Heterogeneity of Putative Surface Layer Proteins in Lactobacillus helveticus

Monica Gatti,¹* Lia Rossetti,² Maria Emanuela Fornasari,² Camilla Lazzi,¹ Giorgio Giraffa,² and Erasmo Neviani¹

Department of Genetic Anthropology Evolution, University of Parma, 43100 Parma, Italy,¹ and Istituto Sperimentale Lattiero Caseario, 26900 Lodi, Italy²

Received 29 October 2004/Accepted 8 July 2005

The S-layer-encoding genes of 21 *Lactobacillus helveticus* strains were characterized. Phylogenetic analysis based on the identified S-layer genes revealed two main clusters, one which includes a sequence similar to that of the *slpH1* gene of *L. helveticus* CNRZ 892 and a second cluster which includes genes similar to that of *prtY*. These results were further confirmed by Southern blot hybridization. This study demonstrates S-layer gene variability in the species *L. helveticus*.

Lactobacillus helveticus is a homofermentative thermophilic lactic acid microorganism widely used in the manufacture of Italian cheese (15, 20, 21). Large differences in the metabolic activity and technological performance of *L. helveticus* have often been reported (8, 10, 22). The presence of different *L.* helveticus biotypes was shown in dairy starter cultures through a variety of genotypic and phenotypic methods (13, 16, 17, 19, 23). It has often been suggested that the selection and prevalence of dominant *L. helveticus* subpopulations could be linked to their environmental origin and affect the different technological conditions of milk processing (12, 14, 18). Surface protein patterns allowed us to distinguish at least six *L. helveticus* strains, which were, to some extent, specific to the source of isolation (11–13).

S-layers are two-dimensional crystalline arrays formed by proteinaceous subunits that cover the outer surfaces of many unicellular organisms (3) and represent 10 to 20% of the total cell protein content (6, 20, 27). Thus far the function of *Lactobacillus* S-layer proteins is largely unknown. However, in *Lactobacillus crispatus* the S-layer protein is involved in adhesion to collagen (29). Numerous genes encoding S-layer proteins having quite different taxonomic positions have been cloned and sequenced from different organisms (26). Conversely, the amino acid sequences of only a few *Lactobacillus* S-layer proteins are available (4–7, 26, 29). While two S-protein-encoding genes have been observed (5) in the genomes of a few *Lactobacillus* species (e.g., *L. acidophilus*), to date there is still no evidence of S-layer gene variability in the species *L. helveticus*.

The aim of this work is to investigate the variability of the S-layer gene from different *L. helveticus* biotypes.

Primer design and PCR amplification. Different pairs of primers were designed to amplify the S-layer gene from 21 *L. helveticus* strains used in this study. Six strains (Lh4, Lh5, Lh12, Lh23, Lh28, and FBT9) were isolated from Grana Padano natural whey starters; six strains (SF2, SiS14, SiM3, SiM7, SiO7, and SiS18) were isolated from Parmigiano Reg-

giano natural whey starters; seven strains (CBT17, DBT16, ABT5, Lh52, Lh59, Lh64, and Lh70) were isolated from Provolone natural whey starters; and two strains (IBT5 and V2BT10) were isolated from semihard Italian cheeses. *L. helveticus* CNRZ 892 was included as the reference strain because of its well-established ability to express an S-layer protein (7). These strains have already been identified in previous investigations (12, 18, 19).

The whole *slpH1* gene of the 21 strains was amplified, using primers 7 FOR and 8 REV (or 6 REV), as reported by Callegari et al. (7) (NCBI accession number X91199), for all strains except Lh4, for which P16 instead of P7 was used as a forward primer (Table 1). Total DNA from different lactobacilli was extracted from 5-ml samples of fresh overnight MRS broth cultures according to an alkaline lysis method used by de los Reyes-Gavilàn et al. (9) The quantity and purity of DNA were assessed by reading optical density at 260 and 280 nm, as described by Sambrook et al. (25).

The amplification products were observed for all the strains, although the molecular size was not always as expected (data not shown). For 14 strains, PCR amplicons were obtained using primers P7 FOR and P6 REV, and for 7 strains, P8 REV was used as the reverse primer instead of P6 REV (Table 1).

To study sequence polymorphisms in the *slpH1* gene, we first used pairs of primers designed for the reported sequence of

TABLE 1. Sequences of the primers used in this study

Primer	Sequence $(5' \rightarrow 3')$	References
P16 FOR	GGAGGAAAGACCACATGAAG	
7 FOR	CATTATAGGCTCCTTTCTCATG	7, 33
P469 REV	GGGTTACGTGAACCAATAGT	,
P469a REV	TGACTCTTAGTGGTGTAGGC	
P435 FOR	GCAGGTAAGGAAATTACTAT	
P435a FOR	GCAGGTGATGACTACACTAT	
P435b FOR	GCAGGTAAGGAATACACTAT	
P435c FOR	GCAGGTAAGAAATACACTAT	
P831 REV	GGAGCCTTGAAGTAACCTAC	
P850a REV	CCACCTTGATACCAGCCTTT	
P806 FOR	GGCTCAAAAGATTGAAGTAA	
8 REV	GCCTTGCCGTTTTCGATTAC	7, 33
6 REV	TTATTCAAAGTTAGCAACCTTAAC	7, 33

^{*} Corresponding author. Mailing address: Department of Genetic Anthropology Evolution, Parco Area delle Scienze 11A, University of Parma, 43100 Parma, Italy. Phone: 390521905479. Fax: 390521905604. E-mail: monica.gatti@unipr.it.

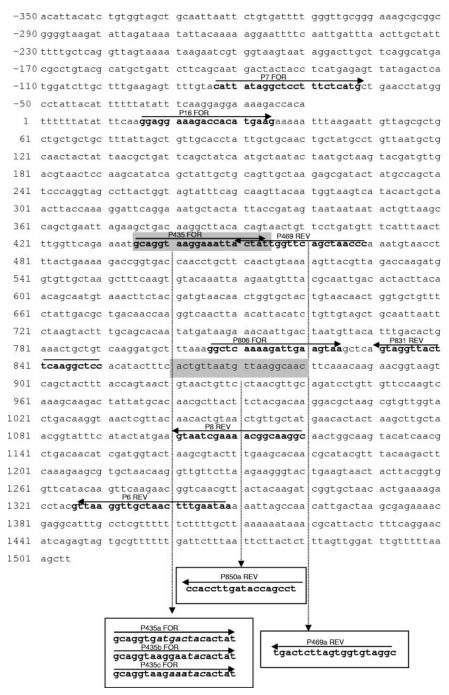


FIG. 1. Primers designed according to the sequence of the *slpH1* gene (NCBI accession number X91199). Primers designed after a gene-walking strategy are in gray boxes.

the gene (Fig. 1). The 5' end (approximately 550 bp in size) was amplified from 13 strains, using the primer pair P7 FOR (or P16 FOR for Lh4)-P469 REV, and for the other 9 strains, the primer pair P7 FOR-P469a REV was used. The 3' end (approximately 550 or 320 bp) was amplified from all the strains, using the primer pair P806 FOR-P6 REV (or P8 REV for Lh4). To obtain the complete sequence, internal fragments (approximately 420 bp) were obtained for only 13 strains, using the primer pair P435 FOR-P831 REV, and for the remaining

9 strains, a gene-walking strategy was necessary. Primer walking allowed us to design new primer pairs, P435a FOR-P850a REV for strain CBT17; P435b FOR-P850a REV for strains V2BT10, SF2, Lh28, and SIM7; and P435c FOR-P850a REV for strains IBT5, SIM3, SIO7, and DBT16 (Fig. 1).

The reaction mixtures contained 50 ng template DNA, 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 1.25 U Ampli*Taq* DNA polymerase (Applera Italia, Monza, Italy), 25 mM MgCl₂, 0.2 mM deoxynucleoside triphosphate, and 0.4 μ M

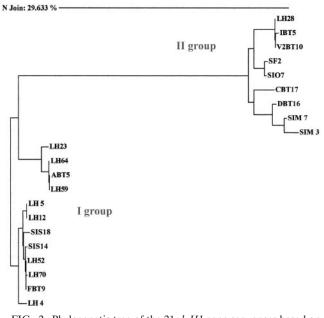


FIG. 2. Phylogenetic tree of the 21 *slpH1* gene sequences based on the neighbor-joining method. The sequence alignment, distance matrix, and phylogenetic tree were calculated with Sequence Navigator software (version 1.0.1; Applied Biosystems). Groups I and II include sequences highly homologous to the published *slpH1* (NCBI accession number X91199) and *prt* (NCBI accession number AB026985) gene sequences, respectively.

each primer (Biotez, Berlin, Germany) in a final reaction volume of 50 μ l. The amplification reactions were performed in a 9700 Perkin Elmer thermal cycler (Applera Italia) with the following conditions: after one cycle at 94°C for 2 min, 25 cycles of 95°C for 60 s, 55°C for 60 s, and 72°C for 2 min were performed. A final extension was carried out at 72°C for 7 min.

Determination of DNA sequences. Amplified products were purified with Microcon-PCR filter units (Millipore, Milan, Italy) according to the manufacturer's recommendations. Forward and reverse sequencing reactions were performed for each amplified product according to the manufacturer's instructions. Nucleotide sequencing on both strands of the PCR amplicons was performed, using an ABI PRISM 310 sequencing machine (Applied Biosystems, Foster City, CA). The sequencing data assembly, analysis, and alignment were performed with Sequence Navigator software (version 1.0.1) and MicroSeq software (version 1.36), both from Applied Biosystems (Foster City, CA).

The sequences of the hypothetical S-layer *slpH1* genes were highly variable among the *L. helveticus* strains. A phylogenetic analysis based on the 21 sequences revealed two main groups of S-layer genes (Fig. 2). The sequence similarity between the two groups was very low. Group I included S-layer genes obtained using primers based on the published *slpH1* gene sequence. Within this group, two subgroups of very similar sequences were detected. After a BLASTN search (http://www .ncbi.nlm.nih.gov/BLAST/), the sequences of group I showed a high level of similarity (>94%) with those of the published *L. helveticus slpH1* gene (7). The remaining S-layer sequences were included in a second group (group II), which exhibited a

	5'		3'
strains	1bp 209bp	210bp 799bp	800 bp 1325 bp
X91199			
ABT5	99%	88%	90%
FBT9	99%	99 %	93%
Lh4	99%	99%	89%
Lh 5	99%	99%	93%
Lh 12	99%	99%	93%
Lh 23	100%	89%	90%
Lh 52	99%	99%	93%
Lh 59	99%	88%	90%
Lh 64	99%	88%	90%
Lh 70	96%	99%	93%
SIS14	99%	99%	93%
SIS18	99%	98%	93%

В

A

	5'		3'
strains	1bp 120bp	121bp 850bp	851bp 1180bp
AB026985			
CBT17	97%	83%	87%
DBT16	97%	79%	87%
IBT 5	97%	81%	96%
Lh28	97%	81%	87%
SF2	97%	77%	97%
SIM3	95%	81%	96%
SIM7	95%	80%	96%
SIO7	97%	78%	97%
V2BT10	97%	81%	97%

FIG. 3. Comparison between group I sequences and the *slpH1* gene sequence (A) and group II sequences and the *prt* gene for proteinase (B) made using MicroSeq software (version 1.36; Applied Biosystems). Similarity, based on the neighbor-joining method, is indicated by shading.

FIG. 4. Multiple sequence alignment of deduced amino acid sequences. The *L. helveticus* sequence for SlpH1 (NCBI protein identification number CAA62606.1) and the *L. helveticus* CP790 proteinase sequence (NCBI protein identification number BAA86287) are reported as references. Asterisks denote residues conserved among all sequences, and dots represent gaps introduced into a sequence for alignment. The highly conserved NH₂- and COOH-terminal regions are shown in boxes. The double underline denotes the prosequence within the proteinase-like sequences. The roman numerals in parentheses following the strain names refer to the phylogenetic tree (I or II). The black- and white-filled arrows indicate the signal peptide and prosequence cleavage sites, respectively.

FIG. 4. Mi	ABT5 LH23 FBT9 LH70	ABT5 LH23 FBT9 LH70 LH70 LH70 LH70 LH70 LH70 LH70 LH70	ABT5 LH23 FBT9 LH70 LH70 LH70 LH70 LH70 LH70 LH12 LH12 LH12 LH12 LH12 LH12 LH4 DBT16 SIN3 SIO7 LH2 LH4 P roteinase	ABT5(1) LH23(1) FBT9(1) LH27(1) LH52(1) S-Layer S1518(1) LH5(1) LH5(1) LH12(1) DBT16(11) SIM3(11) SIM3(11) Profeinase
Multiple sequence alignment of deduced amino acid sequences. The L. helveticus sequence for SlpH1 (NCBI protein identification number CAA62606.1) and the	KD GRPVTGPTAHVKLDQNGVATLSSVQIKNVYAVNTTDNRDVNPYDVTTGATVTTGATVXTGAVSI 213 GTGSTVKVFLDQDGVAKLSSVFLKUVYAVNTDNRDVNPYDVTTGATVTTGAVSI 208 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 211 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQDGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQGGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 212 GDQPASTVKVTLDQGGVAKLSSVFLKUVYALDTTYNSNVNPYDVTTGATVTTGAVSI 213 GANSNDKVLLDQGGNATULKEVISINAVDSANTNAVSFYDKKSGLVATQGSYMT 224 DTSNDKVLLDQGGNATULKEVISINAVDSANTNAVSFYDKKSGLVATQGSYMT 223 GANYNGGKVTDLKFTLK-VKAYDSANTNAVSFYDKKSGLVATQGSYMT 223 GANYNGGKVTDLKFTLK-VKAYDSANTNAVSFYDKKSGLVATQGSYMT 223 GANYNGGKVTDLKFTLK-VKAYDSANTNAVSFYDKKSGLVATGGSYMT 223 GANYNGGKVTDLKFTLK-VKAYDSANTNAVSFYDKKSGLVATGGSYMT 223	TVKPADLEYGKPYKVAVPDVSENEGSENAGKEITIGSANQNUTETTKD 153 TVKPADLEYGKPYKVTVPDVSENEGSENAGKEITIGSANQNUTETEKT 155 TVKPAELEAGKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVKPAELEADKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVKPAELEADKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVKPAELEADKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVKPAELEADKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVKPAELEADKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVKPAELEADKAYTVTVPDVSENEGSENAGKEITIGSANENVTETEKT 155 TVK	YDVDVTPSISALAAKVIGTG-ITIPGSLTGSISASYNGKSYTANLPKDSENATITNSNNE 105 YDVDVTPSISALAAVAGDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITNSNNE 105 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISALAAVAKSDTMPAIPGSLTGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISTLTSALPAKGVPAITGGSIFGSISASYNGKSYTANLPKDSENATITDSNNN 107 YDVDVTPSISTLTSALPAKGVPAITGGSIFGSISASYNGKSYTANLPKDSENATITDSNNN 104 DTVMVTPSFTLTSALPAKGVPAITGGSIFGSISASYNGKSYTANLPKDSENATITDSNNN 104 DTVMVTPSFTLTSALPAKGVPAITGGSIFGSISASUNGTSVTADVADVAKUVTLKDDGGK 113 DTVMVTPSFTLTSALPAKGVPAITGGSIFGSISASUNGTSVTADVADVAKUVTLFDGNK 112 DTVMVTPSFTLTSALPAKGVPAITGGSIFASINGTSVTADVADVAKUVTLFDGKK 112 DTVMVTPSFTLTSALPAKGVPAITGGSIFASINGTSVTADVADVAKUVTLFDGKK 112 DTVMVTPSFTLTSALPAKGVPAITGGSIFASINGTSVTADVADVAKUVTLFDGKK 112 DTVMVTPSFTLTSALPAKGVPAITGGSIFASINGTSVTADVADVAKUVTLFDGKK 112	WERKULE IVSAAAAALLAVAE VAATAMEVVAATTI INADSAINTNINNAK 47 MERKULE IVSAAAAALLAVAE IAATAMEVVAATTI INADSAINTNINNAK 47 MERKULE IVSVAAAALLAVAE VAATAMEVVAATTI INADSAINTNINNAK 47 MERKULE IVSVAAAALLAVAE VAATAMEVVAATTI INADSAINANINNAK 47 MERKULE IVSVAAAALLAVAE IAATAMEVVAATTAINATTAITI TINDSAINANINNAK 47 MERKULE IVSVAAAAALLAVAE IAATAMEVVAATTAITAITATTI TINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVVAATTAITATTESETINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVVAATTAITSTETINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVVAATTAITSTETINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVVAATAMEVVAATTAITSTETINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVVAATAMEVAATTAITSTESETINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVAATAMEVAATTAITSTESTINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEVAATAMEVAATAMEVAATAMEVAATAMEVAATTAITSTESETINKE TVDLSSAGSVESEK 57 MERKULE IVSAAAAALLAVAE IAATAMEV
ticus sequence for	ABT5 1H23 FBT9 1H70 1H52 SIS18 SIS18 1H5 1H12 1H1 DBT16 SIM3 SIO7 SIO7 1H28	ABT5 1.H23 FBT9 1.H70 1.H70 S-1.52 S-1.52 S-1.52 S-1.52 S-1.52 S-1.52 S-1.52 S-1.52 S-1.52 S-1.52 I.H71 SIS116 SIS717 SIS7	ABT5 1H23 FBT9 1H70 1H70 S-159er S-159 er S-159 er S-1518 1H12 1H12 1H12 1H4 SIT16 SIT16 SIT17 SIT07 LH28	ABT5 1423 FBT9 1470 1470 1470 1470 1470 1412 1412 1414 1412 1414 1412 1414 1415 1414 1415 1416 1417 1418 1418 1418 1418 1418 1418 1418
SlpH1 (NCBI protein identification number CAA62606.1) and the	NAYVYKISKKRANKVVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 435 NAYVYKISKKRANKVVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 422 NAYVYKISKKRANKVVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 432 NAYVYKISKKRANKVVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 432 NAYVYKISKKRANKVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 432 NAYVYKISKRANKVVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 432 NAYVYKISKRANKVVLKKGTEVTTYGGSYKFKNGQRYYKIGANTEKTY 432 NGIRLQDFKEAVLTRLFLR 407 MGIRLQDFKEAVLTRLFLR 407 NGIRLQDFKEAVLTRLFLR 407 NGIRLQDFKEAVLTRLFLR 407 NGIRLQDFKEAVLTRLFLR 407 NAYYYESSKKRANKVVLKKGTEVTTYGNEYTFKNGKKYYKIGANTEKTY 407 NAYYYESSKKRANKVVLKKGTEVTTYGNEYTFKNGKKYYKIGANTEKTY 407	RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 36 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 31 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 32 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 33 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 33 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 33 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 33 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 34 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 34 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 367 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 367 RNAKRWGTDKVTRVNTVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 367 RNAKRWGTDKVTRSNITVTVAMNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 367 RNAKRWGTDKVTRSNITVTANNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 367 RNAKRWGTDKVTRSNITVTANNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 367 RNAKRWGTDKVTRSNITVTANNITTKLANGISYYEVIENGKATGKYINADNIDGTKRTLKH 402	EVNSVGYEKAPHTETVINVKATSNVNSKSATLEVTVTVENVALPTVESSKIVMHINATEVD EVNSVGYEKAPHTETVINVKATSNKNSKSATLEVTVTVENVALPTVESSKIVMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD EVSSVGYEKAPHTETVINVKATSNKNGKSATLEVTVTVENVALPTVESGKKTIMHINATEVD 324 KVYAAGNEEAPDTETVILIANKSDVNGKTASLEVVVTVENGKSTVVESGSKTIMHINATEVD 331 KVDAAGNEEAPDTETVILIAKSDVNGKTASLEVVVTVENGKSTVVESGSKTIMHINATEVD 331 KVDAAGNEEAPDTETVILIAKSDVNGKTASLEVVVTVENGKSTVVESGSKTIMHINATEVD 331 KVDAAGNEEAPDTETVILIAKSDVNGKTASLEVVVTVENGKSTVVESGSKTIMHINATEVD 331 KVDAAGNEEAPDTETVILIAKSDVNGKTASLEVVVTVENGKSTVVESGSKTIMHINATEVD 331	DADNQGQLNTASVVAAITSKYFAAQYANKQLSQDNVVVNTETA-VKDALKAQKI 26 DADNQGQLNTASVVAAITSKYFAAQYANKQLSQDNVVVNTETA-VKDALKAQKI 261 DANNQGQLNTTSVVAAINSKYFAAQYDKKQLINVTEDIDTA-VKDALKAQKI 262 DADNQGQLNTTSVVAAINS

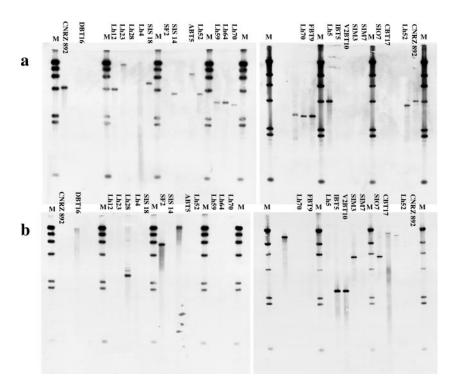


FIG. 5. Southern blot DNA hybridization of internal PCR-amplified fragments of the putative *slpH1* gene of approximately 400 bp, with total genomic DNA from all the *L. helveticus* strains following restriction with EcoRI. The two fragments used as probes were amplified from strains CNRZ 892 (a) and Lh28 (b) by using the pairs of primers P435 FOR-P831 REV and P435b FOR-P850a REV, respectively (Fig. 1). M, marker DNA/HindIII fragments (Invitrogen, Italy).

high level of sequence variation. BLASTN searches revealed that these sequences were highly homologous to those of a prtY gene (NCBI accession number AB026985) coding for a shorter, inactive extracellular *L. helveticus* proteinase (34). No relation between the source of isolation and the two groups was observed.

To complement BLASTN analysis, DNA sequences were further analyzed. According to the multiple alignment, the sequences of group I were divided into three stretches: nucleotides (nt) 1 to 209, nt 210 to 799, and nt 800 to1325, and compared to those of the reference *slpH1* gene. The 5' sequence stretch (1 to 209) was highly conserved (99 to 100% similarity) within all the strains with the exception of Lh70 (96% similarity). The internal sequence stretch (210 to 799) was highly conserved for 8 out of 12 strains, whereas for the other 4 strains (ABT5, Lh23, Lh64, and Lh59) its similarity to the corresponding stretch located on the reference *slpH1* gene was less marked. Finally, the 3' sequence stretch (800 to 1325) was the least conserved (89 to 93% similarity) with respect to *slpH1* (Fig. 3A).

The same approach was followed for the sequences of group II, which were compared with those of the reference prtY gene. Multiple alignments showed three sequence stretches corresponding to nt 1 to 120, nt 121 to 850, and nt 851 to 1180. As was observed for the sequences belonging to group 1, the central gene region appeared the most variable (77 to 83%), and its similarity with respect to the same stretch located on the *slpH1* gene was even lower (<60%). The 5' (1 to 120) and 3' (851 to 1180) flanking regions were highly similar (87 to 97%) and more conserved (Fig. 3B).

Analysis of primary amino acid sequences and multiple sequence alignment. Primary amino acid sequence predictions from the *slpH1* gene sequences were performed with the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (http://www.expasy.org/). Multiple sequence alignments of amino acid sequences were performed with ClustalW of the EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/clustalw/).

The deduced amino acid sequences were compared with the surface layer protein of L. helveticus (NCBI protein identification number CAA62606.1) and with the shorter and inactive proteinase PrtY of L. helveticus CP790 (NCBI protein identification number BAA86287) (Fig. 4). Multiple-sequence alignments showed highly conserved N-terminal and C-terminal regions. The previously described signal peptide of 30 amino acids located in the N terminus and the predicted cleavage site located between alanine residues 30 and 31 were observed for all the sequences. The function of this type of signal sequence is to govern the protein transport outside the cell membrane (2, 30). The sequence of the prosequences ranged from position 30 to 57, but the cleavage site was located at positions Ser37 and Thr38 (indicated as a double underline in Fig. 4). Thus, there is a difference between the lengths of the SlpH1 and PrtY prosequences. Moreover, strain SIM3 showed a 3-residue insertion (region 39 to 41) with respect to the other deduced proteins.

Highly variable amino acid composition in the central parts of the protein sequences was shown. The highest variability corresponded to the locations of insertion or deletion events (i.e., regions 39 to 48, 76 to 79, and 162 to 165) between the two groups of proteins. Many point mutations were observed, both within and between the two groups of proteins. Both silent and missense (both conservative and nonconservative) amino acid changes were observed. Examples of nonconservative amino acid changes were common: i.e., strains of group II showed substitutions of Asn-43 and Asn-45 with Ala (or Val) and Glu residues, respectively. Altogether, 99 nonconservative amino acid changes were observed in the most variable zone of the group II (proteinase-like) sequences with respect to the PrtY proteinases of *L. helveticus* CP790 (34).

DNA hybridization experiments. EcoRI-generated fragments of total DNA from the 21 *L. helveticus* strains were separated by electrophoresis on 1% (wt/vol) agarose gels. Restrictions were carried out for 2 h at 37°C in 20-µl volumes of incubation buffer (Life Technologies) containing 10 U of EcoRI restriction enzyme and 0.25 µg of total DNA. The Southern blotting technique was used to transfer DNA fragments to an Hybond N⁺ membrane (Amersham Pharmacia Biotech Italia, Milan, Italy) under alkaline conditions (0.4 N NaOH).

Two fragments of the putative *slpH1* gene of approximately 400 bp were amplified from strains CNRZ 892 and Lh28 by using the primers P435 FOR-P831 REV and P435b FOR-P850a REV, respectively (Fig. 1). Amplification reactions were performed as described above. The two 400-bp amplified fragments were used to probe EcoRI-digested DNAs in different hybridization experiments. DNA hybridization was performed with enhanced chemiluminescence-direct nucleic acid labeling and detection systems (Amersham Pharmacia Biotech Italia) according to the supplier's instructions. Overnight hybridization was carried out at 42°C. After signal generation and detection, autoradiography films (Hyperfilm-ECL, Amersham Pharmacia Biotech Italia) were exposed to generate light according to the manufacturer's instructions. The probe from CNRZ 892 gave positive signals only with EcoRI-generated fragments of total DNA from strains belonging to group I (Fig. 5a). The probe from Lh28 gave positive signals only with EcoRI-generated fragments of total DNA from strains belonging to group II (Fig. 5b).

Conclusion. The protein structure and function of S-layers have been extensively studied in *Lactobacillus acidophilus* (31, 32) and *L. crispatus* (1). Smit et al. (31) found that the N terminus, which encodes two-thirds of the S-layer protein of *L. acidophilus* ATCC4356, is involved in the self-assembly process of the S-layer protein, while one-third of the C terminus plays an essential role in the binding of the S-layer protein to the cell wall. This structure-function relationship was confirmed for the S-layer protein of *L. crispatus*, where the self-assembly domain was defined more precisely (1, 28).

We found a high homology to cell wall-anchoring domains within proteins of group I strains (SlpH1-like), which appeared well conserved, and the same motif was still present, although shorter, within proteins of group II strains (PrtY-like).

Intriguingly, the presence of either *slpH1* or proteinase was noted in the strains tested, as clearly suggested by hybridization experiments. This indicates that the presence of the *slpH1* gene cannot be used as a molecular marker for *L. helveticus*, as suggested by Ventura et al. (33). In contrast, the *prtY* gene can also be amplified using the same S-layer primers. A first consequence of the erroneous application of these primers can be

found in the work of Saito et al. from 2001 (24), which indicated amplification of the S-layer gene of *L. helveticus* JCM1003 in human feces, whereas, according to our BLAST analysis, the corresponding sequence reported in NCBI (accession number AB061776) was more similar to that of the *prtY* gene.

Even if the relationship between the proteinase and the surface layer protein is still not clear, this study shows the existence of heterogeneity within the *L. helveticus* S-layer sequences and contributes to a better understanding of the possible biological role of the different cell surface proteins in *L. helveticus* populations.

We are grateful to Marco Ventura and J. G. Kenny (Department of Microbiology, National University of Ireland, Cork) and Diego Mora (DISTAM, University of Milano, Milano, Italy) for critical readings of the manuscript and useful suggestions.

This work was partly supported by the Italian Agricultural Ministry within the Program "Valorizzazione e salvaguardia della microflora autoctona caratteristica delle produzioni casearie Italiane," article submission number 26.

REFERENCES

- Antikainen, J., L. Anton, J. Sillanpää, and T. K. Korhonen. 2002. Domains in the S-layer protein Cbsa of *Lactobacillus cispatus* involved in adherence to collagens, laminin and lipoteichoic acids and in self-assembly. Mol. Microbiol. 2:381–394.
- Bahl, H., H. Scholz, N. Bayan, M. Chami, G. Leblon, T. Gulik-Krzywicki, E. Shechter, A. Fouet, S. Mesnage, E. Tosi-Couture, P. Gounon, M. Mock, E. Conway de Macario, A. J. L. Macario, L. A. Fernandez-Herrero, G. Olabarria, J. Berenguer, M. J. Blaser, B. Kuen, W. Lubitz, M. Sara, P. H. Pouwels, C. P. A. M. Kolen, H. J. Boot, A. Palva, M. Truppe, S. Howwrka, G. Schroll, S. Lechleitner, and S. Resch. 1997. Molecular biology of S-layer. FEMS Microbiol. Rev. 20:47–98.
- Beveridge, T. J., P. H. Pouwels, M. Sára, A. Kotiranta, K. Lounatmaa, K. Kari, E. Kerosuo, M. Haapasalo, E. M. Egelseer, I. Schocher, U. B. Sleytr, L. Morelli, M. L. Callegari, J. F. Nomellini, W. H. Bingle, J. Smith, E. Leibovitz, M. Lemaire, I. Miras, S. Salamitou, P. Beguin, H. Ohayon, P. Gounon, M. Matuschek, and S. F. Koval. 1997. Functions of S-layer. FEMS Microbiol. Rev. 20:99–149.
- Boot, H. J., C. P. A. M. Kolen, J. M. van Noort, and P. H. Pouwels. 1993. S-layer protein of *Lactobacillus acidophilus* ATCC 4356: purification, expression in *Escherichia coli*, and nucleotide sequence of the corresponding gene. J. Bacteriol. 175:6089–6096.
- Boot, H. J., C. P. A. M. Kolen, B. Pot, K. Kersters, and P. H. Pouwels. 1996. The presence of two S-layer-protein-encoding genes is conserved among species related to *Lactobacillus acidophilus*. Microbiology 142:2375–2384.
- Boot, H. J., and P. H. Pouwels. 1996. Expression, secretion and antigenic variation of bacterial S-layer proteins. Mol. Microbiol. 21:117–1123.
- Callegari, M. L., B. Riboli, J. W. Sanders, P. S. Cocconcelli, J. Kok, G. Venema, and L. Morelli. 1998. The S-layer gene of *Lactobacillus helveticus* CNRZ 892: cloning, sequence and heterologous expression. Microbiology 144:719–726.
- Carminati, D., L. Mazzuccotelli, G. Giraffa, and E. Neviani. 1997. Incidence of inducible bacteriophage in *Lactobacillus helveticus* strains isolated from natural whey starter cultures. J. Dairy Sci. 80:1505–1511.
- de los Reyes-Gavilàn, C., G. K. Y. Limsowtin, P. Tailliez, L. Séchaud, and J. P. Accolas. 1992. A *Lactobacillus helveticus*-specific DNA probe detects restriction fragment length polymorphism in this species. Appl. Environ. Microbiol. 58:3429–3432.
- Fortina, M. G., G. Nicastro, D. Carminati, E. Neviani, and P. L. Manachini. 1998. *Lactobacillus helveticus* heterogeneity in natural cheese starters: the diversity in phenotypic characteristics. J. Appl. Microbiol. 84:72–80.
- Gatti, M., E. Fornasari, and E. Neviani. 1997. Cell wall protein profiles of dairy thermophilic lactobacilli. Lett. Appl. Microbiol. 25:345–348.
- Gatti, M., G. Contarini, and E. Neviani. 1999. Effectiveness of chemometric techniques in discrimination of *Lactobacillus helveticus* biotypes from natural dairy starter cultures on the basis of phenotypic characteristics. Appl. Environ. Microbiol. 65:1450–1454.
- Gatti, M., C. Lazzi, L. Rossetti, G. Mucchetti, and E. Neviani. 2003. Biodiversity in *Lactobacillus helveticus* strains in natural whey starter used for Parmigiano Reggiano cheese. J. Appl. Microbiol. 95:463–470.
- Gatti, M., C. Trivisano, E. Fabrizi, E. Neviani, and F. Gardini. 2004. Biodiversity within *Lactobacillus helveticus* isolated from different natural whey starter cultures as revealed by classification trees. Appl. Environ. Microbiol. 70:182–190.

- Giraffa, G., G. Mucchetti, F. Addeo, and E. Neviani. 1997. Evolution of lactic acid microflora during Grana cheese-making and ripening. Microbiol. Aliments Nutr. 15:115–122.
- Giraffa, G., P. De Vecchi, P. Rossi, G. Nicastro, and M. G. Fortina. 1998. Genotypic heterogeneity among *Lactobacillus helveticus* strains isolated from natural cheese starters. J. Appl. Microbiol. 85:411–416.
- Giraffa, G., and E. Neviani. 1999. Different Lactobacillus helveticus strain populations dominate during Grana Padano cheese-making. Food Microbiol. 16:205–210.
- Giraffa, G., M. Gatti, L. Rossetti, L. Senini, and E. Neviani. 2000. Molecular diversity within *Lactobacillus helveticus* as revealed by genotypic characterization. Appl. Environ. Microbiol. 66:1259–1265.
- Lombardi, A., L. Dal Maistro, P. De Dea, M. Gatti, G. Giraffa, and E. Neviani. 2002. A poliphasic approach to highlight genotypic and phenotypic diversities of *Lactobacillus helveticus* strains isolated from dairy starter cultures and cheeses. J. Dairy Res. 69:139–149.
- Navarre, W. W., and O. Schneewind. 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to cell wall envelope. Microbiol. Mol. Biol. Rev. 63:174–229.
- Neviani, E., and S. Carini. 1994. Microbiology of Parmesan cheese. Microbiol. Aliments Nutr. 12:1–8.
- Neviani, E., R. Divizia, E. Abbiati, and M. Gatti. 1995. Acidification activity of thermophilic lactobacilli under the temperature gradient of Grana cheese making. J. Dairy Sci. 78:1248–1252.
- Quiberoni, A., P. Tailliez, P. Quénée, V. Suarez, and J. Reinheimer. 1998. Genetic (RAPD-PCR) and technological diversities among wild *Lactobacillus helveticus* strains. J. Appl. Microbiol. 85:591–596.
- 24. Saito, Y., Y. Hamanaka, S. Takizawa, and Y. Benno. Rapid detection of Lactobacillus helveticus in fecal samples of healthy subjects administered L. helveticus yogurt using s-layer gene-targeted primers. Unpublished data.

- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed., vol. 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sára, M., and U. B. Sleytr. 2000. S-layer proteins. J. Appl. Bacteriol. 182: 859–868.
- Schaer-Zammaretti, P., and J. Ubbink. 2003. Imaging of lactic acid bacteria with AFM—elasticity and adhesion maps and their relationship to biological and structural data. Ultramicroscopy 97:199–208.
- Siezen, R. J. 1999. Multi-domain, cell-envelope proteinases of lactic acid bacteria. Antonie Leeuwenhoek 76:139–155.
- Sillanpää, J., B. Martínez, J. Antikainen, T. Toba, N. Kalkkinen, S. Tankka, K. Lounatmaa, J. Keränen, M. Höök, B. Westerlund-Wikström, P. H. Pouwels, and T. K. Korhonen. 2000. Characterization of the collagen-binding S-layer protein CbsA of *Lactobacillus crispatus*. J. Bacteriol. 182:6440–6450.
- Sleytr, Ü. B., and T. J. Beveridge. 1999. Bacterial S-layers. Trends Microbiol. 7:253–260.
- Smit, E., F. Oling, R. Demel, B. Martinez, and P. H. Pouwels. 2001. The S-layer protein of *Lactobacillus acidophilus* ATCC 4356: identification and characterizations of domains responsible for S-protein assembly and cell wall binding. J. Mol. Biol. 305:245–257.
- Smit, E., D. Jager, B. Martinez, F. J. Tielen, and P. H. Pouwels. 2002. Structural and functional analysis of the S-layer protein crystallisation domain of *Lactobacillus acidophilus* ATCC 4356: evidence for protein-protein interaction of two subdomains. J. Mol. Biol. 324:953–964.
- Ventura, M., M. L. Callegari, and L. Morelli. 2000. S-layer gene as a molecular marker for identification of *Lactobacillus helveticus*. FEMS Microbiol. Lett. 189:275–279.
- 34. Yamamoto, N., T. Shinoda, and T. Takano, 2000. Molecular cloning and sequence analysis of a gene encoding an extracellular proteinase from *Lac-tobacillus helveticus* CP790. Biosci. Biotechnol. Biochem. 64:1217–1222.