Published in final edited form as:

Small. 2008 October ; 4(10): 1728–1740. doi:10.1002/sml.200701215.

Recombinant Glycans on an S-Layer Self-Assembly Protein: A New Dimension for Nanopatterned Biomaterials

Dr. Kerstin Steiner^[+],

University of Natural Resources and Applied Life Sciences Center for NanoBiotechnology
Gregor-Mendel-Strasse 33, A-1180 Wien (Austria)

Angelika Hanreich,

University of Natural Resources and Applied Life Sciences Center for NanoBiotechnology
Gregor-Mendel-Strasse 33, A-1180 Wien (Austria)

Birgit Kainz,

University of Natural Resources and Applied Life Sciences Center for NanoBiotechnology
Gregor-Mendel-Strasse 33, A-1180 Wien (Austria)

Dr. Paul G. Hitchen,

Division of Molecular Biosciences, Faculty of Life Sciences Imperial College London, London
SW7 2AZ (UK)

Prof. Anne Dell,

Division of Molecular Biosciences, Faculty of Life Sciences Imperial College London, London
SW7 2AZ (UK)

Prof. Paul Messner, and

University of Natural Resources and Applied Life Sciences Center for NanoBiotechnology
Gregor-Mendel-Strasse 33, A-1180 Wien (Austria)

Prof. Christina Schäffer*

University of Natural Resources and Applied Life Sciences Center for NanoBiotechnology
Gregor-Mendel-Strasse 33, A-1180 Wien (Austria)

Abstract

Crucial biological phenomena are mediated through carbohydrates that are displayed in a defined manner and interact with molecular scale precision. We lay the groundwork for the integration of recombinant carbohydrates into a “biomolecular construction kit” for the design of new biomaterials, by utilizing the self-assembly system of the crystalline cell surface (S)-layer protein SgsE of *Geobacillus stearothermophilus* NRS 2004/3a. SgsE is a naturally O-glycosylated protein, with intrinsic properties that allow it to function as a nanopatterned matrix for the periodic display of glycans. By using a combined carbohydrate/protein engineering approach, two types of S-layer neoglycoproteins are produced in *Escherichia coli*. Based on the identification of a suitable periplasmic targeting system for the SgsE self-assembly protein as a cellular prerequisite for

protein glycosylation, and on engineering of one of the natural protein O-glycosylation sites into a target for N-glycosylation, the heptasaccharide from the AcrA protein of *Campylobacter jejuni* and the O7 polysaccharide of *E. coli* are co- or post-translationally transferred to the S-layer protein by the action of the oligosaccharyltransferase PglB. The degree of glycosylation of the S-layer neoglycoproteins after purification from the periplasmic fraction reaches completeness. Electron microscopy reveals that recombinant glycosylation is fully compatible with the S-layer protein self-assembly system. Tailor-made (“functional”) nanopatterned, self-assembling neoglycoproteins may open up new strategies for influencing and controlling complex biological systems with potential applications in the areas of biomimetics, drug targeting, vaccine design, or diagnostics.

Keywords

biomimetics; composites; protein engineering; self-assembly; S-layers

1. Introduction

Molecular self-assembly systems that exploit the manufactory precision of biological systems are prime candidates for use in nanobiotechnology. Crystalline-cell surface layers (S-layers) of prokaryotic organisms are a very potent self-assembly system, which can be used in bottom-up processes as a patterning element for a “biomolecular construction kit”.^[1–5] The unique property of isolated S-layer protein subunits (monomers) to reassemble into two-dimensional (2D) crystals that are identical to those found on intact bacterial cells, either in suspension, on diverse solid supports, on liposomes, or on various interfaces, opens a wide range of S-layer-based applications.^[6–9] In previous studies, it has been demonstrated that several functional domains introduced into S-layer proteins at distinct positions by genetic engineering do not interfere with the intrinsic self-assembly property of the S-layer system, while simultaneously fully retaining their specific biological properties.^[2,4,10]

A novel line of development is directed toward the integration of carbohydrate compounds (glycans) into this biomolecular construction kit through co- or post-translational fusion of the carbohydrates to the self-assembly S-layer protein matrix, leading to the production of self-assembled S-layer neoglycoproteins.^[11] Considering that glycans are ubiquitous biomolecules, which, in many cases, are key to protein function,^[12–14] it is evident that glycans are useful means for addressing various questions in the fields of basic and applied research, relating to the areas of biomimetics, drug targeting, vaccine design,^[15] or diagnostics.^[16,17] Vital cellular functions that are regulated or influenced by glycans include, for instance, recognition, signaling, trafficking, biological half life, and adhesion events; several immunological phenomena are also enabled and enhanced through glycan “signals” on proteins. Thus, engineering of tailor-made (“functional”) glycoproteins (neoglycoproteins) will decisively change our capability to influence and control complex biological systems. In this context, high-density, periodic, and controllable display of glycans, which through random chemical coupling reactions of glycans to various supports

cannot be fully accomplished, will play a pivotal role. For this demand, the S-layer protein self-assembly system offers an attractive solution.

In this study, the S-layer protein SgsE from *Geobacillus stearothermophilus* NRS 2004/3a (GenBank accession number AF328862) is exploited as a matrix for the high-density display of glycans in a nanometer-scaled, periodic, way using a combined carbohydrate/protein engineering approach. SgsE is a 903-amino acid protein, which includes a 30-amino acid signal peptide aligned by an entropy-driven process into a 2D crystalline array^[8] with oblique (p2) symmetry exhibiting nanometer-scaled periodicity (lattice parameters: $a = 11.6$ nm, $b = 9.4$ nm, $\gamma \approx 78^\circ$).^[18,19] SgsE was chosen for proof of concept because it is naturally O-glycosylated with long-chain poly-L-rhamnans linked via a β -D-galactose residue to the amino acids threonine₅₉₀, threonine₆₂₀, and serine₇₉₄ of the S-layer protein.^[19,20] The protein-inherent glycosylation sites are predicted to form a loop structure that is spatially accessible to the glycosylation machinery of the bacterium. Thus, these sites offer ideal targets for engineering of specific glycosylation sequences that would be recognized by heterologous oligosaccharyl:protein transferases, as required for S-layer neoglycoprotein design.

Glycans are generally synthesized in a complex, template-independent way that has let them escape (nano)biotechnological applications so far. Our understanding of the molecular details of the protein glycosylation process and the availability of the biosynthesis machineries for several glycans in the form of chromosomal gene clusters are prerequisites for the production of S-layer neoglycoproteins in bacterial hosts.^[11] Recent demonstrations of the functional transfer of protein glycosylation pathways into the experimental model organism *Escherichia coli* at the genetic level opens new avenues for carbohydrate engineering of S-layer proteins.^[21,22] Based on the availability of molecular tools, the Pgl protein N-glycosylation system from *Campylobacter jejuni*^[23] and the *E. coli* O7 antigen biosynthesis system^[24] were used for the design of SgsE-neoglycoproteins. The *C. jejuni* Pgl enzymes synthesize a heptasaccharide with the structure D-GalNAc- α 1,4-D-GalNAc- α 1,4-(D-Glc- β 1,3)-D-GalNAc- α 1,4-D-GalNAc- α 1,4-D-GalNAc- α 1,3-D-Bac, where Bac is 2,4-diacetamido-2,4,6-trideoxy-D-Glc,^[25] involving the lipid carrier undecaprenylpyrophosphate.^[26] The heptasaccharide is transferred by the PglB oligosaccharyl:protein transferase to asparagine residues present in the target protein consensus sequence D-X-N-Z-S/T (where X and Z are any amino acid except proline).^[27,28] Using *E. coli* as a host, it was shown that PglB has relaxed substrate specificity, transferring several O-antigen polysaccharides, carrying a 2-acetamido modification at the reducing-end sugar of the glycan, to distinct protein glycosylation sites.^[22] Among the glycans tested in the course of this previous study was also the *E. coli* O7 antigen with the repeating unit structure [3-D-VioNAc- β 1,2-(L-Rha- α 1,3)-D-Man- α 1,4-D-Gal- α 1,3-D-GlcNAc- α 1-]_n, where VioNAc is 2-acetamido-2,6-dideoxy-D-Glc. In Gram-negative bacteria, the oligosaccharyl:protein transferase is proposed to act in the periplasm^[21] in a mechanism, which is, in contrast to eukaryotes, independent of the protein translocation machinery, and is able to glycosylate folded and unfolded proteins.^[29]

The aim of this study is the transfer of the *C. jejuni* heptasaccharide and the *E. coli* O7 polysaccharide onto the SgsE S-layer protein of *G. stearothermophilus* NRS 2004/3a, and

the successful expression of the S-layer neoglycoproteins in *E. coli*. For proof of concept we specifically deal with: i) the establishment of a periplasmic targeting system for the S-layer protein as a prerequisite for S-layer neoglycoprotein production in *E. coli*,^[30] ii) engineering of an *N*-glycosylation site in the naturally *O*-glycosylated S-layer protein to serve as a potential target for the PglB oligosaccharyl:protein transferase to act on, iii) the unequivocal proof of the neoglycoprotein nature of the engineered products, and, most importantly, iv) the demonstration of the self-assembly capability of the produced nanostructured composites.

2. Results and Discussion

2.1. Periplasmic Targeting of SgsE

The export of SgsE to the periplasm of *E. coli* is a prerequisite for the biosynthesis of S-layer neoglycoproteins, because in Gram-negative bacteria the final step of protein glycosylation takes place in the periplasm. Currently, there is nothing known about the mode of translocation used by the S-layer protein SgsE to travel from the cytoplasm across the cytoplasmic membrane of its native host, which is *G. stearotheophilus* NRS 2004/3a, to reach the cell surface as its final destination.

At the point of the conception of this study it was clear from recombinant SgsE production in heterologous organisms such as *E. coli* or *Lactococcus lactis*^[31] that the intrinsic force to promote S-layer protein self-assembly was so strong that the cytoplasm of these organisms was fully packed with assembled protein. This, however, is undesired for recombinant S-layer protein glycosylation. Thus the suitability of different signal peptides to export SgsE from the cytoplasm, where it is synthesized, to the periplasm was investigated, including signal peptides from different Gram-negative organisms, which generally mediate the export of either unfolded proteins (Sec-secretory pathway) or folded proteins (twin arginine translocation (TAT) secretory pathway). More specifically, the signal sequence of pectate lyase B (ssPelB) from *Erwinia carotovora*, the maltose-binding protein (MBP) of *E. coli* including its signal peptide, which both mediate protein export via the Sec-secretory pathway, as well as the signal peptide TorA (ssTorA) of the trimethylamine N-oxide (TMAO) reductase from *E. coli*, which supports the TAT-secretory pathway, were translationally fused to the N-terminus of the S-layer protein. In the course of this study, two distinct forms of SgsE were compared with regard to their suitability as a protein target for recombinant glycosylation. These are the forms A_SgsE, corresponding to the mature S-layer protein devoid of its native signal peptide (A_SgsE₃₁₋₉₀₃), and the truncated protein G_SgsE, corresponding to a 330-amino acid N-terminal truncation (G_SgsE₃₃₁₋₉₀₃). Either form of SgsE has been shown to possess an excellent self-assembly capability.^[2] G_SgsE was included in this study because it can be imagined that, due to its smaller size compared to the mature protein, its export to the periplasm will be facilitated. In another neoglycoprotein production experiment, ssPelB and ssTorA were successfully used for the transfer of the *N*-glycosylation target protein AcrA of *C. jejuni* to the periplasm of *E. coli*, demonstrating that bacterial *N*-glycosylation can occur independently of the protein translocation machinery.^[21,29] Concerning periplasmic targeting of S-layer proteins in

general, MBP has been reported to export the S-layer protein SbsA of *G. stearothermophilus* PV/72 to the periplasm of *E. coli*.^[32]

The different S-layer fusion proteins were expressed in *E. coli* BL21 Star (DE3) carrying the respective expression plasmids. Based on the masses of A_SgsE and G_SgsE of 93.7 and 61.0 kDa, respectively, the masses of the different fusion proteins were calculated to be 95.9 and 63.2 kDa for the ssPelB constructs, 139.2 and 106.5 kDa for the MBP constructs, and 98.4 and 65.7 kDa for the ssTor constructs. This was in accordance with the sodium dodecylsulfate–polyacrylamide electrophoresis (SDS–PAGE) evidence (Figure 1, lanes 3–8). Using ssPelB and ssTor, both A_SgsE and G_SgsE could be expressed in high yield, whereas the fusion with MBP resulted in lower protein yield, with the A_SgsE form being detectable only by Western blotting using anti-SgsE antibody (not shown). To investigate if the native signal peptide of SgsE could mediate export of the S-layer protein into the periplasm of *E. coli*, full length SgsE including its signal peptide was introduced into the expression vector pET28a and produced in *E. coli* BL21 Star (DE3). According to SDS–PAGE analysis, the S-layer protein possessed the expected mass of 96.6 kDa and was produced in high yield (Figure 1, lane 2).

The subcellular localization of the S-layer protein was investigated by electron microscopy after immunogold-labeling (anti-SgsE antibody) of thin-sectioned cells of *E. coli* BL21 Star (DE3) expression cultures. To compare the export efficiency mediated via the different signal peptides, gold particles present in the cytoplasmic and in the periplasmic compartment were counted from electron micrographs. Averaging the values obtained from five cells, each from separate labeling experiments, the approximate periplasm to cytoplasm ratio is 0.3:1 for the ssTor constructs, 1:1 for full-length SgsE possessing its native signal peptide, 2:1 for the ssPelB constructs, and 3:1 for the MBP constructs. The periplasmic-to-cytoplasmic ratio was invariant for different cell sizes. Interestingly, no significant difference in the export efficiency between A_SgsE and G_SgsE was observed, indicating that S-layer protein length does not affect S-layer protein translocation through the cytoplasmic membrane. Thus, SgsE can be targeted to the periplasm by the PelB- and the MBP-signal peptides as well as by its native signal peptide, but not with ssTorA. This indicates that SgsE is exported in an unfolded state via the Sec-secretory pathway. As the native signal peptide was also recognized by the export machinery of *E. coli*, we assume also that in the in vivo background of *G. stearothermophilus* NRS 2004/3a the export of SgsE follows the Sec-pathway. This has already been described for the S-layer proteins of other Gram-positive organisms such as *Clostridium difficile*^[33] and *Corynebacterium glutamicum*.^[34]

For the following S-layer neoglycoprotein production, the signal peptide of PelB was chosen for periplasmic expression of SgsE, because it was the best compromise between export efficiency and product yield (Figure 2). In addition, ssPelB is a small peptide of only 22 amino acids, which should be cleaved off in vivo after the export of the protein to the periplasm and, importantly, the ssPelB–SgsE fusion proteins were shown to fully maintain the self-assembly capability of the native S-layer protein (this study, not shown).

2.2. Design of an SgsE Neoglycoprotein Carrying a *C. jejuni* Heptasaccharide

For the first time, an S-layer neoglycoprotein has been designed, comprising the S-layer protein SgsE from *G. stearotheophilus* NRS 2004/3a, including the PelB signal peptide for periplasmic targeting, and the heptasaccharide Glc(GalNAc)₅Bac of *C. jejuni*. For this purpose, the plasmid pACYCp_{gl} harboring the complete *pgl* gene cluster of *C. jejuni* that is responsible for heptasaccharide biosynthesis and its transfer to the protein was transformed into *E. coli* BL21 Star (DE3).^[21] As a positive control, to confirm that the glycosylation machinery of *C. jejuni* functions under the chosen conditions, the plasmid pEC(AcrA-per), expressing periplasmic soluble AcrA of *C. jejuni* with the PelB signal peptide and a C-terminal hexahistidine tag, was transferred to electrocompetent *E. coli* BL21 Star (DE3) cells containing pACYCp_{gl}.^[29] Analysis of the expression product by Western immunoblotting, using a heptasaccharide specific antibody (anti-p_{gl} antibody), confirmed glycosylation of AcrA (Figure 3, lane 8).

In *C. jejuni*, Glc(GalNAc)₅Bac is naturally linked to distinct *N*-glycosylation sites of the protein AcrA,^[21] while the S-layer protein SgsE possesses *O*-glycosidically linked glycan chains. Inspection of the S-layer protein sequence revealed the presence of one putative *N*-glycosylation site at position N₈₇₉ (DVNQT; the glycosylated asparagine residue is underlined), conforming with the amino acid sequence requirement of the oligosaccharyl:protein transferase PglB of *C. jejuni*, which is the key enzyme for protein *N*-glycosylation.^[22,28,35] Thus, in a first glycosylation approach, the potential of PglB to glycosylate native SgsE with its endogenous heptasaccharide Glc(GalNAc)₅Bac substrate in *E. coli* was investigated by expression of ssPelB–A_SgsE from plasmid pET22b–A_SgsE in *E. coli* BL21 Star (DE3) harboring pACYCp_{gl}.^[21] However, no glycosylated protein was visible on a Western immunoblot, implying that the putative *N*-glycosylation site of SgsE at position N₈₇₉ was not recognized by PglB (not shown). The same negative result was obtained when testing a second putative *N*-glycosylation site (dNNVS; the mutated amino acid is underlined) located at the rather C-terminal position of SgsE (amino acid N₈₉₃), which was created by site-directed mutagenesis of G₈₉₁D using a mutant primer (with pET22b–A_SgsE_G₈₉₁D being the respective expression plasmid). Another set of *N*-glycosylation experiments utilized the *N*-glycosylation consensus sequence DFNRS of the glycoprotein AcrA of *C. jejuni*, or the longer sequence ASKDFNRSKALFS, instead of SgsE inherent sequences. Initially, either sequence was translationally fused to the N-terminus of the truncated S-layer protein G_SgsE, based on the assumption that the N-terminus of SgsE would be spatially accessible for the PglB enzyme. This is because the N-terminus is proposed to mediate anchoring of the S-layer protein to the cell wall in *G. stearotheophilus* NRS 2004/3a, implying its spatial exposure to the environment. However, when introducing the respective fusion proteins encoded by pET22b–DFNRS_G_SgsE and pET22b–long_G_SgsE into *E. coli* BL21 Star (pACYCp_{gl}), again no S-layer neoglycoprotein production occurred (not shown). Finally, the endogenous *O*-glycosylation sites of SgsE were engineered to become *N*-glycosylation target sequences. By this strategy, the amino acid sequence adjacent to the naturally *O*-glycosylated amino acids Thr₆₂₀ or Ser₇₉₄ of SgsE, or both of them, was replaced by the *N*-glycosylation consensus sequence DFNRS or ASKDFNRS–KALFS, respectively. For this purpose, an intermediate form of SgsE lacking the sequence between amino acids L₆₁₀ and E₈₀₄ was constructed,

with the corresponding nucleotide sequence having an artificial KpnI restriction site introduced between the amino acids at the position 611 and 612. After introducing the respective *N*-glycosylation sequence, while simultaneously deleting part of *sgsE* by the use of megaprimers for polymerase chain reaction (PCR) and subsequent ligation into the intermediate *sgsE* form provided through plasmid pET22b-A_SgsE_inter, the sequence DFNRS replaced K₆₁₈TTSD₆₂₂ (pET22b-A_SgsE_T5) or L₇₉₂TSAD₇₉₆ (pET22b-A_SgsE_S5), or both of them (pET22b-A_SgsE_S,T5), and the sequence ASKDFNRSKALFS replaced A₆₁₅DLKTTSDNFKLY₆₂₇ (pET22b-A_SgsE_T12), or T₇₈₉ATLTSADVIRVD₉₀₁ (pET22b-A_SgsE_S12), or both of them (pET22b-A_SgsE_S,T12). Due to the cloning procedure, all chimeric S-layer proteins contained a C-terminal hexahistidine tag, which was used for determination of the molecular mass of the constructs in Western blot analysis. The signals obtained with anti-His-antibody at approximately 96 kDa were in accordance with the expected masses (not shown).

S-layer neoglycoprotein production was induced by isopropyl- β -thiogalactopyranoside (IPTG) in *E. coli* BL21 Star (pACYCpgl) harboring the different expression plasmids coding for the modified S-layer proteins. According to the Western immunoblot evidence using anti-pgl antibody (Figure 3, lanes 2–7), after an expression period of 20 h, the best *N*-glycosylation result was obtained with the S-layer protein in which the endogenous Thr₆₂₀-*O*-glycosylation site had been replaced by the extended *C. jejuni* *N*-glycosylation sequence ASKDFNRSKALFS (Fig. 3, lane 5). After prolonged development of the Western blot, weak bands were also visible for the S-layer protein forms, where this site had been replaced by DFNRS or in which the Thr₆₂₀ and Ser₇₉₄ sites had been simultaneously replaced by the *N*-glycosylation target sequence (data not shown). Unexpectedly, no glycosylation signal was obtained when using the S-layer protein forms in which the S₇₉₄ glycosylation site had been replaced by the *N*-glycosylation target sequence (Fig. 3, lanes 6 and 7). The same results were obtained for the modified forms of G_SgsE for which, due to the higher expression level, a higher probability of glycosylation could have been expected (data not shown). This indicates that the consensus sequence alone is not sufficient for *N*-glycosylation, but the spatial environment is also important.^[29,36] For the bacterial *N*-glycosylation system it is proposed that the glycosylation sites are located in flexible parts of folded proteins.^[29] Rangarajan and coworkers^[36] elucidated the structure of PEB3, a glycosylated adhesin from *C. jejuni*, and showed that the glycosylation site is located within a surface-exposed loop joining two structural elements. While the tertiary structure of SgsE is unknown, secondary structure prediction using PSIPREDView^[37] revealed that Thr₆₂₀ and Ser₇₉₄ are located in a loop between two protein strands, thereby fulfilling the general requirement for bacterial glycosylation. Considering the negative result when using the engineered Ser₇₉₄ *N*-glycosylation target sequence, it is interesting to note that the pI value of the native amino acid sequence surrounding Ser₇₉₄ of 4.2 (T₇₈₉ATLTSADVIRVDFS₉₀₃) is markedly different from that of the introduced ASKDFNRSKALFS *N*-glycosylation sequence (pI = 10.0; A₇₈₉SKDFNRSKALFS₉₀₃). In contrast, for the engineered consensus sequence around the Thr₆₂₀ site that was recognized by the heterologous PglB enzyme, there was only a slight difference in pI values between the native (pI = 8.4; K₆₁₁AVGADLKTSDNFKLY₆₂₇) and the engineered (pI = 9.9; G₆₁₁TVGASKDFNRSKALFS₆₂₇) S-layer protein.

These results indicate that the requirements for protein glycosylation in heterologous systems are complex. Thus, the availability of a well-characterized naturally glycosylated S-layer protein such as SgsE from *G. stearotherophilus* NRS 2004/3a is very valuable for engineering of tailor-made glycosylation target sequences that will be recognized by oligosaccharyltransferases from diverse glycosylation systems to produce self-assembly S-layer neoglycoproteins.

2.3. Purification of the SgsE Neoglycoprotein Carrying the *C. jejuni* Heptasaccharide

The SgsE neoglycoprotein carrying the *C. jejuni* heptasaccharide (named A_SgsE_T12-pgl) was isolated according to a standard S-layer purification protocol that had been adjusted to the periplasmic location of the S-layer neoglycoprotein, and subsequently purified by gel-filtration chromatography on a Superdex 200 prep grade column using 6 M urea as eluent. The purification process was monitored by SDS-PAGE and Western blot analyses using anti-SgsE and anti-pgl antibodies (Figure 4). It was clearly visible that the S-layer neoglycoprotein could be enriched upon purification, resulting in a final yield of 13.4 mg L⁻¹ of expression culture. In intact cells of an *E. coli* BL21 Star (pACYCpgl, pET22b-A_SgsE_T12) expression culture (Figure 4, lane 2) the amount of S-layer neoglycoprotein seems to be low, whereas the protein band in the periplasmic fraction and in the subsequently separated fractions of the dialysate and the soluble material are similarly strong, upon Coomassie staining as well as upon development with both anti-SgsE and anti-pgl antibodies (Figure 4, lanes 3–5). This result indicates on the one hand that the A_SgsE_T12 target sequence is not fully exported to the periplasm, but on the other hand it also clearly shows that the exported S-layer protein is almost fully glycosylated. To verify that periplasmic targeting was mediated by ssPelB, the purified S-layer neoglycoprotein was subjected to N-terminal sequencing. Unexpectedly, the mature S-layer neoglycoprotein revealed the sequence AQPA, corresponding to the four C-terminal amino acids of the signal peptide, which means that the PelB signal peptide was only partially cleaved off upon export of the protein to the periplasm. This finding, however, has no consequence for the functionality of the established S-layer neoglycoprotein production system.

2.4. MS Analysis of the SgsE Neoglycoprotein Carrying a *C. jejuni* Heptasaccharide

Mass spectrometry was performed in order to confirm that the complete Glc(GalNAc)₅Bac N-glycan had been transferred onto the N-glycosylation site of A_SgsE_T12. Purified A_SgsE_T12-pgl neoglycoprotein was digested with trypsin and the resulting peptides were analyzed by nanoliquid-chromatography (nano-LC) matrix-assisted laser-desorption ionization-time of flight (MALDI TOF)/TOF mass spectrometry (MS). The resulting MS and MS/MS data were analyzed for the predicted A_SgsE_T12-pgl tryptic glycopeptide. A signal observed in the MS with a mass-to-charge ratio (m/z) of 1956.9¹⁺ was identified as having the correct mass for the modified peptide (data not shown). Analysis of the resulting product ion spectrum (Figure 5) clearly shows a series of singly charged fragment ions consistent with glycosidic cleavage products from the tryptic peptide DFNR modified with a heptasaccharide glycan moiety, with sufficient information to assign the sequence and branching pattern as observed for the *C. jejuni* N-glycan (Fig. 5, inset).^[21]

2.5. Self-Assembly of SgsE Neoglycoprotein Carrying the *C. jejuni* Heptasaccharide

After the glycoprotein nature of the A_SgsE_T12-pgl neoglycoprotein was confirmed, self-assembly of the recombinant construct was investigated. Following a protocol elaborated in the course of a recent study,^[2] self-assembly of the monomeric S-layer neoglycoprotein as obtained after purification was triggered upon removal of the chaotropic agent, by dialysis against distilled water. Self-assembly was investigated by electron microscopy using negatively stained preparations. An electron micrograph of the negatively stained A_SgsE_T12-pgl neoglycoprotein is shown in Figure 6. Due to the low signal-to-noise ratio only one base vector pair is revealed. Since only second order reflections were unambiguously detectable, no reconstruction of the ultrastructure of this neoglycoprotein meshwork was possible at high resolution. The length of the base vectors, however, compared well with previous results of negatively stained native S-layer glycoprotein from *G. stearotherophilus* NRS 2004/3a^[18] as well as with the lattice parameters of self-assembled rSgsE protein ($a = 11.6$ nm, $b = 9.4$ nm, $\gamma \approx 78^\circ$).^[2] Comparing that material with the S-layer neoglycoprotein produced in the course of the present study, a slightly changed staining behavior of the latter was observed on the micrographs, with the lattice of the neoglycoprotein apparently blurred to some extent. Despite this fact, in self-assembled A_SgsE_T12-pgl the SgsE S-layer protein matrix itself kept its crystallinity. Thus, recombinant SgsE S-layer protein glycosylation is fully compatible with the intrinsic self-assembly property of the protein matrix.

Due to the lack of a suitable microscopic method for visualization of the periodic, nanometer-scale, naturally outwardly-orientated display of heptasaccharides on a self-assembled A_SgsE_T12-pgl neoglycoprotein monolayer, which is conferred to this new nanostructured composite by the S-layer protein portion, we used a combined electron microscopy–modeling approach. In this approach, a negatively stained preparation of the S-layer protein self-assembled in solution served as a base for image reconstruction, and space-filling models of the Glc(GalNAc)₅Bac heptasaccharides were positioned onto the subunits of the SgsE nanolattice (Figure 7A). The resulting model is an in-scale visualization of the first self-assembled S-layer neoglycoprotein nanolattice. In this nanolattice, each constituting subunit carries one heptasaccharide at the engineered Thr₆₂₀ position (Thr₆₂₀ was assumed at a surface-exposed, but fictive, position in the model). The periodicity of glycan display results from the base vectors of the SgsE p2 lattice, which are 11.6 and 9.4 nm, respectively. From detailed self-assembly studies of the SgsE matrix and of functionalized SgsE S-layer protein, it can be expected that self-assembly S-layer neoglycoprotein nanolattices do not only work in solutions, but that they are also suitable for coating of various planar solid supports, liposomes,^[2] or even porous structures like membranes.^[38] This flexibility offers a wide repertoire of opportunities for integration of glycan-mediated interactions in vitro as well as in vivo systems, as may be relevant to address basic and applied questions both in life and nonlife sciences. A potent example for in vivo carbohydrate engineering with a direct applicational aspect was the construction of a recombinant *E. coli* bacterium that displayed a modified lipopolysaccharide (LPS) mimicking a Shiga toxin receptor on its cell surface.^[39]

2.6. Transfer of *E. coli* O7 Antigen to the Engineered N-Glycosylation Site of SgsE

To demonstrate the wider applicability of the engineered SgsE self-assembly matrix for recombinant glycosylation, the O7 antigen from *E. coli* was transferred onto the S-layer protein A_SgsE-T12. This approach is based on the finding of Feldman and coworkers,^[22,35] who discovered that the oligosaccharyl:protein transferase PglB of *C. jejuni* is able to transfer glycan chains with different 2-acetamido sugars at the reducing end to the N-glycosylation sites of proteins.^[22,29,35] In our study, we investigated the possibility that PglB can also transfer glycan chains with GlcNAc at the reducing end (as present in O7) onto the engineered Thr₆₂₀ N-glycosylation site on SgsE.

In this experiment, the *E. coli*-K12 mutant strain CLM24 that lacks the O-antigen ligase and, in addition, does not synthesize O-antigen due to an inactivating insertion in *wbbL*, the gene that encodes a rhamnosyl transferase necessary for the transfer of the second sugar of the O7 antigen, was used as expression host. This strain should favor the PglB-mediated transfer of O-antigen from its carrier to the protein acceptor, because of accumulation of undecaprenylpyrophosphate-linked polysaccharide.^[22] PglB was expressed under the control of the arabinose-inducible BAD promoter from plasmid pMAF10.^[22] The two modified S-layer protein forms A_SgsE_T12 and G_SgsE_T12 with the PelB signal peptide, the extended N-glycosylation site from *C. jejuni* replacing the O-glycosylation site T₆₂₀ and a C-terminal hexahistidine-tag were cloned from pET22b into pBAD24 vector, which is arabinose inducible. The gene cluster necessary for the synthesis of *E. coli* O7 antigen provided through plasmid pJHCV32^[24] was introduced in *E. coli* CLM24 together with PglB and A_SgsE_T12 or G_SgsE_T12. After induction of protein expression with arabinose over night, the SgsE neoglycoprotein was analyzed by Western immunoblotting using anti-SgsE-antibody. In addition to bands at 97.0 and 64.0 kDa, corresponding to nonglycosylated A_SgsE_T12 and G_SgsE_T12, respectively, a ladder-like banding pattern in a higher molecular weight range of ≈ 130 –150 kDa for A_SgsE_T12-O7 and of 75–100 kDa for G_SgsE_T12-O7 could be detected (Figure 8, lanes 1 and 2). The specific banding pattern of each neoglycoprotein displays differences in the degree of polymerization of individual O7 repeating units, which is characteristic of O-antigens.

We have visualized the display of polymerized *E. coli* [VioNAc(Rha)ManGalGlcNAc]_n O7 antigen via the SgsE self-assembly matrix using the same approach as described above for the *C. jejuni* heptasaccharide (Figure 7B). The recombinant transfer not only of short-chain carbohydrates, such as the *C. jejuni* heptasaccharide, but also of elongated glycan chains, such as O-antigens, upon maintenance of the self-assembly property of the S-layer protein matrix, is highly relevant for the future conceptualization of functional S-layer neoglycoproteins. It can be imagined that this approach is mimicking the natural situation in *G. stearothermophilus* NRS 2004/3a, where long-chain poly-L-rhamnans are attached to the distinct O-glycosylation sites of the S-layer protein. Benefits of self-assembly S-layer neoglycoproteins carrying elongated functional glycan chains may be expected in the field of disease intervention and prevention. In general, the finding that several human pathogens contain polysaccharides as surface decoration has opened up a rapidly developing area of biomedical research.^[40] Investigations over the past decade have shown that carbohydrates

possess an enormous potential as lead structures for drug discovery, aiming at antagonizing the interaction of the physiological carbohydrate ligands with their receptor proteins.

To demonstrate the principal possibility of control over glycan display density by coassembly of S-layer *neoglycoproteins* with unmodified SgsE monomers, we show in this model a self-assembled A_SgsE_T12-O7 monolayer, in which only every second subunit is carrying the O7 modification, corresponding to a 1:1 mixture of A_SgsE_T12 and A_SgsE_T12-O7 (Figure 7B). In addition, coassembly of engineered SgsE-based proteins offers also an attractive option for producing multifunctional self-assembly nanolattices.

3. Conclusions

Carbohydrates are an important class of biomolecules that are key components for the mediation of crucial cellular phenomena, with density of glycan display and molecular scale precision of interaction being most relevant aspects. Despite the fact that carbohydrates are promising lead components for diverse in vivo as well as in vitro applications, they have escaped (nano)biotechnological applications so far due to the complexity of their biosynthesis. Using a carbohydrate/protein engineering approach in the experimental host *E. coli*, we aimed at integration of carbohydrates into a nanobiotechnological concept through combination with the established S-layer protein self-assembly system SgsE of *G. stearotherophilus* NRS 2004/3a. First S-layer *neoglycoproteins* were obtained after engineering of the natural protein *O*-glycosylation site threonine₆₂₀ to become a target for *N*-glycosylation through insertion of the *N*-glycosylation consensus sequence of *C. jejuni*, and the subsequent recombinant transfer of a heptasaccharide from *C. jejuni* and of the O7 polysaccharide from *E. coli* onto the S-layer protein by the action of the oligosaccharyltransferase PglB. Due to the intrinsic self-assembly property conferred to these *neoglycoproteins* by the S-layer matrix, these novel nanostructured composites allow nanometer-scale and periodic display of carbohydrates. It can be assumed that due to the recombinant production process and the utilization of the S-layer self-assembly matrix, this concept is superior to conventional, random chemical carbohydrate immobilization.

Tailor-made (“functional”) nanopatterned, self-assembled *neoglycoproteins* follow the current trend for exploiting means for organizing biological functions at the nanometer level aiming at the development of novel concepts for life and nonlife sciences.

4. Experimental Section

Bacterial strains, growth conditions, and plasmids

G. stearotherophilus NRS 2004/3a was obtained from the N. R. Smith Collection, US Department of Agriculture (Peoria, IL) and was grown on modified S-VIII medium (1% peptone, 0.5% yeast extract, 0.5% meat extract, 0.13% K₂HPO₄, 0.01% MgSO₄, 0.06% sucrose) at 55 °C. *E. coli* DH5α (Invitrogen, Lofer, Austria) was used for plasmid propagation. *E. coli* BL21 Star (DE3) and *E. coli* CLM24^[22] were used for (neoglyco)protein expression. *E. coli* strains were grown at 37 °C in Luria–Bertani (LB) broth supplemented with ampicillin (Amp^r; 50 µg mL⁻¹), kanamycin (Km^r; 50 µg mL⁻¹), chloramphenicol (Cm^r; 30 µg mL⁻¹), trimethoprim (Tnp^r; 50 µg mL⁻¹), or tetracycline

(Tet^r; 20 µg mL⁻¹), when appropriate. Plasmids pACYC184^{p_{gl}},^[21] pJHCV32,^[24] pEC(AcrA-per),^[28] and pMAF10^[22] have been described previously. All strains and plasmids are listed in Table 1. Oligonucleotide primers used for PCR amplification are summarized in Table 2.

General methods

Genomic DNA of *G. stearotheophilus* NRS 2004/3a was isolated with the Qiagen Genomic-tip 100 kit (Qiagen, Hilden, Germany) according to the manufacturer's instruction. Restriction enzymes and T4 DNA ligase were purchased from Invitrogen, and calf intestinal alkaline phosphatase was obtained from Roche (Vienna, Austria). The Qiagen MinElute gel extraction kit was used to purify DNA fragments from agarose gels and the Qiagen MinElute reaction cleanup kit was used to purify digested oligonucleotides and plasmids. Plasmid DNA from transformed cells was isolated with the Qiagen plasmid miniprep kit. Agarose gel electrophoresis was performed as described by Sambrook and Russell.^[41] PCR (Sprint Thermocycler, Hybaid, Ashford, UK) was performed using Super Yield Pwo polymerase (Roche). PCR conditions were optimized for each primer pair and amplification products were purified using the Qiagen MinElute PCR purification kit. *E. coli* transformation was done according to the manufacturer's instructions (Invitrogen). Transformants were screened by in situ-PCR reactions using RedTaq ReadyMix PCR Reaction Mix (Sigma-Aldrich, Vienna, Austria); recombinant clones were analyzed by restriction mapping and confirmed by sequencing (Agowa, Berlin, Germany).

SDS-PAGE and visualization of protein bands with Coomassie Blue R-250 staining reagent was carried out according to references.^[42] Western immunoblotting using anti-SgsE antibody and anti-pgl-glycan antibody was done as described elsewhere.^[43] Anti-SgsE antibody was available in our laboratory from previous studies and anti-pgl-glycan antibody was kindly provided by Prof. Markus Aebi (ETH Zurich, Switzerland). N-terminal sequencing of blotted protein bands that have been excised from the polyvinylidene fluoride (PVDF) membrane after visualization with Coomassie Blue R-250 staining reagent was performed by Prof. Herbert Lindner, University of Innsbruck, Austria.

Construction of plasmids for periplasmic targeting of SgsE

To identify a suitable system for the export of SgsE into the periplasm of *E. coli*, the S-layer protein was translationally fused to different signal peptides of Gram-negative bacteria directing the protein to the Sec-pathway (i.e., pectate lyase, PelB) signal peptide from *E. carotovora* and MBP including its signal peptide (i.e., periplasmic targeting peptide) from *E. coli*, and to the TAT-responsive ssTor signal peptide of the *E. coli* TMAO reductase. Depending on the target vector, the sequences encoding rSgsE₃₁₋₉₀₃, corresponding to the mature S-layer protein of *G. stearotheophilus* NRS 2004/3a devoid of its signal peptide (named A_SgsE) and an N-terminally truncated form thereof (rSgsE₃₃₁₋₉₀₃, named G_SgsE) were amplified by PCR with the primer pairs A_sgsE_for(NcoI)/sgsE_rev(XhoI) or G_sgsE_for(NcoI)/sgsE_rev(XhoI) (for the PelB construct), A_sgsE_for(EcoRI)/sgsE_rev(XbaI) or G_sgsE_for(EcoRI)/sgsE_rev(XbaI) (for the MBP construct), and A_sgsE_for(EcoRI)/sgsE_rev(XhoI) or G_sgsE_for(EcoRI)/sgsE_rev(XhoI) (for the ssTor construct). Amplification products were digested with NcoI/XhoI, EcoRI/XbaI, and EcoRI/

XhoI, respectively, and inserted into the dephosphorylated expression vectors pET22b, pMAL-p2x, or pET28a-ssTor. The latter vector was constructed by PCR amplification of the Tor signal peptide from chromosomal DNA of *E. coli* W3110 with the primer pair ssTorA(*NcoI*)_for/ssTorA(*EcoRI*)_rev, digestion with *NcoI/EcoRI* and ligation of the fragment into pET28a(+). In addition, SgsE possessing its endogenous signal peptide from *G. stearothermophilus* NRS 2004/3a was cloned. The sequence encoding rSgsE₁₋₉₀₃ of *G. stearothermophilus* NRS 2004/3a was amplified by PCR with the primer pairs sgsE-K1/sgsE_rev(*XhoI*) and ligated into vector pET28a(+) using the restriction sites *NcoI/XhoI*. The resulting plasmids for over-expression of periplasmic rSgsE protein with the different N-terminal signal peptides were named pET22b-A_SgsE and pET22b-G_SgsE (PelB constructs), pMAL-A_SgsE and pMAL-G_SgsE (MBP constructs), pET28a-ssTor_A_SgsE and pET28a-ssTor_G_SgsE (ssTor constructs), and pET28a-SgsE (endogenous signal peptide).

Engineering of N-glycosylation sites on SgsE

For the introduction of a distinctive protein N-glycosylation consensus sequence from the *C. jejuni* AcrA N-glycoprotein replacing the naturally *O*-glycosylation sites of the S-layer protein,^[19,20] a modified *sgsE* gene containing an artificial *KpnI* site at position nt 1830 (corresponding to amino acid position 611), coding for a protein with a PelB signal peptide and lacking the amino acid sequence between L₆₁₀ and E₈₀₄ was constructed. N- and C-terminal parts of *sgsE* were amplified separately by PCR using the primer pairs A_sgsE_for(*NcoI*)/sgsEteil_rev(*KpnI*) and sgsE-teil_for(*KpnI*)/sgsE_rev(*XhoI*), respectively. After digestion with *KpnI*, the two parts were ligated and the resulting product was amplified with the primer pair A_sgsE_for(*NcoI*)/sgsE_rev(*XhoI*,-Stop). The amplification product was digested with *NcoI* and *XhoI* and ligated with vector pET22b, yielding vector pET22b-SgsEinter that additionally equips SgsE with a C-terminal hexahistidine tag. Subsequently, either a five or a twelve amino acid sequence containing the N-glycosylation consensus sequence DXNZS (DFNRS and ASKDFNRSKALFS, respectively) was introduced into the *sgsE* sequence of *G. stearothermophilus* NRS 2004/3a by PCR amplification reactions using megaprimers. Different combinations of the megaprimers sgsE_T_N-glyk(DFNRS)_for(*KpnI*), sgsE_T_N-glyk(DFNRS)_rev(*KpnI*), sgsE_T_N-glyk(long)_for(*KpnI*), sgsE_T_N-glyk(long)_rev(*KpnI*), sgsE_S_for(*KpnI*), and sgsE_T_rev(*KpnI*) were used to yield six distinctive amplification products. These products were ligated into the vector pET22b-SgsEinter via the *KpnI* site. The resulting plasmids were named pET22b-A_SgsE_T5, pET22b-A_SgsE_T12, pET22b-A_SgsE_S5, pET22b-A_SgsE_S12, pET22b-A_SgsE_S,T5, and pET22b-A_SgsE_S,T12, respectively, and coded for modified SgsE proteins, in which either the natural *O*-glycosylation site Thr₆₂₀, Ser₇₉₄, or both sites were replaced by either DFNRS or ASKDFNRSKALFS. These six plasmids were used as templates to create G_SgsE forms with the same N-glycosylation sites by PCR with the primer pair G_sgsE_for(*NcoI*)/sgsE_rev(*XhoI*,-Stop) and ligation into pET22b. In another approach for N-glycosylation of SgsE the sequon DVN₈₉₃VS was created by site-directed mutagenesis of G₈₉₁D using A_sgsE_for(*NcoI*) and the modified reverse primer sgsE_rev(G₈₉₁D,*XhoI*,+Stop) for PCR. Cloning into vector pET22b was performed via *NcoI/XhoI* sites and the resulting plasmid was named pET22b-A_SgsE_G₈₉₁D. In addition, the DFNRS and ASKDFNRSKALFS N-glycosylation sequence were translationally fused

to the N-terminus of SgsE using the primer pairs G_sgsE_for(*Nco*I, long)/sgsE_rev(*Xho*I,-Stop) and G_sgsE_-for(*Nco*I, DFNRS)/sgsE_rev(*Xho*I,-Stop), and after ligation via *Nco*I/*Xho*I to pET22b and the resulting plasmids were named pET22b-DFNRS_G_SgsE and pET22b-long_G_SgsE. All the plasmids described so far are inducible by IPTG (Fermentas, St. Leon-Rot, Germany).

To obtain an expression system for the S-layer protein that was inducible with arabinose, the periplasmic SgsE forms PelB_A_SgsE_T12 and PelB_G_SgsE_T12 including the C-terminal hexahistidine tag were cloned from pET22b into vector pBAD24, using the primer pair PelB-SgsE_for(*Eco*RI)/SgsE-His_rev(*Xba*I) for PCR. The corresponding expression plasmids were named pBAD24-PelB_A_SgsE_T12 and pBAD24-PelB_G_SgsE_T12, respectively.

Production and purification of an SgsE neoglycoprotein with a *C. jejuni* glycan

E. coli BL21 Star (DE3) cells containing pACYCpgl carrying the *C. jejuni* glycosylation apparatus that is constitutively expressed were transformed with a pET22b-based plasmid encoding a periplasmic SgsE acceptor protein. As a control, the soluble periplasmic *C. jejuni* protein AcrA expressed from pEC(AcrA_per)^[28] was used in combination with pACYCpgl. pEC(AcrA_per) was kindly provided by Prof. Markus Aebi; pACYCpgl^[21] was obtained from Prof. Brendan Wren. Cells were grown in 400 mL cultures. At an OD₆₀₀ of ≈ 0.8 , IPTG was added to a final concentration of 1 mM and cultivation was continued for additional 20 h. The biomass was collected by centrifugation at 4500 g, at 4 °C, for 15 min.

For S-layer neoglycoprotein purification, 1 g of biomass (wet weight) was suspended in 10 mL of buffer A (10 mM MgCl₂ in 50 mM Tris/HCl, pH 7.5) containing 500 U of benzonase (Merck, Darmstadt, Germany) and 4 mg of lysozyme (Merck), and incubated for 20 min at 37 °C. The spheroplast fraction was removed by centrifugation at 1620 g and the supernatant, containing the periplasmic fraction, was concentrated using Microcon centrifugal filter units (cut-off 50 kDa; Millipore). To the retentate, urea (Fluka, Buchs, Switzerland) was added to a final concentration of 6 M to disintegrate the water-insoluble S-layer protein. This extract was loaded onto a Superdex 200 prep grade XK16 FPLC-column (1.6 × 60 cm²; GE Healthcare, Uppsala, Sweden). Elution was performed at a flow rate of 1 mL min⁻¹ using 6 M urea in buffer A. Fractions containing the desired proteins were pooled and dialyzed four times against distilled water, 3 L each (dialysis tubing cut-off 20 kDa; Millipore), to promote self-assembly.

Production of an SgsE neoglycoprotein with *E. coli* O7 glycosylation

E. coli CLM24 cells containing the plasmids pMAF10 (encoding the oligosaccharyl/protein transferase PglB from *C. jejuni*,^[22] pJHCV32 (encoding *E. coli* O7 antigen^[24]), and pBAD24-PelB_A_SgsE_T12 or pBAD24-PelB_G_SgsE_T12 (encoding periplasmic SgsE forms), were induced by the addition of arabinose to 0.02% (wt vol⁻¹). After induction at 37 °C for 5 h, arabinose was added again to ensure protein expression when the carbon source becomes limiting (as arabinose can be metabolized by the cells).

Electron microscopy

Periplasmic targeting and self-assembly of S-layer neoglycoproteins was investigated by transmission electron microscopy using a Philips model CM 12 electron microscope (Philips, Eindhoven, NL) operated at 80 kV using a 50 μm objective aperture.^[44,45]

Periplasmic targeting of SgsE was investigated after immuno-gold-labeling of ultrathin sectioned *E. coli* BL21 Star (DE3) cells expressing SgsE to which different targeting signals have been fused. Cell pellets from 5 mL expression cultures grown to the late logarithmic growth phase were fixed over night at 4 °C in 0.1 M sodium cacodylate buffer (pH 7.4) containing 2.5% glutaraldehyde, 2.5% paraformaldehyde, and 2.5 mM CaCl_2 . After a washing step with cacodylate buffer, dehydration was performed by incubating the cells in ethanol solutions of increasing concentrations (70, 80, 90, and 96%). Subsequently, LR-White resin mixed with ethanol (50%) was added to the cell pellet and the sample was agitated for 30 min. The sample was incubated with 200 μL of resin overnight at 4 °C. New resin was added to the cell pellet, and finally the pellet was transferred to a dry gelatine capsule, filled with LR-Whit resin, and polymerized at 55 °C for 24 h. Thin sections were cut with an Ultracut ultramicrotome (Reichert-Jung, Vienna, Austria) and placed on 200-mesh copper grids coated with pioloform and carbon films. For immunolabeling of thin sections the grids were incubated face-down on drops of anti-SgsE antibody raised in rabbit, diluted 1:1000 with a solution containing 15 mM Tris, 150 mM NaCl, 27 mM ethylenediaminetetraacetic acid (EDTA), 2.9% gelatine from cold-water fish skin (Sigma-Aldrich), 1% Top Block (Sigma-Aldrich), and 0.5% Brij 35 (Sigma-Aldrich) for 1 h at room temperature. After one washing step with buffer, the grids were incubated for 1 h with goat-anti-rabbit antibody labeled with 10 nm gold particles (Sigma-Aldrich). The grids were repeatedly rinsed with distilled water and the thin sections were finally stained with 1% uranyl acetate in Tris-buffered saline (TBS, pH 7.5) for 45 s. Self-assembly formation of S-layer neoglycoproteins was examined after negative staining of purified proteins with 1% uranyl acetate in TBS. For image processing, electron microscopy film negatives were digitized at 4000 dpi using a Super Coolscan 4000 ED film scanner (Nikon, Tokyo, Japan). The S-layer lattice parameters (base vector lengths and base angle) were determined in the Fourier spectra using software developed in house (Prof. Dietmar Pum, personal communication).

Modeling of nanopatterned neoglycoproteins

Image reconstruction of the SgsE(p2) matrix using Cinema 4D was based on a Fourier transformed transmission electron micrograph of a negatively stained preparation of the S-layer protein self-assembled in solution. Pdb data of the short-chain $\text{Glc}(\text{GalNAc})_5\text{Bac}$ glycan of *C. jejuni* and the polymerized, long chain *E. coli* O7 antigen $[\text{VioNAc}(\text{Rha})\text{ManGalGlcNAc}]_n$ were generated with Sweet at <http://www.glycosciences.de/>. The glycans were positioned onto the S-layer matrix in arbitrarily chosen positions and at arbitrarily chosen angles. For simplification of the model, and to avoid an overlapping of glycans, an O antigen of homogeneous length is shown and only one subunit of the p2 unit cell carries a glycan modification.

MS analysis

Purified A_SgsE_T12 glycosylated with *C. jejuni* heptasaccharide was separated by SDS–PAGE and stained using Novex colloidal blue reagent (Invitrogen), and the desired protein was excised, lyophilized, and digested with trypsin (E.C.3.4.21.4, Promega, Southhampton, UK) overnight, followed by extraction of peptides from the gel pieces. Nano-LC was performed on an Ultimate 3000 using a PepMap 100 75 $\mu\text{m} \times 15\text{ cm}$ fused silica C18 analytical column (LC Packings, Dionex, Sunnyvale, CA), coupled to a Probot for fraction collection and matrix addition, with 2,5-dihydrobenzoic acid as the matrix. A gradient of 2–60% acetonitrile in 0.1% trifluoroacetic acid was delivered over 60 min at a flow rate of 0.300 nL min^{−1}. MALDI TOF/TOF MS was performed using a Applied Biosystems 4800 mass spectrometer (Foster City, CA) in the positive reflectron mode with delayed extraction. MS precursor acquisition was followed by interpretation and data-dependent MS/MS acquisition with the CID on. Data interpretation was configured to select a maximum of 10 precursor ions per fraction with a minimum signal-to-noise ratio of 50.

Acknowledgements

The authors thank Prof. Markus Aebi from ETH Zurich and Prof. Brendan W. Wren from London School of Hygiene and Tropical Medicine, University of London, for providing the tools for working with the *C. jejuni* glycosylation system. *E. coli* CLM24 and pJHCV32 were kindly obtained from Dr. Cristina Marolda and Prof. Miguel A. Valvano (University of Western Ontario, London, ON, Canada). We also acknowledge Prof. Dietmar Pum and Markus Gossmann (Center for NanoBiotechnology, University of Natural Resources and Applied Life Sciences, Vienna) for image processing and help with the construction of the S-layer neoglycoprotein model. Financial support came from the Austrian Science Fund, projects P19047-B12 (to C. S.) and P18013-B10 (to P. M.), the Hochschuljubiläumsstiftung der Stadt Wien, project H-1809/2006 (to C. S.), the Biotechnology and Biological Sciences Research Council and the Wellcome Trust (to A. D.). A. D. was a Biotechnology and Biological Sciences Research Council Professorial Research Fellow.

References

- [1]. Sára M, Pum D, Schuster B, Sleytr UB. *J. Nanosci. Nanotechnol.* 2005; 5:1939. [PubMed: 16430129]
- [2]. Schäffer C, Novotny R, Küpcü S, Zayni S, Scheberl A, Friedmann J, Sleytr UB, Messner P. *Small.* 2007; 3:1549. [PubMed: 17786898]
- [3]. Sleytr UB, Messner P, Pum D, Sára M. *Angew. Chem. Int. Ed.* 1999; 38:1034.
- [4]. Sleytr UB, Huber C, Ilk N, Pum D, Schuster B, Egelseer EM. *EMS Microbiol. Lett.* 2007; 267:131.
- [5]. Sleytr UB, Egelseer EM, Ilk N, Pum D, Schuster B. *FEBS J.* 2007; 274:323–334. [PubMed: 17181542]
- [6]. Åvall-Jääskeläinen S, Palva A. *FEMS Microbiol. Rev.* 2005; 29:511. [PubMed: 15935509]
- [7]. Schäffer C, Messner P. *Glycobiology.* 2004; 14:31R.
- [8]. Sleytr, UB.; Sára, M.; Pum, D.; Schuster, B.; Messner, P.; Schäffer, C. *Biopolymers*. Steinbüchel, A.; Fahnestock, SR., editors. Vol. 7. Wiley-VCH; Weinheim, Germany: 2002. p. 285
- [9]. Schuster B, Pum D, Sára M, Sleytr UB. *Mini Rev. Med. Chem.* 2006; 6:909. [PubMed: 16918497]
- [10]. Sára, M.; Egelseer, EM.; Huber, C.; Ilk, N.; Pleschberger, M.; Pum, D.; Sleytr, UB. *Biological and Pharmaceutical Nanomaterials*. Kumar, C., editor. Wiley-VCH; Weinheim, Germany: 2006. p. 219
- [11]. Messner P, Steiner K, Zarschler K, Schäffer C. *Carbohydr. Res.* doi:10.1016/j.carres.2007.12.025.
- [12]. Apweiler R, Hermjakob H, Sharon N. *Biochim. Biophys. Acta.* 1999; 1473:4. [PubMed: 10580125]
- [13]. Spiro RG. *Glycobiology.* 2002; 12:43R.

- [14]. Ohtsubo K, Marth JD. Cell. 2006; 126:855. [PubMed: 16959566]
- [15]. Jones C. An. Acad. Bras. Cienc. 2005; 77:293. [PubMed: 15895165]
- [16]. Ratner DM, Seeberger PH. Curr. Pharm. Des. 2007; 13:173. [PubMed: 17269926]
- [17]. Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, Wilson IA. Science. 2006; 312:404. [PubMed: 16543414]
- [18]. Messner P, Pum D, Sleytr UB. J. Ultrastruct. Mol. Struct. Res. 1986; 97:73. [PubMed: 3453374]
- [19]. Schäffer C, Wugeditsch T, Kählig H, Scheberl A, Zayni S, Messner P. J. Biol. Chem. 2002; 277:6230. [PubMed: 11741945]
- [20]. Steiner K, Pohlentz G, Dreisewerd K, Berkenkamp S, Messner P, Peter-Katalini J, Schäffer C. J. Bacteriol. 2006; 188:7914. [PubMed: 16963578]
- [21]. Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW, Aebi M. Science. 2002; 298:1790. [PubMed: 12459590]
- [22]. Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA, Aebi M. Proc. Natl. Acad. Sci. USA. 2005; 102:3016. [PubMed: 15703289]
- [23]. Szymanski C, Yao R, Ewing CP, Trust TJ, Guerry P. Mol. Microbiol. 1999; 32:1022. [PubMed: 10361304]
- [24]. Marolda CL, Feldman MF, Valvano MA. Microbiology. 1999; 145:2485. [PubMed: 10517601]
- [25]. Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St, Michael F, Aberg E, Szymanski CM. J. Biol. Chem. 2002; 277:42530. [PubMed: 12186869]
- [26]. Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M, Wren BW. Mol. Microbiol. 2005; 55:1695. [PubMed: 15752194]
- [27]. Nita-Lazar M, Wacker M, Schegg B, Amber S, Aebi M. Glycobiology. 2005; 15:361. [PubMed: 15574802]
- [28]. Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M, Aebi M. EMBO J. 2006; 25:1957. [PubMed: 16619027]
- [29]. Kowarik M, Numao S, Feldman MF, Schulz BL, Callewaert N, Kiermaier E, Catrein I, Aebi M. Science. 2006; 314:1148. [PubMed: 17110579]
- [30]. Cristobal S, De Gier JW, Nielsen H, von Heijne G. EMBO J. 1999; 18:2982. [PubMed: 10357811]
- [31]. Novotny R, Scheberl A, Giry-Laterriere M, Messner P, Schäffer C. FEMS Microbiol. Lett. 2005; 242:27. [PubMed: 15675069]
- [32]. Riedmann E, Kyd JM, Smith AM, Gomez-Gallego S, Jalava K, Cripps AW, Lubitz W. FEMS Immunol. Med. Microbiol. 2003; 37:185. [PubMed: 12832124]
- [33]. Mukherjee K, Karlsson S, Burman LG, Akerlund T. Microbiology. 2002; 148:2245. [PubMed: 12101311]
- [34]. Houssin C, Nguyen DT, Leblon G, Bayan N. FEMS Microbiol. Lett. 2002; 217:71. [PubMed: 12445648]
- [35]. Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA, Whitfield C, Aebi M. Proc. Natl. Acad. Sci. USA. 2006; 103:7088. [PubMed: 16641107]
- [36]. Rangarajan ES, Bhatia S, Watson DC, Munger C, Cygler M, Matte A, Young NM. Protein Sci. 2007; 16:990. [PubMed: 17456748]
- [37]. Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT. Nucl. Acids Res. 2005; 33(Web Server issue):W36. [PubMed: 15980489]
- [38]. Tschiggerl H, Breitwieser A, de Roo G, Verwoerd T, Schäffer C, Sleytr UB. J. Biotechnol. 2008; 133:403. [PubMed: 18035441]
- [39]. Paton AW, Morona R, Paton JC. Nat. Methods. 2000; 6:265.
- [40]. Benz I, Schmidt MA. Mol. Microbiol. 2002; 45:267. [PubMed: 12123443]
- [41]. Sambrook, J.; Russell, DW. Molecular Cloning: A Laboratory Manual. 3rd Edn. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2001.
- [42]. Laemmli UK. Nature. 1970; 227:680. [PubMed: 5432063]

- [43]. Steiner K, Novotny R, Patel K, Vinogradov E, Whitfield C, Valvano MA, Messner P, Schäffer C. J. Bacteriol. 2007; 189:2590. [PubMed: 17237178]
- [44]. Messner P, Hollaus F, Sleytr UB. Int. J. Syst. Bacteriol. 1984; 34:202.
- [45]. Guzman LM, Belin D, Carson MJ, Beckwith J. J. Bacteriol. 1995; 177:4121. [PubMed: 7608087]

Figure 1.

Periplasmic targeting of the S-layer protein SgsE according to SDS-PAGE analysis (10% gel) after Coomassie Blue staining. Lane 1: Bench-mark ladder (Invitrogen); lane 2: SP-SgsE; lane 3: ssPelB-A_SgsE; lane 4: ssPelB-G_SgsE; lane 5: MBP-A_SgsE; lane 6: MBP-G_SgsE; lane 7: ssTor-A_SgsE; lane 8: ssTor-G_SgsE; lane 9: Benchmark ladder.

Figure 2.

Transmission electron micrograph of an ultrathin, cross-sectioned and immunogold-labeled cell of *E. coli* BL21 Star (DE3) expressing the S-layer protein SgsE, to which the PelB signal peptide has been fused (ssPelB-A_SgsE). The presence of the high number of gold particles in the periplasmic space indicates the efficiency of this export mechanism for the S-layer protein. P, periplasmic space; C, cytoplasm.

Figure 3.

Engineering of *N*-glycosylation sites on the S-layer protein SgsE as determined by recombinant glycosylation with the *C. jejuni* heptasaccharide (*pgl*) and Western immunoblot detection using a glycan-specific antibody. Lane 1: Precision Plus All Blue Protein Standard (BioRad); lane 2: A_SgsE_S,T5-pgl; lane 3: A_SgsE_S,T12-pgl; lane 4: A_SgsE_T5-pgl; lane 5: A_SgsE_T12-pgl; lane 6: A_SgsE_S5-pgl; lane 7: A_SgsE_S12-pgl; lane 8: glycosylated AcrA protein (control). Successful *neoglycoprotein* production is marked by a circle.

Figure 4.

Monitoring of the purification procedure of A_SgsE_T12-pgl carrying the *C. jejuni* heptasaccharide. SDS-PAGE after A) Coomassie staining and Western immunoblot analysis using antibodies raised against B) the S-layer protein SgsE or C) the *C. jejuni* heptasaccharide. Lane 1: A) Bench-mark ladder (Invitrogen); B), (C): Precision Plus All Blue Protein Standard; lane 2: intact cells of *E. coli* BL21 Star (DE3) expressing the S-layer *neoglycoprotein*; lane 3: periplasmic fraction; lane 4: dialysate; lane 5: soluble fraction. It is evident that the S-layer *neoglycoprotein* is enriched in the periplasm of *E. coli* BL21 Star (DE3) cells; according to Western blot analyses the glycosylation degree of this fraction reaches completeness.

Figure 5. MALDI TOF/TOF product ion spectrum of $[M + H]^+ m/z$ 1956.9 derived from the tryptic digest of the S-layer neoglycoprotein A_SgsE_T12 carrying the *C. jejuni* heptasaccharide. The fragment pattern for the glycan is shown inset.

Figure 6. Electron micrograph of self-assembled S-layer neoglycoprotein A_SgsE_T12 carrying the *C. jejuni* heptasaccharide after negative staining. The inset is the Fourier spectrum of the marked area. The spectrum shows more than one reciprocal lattice. Due to the low signal-to-noise ratio only one base vector pair is shown.

Figure 7.

Model of a self-assembled SgsE(p2)-neoglycoprotein monolayer displaying engineered glycans in a nanometer-scaled, periodic way. Image reconstruction using Cinema 4D is based on a negatively stained preparation of the S-layer protein self-assembled in solution and on the pdb data of the glycans generated with Sweet at <http://www.glycosciences.de/>. The short-chain glycan of *C. jejuni* (A) and the polymerized, long chain *E. coli* O7 antigen (B) that have been used throughout this study were modeled onto the patterning S-layer matrix.

Figure 8.

Glycosylation of the S-layer protein A_SgsE_T12 and G_SgsE_T12 with *E. coli* O7 polysaccharide. The S-layer neoglycoprotein was analyzed by SDS-PAGE (10% gel) and transfer of the samples to a PVDF membrane followed by immunodetection with an antibody recognizing SgsE. Lane 1: A_SgsE_T12-O7; lane 2: G_SgsE_T12-O7. The banding pattern of each neoglycoprotein (at ≈ 130 – 150 and ≈ 75 – 100 kDa, respectively) displays differences in the degree of polymerization of individual O7 repeating units. The prominent band at ≈ 97 and ≈ 64 kDa, respectively, corresponds to nonglycosylated protein; lane 3: Precision Plus All Blue Protein Standard.

Table 1

Strains and plasmids used in this study.

Strain or plasmid	Description	Source
<i>G. stearothermophilus</i> NRS 2004/3a	Wild-type covered by an S-layer	N. R. Smith collection
<i>E. coli</i> DH5a	F ⁻ ϕ 80lacZM15 (lacZYA-argF) U169recA1 endA1 hsdR17 (r_k^- , m_k^+) phoA supE44 thi-1 gyrA96 relA1 λ^-	Invitrogen
<i>E. coli</i> BL21 Star (DE3)	F ⁻ ompT hsdS _B (r_B^- m_B^-) gal dcm rne131 (DE3)	Invitrogen
<i>E. coli</i> W3110	rph-1 1N(rrnD-rrnE)	M. Valvano
<i>E. coli</i> CLM24	W3110, waalL	[22]
pET28a(+)	<i>E. coli</i> expression vector, IPTG inducible, Km ^r	Novagen
pET22b(+)	<i>E. coli</i> expression vector; Amp ^r , used for introducing an N-terminal PelB signal peptide	Novagen
pMAL-p2x	<i>E. coli</i> expression vector; Amp ^r , used for introducing N-terminally MBP including its signal peptide	NEB
pET28a-ssTor	pET28a(+) containing the signal peptide of TorA from <i>E. coli</i> W3110, Km ^r	This study
pET28a-SgsE	pET28a(+) containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a including its signal peptide, Km ^r	This study
pET22b-A_SgsE	pET22b(+) containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a, devoid of its signal peptide (aa 1–30), N-terminal PelB signal peptide; Amp ^r	This study
pET22b-G_SgsE	pET22b(+) containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a, devoid of the first 300 amino acids of the mature SgsE, N-terminal PelB signal peptide; Amp ^r	This study
pMAL-A_SgsE	pMAL-2px containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a, devoid of its signal peptide (aa 1–30), N-terminal MBP; Amp ^r	This study
pMAL-G_SgsE	pMAL-2px containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a, devoid of the first 300 amino acids of the mature SgsE, N-terminal MBP; Amp ^r	This study
pET28a-ssTor_A_SgsE	pET28a-ssTor containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a, devoid of its signal peptide (aa 1–30), N-terminal TorA signal peptide; Km ^r	This study
pET28a-ssTor_G_SgsE	pET28a-ssTor containing SgsE from <i>G. stearothermophilus</i> NRS 2004/3a, devoid of the first 300 amino acids of the mature SgsE, N-terminal TorA signal peptide; Km ^r	This study
pET22b-A_SgsE_G ₈₉₁ D	pET22b-A_SgsE with a mutation at position G ₈₉₁ ; Amp ^r	This study
pET22b-SgsE_inter	pET22b-A_SgsE, devoid of aa L ₆₁₀ to E ₈₀₄ , and an additional Kpn site at position 611; C-terminal His ₆ tag; Amp ^r	This study
pET22b-A_SgsE_S,T5	pET22b-A_SgsE, the sequence DFNRS replaces K ₆₁₈ TTSD ₆₂₂ and L ₇₉₂ TSAD ₇₉₆ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-A_SgsE_S,T12	pET22b-A_SgsE, the sequence ASKDFNRSKALFS replaces A ₆₁₅ DLKTTSDNFKLY ₆₂₇ and T ₇₈₉ ATLTSADVIRVD ₉₀₁ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-A_SgsE_T5	pET22b-A_SgsE, the sequence DFNRS replaces K ₆₁₈ TTSD ₆₂₂ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-A_SgsE_T12	pET22b-A_SgsE, the sequence ASKDFNRSKALFS replaces A ₆₁₅ DLKTTSDNFKLY ₆₂₇ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-A_SgsE_S5	pET22b-A_SgsE, the sequence DFNRS replaces L ₇₉₂ TSAD ₇₉₆ ; C-terminal hexa-His tag; Amp ^r	This study
pET22b-A_SgsE_S12	pET22b-A_SgsE, the sequence ASKDFNRSKALFS replaces T ₇₈₉ ATLTSADVIRVD ₉₀₁ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-G_SgsE_S,T5	pET22b-G_SgsE, the sequence DFNRS replaces K ₆₁₈ TTSD ₆₂₂ and L ₇₉₂ TSAD ₇₉₆ ; C-terminal His ₆ tag; Amp ^r	This study

Strain or plasmid	Description	Source
pET22b-G_SgsE_S,T12	pET22b-G_SgsE, the sequence ASKDFNRSKALFS replaces A ₆₁₅ DLKTTSDNFKLY ₆₂₇ and T ₇₈₉ ATLTSADVIRVD ₉₀₁ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-G_SgsE_T5	pET22b-G_SgsE, the sequence DFNRS replaces K ₆₁₈ TTSD ₆₂₂ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-G_SgsE_T12	pET22b-G_SgsE, the sequence ASKDFNRSKALFS replaces A ₆₁₅ DLKTTSDNFKLY ₆₂₇ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-G_SgsE_S5	pET22b-G_SgsE, the sequence DFNRS replaces L ₇₉₂ TSAD ₇₉₆ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-G_SgsE_S12	pET22b-G_SgsE, the sequence ASKDFNRSKALFS replaces T ₇₈₉ ATLTSADVIRVD ₉₀₁ ; C-terminal His ₆ tag; Amp ^r	This study
pET22b-DFNRS_G_SgsE	pET22b containing G_SgsE with DFNRS fused to the N-terminus, N-terminal PelB signal peptide; C-terminal His ₆ tag; Amp ^r	This study
pET22b-long_G_SgsE	pET22b containing G_SgsE with ASKDFNRSKALFS fused to the N-terminus, N-terminal PelB signal peptide; C-terminal His ₆ tag; Amp ^r	This study
pACYCpgl	pACYC184 containing the C. jejuni pgl cluster; Cm ^r	[21]
pEC(AcrA-per)	pEC415, expressing periplasmic soluble AcrA with PelB signal peptide and C-terminal His ₆ tag; Amp ^r	[28]
pBAD24	E. coli expression vector, arabinose inducible, Amp ^r	[45]
pBAD24-PelB_A_SgsE_T12	pBAD24, expressing periplasmic A_SgsE_T12 with PelB signal peptide and C-terminal His ₆ tag; Amp ^r	This study
pBAD24-PelB_G_SgsE_T12	pBAD24, expressing periplasmic G_SgsE_T12 with PelB signal peptide and C-terminal His ₆ tag; Amp ^r	This study
pMAF10	hemagglutinin-tagged PglB cloned in pMLBAD; Tmp ^r	[22]
pJHCV32	Encodes the O7 antigen cluster from E. coli; Tet ^r	[24]

Table 2

Oligonucleotide primers used for PCR amplification reactions.

Primer	Nucleotide sequence (5'→3') ^[a]
bsE-K1	TAGGCTCCATGG[a]ACAAAAAGAAAGC
sgsE_rev (<i>XhoI</i>)	AATCACTCGAGGGATACATGTGCGGTACAAGAAAGC
sgsE_rev(<i>XhoI</i> , -Stop)	AATCACTCGAGTTTTGCTACGTTTACAACAGTAGC
A_sgsE_for(<i>NcoI</i>)	AATCACCATGGCGGACGTGGCGACGGTCG
G_sgsE_for(<i>NcoI</i>)	AATCACCATGgTAAATTAGTGGTTGATGGCGC
A_sgsE_for(<i>EcoRI</i>)	AATCAGAATTCGCAACGGACGTGGCGACGGTCG
G_sgsE_for(<i>EcoRI</i>)	AATCAGAATTCATGTTAAATTAGTGGTTGATGGCGC
sgsE_rev (<i>XbaI</i>)	AATCATCTAGAGGATACATGTGCGGTACAAGAAAGC
ssTorA(<i>NcoI</i>)_for	AATCACCATGgACAATAACGATCTCTTTCAG
ssTorA(<i>EcoRI</i>)_rev	AATCAGAATTCGCTTGGCGCCGAGTC
sgsE_rev(G891D, <i>XhoI</i> , +Stop)	AATCACTCGAGTTATTTTGCTACGTTTACAACAGTAGCAGAAACATTGTTGtCAAC
sgsEteil_rev(<i>KpnI</i>)	AATCAGGTACCAAGTACATTGTTTTCGTCAATATACCTTATC
sgsEteil_for(<i>KpnI</i>)	AATCAGGTACCGAACCAGGTGGCAATGCGACG
sgsE_T_N-glyk(DFNRS)_for(<i>KpnI</i>)	AATCAGGTACCGTCGGCGCAGATTGGATTTTAATCGTTCTAATTCAAATTATATCTGCCGACAGACGAAAAATCG
sgsE_S_N-glyk(DFNRS)_rev(<i>KpnI</i>)	AATCAGGTACCGTCGACACGGATGACAGAACGATTAAATCGGTCGCTGTGAACGTCGGAGCCACGTTTTCTTTC
sgsE_T_N-glyk(lang)_for(<i>KpnI</i>)	AATCAGGTACCGTCGGCGCAAGCAAGGATTTTAATCGTTCTAAAGCTCTTTTAGTCTGCCGACAGACGAAAAATCGAAA TCGG TTGCTTTG
sgsE_S_N-glyk(lang)_rev(<i>KpnI</i>)	AATCAGGTACCACTAAAAAGAGCTTTAGAACGATTAAATCCTTGCTTGCGAACGTCGGAGCCACGTTTTCTTTCATGTATTC
sgsE_T_rev(<i>KpnI</i>)	AATCAGGTACCGTCGACACGGATGACGCTCTGCGCTGGTTAAG
sgsE_S_for(<i>KpnI</i>)	AATCAGGTACCGTCGGCGCAGATTGAAAACAACCTAGCGAC
G_sgsE_for (<i>NcoI</i> , long)	AATCACCAiGGCAAGCAAGGATTTTAATCGTTCTAAAGCTCTTTTAGTTTAAATTAGTGGTTGATGGCGC
G_sgsE_for (<i>NcoI</i> , DFNRS)	AATCACCAiGGATTTTAATCGTTCTTTAAATTAGTGGTTGATGGCGC
PelB-SgsE_for(<i>EcoRI</i>)	AATCAGAATTCATGAAATACCTGCTGCCGAC
SgsE-His_rev(<i>XbaI</i>)	AATCATCTAGATCAgTGGTGGTGGTGGTGG

^[a] Artificial restriction sites are underlined; lowercase letters indicate changes in the chromosomal DNA sequence. The artificial glycosylation sites are marked in bold. The triples corresponding to the initiation and termination codons in the primer sequence are shown in boxes.