CrossMark

ORIGINAL PAPER

Lactobacillus kefiri shows inter-strain variations in the amino acid sequence of the S-layer proteins

Mariano Malamud · Paula Carasi · Sílvia Bronsoms · Sebastián A. Trejo · María de los Angeles Serradell

Abstract The S-layer is a proteinaceous envelope constituted by subunits that self-assemble to form a two-dimensional lattice that covers the surface of different species of *Bacteria* and *Archaea*, and it could be involved in cell recognition of microbes among other several distinct functions. In this work, both proteomic and genomic approaches were used to gain knowledge about the sequences of the S-layer protein (SLPs) encoding genes expressed by six aggregative and sixteen non-aggregative strains of potentially probiotic *Lactobacillus kefiri*. Peptide mass fingerprint (PMF) analysis confirmed the identity of SLPs extracted from *L. kefiri*, and based on the homology with phylogenetically related species, primers located outside and inside the SLP-genes were employed to amplify genomic DNA. The O-glycosylation site SASSAS was found in all *L. kefiri* SLPs. Ten strains were selected for sequencing of the complete genes. The total length of the mature proteins varies from 492 to 576 amino acids, and all SLPs have a calculated pI between 9.37 and 9.60. The N-terminal region is relatively conserved and shows a high percentage of positively charged amino acids. Major differences among strains are found in the C-terminal region. Different groups could be distinguished regarding the mature SLPs and the similarities observed in the PMF spectra. Interestingly, SLPs of the aggregative strains are 100% homologous, although these strains were

M. Malamud · P. Carasi ·
M. de los Angeles Serradell (✉)
Cátedra de Microbiología, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata (UNLP), 47 y 115,
CP 1900 La Plata, Argentina
e-mail: maserr@biol.unlp.edu.ar

M. Malamud · P. Carasi · S. A. Trejo ·
M. de los Angeles Serradell
CCT-La Plata, CONICET, Buenos Aires, Argentina

S. Bronsoms · S. A. Trejo
Unidad de Proteómica del Servicio de Proteómica y Biología Estructural (SePBioEs), Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

S. A. Trejo
Instituto Multidisciplinario de Biología Celular (IMBICE), CONICET, La Plata, Argentina

M. de los Angeles Serradell
Instituto de Ciencias de la Salud, Universidad Arturo Jauretche (UNAJ), Florencio Varela, Argentina

🖄 Springer

isolated from different kefir grains. This knowledge provides relevant data for better understanding of the mechanisms involved in SLPs functionality and could contribute to the development of products of biotechnological interest from potentially probiotic bacteria.

Introduction

The diversity observed in the molecular architecture of the outermost surface layer of prokaryotic microorganisms is an example of the evolutionary adaptation of an organism to specific ecological and environmental conditions. The S-layer, a proteinaceous envelope that is present in many species belonging to the prokaryotic domains of *Bacteria* and *Archaea*, is regarded as the most ancient biological membrane that has remained through the evolution of microbes (Claus et al. 2005; Messner et al. 2013; Zhu et al. 2016). S-layers are monomolecular arrays of protein or glycoprotein subunits that self-assemble to form a two-dimensional lattice that completely covers the organism during all stages of growth (Sára and Sleytr 2000). Despite the abundancy of S-layer proteins (SLPs) in many prokaryotic cells, their functions have been explained only in few cases and many of them are still hypothetical. These lattices could function as protective coats, molecular sieves or ion traps, as structures involved in surface recognition, cell adhesion, inhibition of pathogens or as virulence factors (Sleytr et al. 2014; Gerbino et al. 2015b; Prado Acosta et al. 2016).

SLPs have been shown to possess exceptional physicochemical properties which make them a unique organizational structure with high potential for application in different areas of bionanotechnology such as generation of generally recognized as safe (GRAS) vehicles for the administration of antigens and other molecules of biomedical importance, and development of reactive solid-phase as biocatalysts, diagnostic devices or biosensors (Ilk et al. 2011; Sleytr et al. 2014). Moreover, they are also very useful as models to learn more about the strategies for stabilization, self-organization and functional evolution of proteins (Claus et al. 2005).

The presence of S-layers has been described in many bacterial species, including some of the genus *Lactobacillus* (Hynönen and Palva 2013). Lactobacilli are typically GRAS and have been isolated from different natural environments, including plants, human gastrointestinal and genital tracts, and food products. Years ago, our group demonstrated that both aggregative and non-aggregative strains of *Lactobacillus kefiri*, one of the most abundant lactobacilli present in the probiotic fermented milk named kefir (Garrote et al. 2005), carry a glycosylated S-layer (Garrote et al. 2004, Mobili et al. 2009b). These SLPs mediate the inhibition of *Salmonella enterica* invasion to Caco-2 cells (Golowczyc et al. 2007), the antagonism of *Clostridium difficile* toxins (Carasi et al. 2012), the interaction of *L. kefiri* with yeasts (Golowczyc et al. 2009), they also participate in the adhesion of *L. kefiri* to the gastrointestinal mucus (Carasi et al. 2014a), and protect bacterial cells against the deleterious effect of Pb^{2+} ions (Gerbino et al. 2015a). Despite all these interesting beneficial properties, up to now no information about the amino acid sequence of these *L. kefiri* SLPs has been reported.

In this work, we used both proteomic and genomic approaches to gain knowledge about the sequences of the genes encoding SLPs expressed by both aggregative and non-aggregative cultured strains of potentially probiotic *L. kefiri*. Differences in amino acid sequences lead to changes in secondary and conformational structures which are responsible, at least in part, for distinct surface properties of whole bacterial cells. Moreover, the knowledge of the sequence of these proteins is essential for the design and development of new bionanotechnological tools taking advantage of the extraordinary structural features of the SLPs.

Materials and methods

Bacterial strains and growth conditions

Twenty *L. kefiri* strains isolated from kefir grains (Garrote et al. 2001; Hamet et al. 2013), *L. kefiri* JCM 5818 obtained from the Japanese Collection of Microorganisms (Reiken, Japan) and *L. kefiri* ATCC 8007 were used in this study. Lactobacilli were cultured in deMan-Rogosa-Sharpe (MRS)-broth (DIFCO, Detroit, USA) at 37 °C for 48 h in aerobic conditions. Frozen stock cultures were stored at

−80 °C in skim milk until use. The most relevant characteristics of these microorganisms and their SLPs are shown in Table 1.

Surface protein (S-layer) extraction

S-layer protein extraction from bacterial cells was performed using 5 M LiCl, as previously described

(Carasi et al. 2012). Briefly, 100 ml of MRS culture of *L. kefiri* was harvested at stationary phase, collected by centrifugation (10,000×*g* at 10 °C for 10 min), washed three times with phosphate buffered saline (PBS, KH_2PO_4 0.144 g/l, NaCl 9 g/l, Na_2HPO_4 0.795 g/l, pH 7.2) and bacteria were resuspended in 15 ml of 5 M LiCl (J.T. Baker, Mallinckrodt Baker S.A., Edo de Mexico, Mexico) giving a bacterial

Table 1 Characteristics of *L. kefiri* strains and their S-layer proteins

L. kefiri strains[a]		S-layer proteins	
L. kefiri strain's codes	Autoagreggation capacity[b]	Apparent molecular mass[c] (kDa)	Glycosylation[d]
Strains isolated from CIDCA collection Kefir grains			
Kefir grain AGK1			
CIDCA 8310	−	71	+[e]
CIDCA 8314	−	71	+[e]
CIDCA 8315	−	71	+[e]
CIDCA 8317	−	71	+[e]
CIDCA 83110	−	71	+[e]
CIDCA 83111	−	71	+
CIDCA 83113	−	66	+
CIDCA 83115	+	66	+
CIDCA 83116	−	71	+[e]
Kefir grain AGK2			
CIDCA 8321	+	66	+
CIDCA 8326	−	71	+[e]
Kefir grain AGK3			
CIDCA 8332	−	71	+[e]
CIDCA 8335	−	66	+
Kefir grain AGK4			
CIDCA 8343	−	71	+[e]
CIDCA 8344	+	71	+
CIDCA 8345	+	66	+
CIDCA 8347	+	66	+
CIDCA 8348	+	66	+
Kefir grain AGK8			
CIDCA 8381	−[e]	71[e]	+[e]
CIDCA 8385	−[e]	71[e]	+[e]
Reference strains			
JCM 5818	−	69	+
ATCC 8007	−	71	+

[a] The identity of the strains was previously evaluated (Garrote et al. 2001; Delfederico et al. 2006; Hamet et al. 2013)

[b] Data from Garrote et al. (2004)

[c] Apparent molecular mass calculated using SDS-PAGE migration patterns (Garrote et al. 2004; Mobili et al. 2009b)

[d] Glycosylation analysis by PAS staining (Mobili et al. 2009b)

[e] Glycosylation analysis by PAS staining, this work

suspension of $OD_{550} = 25$. The mixture was incubated in a shaker at $300 \times g$ at 4 °C for 60 min. Then, it was centrifuged ($16,000 \times g$ at 10 °C for 30 min) and the supernatant was filtered through a membrane filter with 0.45 μm pore diameter (Millipore, USA). The supernatant was then dialyzed against PBS with 0.05% (v/v) Tween 20 for 24 h at room temperature, following by three additional buffer changes of 5 l each, using a cellulose membrane (SpectraPor membrane tube, MWCO 6000–8000, Spectrum Medical Industries, Rancho Dominguez, CA, USA). SLPs extracts were visualized by sodium dodecylsulphate–polyacrylamide gel electrophoresis (SDS–PAGE) in 12% separating and 4% stacking gels using the discontinuous buffer system according to Laemmli (1970). Gels were migrated on a BioRad Mini-Protean II (BioRad Laboratories, Richmond, CA, USA), with LMW Marker kit (GE Healthcare, Sweden) as molecular weight reference and were revealed using Colloidal Blue Staining.

Mass spectrometry

In-gel protein digestion

Gel pieces were destained in 3×100 μl washes of 25 mM NH_4HCO_3, 5% acetonitrile (ACN) (pH 8,5), followed by reduction with 100 μl 10 mM DTT in 25 mM NH_4HCO_3 at room temperature for 30 min and then alkylation with 25 mM iodoacetamide in 25 mM NH_4HCO_3, at room temperature for 30 min. Gel pieces were dessicated with 100 μl of ACN 100% for 10 min at room temperature and rehydrated with trypsin (Promega Trypsin Gold, TPCK treated) in 25 mM NH_4HCO_3, at an approximate trypsin to protein ratio of 1:20. The enzymatic reaction was carried out at 37° C for 3 h and peptides were extracted from the gel pieces with 100 μl of 0.2% TFA. The eluted peptides were dried with a Speed-Vac^{TM} and then suspended in 5 μl of 50% ACN, 0.1% TFA. All assays were carried out in MultiScreen solvinert filter plates (Millipore) with a MultiScreen^{TM} Vacuum Manifold 96-well system (MILLIPORE, Billerica, MA).

MALDI–TOF–MS analysis

Peptide mass fingerprint (PMF) analysis was used to determine protein identities. PMF and MS/MS fragmentation spectra were acquired with an UltrafleXtreme MALDI–TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) at the SePBioEs Proteomics Service (UAB, Barcelona, Spain). Samples were spotted on a ground steel target plate (Bruker Daltonics) mixing 0.5 μl of each sample with 0.5 μl of freshly prepared matrix solution of a-cyano-4-hydroxycinnamic acid (Bruker Daltonics) at 10 mg/ml in a 30% ACN and 0.1% TFA aqueous solution. An external calibration was performed using a standard peptide mixture (Bruker Daltonics). Peptide masses were acquired within a range of ca. m/z 800–4000. Protein identification was carried out with Mascot search engine (Matrix Science Inc., Boston, MA) using the following parameters: two missed cleavages, 100 ppm tolerance, cysteine carbamidomethylation and methionine oxidation were set as variable modifications. Searches were performed using the NCBInr database restricted to Firmicutes.

The tandem mass spectra of selected peptides were obtained and de novo sequencing was performed manually using the table of amino acid masses.

Primer design, PCR conditions, sequencing and restriction profiles

The presence of SLP encoding genes was analyzed by specific PCR using purified chromosomal DNA as a template. Based on *Lactobacillus buchneri* CD034 genome, primers located outside (F1, F2, R1) and inside (TREG-f, TREG-r, GWIY-f) the ORF were constructed for PCR sequencing strategy. Their sequences are shown in Table 2, while their localization in the *L. buchneri* CD034 genome is shown in Online Resource 1. The PCR was performed as follows: one step at 95 °C for 5 min; 35 steps 94 °C for 30 s, 54 °C for 30 s and 72 °C for 3 min; and a final step at 72 °C for 10 min.

Sequencing reactions were performed using a ABI PRISM® BigDyeTM Terminator Cycle Sequencing Kit (Applied Biosystems, USA). The fluorescent-labeled fragments were purified following a BigDye® XTerminator^{TM} purification protocol. The samples were resuspended in distilled water and subjected to electrophoresis in an ABI 3730xl sequencer (Applied Biosystems, USA).

The amplicons obtained using the primers F1 and R1 (or amplicon F2/R1 in the case of *L. kefiri* CIDCA 5818) were cleaved simultaneously with

Table 2 Primers and their nucleotide sequences	Primer code	Sequence 5′–3′	Annealing temperature (°C)
	TREG-R	AACCCAAAGGTCGCCTTCAC	54
	TREG-F	ACCCGTGAAGGCGACCTTTG	54
	GWIY-F	GGCTGGATCTACGGTGGCAA	54
	R1	CTCCTGTTTTCTTTTGCATTCAACC	54
	F1	ACGTTGGCCACAGGAAAGAA	54
	F2	GACCATTATCGAACTAGTGTTCAAA	54

EcoRI and NcoI (New England BioLabs, UK) according to the manufacturer's instructions. The digested DNA fragments were separated on a 1.5% (w/v) agarose gel.

In silico analysis of nucleotide and amino acid sequences

The SLP gene sequences obtained in this study, as well as those from the databases, were analyzed by ClustalOmega (Goujon et al. 2010; Sievers et al. 2011; McWilliam et al. 2013).

The amino acid sequence was deduced from the nucleotide sequence using the ExPASy Translate Tool (http://web.expasy.org/translate/) and in silico tryptic digestion was performed using ExPASy Peptide Mass (http://web.expasy.org/peptide_mass/), both softwares available on ExPASy portal (Artimo et al. 2012). The obtained profiles were compared manually with the experimental spectra previously obtained for each SLP. The theoretical physicochemical properties (pI, Mw, %Asp-Glu, %Arg, Lys) of the amino acid sequences were predicted using the software ProtParam (http://web.expasy.org/protparam/) (Gasteiger et al. 2005) and SignalP 4.0 (Petersen et al. 2011) was used to identify the signal peptide. Secondary structure prediction of SLPs was computed with PSIPRED protein structure prediction server (http://bioinf.cs.ucl.ac.uk/psipred/).

Nucleotide sequence accession numbers

The sequence data of complete genes have been deposited in the EMBL nucleotide sequence database under accession numbers LT601591, LT601592, LT601593, LT601594, LT601595, LT601596, LT601597, LT601598, LT601599 and LT601600.

Results and discussion

Mass spectrometry analysis of SLPs from *L. kefiri* strains

Due to the importance of *L. kefiri* as a potentially probiotic microorganism as well as a component of kefir microbiota (Garrote et al. 2005; Golowczyc et al. 2007; Carasi et al. 2014a, b, 2015), some studies about the biochemical and functional properties of SLPs belonging to *L. kefiri* strains have been conducted by our group in recent years (Mobili et al. 2009a, b; Golowczyc et al. 2007, 2009; Carasi et al. 2012, 2014a; Gerbino et al. 2015a). However, this is the first work reporting the amino acid sequences of SLPs from both aggregative and non-aggregative *L. kefiri* strains.

Lactobacillus SLPs are among the smallest so far described, with molecular masses ranging from 25 to 71 kDa (Hynönen and Palva 2013; Wasko et al. 2014), which agrees with the results presented in this work as well as to previous reports for the SLPs isolated from *L. kefiri* strains (Table 1). The identity of SLPs extracted from *L. kefiri* and digested by trypsinization was confirmed by PMF analysis, finding homology with *L. buchneri* CD034 and *Lactobacillus parafarraginis* F0439 SLPs. The amino acid sequence of several peptides was confirmed by MS/MS fragmentation and subsequent de novo sequencing. Interestingly, some of these fragments also matched with peptides present in SLPs of other phylogenetically related species, such as *Lactobacillus otakiensis* and *Lactobacillus kisonensis* (Endo and Okada 2007; Watanabe et al. 2009). Representative PMFs of the SLPs from *L. kefiri* strains are shown in Fig. 1. In concordance with previous results (Mobili et al. 2009b), SLPs from aggregating strains showed very similar spectral patterns (Fig. 1, spectra A–D), on the other hand more heterogeneity was observed in the various

spectra corresponding to SLPs from non-aggregative strains (Fig. 1, spectra E–S). The analysis of all tryptic peptides detected in the PMF experiments (data not shown), led to identification of two peptides that were present in all samples. The presence of the tryptic peptides m/z 1103.6 and m/z 1200.6 was observed in the mass spectra from the SLPs of all *L. kefiri* strains studied, regardless of their aggregative ability. Moreover, these fragments were also present in the SLPs of *L. buchneri* CD034 and *L. parafarraginis* F0439. Both peptides were further analyzed by MS/MS fragmentation and de novo sequencing, and their sequences could be validated. The 1103.6 Da peptide corresponded to TREGDLWVK sequence, and the 1200.6 Da peptide corresponded to TYRGWIYGGK sequence.

In general, the deduced amino acid sequences of mature *Lactobacillus* SLPs vary considerably, thus a similarity can only be found among related species (Antikainen et al. 2002; Avall-Jääskeläinen and Palva 2005; Hagen et al. 2005). Indeed, the remarkable similarities between SLPs from *L. acidophilus*-related bacteria have led to the proposal of employing their LC–MS/MS analysis for typing strains within this group (Podlesny et al. 2011). Similarly, an exceptionally high conservation was reported for the SLPs of two *Lactobacillus hilgardii* strains (Dohm et al. 2011) as well as for the SLPs of *L. acidophilus* NCFM and ATCC 4356 (Hynönen and Palva 2013), although the strains were clearly distinguishable. In the case of *L. kefiri* strains, although the presence of SLPs in *L. kefiri* strains was described more than eleven years ago (Garrote et al. 2004), and their heterogeneity had been characterized by using specific monoclonal antibodies and MALDI–TOF analysis in our laboratory (Mobili et al. 2009b), we were not able to obtain suitable tools to start to get reliable information about the amino acid sequence of these proteins until the complete genome of *L. buchneri* CD034 was published (Heinl et al. 2012).

L. kefiri SLP-encoding gene amplification

Taking into account the data obtained by mass spectrometry analysis, along with the availability of both the complete genomic sequence of *L. buchneri* CD034 and sequence read archives (SRA) available from other related species (*L. parafarraginis*, *L. kisonensis*, *L. otakiensis*, *L. parabuchneri*, *L. parakefiri*)

(Endo and Okada 2007; Watanabe et al. 2009; Oki et al. 2012), the alignment of the SLP-encoding gene sequences from those lactobacilli was performed using the bioinformatics software platform Geneious (Biomatters Limited). Based on the alignment, the regions showing the highest homology were selected and the set of SLP-specific primers referred to as TREG-r (within the peptide of m/z 1103.6) and GWIY-f (within the peptide of m/z 1200.6) were designed, using the *L. buchneri* CD034 genetic code. Using these primers, an amplicon of 340-bp was obtained from genomic DNA of all *L. kefiri* strains. The fragments were then sequenced, and employing the software available on the ExPASy portal (http://web.expasy.org/translate), nucleotide sequences were translated into peptides. The five aggregative strains (group I) showed no differences in the amino acid sequence of this 106-mer fragment among them, whereas the non-aggregative strains could be separated into four different groups based on the similarities observed in this fragment: group II (strains CIDCA 8314, 8315, 8326, 83111, ATCC 8007), III (strains CIDCA 8310, 8317, 8319, 8332, 8343, 8344, 8381, 8385, 83110, 83116), IV (strains 8335, 83113) and V (strain JCM 5818).

Regarding these results, a total of ten strains belonging to different groups were selected and sequencing of the complete SLP-encoding genes was performed. The partial DNA sequences obtained with the combinations of F1/R1, GWIY-f/TREG-r, GWIY-f/F1 and TREG-r/R1 were assembled to obtain the complete S-layer gene sequences from *L. kefiri* CIDCA 8310, 8314, 8321, 8335, 8343, 8348, 83111, 83113 and 83115. For amplification of DNA from *L. kefiri* JCM 5818, the combinations F1/R1 and GWIY-f/F1 were replaced by F2/R1 and GWIY-f/F2 respectively, because of the absence of F1 region in this strain. All the S-layer gene sequences were deposited in the EMBL database with the corresponding accession numbers given in the "Materials and methods" section.

Since November 2015, two Genome Assemblies and Annotation reports for *L. kefiri* JCM 5818 are available in the NCBI site (Sun et al. 2015). Three different loci encoding hypothetical SLPs can be found. One of them shows 100% nucleotide identity with the sequence reported in this work for *L. kefiri* JCM 5818 (accession AYYV01000002.1; locus KRM54156), which confirms our results, meanwhile
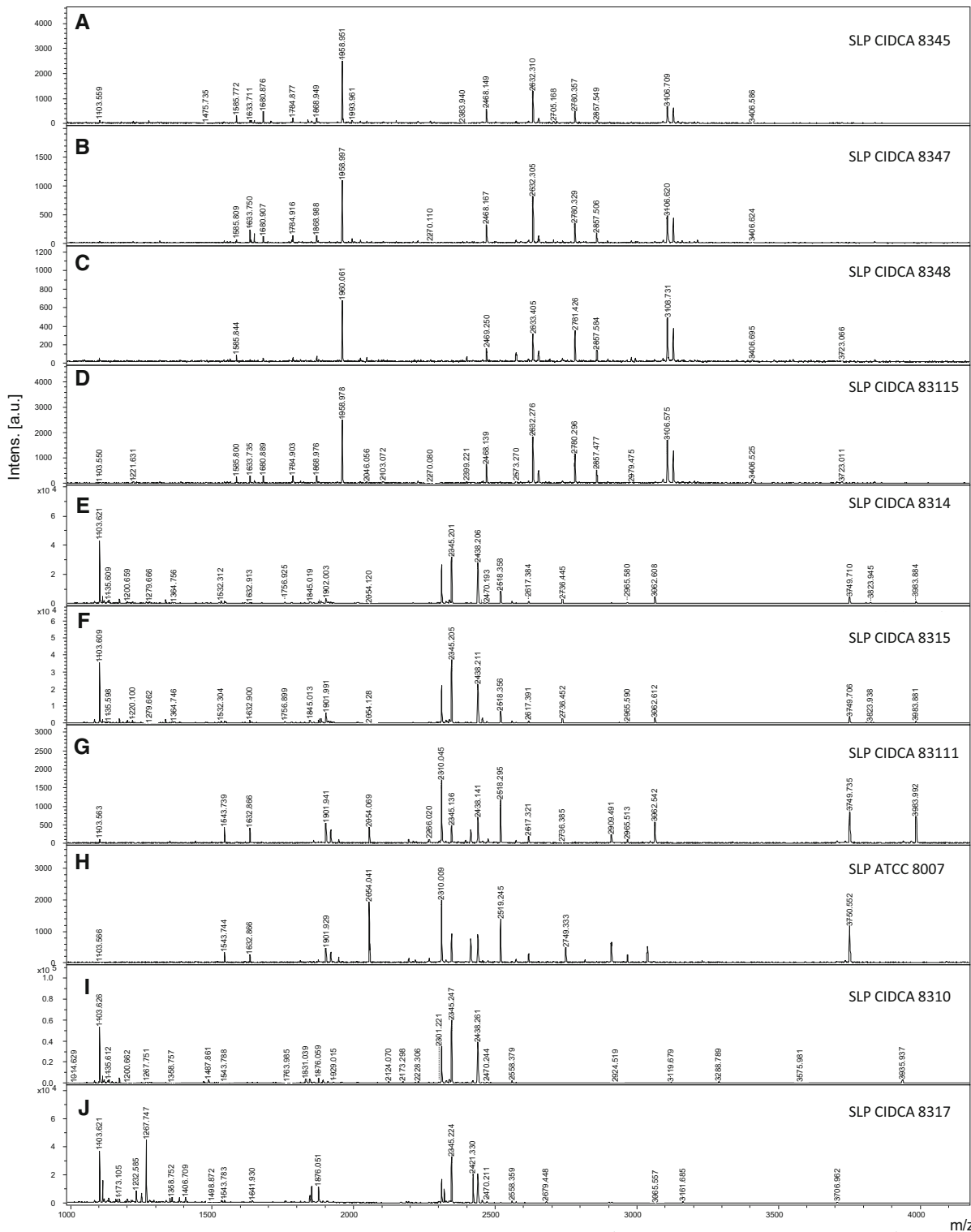
Fig. 1 Peptide mass fingerprints (PMF) of S-layer proteins (SLP) from representative aggregative (*A–D*) (SLP CIDCA 8345, 8347, 8348[*] and 83115) and non-aggregative (*E–S*) (SLP CIDCA 8310, 8314, 8315, 8317, 8332, 8335, 8343, 8381, 8385, 83111, 83113[*], 83116, ATCC 8007 and JCM 5818[*]) strains of *L. kefiri*. *Mass spectra of these SLPs were also reported by Mobili et al. in a previous work (2009b)
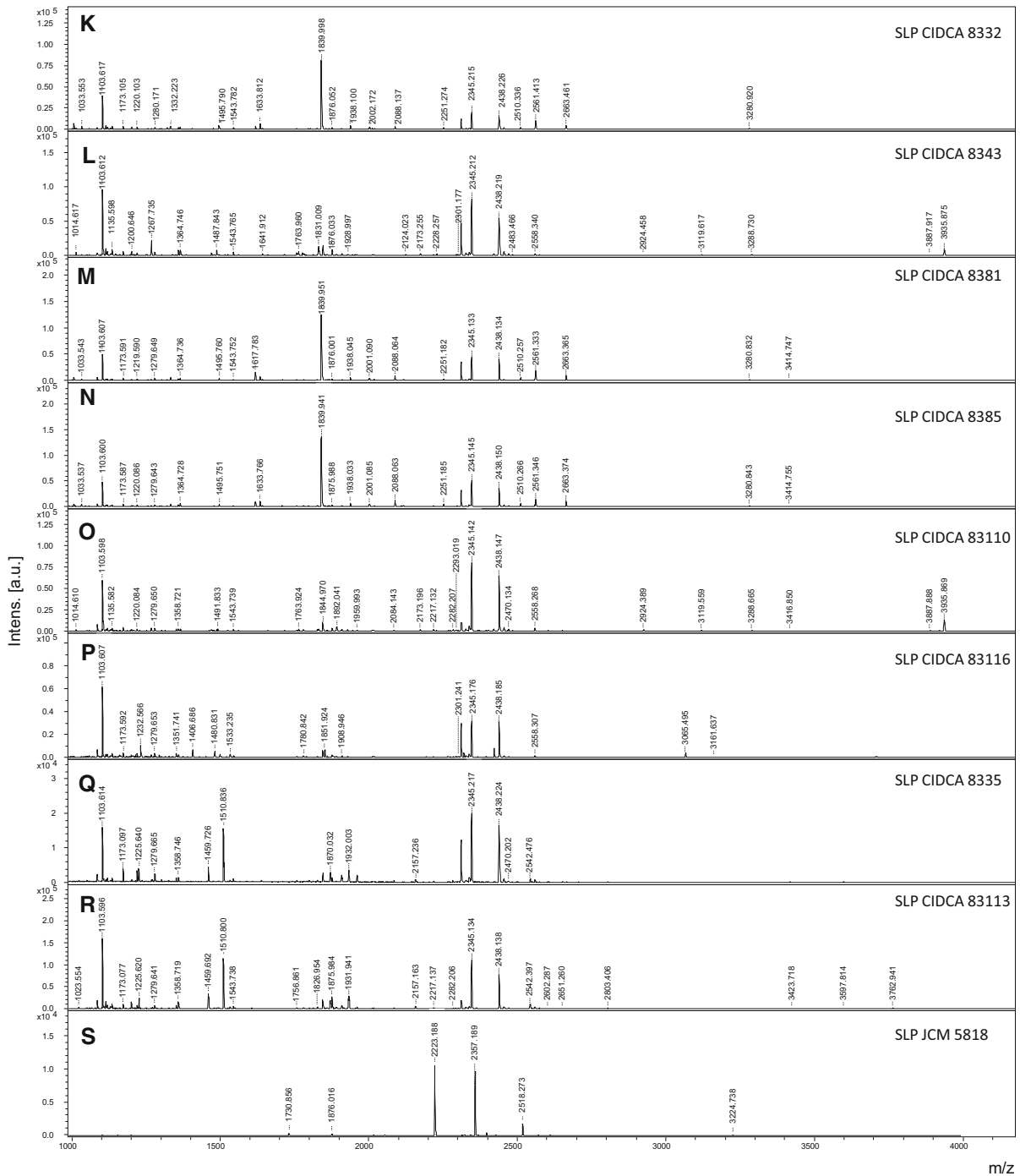
Fig. 1 continued

the other two show 72% (accession AYYV01000 037.1; locus KRM52454) and 60% (accession AYYV01000004.1; locus KRM53976) nucleotide identity, respectively. These findings suggest that at least three SLP-encoding genes are present in the genome of the reference strain JCM 5818. Locus KRM52454 encodes for a polypeptide of 540 amino acids and locus KRM53976 encodes for a polypeptide

of 565 amino acids whereas the protein described in this work shows a total length of 578 amino acids. The presence of more than one SLP-encoding gene has been described in different lactobacilli species, including *Lactobacillus brevis* ATCC 14869 (Jakava-Viljanen et al. 2002); *L. acidophilus* ATCC 4356 (Boot et al. 1995; Palomino et al. 2016) and *L. acidophilus* NCFM (Goh et al. 2009). Moreover, it was reported that the expression level of different SLP's genes could change depending on environmental conditions such as high salt concentration or gastrointestinal tract milieu (Ramiah et al. 2009; Palomino et al. 2016).

Although the primers located inside the ORF (TREG-f, GWIY-f) could generate amplification products of the three mentioned putative genes, it is important to note that the primers used for sequencing, which are located outside the ORF (F2 and R1), only match to the flanking region of the locus KRM54156 of *L. kefiri* JCM 5818. This finding allows us to assume that the amplicons obtained correspond to a gene located in that region of the genome. The presence of other SLP encoding genes in the *L. kefiri* strains of our collection as well an analysis of differential expression patterns will be subjects of future studies.

Besides F1/R1 amplicons (or F2/R1 amplicon in the case of *L. kefiri* JCM 5818) of 2500–3000 bp-length were obtained from genomic DNA of these *L. kefiri* strains. The digestion of these amplicons with a mix of EcoRI and NcoI, resulted in different fragmentation patterns that match with the distinct groups described above (Fig. 2).

Amino acid sequences from *L. kefiri* SLPs

The SLP encoding genes of the selected *L. kefiri* strains were translated into the corresponding protein sequences, and then confirmed by comparison of each experimental PMF spectrum with the theoretical tryptic digestion obtained with the PeptideMass tool (http://web.expasy.org/peptide_mass/). Even though it is not possible to discard the presence of other contaminant proteins in the gel band used to obtain the PMF, more than 79% of the total intensity of the experimental spectra is covered by the amino acid sequences obtained for each SLP (Online Resource 2). Additionally, due to the presence of sugar residues, there are some peaks corresponding to glycopeptides that cannot be assigned in these analyses. The deduced
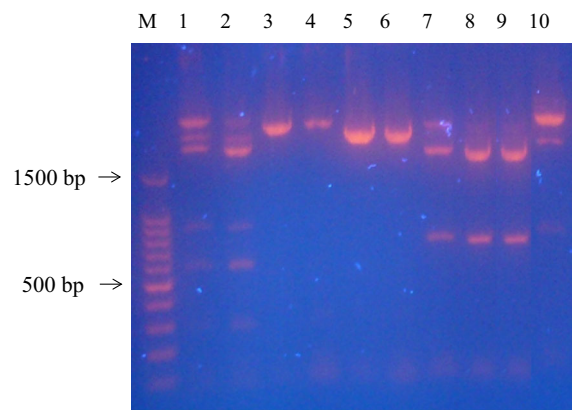
Fig. 2 Fragmentation patterns of amplicons F1/R1 (*1 L. kefiri* CIDCA 8310; *2 L. kefiri* CIDCA 8343; *3 L. kefiri* CIDCA 8314; *4 L. kefiri* CIDCA 83111; *5 L. kefiri* CIDCA 8335; *6 L. kefiri* CIDCA 83113; *7 L. kefiri* CIDCA 8321; *8 L. kefiri* CIDCA 8348; *9 L. kefiri* CIDCA 83115) and F2/R1 (*10 L. kefiri* JCM 5818) with a mix of EcoRI and NcoI. *Lane M* molecular weight marker

amino acid sequence alignment of ten SLPs is shown in Fig. 3.

The total length of the mature proteins varies from 492 to 576 amino acids, with the SLPs from *L. kefiri* CIDCA 8335 and 83113 the shortest and the SLP from *L. kefiri* CIDCA 8310 the longest polypeptides respectively. As expected, cysteine is absent in all *L. kefiri* SLPs analyzed here, and the percentages of hydrophobic and hydroxylated amino acids varies from 34.9 to 38.3% and from 24.6 to 29.2%, respectively among strains (Table 3). The amount of positively charged amino acid residues ranged from 9.5 to 10.5% and is always higher than the amount of negatively charged residues (5.7–7.3%), thus leading to a calculated isoelectric point between 9.37 and 9.60 (Table 3). All these results are similar to those reported for other *lactobacillus* SLPs (Avall-Jääskeläinen and Palva 2005; Wasko et al. 2014), and there are only slight differences between aggregative and non-aggregative strains.

All proteins start with a predicted 31-mer leader peptide for membrane translocation showing the same amino acid sequence in all the tested strains, except for the residue 27, where a threonine present in the non-aggregative strains is changed by a serine in the aggregative strains (Fig. 3). Additionally, the sequence of all the signal peptides includes the A–X–A motif that precedes the cleavage site for type I signal peptidases commonly found in Gram-positive bacteria (van Roosmalen et al. 2004). The similarities

observed between *L. kefiri* SLPs and those from *L. buchneri*, *L. parafarraginis*, *L. farraginis*, *L. otakiensis*, and *L. kisonensis* among others, correlate with the phylogenetical relationship that exists among these species of lactobacilli (Sun et al. 2015). Interestingly, protein BLAST analyses showed that there is almost a 100% identity of their predicted leader peptides with those from strains of phylogenetically related species such as *L. parakefiri*, *L. buchneri*, *L. parabuchneri*, *L. parafarraginis*, *L. farraginis*, *L. sunkii*, *L. kisonensis* and *L. otakiensis*. Additionally, the threonine to serine change at residue 27 was also

Fig. 3 Alignment of ten SLPs sequences from *L. kefiri* strains. The sequence corresponding to GWIY-f/TREG-r fragments (*blue box*) and the O-glycosylation site (*red box*) are marked. *Asterisks* indicates positions which have a single, fully conserved residue, *colon* indicates conservation between groups of strongly similar properties—scoring >0.5 in the Gonnet PAM 250 matrix, *period* indicates conservation between groups of weakly similar properties—scoring ≤0.5 in the Gonnet PAM 250 matrix

```
CIDCA8335    MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYVKVTSNKALTTDATTRNVAVNGTTAL
CIDCA83113   MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYVKVTSNKALTTDATTRNVAVNGTTAL
CIDCA8321    MKKSLKKTLFAGVAALSFVAVAGVSSSNASAKSYAKVTSNKALTTDATTRNVAVNGTTAL
CIDCA8348    MKKSLKKTLFAGVAALSFVAVAGVSSSNASAKSYAKVTSNKALTTDATTRNVAVNGTTAL
CIDCA83115   MKKSLKKTLFAGVAALSFVAVAGVSSSNASAKSYAKVTSNKALTTDATTRNVAVNGTTAL
CIDCA8310    MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYVKVTSNKALTTDATTRNVAVNGTTAL
CIDCA8343    MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYAKVTSNKALTTDATTRNVAVNGTNAL
JCM5818      MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYAKVTSNKALTSDATTRNVAVNGTNAL
CIDCA8314    MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYAKVTSNKALTTDATTRNVAVNGTNAL
CIDCA83111   MKKSLKKTLFAGVAALSFVAVAGVSSTNASAKSYAKVTSNKALTTDATTRNVAVNGTNAL
             **************************:*******.*********:************.**

CIDCA8335    YTKAGTLKGAKTVATKTTLSGLKDSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA83113   YTKAGTLKGAKTVATKTTLSGLKDSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA8321    YTKAGTLKGAKTVATKTTLAGLKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA8348    YTKAGTLKGAKTVATKTTLAGLKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA83115   YTKAGTLKGAKTVATKTTLAGLKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA8310    YTKAGTLKGAKTVATKTTLAGLKDSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA8343    YTKAGTLKGAKTVATKTTLAGLKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
JCM5818      YTKAGTLKGAKTVATKTTLASLKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA8314    YTKAGTLKGAKTVATKTTLAALKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
CIDCA83111   YTKAGTLKGAKTVATKTTLAALKNSKQGQKNFRAYRVATTNRGSVYYKVVSFDKTYRGWI
             *******************:.**:*************************************

CIDCA8335    YGGKSVTAFAGGIASFNTTTAPAAASSASASS-ASSASSTEQTTALTDAQKAATYKITKA
CIDCA83113   YGGKSVTAFAGGIASFNTTTAPAAASSASASS-ASSASSTEQTTALTDAQKAATYKITKA
CIDCA8321    YGGKSVTAFAGGIASFNTTTAPAAANSASASS-ASSASSAEQTTALTDAQKAATYKIATT
CIDCA8348    YGGKSVTAFAGGIASFNTTTAPAAANSASASS-ASSASSAEQTTALTDAQKAATYKIATT
CIDCA83115   YGGKSVTAFAGGIASFNTTTAPAAANSASASS-ASSASSAEQTTALTDAQKAATYKIATT
CIDCA8310    YGGKSVTAFAGGIASFNTTTAPAAANTSSSASSASINTQGQTTALTDAQKNATYKITKA
CIDCA8343    YGGKSVTAFAGGIASFNTTTAPAAANTSSSASSASINTQGQTTALTDAQKNATYKITKA
JCM5818      YGGKSVTAFAGGIASFNTTTAPAAAKSASASS-ASSASSTEQTTALTDAQKAATYKITKA
CIDCA8314    YGGKSVTAFAGGIASFNTTTAPAAAKSASASS-ASSASSTEQTTALTDAQKNATYKITKA
CIDCA83111   YGGKSVTAFAGGIASFNTTTAPAAAKSASASS-ASSASSTEQTTALTDAQKNATYKITKA
             ************************  .:.*.:.::  .:  ********* *****.:.:

CIDCA8335    GTANDGTATTYTYPAWTEYRKGRTVTDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
CIDCA83113   GTANDGTATTYTYPAWTEYRKGRTVTDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
CIDCA8321    GTANDGTQVTYTYPAWTEYKKGRVITDSTPYKDATFKITDQTVRTREGDLWVKIAATDSA
CIDCA8348    GTANDGTQVTYTYPAWTEYKKGRVITDSTPYKDATFKITDQTVRTREGDLWVKIAATDSA
CIDCA83115   GTANDGTQVTYTYPAWTEYKKGRVITDSTPYKDATFKITDQTVRTREGDLWVKIAATDSA
CIDCA8310    GTANDGTATTYTYPAWTEYRKGRTVTDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
CIDCA8343    GTANDGTATTYTYPAWTEYRKGRTVTDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
JCM5818      GTANDGTATTYSYPAWTEYKKGRVIDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
CIDCA8314    GTANDGTATTYTYPAWTEYRKGRTVTDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
CIDCA83111   GTANDGTATTYTYPAWTEYRKGRTVTDATPYANDTFKVTDQTTRTREGDLWVKIADTNAT
             *******  .**:*******:***.: *:*** : ***:****.****:******  *::

CIDCA8335    NGQKINGWIKYSALTAQKPTPAPAP--APIADNQVQVTLHDVYTNKDLGSFLLNKATADL
CIDCA83113   NGQKINGWIKYSALTAQKPTPAPAP--APIADNQVQVTLHDVYTNKDLGSFLLNKATADL
CIDCA8321    N-AKANGWIKFSALKVATPAPT----TTPVADNAVRINFNDPSGKQI-ATLDYVKSNAAK
CIDCA8348    N-AKANGWIKFSALKVATPAPT----TTPVADNAVRINFNDPSGKQI-ATLDYVKSNAAK
CIDCA83115   N-AKANGWIKFSALKVATPAPT----TTPVADNAVRINFNDPSGKQI-ATLDYVKSNAAK
CIDCA8310    NGQKINGWIKYSALTATDQTPAKTPAKTPVADNAVLLNFVDASGKTI-KTVTYTKANAKK
CIDCA8343    NGQKINGWIKYSALTATDQTPAKTPAKTPVADNAVLLNFVDASGKTI-KTVTYTKANAKK
JCM5818      NGQKINGWIKFSALTAQNPTPAP----APIADNAVRLNFIDSQGKTV-KSIDYVKTGAKK
CIDCA8314    NGQKINGWIKYSALTATAQTPATTP-TTPVADNAVRVTFVDASGKQVGKSFDYAKTGAAK
CIDCA83111   NGQKINGWIKYSALTATAQTPATTP-TTPVADNAVRVTFVDASGKQVGKSFDYAKTGAAK
             *  * *****:***..   .:*:      :*:***  *  :.:  *     .:   *: *
```

Fig. 3 continued

```
CIDCA8335     NA-----------------NVTIAHAVNIDPAADINNSLVALGQAHNVRGYSVSGLTDA
CIDCA83113    NA-----------------NVTIAHAVNINPAADINNSLVALGQAHNVRGYSVSGLTDA
CIDCA8321     GTTLGSLQPVKNAAGNITGYQWQLDS-STSDLQAKVDAALHGTGYAFNVQNN--------
CIDCA8348     GTTLGSLQPVKNAAGNITGYQWQLDS-STSDLQAKVDAALHGTGYAFNVQNN--------
CIDCA83115    GTTLGSLQPVKNAAGNITGYQWQLDS-STSDLQAKVDAALHGTGYAFNVQNN--------
CIDCA8310     GDFLGKLGQV--A--GTSNYAWALSSTDQSALQNAIDSALSGTGYSFDLSNT--------
CIDCA8343     GDFLGKLGQV--A--GTSNYAWALSSTDQSALQNAIDSALSGTGYSFDLSNT--------
JCM5818       GDTLGSLTGPKTG--TQDNFSWTLTKVVADDLQSKIDNALSGTGYSFNVSQY--------
CIDCA8314     GATLGTLST------VDGSYQWTLTDAQKSDLKALINTALDGTGYSFDLNNT--------
CIDCA83111    GATLGTLST------VDGSYQWTLTDAQKSDLKALINTALDGTGYSFDLNNT--------
                                      ::  :*  . *  :.::

CIDCA8335     QKSANQTAFQSASYGKNVTFNVTSVKGTSLFADKTKVLP--------GIVAN--------
CIDCA83113    QKSANQTAFQSASYGKNVTFNVTSVKGTSLFADKTKVLP--------GIVAN--------
CIDCA8321     -----QSVLAQSKTGSTI--TLSIAKGDTVYQTLTPTAYNIKTQGQA----QSLKPNPAS
CIDCA8348     -----QSVLAQSKTGSTI--TLSIAKGDTVYQTLTPTAYNIKTQGQA----QSLKPNPAS
CIDCA83115    -----QSVLAQSKTGSTI--TLSIAKGDTVYQTLTPTAYNIKTQGQA----QSLKPNPAS
CIDCA8310     ---ANQSVVAQAKTGEAI--NVTLTKGATVYQQMAAYSSDTTGSFNGGVLANPMSNIAAD
CIDCA8343     ---ANQSVVAQAKTGEAI--NVTLTKGATVYQQMAAYSSDTTGSFNGGVLANPMSNIAAD
JCM5818       -----PSTLAQAQTGSTV--SLPVTKGDVVYQTLKPYVATKDDVTS--THS--LTPNS-K
CIDCA8314     ---ANAAALAQAKTGAET--KIAVTKGSIVYQQLTPYAQTSASDTS--ITASKMSPIANT
CIDCA83111    ---ANAAALAQAKTGAET--KIAVTKGSIVYQQLTPYAQTSASDTS--ITASKMSPIANT
                 :.. .:. *      .:  .**  ::

CIDCA8335     --NNNIT----------DGSN----GAFSPAGTDYVEVNGTLEQANDLADSIAKLGTLYG
CIDCA83113    --NNNIT----------DGLN----GAFSPAGTDYVEVNGTLEQANDLADSIAKLGTLYG
CIDCA8321     VVNTKISKSVLSTKVPADTTNGT----PELDLSLLIDANGAQTN-P------------LK
CIDCA8348     VVNTKISKSVLSTKVPADTTNGT----PELDLSLLIDANGAQTN-P------------LK
CIDCA83115    VVNTKISKSVLSTKVPADTTNGT----PELDLSLLIDANGAQTN-P------------LK
CIDCA8310     AHNNQVA-AKGTATVTKDTANGLN-ADFTFDMTKLLDEKGNANT-TAIKTALNNLLAGKQ
CIDCA8343     AHNNQVA-AKGTATVTKDTANGLN-ADFTFDMTKLLDEKGNANT-TAIKTALNNLLAGKQ
JCM5818       SAGA--F-FKGSFDEGEGP------KATHIDISQLVHTDSDGKV-TDHTNS--------L
CIDCA8314     KHDS--V-YQGEVNVKQDTDNKLPVGGFTLKLSDFVNADGTFQK-DDFKDALGTTKPGFK
CIDCA83111    KNDS--V-YQGEVNVKQDTDNKLPVGGFTLKLSDFVNADGTFQK-DDFKDALGTTKPGFK

CIDCA8335     SNGQTFSGSS------------VE---------K---------AINDANLGT-IYFAKNA
CIDCA83113    SNGQTFSGSS------------VE---------K---------AINDANLGT-IYFAKNA
CIDCA8321     KTVADYVGT-----------DSTQQAKKLQAINDAVATEAAKYYAPIANLQATDLFSGSK
CIDCA8348     KTVADYVGT-----------DSTQQAKKLQAINDAVATEAAKYYAPIANLQATDLFSGSK
CIDCA83115    KTVADYVGT-----------DSTQQAKKLQAINDAVATEAAKYYAPIANLQATDLFSGSK
CIDCA8310     SDGTGSIGEDLSADVSADATKSAEKTQRVTDFNKAIAKAASQTFITPANLTAADLFSGNQ
CIDCA8343     SDGTGSIGEDLSADVSADATKSAEKTQRVTDFNKAIAKAASQTFITPANLTAADLFSGNQ
JCM5818       PTVKAYIDSVAA----------SDKDAALTSLNNTVKVYAANYYLAPANINPSDLFSGTR
CIDCA8314     TTVSNYLNYVDTTGKDKTTTSAEVKAARVQDLNNAFATAATTSYIAPANLSFTDLFSGNN
CIDCA83111    TTVSNYLNYVDTTGKDKTTTSAEVKAARVQDLNNAFATAATTSYIAPANLSFTDLFSGNN
                      .                             **:      *: .

CIDCA8335     AGTILSKDQLNALSANSTVNIVKATLSGV---------NNAYTVGSNPTFSGQYTLNVKN
CIDCA83113    AGTILSKDQLNALSANSTVNIVKATLSGV---------NNAYTVGSNPTFSGQYTLNVKN
CIDCA8321     GSSFSANDVMTYINKHSNLQTLNSPVYPVFNADGSVAHWAQFKFTAQTAYAGSFGQTAPV
CIDCA8348     GSSFSANDVMTYINKHSNLQTLNSPVYPVFNADGSVAHWAQFKFTAQTAYAGSFGQTAPV
CIDCA83115    GSSFSANDVMTYINKHSNLQTLNSPVYPVFNADGSVAHWAQFKFTAQTAYAGSFGQTAPV
CIDCA8310     GSTYTGTDALNFINSRNGLKALKSGAYPVVTVDGTVSGWYEFNLSAASANGGTFGAG-KT
CIDCA8343     GSTYTGTDALNFINSRNGLKALKSGAYPVVTVDGTVSGWYEFNLSAASANGGTFGAG-KT
JCM5818       GSAFSSQDVLNYLAKHSTLNTLKSAIYPVFNADGTVKEWDQLNLNATSANSGTFGSS-PV
CIDCA8314     GATYSRNDVANYINGKAQLKTLKSSVYPVVDSDGNA-TFNQYTFNFNSANGGTFGSG-AT
CIDCA83111    GATYSRNDVANYINGKAQLKTLKSSVYPVVDSDGNA-TFNQYTFNFNSANGGTFGSG-AT
              ..:   *  . :  .  :: :::   *              ..    : .* :

CIDCA8335     NVIYDKASTLASQFAK----------------------
CIDCA83113    NVIYDKASTLASQFAK--------------------
CIDCA8321     NVLYSYDPSTAVT-TALP---------TYSPTGYNPFA
CIDCA8348     NVLYSYDPSTAVT-TALP---------TYSPTGYNPFA
CIDCA83115    NVLYSYDPSTAVT-TALP---------TYSPTGYNPFA
CIDCA8310     QVVYNFNSTNPIKLDAFPTSVTSNSVTPF---------
CIDCA8343     QVVYNFNSTNPIKLDAFPTSVTSNSVTPF---------
JCM5818       QVVYTYGDGSAVS---DPFQNNGSSVNPLTPVPLTPLQ
CIDCA8314     RVVYTYGDPATNV--NYPTKTVSTTVSPFN--------
CIDCA83111    RVVYTYGDPATNV--NYPTKTVSTTVSPFN--------
              .*:*
```

Table 3 Amino acid composition of mature SLPs of *L. kefiri* strains

S-layer protein	MW calculated (kDa)	Length (n° of total amino acids)	pI	% Asp-Glu	% Arg-Lys	% Hydrophobic amino acids	% Amino acids with hydroxyl groups
JCM 5818	57.665	547	9.37	7.3	9.8	36.5	24.6
CIDCA 8314	59.913	568	9.55	6.9	10.6	34.9	29.2
CIDCA 83111	59.939	568	9.52	7.1	10.6	34.9	29.2
CIDCA 8310	59.974	576	9.50	6.8	9.7	36.6	27.1
CIDCA 8343	59.958	564	9.53	6.6	9.7	36.5	26.9
CIDCA 8335	51.464	492	9.53	6.3	9.5	36.9	26.6
CIDCA 83113	51.490	492	9.57	6.1	9.5	37.1	26.4
CIDCA 8321, 8348, 83115[a]	57.039	542	9.60	5.7	9.6	38.4	27.9

Asp aspartic acid, *Glu* glutamic acid, *Lys* lysine, *Arg* arginine

[a] Aggregative strains

observed in the predicted leader peptides of SLPs from the strains *L. buchneri* CD034 and NRRL B-30929, respectively.

Nevertheless, all *L. kefiri* SLPs described here share some specific characteristics with the SLPs from other lactobacilli, such as the absence of cysteines and the highest ratio of positively/negatively charged residues among all the bacterial SLPs (Avall-Jääskeläinen and Palva 2005; Hynönen and Palva 2013; Wasko et al. 2014). These properties result in the absence of intra- or inter-catenary disulfide bonds and also a high predicted isoelectric point comparable to the SLPs from other lactobacilli. Although the SLPs of lactobacilli are mainly reported as non-glycosylated proteins (Hynönen and Palva 2013), the occurrence of glycosylation as a post-translational modification was previously reported for the SLPs of twelve strains of *L. kefiri* by our group (Mobili et al. 2009b) and in the present work it has also been demonstrated for eleven new strains. In this sense, all *L. kefiri* SLPs display at least one glycosylation site located in the N-terminal region of the proteins, which has also been described in SLPs of *L. buchneri* 41021/251 (Möschl et al. 1993) and *L. buchneri* CD034 and NRRLB-30929 (Anzengruber et al. 2014). This agrees with other similarities observed at primary sequence level between SLPs belonging to *L. kefiri* and *L. buchneri* species.

The molecular weight of mature SLPs did not match with that observed in SDS-PAGE for the gel band that was analysed (differences around 10 kDa) (Tables 1, 3). This apparent discrepancy could be explained considering the additional contribution of

sugar moieties to the protein molecular weight and/or other structural characteristics that could influence the electrophoretic mobility of SLPs, even in denaturing conditions.

Most SLPs display two structural regions, i.e. the region involved in the attachment to the cell envelope and the region involved in S-layer assembly. These regions have been characterized at least for seven SLPs from of different lactobacilli strains belonging to *L. acidophilus* (Smit et al. 2001), *Lactobacillus crispatus* (Antikainen et al. 2002; Chen et al. 2009; Hu et al. 2011; Sun et al. 2013), *L. brevis* (Avall-Jääskeläinen et al. 2008) and *L. hilgardii* species (Dohm et al. 2011). The C-terminal region of the SLPs of *L. acidophilus* and *L. crispatus*, and the N-terminal region of the SLPs of *L. brevis* and *L. hilgardii* are the most conserved part of the protein and are responsible for anchoring to the cell envelope. On the other hand, the most variable part of the protein seems to be involved in the self-assembly of the SLPs monomers on the bacterial surface (Smit et al. 2001; Antikainen et al. 2002; Åvall-Jääskeläinen et al. 2008; Dohm et al. 2011; Hu et al. 2011; Sun et al. 2013). These findings, together with the absence of investigations carried out in lactobacilli species more phylogenetically related to *L. kefiri*, highlight the need for additional studies using truncated proteins to characterize these regions in the SLPs of aggregative and non-aggregative *L. kefiri* strains.

A prediction of secondary structure of the *L. kefiri* SLPs was performed using the Psipred software (Table 4). There are no major differences among the

Table 4 Secondary structure prediction of mature SLPs of *L. kefiri* strains

S-layer protein	Secondary structure		
	% α-Helices	% β-Strand	% Random coils
JCM 5818	11.6	28.20	60.20
CIDCA 8314	11.00	28.55	60.45
CIDCA 83111	11.00	27.90	61.10
CIDCA 8310	10.70	25.20	64.10
CIDCA 8343	11.00	24.90	64.10
CIDCA 8335	14.15	26.96	58.89
CIDCA 83113	15.30	28.88	55.82
CIDCA 8321	10.00	29.60	60.40
CIDCA 8348	10.00	29.60	60.40
CIDCA 83115	10.00	29.60	60.40

strains, except for SLPs from the strains *L. kefiri* CIDCA 8335 and CIDCA 83113 which showed a higher percentage of α-helix and a lower percentage of random coil than the other SLPs. Our results are similar to those obtained for the SLP of *L. buchneri* CD034 using the same software. On the other hand, no differences in β-sheet contents were observed between aggregative and non-aggregative strains, in disagreement with the results previously reported by Mobili et al. (2009a, b) using FT-IR spectroscopy. This discrepancy could be explained considering that

glycan residues (that were not included in our predictive analysis) could influence the secondary structure of the whole protein.

Our results show that, unlike the relatively high intra-strain homology observed at amino acid sequence level in the N-terminal region of *L. kefiri* SLPs, the C-terminal part of the proteins shows the most evident differences among strains. Different groups can be distinguished regarding the complete amino acid sequences of the mature SLPs of ten representative strains (Fig. 3). Some similarities found in the SLPs from non-aggregative strains allowed us to group the studied strains as follows: *L. kefiri* CIDCA 8314 and 83111; *L. kefiri* CIDCA 8310 and 8343; *L. kefiri* CIDCA 8335 and 83113; and *L. kefiri* JCM 5818, corresponding to the previously mentioned groups II, III, IV and V respectively.

It should be noted that the analysis of the similarities observed in the PMF spectra (Fig. 1), allowed us to extend this classification to all *L. kefiri* SLPs studied here, and the results were consistent with those coming from the sequencing analyses. Interestingly, the SLPs of the three aggregative strains (*L. kefiri* CIDCA 8321, 8348 and 83115) display a 100% of sequence identity, even though each strain was isolated from a different natural source (kefir grains AGK2, AGK4 and AGK1, respectively) (Table 1). This result strongly suggests that some common structural characteristics of these proteins are responsible, at least in part, for the

Table 5 Amino acid composition of N-terminal and C-terminal regions of mature SLPs of *L. kefiri* strains

S-layer protein	N-terminal				C-terminal			
	Length (n° of amino acids)	pI	% Asp-Glu	%Arg-Lys	Length (n° of amino acids)	pI	% Asp-Glu	%Arg-Lys
JCM 5818	224	10.02	5.3	13.9	323	5.47	8.6	7.1
CIDCA 8314	224	10.02	5.3	13.8	344	8.41	8.2	8.9
CIDCA 83111	224	10.02	5.3	13.8	344	7.58	8.2	8.5
CIDCA 8310	225	10	5.3	13.3	351	5.86	7.7	7.4
CIDCA 8343	225	10.05	4.9	13.3	339	5.86	7.7	7.4
CIDCA 8335	224	9.94	5.8	13.4	268	6.32	6.7	6.4
CIDCA 83113	224	9.94	5.8	13.4	268	7.02	6.3	6.4
CIDCA 8321, 8348, 83115[a]	223	10.02	5.3	13.9	319	8.52	5.9	6.6

Asp aspartic acid, *Glu* glutamic acid, *Lys* lysine, *Arg* arginine

[a] Aggregative strains

aggregative ability of the whole bacteria. However, it cannot be discarded that sugar residues bound to the polypeptide skeleton could also contribute to this surface characteristic. Moreover, other non-covalently bound exoproteome components, named as SLAPs in other lactobacilli (Johnson et al. 2015), might contribute to aggregative and non-aggregative phenotype of each strain. Further studies will be necessary to elucidate the specific SLPs domains involved in this phenomenon.

Comparing N-terminal versus C-terminal regions of *L. kefiri* SLPs, some differences were found regarding their amino acid composition (Table 5) and the sequence homology among strains (Fig. 3). Clearly, in the N-terminal region of the proteins, the percentage of positively charged amino acids is nearly 2.5-fold higher than the percentage of negatively charged residues, which correlates with their high predicted pI values. Moreover, the N-terminal region is relatively conserved among strains, showing a low intra-species variability in this portion of the SLPs. On the other hand, the C-terminal region behaves quite differently. Positively and negatively charged amino acids are almost equally distributed along this region and they appear in very similar percentages in all the SLPs analyzed. Interestingly, the C-terminal region of the SLPs harbors the major differences in the amino acid sequence among all the strains studied here.

The knowledge of the amino acid sequence of the SLPs of different *L. kefiri* strains provides relevant data not only for a better understanding of the mechanisms involved in the functionality of these proteins with exceptional physicochemical properties, but also to contribute to the development of products of biotechnological interest from safe and potentially probiotic lactic acid bacteria.

References

Antikainen J, Anton L, Sillanpää J, Korhonen TK (2002) Domains in the S-layer protein CbsA of *Lactobacillus crispatus* involved in adherence to collagens, laminin and lipoteichoic acids and in self-assembly. Mol Microbiol 46:381–394

Anzengruber J, Pabst M, Neumann L, Sekot G, Heinl S, Grabherr R, Altmann F, Messner P, Schäffer C (2014) Protein O-glucosylation in *Lactobacillus buchneri*. Glycoconj J 31:117–131. doi:10.1007/s10719-013-9505-7

Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H (2012) ExPASy: SIB bioinformatics resource portal. Nuclei Acids Res 41:W597–W603. doi:10.1093/nar/gks400

Avall-Jääskeläinen S, Palva A (2005) *Lactobacillus* surface layers and their applications. FEMS Microbiol Rev 29:511–529

Åvall-Jääskeläinen S, Hynönen U, Ilk N, Pum D, Sleytr UB, Palva A (2008) Identification and characterization of domains responsible for self-assembly and cell wall binding of the surface layer protein of *Lactobacillus brevis* ATCC 8287. BMC Microbiol 8:165. doi:10.1186/1471-2180-8-165

Boot HJ, Kolen CP, Pouwels PH (1995) Identification, cloning, and nucleotide sequence of a silent S-layer protein gene of *Lactobacillus acidophilus* ATCC 4356 which has extensive similarity with the S-layer protein gene of this species. J Bacteriol 177:7222–7230

Carasi P, Trejo FM, Pérez PF, De Antoni GL, de Serradell ML (2012) Surface proteins from *Lactobacillus kefiri* antagonizes in vitro cytotoxic effect of *Clostridium difficile* toxins. Anaerobe 18:135–142. doi:10.1016/j.anaerobe.2011.11.002

Carasi P, Ambrosis NM, De Antoni GL, Bressollier P, Urdaci MC, de Serradell ML (2014a) Adhesion properties of potentially probiotic *Lactobacillus kefiri* to gastrointestinal mucus. J Dairy Res 81:16–23. doi:10.1017/S0022029913000526

Carasi P, Díaz M, Racedo SM, De Antoni GL, Urdaci MC, de Serradell ML (2014b) Safety characterization and antimicrobial properties of kefir-isolated *Lactobacillus kefiri*. Biomed Res Int. doi:10.1155/2014/208974

Carasi P, Racedo SM, Jacquot C, Romanin DE, Serradell MA, Urdaci MC (2015) Impact of kefir derived *Lactobacillus kefiri* on the mucosal immune response and gut microbiota. J Immunol Res. doi:10.1155/2015/361604

Chen X, Chen Y, Li X, Chen N, Fang W (2009) Characterization of surface layer proteins in *Lactobacillus crispatus* isolate ZJ001. J Microbiol Biotechnol 19:1176–1183

Claus H, Akça E, Debaerdemaeker T, Evrard C, Declercq J-P, Harris JR, Schlott B, König H (2005) Molecular organization of selected prokaryotic S-layer proteins. Can J Microbiol 51:731–743

Delfederico L, Hollmann A, Martínez M, Iglesias NG, De Antoni G, Semorile L (2006) Molecular identification and typing of lactobacilli isolated from kefir grains. J Dairy Res 73:20–27

Dohm N, Petri A, Schlander M, Schlott B, König H, Claus H (2011) Molecular and biochemical properties of the S-layer protein from the wine bacterium *Lactobacillus hilgardii* B706. Arch Microbiol 193:251–261. doi:10.1007/s00203-010-0670-9

Endo A, Okada S (2007) *Lactobacillus farraginis* sp. nov. and *Lactobacillus parafarraginis* sp. nov., heterofermentative lactobacilli isolated from a compost of distilled shochu residue. Int J Syst Evol Microbiol 57:708–712

Garrote GL, Abraham AG, De Antoni GL (2001) Chemical and microbiological characterisation of kefir grains. J Dairy Res 68:639–652

Garrote GL, Delfederico L, Bibiloni R, Abraham AG, Pérez PF, Semorile L, De Antoni GL (2004) Lactobacilli isolated from kefir grains: evidence of the presence of S-layer proteins. J Dairy Res 71:222–230

Garrote GL, Serradell MA, Abraham AG, Añón MC, Fossati CA, De Antoni GL (2005) Development of an immuno-chemical method to detect *Lactobacillus kefir*. Food Agric Immunol 16:221–233

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press Inc, Totowa, NJ, pp 571–607

Gerbino E, Carasi P, Araujo-Andrade C, Tymczyszyn EE, Gómez-Zavaglia A (2015a) Role of S-layer proteins in the biosorption capacity of lead by *Lactobacillus kefir*. World J Microbiol Biotechnol 31:583–592. doi:10.1007/s11274-015-1812-7

Gerbino E, Carasi P, Mobili P, Serradell MA, Gómez-Zavaglia A (2015b) Role of S-layer proteins in bacteria. World J Microbiol Biotechnol 31:1877–1887. doi:10.1007/s11274-015-1952-9

Goh YJ, Azcárate-Peril MA, O'Flaherty S, Durmaz E, Valence F, Jardin J, Lortal S, Klaenhammer TR (2009) Development and application of a upp-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. Appl Environ Microbiol 75:3093–3105. doi:10.1128/AEM.02502-08

Golowczyc MA, Mobili P, Garrote GL, Abraham AG, De Antoni GL (2007) Protective action of *Lactobacillus kefir* carrying S-layer protein against *Salmonella enterica* serovar Enteritidis. Int J Food Microbiol 118:264–273

Golowczyc MA, Mobili P, Garrote GL, de Los Angeles Serradell M, Abraham AG, De Antoni GL (2009) Interaction between *Lactobacillus kefir* and *Saccharomyces lipolytica* isolated from kefir grains: evidence for lectin-like activity of bacterial surface proteins. J Dairy Res 76:111–116. doi:10.1017/S0022029908003749

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, López RA (2010) New bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38(Suppl): W695–W699. doi:10.1093/nar/gkq313

Hagen KE, Guan LL, Tannock GW, Korver DR, Allison GE (2005) Detection, characterization, and in vitro and in vivo expression of genes encoding S-proteins in *Lactobacillus gallinarum* strains isolated from chicken crops. Appl Environ Microbiol 71:6633–6643

Hamet MF, Londero A, Medrano M, Vercammen E, Van Hoorde K, Garrote GL, Huys G, Vandamme P, Abraham AG (2013) Application of culture-dependent and culture-independent methods for the identification of *Lactobacillus kefiranofaciens* in microbial consortia present in kefir grains. Food Microbiol 36:327–334. doi:10.1016/j.fm.2013.06.022

Heinl S, Wibberg D, Eikmeyer F, Szczepanowski R, Blom J, Linke B, Goesmann A, Grabherr R, Schwab H, Pühler A, Schlüter A (2012) Insights into the completely annotated genome of *Lactobacillus buchneri* CD034, a strain isolated from stable grass silage. J Biotechnol 161:153–166. doi:10.1016/j.jbiotec.2012.03.007

Hu S, Kong J, Sun Z, Han L, Kong W, Yang P (2011) Heterologous protein display on the cell surface of lactic acid bacteria mediated by the S-layer protein. Microb Cell Fact 10:86. doi:10.1186/1475-2859-10-86

Hynönen U, Palva A (2013) *Lactobacillus* surface layer proteins: structure, function and applications. Appl Microbiol Biotechnol 97:5225–5243. doi:10.1007/s00253-013-4962-2

Ilk N, Egelseer EM, Sleytr UB (2011) S-layer fusion proteins—construction principles and applications. Curr Opin Biotechnol 22:824–831. doi:10.1016/j.copbio.2011.05.510

Jakava-Viljanen M, Åvall-Jääskeläinen S, Messner P, Sleytr UB, Palva A (2002) Isolation of three new surface layer protein genes (slp) from *Lactobacillus brevis* ATCC 14869 and characterization of the change in their expression under aerated and anaerobic conditions. J Bacteriol 184:6786–6795

Johnson BR, Hymes J, Sanozky-Dawes R, Henriksen ED, Barrangou R, Klaenhammer TR (2015) Conserved S-layer-associated proteins revealed by exoproteomic survey of S-layer-forming lactobacilli. Appl Environ Microbiol 82:134–145. doi:10.1128/AEM.01968-15

Laemmli K (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, López R (2013) Analysis tool web services from the EMBL-EBI. Nuclei Acids Res 41:W597–W600. doi:10.1093/nar/gkt376

Messner P, Schäffer C, Kosma P (2013) Bacterial cell-envelope glycoconjugates. Adv Carbohydr Chem Biochem 69:209–272. doi:10.1016/B978-0-12-408093-5.00006-X

Mobili P, Londero A, Maria TMR, Eusébio MES, De Antoni GL, Fausto R, Gómez-Zavaglia A (2009a) Characterization of S-layer proteins of *Lactobacillus* by FTIR spectroscopy and differential scanning calorimetry. Vib Spectrosc 50:68–77. doi:10.1016/j.vibspec.2008.07.016

Mobili P, Serradell M, Trejo SA, Avilés Puigvert FX, Abraham AG, De Antoni GL (2009b) Heterogeneity of S-layer proteins from aggregating and non-aggregating *Lactobacillus kefir* strains. Antonie Van Leeuwenhoek 95:363–372. doi:10.1007/s10482-009-9322-y

Möschl A, Schäffer C, Sleytr UB, Messner P, Christian R, Schulz G (1993) Characterization of the S-layer glyco-proteins of two lactobacilli. In: Beveridge TJ, Koval SF (eds) Advances in bacterial paracrystalline surface layers. Plenum Press, New York, pp 281–284

Oki K, Kudo Y, Watanabe K (2012) *Lactobacillus saniviri* sp. nov. and *Lactobacillus senioris* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol 62:601–607. doi:10.1099/ijs.0.031658-0

Palomino MM, Waehner PM, Martin JF, Ojeda P, Malone L, Sánchez Rivas C, Prado Acosta M, Allievi MC, Ruzal SM (2016) Influence of osmotic stress on the profile and gene expression of surface layer proteins in *Lactobacillus*

acidophilus ATCC 4356. Appl Microbiol Biotechnol 100:8475–8484. doi:10.1007/s00253-016-7698-y

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786. doi:10.1038/nmeth.1701

Podlesny M, Jarocki P, Komon E, Glibowska A, Targonski Z (2011) LC–MS/MS analysis of surface layer proteins as a useful method for the identification of lactobacilli from the *Lactobacillus acidophilus* group. J Microbiol Biotechnol 21:421–429

Prado Acosta M, Ruzal SM, Cordo SM (2016) S-layer proteins from *Lactobacillus* sp. inhibit bacterial infection by blockage of DC-SIGN cell receptor. Int J Biol Macromol 92:998–1005. doi:10.1016/j.ijbiomac.2016.07.096

Ramiah K, van Reenen CA, Dicks LMT (2009) Expression of the mucus adhesion gene mub, surface layer protein slp and adhesion-like factor EF-TU of *Lactobacillus acidophilus* ATCC 4356 under digestive stress conditions, as monitored with real-time PCR. Probiotics Antimicrob Proteins 1:91–95. doi:10.1007/s12602-009-9009-8

Sára M, Sleytr UB (2000) S-layer proteins. J Bacteriol 182:859–868

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, López R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins D (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using ClustalOmega. Mol Syst Biol 7:539. doi:10.1038/msb.2011.75

Sleytr UB, Schuster B, Egelseer E-M, Pum D (2014) S-layers: principles and applications. FEMS Microbiol Rev 38:823–864. doi:10.1111/1574-6976.12063

Smit E, Oling F, Demel R, Martinez B, Pouwels PH (2001) The S-layer protein of *Lactobacillus acidophilus* ATCC 4356:

identification and characterization of domains responsible for S-protein assembly and cell wall binding. J Mol Biol 305:245–257

Sun Z, Kong J, Hu S, Kong W, Lu W, Liu W (2013) Characterization of a S-layer protein from *Lactobacillus crispatus* K313 and the domains responsible for binding to cell wall and adherence to collagen. Appl Microbiol Biotechnol 97:1941–1952. doi:10.1007/s00253-012-4044-x

Sun Z, Harris HM, McCann A et al (2015) Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. Nat Commun 6:8322. doi:10.1038/ncomms9322

van Roosmalen ML, Geukens N, Jongbloed JDH, Tjalsma H, Dubois JYF, Bron S, van Dijl JM, Anné J (2004) Type I signal peptidases of Gram-positive bacteria. Biochim Biophys Acta 1694:279–297

Waśko A, Polak-Berecka M, Kuzdraliński A, Skrzypek T (2014) Variability of S-layer proteins in *Lactobacillus helveticus* strains. Anaerobe 25:53–60. doi:10.1016/j.anaerobe.2013.11.004

Watanabe K, Fujimoto J, Tomii Y, Sasamoto M, Makino H, Kudo Y, Okada S (2009) *Lactobacillus kisonensis* sp. nov., *Lactobacillus otakiensis* sp. nov., *Lactobacillus rapi* sp. nov. and *Lactobacillus sunkii* sp. nov., heterofermentative species isolated from sunki, a traditional Japanese pickle. Int J Syst Evol Microbiol 59:754–760. doi:10.1099/ijs.0.004689-0

Zhu C, Guo G, Ma Q, Zhang F, Ma F, Xiao D, Yang X, Sun M (2016) Diversity in S-layers. Prog Biophys Mol Biol. doi:10.1016/j.pbiomolbio.2016.08.002