.3.1. Nucleotide sugar biosynthesis enzymes. Based on the identification of the S-layer glycan specific nucleotide sugar genes in the *slg* gene clusters, the encoded

proteins were cloned and overexpressed in *E. coli*. This, together with the establishment of recombinant enzymes functional assays led to the characterization of the biosynthesis pathways for dTDP-β-L-Rha*p*,[84] dTDP-α-D-Fuc*p*3NAc,[86] dTDP-α-D-Qui*p*3NAc,[89] GDP-D-*glycero*-α-D-*manno*-heptose,[87] GDP-α-D-Rha*p*,[85] and dTDP-α-D-Fuc*p*[64] in Gram-positive organisms.

Comparable to lipopolysaccharide O-antigen biosynthesis in Gram-negative bacteria, dTDP-β-L-rhamnose is synthesized by *A. thermoaerophilus* DSM 10155/G$^+$ in a four-step reaction sequence from dTTP and glucose-1-phosphate by the enzymes glucose-1-phosphate thymidylyltransferase (RmlA), dTDP-D-glucose-4,6-dehydratase (RmlB), dTDP-4-dehydrorhamnose-3,5-epimerase (RmlC), and dTDP-4-dehydrorhamnose reductase (RmlD).[84] RmlA and RmlB are also involved in the biosynthesis of dTDP-α-D-Fuc*p*3NAc in *A. thermoaerophilus* L420-91^T, in the production of dTDP-α-D-Qui*p*3NAc in *T. thermosaccharolyticum* E207-71, and in the biosynthesis of dTDP-α-D-fucose in *G. tepidamans* GS5-97^T, producing the important key intermediate dTDP-4-dehydro-6-deoxyglucose[96] (Fig. 5). In the first case, the production of dTDP-α-D-Fuc*p*3NAc is completed by the sequential action of a dTDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase (FdtA), a dTDP-3-keto-6-deoxy-D-galactose aminase (FdtB), and a dTDP-α-D-Fuc*p*3N acetylase (FdtC). FdtA was the first isomerase described that is capable of synthesizing dTDP-3-keto-(dehydro)-6-deoxy-galactose from dTDP-4-dehydro-6-deoxyglucose.[86] In the second case, dTDP-4-dehydro-6-deoxyglucose is processed by the enzymes dTDP-4-keto-6-deoxy-D-glucose-3,4-ketoisomerase (QdtA), dTDP-3-keto-6-deoxy-D-glucose aminase (QdtB), and dTDP-D-Qui*p*3N acetylase (QdtC) to form dTDP-α-D-Qui*p*3NAc.[89] In the last case, the biosynthesis of dTDP-α-D-fucose is completed by the dTDP-4-dehydro-6-deoxyglucose reductase (Fcd), using NADH as a cofactor.[64]

The two enzymes responsible for the biosynthesis of the nucleotide activated form of the D-enantiomer of rhamnose, GDP-α-D-rhamnose, are the GDP-D-mannose dehydratase (Gmd) converting GDP-D-mannose to GDP-4-dehydro-6-deoxy-D-mannose with NADPH as cofactor, and the reductase Rmd catalyzing the reduction of the keto-intermediate to the final product using both NADH and NADPH as hydride donor. Gmd was identified as a novel bifunctional enzyme exhibiting both dehydratase and reductase activities.[85]

For biosynthesis of the nucleotide-activated form of D-*glycero*-D-*manno*-heptose in *A. thermoaerophilus* DSM 10155/G$^+$, the four enzymes GmhA, HddA, GmhB, and HddC are required. The isomerase GmhA catalyzes the conversion of D-sedoheptulose-7-phosphate to D-*glycero*-D-*manno*-heptose-7-phosphate and the phosphokinase HddA adds a phosphate group to form D-*glycero*-D-*manno*-heptose-1,7-bisphosphate. The phosphatase GmhB removes the phosphate at the C-7 position, and the intermediate D-*glycero*-α-D-*manno*-heptose-1-phosphate is finally activated with GTP by the pyrophosphorylase HddC to yield the final product GDP-D-*glycero*-α-D-*manno*-heptose. It should be noted that the heptose residue is synthesized as GDP-D-*glycero*-α-D-*manno*-heptose, whereas ADP-L-*glycero*-β-D-*manno*-heptose is the precursor of the inner core

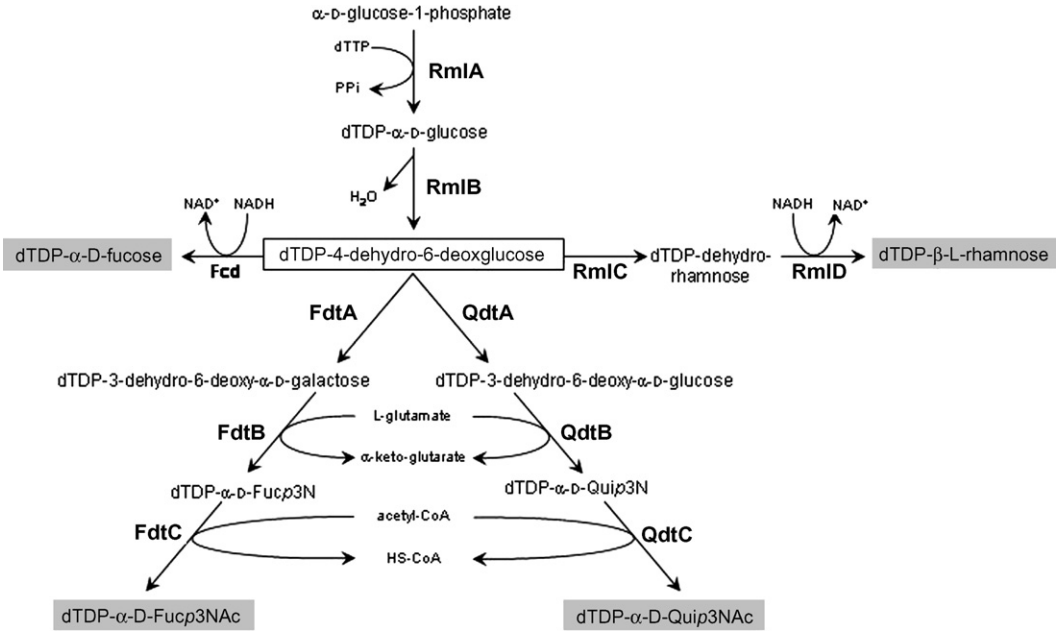

Figure 5. Nucleotide sugar biosynthetic pathways of activated sugars required for S-layer glycoprotein glycan biosyntheses utilizing dTDP-4-dehydro-6-deoxyglucose as key intermediate.

lipopolysaccharide biosynthesis of organisms like *E. coli* or *Salmonella enterica*.[87,88]

From these data, it is obvious that the different S-layer protein glycosylation pathways provide a spectrum of rare enzymes that may be used for glycoengineering purposes in heterologous hosts. Furthermore, some of these enzymes, for example, most Rml enzymes from *A. thermoaerophilus* DSM 10155/G^{+}, exhibit significantly higher stability at 37 °C than the enzymes from the mesophilic strain *S. enterica*, due to their origin from thermophilic organisms. This advantage could lead to the development of improved high-throughput screening systems for specific sugars.[97]

4.3.2. Initiation enzyme. Following the identification of the initiation enzyme as one of the key modules of S-layer protein glycosylation, recently, the initiation enzyme of S-layer glycoprotein biosynthesis has been characterized in the model organism *G. stearothermophilus* NRS 2004/3a.[98] The 471-aa membrane protein WsaP of *G. stearothermophilus* NRS 2004/3a shows high homology to glycosyltransferases, for example, WbaP, which catalyzes the first step in polysaccharide biosynthesis by transferring a hexose-1-phosphate residue from UDP-hexoses (galactose and glucose) to a phosphorylated lipid carrier.[99] To assess the functional domain of the enzyme, different truncated forms of the protein were designed and heterologously expressed in *E. coli*.[98] WsaP is capable of reconstituting K30 antigen and O-antigen biosynthesis in the WbaP-deficient strains *E. coli* CWG 466 and *S. enterica* MSS2, respectively. In vitro assays of isolated membranes of *E. coli* harboring WsaP confirmed the galactosyltransferase activity and the catalytic site is located at the C-terminal half of WsaP, including one transmembrane domain. Undecaprenol phosphate is recognized as an acceptor molecule for WsaP in the *E. coli* background, which is a necessary prerequisite for combining the S-layer protein O-glycosylation system with other polysaccharide biosynthesis pathways for S-layer *neo*glycoprotein production.

5. Proposed S-layer protein glycosylation pathway

It is obvious that the biosynthesis of an S-layer glycoprotein is a very complex process in which the glycosylation event has to be coordinated with the amount of the synthesized S-layer protein, its translocation through the cell wall, and its incorporation into the existing S-layer lattice. Considering that up to 20% of the total cellular

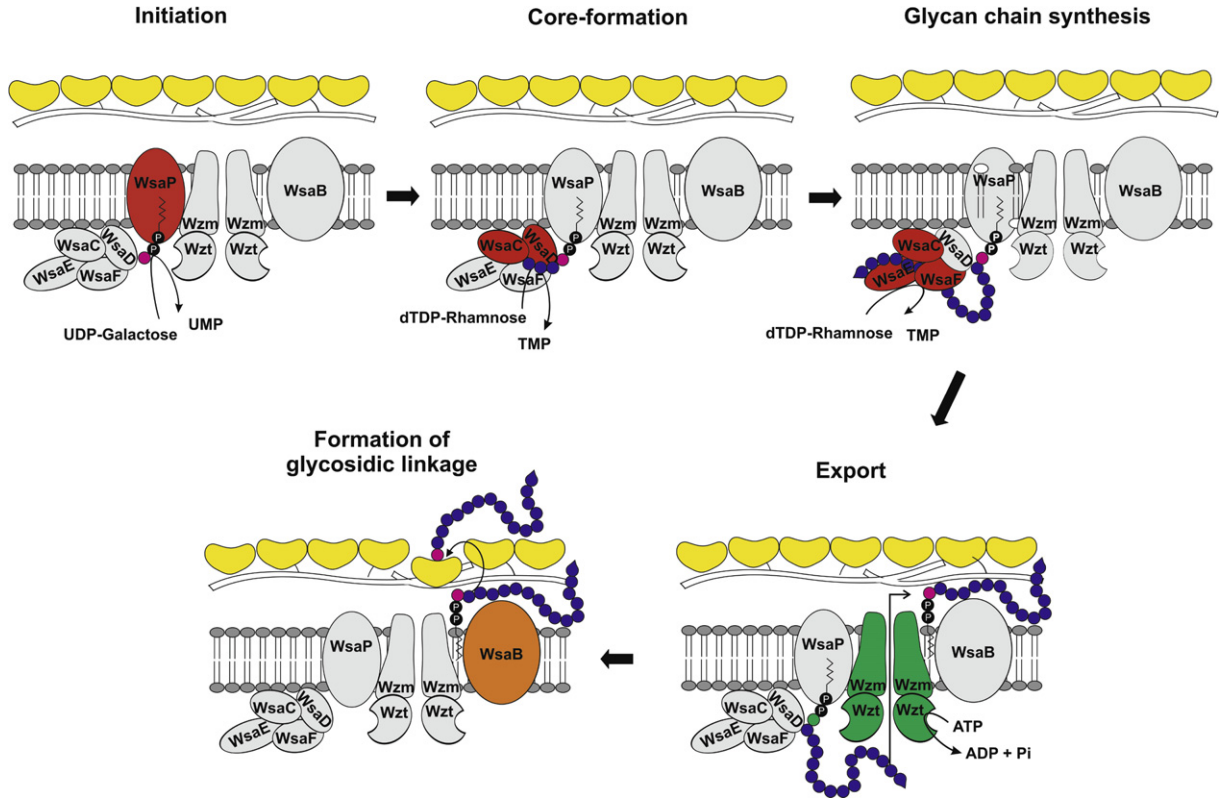

Figure 6. Schematic representation of the proposed biosynthesis route of S-layer glycoproteins, exemplified with *G. stearothermophilus* NRS 2004/3a. The identical color code for the components involved in S-layer glycoprotein glycan biosynthesis is used in Figures 4, 6 and in Table 1: light blue, monosaccharide biosynthesis; red, glycan assembly; orange, glycan transfer; green, export; gray, transposase; black, unknown function; dark blue, elongated glycan chain; pink, linkage glycose; yellow, S-layer protein.

protein synthesis effort of a bacterium may be devoted to S-layer protein production, it is conceivable that the S-layer protein glycosylation machinery has to be very effective. However, preliminary data indicate that the glycosylation event is lagging behind protein production (R. Novotny, A. Scheberl, and C. Schäffer, unpublished data). It is currently neither known whether S-layer glycosylation occurs co- or posttranslationally, nor whether the S-layer target protein is still unfolded or already prefolded when the oligosaccharyl::protein transferase exerts its activity.

Accumulated data on the S-layer nanoglycobiology of the model organism *G. stearothermophilus* NRS 2004/3a have led to the proposal of a pathway for S-layer protein O-glycosylation. The proposed pathway clearly benefits from the knowledge of the two principal routes of LPS O-antigen biosynthesis and of the *C. jejuni* protein N-glycosylation system.[44,49,100] O-Glycosylation of this S-layer glycoprotein generally seems to follow a similar pathway as is described for N-glycosylation of bacterial proteins.[101,102] The presence of a predicted ABC-2-type transporter system and the absence of a putative polymerase in the *slg* gene cluster of *G. stearothermophilus* NRS 2004/3a[83] indicate that the S-layer glycan chains are most probably synthesized in a process comparable to the ABC transporter-dependent pathway of LPS O-polysaccharide biosynthesis, although being initiated by a WbaP homologue instead of a WecA homologue, which usually serves as the initiation enzyme in that pathway.[57] Our current model implicates WsaP in the first step of synthesis whereby galactose is transferred from its nucleotide-activated form (UDP-Gal) to a membrane-associated lipid carrier at the cytoplasmic face of the plasma membrane[98] (Fig. 6). Chain extension presumably would continue in the cytoplasm by processive addition of rhamnose residues from dTDP-β-L-rhamnose to the non-reducing terminus of the lipid-linked glycan chain. Chain growth is predicted to be terminated by 2-O-methylation of the terminal repeating unit, catalyzed by an O-methyltransferase. A similar modification was recently described as chain length termination signal in the biosynthesis of O8 and O9 antigens.[62] The complete glycan chain would then be transported across the membrane by a process involving an ABC transporter and eventually transferred to the S-layer protein by the oligosaccharyl::protein transferase WsaB in a reaction comparable to that described recently for protein N-glycosylation of *C. jejuni*.[44,49,100]

As several other *slg* gene clusters contain an ABC-transporter (Fig. 4),[83] and numerous glycan chains are modified at the non-reducing end (Fig. 2),[60,61,103] the described model for S-layer protein glycosylation might be widely valid. However, the presence of a putative flippase Wzx in the *slg* gene cluster of *T. thermosaccharolyticum* E207-71 might be taken as an indication that S-layer glycoprotein glycans may also be synthesized via a route that is more related to the *wzy*-dependent O-antigen biosynthesis pathway.[57,83]

6. Conclusions

This review article covers recent advances made in the field of bacterial S-layer nanoglycobiology. From the accumulated data, important conclusions can be drawn for the future design of 'functional' glycans on S-layer proteins. The knowledge of the enzyme apparatus involved in S-layer glycoprotein glycan biosynthesis and the understanding of the underlying mechanisms should eventually allow the alteration or the rational design of S-layer protein glycosylation patterns to obtain bioactive S-layer *neo*glycoproteins. In principal, such compounds may be either used for nanobiotechnological bottom-up strategies to build nanoarrays or will be presented by the engineered bacteria in vivo. Controlled surface display of heterologous (glyco)proteins has become an increasingly used strategy in various applications in microbiology, nanobiotechnology, and vaccinology. The common trend of glycoengineering is reflected by several recent review articles on that topic.[104–109] Nanobiotechnology applications of tailored S-layer *neo*glycoproteins may include the fields of receptor mimics, vaccine design, or drug delivery using carbohydrate recognition.

Acknowledgments

We want to acknowledge our technicians Sonja Zayni and Andrea Scheberl, whose contributions to the advancement of S-layer nanoglycobiology by far exceeded routine work. Financial support came from the Austrian Science Fund, Projects P18013-B10 (to P.M.) and P19047-B12 (to C.S.), and the Hochschuljubiläumsstiftung der Stadt Wien, Project H-1809/2006 (to C.S.).

References

1. Houwink, A. L.; Le Poole, J. B. *Physikalische Verhandlungen* **1952**, *3*, 98.
2. Houwink, A. L. *Biochim. Biophys. Acta* **1953**, *10*, 360–366.
3. Sleytr, U. B. *Int. Rev. Cytol.* **1978**, *53*, 1–64.
4. Murray, R. G. E. In Guze, L. B., Ed.; Microbial Protoplasts, Spheroplasts and L-Forms; Williams & Wilkins: Baltimore, 1968; pp 1–18.
5. Glauert, A. M.; Thornley, M. J. *Annu. Rev. Microbiol.* **1969**, *23*, 159–198.
6. Kawata, T. *Electron Microsc.* **1981**, *15*, 115–122.
7. Wieland, F.; Lechner, J.; Bernhardt, G.; Sumper, M. *FEBS Lett.* **1981**, *132*, 319–323.

8. Sleytr, U. B.; Messner, P. *Annu. Rev. Microbiol.* **1983**, *37*, 311–339.
9. Kandler, O.; König, H. In *The Bacteria*; Woese, C. R., Wolfe, R. S., Eds.; Archaebacteria; Academic Press: New York, 1985; Vol. VIII, pp 413–457.
10. Koval, S. F.; Murray, R. G. E. *Microbiol. Sci.* **1986**, *3*, 357–361.
11. Smit, J. In *Bacterial Outer Membranes as Model Systems*; Inouye, M., Ed.; John Wiley & Sons: New York, 1987; pp 343–376.
12. Lechner, J.; Wieland, F. *Annu. Rev. Biochem.* **1989**, *58*, 173–194.
13. Kandler, O.; König, H. In *The Biochemistry of Archaea (Archaebacteria)*; Kates, M., Kushner, D. J., Matheson, A. T., Eds.; Elsevier: Amsterdam, 1993; pp 223–259.
14. Beveridge, T. J. *Curr. Opin. Struct. Biol.* **1994**, *4*, 204–212.
15. Sumper, M.; Wieland, F. T. In *Glycoproteins*; Montreuil, J., Vliegenthart, J. F. G., Schachter, H., Eds.; Elsevier: Amsterdam, 1995; pp 455–473.
16. Mescher, M. F.; Strominger, J. L. *J. Biol. Chem.* **1976**, *251*, 2005–2014.
17. Sleytr, U. B.; Thorne, K. J. I. *J. Bacteriol.* **1976**, *126*, 377–383.
18. Messner, P.; Schäffer, C. In *Progress in the Chemistry of Organic Natural Products*; Herz, W., Falk, H., Kirby, G. W., Eds.; Springer: Wien, 2003; Vol. 85, pp 51–124.
19. Woese, C. R.; Kandler, O.; Wheelis, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 4576–4579.
20. Tsuboi, A.; Uchihi, R.; Tabata, R.; Takahashi, Y.; Hasiba, H.; Sasaki, T.; Yamagata, H.; Tsukagoshi, N.; Udaka, S. *J. Bacteriol.* **1986**, *168*, 365–373.
21. Udaka, S.; Tsukagoshi, N.; Yamagata, H. *Biotechnol. Gen. Eng. Rev.* **1989**, *7*, 113–146.
22. Sleytr, U. B.; Messner, P.; Pum, D.; Sára, M. *Angew. Chem., Int. Ed.* **1999**, *38*, 1034–1054.
23. Sleytr, U. B.; Egelseer, E. M.; Ilk, N.; Pum, D.; Schuster, B. *FEBS J.* **2007**, *274*, 323–334.
24. Sleytr, U. B.; Huber, C.; Ilk, N.; Pum, D.; Schuster, B.; Egelseer, E. M. *FEMS Microbiol. Lett.* **2007**, *267*, 131–144.
25. Sleytr, U. B.; Sára, M.; Pum, D.; Schuster, B.; Messner, P.; Schäffer, C. In *Biopolymers*; Steinbüchel, A., Fahnestock, S. R., Eds.; Polyamides and Complex Proteinaceous Matrices I; Wiley-VCH: Weinheim, 2002; Vol. 7, pp 285–338.
26. Sleytr, U. B.; Sára, M.; Pum, D.; Schuster, B. In *Supramolecular Polymers*; Ciferri, A., Ed., 2nd ed.; Taylor & Francis: Boca Raton, FL, 2005; pp 583–616.
27. Sleytr, U. B.; Sára, M. *Trends Biotechnol.* **1997**, *15*, 20–26.
28. Benz, I.; Schmidt, M. A. *Mol. Microbiol.* **2002**, *45*, 267–276.
29. Sandercock, L. E.; MacLeod, A. M.; Ong, E.; Warren, R. A. J. *FEMS Microbiol. Lett.* **1994**, *118*, 1–8.
30. Logan, S. M. *Microbiology* **2006**, *152*, 1249–1262.
31. Szymanski, C. M.; Logan, S. M.; Linton, D.; Wren, B. W. *Trends Microbiol.* **2003**, *11*, 233–238.
32. Szymanski, C. M.; Wren, B. W. *Nature Rev. Microbiol.* **2005**, *3*, 225–237.
33. Schäffer, C.; Messner, P. *Biochimie* **2001**, *83*, 591–599.
34. Eichler, J.; Adams, M. W. W. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 393–425.
35. Lechner, J.; Sumper, M. *J. Biol. Chem.* **1987**, *262*, 9724–9729.
36. Sumper, M.; Berg, E.; Mengele, R.; Strobl, I. *J. Bacteriol.* **1990**, *172*, 7111–7118.
37. Kalmokoff, M. L.; Jarrell, K. F. *J. Bacteriol.* **1991**, *173*, 7113–7125.
38. Voisin, S.; Houliston, R. S.; Kelly, J.; Brisson, J.-R.; Watson, D.; Bardy, S. L.; Jarrell, K. F.; Logan, S. M. *J. Biol. Chem.* **2005**, *280*, 16586–16593.
39. Schäffer, C.; Messner, P. *Glycobiology* **2004**, *14*, 31R–42R.
40. Schäffer, C.; Wugeditsch, T.; Kählig, H.; Scheberl, A.; Zayni, S.; Messner, P. *J. Biol. Chem.* **2002**, *277*, 6230–6239.
41. Novotny, R.; Pfoestl, A.; Messner, P.; Schäffer, C. *Glyconjugate J.* **2004**, *20*, 435–447.
42. Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta* **1999**, *1473*, 4–8.
43. Spiro, R. G. *Glycobiology* **2002**, *12*, 43R–56R.
44. Wacker, M.; Linton, D.; Hitchen, P. G.; Nita-Lazar, M.; Haslam, S. M.; North, S. J.; Panico, M.; Morris, H. R.; Dell, A.; Wren, B. W.; Aebi, M. *Science* **2002**, *298*, 1790–1793.
45. Feldman, M. F.; Wacker, M.; Hernandez, M.; Hitchen, P. G.; Marolda, C. L.; Kowarik, M.; Morris, H. R.; Dell, A.; Valvano, M. A.; Aebi, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3016–3021.
46. Samuelson, P.; Gunneriusson, E.; Nygren, P.-Å.; Ståhl, S. *J. Biotechnol.* **2002**, *96*, 129–154.
47. Paton, A. W.; Morona, R.; Paton, J. C. *Nat. Med.* **2000**, *6*, 265–270.
48. Kowarik, M.; Young, N. M.; Numao, S.; Schulz, B. L.; Hug, I.; Callewaert, N.; Mills, D. C.; Watson, D. C.; Hernandez, M.; Kelly, J. F.; Wacker, M.; Aebi, M. *EMBO J.* **2006**, *25*, 1957–1966.
49. Wacker, M.; Feldman, M. F.; Callewaert, N.; Kowarik, M.; Clarke, B. R.; Pohl, N. L.; Hernandez, M.; Vines, E. D.; Valvano, M. A.; Whitfield, C.; Aebi, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7088–7093.
50. Schäffer, C.; Novotny, R.; Küpcü, S.; Zayni, S.; Scheberl, A.; Friedmann, J.; Sleytr, U. B.; Messner, P. *Small* **2007**, *3*, 1549–1559.
51. Kandler, O. *Progr. Botany* **1993**, *54*, 1–24.
52. Kärcher, U.; Schröder, H.; Haslinger, E.; Allmaier, G.; Schreiner, R.; Wieland, F.; Haselbeck, A.; König, H. *J. Biol. Chem.* **1993**, *268*, 26821–26826.
53. Lee, S.-W.; Sabet, M.; Um, H.-S.; Yang, J.; Kim, H. C.; Zhu, W. *Gene* **2006**, *371*, 102–111.
54. Fletcher, C. M.; Coyne, M. J.; Bentley, D. L.; Villa, O. F.; Comstock, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2413–2418.
55. Messner, P. *J. Bacteriol.* **2004**, *186*, 2517–2519.
56. Schäffer, C.; Graninger, M.; Messner, P. *Proteomics* **2001**, *1*, 248–261.
57. Raetz, C. R. H.; Whitfield, C. *Annu. Rev. Biochem.* **2002**, *71*, 635–700.
58. Wugeditsch, T.; Zachara, N. E.; Puchberger, M.; Kosma, P.; Gooley, A. A.; Messner, P. *Glycobiology* **1999**, *9*, 787–795.
59. Steiner, K.; Pohlentz, G.; Dreisewerd, K.; Berkenkamp, S.; Messner, P.; Peter-Katalinić, J.; Schäffer, C. *J. Bacteriol.* **2006**, *188*, 7914–7921.
60. Schäffer, C.; Müller, N.; Christian, R.; Graninger, M.; Wugeditsch, T.; Scheberl, A.; Messner, P. *Glycobiology* **1999**, *9*, 407–414.
61. Kählig, H.; Kolarich, D.; Zayni, S.; Scheberl, A.; Kosma, P.; Schäffer, C.; Messner, P. *J. Biol. Chem.* **2005**, *280*, 20292–20299.
62. Clarke, B. R.; Cuthbertson, L.; Whitfield, C. *J. Biol. Chem.* **2004**, *279*, 35709–35718.

63. Steiner, K. *Unravelling Details of the S-Layer Glycoprotein of Geobacillus stearothermophilus NRS 2004/3a*; Doctoral Thesis, University of Natural Resources and Applied Life Sciences: Vienna, Austria, 2006; 146 p.
64. Zayni, S.; Steiner, K.; Pfoestl, A.; Hofinger, A.; Kosma, P.; Schäffer, C.; Messner, P. *Glycobiology* **2007**, *17*, 433–443.
65. Schäffer, C.; Messner, P. *Microbiology* **2005**, *151*, 643–651.
66. Ferner-Ortner, J.; Mader, C.; Ilk, N.; Sleytr, U. B.; Egelseer, E. M. *J. Bacteriol.* **2007**, *189*, 7154–7158.
67. Rünzler, D.; Huber, C.; Moll, D.; Köhler, G.; Sára, M. *J. Biol. Chem.* **2004**, *279*, 5207–5215.
68. Altman, E.; Schäffer, C.; Brisson, J.-R.; Messner, P. *Eur. J. Biochem.* **1995**, *229*, 308–315.
69. Peter-Katalinić, J. *Methods Enzymol.* **2005**, *405*, 139–171.
70. Harvey, D. J. *Exp. Rev. Proteomics* **2005**, *2*, 87–101.
71. Hitchen, P. G.; Dell, A. *Microbiology* **2006**, *152*, 1575–1580.
72. Allmaier, G.; Schäffer, C.; Messner, P.; Rapp, U.; Mayer-Posner, F. J. *J. Bacteriol.* **1995**, *177*, 1402–1404.
73. Bindila, L.; Steiner, K.; Schäffer, C.; Messner, P.; Mormann, M.; Peter-Katalinić, J. *Anal. Chem.* **2007**, *79*, 3271–3279.
74. Küpcü, Z.; März, L.; Messner, P.; Sleytr, U. B. *FEBS Lett.* **1984**, *173*, 185–190.
75. Schäffer, C.; Franck, W. L.; Scheberl, A.; Kosma, P.; McDermott, T. R.; Messner, P. *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 2361–2368.
76. Kosma, P.; Neuninger, C.; Christian, R.; Schulz, G.; Messner, P. *Glycoconjugate J.* **1995**, *12*, 99–107.
77. Kosma, P.; Wugeditsch, T.; Christian, R.; Zayni, S.; Messner, P. *Glycobiology* **1995**, *5*, 791–796, Erratum: Glycobiology, **1996**, *6*, 5.
78. Yoshii, Z. *Proc. 6th Int. Congr. Electron Microsc., Kyoto* **1966**, *2*, 277.
79. Sára, M.; Moser-Thier, K.; Kainz, U.; Sleytr, U. B. *Arch. Microbiol.* **1990**, *153*, 209–214.
80. Altman, E.; Brisson, J.-R.; Messner, P.; Sleytr, U. B. *Biochem. Cell Biol.* **1991**, *69*, 72–78.
81. Messner, P.; Christian, R.; Neuninger, C.; Schulz, G. *J. Bacteriol.* **1995**, *177*, 2188–2193.
82. Whitfield, C. *Annu. Rev. Biochem.* **2006**, *75*, 39–68.
83. Novotny, R.; Schäffer, C.; Strauss, J.; Messner, P. *Microbiology* **2004**, *150*, 953–965.
84. Graninger, M.; Kneidinger, B.; Bruno, K.; Scheberl, A.; Messner, P. *Appl. Environ. Microbiol.* **2002**, *68*, 3708–3715.
85. Kneidinger, B.; Graninger, M.; Adam, G.; Puchberger, M.; Kosma, P.; Zayni, S.; Messner, P. *J. Biol. Chem.* **2001**, *276*, 5577–5583.
86. Pfoestl, A.; Hofinger, A.; Kosma, P.; Messner, P. *J. Biol. Chem.* **2003**, *278*, 26410–26417.
87. Kneidinger, B.; Graninger, M.; Puchberger, M.; Kosma, P.; Messner, P. *J. Biol. Chem.* **2001**, *276*, 20935–20944.
88. Kneidinger, B.; Marolda, C.; Graninger, M.; Zamyatina, A.; McArthur, F.; Kosma, P.; Valvano, M. A.; Messner, P. *J. Bacteriol.* **2002**, *184*, 363–369.
89. Pföstl, A.; Zayni, S.; Hofinger, A.; Kosma, P.; Schäffer, C.; Messner, P. *Biochemical J.* **2008**, *410*, 187–194.
90. Reeves, P. R.; Wang, L. *Curr. Top. Microbiol. Immunol.* **2002**, *264*, 109–135.
91. Jolly, L.; Stingele, F. *Int. Dairy J.* **2001**, *11*, 733–745.
92. Skurnik, M. *Adv. Exp. Med. Biol.* **2003**, *529*, 187–197.
93. Wang, L.; Romana, L. K.; Reeves, P. R. *Genetics* **1992**, *130*, 429–443.
94. Kaplan, J. B.; Perry, M. B.; MacLean, L. L.; Furgang, D.; Wilson, M. E.; Fine, D. H. *Infect. Immun.* **2001**, *69*, 5375–5384.
95. Reeves, P. R.; Hobbs, M.; Valvano, M. A.; Skurnik, M.; Whitfield, C.; Coplin, D.; Kido, N.; Klena, J.; Maskell, D.; Raetz, C. R. H.; Rick, P. D. *Trends Microbiol.* **1996**, *4*, 495–503.
96. Stein, A.; Kula, M.-R.; Elling, R.; Verseck, S.; Klaffke, W. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1748–1749.
97. Ma, Y.; Stern, R. J.; Scherman, M. S.; Vissa, V. D.; Yan, W.; Jones, V. C.; Zhang, F.; Franzblau, S. G.; Lewis, W. H.; McNeil, M. R. *Antimicrob. Agents Chemother.* **2001**, *45*, 1407–1416.
98. Steiner, K.; Novotny, R.; Patel, K.; Vinogradov, E.; Whitfield, C.; Valvano, M. A.; Messner, P.; Schäffer, C. *J. Bacteriol.* **2007**, *189*, 2590–2598.
99. Valvano, M. A. *Front. Biosci.* **2003**, *8*, 452–471.
100. Kowarik, M.; Numao, S.; Feldman, M. F.; Schulz, B. L.; Callewaert, N.; Kiermaier, E.; Catrein, I.; Aebi, M. *Science* **2006**, *314*, 1148–1150.
101. Linton, D.; Dorrell, N.; Hitchen, P. G.; Amber, S.; Karlyshev, A. V.; Morris, H. R.; Dell, A.; Valvano, M. A.; Aebi, M.; Wren, B. W. *Mol. Microbiol.* **2005**, *55*, 1695–1703.
102. Glover, K. J.; Weerapana, E.; Imperiali, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14255–14259.
103. Bock, K.; Schuster-Kolbe, J.; Altman, E.; Allmaier, G.; Stahl, B.; Christian, R.; Sleytr, U. B.; Messner, P. *J. Biol. Chem.* **1994**, *269*, 7137–7144.
104. Endo, T.; Koizumi, S. *Curr. Opin. Struct. Biol.* **2000**, *5*, 536–541.
105. Mendez, C.; Salas, J. A. *Trends Biotechnol.* **2001**, *11*, 449–456.
106. Saxon, E.; Bertozzi, C. R. *Annu. Rev. Cell Dev. Biol.* **2001**, *17*, 1–23.
107. Desvaux, M.; Dumas, E.; Chafsey, I.; Hébraud, M. *FEMS Microbiol. Lett.* **2006**, *256*, 1–15.
108. Campbell, C. T.; Sampathkumar, S.-G.; Yarema, K. J. *Mol. BioSyst.* **2007**, *3*, 187–194.
109. Thibodeaux, C. J.; Melançon, C. E.; Liu, H.-w. *Nature* **2007**, *446*, 1008–1016.