S-layer gene sbsC of Bacillus stearothermophilus ATCC 12980: molecular characterization and heterologous expression in Escherichia coli

Marina Jarosh,1 Eva M. Egelseer,1 Diethard Mattanovich,2 Uwe B. Sleytr1 and Margit Sára1

Author for correspondence: Margit Sára. Tel.: +43 1 47 654 2208. Fax: +43 1 47 89112.
e-mail: sara@edv1.boku.ac.at

1 Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur, 1180 Vienna, Austria
2 Institut für Angewandte Mikrobiologie, Universität für Bodenkultur, 1190 Vienna, Austria

INTRODUCTION

Crystalline bacterial cell surface layers (S-layers) represent the outermost cell envelope component of many bacteria and archaea (for reviews see Sára & Sleytr, 1996; Sleytr et al., 1993, 1999). S-layers exhibit oblique, square or hexagonal lattice symmetry and they are composed of identical protein or glycoprotein subunits. Bacillus stearothermophilus represents a Gram-positive, strictly aerobic species of endospore-forming bacteria. Although S-layers have been identified as the outermost cell envelope component on more than 40 strains (Messner et al., 1984), so far only two S-layer genes have been cloned and sequenced. The S-layer gene sbsA (Kuen et al., 1994, 1995) encodes the S-layer protein of B. stearothermophilus wild-type strains, the nucleotide sequence encoding the S-layer protein SbsC of B. stearothermophilus ATCC 12980 was determined by PCR techniques. The entire sbsC sequence showed an ORF of 3297 bp predicted to encode a protein of 1099 aa with a theoretical molecular mass of 115409 Da and an isoelectric point of 5.73. Primer extension analysis suggested the existence of two promoter regions. Amino acid sequence comparison between SbsC and SbsA, a previously characterized S-layer protein of B. stearothermophilus PV72/p6 which assembles into a hexagonally ordered lattice, revealed an identical secretion signal peptide, 85% identity for the N-terminal regions (aa 31–270) which do not carry any S-layer homologous motifs, but only 21% identity for the rest of the sequences. Affinity studies demonstrated that the N-terminal part of SbsC is necessary for recognition of a secondary cell wall polymer. This was in accordance with results obtained in a previous study for SbsA, thus confirming a common functional principle for the N-terminal parts of both S-layer proteins. The sbsC coding region cloned into the pET3a vector without its own upstream region, the signal sequence and the 3’ transcriptional terminator led to stable expression in Escherichia coli.

Keywords: S-layer protein, Bacillus stearothermophilus, promoter, heterologous expression, secondary cell wall polymer
type. SbsA and SbsB show an overall identity of 25% and only SbsB carries three typical S-layer homologous (SLH) motifs (Engelhardt & Peters, 1998; Lupas, 1996; Lupas et al., 1994) at the N-terminal part. The mechanism leading to oxygen-induced S-layer variation has been described by Scholz et al. (1997).

By sequence comparison, SLH motifs (Lupas et al., 1996; Lupas, 1994) have been identified at the N-terminal part of many S-layer proteins and at the C-terminal end of cell-associated exoenzymes and other exoproteins (for recent review see Engelhardt & Peters, 1998). Due to their wide distribution among Gram-positive bacteria, SLH motifs were suggested to anchor the different types of cell-associated exoenzymes permanently or transiently to the cell surface. In contrast to the original assumption that peptidoglycan functions as binding site (Lupas et al., 1994), it is now evident that secondary cell wall polymers (SCWPs) serve as anchoring structures for the SLH motifs of S-layer proteins (Chauvaux et al., 1999; Ilk et al., 1999; Mesnage et al., 1999) and cell-associated exoenzymes (Brechtl & Bahl, 1999). On the basis of affinity studies, a binding region for a SCWP has also been identified on the N-terminal part of SbsA, the S-layer protein of *B. stearothermophilus* PV72/p6 (Engelser et al., 1998). The SCWPs of *B. stearothermophilus* PV72/p6, ATCC 12980 and NRS 2004/3a have an identical chemical composition, which led to the conclusion that this type of SCWP might be typical of *B. stearothermophilus* wild-type strains (Engelser et al., 1998). More recently, the complete structure was provided by NMR analysis, showing that this type of SCWP is composed of repeating units that contain glucose, N-acetylglucosamine and 2,3-diacetamido-2,3-dideoxy-D-mannuronic acid in a molar ratio of 1 to 1 to 2 (Schäffer et al., 1999).

The S-layer protein of *B. stearothermophilus* ATCC 12980 shows an apparent molecular mass of 122,000 Da on SDS gels, assembles into an oblique lattice type and possesses a binding site for a high molecular mass exoamylase (Engelser et al., 1996). N-terminal sequencing revealed that the N-terminal 14 aa of the whole S-layer protein are identical to those of SbsA. In previous studies, the S-layer protein could bind and recrystallize into the oblique lattice on native peptidoglycan-containing sacculi but not on those extracted with 48% hydrofluoric acid leading to pure peptidoglycan (Engelser et al., 1998). To investigate whether sequence identities and a common structure–function relationship exist in S-layer proteins of *B. stearothermophilus* wild-type strains, the gene *sbsC* encoding the S-layer protein of *B. stearothermophilus* ATCC 12980 was sequenced and cloned, and SbsC was compared with SbsA. The transcription start was determined and a high level expression system was established.

**METHODS**

**Growth of B. stearothermophilus ATCC 12980,** preparation of cell wall fragments, isolation of peptidoglycan-containing sacculi and production of S-layer self-assembly products. *B. stearothermophilus* ATCC 12980 was grown in batch culture on SVIII-medium under conditions described in previous studies (Engelser et al., 1996, 1998). Preparation of cell wall fragments, isolation of the S-layer protein, production of self-assembly products, N-terminal sequencing of the whole S-layer protein and purification of peptidoglycan-containing sacculi were performed as described by Egelseer et al. (1998).

**Other strains, plasmids, culture conditions and DNA manipulations.** *Escherichia coli* TG1 and XL-1 Blue were used for transformations with the plasmids pBluescript II SK(+) and pET3a, respectively. For expression, *E. coli* HMS174(DE3) was chosen as a host strain for derivatives of pET3a as described by Studier et al. (1990). *E. coli* was grown on Luria–Bertani medium (Gibco-BRL Life Technologies) or on modified M9ZB medium (Studier et al., 1990) at 37°C. For selection of transformants harbouring pBluescript II SK(+) or pET3a, ampicillin was added to the medium to a final concentration of 50 µg ml⁻¹. Chromosomal DNA of *B. stearothermophilus* ATCC 12980 was prepared by using Genomic Tips 100 (Qiagen) according to the manufacturer’s instructions. Digestion of DNA with restriction endonucleases, separation of DNA fragments by agarose gel electrophoresis, ligation of DNA fragments and transformation procedures were performed as described by Sambrook et al. (1989). DNA fragments were recovered from agarose gels by using the Qiagen II Gel Extraction kit (Qiagen).

**Isolation of the sbsC gene.** The six N-terminal amino acids of the mature S-layer protein of *B. stearothermophilus* ATCC 12980 were chosen for construction of the degenerate primer SLN 9/2 [5'-GC(T/C/A/G) AC(T/C/A/G) GA(T/C) GT(T/C/A/G) GC(T/C/A/G) AC-3'], which was used as a probe in hybridization studies to select the *sbsC* gene. For isolation of the *sbsC* gene by a PCR-based technique, chromosomal DNA of *B. stearothermophilus* ATCC 12980 was digested with different restriction endonucleases (Fig. 1) and separated on 1% agarose gels. A size range containing the DNA fragment carrying the N-terminus was identified by hybridization with the primer SLN 9/2, excised, recovered and ligated into a dephosphorylated pBluescript II SK(+) vector. PCR products were obtained directly from the ligation mixture by using an insert-specific primer in combination with a vector-specific primer. Derived from the obtained sequence of the PCR product, a new hybridization probe was constructed and used to isolate an adjacent *sbsC*-carrying DNA fragment. By applying this method, four PCR-generated fragments covering the entire *sbsC* sequence, including the putative promoter region and the transcription terminator, were obtained (Fig. 1). PCR amplification of ligated *sbsC* fragments was performed in a 50 µl reaction volume containing 240 µM dATP, dCTP, dGTP and dTTP, 240 nM primer or 2 µM degenerate primer, 1.25 mM MgCl₂, 1 U Taq DNA polymerase (Gibco-BRL Life Technologies) and 10 ng of the inactivated ligation as the template in 1x Taq reaction buffer (Gibco-BRL Life Technologies). Thirty cycles of amplification were performed in a thermocycler (Perkin Elmer Gene Amp 2000 or Hybaid Touch Down Control). Each cycle consisted of a 30 s denaturation step at 95°C, a 45 s annealing step performed by the touchdown method described by Don et al. (1991) with annealing temperatures depending on the calculated *Tₘ* of the oligonucleotides used, and extension times of 60 s per 1000 bp at 72°C.

**DNA sequencing.** PCR product sequencing of the four overlapping *sbsC* gene fragments (Fig. 1) was performed twice, once from each strand, by using the dideoxy chain-termination method of Sanger et al. (1977) with a Perkin Elmer
Isolation of RNA. Total RNA was isolated from exponentially growing cells of *B. stearothermophilus* ATCC 12980 which were harvested at an OD$_{600}$ of 0.78. For enzymic lysis, the cell pellet obtained by centrifugation of 4.5 ml of the bacterial suspension at 5000 g for 5 min at 4°C was resuspended in 0.5 ml 10 mM Tris/HCl buffer, pH 8.0, containing 1 mM EDTA and 5 mg lysozyme (Sigma) ml$^{-1}$. Samples were incubated for 15 min at room temperature. Isolation of total RNA was performed by using the RNeasy Midi kit (Qiagen) according to the manufacturer’s instructions. The RNA concentration was determined spectrophotometrically at 260 nm and samples were stored at $-20^\circ$C.

DNA and RNA hybridization. Northern and Southern blotting were performed as described by Sambrook *et al.* (1989). For hybridization of Southern blots, oligonucleotides were 3' tailed with digoxigenin-11-dUTP (DIG-dUTP; Boehringer Mannheim) and DNA probes were labelled with DIG-dUTP by incorporation during PCR. Hybridization was performed according to the manufacturer’s recommendations. Hybrids were detected using the DIG Luminescent Detection kit (Boehringer Mannheim). For Northern blotting, total RNA from *B. stearothermophilus* ATCC 12980 (5 µg per well) was fractionated by electrophoresis in 1% agarose/formaldehyde gels. Hybridization of Northern blots was carried out with a 3.3 kb DNA fragment comprising the coding region of the *sbsC* gene randomly labelled with DIG-dUTP. The size of the *sbsC* mRNA was estimated from its mobility relative to those of the RNAs in the DIG-dUTP-labelled RNA Molecular Weight Marker II (size range 1.5–6.9 kb; Boehringer Mannheim).

**Primer extension analysis.** Primer extension was carried out with a 5’FAM [5(6)-carboxyfluorescein]-labelled oligonucleotide primer designated 9/2 24 (5’-TTG TCC ATA AAG CCT AAA ATC CCC-3’). Total RNA (5 µg) from exponentially growing cells and the oligonucleotide primer 9/2 24 (10 pmol) were mixed in 1× hybridization buffer (150 mM KCl, 1 mM EDTA, 10 mM Tris/HCl buffer, pH 8.3), incubated at 65°C for 3 h and cooled down to room temperature. The nucleic acids in the annealed hybridization mixture were precipitated with 96% ethanol and resuspended in the primer extension mixture (total volume 20 µl) containing 1 mM dATP, dCTP, dGTP and dTTP, 20 U ribonuclease inhibitor (MBI Fermentas) and 20 U M-MuLV reverse transcriptase (MBI Fermentas) in 1× reaction buffer supplied by the manufacturer. The primer extension reaction was carried out at 37°C for 1.5 h. The reaction was stopped by the addition of 1 µl 500 mM EDTA and 10 ng RNase A (Boehringer Mannheim) as described by Sambrook *et al.* (1989). The reaction mixture was extracted with phenol/chloroform/isoamyl alcohol [25:24:1 (by vol.)] and precipitated with 96% ethanol. The length of the primer extended product was determined by electrophoresis on polyacrylamide/urea gels (IBL).

**Cloning and expression of the *sbsC* gene.** For cloning of the *sbsC* gene, the oligonucleotide primers *sbsC* III (5’-CGG GGA TTC CAT ATG GCA ACG GTG GCC AC-3’) and *sbsC* IV (5’-CGC GGA TTC TTA CGA TGC TGA TTT TGT ACC AAT TTG-3’), which introduced the restriction sites NdeI (at the 5’ end of the coding sequence) and BamHI (at the 3’ end), respectively, were used to amplify the *sbsC* gene from total *B. stearothermophilus* ATCC 12980 DNA preparations. PCR was performed as described above except that 100 ng *B. stearothermophilus* ATCC 12980 chromosomal DNA and 2.5 U Pwo polymerase (Boehringer Mannheim) were used according to the manufacturer’s instructions. The PCR product encoding the mature SbsC protein terminated by its own stop codon was digested with the restriction endonucleases NdeI and BamHI. The fragment was subsequently gel-purified and ligated into the corresponding restriction sites of the pET3a vector. To investigate the stability of the *sbsC* gene in pET3a, the resulting construct was transformed into *E. coli* XL-1 Blue. For expression, the plasmid pET3a carrying the *sbsC* gene was transformed into *E. coli* HMS174(DE3). The plasmid stability test was performed as described by Studier *et al.* (1990). Expression of the *sbsC* gene was induced by the addition of IPTG (Gerbii) to a final concentration of 0.4 mM at an OD$_{600}$ of 0.9. Samples (1.5 ml) were taken 1, 3 and 7 h after induction of *sbsC* gene expression. Preparation of samples and SDS-PAGE were carried out as described by Laemmli (1970). Immunoblotting with polyclonal rabbit antiserum raised against the S-layer protein of *Bacillus stearothermophilus* ATCC 12980 was performed as described by Egelseer *et al.* (1996).

**Preparation of proteolytic cleavage fragments of the S-layer protein from *B. stearothermophilus* ATCC 12980 and affinity studies.** For obtaining proteolytic cleavage fragments, 1 mg lyophilized S-layer self-assembly products was dissolved per ml 2 M guanidine. HCl in 50 mM Tris/HCl buffer (pH 7.8) and 40 µg endoproteinase Glu-C (Sigma P 6181) was added.
performed as described in a previous paper (Egelseer et al. 1996).

RESULTS

Characterization of the sbsC gene

The entire sbsC sequence showed one ORF of 3297 bp predicted to encode a protein of 1099 aa (AF055578) with a calculated theoretical molecular mass of 115409 Da and an isoelectric point (pI) of 5.73. The ORF started with ATG, preceded by a typical Shine-Dalgarno sequence with an appropriate distance of 11 bp to the start codon. Thirty-three nucleotides downstream of the stop codon TAA, a putative rho-independent transcriptional termination signal was identified. The first 90 bp of the structural gene encoded a typical Gram-positive type signal sequence of 30 aa residues.

Northern blotting and primer extension analysis

The size of the mRNA transcribed from the sbsC gene was analysed by Northern blotting using a DNA probe comprising the coding region of the sbsC gene. The probe detected a transcript of approximately 3.2 kb, which was in good agreement with the size of the sbsC gene. Primer extension mapping of the sbsC mRNA revealed two 5' ends with a difference in size of 35 nt. This indicated the existence of two different sbsC mRNA species, designated mRNA1 and mRNA2 (Fig. 3). As estimated from the peak areas of the primer extended products, mRNA2 was predominant (2.4 times higher than mRNA1) in exponentially growing cells. The first transcription start site was found to be located 224 nt upstream from the translation start, whereas the second one could be identified 189 nt upstream from the start codon. The two transcription start points were separated by a short AT-rich region of 34 nt. Examination of the DNA sequence upstream of the two transcription start sites revealed two putative promoter regions, designated P1 and P2. In the distal located promoter P1, the suggested 35 and -10 regions are TTTAAT and TGATT, whereas the putative -35 and -10 regions of the more downstream located promoter P2 are TAGAAT and TATTAT, respectively (Fig. 3).

Cloning and expression of the sbsC gene in E. coli HMS174(DE3)

The sequence encoding the mature SbsC was amplified with PCR by using the primers sbsC III and sbsC IV, cloned into the pET3a vector and the plasmid was established in E. coli HMS174(DE3). Expression of the sbsC gene was induced by the addition of IPTG and samples were taken at distinct points of time. As shown by SDS-PAGE (Fig. 4), a protein band with an apparent molecular mass of 122000 Da had accumulated in the host cells 1 h after induction of sbsC expression (lane 2) and the intensity of this protein band increased with increasing time of induction (lanes 3 and 4). No protein band with the respective molecular mass was detected in uninduced E. coli cells (lane 1). As shown by immunoblotting, SDS extracts of self-assembly products formed...
by the S-layer protein from *B. stearothermophilus* ATCC 12980 gave a strong cross-reaction with the polyclonal rabbit antiserum raised against this S-layer protein (Fig. 4, lane 5). A comparably strong cross-reaction was observed between the polyclonal rabbit antiserum and the high molecular mass protein band formed by the S-layer protein from *B. stearothermophilus* ATCC 12980 diluted 1:30000 was used. Lane 5, S-layer self-assembly products of a single major proteolytic cleavage fragment showing an apparent molecular mass of 122 kDa (Fig. 4, lane 5).

Analysis of the S-layer protein SbsC and comparison with SbsA

SbsC, which was identified as the S-layer protein of *B. stearothermophilus* ATCC 12980, had an amino acid composition typical of this class of secreted proteins (Seyt & Messner, 1983). This includes a high content of acidic and hydrophobic amino acids, lysine as the predominant basic amino acid, a low content of histidine, arginine and methionine and no cysteine. According to the sequence data, the mature SbsC protein has a theoretical molecular mass of 112471 Da and a calculated pl of 5.40.

Analysis of the amino acid distribution within SbsC revealed that lysine, arginine and tyrosine are accumulated at the N-terminal part (aa 31–270). The lysine content of the N-terminal region is 117 mol% and decreases to 88 mol% for the rest of the sequence (aa 271–1099). Furthermore, nine arginine residues were found to be located between aa 120 and 210, which corresponds to 75% of the total number of arginine residues in this S-layer protein. Due to the high density of positively charged amino acids, the N-terminal part reveals a calculated pl of 9.13, which is in clear contrast to the pl of 4.88 for the rest of the sequence. In addition to lysine and arginine, tyrosine is concentrated at the N-terminal part between aa 31 and 180, which represents 10 mol% in this region. For comparison, the tyrosine content of the rest of the sequence (aa 181–1099) is only 24 mol%. Secondary structure prediction according to Rost & Sander (1993) revealed that approximately 70% of the N-terminal 270 aa are organized as α-helices. The middle and C-terminal part of SbsC consists mainly of loops and β-sheets and only five short, but strongly predicted α-helices could be distinguished.

For identifying regions with amino acid sequence similarities within SbsC, the BLAST program (Altschul et al., 1997) was used. An identity of 23% was determined for the segments between aa 447 and 700 (R_m) and aa 844 and 1089 (R_c), being either located in the middle (aa 271–706) or on the C-terminal part (aa 707–1099) of SbsC. Within the middle part, the segment between aa 271 and 600 (R_m) showed 23% identity to that between aa 348 and 706 (R_m). A similar observation was made for the C-terminal part, where the segment between aa 831 and 978 (R_c) had 28% identity to that between aa 931 and 1089 (R_c).

Preparation of proteolytic cleavage fragments of the S-layer protein from *B. stearothermophilus* ATCC 12980 and affinity studies

As previously described, a single major proteolytic cleavage fragment showing an apparent molecular mass of 100000 Da on SDS gels was formed by degradation of the whole S-layer protein from *B. stearothermophilus*...
Fig. 5. Schematic drawing showing comparison of the S-layer proteins SbsA and SbsC from two B. stearothermophilus wild-type strains (PV72/p6 and ATCC 12980) and SbsB from the oxygen-induced strain variant B. stearothermophilus PV72/p2. The repeated amino acid sequences in the middle (aa 271–706) and C-terminal part (aa 707–1099) of SbsC are shown. Dotted lines indicate those segments on the different S-layer proteins which show different length but begin and stop with common sequences.

Fig. 6. SDS-PAGE pattern of the S-layer protein isolated from B. stearothermophilus ATCC 12980 after proteolytic degradation with endoproteinase Glu-C, removing guanidine HCl by dialysis and centrifugation. Lanes: 1, clear supernatant before incubation with native peptidoglycan-containing sacculi; 2, the 100000 Da cleavage fragment missing the N-terminal 227 aa of the mature S-layer protein did not recognize native peptidoglycan-containing sacculi as binding site and remained in the soluble fraction; 3, the whole S-layer protein could bind to native peptidoglycan-containing sacculi and was enriched in the insoluble fraction. The molecular mass is indicated on the right.

ATCC 12980 with endoproteinase Glu-C in 2 M guanidine HCl. Negative staining of the suspension obtained after dialysis of the digested S-layer protein revealed that the 100000 Da cleavage fragment had retained the ability to self-assemble (Egelseer et al., 1998). After centrifugation of the suspension and investigating the pellet and the clear supernatant (Fig. 6, lane 1) by SDS-PAGE, it became evident that both fractions had an identical composition, containing comparable amounts of the 100000 Da cleavage fragment and only traces of the whole S-layer protein. For affinity studies, the soluble fraction (Fig. 6, lane 1) was incubated with native peptidoglycan-containing sacculi. As shown in Fig. 6, the major portion (>90%) of the 100000 Da cleavage fragment remained unbound in the soluble fraction (lane 2), whereas the whole S-layer protein with an apparent molecular mass of 122000 Da was enriched in the pellet (lane 3). The 100000 Da cleavage fragment which did not recognize native peptidoglycan-containing sacculi as binding site had the N-terminus AALTPK. Comparison of this N-terminus with the SbsC sequence revealed that the first amino acid corresponded to alanine in position 258 (228 of the mature S-layer protein). Since neither the whole S-layer protein nor the 100000 Da cleavage fragment could bind to hydrofluoric-acid-extracted sacculi (not shown), it became evident that the N-terminal part recognizes the SCWP but not the peptidoglycan as binding site. These findings were in good agreement with previous results obtained for SbsA (Egelseer et al., 1998), confirming a common functional principle for the N-terminal parts of both S-layer proteins.

DISCUSSION

Up to now, the nucleotide sequences of the genes encoding the S-layer proteins SbsA from B. stearothermophilus PV72/p6, a wild-type strain, and SbsB from B. stearothermophilus PV72/p2, its oxygen-induced variant strain, have been established (Kuen et al., 1994, 1997). In the present study, the sbsC gene encoding the S-layer protein of B. stearothermophilus ATCC 12980, a further wild-type strain, has been sequenced and cloned and the 5' upstream region has been characterized. In contrast to the results obtained in previous studies for sbsA and sbsB (Kuen et al., 1995, 1997), stable cloning of the sbsC gene in E. coli was possible without its signal sequence and its own transcriptional terminator.

By primer extension analysis, two putative promoter regions designated P1 and P2 could be identified within
the 700 bp region upstream of the sbsC ORF of B. stearothermophilus ATCC 12980. In the more downstream located promoter P2, the putative −10 and −35 regions are separated by an optimal spacing of 17 nt, whereas the −10 and −35 regions of the distal promoter P1 show a spacing of 20 nt. Primer extension analysis with RNA from B. stearothermophilus ATCC 12980 indicated that transcripts from the more downstream located promoter (P2) in front of the sbsC gene were predominant, which might indicate more efficient usage of this promoter in exponentially growing cells. Actually, multiple and tandemly arranged promoters have been detected in front of several S-layer genes (Adachi et al., 1989; Boot et al., 1996; Kahala et al., 1997). For Lactobacillus brevis it was reported that the S-layer gene is expressed by two adjacent promoters, P1 and P2 (Kahala et al., 1997). Interestingly, the level of P2-derived transcripts was 10 times higher than that of P1-derived transcripts throughout all growth stages. On the other hand, in Lactobacillus acidophilus (Boot et al., 1996), only one of the two putative promoter regions was active in directing mRNA synthesis. The example of Bacillus brevis 47 demonstrated that at least one of the five tandemly arranged promoters of the cell wall protein gene operon is used during the exponential growth phase (Adachi et al., 1989). The existence of a typical rho-factor-independent transcriptional terminator at the downstream non-coding sequence of the sbsC gene indicated the monocistronic nature of this S-layer gene, which could also be confirmed by Northern blotting. Comparison of the nucleic acid sequences revealed that the transcription terminator of the sbsC gene has 100% identity to the transcription terminator of the sbsA and sbsB genes (Kuen et al., 1994, 1997). However, the 3′ non-translated sequence of the sbsC gene did not show any identity to the 3′ upstream region of all known S-layer genes.

In contrast to whole cells of B. stearothermophilus ATCC 12980, which are completely covered with an oblique lattice composed of the S-layer protein SbsC, the S-layer protein SbsA of B. stearothermophilus PV72/p6 assembles into a hexagonally ordered lattice (Sára et al., 1996). In both S-layer proteins, the N-terminal part (aa 31–270) comprises a binding region for an identical type of SCWP. According to a common functional principle, the N-terminal parts of both S-layer proteins revealed high (85%) identity. The accumulation of positively charged amino acids indicated that direct electrostatic interactions between the N-terminal part and the SCWP possessing a negative net charge play an important role in anchoring the S-layer subunits to the rigid cell wall layer. According to these findings, neither mono nor bivalent cations were required for binding of the S-layer subunits to native peptidoglycan-containing sacculi (Egelseer et al., 1998) and even the presence of EDTA did not influence the recognition and binding process. The accumulation of tyrosine, as was observed for the N-terminal parts of both S-layer proteins, is typical of carbohydrate-binding proteins such as lectins (Transue et al., 1997; Weis, 1997). In such proteins, tyrosine interacts with N-acetylated sugars via hydrogen bonds. In contrast to the high identity observed for the N-terminal parts, the middle and C-terminal parts of both S-layers showed an identity of only 21%. The formation of self-assembly products by the 100000 Da proteolytic cleavage fragment confirmed that the segment comprising the N-terminal 227 aa of the mature SbsC is not involved in the self-assembly process (Egelseer et al., 1998).

With the exception of the N-terminal part of SbsC comprising the binding region for the SCWP, the remaining part of the sequence showed identities of 20–25% to the S-layer proteins SbsB of B. stearothermophilus PV72/p2 (Kuen et al., 1997), Sap of Bacillus anthracis (Etienne-Toumelin et al., 1995), CTC of Bacillus thuringiensis (M. Sun & Z. Yu, GenBank accession no. AJ012290) and OlpA of Bacillus licheniformis (Zhu et al., 1996), which most probably assemble into oblique lattice types. In contrast to SbsC, each of these S-layer proteins possesses three typical SLH motifs at the N-terminal part (Mesnage et al., 1997). Comparison of SbsC with SbsB revealed that the regions of
identity start with aa 262 on SbsC and with aa 204 on SbsB (Fig. 5). Both S-layer proteins possess a conserved endoproteinase Glu-C cleavage site, which is located at aa 400 on SbsC and at aa 330 on SbsB. As indicated in Fig. 5, SbsB possesses a longer middle part with 68 additional amino acids, leading to a nearly identical length for both S-layer proteins at a related sequence which begins with aa 714 on the SbsB (SKSTANGD) and with aa 712 on the SbsB (SSSTVNVQ). Sequence comparison between SbsC, Sap and CTC revealed the existence of common sequences of 4–5 aa (KVES, TYTV, TVKD, ESLN, DQYG, VDANG) over the whole middle and the corresponding C-terminal parts (Fig. 7). However, SbsC possesses additional C-terminal amino acids which approximately represent the complete repeat R₄C. Derived from sequence comparison it could be speculated that the additional C-terminal amino acids on SbsC are not required for the formation of the oblique lattice structure. Recently, experimental evidence was provided for this consideration since a C-terminal truncated form of SbsC was still capable of self-assembling (M. Jarosch, unpublished observation). Considering the longer N-terminal part of SbsC in comparison to that of Sap, CTC and OlpA (Fig. 7), the truncated form approximately corresponds to the lengths of the whole other S-layer proteins.

ACKNOWLEDGEMENTS

This work was supported by the Austrian Science Foundation, project P12938, and by the Ministry of Science and Transports. We thank Christoph Hotzy and Aida Medovic for technical assistance.

REFERENCES


---

Received 23 August 1999; revised 22 October 1999; accepted 29 October 1999.