JOURNAL OF BACTERIOLOGY, Oct. 1993, p. 6089–6096
0021-9193/93/196089-08$02.00/0

Vol. 175, No. 19

S-Layer Protein of *Lactobacillus acidophilus* ATCC 4356: Purification, Expression in *Escherichia coli*, and Nucleotide Sequence of the Corresponding Gene†

HEIN J. BOOT,‡* CARIN P. A. M. KOLEN,‡ JOHANNES M. VAN NOORT, AND PETER H. POUWELS

TNO Medical Biological Laboratory, 2280 HV Rijswijk, The Netherlands

Received 2 April 1993/Accepted 19 July 1993

The cell surfaces of several *Lactobacillus* species are covered by a regular layer composed of a single species of protein, the S-protein. The 43-kDa S-protein of the neotype strain *Lactobacillus acidophilus* ATCC 4356, which originated from the pharynx of a human, was purified. Antibodies generated against purified S-protein were used to screen a λ library containing chromosomal *L. acidophilus* ATCC 4356 DNA. Several phages showing expression of this S-protein in *Escherichia coli* were isolated. A 4.0-kb DNA fragment of one of those phages hybridized to a probe derived from an internal tryptic fragment of the S-protein. The *slpA* gene, coding for the surface layer protein, was located entirely on the 4.0-kb fragment as shown by deletion analysis. The nucleotide sequence of the *slpA* gene was determined and appeared to encode a protein of 444 amino acids. The first 24 amino acids resembled a putative secretion signal, giving rise to a mature S-protein of 420 amino acids (44.2 kDa). The predicted isoelectric point of 9.4 is remarkably high for an S-protein but is in agreement with the data obtained during purification. The expression of the entire S-protein or of large, C-terminally truncated S-proteins is unstable in *E. coli*.

Regular structures on the outsides of bacteria have been found in archaebacteria and in gram-negative and gram-positive eubacteria and are known as S-layers. Such S-layers are composed of one (glyco)protein known as S-protein that has a molecular mass of between 40 and 200 kDa. Surface layer proteins crystallize into a two-dimensional layer on the outside of the bacterial cell wall (for reviews, see references 3 and 29). Several reports have appeared in which functions of S-layers are described or assumed; examples of these functions are a protective sheath against hostile environmental agents (21), a cell shape determinant for archeabacteria (24), and a sheath to mask properties of the underlying cell wall, such as charge (27) and phage receptors (12).

For bacteria in their natural environment, the presence of an S-layer must provide a selective advantage because maintenance of the S-layer requires a large energy input; S-protein can constitute up to 10% of the total protein of cells in the exponential growth phase. Under artificial conditions, e.g., when bacteria are cultivated in laboratory media, S-layer-deficient mutants can outgrow the wild-type strain (15).

The intestinal tracts of animals and humans are inhabited by a variety of microorganisms, including species of *Lactobacillus*. Some *Lactobacillus* strains can adhere to and form a layer of bacteria on the epithelial cells of the gastrointestinal tract. Lactobacilli are thought to benefit the host by exerting adjuvant properties (9). *Lactobacillus acidophilus* strains, isolated from humans or animals, which belong to the DNA homology groups A are reported to possess an S-layer, while the strains which belong to the DNA homology groups B appear not to have an S-layer (13, 19, 20). The influence of the S-layer, which covers the underlying cell wall completely, on the adherence of *Lactobacillus* strains to epithelial cells of the intestinal tract is at present unclear. There are reports that the S-layer of *Aeromonas salmonicida* influences the adhesion of these bacteria to the epithelial cells of the salmon (7). An S-layer-deficient mutant of *A. salmonicida* is no longer virulent for salmon (12).

To gain more knowledge about the structure and function of S-layers and S-proteins of the genus *Lactobacillus*, we have started an investigation aimed at the characterization of S-protein of *L. acidophilus* ATCC 4356 and of the mode of expression of this protein. In this paper we describe the purification of S-protein of *L. acidophilus* ATCC 4356, the isolation and nucleotide sequence of the S-protein-encoding gene, and its expression in *Escherichia coli*.

MATERIALS AND METHODS

Strains. The neotype strain *L. acidophilus* ATCC 4356 was obtained from the American Type Culture Collection and was cultivated in ILS medium (8) under anaerobic conditions at 37°C. *E. coli* JM109 was used for all transformations with derivatives of pUC19 (35) or pBluescript II SK+ (Stratagene) vectors. *E. coli* NM538 and NM539 were used to grow wild-type and recombinant λ EMBL3 (10), respectively.

Isolation of S-protein. The S-layer protein of *L. acidophilus* ATCC 4356 was isolated after inoculation of an overnight culture into 500 ml of fresh medium and cultivated until the optical density at 695 nm (OD_{695}) had reached 0.7. Cells were harvested by centrifugation at 15,000 × *g* for 15 min at 4°C. The cells were washed twice with 500 ml of ice-cold water. Cells were resuspended in 10 ml of 4.0 M guanidine hydrochloride (pH 7.0), kept for 1 h at 37°C, and centrifuged at 18,000 × *g* for 15 min. The supernatant was dialyzed against water at 4°C. The dialyzed material was

* Corresponding author.

† This paper is dedicated with great appreciation to Dr. Frits Berends on the occasion of his retirement as Head of the Biochemistry Department of the TNO Medical Biological Laboratory.

‡ Present address: Department of Molecular Cell Biology, Section Molecular Biology, University of Amsterdam, Kruislaan 318, 1098 SM Amsterdam, The Netherlands.

freeze-dried and resuspended in 1.0 ml of 8.0 M urea–50 mM Tris (pH 7.5) (Q buffer).

Purification of S-protein. S-protein was purified from the crude cell extract by use of a cation-exchange column (Econo-Pac S cartridge; Bio-Rad). The crude extract was loaded onto the column, which had been equilibrated in Q buffer. S-protein was eluted with a gradient of 0.0 to 1.0 M NaCl in Q buffer at a flow rate of 2 ml/min. Fractions (5 ml) containing the S-protein were separately dialyzed against 50 mM Tris (pH 7.5) at 4°C, freeze-dried, dissolved in 50 mM Tris (pH 7.5), and stored at −20°C until further use.

Antibodies. Female BALB/c mice (12 to 14 weeks old) were immunized twice with a 4-week interval between immunizations. Purified S-protein (10 μg per dose) solubilized in phosphate-buffered saline (pH 7.2) was administered in a 200-μl volume intraperitoneally and was emulsified (11/9, vol/vol) in Freund complete adjuvant (Difco Laboratories) for priming and in Freund incomplete adjuvant (Difco Laboratories) for a booster immunization. Serum samples were collected 5 and 7 days after each immunization.

Amino acid analysis. Purified S-protein was hydrolyzed under HCl vapor at 112°C for 24 h and derivatized with phenylisothiocyanate for determination of the amino acid composition. The resulting phenylthiocarbamyl-amino acids were quantitated at 250 nm by reversed-phase high-performance liquid chromatography (HPLC) as previously described by Van Noort et al. (31). Tryptic peptide fragments of purified S-protein were generated by digestion of 200 μg of S-protein in 100 μl of 100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 7.4) with 2 μg of trypsin (*N*-tosyl-L-phenylalanine chloromethyl ketone treated; Boehringer Mannheim), with incubation at 37°C for 24 h. The reaction was stopped by addition of acetic acid to 50% (vol/vol). Preparative reversed-phase HPLC was performed by using a Superpak PepS column (Pharmacia/LKB) for the purification of the total tryptic digest. Peptide fragments were eluted by a linear gradient of 10 to 70% buffer B (70% aqueous acetonitrile, 0.1% trifluoroacetic acid) in buffer A (5% aqueous acetronitrile, 0.1% trifluoroacetic acid) during 40 min with a flow rate of 1.0 ml/min. The A_{214} of the effluent was monitored.

N-terminal amino acid sequences of purified S-protein and of four tryptic peptide fragments of S-protein were determined with an Applied Biosystems model 470A protein sequencer on-line equipped with a model 120A phenylthiohydantoin analyzer. Analyses were performed at the SON sequencing facility, State University Leiden.

Isolation of genomic DNA. ILS medium (100 ml) was inoculated with an overnight culture (5.0 ml) which was then cultivated to an OD_{695} of 1.0, centrifuged for 5 min at 4°C and 3,000 × *g*, resuspended in 10 ml of 20 mM Tris (pH 8.2), centrifuged for 15 min at 4°C and 3,000 × *g*, and resuspended in 2.5 ml of Tris (pH 8.2). Five milliliters of 24% polyethylene glycol 20000 and 2.5 ml of lysozyme (Boehringer Mannheim) solution (4 mg/ml) were added and mixed, and the mixture was incubated at 37°C for 60 min. After incubation, 5.0 ml of 0.2 M Na_2EDTA was added and mixed, and the mixture was centrifuged at 4°C and 3,000 × *g* for 15 min. The cells were resuspended in 10 ml of 20.0 mM Tris (pH 8.2) and lysed by addition of 1.5 ml of 9% Sarkosyl and 3.0 ml of 5.0 M NaCl. Chromosomal DNA (1.7 mg) was prepared from the lysate by ethanol precipitation, repeated phenol extractions, and RNase and proteinase K treatments as described by Lokman et al. (17).

Immunological detection of S-protein. The λ library of the genome of *L. acidophilus* ATCC 4356 was screened as described by Sambrook et al. (25) by using anti-SP(LA4356) as the first antiserum and anti-mouse immunoglobulin G–alkaline phosphatase conjugate (Promega) (0.2 μg/ml) as a second antiserum. Isolated recombinant phages were rescreened until a pure phage was obtained. S-protein of *L. acidophilus* ATCC 4356 or S-protein expressed in recombinant *E. coli* strains was detected by Western blotting (immunoblotting) with the same procedure.

Cloning procedures. Chromosomal DNA of *L. acidophilus* ATCC 4356 was partially digested with restriction endonuclease *Sau*3AI and separated on a 0.8% (wt/vol) low-melting-point agarose gel. Chromosomal DNA of 9 to 23 kb was recovered from the gel after β-agarase I (New England Biolabs) treatment according to the manufacturer's instructions and was ligated in λ EMBL3 which was previously digested with the restriction endonucleases *Bam*HI and *Eco*RI (10). The molar ligation ratio of λ EMBL3 arms to the chromosomal DNA fragment was 2:1 to minimize the chance of two chromosomal DNA fragments in one recombinant phage. The Gigapack II gold packaging kit (Stratagene) was used for in vitro packaging of recombinant phage DNA, yielding 5.0×10^5 recombinant phages. The 4.0-kb *Hin*dIII-*Sal*I fragment of phage SP-IV (see Fig. 2) was ligated in pUC19 and pBluescript, which were previously digested with the restriction endonucleases *Hin*dIII and *Sal*I and precipitated with isopropanol to remove the short linker fragment, yielding pBK-1 and pBK-2, respectively. pBK-1 and -2 were used to construct a subset of overlapping deletion clones by using the Erase-a-Base kit (Promega). pBK-1 was cut with restriction endonucleases *Kpn*I and *Xba*I; pBK-2 was cut with restriction endonucleases *Sst*I and *Hin*dIII (see Fig. 5). Religated deletion plasmids were transferred by electroporation to competent *E. coli* cells. Deletion plasmids from transformants were screened for length, as determined by agarose gel electrophoresis, and for the presence of the restriction endonuclease sites.

Southern blotting. Southern blotting of digested phage SP-IV DNA was performed essentially as described by Southern (30), with 33°C as the hybridization temperature and 6× SSC–0.1% sodium dodecyl sulfate (SDS) (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) as the washing buffer (25). Probe 3-c (5′-TGG CAA TAC GCT TAC TGA-3′) was designed according to the amino acid sequence SVSVLP and appeared to match the DNA sequence of the *slpA* gene exactly (see the underlined sequence in Fig. 6).

Nucleotide sequence analysis. Nucleotide sequencing was performed by the dideoxy chain termination method of Sanger et al. (26), partially by using a 373a DNA sequencer (Applied Biosystems) with a dye terminator kit and partially by using an S2 Bio-Rad gel electrophoresis system with the T7 DNA polymerase sequencing system (Promega) and [α-^{35}S]dATP (Amersham). The M13 universal and reverse primers (Boehringer Mannheim) and sequence-specific primers were used for the sequence reactions.

Nucleotide sequence accession number. The EMBL data library accession number of the DNA fragment shown in Fig. 6, which contains the *slpA* gene of *L. acidophilus* ATCC 4356, is X71412.

RESULTS

Purification of S-protein. Surface proteins of *L. acidophilus* ATCC 4356 were extracted by treatment of whole cells with 4 M guanidine hydrochloride. When this extract was analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), one dominant band of 43 kDa (which is known

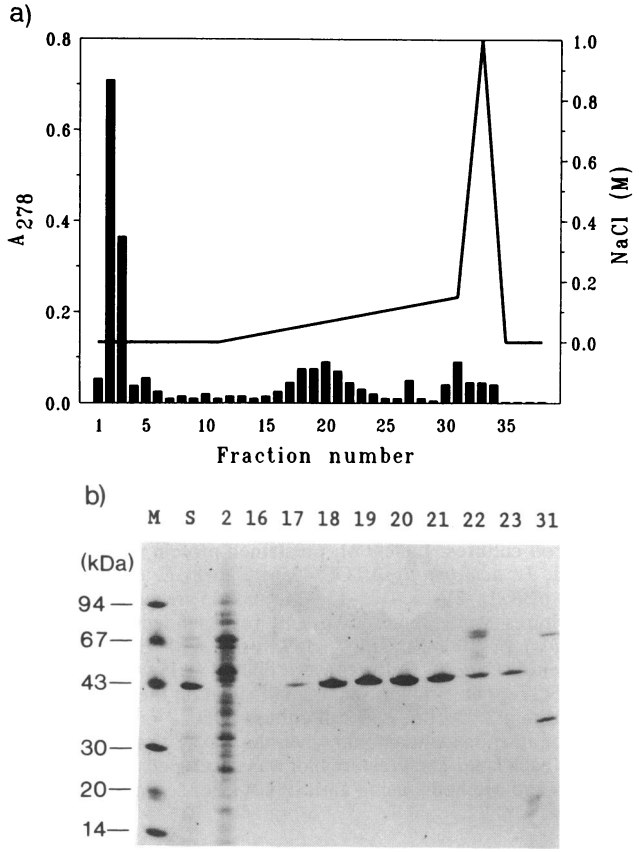

FIG. 1. Purification of the S-protein of *L. acidophilus* ATCC 4356. (a) Surface proteins of *L. acidophilus* ATCC 4356 were extracted as described in Materials and Methods. S-protein was purified from the extract with a cation-exchange column, which was eluted with a gradient of 0.0 to 1.0 M NaCl in Q buffer. The A_{278} of the different fractions was measured. The A_{278} for fractions 5 to 38 was multiplied by 5. The A_{278} of S-protein is rather low, as this protein contains no tryptophan (see Fig. 6). (b) SDS-PAGE gel (10 to 15% polyacrylamide) of fractions from the cation-exchange column, which were dialyzed against 50 mM Tris (pH 7.5). The amount of protein loaded in each lane was 0.1% of the amount of protein in the corresponding column fraction. The gel was stained with Coomassie blue. M, protein markers (Pharmacia); S, starting material. Lane numbers correspond to fraction numbers in panel a.

as the S-protein [20]) and a few faint bands were visible (Fig. 1b, lane S). The S-protein was purified by use of a cation-exchange column to homogeneity as shown by Coomassie blue staining (Fig. 1). The amino acid composition was determined; it closely resembles the amino acid composition of the S-protein of *L. acidophilus* ATCC 4357 (20) (Table 1). Highly purified S-protein of *L. acidophilus* ATCC 4356 was used to raise antibodies in mice as described in Materials and Methods. Purified S-protein was also used for determination of the N-terminal amino acid sequences of the whole protein and of tryptic fragments of the protein (see below).

Cloning of the *slpA* gene. Chromosomal DNA of *L. acidophilus* ATCC 4356 was used to make a library in λ EMBL3. Recombinant phages (5.0×10^4) were plated and screened with antiserum against S-protein, which yielded approximately 60 positive recombinant phages. Several phages were isolated and rescreened until pure phages were obtained. DNA was isolated from four phages (phages SP-I to SP-IV)

TABLE 1. Amino acid compositions of S-proteins of *L. acidophilus* strains

Amino acid	mol% in:		
	L. acidophilus ATCC 4357 (determined)[a]	*L. acidophilus* ATCC 4356	
		Determined[b]	Deduced[c]
Asp			4.76
Asn			12.14
Asx	16.4	17.27	
Glu			1.90
Gln			1.90
Glx	5.2	4.19	
Ser	6.7	8.05	7.86
Thr	10.8	12.16	12.86
Gly	6.4	6.27	5.48
Pro	2.2	2.19	1.90
Ala	13.3	13.01	13.33
Val	10.7	11.52	12.38
Cys	0.2	—[d]	0.00
Met	0.5	0.10	0.24
Ile	3.7	3.68	3.57
Leu	3.9	3.32	3.10
Tyr	6.2	3.93	5.48
Phe	2.3	2.62	2.14
Trp	—	—	0.00
His	0.9	0.77	0.71
Lys	8.8	9.10	8.57
Arg	1.8	1.80	1.67

[a] From reference 20.
[b] As described in Materials and Methods.
[c] Predicted mature S-protein as deduced from the DNA sequence of the *slpA* gene (Fig. 6).
[d] —, not determined.

which remained positive after successive rescreening. The insert length of *L. acidophilus* ATCC 4356 chromosomal DNA in those four recombinant phages varies from 13.6 to 15.5 kb. The orientations of the inserts are the same, and they all contain the same 12.1-kb fragment (Fig. 2). The DNA of phage SP-IV was digested with restriction endonuclease *Sal*I or *Hin*dIII and with a combination of both enzymes. A Southern blot of this DNA with a probe derived from a tryptic fragment of the S-protein is shown in Fig. 3. The 4.0-kb *Hin*dIII-*Sal*I fragment, which gave a positive signal in the Southern blot, was ligated in pUC19 and pBluescript, yielding pBK-1 and pBK-2, respectively. Protein samples obtained from recombinant *E. coli* strains containing these plasmids were separated by SDS-PAGE. A protein with a molecular weight slightly higher than that of S-protein from *L. acidophilus* ATCC 4356 was observed in a

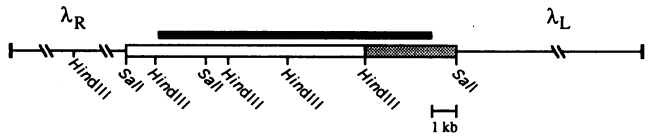

FIG. 2. DNA map of phage SP-IV. The black bar represents the 12.1-kb DNA fragment of *L. acidophilus* ATCC 4356, which was present in the same orientation in all four phages (phages SP-I to -IV). λ_R and λ_L represent the right and left arms of λ EMBL3, respectively. The open and stippled boxes represent DNA originating from *L. acidophilus* ATCC 4356. The stippled box represents DNA which gave a positive signal in a Southern blot with probe 3-c (see Fig. 3) and which was subcloned to yield plasmids pBK-1 and pBK-2.

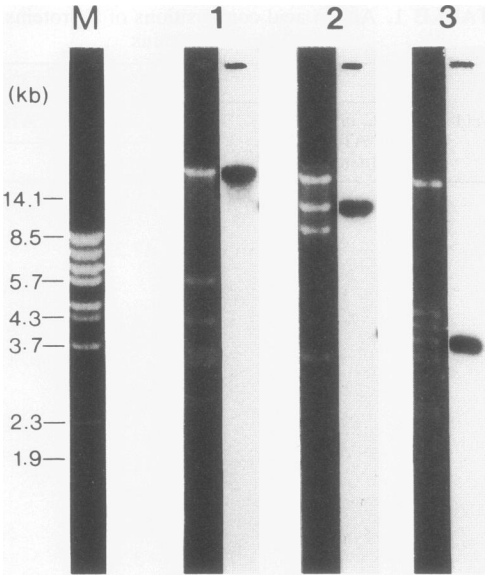

FIG. 3. DNA of phage SP-IV was digested with different restriction endonucleases, separated on a 1% agarose gel and transferred to nitrocellulose. Lanes 1 to 3, agarose gels with corresponding autoradiograms of Southern blots which were incubated with [γ-^{32}P]dATP-labelled probe 3-c. Lane M, marker DNA (wild-type λ digested with restriction endonuclease *Bst*EII). Phage SP-IV was digested with *Hin*dIII (lane 1), *Sal*I (lane 2), or *Hin*dIII and *Sal*I (lane 3).

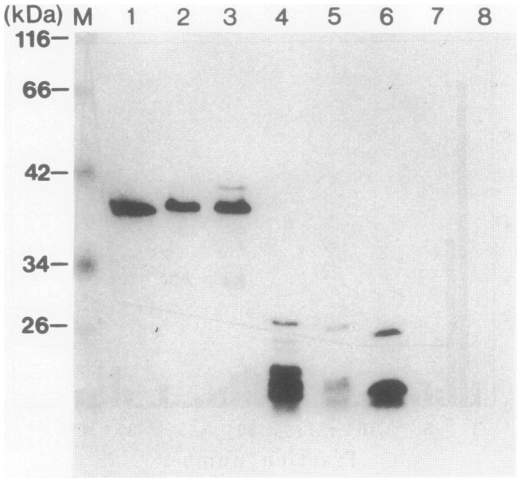

FIG. 4. Western blot of SDS-PAGE gel (12.5%, homogeneous) with protein samples from *L. acidophilus* ATCC 4356 and recombinant *E. coli* cultures. Lanes: M, prestained protein markers (Stratagene); 1, *L. acidophilus* ATCC 4356; 2 to 8, *E. coli* containing plasmids pBK-1, -2, -11, -12, -13, -14, and -15, respectively. Overnight cultures were spun down and resuspended in water (*L. acidophilus*) or 10 mM Tris (pH 7.8)–0.1 mM EDTA–0.1 mM dithiothreitol (*E. coli*). Cultures were diluted (1:1) with 2× sample buffer (25), kept at 95°C for 5 min, and loaded on the gel. Total cell protein from a 3.75-μl *E. coli* cell culture (OD_{695} = 1.0, 1-cm path length), or an equivalent amount when the OD_{695} was different, was loaded in each lane. The Western blot was developed as described in Materials and Methods, using anti-SP(LA4356) serum.

Western blot using anti-SP(LA4356) serum (compare lane 1 with lanes 2 and 3 of Fig. 4). Recombinant *E. coli* strains containing plasmid pBK-1 or -2 have a reduced growth rate (data not shown). The expression of S-protein by these recombinant *E. coli* cultures was frequently lost, indicating that these plasmids have a reduced stability.

Sequence analysis. Plasmids pBK-1 and -2 were used to create a set of overlapping clones by using the Erase-A-Base kit from Promega. By creating such a set, we could use universal oligonucleotides for sequence analysis. We succeeded only partially in creating an overlapping deletion set (Fig. 5). Protein samples from *E. coli* strains which contain deletion plasmids originating from pBK-1 were analyzed by SDS-PAGE followed by Western blotting with antiserum against S-protein of *L. acidophilus* ATCC 4356. Proteins of decreasing molecular weight were detected as the deleted DNA was increased (Fig. 4). By using this data, the position and orientation of the gene for the S-protein on plasmid pBK-1 could be deduced (Fig. 5).

The complete gene for the S-protein of *L. acidophilus* ATCC 4356, designated *slpA*, was sequenced in both directions, partially by using deletion plasmids and partially by using sequence-specific oligonucleotides when no deletion plasmids could be used (Fig. 5). The sequence revealed one large open reading frame (ORF), which consists of 1,332 nucleotides. A putative ribosome binding site, AGGAGG, is located 9 bases upstream of the ATG start codon of this ORF. The ORF ends with two adjacent stop codons (TAATAA) (Fig. 6). All of the more than 50 known *Lactobacillus* genes end with one stop codon, except for the S-protein genes of *L. acidophilus* ATCC 4356 and *Lactobacillus brevis* ATCC 8287 (32), both of which end with two stop codons (TAATAA). Having two adjacent stop codons may provide a stronger stop signal, necessary to diminish readthrough of these probably very efficiently transcribed genes. A Rho-independent transcription terminator (ΔG, −16.3 kcal [ca. −68.2 kJ]/mol) is found 25 bp downstream of these two stop codons by the method of Brendel and Trifonov (5). The stem of the hairpin consists of seven AU pairs, five GC pairs, and two GU pairs. The loop of the hairpin consists of four unpaired bases. The hairpin is followed by an AU-rich region.

The surface layer protein. The determination of the N-terminal amino acid sequence of the mature protein gave no positive result. Treatment of purified S-protein with trypsin generated several peptide fragments which were separated by reversed-phase HPLC. The N-terminal amino acid sequences of four purified peptides were determined; results for two peptide fragments gave unique sequences, which fully agree with parts of the amino acid sequence as deduced from the nucleotide sequence analysis of the *slpA* gene (Fig. 6, underlined sequences).

As deduced from the nucleotide sequence, the protein encoded by the *slpA* gene has in its N-terminal part a characteristic prokaryotic signal sequence, consisting of positively charged region (M-K-K-N-L-R) followed by a hydrophobic region of 13 amino acids (which is predicted to fold into an α-helix [6]) which ends with a proline. By the method of Von Heijne (33), the cleavage of the signal sequence is predicted to occur five amino acids after this proline (i.e., after serine 24). From this prediction, the mature S-protein consists of 420 amino acids, yielding a protein of 44.2 kDa, which is in good agreement with the molecular mass of 43 kDa estimated by SDS-PAGE (Fig. 1). The protein has a charge of +15, which yields an isoelectric point of 9.40 as predicted by computer analysis (MacVector). The amino acid compositions determined by amino acid

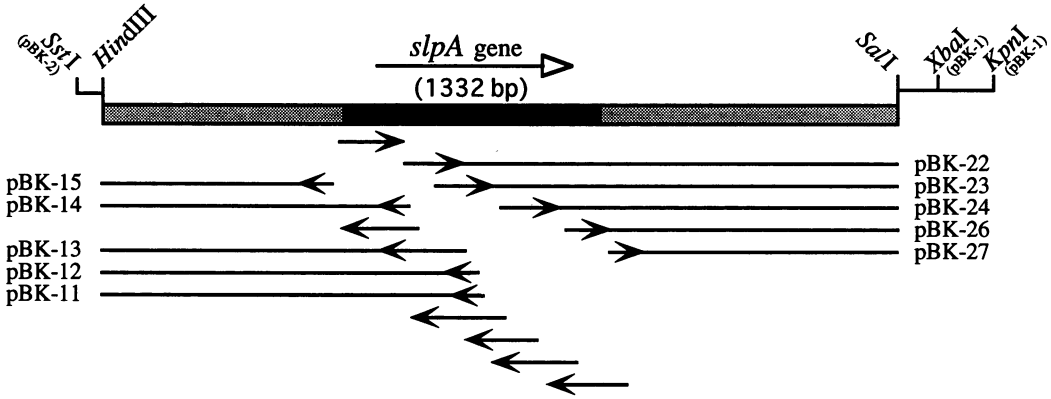

FIG. 5. Strategy for nucleotide sequence determination of the *slpA* gene of *L. acidophilus* ATCC 4356. The stippled box represents the same *Hin*dIII-*Sal*I DNA fragment as the stippled box in Fig. 2. The black box represents the 1,332-bp open reading frame of the *slpA* gene (see Fig. 6 for sequence), and the open arrow indicates the transcription direction. Arrowheads within lines indicate nucleotides determined with universal primers and deletion plasmids; arrows indicate nucleotides determined with sequence-specific primers and pBK-1 as the template.

analysis and predicted by the DNA sequence are given in Table 1.

DISCUSSION

Although the primary sequences of S-proteins are known for only a few bacteria, S-proteins display some common characteristic features. The hydrophobic amino acid content of these proteins is high (in most cases up to 45%), the hydroxy amino acid content is in most cases higher than 15%, and sulfur-containing amino acids are very rare or absent. For a long time it was believed that S-proteins are acidic proteins (11, 14, 23, 28), but a positively charged 45-kDa S-protein of *L. brevis* ATCC 8287 has recently been described (32). This 45-kDa S-protein has an isoelectric point of 9.37 (predicted with MacVector). In this article we describe the positively charged S-protein of *L. acidophilus* ATCC 4356, which has a predicted isoelectric point of 9.40. The high isoelectric point is in agreement with the data we obtained during purification of the S-protein of *L. acidophilus* ATCC 4356. The protein binds to a cation-exchange column at pH 7.5 when solubilized in 8 M urea–50 mM Tris. The S-proteins of *L. brevis* ATCC 8287 and *L. acidophilus* ATCC 4356 are the only *Lactobacillus* S-proteins for which the DNA sequence is currently known. Whether other *Lactobacillus* strains or strains from other genera harbor a positively charged S-protein needs to be established. A fast and convenient method for purifying a positively charged S-protein is described in Materials and Methods.

A genomic library of DNA of *L. acidophilus* ATCC 4356 was made in λ EMBL3. S-protein-encoding fragments could be subcloned on multicopy plasmids. The presence of the *slpA* gene on recombinant phages or plasmids could be assessed by immunoscreening using anti-SP(LA4356) serum. A protein of the expected size was observed, indicating that the entire *slpA* gene had been cloned. The promoter of the *slpA* gene must be active in *E. coli*, as no promoter is present for the transcription of inserted DNA in λ EMBL3 and no induction of the *lac* promoter with isopropyl-β-D-thiogalactopyranoside (IPTG) was used in pBK-1 and -2. Derivatives of pBK-1 containing deletions of increasing size expressed truncated S-proteins of decreasing size, suggesting that the deletions went from the C-terminal end towards the N-terminal end of S-protein. This result was confirmed by nucleotide sequence analysis.

For DNA sequence analysis a set of overlapping clones was constructed. The preparation of this set of clones succeeded only partially. The earlier time points of the exonuclease reaction of plasmid pBK-1 yielded mainly plasmids containing small deletions, as was revealed by agarose gel electrophoresis (data not shown). After religation and electroporation of competent *E. coli* cells, transformants containing plasmids with only larger deletions were obtained. At later time points the sizes of the plasmids isolated from *E. coli* transformants corresponded to those of the plasmid DNA molecules after exonuclease treatment. Transformants which gave a positive signal in an immunoscreening appeared to have a plasmid encoding a truncated S-protein of approximately 25 kDa or less, as was concluded from a Western blot containing protein samples from those transformants. Since nucleotides are removed from the S-protein-encoding strand in the 3′-to-5′ direction in pBK-1, the truncated proteins represent the N-terminal part of the S-protein. Apparently, *E. coli* transformants containing plasmids with larger deletions, leading to the expression of truncated S-protein smaller than 25 kDa, are preferentially found if the ligation mixture contains deletion plasmids of different sizes. A similar phenomenon was observed with the deletion clones made from plasmid pBK-2. In this plasmid the S-protein-encoding strand is degraded in the 5′-to-3′ direction. The smallest deletion plasmids that could be recovered from *E. coli* transformants contained a deletion which extended into the N-terminal region of S-protein. The selection against deletion plasmids which code for large (>25-kDa) S-protein or the entire (pre-)S-protein is probably a result of the expression of those (pre-)S-proteins in *E. coli* transformants.

In several studies involving the cloning of S-protein genes from different bacterial species, structural instability or even inviability of *E. coli* transformants with S-protein genes has been observed (these species include *Acetogenium kivui* DSM 2030 [22], *A. salmonicida* [2], *Bacillus brevis* 47 [34], *Bacillus sphaericus* 2362 [4], *Halobacterium halobium* [16], and *L. brevis* ATCC 8287 [32]). An explanation for the difficulties in cloning DNA encoding (pre-)S-proteins in *E. coli* on multicopy plasmids might involve the efficiency with

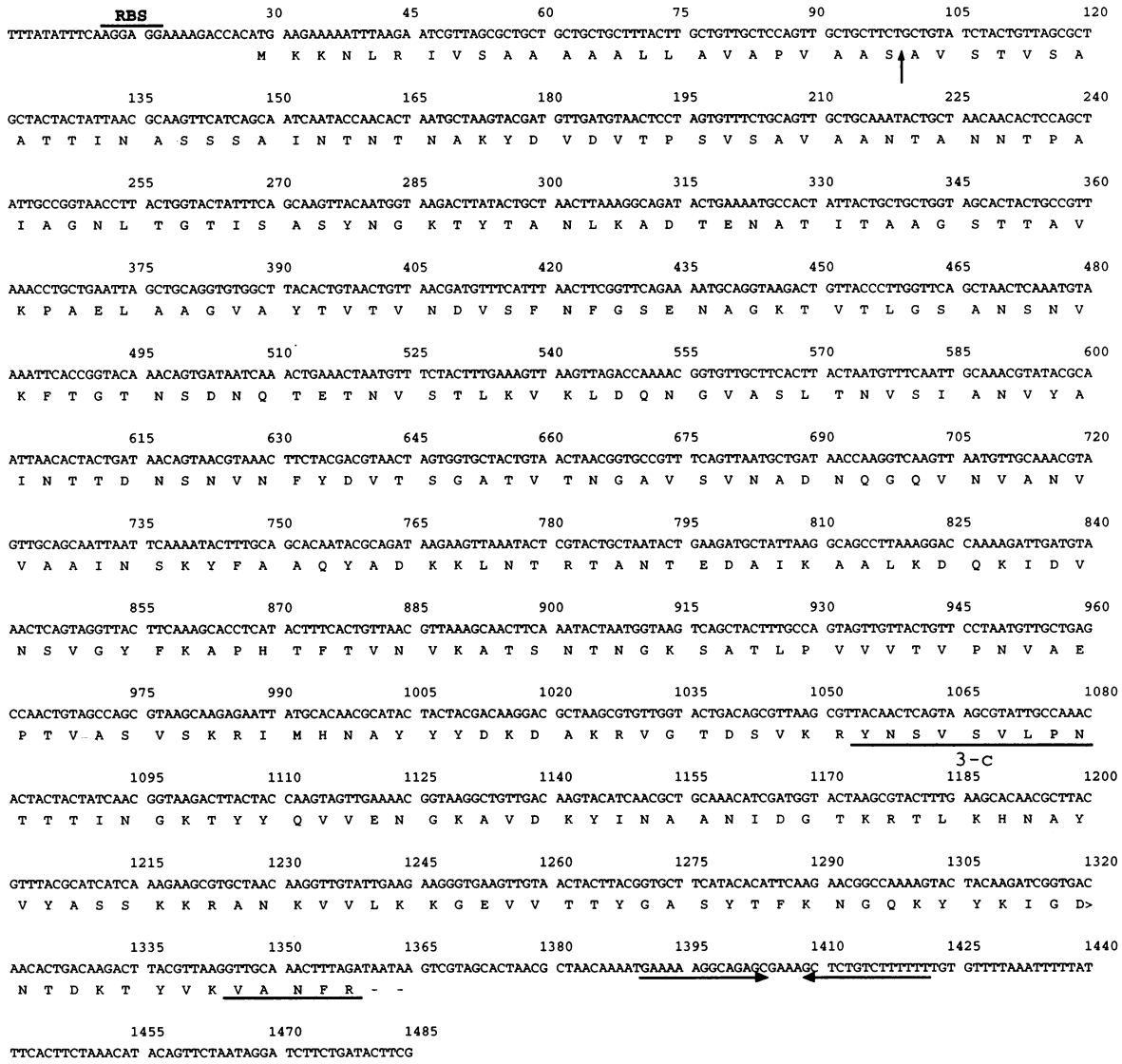

FIG. 6. DNA sequence of the *slpA* gene of *L. acidophilus* ATCC 4356 and deduced amino acid sequence. The N-terminal amino acid sequences of two tryptic peptides are underlined, and the ribosome binding site (RBS) is indicated. The predicted cleavage site after Ser-24 is indicated with a vertical arrow. Horizontal arrows indicate an inverted repeat.

which these proteins are normally produced. Efficient production of such S-proteins in *E. coli*, if not accompanied by correspondingly efficient processing (including degradation) or secretion, might result in a reduced growth rate or even in lethality.

The S-protein of *L. acidophilus* ATCC 4356 expressed in *E. coli* carrying pBK-1 or -2 is slightly larger than the native S-protein of *L. acidophilus* ATCC 4356 (Fig. 4). This difference in molecular mass might be due to a lack of or incorrect cleavage of pre-S-protein in *E. coli*. The predicted leader has a molecular mass of 2.3 kDa, which means that a protein of 46.6 kDa would result if cleavage of the pre-S-protein did not occur. The truncated (pre-)S-proteins which are encoded by deletion plasmids pBK-11 to -13 are less stable than the entire (pre-)S-protein in *E. coli*. This can be concluded from Fig. 4, since hardly any degradation products can be found in the cases of pBK-1 and -2 but degradation products are abundant in the cases of pBK-11 to -13. A less stable protein may be less deleterious for *E. coli*, thus providing a selective advantage for *E. coli* transformants containing larger deletions, as was observed. Although a small N-terminal part (85 amino acids) of the (pre-)S-protein is encoded by pBK-14 and is likely to be expressed by the *E. coli* transformant containing this plasmid, no polypeptide or degradation product could be found in a Western blot using anti-SP(LA4356) serum (Fig. 4). Whether the absence of a positive signal is due to rapid degradation of this small polypeptide or whether this small polypeptide is not recognized by the polyclonal anti-SP(LA4356) serum is at present unclear.

The first 23 amino acids of the S-protein of *Lactobacillus helveticus* ATCC 12046 have been determined (18) and appear to be nearly identical to those of *L. acidophilus* ATCC 4356 (Fig. 7). Homology searches comparing the S-protein of *L. acidophilus* ATCC 4356 with a protein library (SWISSPROT) revealed similarities over the entire 420 amino acids with the 125-kDa S-protein (42.5%) and a

```
     1                                31
(LA) AVSTVSAATTINASSSAINTNTNAKYDVDVT
           |||||| |||| ||||||||||||
(LH)       ATTINAD-SAINANTNAKYDVDVT
           1                      23
```

FIG. 7. N-terminal amino acid sequences of S-proteins of *L. acidophilus* ATCC 4356 (LA) and *L. helveticus* ATCC 12046 (LH). The N-terminal sequence of the mature S-protein of *L. acidophilus* ATCC 4356 was predicted by the method of Von Heijne (33) from the DNA sequence of the *slpA* gene as presented in Fig. 6. The N-terminal sequence of the S-protein of *L. helveticus* ATCC 12046 is based on N-terminal amino acid analysis (23 amino acids) and was published by Lortal et al. (18). Numbers indicate the amino acid residue numbers, counted from the N-terminal side. Vertical lines indicate identical amino acids, dots indicate conserved replacements, and the horizontal line indicates a gap.

cryptic 80-kDa S-protein (38.3%) of *B. sphaericus* 2362 (4) and with the S-protein of *L. brevis* ATCC 8287 (35.7%) (32). The similarity between the S-proteins of *L. acidophilus* ATCC 4356 and *L. helveticus* ATCC 12046 is probably much greater, as the N termini of these S-proteins are nearly identical. We are currently screening several *Lactobacillus* species with the anti-SP(LA4356) serum to see whether related *Lactobacillus* species possess related S-proteins. S-proteins are expressed at a high level. For *B. brevis* 47 it was shown that three promoters probably are used for the production of S-protein mRNA. These promoters are used at different efficiencies during different growth phases (1, 34). We are presently elucidating the DNA sequence upstream of the *slpA* gene of *L. acidophilus* ATCC 4356. Further studies will be aimed at determination of the transcription initiation site(s) and control of gene expression during different growth phases.

ACKNOWLEDGMENTS

We gratefully acknowledge Jacqueline Boon for technical assistance, Marc Schellekens for assistance in the preparation of the anti-SP(LA4356) serum, Mariska van Giesen for preparing oligonucleotides, and Michel Boermans for excellent photographic work. We thank Mark Posno and Rob Leer for fruitful and stimulating discussions.

REFERENCES

1. **Adachi, T., H. Yamagata, N. Tsukagoshi, and S. Udaka.** 1989. Multiple and tandemly arranged promoters of the cell wall protein gene operon in *Bacillus brevis* 47. J. Bacteriol. **171:** 1010–1016.
2. **Belland, R. J., and T. J. Trust.** 1987. Cloning of the gene for the surface array protein of *Aeromonas salmonicida* and evidence linking loss of expression with genetic deletion. J. Bacteriol. **169:**4086–4091.
3. **Beveridge, T. J., and L. L. Graham.** 1991. Surface layers of bacteria. Microbiol. Rev. **55:**684–705.
4. **Bowditch, R. D., P. Baumann, and A. A. Yousten.** 1989. Cloning and sequencing of the gene encoding a 125-kilodalton surface-layer protein from *Bacillus sphaericus* 2362 and of a related cryptic gene. J. Bacteriol. **171:**4178–4188.
5. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res. **12:**4411–4427.
6. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. **47:**45–147.
7. **Doig, P., L. Emödy, and T. J. Trust.** 1992. Binding of laminin and fibronectin by the trypsin-resistant major structural domain of the crystalline virulence surface array protein of *Aeromonas salmonicida*. J. Biol. Chem. **267:**43–49.
8. **Efthymiou, C., and P. A. Hansen.** 1962. An antigenic analysis of *Lactobacillus acidophilus*. J. Infect. Dis. **110:**258–267.
9. **Fernandes, C. F., K. M. Shahani, and M. A. Amer.** 1987. Therapeutic role of dietary lactobacilli and lactobacillic fermented dairy products. FEMS Microbiol. Rev. **46:**343–356.
10. **Frischauf, A.-M., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:**827–842.
11. **Hagiya, H., T. Oka, H. Tsuji, and K. Takumi.** 1992. The S layer composed of two different protein subunits from *Clostridium difficile* GAI 1152: a simple purification method and characterization. J. Gen. Appl. Microbiol. **38:**63–74.
12. **Ishiguro, E. E., W. W. Kay, T. Ainsworth, J. B. Chamberlain, R. A. Austen, J. T. Buckley, and T. J. Trust.** 1981. Loss of virulence during culture of *Aeromonas salmonicida* at high temperature. J. Bacteriol. **148:**333–340.
13. **Johnson, M. C., B. Ray, and T. Bhowmik.** 1987. Selection of *Lactobacillus acidophilus* strains for use in "acidophilus products." Antonie van Leeuwenhoek **53:**215–231.
14. **Koval, S. F., and R. G. E. Murray.** 1984. The isolation of surface array proteins from bacteria. Can. J. Biochem. Cell Biol. **62:**1181–1189.
15. **Koval, S. F., and R. G. E. Murray.** 1986. The superficial protein arrays on bacteria. Microbiol Sci. **3:**357–361.
16. **Lechner, J., and M. Sumper.** 1987. The primary structure of a procaryotic glycoprotein. J. Biol. Chem. **262:**9724–9729.
17. **Lokman, B. C., P. van Santen, J. C. Verdoes, J. Krüse, R. J. Leer, M. Posno, and P. H. Pouwels.** 1991. Organization and characterization of three genes involved in D-xylose catabolism in *Lactobacillus pentosus*. Mol. Gen. Genet. **230:**161–169.
18. **Lortal, S., J. van Heijenoort, K. Gruber, and U. B. Sleytr.** 1992. S-layer of *Lactobacillus helveticus* ATCC 12046: isolation, chemical characterization and re-formation after extraction with lithium chloride. J. Gen. Microbiol. **138:**611–618.
19. **Masuda, K.** 1992. Heterogeneity of S-layer proteins of *Lactobacillus acidophilus* strains. Microbiol. Immunol. **36:**297–301.
20. **Masuda, K., and T. Kawata.** 1983. Distribution and chemical characterization of regular arrays in the cell walls of strains of the genus *Lactobacillus*. FEMS Microbiol. Lett. **20:**145–150.
21. **Munn, C. B., E. E. Ishiguro, W. W. Kay, and T. J. Trust.** 1982. Role of surface components in serum resistance of virulent *Aeromonas salmonicida*. Infect. Immun. **36:**1069–1075.
22. **Peters, J., M. Peters, F. Lottspeich, and W. Baumeister.** 1989. S-layer protein gene of *Acetogenium kivui*: cloning and expression in *Escherichia coli* and determination of the nucleotide sequence. J. Bacteriol. **171:**6307–6315.
23. **Peters, J., M. Peters, F. Lottspeich, W. Schäfer, and W. Baumeister.** 1987. Nucleotide sequence analysis of the gene encoding the *Deinococcus radiodurans* surface protein, derived amino acid sequence, and complementary protein chemical studies. J. Bacteriol. **169:**5216–5223.
24. **Pum, D., P. Messner, and U. B. Sleytr.** 1991. Role of the S layer in morphogenesis and cell division of the archaebacterium *Methanocopusculum sinense*. J. Bacteriol. **173:**6865–6873.
25. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
27. **Sára, M., and U. B. Sleytr.** 1987. Molecular sieving through S layers of *Bacillus stearothermophilus* strains. J. Bacteriol. **169:** 4092–4098.
28. **Sleytr, U. B., and P. Messner.** 1983. Crystalline surface layers on bacteria. Annu. Rev. Microbiol. **37:**311–339.
29. **Sleytr, U. B., and P. Messner.** 1988. Crystalline surface layers in procaryotes. J. Bacteriol. **170:**2891–2897.
30. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
31. **Van Noort, J. M., J. Boon, A. C. M. van der Drift, A. M. H. Boots, J. P. A. Wagenaar, and C. J. P. Boog.** 1991. Processing by endosomal proteases determines which parts of sperm whale

myoglobin are eventually recognized by T cells. Eur. J. Immunol. **21**:1989–1996.

32. **Vidgrén, G., I. Palva, R. Pakkanen, K. Lounatmaa, and A. Palva.** 1992. S-layer protein gene of *Lactobacillus brevis*: cloning by polymerase chain reaction and determination of the nucleotide sequence. J. Bacteriol. **174**:7419–7427.
33. **Von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
34. **Yamagata, H., T. Adachi, A. Tsubio, M. Takao, T. Sasaki, N. Tsukagoshi, and S. Udaka.** 1987. Cloning and characterization of the 5′ region of the cell wall protein gene operon in *Bacillus brevis* 47. J. Bacteriol. **169**:1239–1245.
35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.