ISBst12, a novel type of insertion-sequence element causing loss of S-layer-gene expression in Bacillus stearothermophilus ATCC 12980

Eva M. Egelseer, Rughia Idris, Marina Jarosch, Thomas Danhorn, Uwe B. Sleytr and Margit Sára

Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur, A-1180 Vienna, Austria

The cell surface of the surface layer (S-layer)-carrying strain of Bacillus stearothermophilus ATCC 12980 is completely covered with an oblique lattice composed of the S-layer protein SbsC. In the S-layer-deficient strain, the S-layer gene sbsC was still present but was interrupted by a novel type of insertion sequence (IS) element designated ISBst12. The insertion site was found to be located within the coding region of the sbsC gene, 199 bp downstream from the translation start of SbsC. ISBst12 is 1612 bp long, bounded by 16 bp imperfect inverted repeats and flanked by a directly repeated 8 bp target sequence. ISBst12 contains an ORF of 1446 bp and is predicted to encode a putative transposase of 482 aa with a calculated theoretical molecular mass of 55562 Da and an isoelectric point of 9.13. The putative transposase does not exhibit a typical DDE motif but displays a His-Arg-Tyr triad characteristic of the active site of integrases from the bacteriophage λ Int family. Furthermore, two overlapping leucine-zipper motifs were identified at the N-terminal part of the putative transposase. As revealed by Southern blotting, ISBst12 was present in multiple copies in the S-layer-deficient strain as well as in the S-layer-carrying strain. Northern blotting indicated that S-layer gene expression is already inhibited at the transcriptional level, since no sbsC-specific transcript could be identified in the S-layer-deficient strain. By using PCR, ISBst12 was also detected in B. stearothermophilus PV72/p6, in its oxygen-induced strain variant PV72/p2 and in the S-layer-deficient strain PV72/T5.

Keywords: insertion-sequence element, S-layer protein, Bacillus stearothermophilus

INTRODUCTION

Crystalline bacterial cell surface layers (S-layers) represent the outermost cell-envelope component of many bacteria and archaea (for reviews see Sára & Sleytr, 2000; Sleytr & Sára, 1997; Sleytr et al., 1993, 1999). S-layers exhibit oblique (p1, p2), square (p4) or hexagonal (p3, p6) lattice symmetry and they are composed of identical protein or glycoprotein subunits. In bacteria, the S-layer subunits are linked to each other and to the underlying cell-envelope layer by non-covalent interactions.

Bacillus stearothermophilus is a Gram-positive, strictly aerobic species of endospore-forming bacteria that can produce large amounts of exoenzymes such as proteases and amylases (Archibald, 1989; Priest, 1981). In previous reports, the putative role of the S-layers from B. stearothermophilus DSM 2358 and ATCC 12980 (DSM 22) as an adhesion site for a high-molecular-mass amylase was confirmed (Egelseer et al., 1995, 1996). To gather information as to whether S-layer lattices of Gram-positive bacteria delineate a kind of periplasmic space and may control the speed of exoenzyme release,
comparative studies were carried out with the S-layer-carrying (S⁺) strain of *B. stearothermophilus* ATCC 12980 and its S-layer-deficient (S⁻) derivative (Egelseer et al., 1996). The S⁻ strain was isolated from SVIII agar slants of the S⁺ strain which were stored for 2 years at 4 °C (Eder, 1983). In freeze-etched preparations, a highly ordered S-layer lattice with oblique symmetry could only be observed on whole cells from the S⁺ strain, whereas the S⁻ strain revealed an amorphous cell surface, typical of strains lacking an S-layer (Egelseer et al., 1996).

Recently, the nucleotide sequence encoding the S-layer protein SbsC from the S⁺ strain of *B. stearothermophilus* ATCC 12980 was determined by PCR techniques (Jarosch et al., 2000). The entire sbsC sequence shows an ORF of 3297 bp predicted to encode a protein of 1099 aa (AF055378) with a theoretical molecular mass of 115409 Da and a pI of 5.73. Primer-extension analysis indicated the existence of two promoter regions which are most probably active in different growth stages (Jarosch et al., 2000).

The loss in S-layer protein synthesis has been described for several organisms during prolonged cultivation under optimal laboratory conditions, indicating that these crystalline arrays provide a selective advantage in competitive natural habitats (Messner & Sleytr, 1992). On the other hand, environmental stress factors such as temperature upshift also caused a loss in S-layer-protein synthesis (Sleytr et al., 1982). The first detailed studies to understand the molecular mechanism leading to the loss in S-layer protein synthesis were carried out with *Aeromonas salmonicida*, a fish-pathogenic organism. It was demonstrated that the insertion of insertion sequence (IS) elements at different positions of the vapA gene encoding the S-layer (A-layer) protein and in its upstream region had occurred (Gustafson et al., 1994).

In general, IS elements are small translocating segments of DNA that are located on the host chromosome and on plasmids with the potential to move within and among bacterial genomes. While these elements have a variety of architectures, there are two essential features common to all transposable elements. Firstly, they are delineated by end sequences that are required in cis for the transposition reaction. Secondly, they encode the functions that facilitate and control their movement. Gene inactivation or macrogenomic rearrangements such as adjacent deletions, inversions and cointegrate formations are often discovered as a result of ISs (Gasson & Fitzgerald, 1993). Today over 300 bacterial ISs isolated from 73 genera representing 159 species of bacteria and archaea have been characterized at the nucleotide-sequence level (Mahillon & Chandler, 1998). It is now evident that transposition does not occur randomly, since in each case studied, some target preference for integration has been observed (Craig, 1997). The integration of an IS element can either repress or activate genes located downstream.

In the present study, we report the sequence determination and the characterization of a novel type of bacterial IS element, designated ISBst12. This IS element was found to be responsible for the loss of S-layer protein synthesis in the S⁻ strain of *B. stearothermophilus* ATCC 12980.

**Methods**

Bacterial growth conditions of the S⁺ and the S⁻ strains of *B. stearothermophilus* ATCC 12980 as well as of *B. stearothermophilus* PV72/p6, PV72/p2 and PV72/T5. The S⁺ and S⁻ strains of *B. stearothermophilus* ATCC 12980 as well as *B. stearothermophilus* strains PV72/p6, PV72/p2 and PV72/T5 (Sára et al., 1996) were grown in 300 ml shaking flasks containing 80 ml SVIII medium (Barrelmus & Perschak, 1987) containing 0-12% glucose at 55 °C until the mid-exponential growth phase was reached. Cell pellets obtained by centrifugation at 14000 g for 5 min at 4 °C were prepared for SDS-PAGE as described by Laemmli (1970).

DNA manipulations and oligonucleotides used for PCR. Chromosomal DNA of the S⁺ and S⁻ strains 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 and separation of DNA fragments by agarose gel electrophoresis were performed as described by Sambrook et al. (1989). DNA fragments were recovered from agarose gels by using the Qiaex II Gel Extraction Kit (Qiagen). Restriction endonucleases were purchased from Roche Molecular Biochemicals. All oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer.

Detection and PCR amplification of ISBst12. The presence of the sbsC gene sequence was investigated in the S⁺ strain by using the primer combinations sbsC3/12, sbsC3/14, sbsC14/22, sbsC21/22 and sbsC12/19 (Table 1). PCR reactions were performed in a 50 µl reaction volume containing 240 µM each dNTP, 240 nM primer, 1-25 mM MgCl₂, 1 U Taq DNA polymerase (Roche Molecular Biochemicals) and 100 ng chromosomal DNA from either the S⁺ or S⁻ strain of *B. stearothermophilus* ATCC 12980 as template in 1× Taq reaction buffer (Roche Molecular Biochemicals) according to the manufacturer's instructions. Thirty cycles of amplification were performed in a thermocycler (Hybaid Touch Down Control). Each cycle consisted of a 30 s denaturation step at 95 °C, a 45 s annealing step with annealing temperatures depending on the calculated T₅₀ of the oligonucleotides used and extension times of 60 s per 1000 bp at 72 °C. PCR fragments generated in at least four separate reaction vials were pooled and the mixture was subjected to DNA sequencing. For investigation of the occurrence of ISBst12 in *B. stearothermophilus* PV72/p6, PV72/p2 and PV72/T5, PCR reactions were carried out as described above, except that 1 µl of the suspensions of exponentially growing cells were used as the template. The primer combination ISBst12-1/2 was used to detect ISBst12 in these organisms (Table 1). As a control, the S⁺ and S⁻ strains of *B. stearothermophilus* ATCC 12980 were included into the study.

DNA sequencing of ISBst12. Nucleotide sequence determination of ISBst12 was carried out on a PCR fragment generated from the genomic DNA of the S⁺ strain of *B. stearothermophilus* ATCC 12980 by using the sbsC-specific primers sbsC3 and sbsC14 (Table 1). PCR-product sequencing was performed twice, once from each strand, by the dideoxy-chain termination method with a Perkin Elmer Applied Biosystems.
**Table 1. Oligonucleotides used for synthesis of PCR fragments**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’–3’)</th>
<th>Position</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>sbsC3</td>
<td>TACAGCCATACGGTAACCGAAACG</td>
<td>154 to 177</td>
<td>Forward</td>
</tr>
<tr>
<td>sbsC12</td>
<td>GATTTTCTTCCGCGAGCTTTTC</td>
<td>3336 to 3334</td>
<td>Reverse</td>
</tr>
<tr>
<td>sbsC14</td>
<td>GTAAGGTCGATGTACGTAGCCAC</td>
<td>384 to 361</td>
<td>Reverse</td>
</tr>
<tr>
<td>sbsC19</td>
<td>TGTAACCTGGTCTAGACTTATGTCG</td>
<td>2661 to 2684</td>
<td>Forward</td>
</tr>
<tr>
<td>sbsC21</td>
<td>GCCTAAAGGTGCTAACATGAG</td>
<td>–78 to –100</td>
<td>Reverse</td>
</tr>
<tr>
<td>sbsC22</td>
<td>ACAGCGCGATTACGCAATGCCAG</td>
<td>–536 to –514</td>
<td>Forward</td>
</tr>
<tr>
<td>ISBst12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISBst12-1</td>
<td>CAACCGTTCTCAAAGTCGATTCG</td>
<td>384 to 407</td>
<td>Forward</td>
</tr>
<tr>
<td>ISBst12-2</td>
<td>GCCAACAGAGACATACCGATTC</td>
<td>1412 to 1389</td>
<td>Reverse</td>
</tr>
</tbody>
</table>

**Isolation of RNA.** Total RNA was isolated from cultures of exponentially growing cells of the S+ and S− strains of *B. stearothermophilus* ATCC 12980 which had reached an OD₆₀₀ of 0.78. For enzymic lysis, the cell pellet obtained by centrifugation of 4·5 ml 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 °C.

**DNA and RNA hybridization.** Northern and Southern blotting was performed as described by Sambrook et al. (1989). Southern blots were carried out with chromosomal DNA from the S+ and S− strains of *B. stearothermophilus* ATCC 12980 digested with either HindIII or SalI. For hybridization of Southern blots, a DNA probe generated by using the primer combination ISBst12-1/2 was labelled with digoxigenin-11-dUTP (DIG-dUTP) by incorporation during PCR. Hybridization was performed according to the manufacturer’s recommendations. For Northern blotting, total RNA from the S+ and S− strains (5 µg per well) was fractionated by electrophoresis through a 1% agarose/0·22 M formaldehyde gel. Hybridization of Northern blots was carried out either with the ISBst12-specific DNA probe or with a 3·3 kb DNA fragment of the sbsC gene generated by using the primer pair sbsC3/12 and randomly labelled with DIG-dUTP. Hybrids were detected using the DIG Luminescent Detection Kit (Roche Molecular Biochemicals). The size of the transcripts was estimated from their mobilities relative to those of the RNAs in the DIG-dUTP labelled RNA Molecular Mass Marker II (size range 1·5–6·9 kb; Roche Molecular Biochemicals).

**RESULTS**

**Comparative studies between the S+ and S− strains of *B. stearothermophilus* ATCC 12980**

The S− strain was isolated from SVIII agar slants of the S-layer carrying wild-type strain *B. stearothermophilus* ATCC 12980 which were stored for 2 years at 4 °C (Eder, 1983). With the exception of the S-layer protein band showing an apparent molecular mass of 122 kDa, both organisms revealed an identical pattern of protein bands on SDS gels (Egelseer et al., 1996). As demonstrated by RAPD fingerprinting using four different RAPD analysis primers, the S+ and S− strains had an identical pattern of randomly amplified DNA fragments (data not shown). Sequencing of 450 bp in the variable region of the 16S rDNA confirmed that the similarity values for the S+ and S− strains of *B. stearothermophilus* ATCC 12980 are 100% (DSMZ, Braunschweig; data not shown).

**Localization and sequence characterization of ISBst12**

By using the sbsC-specific primer combinations sbsC3/12, sbsC3/14 and sbsC14/22, the PCR-generated fragments of the sbsC gene from the S− strain increased in size by approximately 1600 bp compared to that from the S+ strain (Fig. 1, lanes 2, 4 and 6). However, by using the sbsC-derived primer combinations sbsC21/22 and
**Table 2.** Putative transposases or other proteins from different organisms showing significant identities to the putative transposase encoded by ISBst12

<table>
<thead>
<tr>
<th>Accession no.</th>
<th>IS element</th>
<th>Protein</th>
<th>Organism</th>
<th>Identity Score (bits)</th>
<th>Identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAF12669</td>
<td>Putative transposase</td>
<td>Deinococcus radiodurans</td>
<td>1940</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>AAC82847</td>
<td>Gas-vesicle protein (ORF H0698)</td>
<td>Halobacterium sp. plasmid pNRC 100</td>
<td>994</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>BAA33622</td>
<td>IS22-1 Putative transposase</td>
<td>Vibrio cholerae</td>
<td>632</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>AAA98376</td>
<td>IS66 Putative transposase</td>
<td>Agrobacterium tumefaciens Ti-plasmid pTiA66</td>
<td>566</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>AAD33661</td>
<td>ISRm14 Unknown protein</td>
<td>Sinorhizobium meliloti</td>
<td>426</td>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>

---

**Fig. 2.** Schematic drawing of the putative transposase encoded by ISBst12. The numbers indicate the amino acid positions of the respective motifs and regions. The leucine-zipper motif was detected in transposases from the IS3 family as well as from the eukaryotic mariner/Tc family whereas the His-Arg-Tyr triad is characteristic of integrases of the bacteriophage λ Int family and could also be detected in the IS1 transposase. The pIs of the basic and acidic regions are given. The α-helices are represented by barrels. Tpase, transposase.

sbsC12/19, comprising the upstream and the downstream region of the sbsC gene, respectively, the PCR-generated fragments from the S+ and S− strains were the same size (Fig. 1, lanes 7–10). These preliminary results obtained from PCR indicated that the coding region of the sbsC gene in the S− strain carried an insertion of approximately 1600 bp. The nucleotide sequence determination of the insertion was carried out on an 1800 bp PCR fragment generated from genomic DNA of the S− strain by using the primer combination sbsC3/14 (Fig. 1, lane 4). The insertion site of the foreign DNA was found to be located within the coding region of the sbsC gene, 199 bp downstream from the ATG start codon. Due to the insertion of the foreign sequence, a stop codon was introduced into the ORF of the sbsC gene, 199 bp downstream from the ATG start codon. This DNA sequence displayed the structural features of bacterial IS elements and thus was designated ISBst12 (accession no. AF162268). The G+C content of ISBst12 was determined to be 47%. The IS element contained one ORF of 1446 bp, predicted to encode a protein of 482 aa. The ORF started with ATG, preceded by a typical prokaryotic ribosome-binding site (AAGGAGG).

**Characterization of the putative transposase encoded by ISBst12**

According to the sequence data, the putative transposase has a calculated molecular mass of 55562 Da. The calculated value for the pI of this protein is 9.13. At the nucleotide-sequence level, no sequence identities to
other IS elements were identified. However, by using the BLAST program (Altschul et al., 1997), the predicted amino acid sequence of the ORF encoded by ISBst12 revealed sequence identities to a putative transposase of Deinococcus radiodurans (White et al., 1999), a gas-vesicle protein (ORF H0698) encoded by the Halobacterium sp. plasmid pNRC 100 (Ng et al., 1991), a putative transposase encoded by IS66 found in the T-DNA region of the mutant Ti-plasmid pTiA66 of Agrobacterium tumefaciens (IS66 family) (Michida et al., 1984) and a protein encoded by ISMr14 of Sinorhizobium meliloti (Schneiker et al., 1999), a protein encoded by IS22-1 of Vibrio cholerae (Yamasaki et al., 1999), a protein encoded by IS66 found in the T-DNA region of the mutant Ti-plasmid pTiA66 of Agrobacterium tumefaciens (IS66 family) (Michida et al., 1984) and a protein encoded by ISMr14 of Sinorhizobium meliloti (Schneiker et al., 1999) (see Table 2). The results obtained by sequence comparison supported that the ORF of ISBst12 encodes a putative bacterial transposase.

By scanning the protein sequence, two overlapping leucine-zipper motifs [LX(6)LX(6)LX(6)L] (Landschulz et al., 1988) were identified at the N-terminal part of the potential transposase (aa 28–56) (Fig. 2). The leucine zipper is followed by a basic region (aa 57–120) with a calculated pl of 10.37 (Fig. 2). At the very C-terminal end of the protein, an acidic region with a calculated pl of 5.74 was detected (aa 401–482). Furthermore, an amino acid triad represented by His-276, Arg-279 and Tyr-308 was identified which resembles the signature of the highly conserved His-Arg-Tyr triad characteristic of the active site of integrases of the bacteriophage λ Int family (Figs 2 and 3) (Abremski & Hoess, 1992; Argos et al., 1986; Serre et al., 1995). The His-Arg-Tyr triad detected in the ISBst12 transposase was compared to the consensus sequence established for the active site of the integrase-family recombinases by Serre et al. (1995), as shown in Fig. 3. Secondary-structure prediction (Rost & Sander, 1993) of the whole transposase encoded by ISBst12 revealed that 51% of the protein is organized as α-helices, whereas the rest of the amino acid sequence forms random coils (35%) and β-strands (14%). The segment between aa 9 and 61, comprising the leucine zipper, is organized as α-helices with high probability. As shown in Fig. 2, three big clusters of α-helices can be distinguished within the putative transposase and are located at the very N-terminal region (aa 9–61), in the middle part (aa 201–242) and at the very C-terminal region (aa 308–479).

Results from Southern and Northern blotting

After digestion of chromosomal DNA from the S+ and S− strains with HindIII or SalI, neither of which cut within the ISBst12 sequence, several bands were detected by an ISBst12-derived DNA probe, indicating the presence of multiple copies of the IS element in both strains (Fig. 4a, b, lanes 3 and 4). By using HindIII-digested chromosomal DNA, the ISBst12-specific DNA probe detected five distinct bands in the S+ strain (Fig. 4a, lane 3), whereas seven bands could be identified in the S− strain (Fig. 4a, lane 4). This result reflected the insertion of at least two more copies of ISBst12 elsewhere in the genome of the S− strain. Furthermore, Southern blotting was used to identify the copy of ISBst12 inserted into the sbsC gene. Since the sbsC gene carries two HindIII sites, hybridization of HindIII-digested chromosomal DNA with an sbsC-derived probe detected three fragments in the S+ strain (7100, 5490 and 1500 bp) (Fig. 4a, lane 1). In the S− strain, only two bands (7100 and 1500 bp) were detected (Fig. 4a, lane 2). This 7100 bp band consists of two fragments of nearly the same size, since insertion of ISBst12 (1612 bp) into the 5490 bp fragment carrying the first 819 bp of the coding region of the sbsC gene led to a fragment of 7100 bp, too (Fig. 4a, lane 2). Furthermore, hybridization was carried out with chromosomal DNA digested with SalI, which has one restriction site in the coding region of sbsC. After hybridization of SalI digested chromosomal DNA of the S+ and S− strains with the ISBst12-specific probe, one additional band of 5260 bp was detected in the S− strain (Fig. 4b, lane 4). This distinct band carried the copy of ISBst12 inserted into the sbsC gene, since a fragment of
the same size was also detected by the sbsC-specific DNA probe in the S+ strain (Fig. 4b, lane 2).

Northern blotting of total RNA (Fig. 5) from the S+ strain using an sbsC-derived DNA probe revealed no sbsC-specific transcript (Fig. 5b, lane 2). In contrast, in total RNA of the S− strain, an sbsC-specific mRNA of approximately 3.6 kb was detected (Fig. 5b, lane 1). By using an ISbst12-specific DNA probe, an IS-element-specific mRNA of about 1.6 kb was only present in total RNA preparations of the S− strain (Fig. 5a, lane 2) and not in the S+ strain (Fig. 5a, lane 1) suggesting that transcription of ISbst12 is inhibited or considerably reduced in the S+ strain.

Occurrence and distribution of ISbst12

For detection of ISbst12 in other organisms, the primer combination ISbst12-1/2 was used in PCR which, according to the sequence data, should generate an ISbst12-specific fragment of 1029 bp. On agarose gels, a corresponding ISbst12-specific fragment was detectable in B. stearothermophilus PV72/p6, its oxygen-induced strain variant PV72/p2 and in its S-layer-deficient strain PV72/T5 (data not shown). The data were confirmed by using the S+ and S− strains of B. stearothermophilus ATCC 12980 as a control, which led to PCR products of the same size (data not shown).

DISCUSSION

In the present study, it was demonstrated that in the S+ strain of B. stearothermophilus ATCC 12980, the coding region of the sbsC gene was still present but was interrupted by a novel type of bacterial IS element designated ISbst12. Inhibition of S-layer-protein expression by insertion of an IS element was reported for the first time for A. salmonicida (Gustafson et al., 1994).

In the case of this fish-pathogenic organism, 10 independent mutants that exhibited either reduced synthesis of the A-layer protein, synthesis of truncated subunits or complete loss of A-layer gene expression were isolated by raising the growth temperature from 20 °C to 30 °C (Gustafson et al., 1994). These mutations resulted from insertion of two different IS elements (ISAS1 and ISAS2) in the vapA gene and its promoter. While ISAS1 was revealed to be unique among reported IS elements, ISAS2 showed high sequence identity to transposons encoded by the IS30 family. Both IS elements were found to be restricted to two A. salmonicida strains, where they were present in low copy number. Temperature-dependent transposition was also described for IS4712 (accession no. AJ223150), an IS element that inhibits expression of the S-layer gene sbsA in the S-layer-deficient strain of B. stearothermophilus PV72/p6, designated PV72/T5. Integration of IS4712 into the regulatory region of the sbsA gene was achieved by cultivating the organism for several generations at 67 °C instead of 57 °C (Scholz, 1998). However, in the case of the S− strain from B. stearothermophilus ATCC 12980, which was isolated from cultures of the S− strain that were stored for a long period of time at 4 °C, temperature shift did not induce the integration of ISbst12 into the upstream region of the sbsC gene (E. M. Egelseer, unpublished observation). In both organisms, environmental-stress factors may have triggered transposase activity, being in accordance with the findings that IS elements are responsible for the diversity in bacterial populations. On the other hand, it was demonstrated that transposition is coupled to the physiological state of the host cell or to a certain state within the cell cycle (Kleckner, 1990).

As revealed by Southern blotting, ISbst12 was present in the genome of the S− strain as well as in the S+ strain in multiple copies, which was also described for Lactococcus lactis harbouring at least 12 copies of IS905 (Dodd et al., 1994). The hybridization patterns, which were slightly different between the S+ and S− strains, indicated the insertion of at least two more copies of ISbst12 elsewhere in the genome of the S− strain. Furthermore, an ISbst12-specific transcript could exclusively be detected in the S− strain. Taken together, these results could indicate that an active transposase is only present in the S− strain, leading to an increase in ISbst12 copy number compared to the S+ strain in which transcription of ISbst12 seems to be inhibited or considerably reduced. Furthermore, by using PCR, ISbst12 was also detected in B. stearothermophilus PV72/p6, its oxygen-induced strain variant PV72/p2 and the S-layer-deficient strain PV72/T5 (Sára et al., 1996). The latter also carries the IS element IS4712, which was found to be responsible for the loss of S-layer gene expression (Scholz, 1998).

The putative transposase encoded by ISbst12 has a calculated pl of 9.13, which is consistent with a probable DNA-binding property required for transposase activity. The N-terminal region of the transposase displays two overlapping leucine-zipper motifs (Landschulz et
al., 1988). These heptad repeats of leucines were originally described as a protein-dimerization motif for the formation of coiled-coil intertwining of α-helices for several eukaryotic transcriptional regulators and also for some prokaryotic DNA-binding proteins like MetR (Maxon et al., 1990) and σ^4 (Sasse-Dwight & Gralla, 1990). Interestingly, leucine-zipper motifs were also detected in many members of the IS3 family, for example in IS2 (Lei & Hu, 1997), IS911 (Haren et al., 1997) and IS1221 (Zheng & McIntosh, 1995). This is consistent with the observation that some, if not all transposases, including prokaryotic and eukaryotic elements such as retroviruses, have the capacity to generate multimeric forms essential for their activity (Polard & Chandler, 1995). In the putative transposase encoded by ISBst12, a highly basic region was identified adjacent to the leucine-zipper motif that is characteristic of eukaryotic DNA-binding proteins but was absent in the prokaryotic DNA-binding proteins MetR and σ^4 as well as in IS1221 (Zheng & McIntosh, 1995). From the increasing number of different transposable elements isolated and characterized at the nucleotide-sequence level, a general pattern for the functional organization is emerging, namely that the sequence-specific DNA-binding activities are located at the N-terminal region, while the catalytic domain is mostly located at the C-terminal end (Mahillon & Chandler, 1998). One functional interpretation of this arrangement for prokaryotic elements is that it may permit the interaction of a nascent polypeptide chain with its target sequences on the IS, thus coupling expression and activity. This notion is reinforced by the observation that the presence of the C-terminal region of several transposases [IS30 (Weinreich et al., 1993), IS10 (Jain & Kleckner, 1993)] appears to mask the DNA-binding domain and decreases the binding activity, thereby favouring the activity of the protein in cis. This preferential activity in cis reduces the probability that transposase expression from a given element would activate transposition of related copies elsewhere in the genome. This could also be true of the putative transposase encoded by ISBst12 carrying a highly acidic region towards the C-terminal end which after complete folding of the polypeptide chain probably masks the basic DNA-binding region located at the N-terminal part.

Comparative studies between bacterial transposases and retroviral integrases revealed similarities in a region that is thought to form part of the active site, namely the DDE motif (Polard & Chandler, 1995). This highly conserved acidic amino acid triad was found to be intimately involved in catalysis by coordinating divalent metal cations (in particular Mg^2+) implicated in assisting the various nucleophilic attacking groups during the course of the reaction. The majority of the IS elements found in the genus Bacillus were assigned to the IS4 family (Mahillon & Chandler, 1998). This family is quite heterogeneous but it is characterized by a conserved DDE signature (Rezsöhazy et al., 1993). However, the putative transposase encoded by ISBst12 does not exhibit a typical DDE triad. Although members of the DDE family represent the majority of known IS elements, and mutagenic studies clearly underlined the importance of these residues, a significant fraction of IS elements does not exhibit a real or potential DDE triad. Interestingly, a highly conserved His-Arg-Tyr triad was identified in the putative transposase which resembles the signature of the catalytic site of integrases of the bacteriophage λ Int family and was also detected in the C-terminal part of the IS1 transposase (Abremzki & Hoess, 1992; Argos et al., 1986; Serre et al., 1995). For the IS1 transposase it was demonstrated that each of the three amino acid residues of the conserved triad is important for transposase activity (Serre et al., 1995).

While no sequence identities to IS elements were identified at the nucleotide-sequence level, the protein encoded by ISBst12 revealed identity to several putative transposases encoded by bacterial IS elements. The transposase encoded by ISBst12 showed the highest identity value (35%) to a recently identified putative transposase located on a 46 kb plasmid of D. radiodurans (White et al., 1999). This extremely radiation-resistant bacterium contains numerous insertion sequences (52 copies) and small non-coding repeats (247 copies) whose evolutionary significance and role in genome function remain unclear (White et al., 1999; Makarova et al., 1999). Furthermore, a gas-vesicle protein (ORF H0698) that was detected in both large inverted repeats identified on the Halobacterium sp. plasmid pNRC 100 (Ng et al., 1991) revealed identity to the putative ISBst12 transposase. These two large inverted repeats were found to mediate inversion of the intervening single-copy region (Ng et al., 1991). The transposase encoded by ISBst12 showed identity to a putative transposase encoded by IS22-1 of V. cholerae as well as to two members of the IS66 family, namely to a protein encoded by IS66 of A. tumefaciens and to a protein encoded by ISRm14 of S. meliloti. Like ISBst12, members of the IS66 family are flanked by 8 bp direct target repeats and by terminal inverted repeats of 15–27 bp, which are very similar among members of this family. Both 16 bp imperfect inverted repeats of ISBst12 start with 5'-GTAA-3', a sequence which seems to be conserved among the inverted repeats of several members from the IS66 family (Mahillon & Chandler, 1998). The IS66 family was found to be restricted to agrobacteria and rhizobia (Mahillon & Chandler, 1998). With the development of studies on the mechanism of bacterial pathogenesis, an association between IS elements and many pathogenic and virulence functions became evident. Such associations have been observed in animal pathogens like Vibrio, in plant pathogens like Agrobacterium or in symbionts like Rhizobium.

To conclude, ISBst12 represents a novel type of IS element, whose putative transposase exhibits a well defined leucine-zipper motif as well as a His-Arg-Tyr triad and reveals sequence identity to transposases from different IS families detected in distantly related genera of bacteria (Deinococcus, Vibrio, Agrobacterium, Sinorhizobium) and even to a plasmid-encoded halobacterial protein. Since it could be speculated that this novel type
of IS element may represent an ancestral IS, the distribution and abundance among bacterial, archaeal and even eukaryotic genera remains to be investigated.

ACKNOWLEDGEMENTS

We gratefully acknowledge Jacques Mahillon for helpful suggestions and discussions. We thank Christoph Hotzy and Aida Medovic for excellent technical assistance. This work was supported by the Austrian Science Foundation, project P12938 and by the Ministry of Science and Transports.

REFERENCES


Charaterization of ISBst12


Received 15 May 2000; revised 20 June 2000; accepted 23 June 2000.