The Lactobacillus acidophilus S-layer protein gene expression site comprises two consensus promoter sequences, one of which directs transcription of stable mRNA.

Boot, H.J.; Kolen, C.P.A.M.; Andreadaki, F.; Leer, R.J.; Pouwels, P.H.

Published in:
Journal of Bacteriology

DOI:

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The *Lactobacillus acidophilus* S-Layer Protein Gene Expression Site Comprises Two Consensus Promoter Sequences, One of Which Directs Transcription of Stable mRNA

HEIN J. BOOT,1* CARIN P. A. M. KOLEN,1 FOTINI J. ANDREADAKI,1 ROB J. LEER,2 AND PETER H. POUWELS1,2

BioCentrum Amsterdam, University of Amsterdam, 1018 TV, Amsterdam, and TNO Nutrition and Food Research Institute, 2280 HV Rijswijk, The Netherlands

Received 23 April 1996/Accepted 2 July 1996

S-proteins are proteins which form a regular structure (S-layer) on the outside of the cell walls of many bacteria. Two S-protein-encoding genes are located in opposite directions on a 6.0-kb segment of the chromosome of *Lactobacillus acidophilus* ATCC 4356 bacteria. Inversion of this chromosomal segment occurs through recombination between two regions with identical sequences, thereby interchanging the expressed and the silent genes. In this study, we show that the region involved in recombination also has a function in efficient S-protein production. Two promoter sequences are present in the S-protein gene expression site, although only the most downstream promoter (P-1) is used to direct mRNA synthesis. S-protein mRNA directed by this promoter has a half-life of 15 min. Its untranslated leader can form a stable secondary structure in which the 5' end is base paired, whereas the ribosome-binding site is exposed. Truncation of this leader sequence results in a reduction in protein production, as shown by reporter gene analysis of *Lactobacillus casei*. The results obtained indicate that the untranslated leader sequence of S-protein mRNA is involved in efficient S-protein production.

The cell walls of many bacterial species are covered by a crystalline structure known as the S-layer (for reviews, see references 3 and 26). The monomolecular S-layer is composed of a single species of protein, the S-protein, which is capable of crystallizing into a regular array. S-layers have been found on eu- and archaebacteria living in all kinds of environmental conditions and are considered membrane-like structures on the outside of the cell. Several specific functions have been reported for S-layers of different species; however, no general function is known. For some archaebacteria, it was found that the S-layer is involved in maintenance of the cell shape (33). The S-layer plays an important role in the overall virulence of several pathogenic bacteria (16). An S-layer can also act as a shield to cover phage receptors present in the underlying cell wall (16, 29) or prevent invasion by *Bdellovibrio bacteriovorus* (19). Furthermore, S-layers can serve as an attachment structure for extracellular proteins (25). For both pathogenic (11) and nonpathogenic (42) bacteria colonizing the gastrointestinal tracts of mammals, it was found that the S-layer interacts with components of the epithelial cell layer.

*Lactobacilli* are found widely distributed throughout nature and have been used for a long time in food and feed production. Several reports have appeared ascribing probiotic properties to certain *Lactobacillus* species which are naturally found in the gastrointestinal tracts of humans and animals (e.g., *Lactobacillus acidophilus* and *Lactobacillus gasseri*), or species which are used in the production of standard yogurt (i.e., *Lactobacillus delbrueckii* subsp. *bulgaricus* or health promoting yogurt variants (e.g., *Lactobacillus casei*). Effects reported to be induced by these lactobacilli are lowering of levels of cholesterol in serum (12), binding of mutagenic compounds (28), and stimulation of immunoglobulin production (20). In addition, protection of the host against invasion by pathogenic bacteria is reported to be partially due to the presence of lactobacilli.

The type strain of the *L. acidophilus* species was originally isolated from the human pharynx and is considered to be probiotic (18). The *slpA* gene, encoding the S_A-protein which forms the wild-type S-layer of this species, has been cloned and sequenced previously (7). Near this expressed gene, we have identified a silent gene (*slpB*) encoding the S_P-protein (5). The *slpA* gene is interchanged with the *slpB* gene through inversion of a chromosomal fragment in a fraction of an *L. acidophilus* culture (0.3% of the cells grown under laboratory conditions). Such a recombination event is expected to lead to the production of a partially different S-protein and S-layer (S-layer variation) (6). It is estimated that the S-protein represents 10 to 15% of the total cellular protein of the bacterial cell. This implies that the single S-protein gene must be transcribed at a high level and/or give rise to a stable mRNA. Why the S-protein genes of some species are preceded by a multiple promoter structure, whereas the S-protein genes of other species are preceded by a single promoter is presently unclear. To gain further insight into the regulation of the expression of the *L. acidophilus* S-protein gene(s), we have analyzed the expression directed by the S-promoter. We found that although two putative promoter sequences are present, only one is used under all of the tested growth conditions. The long untranslated leader sequence present in the mRNA appears to be involved in efficient production of the S-protein.

**MATERIALS AND METHODS**

**Strains and plasmids.** *L. acidophilus* ATCC 4356 and *L. casei* ATCC 393, which were obtained from the American Type Culture Collection, were cultured anaerobically in MRS (Difco) broth at 37°C. Plasmids pBK-1 and pBK-2

---

* Corresponding author. Present address: Institute for Animal Science and Health (ID-DLO), P.O. Box 65, NL-8200 AB Lelystad, The Netherlands. Phone: 31 320 238881. Fax: 31 320 238668. Electronic mail address: H.J.Boot@id.dlo.nl.
are multicopy plasmids containing the entire slpA gene and flanking sequences (7). Escherichia coli DH5α was used as a host for recombinant plasmids. Isolation of chromosomal DNA and total RNA. Chromosomal DNA and total mRNA of L. acidophilus ATCC 4356 were isolated essentially as described by Sambrook (5).

Northern (RNA) and Southern (DNA) blot analysis. Total RNA was separated in either a 1.0 × or a 1.0% TBE (10× TBE is 10× g of Tris, 55 g of boric acid, and 10 g of Na₂EDTA) agarose gel without ethidium bromide. The gel was soaked twice for 20 min in 0.2 M sodium acetate (pH 5.0) and transferred to a Hybond filter (Amersham) by blotting with the same solution. Southern analysis was performed essentially as described by Southern et al. (40). Restriction enzyme-digested chromosomal DNA was separated on a 1% agarose gel and transferred to a Hybond filter. RNA and DNA were cross-linked to the filter by short ultraviolet (UV) light exposure for 5 min. Prehybridization and hybridization of the filters were performed in 10 ml of 5× SSC (1× SSC is 0.15 M NaCl plus 0.15 M sodium citrate)–5× Denhardt solution (35–0.1% sodium dodecyl sulfate (SDS))–1.0 mg of sonicated salmon sperm DNA at 65°C. The filters were washed twice with 2× SSC–0.1% SDS and three times with 0.5× SSC–0.1% SDS at 65°C. Probe S was a restriction fragment of plasmid pBK-2 and was located entirely within the coding sequence of slpA (nucleotides [nt] +353 to +1135, relative to the start of the coding sequence (7)). Probes Sₐ, Sₚ, and S₂ were generated by PCR using the specific oligonucleotides as indicated in Fig. 1. Probe Sₐ is a 701 bp restriction fragment for mapping the entire coding sequence of the slpA gene (nt +99 to +1445). The EF-Tu probe was generated with primers against part of the gene encoding conserved amino acid sequences of known EF-Tu proteins of other bacterial species (22). These primers, ML6-1 (−AT(A,G).TAT.TC(A,G).TC(A,T,G).AC(A,T,G).GT(A,G).TCC.AT-3) and ML6-4 (5′-AT(AG).TAT.TC(AG).TC(AT.G).AC(T,G).GT(G,A).TCC-3′), generated a DNA fragment of the expected length (140 bp) in a PCR experiment when chromosomal DNA of L. acidophilus ATCC 4356 was used as a template. PCR conditions were taken from the restriction fragment-based probe screening conditions with optimization of the annealing temperature (44°C). The PCR products were self-ligated and used to transform E. coli DH5α, yielding plasmid pBK-62. Plasmid DNA was used to transform L. casei ATCC 393 by electroporation, and recombinant bacteria were selected by plating the transformation mixture on MRS plates containing 5 μg of erythromycin per ml.

Determination of CAT activity. L. casei ATCC 393 bacteria which harbor pBK-62 derivatives of this plasmid were grown overnight anaerobically in MRS broth at 37°C in the presence of 5 μg of erythromycin per ml. Fresh MRS was inoculated (1:100) with this pre-culture, and the mixture was incubated until an optical density at 600 nm of between 0.5 and 1.0 (±0.2) was reached by centrifugation (5.000 × g, 30 min). The cell pellet was washed with 1.0 ml of 20 mM Hepes (pH 7.4) and resuspended in 50 μl of 20 mM N-2-hydroxypropyicarbene-N′-2-ethanesulfonic acid (HEPES; pH 7.4), and 0.5 mg of glass beads (9 × 0.45 mm) was added. This suspension was vortexed for 1 min and centrifuged (5.000 × g, 1 min). The supernatant was used as a total-cell extract for the determination of CAT activity. The CAT activity was determined in three independent cultures for each plasmid which were determined twice spectrophotometrically with the 5.5′-dithiobis-(2-nitrobenzoic acid) (DTNB) assay described by Burtis and Ashwood (39), with a Cobas Bio automatic analyzer (Hoffmann-La Roche). Specific CAT activities were calculated after determination of protein contents of the total-cell extracts by the bicinchoninic acid (Sigma) method (39). A Southern blot experiment with total DNA of the L. casei ATCC 393 recombinant strains indicated that the copy numbers of all pBK-50-based plasmids are approximately the same (data not shown).

RESULTS

Mapping of the transcription start. Two putative promoter sequences (Fig. 1 [P-1 and P-2]) are present in the 500-bp region upstream of the slpA ORF of wild-type L. acidophilus ATCC 4356 cells. Total RNA of an exponentially growing L. acidophilus ATCC 4356 culture was extracted. This RNA was purified in an agarose gel and then used as a template for primer extension analysis (1.0× TBE) or prevent secondary-structure formation (1.0× TBE) (14) and was used for Northern blot analysis. mRNA encoding the S₂ protein was detected by using a radioactively labelled part of the slpA gene (probe S; see Materials and Methods). A sharp band was visible on the autoradiogram of the Northern blot when secondary-structure formation was not prevented at a position below that of the 16S rRNA (Fig. 2A). When conditions were used which resulted in the loss of secondary structure of the RNA, a broad band was visible at the position corresponding to the position of 16S rRNA (Fig. 2A). These results indicate that the size of the slpA mRNA is approximately the same as that of the 16S rRNA (1,500 nt). A transcription terminator-like hairpin structure is found behind the slpA gene, which ends 63 nt downstream from the transcriptional stop codons (7). Since the terminal gene encodes an extracellular acid resistant, this implies that the slpA gene is monocistronic and that the start point of the mRNA is located approximately 100 nt upstream from the translation start codon. The exact position of the transcription start point was identified by a primer extension analysis with the same RNA preparation. A single
signal at position −191 nt relative to the start of the coding sequence (Fig. 2B) and 8 nt downstream of promoter P-1 was detected (Fig. 1). No extension signal is present at the position which corresponds to slpA mRNA directed by the putative promoter P-2, indicating that this promoter is not active during exponential growth under the applied growth conditions.

**Promoter activity during different growth stages.** Only promoter P-1 was found to be used for transcription of S-protein mRNA in all RNA samples derived from bacteria in different phases of growth (Fig. 3). The first RNA sample (RNA-1) was obtained from cells in the mid-exponential growth phase, the second sample (RNA-2) was derived from cells at the end of the exponential growth phase, and the third sample (RNA-3) was derived from cells at the beginning of the stationary growth phase.
phase (Fig. 3A). To determine whether specific induction of S-protein mRNA will occur in L. acidophilus bacteria deprived of the S-layer (e.g., autoregulation), we isolated RNA from an exponentially growing culture which was treated with lithium chloride and cultivated for an additional 30 min in fresh MRS broth (RNA-L; see Materials and Methods). The lithium chloride treatment specifically extracts the S-protein from the cell wall without being deleterious (21). RNAs isolated from these culture samples were used for Northern blotting, and specific mRNAs were detected with different probes. Three probes derived from different parts of the slpA region were used to detect mRNA directed by the S-promoter. Probe S is entirely located within the coding region of the slpA gene, and probe SP-1 overlaps the transcription start of mRNA derived from both promoters P-1 and P-2, whereas probe SP-2 overlaps only the transcription start of mRNA derived from promoter P-2 (Fig. 1 and legend). Comparison of the three probes for the slpA mRNA shows that no hybridization signal is found when the SP-2 probe is used, while probes S and SP-1 show a hybridization signal for all RNA samples. Probes SP-1 and SP-2 are equally labelled, which can be concluded from the intensity of the signals in the Southern blot control experiment (Fig. 3C). Furthermore, we used a probe directed against mRNA for the EF-Tu protein as a control for mRNA detection in the different samples. The amount of mRNA for the EF-Tu protein is, at least for E. coli, known to vary in direct proportion to the growth rate (43). The hybridization signals obtained for L. acidophilus mRNA by using an EF-Tu probe are consistent with this idea (Fig. 3B). Primer extension analysis confirmed that only promoter P-1 is used to direct the synthesis of slpA mRNA. Only signals representing slpA mRNA directed by promoter P-1 were found when RNAs from the different stages of growth were used (data not shown).

Stability of S-protein mRNA. Computer analysis shows that the untranslated leader sequence of the mRNA which is directed by promoter P-1 can fold into a large stem-loop structure (Fig. 4) with an energy of −191 kJ/mol (MFold; Genetics Computer Group). A stably folded 5' end can protect mRNA from 5'→3' degradation, thereby increasing the half-life. The half-life of the mRNA for the S$_{A}$-protein of L. acidophilus ATCC 4356 was determined by adding rifampin to an exponentially growing culture, thereby blocking further RNA synthesis. Total RNA was isolated at 5-min time intervals from samples of this culture and analyzed in a Northern blot (Fig. 5A) by using a probe which comprises the entire coding region of the slpA gene (probe S; see Materials and Methods). The PhosphorImager data of this experiment are presented in Fig. 5B. From these data, it can be deduced that mRNA of the S$_{A}$-protein has a half-life of 15 min.

S-promoter activity in L. casei ATCC 393. The intact untranslated leader of the S-protein mRNA gives rise to a reporter protein level higher than that of a truncated leader sequence (Fig. 6). The S-promoter and the untranslated leader sequences were cloned before a CAT reporter gene of a broad-host-range plasmid (see Materials and Methods), yielding pBK-61. To determine whether the untranslated leader sequence of the transcript directed by the S-promoter influences the reporter protein level, we deleted more than half of the DNA sequence (147 nt [Fig. 4; see Materials and Methods]) encoding this region to construct the S$_{AUL}$-promoter (pBK-62). This truncated untranslated leader sequence is no longer able to fold into a stable stem-loop structure. Furthermore, we used the reporter plasmid without any promoter (pBK-50) and with the promoter of the lactate dehydrogenase gene (LDH-promoter, pBK-60) of L. casei ATCC 393 (17). The reporter protein levels of the above-mentioned plasmids were analyzed

FIG. 4. Computer prediction (MFold; Genetics Computer Group) of the secondary structure of the untranslated leader sequence of the S$_{A}$-protein mRNA. The RBS, 5'→3' end of the mRNA, and the position of the SphI restriction site in the corresponding DNA sequence are indicated. Base pairs consisting of guanine and adenine are indicated by dots. Prediction of secondary-structure formation with the RNAfold program gives essentially the same structure (data not shown), except that the region of the RBS is predicted to fold slightly differently. The 5' end is according to the RNAfold program still involved in stem formation, whereas the RBS is exposed.
in *L. casei*, a species which is related to *L. acidophilus* but which possesses no S-protein-encoding genes and is therefore not expected to have trans-acting regulatory factors specific for the S-promoter. Total-cell extracts of exponentially growing cultures harboring these plasmids were analyzed for CAT activity. The CAT activity of *L. casei* cells transformed with the *S*$_{\text{casei}}$-promoter plasmid appeared to be approximately half of the CAT activity of *L. casei* cells containing the plasmid with the complete untranslated region behind the S-promoter (Fig. 6), which itself was threefold higher than that of bacteria transformed with a plasmid containing the LDH-promoter.

**DISCUSSION**

Electron microscopic photographs of bacteria which possess an S-layer show that the complete cell surface is covered with an S-layer during the entire growth cycle. In addition, the presence of free S-protein in the growth medium, besides an intact S-layer, has been reported for several bacteria (1, 4, 31). Information about growth-dependent regulation of S-protein gene expression at transcriptional or translational levels is lacking for most of the known S-proteins. To gain further insight into factors influencing S-protein production, we analyzed the production of the S-protein of *L. acidophilus* ATCC 4356.

Multiple promoter structures have been found before several S-protein genes, including the S-protein of *Lactobacillus brevis* (44). For *Bacillus brevis*, it was shown that at least one of the five identified promoters before the S-protein-encoding operon is active only during the exponential growth phase (1), whereas another promoter is active during all growth phases. The latter promoter is repressed in the presence of divalent cations (2). Sequence analysis of the promoter region of the *L. acidophilus* ATCC 4356 *slp* expression site revealed two potential promoters, P-1 (−228 to −198) and P-2 (−335 to −303) (Fig. 1). These two potential promoters have the same −10 and −35 sequences, showing extensive sequence identity with consensus −10 and −35 promoter sequences of lactobacilli (24, 32). Both the P-1 and P-2 −10 and −35 promoter sequences are preceded by an A+T-rich region (20 bp). This region is known as the UP element and was found to stimulate transcription through interaction with the α subunit of the RNA polymerase in *E. coli* and *Bacillus subtilis* (34). The only difference between the P-1 and the P-2 promoters is the spacing of the −10 and −35 sequences. The −10 and −35 sequences of promoter P-1 have the optimal spacing of 17 nt, whereas the spacing of the P-2 −10 and −35 sequences have the suboptimal spacing of 20 nt (41). Although it is suboptimal, at least one active *L. delbrueckii* promoter is reported to possess a spacing of 20 nt (24). Transcription directed by a multiple promoter structure might contribute to an amount of mRNA higher than that of transcription directed by a single promoter. Another reason for the observed multiple promoter structure might be that the two promoters are active under different conditions. Both primer extension (Fig. 2B) and Northern blot (Fig. 3) analyses showed that only the most downstream promoter (P-1) is active during all growth conditions analyzed. Even in RNA of *L. acidophilus* cells which had been cultivated for 30 min in fresh medium after the S-layer had been extracted by a nonlethal procedure, no transcript from the more distal promoter (P-2) was detected (Fig. 3). The observation that only one of two promoter sequences of the *L. acidophilus* S-protein gene is active is in contrast to the situation described for *L. brevis*, in which the two identified promoters before the S-protein gene are about equally active during exponential growth (44).

Regulation of promoters is often found to take place by binding of a trans-acting factor (DNA-binding protein) to a hairpin structure which regulates the expression positively (activator) or negatively (repressor). The sequence capable of hairpin formation which is present before promoter P-2 of the

FIG. 5. Half-life determination of *S*$_{\text{A}}$-protein mRNA of *L. acidophilus* ATCC 4356. Samples were taken at 5-min intervals after addition of rifampin to a culture grown at 37°C in MRS broth. (A) Total RNA was extracted, and equal amounts of RNA were analyzed by Northern blot analysis with a probe (*S*$_{\text{II}}$) which completely overlaps the coding region of the *S*$_{\text{A}}$-protein mRNA. (B) Two Northern blot analyses were performed with independent RNA preparations, and the resulting signals were quantified with a PhosphorImager apparatus. The half-life (15 min) of the *S*$_{\text{A}}$-protein mRNA can be deduced from the graphic representation of the PhosphorImager data.

![Image](58x297)

**FIG. 6.** Reporter protein levels resulting from different promoter constructs in extracts of *L. casei* were analyzed. A broad-host-range plasmid containing a CAT reporter gene was used for expression analysis. The numbers at the bottom of the schematic drawings of the promoter regions represent nucleotides relative to the start of the different ORFs. *L. casei* was transformed with these plasmids. Purified transformants were cultivated in MRS broth, and cell extracts of exponentially growing cultures were analyzed for CAT activity. Values in parentheses are standard deviations of the mean values of CAT activities ($n = 6$).
slp expression site (Fig. 1) could be part of such a regulation mechanism. Regulation involving a transacting factor (AbcA, a leucine zipper-containing protein) and a potential inverted DNA sequence in the vicinity of one of the promoters of an S-protein gene has recently been described for *Aeromonas salmonicida* (27). It is not yet clear whether the *L. acidophilus* ATCC 4356 promoter P-2 is detectably activated or derepressed under any specific growth condition or whether this potential promoter structure is never used to direct transcription.

Efficient production of a protein encoded by a single gene is often found to be directed by an mRNA which is considerably more stable than average mRNAs of the same species. The half-lives of mRNAs encoding S-proteins from *Caulobacter crescentus* (13) and *A. salmonicida* (10) have been determined and were found to be 10 to 15 min and 22 min, respectively. We found that the mRNA for the S-protein of *L. acidophilus* ATCC 4356 also has a relatively long half-life, i.e., 15 min (Fig. 5). Formation of a hairpin-like structure by the 5' untranslated leader of the *S*-protein gene has recently been described for *E. coli* (30). Computer analysis predicts that the 191-nt-long untranslated leader of the *S*-protein mRNA, derived from P-1, will fold into an extensive hairpin-like structure (Fig. 4). According to this computer prediction of the secondary structure, the first nucleotides of the 5' end are involved in stem formation. Base pairing of the 5' end of mRNA has, for *E. coli*, been shown to be a prerequisite for preventing fast degradation of the mRNA by RNase E (8). The ribosome-binding site (RBS) is in the predicted secondary structure exposed in a loop. Exposure of the RBS is expected for a highly translated mRNA, since it has been shown that only an exposed RBS can hybridize with the complementary sequences of the 16S rRNA subunit of the ribosome (37). On the basis of computer analysis, the untranslated leader sequence of mRNA directed by promoter P-2 will also fold into a stable structure. However, in this prediction, the 5' end of the hypothetical mRNA is not part of a stem structure, and the nucleotide sequence which constitutes the RBS forms an internal duplex (data not shown).

The *L. acidophilus* S-promoter can be used in *L. casei* ATCC 393 for heterologous protein production. Bacteria belonging to the *L. casei* species do not possess an S-layer (23) and, in contrast to *L. acidophilus*, can efficiently be transformed with plasmids. Results based on enzymatic detection of a reporter protein (CAT) showed that the S-promoter is highly active in this host, since an activity approximately three times higher than that of the most active homologous *L. casei* promoter known to us was found with this promoter. This efficient production is in part due to the presence of the untranslated leader sequence of the S-protein mRNA, since deletion of part of this sequence diminished the reporter protein activity by nearly twofold (Fig. 6).

It appears that the region encoding the long untranslated leader sequences of the S-protein mRNA has a dual function. In earlier studies, we have shown that the DNA encoding the untranslated leader mRNA is involved in chromosomal recombination (6). A 6.9-kb chromosomal segment (slp segment), containing two S-protein-encoding genes in opposite orientations, is inverted in a fraction of the bacterial population. This inversion occurs between two regions with identical sequences (280 nt), which comprise the region encoding the untranslated leader sequence (180 nt) of both S-protein encoding genes. In this study, we have shown that this region is also important for efficient production of the S-protein, since it generates a long untranslated leader sequence in S-protein mRNA which can fold into a stable secondary structure.

REFERENCES


