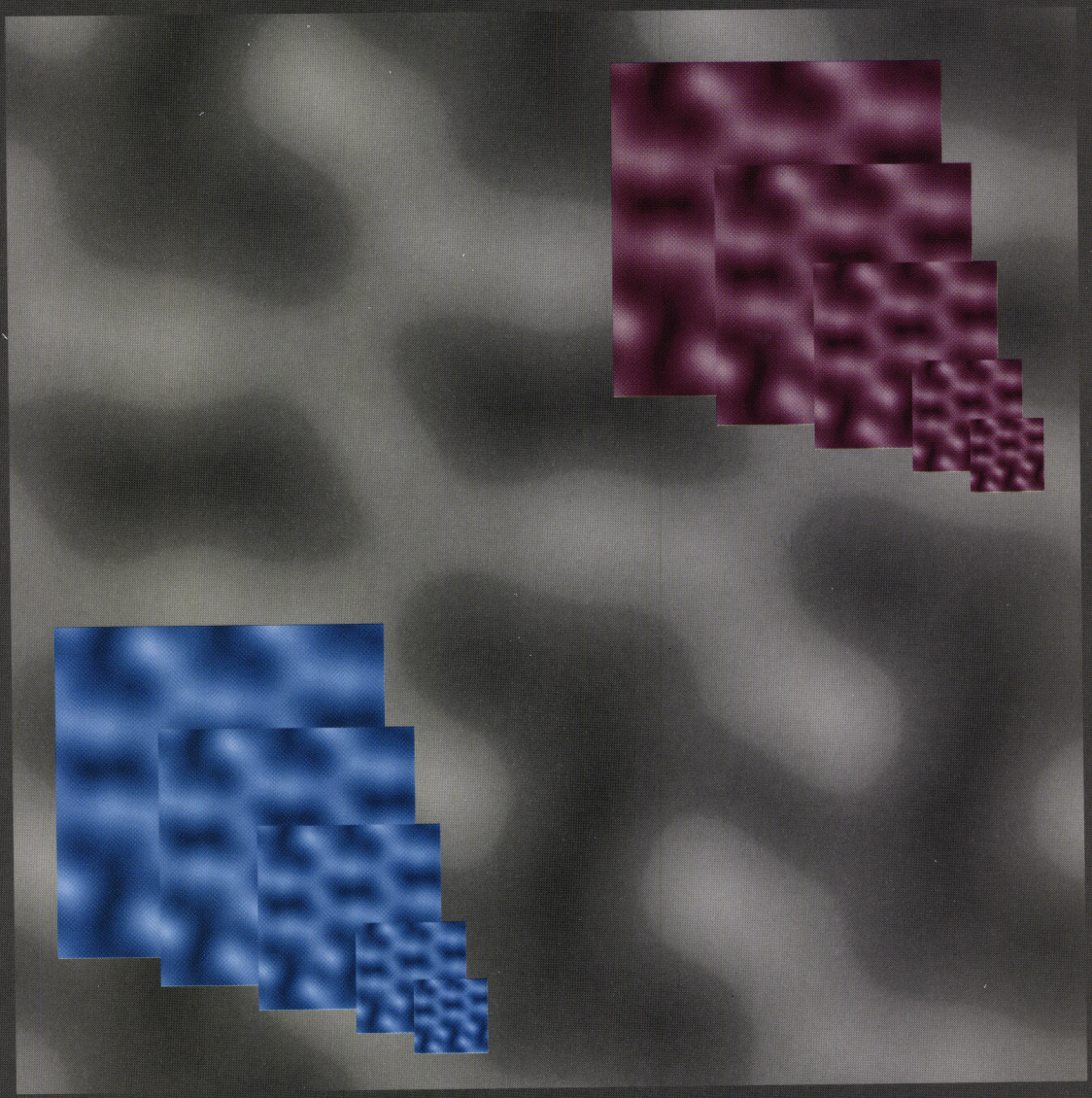
surface layer proteins

Structure-function relationship and application potential



Egbert Smit

Lactobacillus S-layer proteins

Structure-function relationship

and application potential

Front cover: From large to small. An artist's adaptation of figure 2h, chapter 2, symbolizing a scientist's search for ever greater detail.

The work described in this thesis was performed in the Molecular Genetics and Gene Technology department at the TNO Medical Biological Laboratory in Rijswijk, The Netherlands, and in the Applied Microbiology and Gene Technology department at the TNO Nutrition and Food Research Institute, Zeist, The Netherlands.

Publication of this thesis was financially supported by the TNO Nutrition and Food Research Institute, Zeist, The Netherlands.

Printed by : Ponsen & Looijen, Wageningen

ISBN : 90-9015872-3

Lactobacillus S-layer proteins

Structure-function relationship

and application potential

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit van Amsterdam op gezag van de Rector Magnificus Prof. Mr. P.F. van der Heijden ten overstaan van een door het college voor promoties ingestelde commissie, in het openbaar te verdedigen in de Aula der Universiteit

op dinsdag 25 juni 2002

te 10.00 uur

door

Egbert Smit

geboren te Eindhoven

Promotie-commissie

Promotor: Prof. Dr. P.H. Pouwels

Faculteit: Natuurwetenschappen, Wiskunde en Informatica

Leden: Prof. Dr. K. van Dam
Prof. Dr. K. Hellingwerf
Prof. Dr. J. Delcour
Dr. B. Martinez
Dr. H. Boot

Voor Renée

Contents

Abbreviations

GRAS	Generally Recognized As Safe
LAB	Lactic acid bacteria
ECM	Extracellular matrix
GSP	General secretory pathway
GIT	Gastro-intestinal tract
PG	Peptidoglycan
PS	Polysaccharide
CM	Cytoplasmic membrane
CWAP	Cell wall-associated polymer
LTA	Lipoteichoic acid
TA	Teichoic acid
LPS	Lipopolysaccharide
SLH domain	Surface layer homology domain
S_A-protein	wild type *slpA* gene product
S_{HA}-protein	S_A-protein with an N-terminal 6 histidine tag
SAN	S_A-protein N-terminal region
SAC	S_A-protein C-terminal region
SAC1	N-terminal SAC repeat
SAC2	C-terminal SAC repeat
GFP	Green fluorescent protein
HGFP	GFP with an N-terminal six histidine tag
CNBr	Cyanogenic bromide
LiCl	Lithium chloride
GHCl	Guanidinium hydrochloride
EM	Electron microscopy
CWF	Cell wall fragment
CWBD	Cell wall binding domain
SPR	Surface plasmon resonance
kDa	Kilo Dalton
Mr	Relative molecular mass
SUMS	S-layer ultrafiltration membranes
CBB	Coomassie brilliant blue

Chapter 1

General Introduction

1. Lactobacillus

1.1. Taxonomy of Lactobacillus. Lactobacilli are Gram-positive, non-spore-forming rods or coccobacilli. They are strictly fermentative, aero-tolerant or anaerobic, aciduric or acidophilic and require complex media for growth (Hammes & Vogel, 1995). Lactobacilli occur in a wide variety of carbohydrate-rich habitats including human and animal mucosal surfaces (oral and nasal cavities and gastro-intestinal and uro-genital tracts), on plant-derived material, sewage, and in fermented food products. They ferment carbohydrates and produce lactic acid as the main end product. As such lactobacilli are members of a heterogeneous group called lactic acid bacteria (LAB) which also amongst others includes the genera *Lactococcus*, *Leuconostoc* and *Pediococcus.* Of the genus *Streptococcus* only *S. thermophilus* is included in the LAB, whereas all others have been excluded since they are often pathogens (*S. pneumoniae, S. pyogenes*) and are not involved any in food or feed fermentative process.

The genus *Lactobacillus* is very heterogeneous (GC content 36-52%) and can be subdivided into three evolutionary related groups (a, b and c) based on 16S ribosomal RNA analysis namely: the *Lactobacillus* (*L.*) *delbrueckii* group (a), *L. casei-Pediococcus* group (b) and the *Leuconostoc* (c) group (Collins *et al.*, 1991; Yang & Woese, 1989).

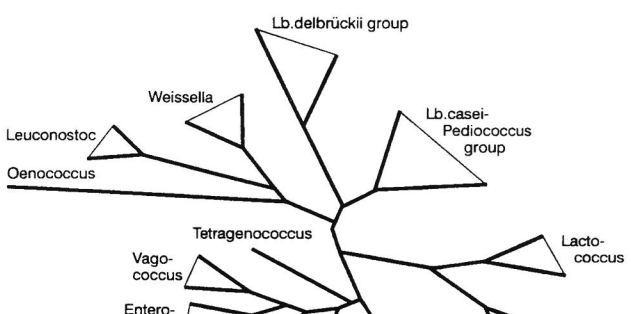


Figure 1. Schematic, unrooted phylogenetic tree of the lactic acid bacteria, including some aerobic and facultatively anaerobic Gram-positive bacteria of the low G+C subdivision. Note: evolutionary distances are approximate (Axelsson, 1998).

Figure 1 depicts the three groups in relation to other lactic acid bacteria and several other Gram-positive bacteria (Axelsson, 1998). Another division into three groups (A, B and C) was made according to the fermentation characteristics of lactobacilli: obligately

homofermentative lactobacilli (A), facultatively heterofermentative lactobacilli (B), and obligately heterofermentative lactobacilli (C) (Kandler & Weiss, 1986). *L. acidophilus*, the subject of the work described in this thesis, belongs to the *L. delbrueckii* group and is obligate homofermentative (grouping: Aa). *L. acidophilus* and several closely related species were placed in the so-called *L. acidophilus* group by Johnson, based on various criteria such as G+C content, DNA homology, cell wall sugar composition and type of lactic acid produced (Johnson *et al.*, 1980). This group consists of type A strains *L. acidophilus* (A1), *L. crispatus* (A2), *L. amylovorus* (A3), *L. gallinarum* (A4) and type B strains *L. gasseri* (B1), *L. johnsonii* (B2).

1.2. Application of lactobacilli in food and feed. Lactobacilli have since human memory been used in the preservation and fermentation of food products. Since lactobacilli are a common constituent of our daily diet many strains have been recognized as a safe ingredient and been given the Generally Recognized As Safe (=GRAS) status. Over the last twenty years, understanding of the processes yielding fermented food products has greatly increased. At present many fermented food products are produced in controlled industrial processes which make use of defined starter cultures. Beside lactobacilli such starters may also contain *S. thermophilus*, *Leuconostoc* or *Lactococcus* species. Food products made and preserved using lactobacilli, include cheese, yogurt, sour bread, salami-type sausages, pickles and animal feed such as silage (Märyä-Mäkinen & Bigret, 1998). By the production of anti-microbial compounds (lactic acid, hydrogen peroxide, bacteriocins) these bacteria limit the growth of spoilage organisms and increase the shelf-life of the product (Ouwehand, 1998). These and other secreted compounds like extracellular polysaccharides (EPS's) also often contribute to product taste and texture.

1.3. Lactobacilli and health. In addition to the use of lactobacilli for the production and conservation of many types of fermented food products, they are known to possess health-promoting properties. It is well established that consumption of certain bacteria can have beneficial effects for the host (Goldin & Gorbach, 1992). The term 'probiotic', presently used to describe the health-promoting effect that consumption of certain bacteria can have, was first defined as '*live microbial food supplements which beneficially affect the host by improving the intestinal microbial balance*' (Fuller, 1989). The probiotic activity of certain bacteria is considered amongst others to lie in their capacity to interfere with the mucosal colonization of or to remove already present unwanted bacteria, from the gastro-intestinal tract (GIT). Probiotic bacteria are able to do this through the production of antimicrobial substances, by competing for the same mucosal

attachment sites and by modulating or stimulating the immune system. Documented health effects further include increasing lactose tolerance, lowering of blood cholesterol levels and neutralizing carcinogenic compounds (Salminen *et al.*, 1998).

For a bacterium to be labeled as probiotic it should meet the following criteria: (1) production of antimicrobials, (2) resistance to gastric acid and bile salts, (3) adherence to gut epithelium and (4) persistence, at least for short periods of time, in the gastro-intestinal tract. The ability to modulate immune responses, to influence metabolic activities (cholesterol assimilation, lactase activity, vitamin production) and to interact with other intestinal residents are also important characteristics (Dunne *et al.*, 1999). From an industrial point of view criteria such as stability to preparation, storage, and delivery conditions and suitability for large-scale industrial production processes need to be met also (Kullen & Klaenhammer, 2000).

1.4. The cell envelope, cell surface properties and adherence. Bacterial cell surface properties are of extreme importance since these determine the physical interactions with the environment, e.g. adherence to solid substrates and mucosal tissues. For probiotic lactobacilli adherence to mucosal surfaces may determine whether they are able to successfully persist on or colonize these surfaces. Persistence and colonization in turn are thought to be of great importance for the effectiveness of a probiotic (Salminen *et al.*, 1996). Gram-positive bacteria (such as lactobacilli) possess a thick cell wall surrounding the cytoplasmic membrane. The main constituent of the cell wall is peptidoglycan (PG) which is a polymer of the repeating disaccharide subunit N-acetylmuramic acid-(ß1-4)-N-acetyl-glucosamine (MurNAc-GlcNAc). Individual PG chains are cross-linked via pentapeptides attached to the MurNAc moiety which results in the formation of the extensively cross-linked three-dimensional cell wall skeleton. Besides PG other compounds such as teichoic acid, lipoteichoic acid or polysaccharides are equally important cell wall constituents. These compounds are found associated non-covalently or covalently with the PG-backbone and may bind divalent cations or specific proteins that are involved in cell wall hydrolysis or bacterial adherence. Exterior of the cell wall various proteins, protein complexes (fimbriae, flagella, S-layer) and sometimes a polysaccharide capsule may be observed (Cerning, 1990; Navarre & Schneewind, 1999; Sleytr & Messner, 1988). All of these surface located molecules may mediate contact of the bacterium with the environment. In figure 2 a schematic overview of a typical Gram-positive cell wall structure is shown.

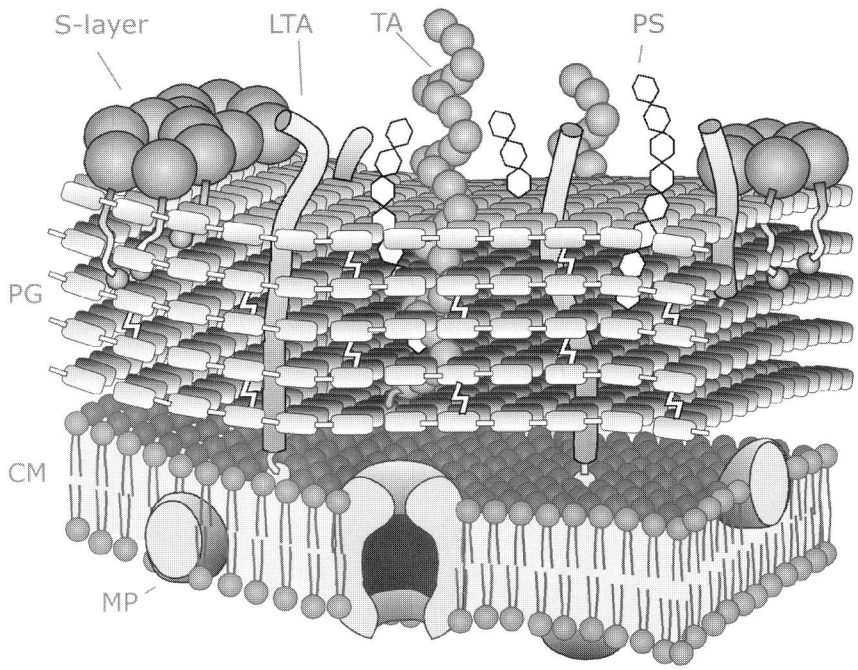


Figure 2. Artist's impression of the cell wall of a Gram-positive bacterium. The cytoplasmic membrane with embedded membrane proteins (MP) is covered by a multi-layered peptidoglycan shell decorated with polysaccharides (PS), lipoteichoic acids (LTA) and teichoic acids (TA), surrounded by an outer envelope of S-layer proteins. For the sake of clarity other cell wall-associated proteins are not depicted (Delcour *et al.*, 1999).

Bacterial adherence to host tissues is the most important initial step in host colonization (Hasty *et al.*, 1992). Adherence may occur via specific and non-specific mechanisms, whereby non-specific recognition involves macroscopic surface properties such as hydrophobicity and charge, while specific recognition requires the interaction between complementary molecules (Beachey, 1981; Millsap *et al.*, 1997; Whittaker *et al.*, 1996).

Pathogenic bacteria use different types of cell surface molecules to adhere to epithelial tissues. Non-proteinaceous (lipoteichoic acid (LTA), polysaccharides (PS)) or proteinaceous cell surface components (fimbriae, flagella, S-layers) may mediate adherence directly with target molecules on the host mucosal tissues of proteinaceous, carbohydrate or lipid nature. Alternatively, secreted molecules may serve as a bridge between bacterial and host surface molecules. An example of a Gram-positive bacterium that uses various cell surface located molecules for adherence to host tissue is *Streptococcus pyogenes*, a group A streptococcus. Tissue adherence by these pathogenic

bacteria is thought to occur via at least two steps. An initial weak and a-specific interaction is followed by a second adhesion event in which multiple tissue specific adherence factors (adhesins) are involved. Adhesins of group A streptococci include LTA and several protein factors, of which the M protein is the best known. The target molecules for these adhesins are fibronectin, collagen or carbohydrates (Cunningham, 2000). Thus adherence to epithelial tissue and persistence (or colonization) can be directly related to the cell surface properties of the bacterium. Epithelial surfaces of the GIT are covered by a thick mucus layer which provides yet another site for adherence.

One of the important prerequisites for a probiotic bacterium, as mentioned above, is the ability to adhere to and colonize mucosal surfaces. Thus it seems obvious that the molecular processes involved in adherence and colonization of the mucosa by probiotic bacteria should be investigated. Many *in vitro Lactobacillus* adherence studies have yielded valuable data. For example adherence of lactobacilli to colonic cell lines such as Caco-2 and HT-29, or isolated colonic mucus is mediated by proteinaceous factors as well as non-proteinaceous molecules such as LTA (Adlerberth *et al.*, 1996; Granato *et al.*, 1999; Neeser *et al.*, 2000). *In vitro* studies on the physico-chemical surface properties of various lactobacilli have shown that cell surface characteristics such as hydrophobicity and charge are important for non-specific adherence (Boonaert & Rouxhet, 2000; Millsap *et al.*, 1997).

Several *Lactobacillus* cell surface molecules have been described that may serve as adhesins. *L. johnsonii* La1 LTA has been shown to mediate adhesion to Caco-2 cells (Granato *et al.*, 1999). This strain has also been shown to possess carbohydrate-binding characteristics similar to those of various enteropathogenic bacteria (Neeser *et al.*, 2000). Certain cell surface proteins such as a collagen binding S-layer protein (CbsA) of *L. crispatus* JCM 5810 and the S-layer protein of *L. acidophilus* (Schneitz *et al.*, 1993; Toba *et al.*, 1995) have also been suggested to act as adhesins to sub-epithelial collagens and avian intestinal epithelial cells, respectively. The role of the *L. acidophilus* S-layer in adherence has been questioned recently by Greene *et al.*, who demonstrated that the *L. acidophilus* S-layer protein is not involved in Caco-2 adhesion (Greene & Klaenhammer, 1994). They suggested that a mixture of carbohydrate and proteinaceous factors is responsible for adhesion.

In vivo studies however provide the best indication whether a certain cell surface component is involved in adherence and colonization, but such studies remain scarce (Alander *et al.*, 1999; Jacobsen *et al.*, 1999). Mucus adhering protein (MapA), a highly basic cell surface protein from *L. reuteri* 104R, was shown to promote bacterial

aggregation and persistence of the strain in chickens (E. Satoh, personal communication). A high molecular mass mucus binding (Mub) protein of *L. reuteri* 1063 has also been implicated in *in vivo* persistence of the bacterium (Roos & Jonsson, 2002). Despite the many reports of probiotic effects there is still a need for further research to broaden the scientific base for these effects by systematically performing well-controlled *in vivo* experiments (Sanders & Klaenhammer, 2001).

1.5. Genetics of lactobacilli. In order to analyze and, where possible, exploit favorable traits of LAB the study of the genetics of LAB was initiated in the early seventies. This started with the analysis of plasmid encoded traits such as lactose utilization and proteolysis by group N streptococci (later renamed lactococci). Construction of cloning and integration vectors as well as development of efficient transformation systems have boosted LAB genetic research (Chassy & Flickinger, 1987; Luchansky *et al.*, 1988). Although work on the genetics of lactobacilli lagged behind developments in the lactococcal field, rapid progress has been made in recent years.

To date over 200 *Lactobacillus* genes, often involved in carbohydrate metabolism or proteolysis, have been isolated. Analysis of these genes has resulted in the identification and characterization of elements for the control of gene expression and targeting of gene products. Plasmid vectors for constitutive or controlled gene expression have either been adapted from existing streptococcal and lactococcal vectors or have in some cases been newly developed using *Lactobacillus* sequences (de Vos, 1999; Kuipers *et al.*, 1997; Leer, 1993; Pouwels *et al.*, 2001; Wright & Sibakov, 1998). At present the genetic modification of species like *L. casei*, *L. plantarum* and *L. reuteri*, and the expression of foreign genes on plasmid vectors or integrated into the chromosome have become routine.

Despite these rapid developments many *Lactobacillus* species remain poorly accessible, such as members from the *L. acidophilus* homology groups A and B. With respect to these latter groups Russell *et al.* (2001) recently described a system for efficient homologous recombination into the *L. acidophilus* and *L. gasseri* chromosome, which may make genetic modification of these and related strains easier (Russell & Klaenhammer, 2001). Such tools will be of great value for further characterization of genes from *Lactobacillus* species. Especially, now that the many current *Lactobacillus* genome sequencing projects will provide the sequence of numerous new genes.

2. Surface layer (S-layer) proteins

2.1. Occurrence and function. S-layers form very characteristic two-dimensional crystalline arrays covering the surface of many bacteria. They are found in the Archaea, where they are almost universally present, and in Eubacteria where they are ubiquitous. In Archaea the S-layer is often an integral and essential part of the cell envelope, determining and maintaining cell shape. For S-layers of Eubacteria, however, a general function cannot be easily given. Eubacterial S-layers can have a role in virulence (evading the host immune system, see S-layer variation), bacterial adherence (to intestinal epithelium or extracellular matrix (ECM)), and as molecular sieve or a scaffold for extracellular enzymes (Beveridge, 1994; Egelseer *et al.*, 1995; Kay & Trust, 1991; Sára & Sleytr, 1987a). S-layers, as the outermost component of the cell envelope, contribute to the physico-chemical properties of the bacterial surface and may determine the interaction of the bacterial cell with its environment.

2.2. Secretion of S-layer proteins. Several systems for the transport of bacterial proteins across the cytoplasmic membrane exist and the signal peptide/Sec-dependent general secretory pathway (GSP), or Type II secretion system, is commonly used for protein secretion both in Gram-positive and Gram-negative bacteria (Pugsley, 1993). Gram-negative bacteria further possess systems (specialized terminal branches of the GSP) for translocation of proteins across the outer membrane. S-layers are metabolically expensive products and may constitute between 10 and 15 % of total cellular protein. The S-layer on an average-sized bacterium may contain 5 x 10^5 S-protein molecules and during growth as many as 500 molecules have to be synthesized, translocated and incorporated in the S-layer lattice per second (Sleytr *et al.*, 1996). Little is yet known about S-protein secretion but many S-proteins contain an N-terminal signal peptide sequence for Sec-dependent export via the GSP. The S-proteins of the Gram-negative *Aeromonas salmonicida* and *A. hydrophila* are exported across the cytoplasmic membrane via the GSP and then across the outer membrane by an S-protein dedicated system (Boot & Pouwels, 1996). Other Gram-negatives such as *Campylobacter fetus* and *Caulobacter crescentus* employ a sec-independent Type I secretion system for the simultaneous translocation of their S-protein across the cytoplasmic and outer membrane (Awram & Smit, 1998; Thompson *et al.*, 1998).

2.3. S-layer variation. Some S-layer-carrying bacteria have shown the capacity to produce different S-proteins depending on the environmental conditions. In these cases the presence of two or more S-protein encoding genes was demonstrated. S-layer proteins of pathogenic bacteria are important antigenic determinants and for these

bacteria it is especially favorable to be able to vary their cell surface characteristics according to environmental demands. The best described case of S-layer variation is that of *C. fetus,* a human pathogen. S-protein variation in this bacterium is triggered by the host immune response (Wang *et al.*, 1993). This species was shown to actively recombine S-protein gene sequences resulting in the production of a different S-protein each time. As many as eight variant S-protein genes have been detected in *C. fetus* (Dworkin & Blaser, 1996). Apart from *C. fetus* several other bacteria have shown S-layer variation such as *B. stearothermophilus* PV72/p6 which upon changes in environmental oxygen concentration synthesizes a variant S-protein and *B. stearothermophilus* ATCC 12980 that produces an S-protein variant at elevated cultivation temperature (Egelseer *et al.*, 2001; Schuster *et al.*, 1997). *L. acidophilus* ATCC 4356 was found to possess a silent S-protein gene designated *slpB*. Although the protein encoded by this gene has never been detected in cells, genetic analyses have shown that via inversion of a chromosomal DNA segment the *slpB* gene is placed downstream of the *slp* promoter (Boot *et al.*, 1995). Whether the *slpB* gene product plays a role in S-layer formation or S-layer variation remains unanswered. The presence of an expressed as well as a silent S-layer gene was confirmed for several other S-layer-carrying lactobacilli including *L. crispatus*, *L. amylovorus* and *L. gallinarum* (Boot *et al.*, 1996).

2.4. Double S-layers. Several bacteria have been reported to possess a double S-layer structure at their cell surface. *Brevibacillus brevis* 47 for example possesses a middle wall protein (MWP) that is attached to the cell surface. The MWP layer is covered by an outer wall protein (OWP) layer which does not possess a known cell wall anchoring domain and may be attached directly to the MWP layer. Also for several *Clostridium difficile* strains it was shown that these bacteria possess two dominant cell surface proteins (Calabi *et al.*, 2001). These proteins also formed two superimposed layers (Cerquetti *et al.*, 2000). Interestingly, the largest of the (putative) S-proteins showed a relatedness to the *B. subtilis* amidases LytC and LytB and displayed amidase activity (Calabi *et al.*, 2001). Finally, *B. anthracis* simultaneously produces two S-layer proteins, Sap and EA1. These two proteins probably form a mosaic structure since EA1 is the major and SAP the minor S-layer component (Mesnage *et al.*, 1997). A superimposed S-layer structure would require more or less equal amounts of both proteins.

2.5. S-protein genes and analysis of S-protein primary structure. Over the past 20 years many S-protein encoding genes have been cloned and sequenced. Here we will only deal with Eubacterial S-proteins which have been studied structurally and/or functionally in some detail. Table 1 lists these proteins and several important characteristics concerning their structure and function. Many attempts have been made

to extract relevant structural information from S-proteins by alignment of S-protein primary amino acid sequences of members from closely related but also more distantly related bacteria. These studies included genes from *Aeromonas*, *Caulobacter*, *Bacillus* and *Lactobacillus* species, but in general identities above 25% have not been detected (Kuen & Lubitz, 1996; Lemaire *et al.*, 1998). There is one important exception, namely the Surface-Layer Homology (SLH) domain for cell surface attachment, that was first found in the S-proteins MWP of *Brevibacillus brevis* and *Thermoanaerobacter kivui*. Over the past five years many additional S-protein genes have been sequenced and alignment of these sequences, especially from closely related species, may provide new conserved S-protein motifs. The overall conclusion, however, remains that the S-proteins of bacteria from different genera vary greatly in primary amino acid sequence and size, and that homology is weak. Since no universal S-protein motifs have been found it has been difficult to clarify S-layer structure and function (Engelhardt & Peters, 1998).

3. S-layer structure

3.1. S-layer ultrastructure. S-layers possess a highly regular structure and usually consist of a single (glyco-) protein subunit varying in size from 40 to 200 kilo Dalton (kDa). The structure of S-layers is traditionally studied by electron microscopy (EM) of freeze-etched whole bacteria or negatively stained or frozen preparations consisting of isolated or reassembled S-layers. S-layer lattices show either oblique (p1, p2), square (p4) or hexagonal lattice symmetry (p3, p6) (figure 3). Depending on the lattice type the morphological units constituting the crystal lattice consists of one, two, three, four or six subunits and have center-to-center spacing varying between 3 and 30 nm. S-layers are highly porous structures resulting in a porosity of between 30 and 70 % as estimated from electron micrographs. Pores are uniformly shaped and their size may vary from 2 to 8 nm.

More structural detail may be obtained from high-resolution EM images using computer image reconstruction techniques (Baumeister, 1987; Beveridge & Graham, 1991). Two-dimensional and three-dimensional (3D) reconstructions have been published over the years for many Archaeal and Eubacterial S-layer lattices (Beveridge, 1994). These studies have shown that S-layers differ greatly in mass distribution and subunit-subunit linkages and that no common morphological organization exists (Sleytr *et al.*, 1996). There have also been reports that S-proteins share one common feature, the presence of two morphological domains (Baumeister *et al.*, 1989; Saxton & Baumeister, 1986). The maximum structural detail reached in EM studies, typically is between 10 and 20 Å. Atomic structures of S-proteins or S-protein domains have not been obtained to date.

Table 1. Origin and characteristics of S-proteins available in Genbank sequence database

Organism	Strain	Gene/ protein	Protein size[1]	Lattice[2]	3D[4]	SLH[5]	Accession no.
Bacteria							
Aeromonas hydrophila	TF7	*vapA*	467 (19)	s	+	-	L37348
Aeromonas salmonicida	A449/A450	*ahsA*	502 (21)	s	+	-	M64655
Bacillus anthracis	Sterne	*eag*	862 (29)	o	-	3	X99724
	Sterne 9131	*sap*	814 (29)	h	-	3	Z36946
Bacillus stearothermophilus	PV72/p6	*sbsA*	1228 (30)	h	+	-	X71092
	PV72/p2	*sbsB*	920 (31)	o	-	3	X98095
	ATCC 12980	*sbsC*	1099 (30)	o	-	-	AF055578
	ATCC 12980	*sbsD*	903 (30)	n.d.	-	-	AF228338
Bacillus licheniformis	NM105	*slap*	874(29)	s	-	3	U38842
Bacillus pseudofirmus	OF4	*slpA*	931 (31)	o	-	3	AF242295
Bacillus sphaericus	2362		1176 (30)	s	-	3	M28361
	P-1		1252 (30)	s	+	3	A45814
	CCM2177	*sbpA*	1268 (30)	s	-	3	AF211170
Bacillus thuringiensis galleriae	NRRL 4045	*slpA*	821 (29)	o	-	3	AJ249446
mexicanensis#	TkD2-14		823 (29)	n.d.	-	3	D86346
finitimus#	CTC		816 (29)	n.d.		3	AJ012290
Brevibacillus brevis	47	OWP	1004 (24)	h	-	-	M14238
	47	MWP	1053 (23)	h	+	2	M19115
	HPD31	HWP	1087 (23/53)	h	-	2	D90050
Clostridium thermocellum	NCIB 10682	*slpA*	1036 (26)	o	-	1.5	U79117
Clostridium difficile	R7404	*slpA*	340 / 374 (24)	s + h	-	-	AJ300677
	R8366	*slpA*	341 / 415 (24)	s + h	-	-	AJ300676
	C253	*slpA*	719 (24)	s + h	-	-	AJ291709
	79685	*slpA*	717 (24)	s + h	-	-	AY004256
Caulobacter crescentus	ATCC 19089	*rsaA*	1026	h	+	-	M84760
Campylobacter fetus	23D	*sapA*	933	h	+	-	L15800
Campylobacter rectus	ATCC 33238	*Slp*	1361	h	-	-	AB001876
	ATCC 314	*crs*	1361	h	-	-	AF010143
Corynebacterium glutamicum	BI15	PS2	510 (30)	o	-	-	X69103
Deinococcus radiodurans		*hpi*	1036 (31-59)	h	+	-	M17895
Lactobacillus brevis	ATCC 8287	*slpA*	465 (30)	s	-	-	Z14250
*Lactobacillus acidophilus**	ATCC 4356	*slpA*	444 (31)	o	-	-	X71412
	JCM 1038#	HAP50	488 (30)	n.d.	-	-	AF250229
*Lactobacillus crispatus**	JCM 5810	*cbsA*	440 (30)	o	-	-	AF001313
	LMG 12003#	*slpNA*	458 (30)	n.d.	-	-	AF253043
Lactobacillus helveticus[3]	CNRZ 892	*slpH*	439 (30)	o	-	-	P38059
Rickettsia conorii#	Malish 7	*ompB*	1655	n.d.	-	-	Q9KKA3
Rickettsia prowazekii	BREINL	*spaP*	1643	s	-	-	M37647
Rickettsia japonica#	YH	*ompB*	1656	n.d.	-	-	AB003681
Rickettsia rickettsii#	R	*ompB*	1654	n.d.	-	-	X16353
Rickettsia typhi#	Wilmington	*slpT*	1645	n.d.	-	-	L04661
Serratia marcescens	SR41	*slaA*	1004	n.d.	-	-	AB007124
Thermoanaerobacter kivui	DSM 2030		762 (26)	o	+	2.5	M31069
Thermus thermophilus	HB8	P100	917 (27)	h	+	1	X57333
Archaea							
Haloferax volcanii			827 (34)	h	+	-	M62816
Methanococcus jannaschii	ATCC 43067		558 (28)	h	-	-	Q58232
Methanococcus thermolithotrophicus	DSM 2095		559 (28)	n.d.	-	-	AJ308554
Methanococcus voltea	DSM 1537		576 (23)	h	-	-	M59200
Methanococcus vannielii	DSM 1224		566 (28)	h	-	-	AJ308553
Methanosarcina mazei	S-6		1673	n.d.	-	-	AF394229
Methanothermus fervidus	DSM 2088		593	h	-	-	P27373
Methanothermus sociabilis	DSM 3496		593	h	-	-	P27374
Staphylothermus marinus	F1		1524 (26)	n.d.	-	-	S68553

annotation of protein function is based on sequence homology only
* silent slp gene(s) present in the bacterium
1: value between brackets indicates the length of the N-terminal secretion signal
2: o = oblique lattice symmetry, s = square, h = hexagonal
3: multiple sequences with > 97% identity have been deposited in the database but are not listed here
4: 3D indicates that the three dimensional ultrastructure (not atomic structure) of this S-layer has been determined by electron microscopical techniques
5: SLH = Surface Layer Homology domain for cell wall attachment
N.d.: not determined

3.2. Extraction and reassembly *in vitro*. In an S-layer two types of non-covalent interactions can be defined: (i) subunit-subunit interactions (responsible for formation of the two-dimensional crystalline network) and (ii) subunit-cell envelope interactions (anchoring the S-layer to the bacterial cell surface) (figure 4). The interactions between subunits are usually stronger than those between a subunit and cell surface as was shown by many extraction and dissociation experiments. These interactions may occur through electrostatic forces (via divalent cations or directly between polar groups), hydrogen bonding and/or hydrophobic interactions.

The S-layer of Gram-positive bacteria can often be efficiently extracted from whole cells or from purified cell wall fragments with hydrogen-bond breaking agents (i.e. urea, guanidinium hydrochloride) that disassemble the S-layer into S-protein monomers. The use of lithium chloride (LiCl) proved efficient for extraction of the *L. helveticus* S-protein (Lortal *et al.*, 1992). For Gram-negative bacteria less harsh treatments with acidic buffers, metal-chelating agents, detergents, or even de-ionized water, may be used. Many S-layers depend on divalent cations for structural stability of the S-layer and/or cell surface attachment. Upon removal of the extracting agent monomers are often able to reassemble into S-layer sheets identical to those observed on the bacterium.

3.3. Crystallization domains of S-layer proteins. Within an S-layer multiple subunit-subunit interactions exist resulting in the formation of the S-layer, an interconnected two-dimensional crystalline network. Based on EM studies and biochemical data it was suggested that many S-proteins consist of two morphological domains which could correspond to two functional domains, i.e. a domain for cell surface attachment and one for formation of the crystalline network. Conclusive evidence that morphological and functional domains coincide has not been reported. For the *A. salmonicida* and *A. hydrophila* S-proteins it has been shown that the two functions cannot be attributed to separate domains. In the case of S-proteins for which the cell wall binding domain has been identified it was deduced, but not shown that the remaining part of the protein would be responsible for crystal formation. No study to date has demonstrated unequivocally the presence of a distinct crystallization domain (Deblaere, 1995; Hastie & Brinton, 1979). In chapters 5 and 7 we will describe the characterization of the crystallization domain of the *L. acidophilus* S_A-protein and the implications for the S-layer field in general.

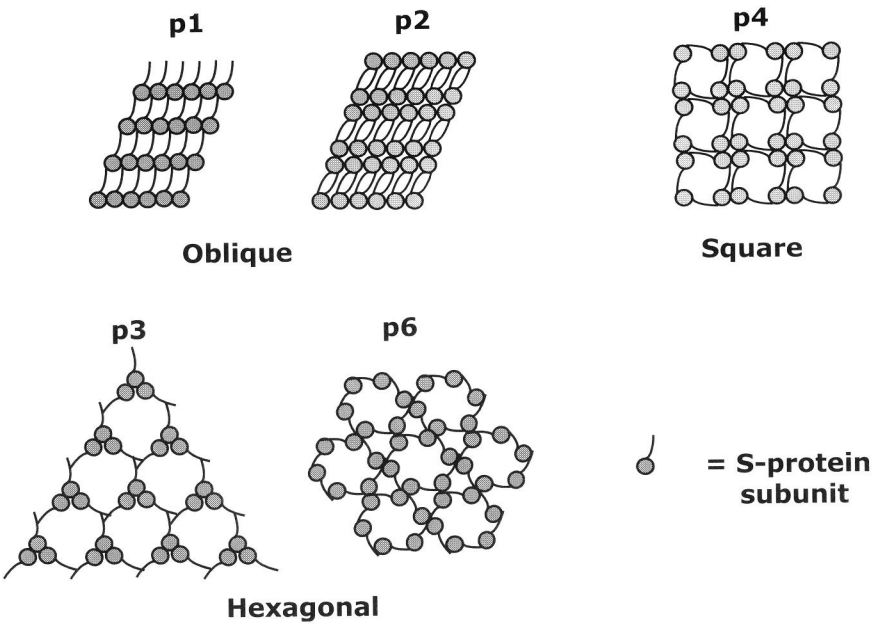


Figure 3. (Top) S-layer lattice types according to the International Tables for X-Ray Crystallography. Adapted from Sleytr *et al.* (1996).

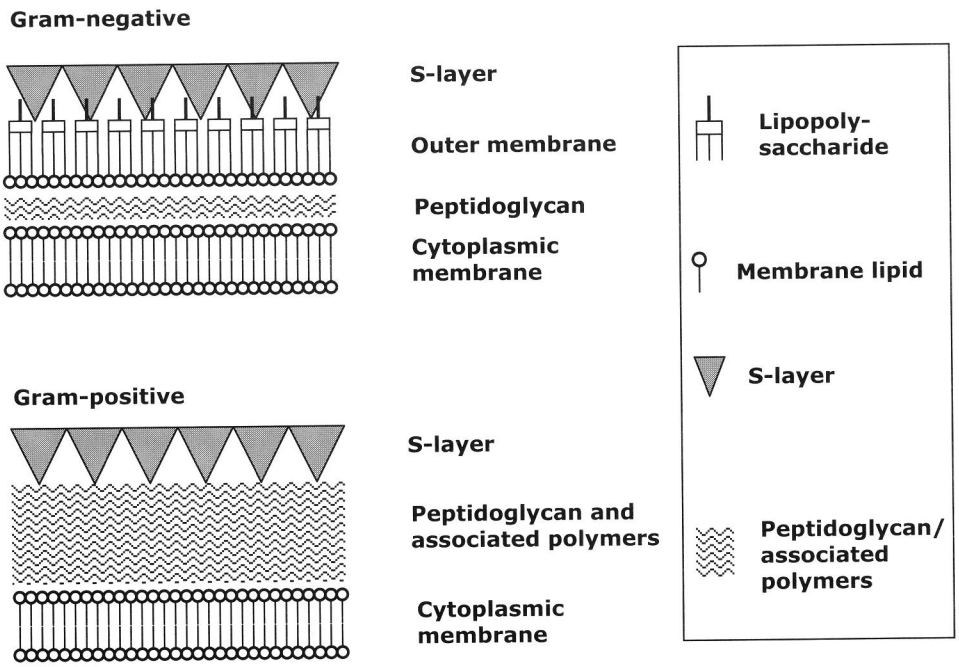


Figure 4. Schematic cross-section of the Gram-negative and Gram-positive cell envelope. Adapted from Sleytr *et al.* (1996).

4. Cell surface attachment of S-proteins

Bacterial proteins are associated with several different cellular compartments/cell envelope structures depending on the type of bacterium, i.e. Gram-positive or Gram-negative, which possess quite different cell envelope structures. Targeting of these proteins occurs by covalently or non-covalently anchoring to the cell envelope through a variety of mechanisms involving specific sequence motifs and receptor molecules (Navarre & Schneewind, 1999; Pugsley, 1993). Proteins may be inserted in the cytoplasmic or outer membranes through one or more transmembrane domains (TMD)'s as is the case for integral membrane proteins. Proteins may also be anchored covalently to the cytoplasmic membrane via a mechanism involving N-acyl-diglyceride modification of a cysteine residue directly downstream of the signal sequence. This mechanism is found in bacterial lipoproteins such as the major lipoprotein (Lpp) of *E. coli* and PrtM of *L. lactis* (Pugsley, 1993).

Anchoring of proteins to PG or other cell wall components may involve various mechanisms. Proteins can e.g. be covalently attached to the PG via a cell wall C-terminal sorting signal consisting of a charged cytoplasmic tail, a hydrophobic membrane spanning region and a conserved sequence motif (LPXTG). Cleavage between the threonine and glycine residues of the LPXTG motif is followed by covalent coupling of the threonine to the PG backbone (Navarre & Schneewind, 1999). *Staphylococcus aureus* protein A is coupled to the cell wall via this mechanism.

Non-covalent coupling of proteins to the cell wall may be mediated by different repeated motifs such as choline-binding domains found in the *Streptococcus pneumoniae* cell surface proteins PspA and LytA. These repeats mediate binding of PspA and LytA to choline moieties of the *S. pneumoniae* PG-associated LTA (Garcia *et al.*, 1998). Other repetitive cell wall binding motifs include those observed in many cell wall hydrolases such as *L. lactis* AcmA and bacterial S-layer proteins.

4.1. Gram-positive cell surface attachment

4.1.1. The SLH domain. Although S-layer proteins of different species do not share much homology in their primary amino acid sequences one important exception exists, the SLH domain. Conservation of the SLH sequence was first observed for the S-protein of *Thermoanaerobacter kivui* and the middle wall protein (MWP) of *Brevibacillus brevis*. Since then these domains have been found in many S-proteins and other extracellular

bacterial proteins (Peters *et al.*, 1989). SLH domains consist of a conserved sequence of about 55 amino acids, are modular, and can be found in 1 to 3 copies in the various proteins. The amino acid composition of SLH domains shows the presence of amino acids generally thought to be important for carbohydrate-binding such as arginine, lysine and aromatic amino acids. Secondary structure predictions have shown that the SLH domain probably has a helix-turn-helix structure (Engelhardt & Peters, 1998). SLH domains were initially suggested to associate through non-covalent interactions with peptidoglycan (PG) (Lupas *et al.*, 1994) but recent reports have shown that SLH domains found in the S-proteins of several *Bacillus* species bind to teichuronic acids or polysaccharides rather than PG (Egelseer *et al.*, 1998; Mesnage *et al.*, 1999b).

Recently, SLH-mediated cell wall attachment of *B. anthracis* S-proteins, Sap and Eag, was demonstrated to completely depend on pyruvylation of a cell wall polysaccharide covalently attached to PG. The authors further showed that such a mechanism probably exists in several related bacteria such as *B. cereus*, *B. thuringiensis* and the distantly related *Thermotoga maritima* (Mesnage *et al.*, 2000). The SbsB protein of *B. stearothermophilus* PV72 contains three SLH domains that were shown to bind to a PG-associated polymer consisting of N-acetyl-glucosamine and N-acetyl-mannosamine substituted with pyruvic acid, supporting the findings of Mesnage *et al.* (Sára, 2001).
It is unknown whether all SLH domain-containing bacterial cell surface proteins attach to PG-associated polymers via the same mechanism. The Sap, Eag and SbsB S-proteins possess three copies while other S-proteins contain 1 or 2 copies of the SLH domain, which could suggest that they employ different modes of attachment. The S-protein of *Thermus thermophilus* for example was shown to bind to peptidoglycan through one SLH domain (Olabarría, 1996). Since *T. thermophilus* is far removed from *Bacillus* in an evolutionary sense it is possible that this bacterium uses an altogether different mechanism for SLH-mediated surface attachment.

4.1.2. *Bacillus stearothermophilus* SbsA and SbsC proteins. The cell wall binding (CWB) domain of the S-layer proteins SbsA and SbsC of *B. stearothermophilus* PV72 and ATCC 12980, respectively, consists of a region of around 230 amino acids near the N-terminus of the proteins which are about 85 % identical. They do not show homology to SLH domains such as those of the SbsB protein from the PV72 strain nor to any other S-protein CWB domain. In analogy with the SLH domain, the CWB domain of SbsA and SbsC showed an amino acid composition characteristic of lectins and other carbohydrate-binding proteins or domains (Jarosch *et al.*, 2000). Both proteins were shown to specifically bind via their CWB domain to an anionic cell wall polysaccharide fraction

(classified as teichuronic acid) consisting of glucose, N-acetylglucosamine and N-acetylmannuronic acid.

4.1.3. *Lactobacillus* S-proteins. The C-terminal region (~120 amino acid residues) of the S-proteins from *L. acidophilus* and related species was implicated in cell wall binding since it is highly conserved in these proteins (Boot *et al.*, 1995). It is a strongly hydrophilic region and shows an abundance of basic amino acids. It does not show homology to the S-protein from *L. brevis* suggesting that this protein uses a different sequence for cell wall attachment. Cell wall attachment of *Lactobacillus* S-layer proteins will be discussed in detail in chapters 2, 3 and 7.

4.1.4. *Corynebacterium glutamicum* S-protein. In the S-protein PS2 of *Corynebacterium glutamicum* a distinct CWB domain was identified. The S-layer of this bacterium cannot be extracted with hydrogen-bond breaking agents, but can be stripped from the cell surface with detergents, indicating that hydrophobic interactions are involved in cell wall attachment. It has been proposed that a domain comprising the C-terminal 21, mainly hydrophobic, amino acids interacts with a hydrophobic cell wall layer consisting of mycolic acids, which are a characteristic component of the *Corynebacterium* cell wall (Chami *et al.*, 1997).

4.2. Gram-negative cell surface attachment

4.2.1. Lipopolysaccharide-binding domains. Cell wall attachment of S-protein from Gram-negative eubacteria has been studied for the S-proteins from *C. fetus*, a human pathogen, *A. salmonicida* and *A. hydrophila*, both fish pathogens, and *C. crescentus*, a bacterium occurring in fresh water (Blaser *et al.*, 1988; Dooley, 1989; Dooley *et al.*, 1988; Stahl *et al.*, 1992). Many *in vitro* reconstitution experiments have been performed with the S-proteins of these bacteria, which showed a requirement for divalent cations (usually Ca^{2+}) for proper attachment. The S-proteins of these bacteria attach to the lipopolysaccharides (LPS's) of the outer membrane. The *C. fetus* and *C. crescentus* S-proteins attach with their N-terminal regions and the *A. salmonicida* and *A. hydrophila* proteins with their C-terminal regions to O-polysaccharide side chains of the LPS (Noonan & Trust, 1997; Walker *et al.*, 1992). The O-polysaccharide side chain of the *C. crescentus* LPS was shown to be a heterogeneous polymer of N-acetyl-dideoxy-aminohexose.

5. Biotechnological application of S-layer proteins

Soon after their discovery S-layers were recognized as ideal systems for various biotechnological applications. In table 2 S-layer-based applications and the exploited S-layer/S-protein properties are summarized.

Table 2. Properties and applications of S-layers

Application	Form	Property used	Reference
Ultrafiltration membranes	APS	Isoporosity / antifouling	(Sára & Sleytr, 1987b)
Dipstick immuno assays	APS	chemical modification of S-protein	(Breitwieser *et al.*, 1996)
Affinity purification of antibodies	APS	chemical modification of S-protein	(Sleytr *et al.*, 1997a)
Conjugate vaccines / immunotherapy - tumor glycan antigens - bacterial glycan antigens - allergens	APS	low immunogenicity, adjuvanticity, chemical modification of S-protein	(Sleytr *et al.*, 1997a)
Bacterial vaccine carriers	LB/ rBG	Production of modified S-protein, adjuvanticity of AP and bacteria	(Lubitz *et al.*, 1999; Pouwels *et al.*, 2001; Simon *et al.*, 2001)
Heterologous protein production and secretion	LB	*slp* expression, secretion systems	(Kahala & Palva, 1999)
Planar supported lipid membranes - electro-physical sensors - biomimetics	API	stabilization of lipid layers	(Pum & Sleytr, 1996; Sleytr *et al.*, 1997b)
Nanopatterning	API	Regular assembly at solid interfaces	(Pum & Sleytr, 1996)

AP: S-layer Assembly Products; APS: AP crystallized in Solution; API: AP crystallized at Interfaces; rBG: recombinant Bacterial Ghosts; LB: Live Bacteria

Since S-proteins are efficiently produced and exported, their expression and secretion signals may be used to synthesize and export non-native proteins. This was indeed done for *L. brevis* and recently a recombinant protein production/secretion and purification system based on the S-protein of *C. crescentus* was commercialized (Bingle *et al.*, 1997; Savijoki *et al.*, 1997).

Application of S-layers and S-layer carrying bacteria in vaccine technology has been the subject of many studies over recent years. S-protein assembly products, when conjugated with specific antigens, elicit good immune responses and were shown to possess intrinsic adjuvant properties. Besides applications based on isolated and recrystallized S-proteins, live S-layer-carrying bacteria have also been the subject of recent vaccine research.

Especially, GRAS organisms have been studied as whole bacterial vaccine carriers in recent years. Efficient secretion of a viral glycoprotein by *Caulobacter crescentus* was used to develop a subunit vaccine for immunization of salmonid fish against infection by hematopoietic necrosis virus (Simon *et al.*, 2001).

Lactobacilli, as common food fermentation organisms and inhabitants of the human GIT, also have GRAS status and have been under investigation as oral vaccine carriers (Medaglini *et al.*, 1998; Mercenier *et al.*, 2000; Pouwels *et al.*, 1998). Lactobacilli offer additional advantages as vaccine carrier, such as intrinsic adjuvant and probiotic properties as well as the ability to colonize host mucosal surfaces (Pouwels *et al.*, 1998). Several *Lactobacillus* antigen production and targeting systems in combination with several immunization strategies have been developed. Antigens can be produced in cytoplasmic, secreted or cell surface coupled form. Immunization strategies include triggering of the mucosal immune system via the nasal or oral routes (Grangette *et al.*, 2001; Pouwels *et al.*, 2001).

Results of immunization trials have shown that serum as well as mucosal immune responses can be elicited via both the nasal and oral route by intracellularly produced antigen (Grangette *et al.*, 2001; Shaw *et al.*, 2000). Much is expected of the application of S-layer-carrying lactobacilli as oral vaccine carriers. S-layer proteins occur in over 10^5 copies at the surface of the cell, i.e. in much higher numbers than other cell surface proteins. Thus, cell-surface display of antigens using S-layer proteins may result in a higher antigen dose and thus stronger immune responses. Chapter 4 describes the development of epitope surface display using the S-layer of *L. acidophilus*.

Various other properties of S-protein assembly products have been exploited. One example are the S-layer ultrafiltration membranes (SUMS) in which the uniform pore size is used for filtration purposes. SUMS have also proven useful in the construction of biosensors consisting of SUMS to which for example enzymes, antibody fragments or antigens were coupled. The antifouling properties (i.e. low adsorption of proteins and other compounds) and high density coupling possibilities of SUMS have resulted in the construction of biosensors with excellent sensitivity and high signal-to-noise ratios.

Reassembly of S-proteins at specific solid surfaces such as gold, glass or mica is also used in biomimetics and nano-patterning in which reassembled S-layer either function as stabilizing structures or as templates for the patterned deposition metal atoms on solid supports (Sleytr *et al.*, 1997b; Pum & Sleytr, 1996). This property is also exploited in the production of highly stable and versatile SPR sensor chip coatings (Chapter 6).

6. Outline of this thesis

The work in this thesis describes the structure-function relationship of the S-layer protein of *L. acidophilus* ATCC 4356. Due to its high abundance and cell surface location this S-protein may be involved, directly or indirectly, in bacterial adhesion, an important selection criterion for probiotic bacteria. Understanding of the structure of the *L. acidophilus* S-layer protein (S_A-protein) may also provide opportunities for application of S-layer carrying lactobacilli in vaccine research.

In **Chapter 2** the domain organization of the *L. acidophilus* S_A-protein is outlined. S_A-protein possesses a C-terminal cell wall binding domain (SAC) consisting of a tandemly repeated sequence. SAC is conserved in S-proteins of related lactobacilli as well as some *Lactobacillus* extracellular proteinases. Using electron crystallographic techniques we determined the projected two-dimensional structure of S_A-protein and that of the N-terminal crystallization domain (SAN) and demonstrated that they form identical two-dimensional crystals.

Chapter 3 describes the functional analysis of the SAC domain and its two constituent repeats, SAC1 and SAC2. We show that the SAC2 repeat can be deleted entirely without seriously affecting domain function and that SAC2 itself does not show any cell wall binding properties. Chemical extraction of isolated cell wall fragments shows that cell wall-bound teichoic acid most likely is the receptor for SAC.

Chapter 4 describes the insertion mutagenesis of the S_A-protein and the surface presentation of the myc epitope in *L. acidophilus*. S_A-protein has a central region in which additional sequences can be inserted without affecting protein crystallization. This region proved suitable for S-layer surface presentation of epitopes *in vivo* and this system has potential for application of *Lactobacillus* as carrier component in vaccines.

In **Chapter 5** the structural organization of the S_A-protein crystallization domain, SAN, is investigated. Deletion analysis shows that the domain can be divided into two subdomains that are able to form stable dimers. The boundary between the two subdomains is located in the surface exposed region identified previously. Extensive S_A-protein sequence alignment, followed by secondary structural analysis, provided a further refinement of the model of the structural organization. The SAN domain consists of several conserved sequence motifs alternating with variable regions. Conservation of these motifs in a *L. helveticus* extracellular proteinase (PrtY) shows that this proteinase is

related to S-proteins and that the conserved motifs most likely represent sequences important for monomer structural integrity.

In **chapter 6** the *in vitro* immobilization of S_A-protein on solid surfaces as determined by surface plasmon resonance is described. We show that S_A-protein can be immobilized permanently on a gold surface indicating crystal formation at the water-gold interface. Modification of functional groups in the S_A-protein does not affect the capacity for permanent immobilization, suggesting that these groups are not important for protein-protein interaction and protein-gold interaction. S_A-protein-coated gold chips were used for the development of a highly stable biosensor coating.

Finally, **Chapter 7** summarizes our findings and discusses their implications for the S-layer field and possibilities for future research.

Chapter 2

The S-layer protein of *Lactobacillus acidophilus* ATCC 4356: Identification and characterization of domains responsible for S-protein assembly and cell wall binding

Egbert Smit, Frank Oling, Rudy Demel, Beatriz Martínez, and Peter H. Pouwels

J. Mol. Biol. (2001) **305**, 245-257

Summary

L. acidophilus, like many other bacteria, harbors a surface layer consisting of a protein (S_A-protein) of 43 kDa. S_A-protein could be readily extracted and crystallized *in vitro* into large crystalline patches on lipid monolayers with a net negative charge but not on lipids with a net neutral charge. Reconstruction of the S-layer from two-dimensional crystals grown on dioleoylphosphatidyl serine (DOPS) indicated an oblique lattice with unit cell dimensions (a = 118 Å; b = 53 Å, and γ = 102°), resembling those determined for the S-layer of *L. helveticus* ATCC 12046. Sequence comparison of S_A-protein with S-proteins from *L. helveticus*, *L. crispatus* and the S-proteins encoded by the silent S-protein genes from *L. acidophilus* and *L. crispatus* suggested the presence of two domains, one comprising the N-terminal two-third (SAN), and another made up of the C-terminal one-third (SAC) of S_A-protein. The sequence of the N-terminal domains is variable, while that of C-terminal domain is highly conserved in the S-proteins of these organisms and contains a tandem repeat. Proteolytic digestion of S_A-protein showed that SAN was protease resistant, suggesting a compact structure. SAC was rapidly degraded by proteases and therefore has a more susceptible structure. DNA sequences encoding SAN or Green Fluorescent Protein fused to SAC (GFP-SAC) were efficiently expressed in *E. coli*. Purified SAN could crystallize into mono- and multi-layered crystals with the same lattice parameters as found for authentic S_A-protein. A calculated S_A-protein minus SAN density-difference map revealed the probable location, in projection, of the SAC domain, which is missing in the truncated SAN peptide. The GFP-SAC fusion product was shown to bind to the surface of *L. acidophilus*, *L. helveticus and L. crispatus* cells from which the S-layer had been removed, but not to non-stripped cells or *L. casei*.

Introduction

Crystalline surface (S-)layers are cell envelope structures that can be found in many Gram-positive and Gram-negative bacteria, and in *Archaea*, in which they are almost universally present. S-layers exist as two-dimensional regular arrays completely covering the bacterial surface. They are composed of only one (glyco-) protein subunit, the S-protein, varying in size from 40 to 200 kilo Dalton (kDa), and are assembled into arrays with either oblique, square or hexagonal symmetry (Sleytr & Messner, 1983). S-layers are entropy-driven assembly systems in which all information for crystallization into a regular array resides within the monomers. Under appropriate conditions monomers can reassemble in solution or on solid supports into S-layer-like structures with the authentic symmetry pattern (Sleytr & Glauert, 1975). Biological functions of S-layers include (i) determination and maintenance of cell shape, (ii) cell adhesion and infection (iii) molecular sieve and (iv) scaffold for enzymes (Beveridge, 1994), (Brechtel *et al.*, 1999; Engelhardt & Peters, 1998; Kay & Trust, 1991; Sára & Sleytr, 1987a).

Most S-proteins are acidic (pI 4 to 6) and contain about 40% hydrophobic amino acids but very few if any sulphur-containing amino acids. S-proteins differ markedly in their primary structures and only very low levels of homology have been found between S-proteins from different species. From secondary structure data of S-proteins known to date it was concluded that S-proteins are in general composed of β-pleated sheets with some α-helices. How these secondary structure components are organized into tertiary structures is known in detail only in few instances (for a review, see Engelhardt & Peters, 1998).

Electron microscopy and computer image analysis procedures have been applied to obtain structural information of S-layers down to the nanometer range. These ultra structural studies have shown that all S-proteins analyzed so far share a common feature, the presence of two morphological domains (Baumeister *et al.*, 1989; Saxton & Baumeister, 1986). Biochemical and electron microscopic studies have indicated that the two domains might have different functions (Chu *et al.*, 1991; Egelseer *et al.*, 1996; Thomas *et al.*, 1992). Evidence for a role of one of the domains in interaction with cell wall components was provided by studies showing the occurrence of multiple copies of a conserved sequence, the S-layer homology region (SLH), in S-proteins from different species (Lupas *et al.*, 1994).

Sequence repeats have also been implied in cell wall attachment of extracellular enzymes, e.g. proteinases PrtB and PrtH from *Lactobacillus* species (Siezen, 1999).

Binding studies have confirmed the importance of SLH containing domains in cell wall binding (Lemaire *et al.*, 1998; Mesnage *et al.*, 1999b; Olabarria, 1996). In contrast, there is no conclusive evidence that a single domain determines the interaction between S-protein monomers.

S-layers have been found in several *Lactobacillus* species (Johnson *et al.*, 1980; Masuda, 1992; Masuda & Kawata, 1983). Although the function of S-proteins in lactobacilli is not known, they may play an important role in cell adhesion (Toba *et al.*, 1995). S-layer proteins of *Lactobacillus* species are among the smallest S-proteins known and are different from other S-proteins because they are highly basic (pI >9) instead of acidic proteins. *Lactobacillus acidophilus* ATCC 4356 possesses an S-layer composed of a single ~43 kDa hydrophobic, non-glycosylated protein, S_A-protein (Boot *et al.*, 1993). S_A-protein shows considerable similarity to the putative product of *slpB* encoded by a silent S-protein gene of *L. acidophilus,* and to S-proteins from *Lactobacillus helveticus* and *Lactobacillus crispatus*, but not to other S-proteins (Boot *et al.*, 1996; Boot *et al.*, 1995). The C-terminal one-third of the molecules shows 70%-90% similarity. Far less homology was observed for the N-terminal two-third of the S-proteins, <40% (Boot *et al.*, 1993; Callegari *et al.*, 1998). These results suggested that the N-terminal and C-terminal parts of these S-proteins may constitute different domains.

This work describes the 2D structure of the crystalline S-layer of *L. acidophilus* ATCC 4356 and the identification in S_A-protein of two distinct domains. The first domain, SAN, which comprises the N-terminal two-third of the protein, was identified as the crystallization domain since it could reassemble *in vitro* in the absence of the C-terminal domain. The C-terminal one-third of S_A-protein, SAC, was found to be the cell wall binding domain, since it could bind specifically to *L. acidophilus* cells stripped of S-protein in the absence of the N-terminal domain. Furthermore, by calculating a S_A-protein minus SAN density-difference map based on the 2D structures of S_A-protein and SAN, we were able to determine the probable location of the SAC domain in the wild type crystal.

Materials and Methods

Bacterial strains and growth conditions. *Lactobacillus acidophilus* ATCC 4356 and *L. casei* ATCC 393 were cultivated anaerobically in MRS broth (Difco) at 37°C. Generally, 500 ml of pre-warmed MRS medium was inoculated 1:100 (v/v) with an over night (ON) culture and bacteria were harvested during the exponential growth phase (OD_{695} ~ 0.7),

unless indicated otherwise. *Escherichia coli* M15 [pREP4] (Qiagen) and *E. coli* DH5α (Phabagen, The Netherlands) were cultivated aerobically in Luria Bertani medium (LB medium) at 37°C. When appropriate, media were supplemented with 1,5% agarose and for *E. coli* with ampicillin (100 µg/ml) and/or kanamycin (25 µg/ml).

Extraction of S_A-protein and preparation of S_A-protein self-assembly products. S_A-protein was extracted from *L. acidophilus* after harvesting of cells by centrifugation (10,000 x g, 20 min, 4°C). The cell pellet was washed once with 1 volume of chilled distilled water (w/v), extracted with a 0.1 volume of 1 M lithium chloride (LiCl) solution for 30 min at room temperature and centrifuged. The cell pellet was then extracted with a 0.1 volume of 5 M LiCl for 1 h at room temperature and centrifuged once more. Both LiCl extracts were filtered through an 0.2 µm filter and exhaustively dialyzed against 50 mM Tris-HCl (pH 7.5). The dialyzed extracts were centrifuged (20,000 x g, 20 min, 4°C) and dialyzate pellet (DP) and dialyzate supernatant (DS) were kept at 4°C until further use.

Crystallization of S_A-protein at phospholipid and air-water interfaces. Lipid monolayers were formed on a 1.4 ml sub-phase (50 mM Tris-HCl, pH 7.5) at an initial surface pressure of around 45 mN/m (Demel, 1994). Crystallization at lipid surfaces was performed with the DS S_A-protein fraction concentrated to 700 µg/ml. S_A-protein solution was added to the sub-phase through an injection hole at a final protein concentration of 50 µg/ml. Crystallization was performed at dioleoylphosphatidyl-serine (DOPS), dioleoylphosphatidyl ethanolamine (DOPE), dioleoylphosphatidyl glycerol (DOPG), dipalmitoylphosphatidyl choline (DPPC) and dipalmitoylphosphatidyl serine (DPPS) monolayers. S_A-protein was also allowed to crystallize at the air-water interface.

Electron microscopy and computer image processing. Samples for electron microscopy were obtained from lipid layer crystallization experiments by depositing carbon coated or formvar-carbon coated electron microscopical grids on the lipid surface and horizontally lifting them after 1 min. Special care was taken to maintain the same orientation of the grid during sampling and electron microscopy. All other samples were applied directly to the grids. Samples were washed once with AD and then stained with 2% uranyl acetate for 2 min. Screening of grids for S_A-protein or SAN crystals was performed with a Philips CM10 transmission electron microscope operated at 80 kV. Electron microscopy for structural analysis was performed with a Philips CM120 Transmission Electron Microscope operated at 120 kV. Micrographs were recorded under low dose conditions with a slow scan CCD camera (GATAN type 794) with a scanning step size of 24 µm (corresponding to a pixel size of 5 Å at the specimen level) at a

nominal magnification of 45,235 times. CCD image analyses were performed with a modified Imagic software package (Keegstra and Brisson, unpublished results).

Proteolytic and chemical degradation of S_A-protein and N-terminal sequencing. Cyanogenic bromide (CNBr) digestion of S_A-protein was performed for 1 h at room temperature in 1 ml 6M guanidinium hydrochloride (GHCl), 0.2 M hydrochloric acid (HCl) containing 1 mg of protein and 50 mg of CNBr. After digestion the solution was dialyzed against a large volume of 50 mM Tris-HCl (pH 7.5). S_A-protein (~1 mg) was proteolytically digested for 1 h at 37°C with 10 µg trypsin (Sigma), 5 µg *Bacillus globigii* protease (Lys-C, Boehringer), 10µg chymotrypsin (Sigma) in 25 mM Tris-HCl (pH 8) or 10µg *Staphylococcus aureus* V8 protease (V8 protease, Sigma) in 25 mM ammonium carbonate (pH 7.8). Digestion was arrested by boiling for 10 min. *L. acidophilus* cells from 1 ml log phase culture were washed with 50 mM Tris-HCl, pH 7,5 and resuspended in 100 µl digestion buffer containing either 5 µg trypsin or 5 µg Lys-C. The suspension was incubated for 1 h at 37 °C followed by 10 min boiling to inactivate the enzymes. Digestion products were analyzed by SDS-PAGE. N-terminal sequencing by Edman degradation was performed on peptides that were separated by SDS-PAGE, blotted to PVDF (Millipore) in CAPS/NaOH buffer and stained with Coomassie brilliant blue R-250.

Analysis of primary amino acid sequences: multiple sequence alignment and secondary structure prediction. Multiple sequence alignments of derived amino acid sequences of *slpA* and *slpB* (*L. acidophilus* ATCC 4356, Genbank X71412 and X89376), *slpH* (*L. helveticus* CNRZ 892, X91199), *cbsA* (*L. crispatus* JCM 5810, AF001313), *prtH* (*L. helveticus* CNRZ 32, AF133727), *prtB* (*L. delbrueckii* subsp. *bulgaricus* CNRZ 397, L48487) and the HAP50 protein (*L. acidophilus* JCM 1038, AF250229) and analysis of hydrophilicity pattern were performed with the Genetics Computer Group (Madison, Wisconsin) sequence analysis software. Alignments for secondary structure prediction were performed as described by Rost and Sander (Rost, 1996; Rost & Sander, 1993).

PCR amplification and cloning of two S_A-protein fragments in *E. coli*. DNA fragments encoding the S_A-protein N-terminal two-third (SAN: S_A-protein N-terminal fragment) and C-terminal one-third (SAC: S_A-protein C-terminal fragment) of mature S_A-protein were amplified by PCR. Oligonucleotides CEA1 (5'-GGG GGG ATC CGG TAC CGC TAC TAC TAT TAA CGC AAG TTC-3', forward) and SAN1 (5'-CCC CGA ATT CAA GCT TTT ATT AAA TTC TCT TGC TTA GCT GGC TAC-3', reverse) were used to amplify the SAN fragment and oligonucleotides SAC1 (5'-GAT CGG ATC CCT CGA GAT GCA CAA CGC ATA CTA CTA CGA-3', forward) and CEA2 (5'-CCC GGA TCC AAG CTT ATC GAA TAT CAG AAG ATC CTA TT-3', reverse) were used to amplify the SAC fragment with plasmid pTA10S (a

pBR322 derivative containing a DNA fragment encoding mature S_A-protein) as template. eGFP was amplified from pEGFP-C1 (Stratagene) using oligonucleotides GFPBAA (5'-GGG GGG ATC CAT GGT GAG CAA GGG CGA GC-3', forward) and GFP-STOPHX (5'-CCC CTC GAG AAG CTT TTA CTT GTA CAG CTC GTC CAT GCC-3', reverse) or GFPXrev2 (5'-CCC CCT CGA GCT TGT ACA GCT CGT CCA TGC C-3', reverse). Restriction endonuclease sites were introduced for *Bam*HI (5' of SAN and eGFP), *Xho*I (5' of SAC and 3' of eGFP) and *Hin*dIII sites (3' of SAN, SAC and eGFP) to facilitate cloning. PCR was performed in triplicate, reaction mixtures were pooled, purified from gel and digested with appropriate restriction endonucleases. The SAN and eGFP fragments were ligated directly in pQE30dXN (a derivative of pQE30 (Qiagen) from which the *Xho*I and *Nco*I sites were deleted) and the SAC fragment was ligated with the eGFP fragment in the same vector generating an in-frame GFP-SAC fusion. In this way the expression plasmids pHSAN, pHGFPSAC and pHGFP were obtained.

Production and purification of HSAN, HGFP-SAC and HGFP. The HSAN, HGFP-SAC and HGFP proteins were produced as follows. *E. coli* M15 containing the expression plasmids were grown ON and inoculated 1:50 (v/v) into 50 ml fresh pre-warmed medium and incubated at 37°C. At an OD_{695} of 0.7 IPTG was added to a final concentration of 1 mM and incubation was continued for another hour at 37 °C (for HSAN and HGFP) or at 25 °C (for HGFP-SAC). Following induction, cells were collected by centrifugation (3,000 x g, 30 min, 4°C) and frozen at –20°C. HSAN protein was purified in the presence of 6 M GHCl and HGFP-SAC and HGFP were purified under native conditions according to the instructions of the supplier. To minimize proteolytic degradation of HGFP-SAC 0.1 mM phenyl methyl sulfonyl fluoride (PMSF) was included in the expression medium and all purification buffers. After purification HSAN-containing fractions were dialyzed ON at 4°C against 50 mM Tris-HCl, pH 7.5. For HGFP-SAC and HGFP elution buffer was exchanged for 50 mM Tris-HCl, pH 7.5 using Econo-Pac 10DG columns (BioRad).

Binding of HGFP-SAC, HGFP and HSAN to *L. acidophilus, L. helveticus, L. crispatus* and *L. casei*. 100 ml cultures of *L. acidophilus, L. helveticus, L. crispatus* and *L. crispatus* were grown to an OD_{695} of 0.7 and harvested (3,200 x g, 25 min, 4°C). The pellets were resuspended in 10 ml 20 % glycerol and frozen in 1 ml aliquots at -80 °C. Prior to the binding assay one aliquot of each strain was extracted twice with 0.1 volume of 5 M LiCl for 1 hour at room temperature and one aliquot was left untreated. LiCl extracted and untreated cells were washed 3 times with 1 ml 50 mM Tris-HCl, pH 7.5 and finally resuspended in 0.5 ml Tris-HCl buffer, pH 7.5. 50 µl of cell suspension was mixed with 50 µl of purified HGFP-SAC (1 mg/ml), HGFP (1 mg/ml) or HSAN (0,25

mg/ml) and incubated for 1 h at RT. After incubation the cells were collected by centrifugation and washed 3 times with 50 mM Tris-HCl, pH 7.5 (the last wash supplemented with 0.2 µg/ml propidium iodide). Cells were analyzed for the presence of cell-bound HGFP-SAC, HGFP or HSAN by immuno-fluorescence or fluorescence microscopy and SDS-PAGE followed by Western blot analysis.

DNA manipulation. Plasmid DNA was isolated using the alkaline lysis procedure described by (Birnboim & Doly, 1979) or using plasmid Tip-20 or Tip-100 columns (Qiagen) according to the instruction of the supplier. Restriction endonucleases, T4 DNA polymerase and T4 DNA ligase were used according to the instructions of the manufacturers (Gibco BRL Life Technologies or New England Biolabs). Separation of DNA fragments by agarose gel electrophoresis, ligation of DNA fragments and bacterial transformation were performed as described by (Sambrook *et al.*, 1989). DNA fragments were isolated from agarose gels with a Qiaquick Gel Extraction Kit (Qiagen). PCR was performed with SuperTaq DNA polymerase according to standard procedures. All PCR products were checked by DNA sequencing.

Protein analysis. Protein samples were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) performed according to (Laemmli, 1970) using a minigel system (BioRad) or Phastsystem (Pharmacia). Proteins were visualized by Coomassie Brilliant Blue staining. S_A-protein was transferred to nitrocellulose membranes and detected with a mouse polyclonal antiserum, as described (Boot *et al.*, 1993). Histidine-tagged proteins were detected with anti RGS-His mouse polyclonal antiserum (Qiagen). Protein concentrations were determined in triplicate using BSA as a standard (Lowry, 1951).

Results

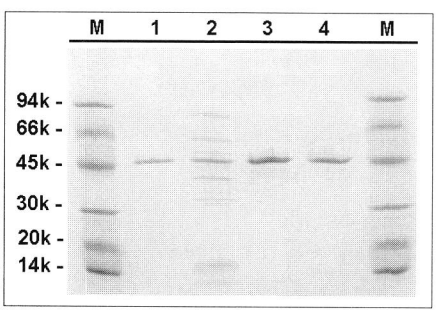


Figure 1. SDS-PAGE of dialyzed S_A-protein LiCl extracts from *L. acidophilus* (1) 1 M LiCl extract dialyzate supernatant fraction; (2) 1 M LiCl extract dialyzate pellet fraction; (3) 5 M LiCl extract dialyzate supernatant fraction; (4) 5 M LiCl extract dialyzate pellet fraction; M = protein marker (marker sizes in Dalton are indicated in the left margin).

Extraction of S_A-protein and analysis by SDS-PAGE. It has been known for a long time that bacteria of the *Lactobacillus acidophilus* A group harbor an S-layer (Johnson *et al.*, 1980; Masuda & Kawata, 1983) but the 2D structure of the S-layer of these bacteria has not previously been investigated. Using freeze-etch electron microscopy, we observed that the S-layer of this strain was similar in appearance to that of *L. helveticus* which has an oblique lattice structure (Boot, 1996; Lortal *et al.*, 1992). To perform a 2D-structure analysis of the S-protein of the *L. acidophilus* type strain, S_A-protein was extracted from whole cells by a two-step procedure, using 1 M LiCl and 5M LiCl consecutively. With 1 M LiCl several *L. acidophilus* surface-associated proteins were released including some S_A-protein. Most of the S_A-protein however remained on the cell surface and could be extracted in a highly pure form with 5 M LiCl (figure 1).

***In vitro* crystallization of S_A-protein in solution and on lipid monolayers.** S_A-protein extracted from cells with 5 M LiCl aggregated readily upon removal of the salt by dialysis forming a white precipitate. Analysis of this precipitate by EM showed that it was composed exclusively of double or multiple-layered sheets and cylinders. Monolayered sheets were not observed under these conditions. Comparison of the 5 M LiCl extract dialyzate supernatant (DS) and pellet (DP) fractions by SDS-PAGE showed that a large fraction of S_A-protein remained in solution. Following injection of this soluble S_A-protein fraction below a phospholipid monolayer, the protein was found to easily crystallize at the phospholipid surface. The S_A-protein formed crystalline sheets covering a surface of several square cm on DOPS (figure 2a) and on DPPS, DOPE and DOPG monolayers (data not shown). However, on DPPC monolayers only small crystalline patches were formed. Analysis and comparison of power spectra of the crystals formed on the different lipids demonstrated that S_A-protein crystallized with the same orientation in all cases. Comparison of the power spectra of S_A-protein crystals formed on DOPS with that of the S-layer found on *L. acidophilus* cells showed that the S-layer was oriented with its inner face towards the lipid head groups. On an air-water interface S_A-protein formed only small crystals. From the power spectra it was concluded that they had the opposite orientation with respect to the grid compared to crystals on phospholipid monolayers. These results suggest that the outer face of the S-protein is relatively hydrophobic while the inner face is more hydrophilic.

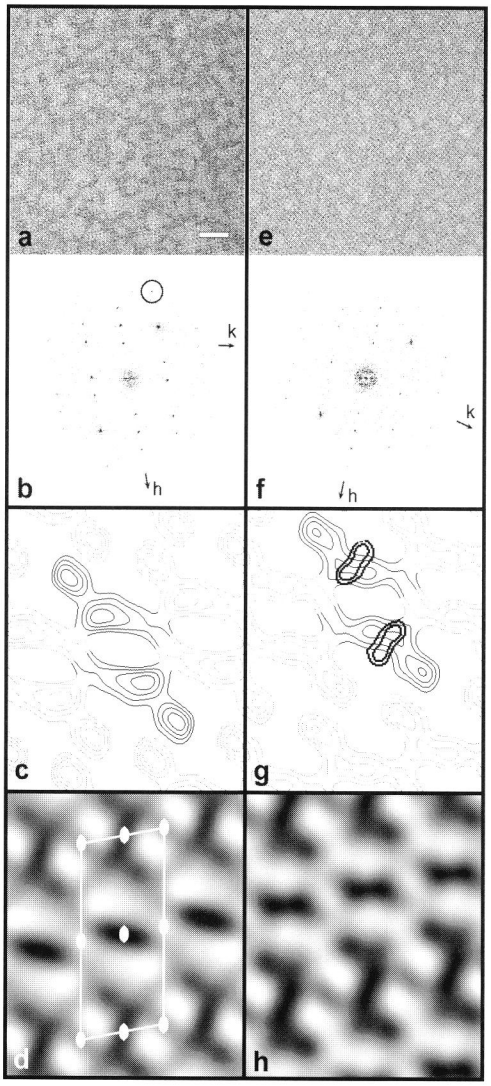


Figure 2. Electron image analysis of the 2D crystals formed by *L. acidophilus* ATCC 4356 S_A-protein (a to d) and by the C-terminally truncated peptide SAN (e to h). Low-dose micrograph of a negatively stained crystalline area of S_A-protein on DOPS (a), and SAN (e) deposited on a carbon support film. Power spectrum of a micrograph of a negatively stained S_A-protein (b) and SAN (f) 2D crystal, showing the oblique reciprocal lattice. The reciprocal lattice axes are indicated by h and k. A peak at 1/22.6 $Å^{-1}$ resolution is circled (b). Contour plot of the S_A-protein (c) and SAN (g) crystal at around 17 Å resolution. One repeating feature formed by two-fold related motifs is emphasized. Two two-fold related S_A-protein minus SAN density-difference peaks are indicated in bold character in (g). 2D projection density map of S_A-protein (d) and SAN (h). The unit cell is drawn in (d) (a=118 Å, b=53 Å, γ =102°), together with the p2 symmetry elements. The crystal shows rows of two-fold related motifs along the smallest unit cell direction.

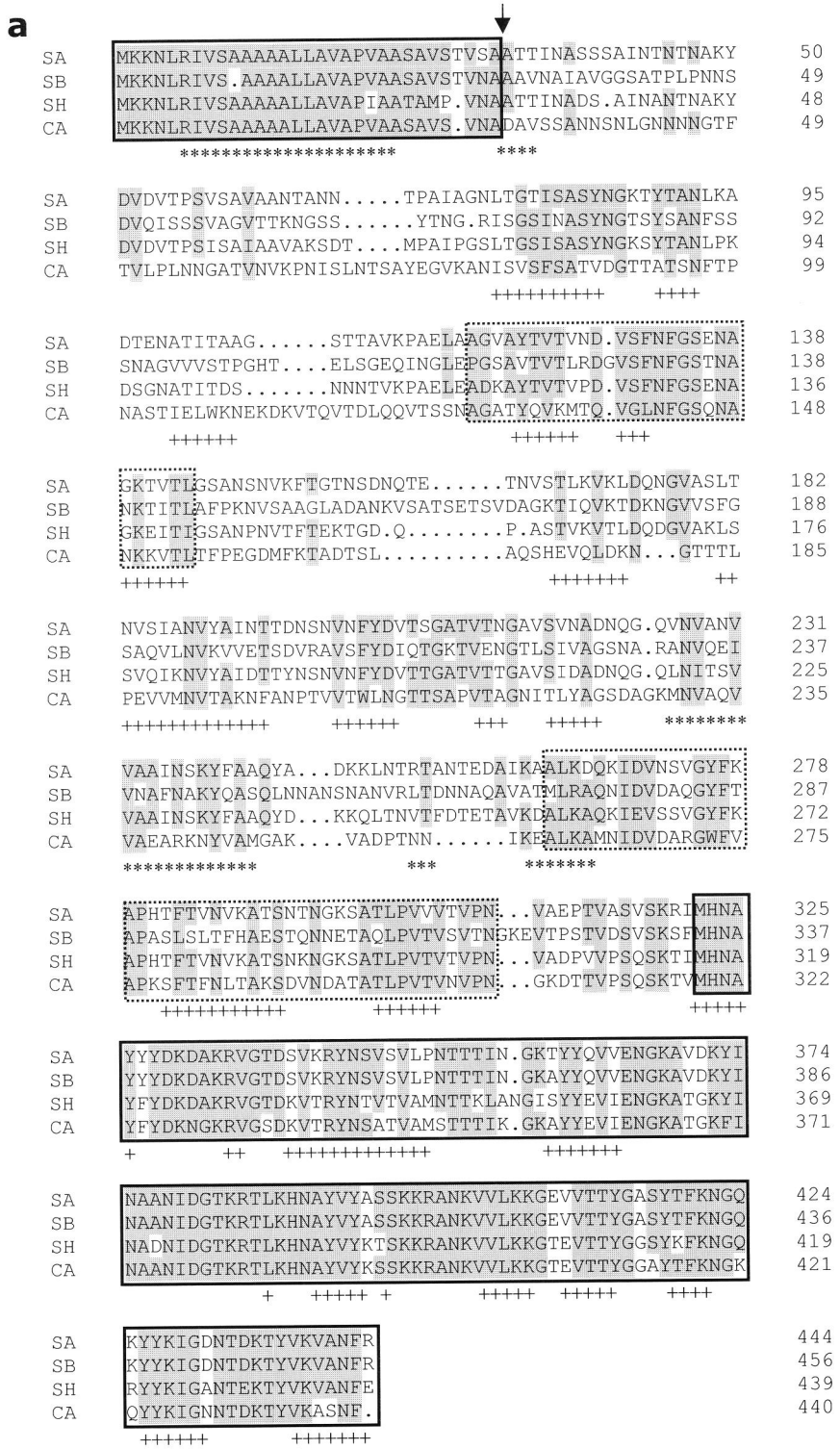
Projection map of S_A-protein bound to a DOPS monolayer. After comparing the power spectra of micrographs of negatively stained 2D crystals formed on various lipid monolayers, we found that S_A-protein 2D crystals were best ordered on DOPS, and they were used for calculating a projection map. Figure 2a shows the oblique lattice and in figure 2b a power spectrum is shown. Lattice parameters (averaged over several images) are a= 118 Å, b= 53 Å, γ = 102° (plane group = p2). Peaks up to a resolution of 17 Å were included for calculation of the projection map. The map was symmetrized according to p2 and is shown as a contour plot (figure 2c) and as a density plot (figure 2d). One repeating unit, which contains two-fold related motifs, is emphasized in figure 2c. One unit cell with p2 symmetry elements is drawn in figure 2d. Each repeating unit shows a spherical and a more elongated density.

Multiple sequence alignment and primary amino acid sequence analysis of *Lactobacillus* S-layer proteins. Sequence alignment, secondary structure prediction, hydrophilicity determination and finally analysis of the distribution of charged amino acid residues revealed that S_A-protein is composed of two regions with quite different characteristics. Alignment of four *Lactobacillus* S-protein sequences showed a highly conserved signal sequence of 31 amino acids, an N-terminal region (~290 amino acids) of variable amino acid composition (31-72% identity) and a C-terminal region (~123 amino acids) of conserved composition (77-99% identity; figures 3a, 3b). Although the N-terminal regions of the four proteins vary greatly in overall primary sequence, several stretches could be observed showing significant conservation (amino acid residues 118-144 and 263-308 in the alignment, figure 3a) in all four proteins. Alignment-based secondary structure prediction (figure 3a) showed the presence in the N-terminal region of beta strands and alpha helices, whereas the C-terminal region contains only beta strands. The hydrophilicity plot of S_A-protein (figure 3c) showed that the N-terminal region consists of alternating stretches of hydrophilic and hydrophobic amino acids, while the C-terminal region is almost exclusively comprised of hydrophilic amino acids. This difference in hydrophilicity is also reflected in the distribution of positively and negatively charged amino acid residues (figure 3d). The N-terminal region contains about as many negatively as positively charged residues, but the C-terminal region has a surplus of positive charges (+ 14). Multiple alignment of the C-terminal regions of the four S-layer proteins and of the cell surface-associated proteinases PrtB and PrtH (Gilbert *et al.*, 1996; Pederson *et al.*, 1999) showed the presence of a tandemly repeated sequence with two conserved, adjacent tyrosine residues at similar positions in the repeated sequence (figure 4).

Proteolytic and chemical breakdown of S_A-protein and N-terminal sequencing.
S_A-protein (DS) was proteolytically or chemically digested and analyzed by SDS-PAGE and Western blotting (figure 5). As judged by SDS-PAGE, treatment of S_A-protein with trypsin or chymotrypsin yielded a large peptide fragment with a relative molecular mass (Mr) of ~36 kDa and also smaller peptides with Mr's of 31, 25 and 18 kDa. Treatment with V8 protease mainly yielded a fragment with a Mr of ~36 kDa and, in addition, one slightly larger fragment and some fragments smaller than 36 kDa. Western blot analysis revealed also some fragments smaller than 14 kDa (not shown). The N-termini of the three 36 kDa protease fragments were determined and found to be the same, ATTIN. This corresponds to the N-terminus of mature S_A-protein (Boot *et al.*, 1995). The largest V8-peptide had the same N-terminus, while the N-termini of two smaller V8-peptides were NATIT and TNVST, corresponding to amino acids 68-72 and 133-137 of S_A-protein.

CNBr treatment of S_A-protein (S_A-protein contains a single Met residue at position 291) yielded two fragments, as expected, of 29.7 kDa and 13.8 kDa (corresponding to the N-terminal two-third and the C-terminal one-third of S_A-protein, respectively). Antibodies raised against S_A-protein reacted strongly in Western blots with the large but not with the small fragment (not shown). Interestingly, the 29.7 kDa CNBr fragment had a relative molecular mass in SDS-PAGE of ~36 kDa suggesting a relatively low SDS binding capacity of the fragment. The C-terminal CNBr fragment of 13.8 kDa showed no aberrant mobility (figure 5).

To determine the accessibility of lysine residues in S_A-protein at the bacterial surface, *L. acidophilus* whole cells were treated with trypsin and Lys-C, which have comparable specificity but a different relative molecular mass (23,5 versus 30 kDa, respectively). Digestion with trypsin liberated a peptide with an apparent molecular mass of 36 kDa in SDS-PAGE, showing that this protease is able to cleave S_A-protein on whole cells at the same position as in purified S_A-protein. Digestion with Lys-C did not produce any visible breakdown products as observed by Coomassie staining and Western blot analysis (results not shown). Extracted S_A-protein, however was degraded by Lys-C.


a
SA MKKNLRIVSAAAAALLAVAPVAASAVSTVSAATTINASSSAINTNTNAKY 50
SB MKKNLRIVS.AAAALLAVAPVAASAVSTVNAAAVNAIAVGGSATPLPNNS 49
SH MKKNLRIVSAAAAALLAVAPIAATAMP.VNAATTINADS.AINANTNAKY 48
CA MKKNLRIVSAAAAALLAVAPVAASAVS.VNADAVSSANNSNLGNNNNGTF 49
SA DVDVTPSVSAVAANTANN.....TPAIAGNLTGTISASYNGKTYTANLKA 95
SB DVQISSSVAGVTTKNGSS.....YTNG.RISGSINASYNGTSYSANFSS 92
SH DVDVTPSISAIAAVAKSDT....MPAIPGSLTGSISASYNGKSYTANLPK 94
CA TVLPLNNGATVNVKPNISLNTSAYEGVKANISVSFSATVDGTTATSNFTP 99
SA DTENATITAAG......STTAVKPAELAAGVAYTVTVND.VSFNFGSENA 138
SB SNAGVVVSTPGHT....ELSGEQINGLEPGSAVTVTLRDGVSFNFGSTNA 138
SH DSGNATITDS.......NNNTVKPAELEADKAYTVTVPD.VSFNFGSENA 136
CA NASTIELWKNEKDKVTQVTDLQQVTSSNAGATYQVKMTQ.VGLNFGSQNA 148
SA GKTVTLGSANSNVKFTGTNSDNQTE......TNVSTLKVKLDQNGVASLT 182
SB NKTITLAFPKNVSAAGLADANKVSATSETSVDAGKTIQVKTDKNGVVSFG 188
SH GKEITIGSANPNVTFTEKTGD.Q.....P.ASTVKVTLDQDGVAKLS 176
CA NKKVTLTFPEGDMFKTADTSL.....AQSHEVQLDKN...GTTTL 185
SA NVSIANVYAINTTDNSNVNFYDVTSGATVTNGAVSVNADNQG.QVNVANV 231
SB SAQVLNVKVVETSDVRAVSFYDIQTGKTVENGTLSIVAGSNA.RANVQEI 237
SH SVQIKNVYAIDTTYNSNVNFYDVTTGATVTTGAVSIDADNQG.QLNITSV 225
CA PEVVMNVTAKNFANPTVVTWLNGTTSAPVTAGNITLYAGSDAGKMNVAQV 235
SA VAAINSKYFAAQYA...DKKLNTRTANTEDAIKAALKDQKIDVNSVGYFK 278
SB VNAFNAKYQASQLNNANSNANVRLTDNNAQAVATMLRAQNIDVDAQGYFT 287
SH VAAINSKYFAAQYD...KKQLTNVTFDTETAVKDALKAQKIEVSSVGYFK 272
CA VAEARKNYVAMGAK....VADPTNN.....IKEALKAMNIDVDARGWFV 275
SA APHTFTVNVKATSNTNGKSATLPVVVTVPN...VAEPTVASVSKRIMHNA 325
SB APASLSLTFHAESTQNNETAQLPVTVSVTNGKEVTPSTVDSVSKSFMHNA 337
SH APHTFTVNVKATSNKNGKSATLPVTVTVPN...VADPVVPSQSKTIMHNA 319
CA APKSFTFNLTAKSDVNDATATLPVTVNVPN...GKDTTVPSQSKTVMHNA 322
SA YYYDKDAKRVGTDSVKRYNSVSVLPNTTTIN.GKTYYQVVENGKAVDKYI 374
SB YYYDKDAKRVGTDSVKRYNSVSVLPNTTTIN.GKAYYQVVENGKAVDKYI 386
SH YFYDKDAKRVGTDKVTRYNTVTVAMNTTKLANGISYYEVIENGKATGKYI 369
CA YFYDKNGKRVGSDKVTRYNSATVAMSTTTIK.GKAYYEVIENGKATGKFI 371
SA NAANIDGTKRTLKHNAYVYASSKKRANKVVLKKGEVVTTYGASYTFKNGQ 424
SB NAANIDGTKRTLKHNAYVYASSKKRANKVVLKKGEVVTTYGASYTFKNGQ 436
SH NADNIDGTKRTLKHNAYVYKTSKKRANKVVLKKGTEVTTYGGSYKFKNGQ 419
CA NAANIDGTKRTLKHNAYVYKSSKKRANKVVLKKGTEVTTYGGAYTFKNGK 421
SA KYYKIGDNTDKTYVKVANFR 444
SB KYYKIGDNTDKTYVKVANFR 456
SH RYYKIGANTEKTYVKVANFE 439
CA QYYKIGNNTDKTYVKASNF. 440

b

	S_A		S_B		S_H		CbsA	
	N	C	N	C	N	C	N	C
S_A	-	-	38	99	72	78	33	77
S_B	38	99	-	-	36	31	31	77
S_H	72	78	36	78	-	-	30	81
CbsA	33	77	31	77	30	81	-	-

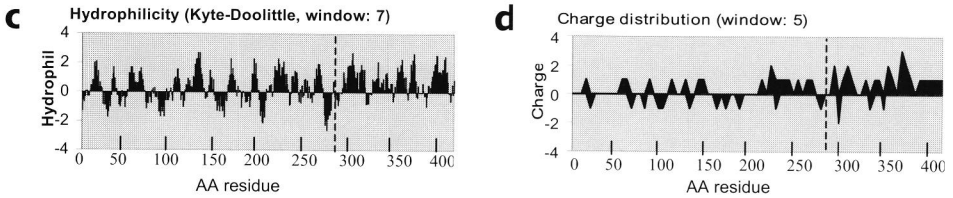


Figure 3. (a) (Previous page) Alignment and secondary structure prediction based on the alignment of four *Lactobacillus* S-protein sequences. Well-conserved amino acid residues (≥ 75%) are shaded in gray. ++ indicate predicted β-strands and ** indicate predicted α-helices. Signal peptide cleavage site is indicated with an arrow. Conserved N-terminal signal peptide and C-terminal sequences are boxed with a solid line and regions of relatively high sequence identity are boxed with a dashed line; (b) the percentages of identity between the different protein sequences for the N-terminal and C-terminal region separately; (c) hydrophilicity of S_A-protein according to Kyte-Doolittle calculated with a window of seven amino acid residues; (d) charge distribution for S_A-protein averaged over a window of five amino acids; Signal sequence amino acids were not included in (b), (c) and (d). SA = S_A-protein of *L. acidophilus* ATCC 4356, SB = S_B-protein of *L. acidophilus* ATCC 4356, SH = S-protein of *L. helveticus* CNRZ 892, CA/CbsA = S-protein of *L. crispatus* JCM 5810.

Properties of N-terminal and C-terminal S_A-protein fragments synthesized in *E. coli*. Three recombinant plasmids were constructed, one encoding the N-terminal two-third fragment of S_A-protein (SAN), one encoding the C-terminal fragment (SAC) fused to Green Fluorescent Protein (GFP) and one containing GFP only. The genes encoding these proteins were efficiently expressed in *E. coli* as His-tagged proteins which were purified by metal affinity chromatography. No transformant could be obtained that expressed the C-terminal region only. This may be due to the very basic properties of SAC that might interfere with cellular processes. Alternatively, an expressed SAC fragment may not be detectable due to its sensitivity to proteolytic attack.

The SAN peptide was found as inclusion bodies and could only be purified under denaturing conditions. During dialysis a precipitate was formed composed exclusively of

SAN as shown by SDS-PAGE (figure 6). Purified SAN yielded mono- and multi-layered 2D crystals and the mono-layered crystals were used for structure analysis. SAN crystals formed 2D crystals with lattice parameters identical to those of intact S-protein (figures 2e, 2f). An S_A-protein minus SAN density-difference map was calculated and showed one large peak that is superimposed on the SAN contour plot. In figure 2g two, p2 related, difference peaks are shown. The more elongated density of the repeating subunit is clearly less pronounced in SAN than in S_A-protein (2g, 2h).

```
Repeat 1
Sa1     TVASVSKRIMHNAYYYDKDAKR..VGTDSVKRYNSVSVLPNTTTI.NGKTYYQVVENGKAVDKYINA.A.... 65
Sb1     TVdSVSKsfMHNAYYYDKDAKR..VGTDSVKRYNSVSVLPNTTTI.NGKaYYQVVENGKAVDKYINAa..... 65
Sh1     vVpSqSKtIMHNAYfYDKDAKR..VGTDkVtRYNtVtVamNTTklaNGisYYeViENGKAtgKYINAd..... 66
Cbsa1   TVpSqSKtvMHNAYfYDKngKR..VGsDkVtRYNSatVamsTTTI.kGKaYYeViENGKAtgKfINAa..... 65
Hap501  TtpSdSnaetHtvmvdsraydkdgnylghmyyaydnidivpTvvtiNGKTYYkVankdeyVsvt......... 64
Prth1   TeegkdtkvMfksvlYtKDlKktrstaqaysslklVtekgklkvytfkghYfykVvdrnAyvrvr........ 65
Prtb1   .Vkktktykkvkltkltkvynkk.gkvvgklkkktsvkLlskkqklhGKyYYrVg.....knrYIlAsslp.. 64

Repeat 2
Sa2     nidgtkrtlkHNAYvYasskKRa.nkvvlkK.gevVttygasyTfkNGqkYYkig..dntdktYvkvanf.r. 67
Sb2     nidgtkrtlkHNAYvYasskKRa.nkvvlkK.gevVttygasyTfkNGqkYYkig..dntdktYvkvanfr.. 67
Sh2     nidgtkrtlkHNAYvYktskKRa.nkvvlkK.gteVttyggsykfkNGqrYYkig..antektYvkvanf..e 67
Cbsa2   nidgtkrtlkHNAYvYksskKRa.nkvvlkK.gteVttyggayTfkNGKqYYkig..nntdktYvkAsnf... 66
Hap502  nitghqrtlrHNAYiYwssyrRt.pGTgkmyRgqtVttygpamrfkNGKkYYriqgcrnnnkrYIkAanfy.. 70
Prth2   nVtgtkatlkrNsfvYqsngKkasrkllkkgttitVygdqykalkhykKyaYrigEgryiksvnvNrvdlvk. 72
Prtb2   ..kktkKvkqvrArknaKvynkk.gkvvghlkkkqkvkLlskkqklhGKyYYrig.....knrYvNAnvl... 62
Consensus GTK T  HNAY Y K  KR   K         VT          NGK YY IG        Y  A

Consensus (von Eichel-Streiber)                    IDGkwYYFD N G
Consensus (Wren)                                   KAVTGWxTIxGxxYYFxxNGx
```

Figure 4. Alignment of two adjacent repeats present in the C-terminal region of four S-proteins, proteinases PrtH and PrtB and haem-agglutinating surface layer protein HAP50. The consensus is given in the last line in bold face. The consensus sequence of repetitive C-terminal ligand-binding sites of clostridial toxins and streptococcal glucosyltransferases (von Eichel-Streiber *et al.*, 1992; Wren, 1991) are shown. Sa = S_A-protein of *L. acidophilus* ATCC 4356, Sb = S_B-protein of *L. acidophilus* ATCC 4356, Sh = S-protein of *L. helveticus* CNRZ 892, Cbsa = S-protein of *L. crispatus* JCM 5810, Prth = cell surface associated proteinase PrtH of *L. helveticus* CNRZ 32, Prtb = cell surface associated proteinase PrtB of *L. delbrueckii* subsp. *bulgaricus*, Hap50 = haem-agglutinating surface layer protein HAP50 of *L. acidophilus* JCM 1038. Suffixes 1 and 2 indicate repeat one and two, respectively.

The GFP-SAC peptide was present in *E. coli* in a soluble form and could be purified under native conditions (figure 7e). GFP-SAC was shown to bind to *L. acidophilus* cells from which the S-layer was removed by treatment with 5 M LiCl, but not to untreated *L. acidophilus*, or to *L. casei* (figure 7). Fluorescence microscopy studies showed that the bacteria were densely covered with GFP-SAC (not shown). GFP-SAC was also shown to bind to LiCl-treated *L. helveticus* and *L. crispatus*. Neither purified N-terminal fragment (SAN) nor GFP bound to stripped *L. acidophilus* or to *L. casei* (not shown).

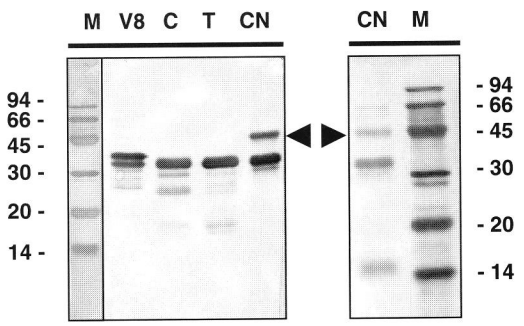


Figure 5. Western blot analysis (left panel) and Coomassie staining (right panel) of peptides obtained after proteolytic and chemical digestion of S_A-protein. S_A-protein was digested with V8-protease (V8), chymotrypsin (C), trypsin (T) and CNBr (CN). Peptides were detected with anti-S_A polyclonal antiserum. M = reference proteins (Mr in kDa indicated in the margin). The triangles indicates the position of S_A-protein (43 kDa).

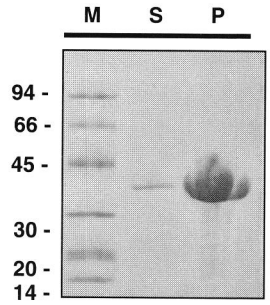


Figure 6. SDS-PAGE of dialyzed fractions of purified SAN peptide. P = pellet fraction, S = supernatant fraction, (M) reference proteins (Mr in kDa indicated in the left margin).

Discussion

S-layers of different bacteria display different types of symmetry pattern but all S-proteins investigated so far have in common that they are composed of two morphological domains (Baumeister *et al.*, 1989). In the present study we show that the S-layer of *L. acidophilus* ATCC 4356 comprised of S_A-protein has a p2 lattice structure. We also show that S_A-protein consists of two differently structured domains with different functions, a crystallization domain consisting of the N-terminal two-third and a cell wall binding domain formed by the C-terminal one-third of the molecule. To our knowledge this is the first S-protein for which both the crystallization and cell wall binding domain have been identified.

Electron image analysis of S_A-protein crystals grown on DOPS showed that the S-layer is organized in an oblique lattice with p2 symmetry. The lattice parameters correspond to those found for the S-layer present on *L. helveticus* ATCC 12047 cells (Lortal *et al.*, 1992). The symmetry pattern (p4) of the S-layer of *L. brevis* is however different, and this difference corresponds with the absence of detectable sequence similarity between the S-proteins of these species (Vidgrén *et al.*, 1992).

The presence of two domains in S_A-protein was inferred from an alignment of four S-protein sequences from related lactobacilli. S_A-protein encompasses an N-terminal region (~ 290 amino acids) that is highly variable among the four S-proteins and consists of alternating regions of hydrophilic and hydrophobic amino acids, and a highly conserved C-terminal region (~123 amino acids) that is almost exclusively comprised of hydrophilic, mainly basic, amino acids (figure 3).

By proteolytic treatment we were able to divide S_A-protein into two regions with sizes corresponding to those predicted by our computer analysis. Treatment of S_A-protein with

trypsin or chymotrypsin resulted in the complete removal of the C-terminal region, yielding a ~30 kDa peptide resistant to further degradation, despite the presence of many potential trypsin and chymotrypsin cleavage sites. Target sites in the C-terminal region apparently are freely accessible to these proteolytic enzymes, since neither the C-terminal fragment itself nor degradation products could be detected in Western blots, even after mild enzymatic treatment.

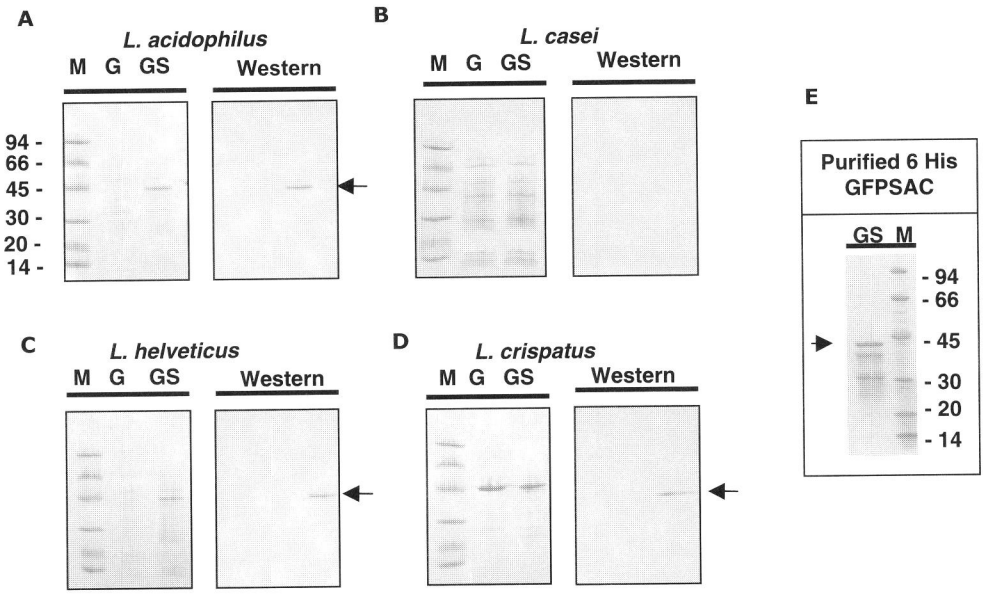


Figure 7. Cell wall binding of GFP-SAC to LiCl-treated *L. acidophilus* (A), *L. casei* (B), *L. helveticus* (C) and *L. crispatus* (D). G = 6 His GFP; GS = 6 His GFP-SAC; left panels SDS-PAGE; right panels Western blots with anti 6 His. E: SDS-PAGE of purified GFP-SAC (GS). The arrow indicates the intact GFP-SAC fusion protein. (M) reference proteins (Mr indicated in the right margin in kDa).

The C-terminal region was digested by trypsin but not by Lys-C, when S_A-protein was present on the bacterial surface. This result suggests that the C-terminal domain is buried below the N-terminal domain and is accessible only to molecules small enough to pass through the pores of the S-layer. The presence of protease-susceptible and protease-refractile regions has been demonstrated for several S-proteins (Chami *et al.*, 1997; Chu *et al.*, 1991; Egelseer *et al.*, 1998; Hastie & Brinton, 1979; Thomas *et al.*, 1992). In general, protease-susceptible regions fold into open structures (i.e. surface loops, domain linking regions), whereas more refractile regions fold into compact structures. The N-terminal two-third of S_A-protein thus probably forms a domain with a compact structure in which target residues are protected from proteolytic attack, while

the target residues in the C-terminal domain are probably located at the domain surface and so are more susceptible to proteolytic attack.

Final proof for the presence of two domains was obtained by a functional analysis of the two regions. A truncated S_A-protein, comprising the N-terminal 290 amino acids, synthesized in *E. coli* formed large, mono- and multi-layered crystalline arrays with the same lattice parameters as those of S_A-protein, indicating that all information necessary for the formation of S-layer crystals is contained within the N-terminal domain. The contour plot of SAN shows a spherical and a more elongated density that might represent two SAN subdomains (figure 2c, 2g). For the elongated SAN subdomain less electron density was observed than the corresponding region in S_A-protein reconstruction indicating that the SAC domain in S_A-protein is located above or below, but most likely below this subdomain.

Digestion of the S-protein of *B. stearothermophilus* ATCC 12980 with *S. aureus* V8 protease and of *B. sphaericus* NTCC 9602 after prolonged storage yielded protein fragments, that could reassemble *in vitro* into sheets exhibiting an ordered 2D structure. Since the fragments have not been characterized, the identity of the crystallization domain could not be determined. Also no conclusions could be drawn whether the capacity to reassemble resides in a subunit different from the one needed for cell wall assembly (Egelseer *et al.*, 1998; Hastie & Brinton, 1979).

The presumed role of the C-terminal domain of S_A-protein in cell wall attachment was assessed by binding experiments. The C-terminal 123 amino acids of S_A-protein (SAC) fused with GFP efficiently bound to *L. acidophilus, L. helveticus* and *L. crispatus* cells from which the S-layer had been stripped, but not to non- stripped *L. acidophilus* or *L. casei.* From these results we conclude that the C-terminal region of S_A-protein is necessary and sufficient for cell wall anchoring. Interestingly, the C-terminal region (W-domain) of the cell-envelope proteinases PrtH of *L. helveticus* CNRZ 32 and PrtB of *L. delbrueckii subsp. bulgaricus* showed strong similarity to the C-terminal region of *Lactobacillus* S-proteins (Gilbert *et al.*, 1996; Pederson *et al.*, 1999). This might mean that these proteinases use a similar domain for attachment to the cell surface (Siezen, 1999). The SAC domain and corresponding domains of related S-proteins consist of a tandem repeat as was found for PrtB and PrtH. In each repeat two adjacent tyrosines are present. A database search revealed that the tyrosine-containing regions are highly conserved in the carbohydrate-binding domains of extracellular glycosyltransferases, in *Clostridium difficile* toxin A and in haem-agglutinating protein HAP50 from *L. acidophilus* (Genbank AF250229) (von Eichel-Streiber *et al.*, 1992; Wren, 1991). The spacing of the

two Tyr pairs is almost the same suggesting that the motif arose by gene duplication. Our data indicate that, like SLH-containing domains, this domain is repetitive. It is also clear that the domain is modular, supporting the conclusion that it forms an independent domain involved in cell wall binding. Finally, these findings suggest that carbohydrate-binding is important in cell wall attachment of *Lactobacillus* surface proteins.

An important difference between S-proteins of lactobacilli and those of other bacteria is the difference in overall charge. S-proteins of lactobacilli are highly basic (pI >9.4), while S-proteins of most other bacteria, including bacilli are slightly acidic (pI 4-6). Our lipid-layer crystallization experiments suggest that positively charged amino acids of S_A-protein interact with negative charges of the lipid head groups. No divalent cations were necessary for crystallization. Since positively charged amino acids - mainly lysine and some arginine – were predominantly found in the C-terminal region of S_A-protein, these findings support the hypothesis that, besides carbohydrate-binding, electrostatic interactions play a role in binding of the C-terminal region of S_A-protein to components of the cell wall. The cell envelope of lactobacilli and of many other Gram-positive, bacteria contains anionic polymers, like cell wall teichoic acid and lipoteichoic acid (LTA). LTA can specifically bind proteins, as was recently reported for *Listeria monocytogenes* (Jonquières *et al.*, 1999).

The component(s) in the cell wall of *L. acidophilus* interacting with S_A-protein is/are as yet unidentified but one may assume that they must be very similar if not identical in *L. acidophilus, L. helveticus* and *L. crispatus,* since the C-terminal ends of their S-proteins differ only slightly. This hypothesis is supported by experiments showing that the C-terminal fragment of S_A-protein binds to stripped *L. helveticus* and *L. crispatus* cells. In addition, the S-protein encoded by *L. crispatus cbsA,* remained attached to the cell wall when synthesized in *L. acidophilus* but was secreted into the culture medium by *L. casei* (Martinez *et al.*, 2000).

The difference in primary structure of the N-terminal region of the S-proteins of *L. acidophilus*, *L. crispatus* and *L. helveticus* predicts that the surface properties of these bacteria are different. Our finding that antibodies against S_A-protein do not react with S-protein of *L. crispatus* supports this view. Moreover, S-protein of *L. crispatus* can mediate bacterial adhesion to immobilized collagen type IV, whereas S-protein from *L. acidophilus* lacks this property (Toba *et al.*, 1995). The capacity to bind to collagen was found to reside in amino acids 1-287 of the *L. crispatus* S-protein, implying that the N-terminal domain is exposed to the external milieu (Sillanpää *et al.*, 2000). Despite clear differences in primary structure, some amino acid sequences (indicated in bold face in

Figure 3a) in the N-terminal domain were found to be conserved. These amino acids may play a crucial role in S-protein assembly or in maintaining the tertiary structure of S-protein.

In conclusion, our data show that S_A-protein is composed of two structural domains. Interaction between monomers of the *L. acidophilus* S-layer is determined by the N-terminal ~290 amino acids only, while the C-terminal ~123 amino acids define the capacity to attach to the cell wall. EM data show the probable location in projection of the SAC domain in the S_A-protein 2D structure. The N-terminal domain faces the environment and thus contributes to the physical and functional properties of the bacterium. The C-terminal domain, however, is not exposed to the environment.

Acknowledgements. We thank Ton Muijsers for performing the N-terminal sequencing and Alain Brisson for providing EM facilities and critical reading of the manuscript.

Chapter 3

One repeat of the cell wall binding domain is sufficient for the anchoring of the *Lactobacillus acidophilus* surface layer protein

Egbert Smit and Peter H. Pouwels

J. Bacteriol. in press

Summary

The S-layer protein of *Lactobacillus acidophilus* ATCC 4356 is composed of two domains, a crystallization domain, SAN, comprising the first 290 amino acids of the protein, and a cell wall binding domain, SAC, consisting of the last 123 amino acid residues. SAC, which consists of two tandemly repeated sequences, SAC1 and SAC2, represents a new member of the superfamily of ligand-binding proteins. SAC, SAC1 and SAC2, were efficiently produced in *Escherichia coli*, as fusions to the C-terminus of His-tagged Green Fluorescent Protein (HGFP), and purified by metal-affinity chromatography. HGFP-SAC and HGFP-SAC1 bound very efficiently to native *L. acidophilus* cell wall fragments (CWF's), whereas HGFP-SAC2 did not. When bound to CWF's, HGFP-SAC and HGFP-SAC1 could form dimers. The binding of GFP-SAC and GFP-SAC1 was completely abolished by HF treatment of CWF's, a method which extracts peptidoglycan-associated polymers. Native CWF's contained high amounts of phosphorus, indicative of the presence of (lipo)teichoic acids ((L)TA's), and glucose, galactose and an unidentified sugar. The complete extraction of phosphorus and decrease of the galactose content of CWF's by HF treatment indicated the removal of (L)TA. Specific removal of LTA from CWF's with phenol did not affect binding of SAC or SAC1. These data indicate that SAC anchors the S_A-protein to a cell wall teichoic acid, possibly substituted with galactosyl residues. We conclude that SAC1 is necessary and sufficient for cell wall binding. SAC2 probably plays an accessory role in the binding process, e.g. by increasing the specificity of interaction.

Introduction

The cell surface of Gram-positive bacteria is decorated with many different proteins with a wide range of functions. A special category of surface proteins is formed by the so-called Surface layer or S-layer proteins, which represent the most abundant cell-surface proteins in many bacteria. S-layers are two-dimensional paracrystalline arrays usually found as the outermost layer of the cell envelope. They are present in most of the Archaea, where they primarily serve as a shape-determinant, and in many Eubacteria, where they can serve as a molecular sieve, or a scaffold for extracellular enzymes, or can play a role in bacterial adhesion and pathogenesis (Baumeister & Lembcke, 1992; Noonan & Trust, 1997; Sára & Sleytr, 1987a).

S-layers are composed of a single protein or glycoprotein, which self-assembles at the cell surface into a paracrystalline layer with p1, p2, p3, p4 or p6 symmetry. The S-protein subunits vary in size from 40 to ~200 kDa. S-layers self-assemble in an entropy-driven process during which multiple, non-covalent interactions are formed between individual S-protein monomers and with the underlying cell surface. In some S-proteins the two types of interaction can be assigned to two separate domains (Jarosch *et al.*, 2000; Smit *et al.*, 2001).

While relatively little is known about the domains responsible for the interactions between S-protein monomers, a wealth of information has been collected about cell wall binding domains (CWBD) of S-layer proteins. The Surface Layer Homology (SLH) domain, found in S-proteins of *Bacillus*, *Thermophilus*, *Thermoanaerobacter* and *Clostridium* species, represents the most extensively studied CWBD of S-layer proteins. SLH domains are also found near the N-terminal end of extracellular enzymes and in outer membrane proteins of Gram-negative Bacteria. They consist of ~ 55 amino acid residues and can be found in 1 to 3 copies in different proteins (Brechtel *et al.*, 1999; Engelhardt & Peters, 1998; Matuschek *et al.*, 1996; Mesnage *et al.*, 1997). First, thought to be a peptidoglycan–binding domain, the SLH domain is now known to bind PG-associated cell wall polymers. The mechanism by which *B. anthracis* S-layer proteins bind to the cell wall is well documented and their SLH domains bind to a pyruvylated PG-associated cell wall polysaccharide through a mechanism shown to be conserved among related bacteria (Mesnage *et al.*, 2000).

In other bacteria the mechanism by which S-proteins are anchored to the cell wall may be different. The S-protein of *Corynebacterium glutamicum* possesses a C-terminal hydrophobic anchor of ~ 79 amino acid residues, which was suggested to interact with a

hydrophobic layer composed of mycolic acids present in the cell wall (Chami *et al.*, 1997). The S-layer protein of *L. acidophilus* ATCC 4356 is anchored to the cell wall by means of a C-terminal CWBD, SAC, which consists of a tandemly repeated ~60 amino acid sequence with a conserved tyrosine doublet. The doublet region showed homology to carbohydrate binding regions of *Clostridium difficile* toxins and extracellular glycosyltransferases and cell wall-associated proteinases of lactic acid bacteria, suggesting a mechanism of cell wall binding different from either SLH- or hydrophobic anchor-mediated cell surface attachment (Smit *et al.*, 2001).

In this study we describe the cell wall binding properties of the *L. acidophilus* SAC domain and its two constituent repeats, SAC1 and SAC2. SAC1 binds to cell wall fragments in similar numbers as SAC, but SAC1 was more easily removed with NaCl. SAC2 however does not bind to the *L. acidophilus* cell wall. We further demonstrate that the ligand for SAC is a PG-associated cell wall polymer, most likely a cell wall teichoic acid, possibly substituted with galactosyl residues.

Materials and Methods

Bacterial strains and culture conditions. *E. coli* strain DH5α, used as host for the cloning of DNA fragments, and strain M15 [pREP4] (Qiagen), used for synthesis of recombinant S_A-protein or S_A-protein fragments, were cultivated in Luria (L) broth. For the production of fusion proteins *E. coli* M15 [pREP4] was first grown to an optical density (OD_{695}) of 0.6 at 37°C followed by cooling of the culture to room temperature (RT) after which IPTG was added to a final concentration of 1 mM and cultivation was continued over night (ON). *L. acidophilus* and *L. casei* were cultivated anaerobically in MRS broth (Difco). When necessary, media were supplemented with 100 µg/ml ampicillin, 25 µg/ml kanamycin, 10 µg/ml chloramphenicol and agar (1.5%).

Cloning, synthesis and purification of HGFP-SAC, HGFP-SAC1 and HGFP-SAC2 fusion proteins. The sequences encoding the two SAC domain repeats, SAC1 and SAC2 were amplified from plasmid pBK1 (Boot *et al.*, 1993) by PCR using oligonucleotides SP2 (5'GAT CGG ATC CCT CGA GAT GCA CAA CGC ATA CTA CTA CGA 3') and SP3 (5' CCC CAA GCT TTT ATT ATG CAG CGT TGA TGT ACT TGT C 3') for SAC1 and oligonucleotides SP4 (5'GGG GCT CGA GAA CAT CGA TGG TAC TAA GCG TAC 3') and SP5 (5' CCC CGG ATC CAA GCT TAT CGA AGT ATC AGA AGA TCC TAT T 3') for SAC2. The PCR products were cloned in pGEM-T and sequenced using universal primers. They were then transferred as *Xho*I-*Hin*dIII fragments to *Xho*I-*Hin*dIII digested pHGFPSAC, replacing the

SAC fragment and yielding expression vectors pHGFPSAC1 and pHGFPSAC2 (figure 1). Purification of his-tagged fusion proteins was performed as described previously for HGFP-SAC (Smit *et al.*, 2001). After purification, the buffer for elution of the his-tagged proteins was exchanged for 50 mM Tris-HCl, pH 7.5 using PD-10 desalting columns (AP Biotech, Sweden) and protein concentrations were adjusted to 1 mg/ml.

Preparation of *L. acidophilus* native, HF- and phenol extracted cell wall fragments. *L. acidophilus* and *L. casei* were grown ON and harvested by centrifugation. About 5 g (wet weight) cells were washed once with a physiological salt solution and resuspended in 30 ml of 1% SDS. Cells were broken by ultrasonic treatment (ten 30 s pulses, the temperature was kept below 60°C via intermittent cooling at 0°C) using a Branson sonifier (New Brunschwick Scientific). Cell wall fragments (CWF's) were collected by centrifugation (20 min, 20.000 x g, 4°C), resuspended in 1 % SDS and incubated at 100°C for 30 min. This was repeated three times followed by extensive washing with distilled water (AD). The removal of cell wall-associated proteins was checked by SDS-PAGE and silver staining. CWF's were then resuspended in 30 ml of 20 mM Tris-HCl (pH 8) containing 2 mM $MgCl_2$, 25 µg/ml RNaseA (Sigma), 25 µg/ml DNAseI (Sigma) and incubated for 30 min at 37°C. RNase and DNase were removed by 1% SDS treatment as described above. Finally, CWF's were washed five times with AD and lyophilized. The CWF preparation obtained at this point will be referred to as native CWF's. To remove peptidoglycan (PG)-associated cell wall components *L. acidophilus* CWF's (100 mg dry weight) were treated with 40 % (wt/vol) aqueous hydrofluoric acid (HF) for 96 h at 4°C. After collecting the CWF's by centrifugation (20 min, 20.000 x g, 4°C) the pellets were washed once with 40% HF, washed five times with AD and lyophilized. The preparations obtained are referred to as HF-CWF's. Native CWF's were also treated with aqueous phenol (80%, 65°C, 1 h) to specifically extract lipoteichoic acid (LTA) (Fischer *et al.*, 1983). After centrifugation (20 min, 20.000 x g, 4°C) CWF's were removed from the phenol-water interface and washed once with chloroform/water (1:1, v/v) to remove remaining phenol. After extensive washing with AD CWF's were lyophilized. These CWF's are referred to as phenol-CWF's.

Binding of HGFP-SAC, HGFP-SAC1 and HGFP-SAC2 to *L. acidophilus* and *L. casei* cell wall fragments. Fusion proteins HGFP-SAC, HGFP-SAC1 and HGFP-SAC2 (~50 µg) were incubated with ~50 µg of *L. acidophilus* native-, HF- or phenol-CWF's or native *L. casei* CWF's in 100 µl binding buffer for 1 h at 37 °C. After incubation, CWF's with bound proteins were collected by centrifugation to separate them from soluble, non-bound, protein. The pellet and supernatant fractions were saved for further treatment. The pelleted CWF's were washed three times with incubation buffer and suspended in SDS

sample buffer. Non-bound protein was precipitated from the supernatant with 3 volumes of 100 % (w/v) ammonium sulfate for 30 min at 0°C, centrifuged (15 min 20.000 x g, 4°C) and dissolved in SDS sample buffer. Non-bound and bound protein fractions were assessed by SDS-PAGE. In initial binding experiments the NaCl concentration of the buffer was varied to determine the optimal buffer composition for binding. A Standard Incubation Buffer (SIB) containing 50 mM Tris-HCl, pH 7.5 and 150 mM NaCl was chosen and used in all subsequent binding experiments. SIB was supplemented with monosaccharides and detergents to determine their effect on binding. Galactose, glucose and N-acetylglucosamine were added at concentrations up to 500 mM and non-ionic detergents Tween-20, Tween-80 or Triton X-100 up to 1% (w/v). Binding to CWF's was performed as described above.

Compositional analysis of native, HF- and phenol-extracted cell wall fragments, and HF extracted secondary cell wall component. The phosphorus content of native, HF- and phenol-extracted CWF's as well as HF-solubilized low molecular weight material was determined according to Chen (Chen *et al.*, 1956). The monosaccharide composition of native and HF-CWF's was determined by acid hydrolysis (2N H_2SO_4, 100°C, 1 h) followed by Dionex anion-exchange chromatography using glucose, galactose, mannose and N-acetyl-glucosamine as standards.

DNA and protein techniques. DNA isolation, restriction, agarose gel electrophoresis, ligation and introduction of DNA into *E. coli* by electroporation were performed according to standard procedures (Sambrook *et al.*, 1989). PCR and DNA sequence analysis were performed according to standard procedures. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli (Laemmli, 1970) using the Miniprotean™ system (BioRad). Protein concentrations were measured according to Lowry (Lowry *et al.*, 1951). Blotting of proteins to nitrocellulose membranes and Western blot analysis were performed according to standard procedures. Alkaline phosphatase chromogenic substrates NBT and BCIP were obtained from Sigma.

Results

Production and purification of HGFP-SAC, HGFP-SAC1 and HGFP-SAC2. *E. coli* transformants harboring pHGFP-SAC1, pHGFP-SAC2 or pHGFP-SAC were used for the synthesis of the corresponding fusion proteins HGFP-SAC, HGFP-SAC1 and HGFP-SAC2. High yields were obtained for all three peptides, which were of the expected M_r 43 kDa (HGFP-SAC), 37 kDa (HGFP-SAC1) and 37 kDa (HGFP-SAC2), respectively. The inset in

figure 1 shows the purified HGFP-SAC, HGFP-SAC1 and HGFP-SAC2 preparations. Several C-terminal deletion products were co-purified with each peptide probably due to *E. coli* proteases. Two major degradation products were co-purified with HGFP-SAC with a M_r of ~ 37 kDa (corresponding to HGFP-SAC1) and ~ 30 kDa (consisting of only GFP). A peptide corresponding to GFP was also co-purified with HGFP-SAC2. Interestingly, HGFP-SAC1 was much less sensitive to proteolytic degradation than HGFP-SAC2.

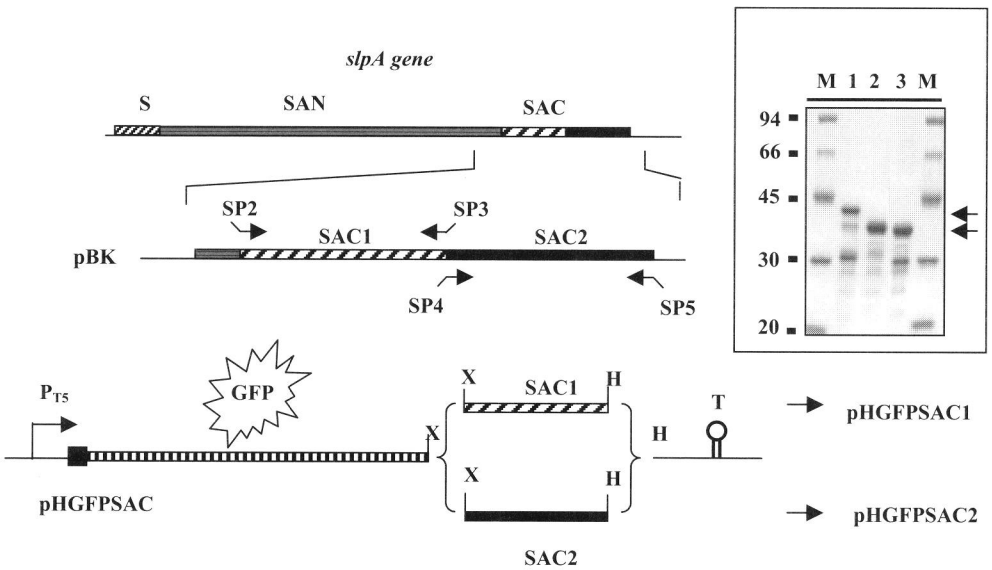


Figure 1. Construction of vectors for the production of the HGFP-SAC1 and HGFP-SAC2 fusion proteins. The *slp*A gene with the signal sequence (S) and functional domains SAN and SAC is shown. PCR fragments (based on primers SP2, SP3, SP4 and SP5) encoding SAC1 and SAC2 were substituted for SAC in expression plasmid pHGFP-SAC. The black square indicates the 6 histidine tag included for purification purposes. X: *Xho*I; H: *Hin*dIII; P_{T5}: hybrid inducible phage T5 promoter; T: transcription terminator. The inset shows purified preparations of HGFP-SAC (lane 1), HGFP-SAC1 (2) and HGFP-SAC2 (3). The top arrow indicates the position of HGFP-SAC, the bottom one that of HGFP-SAC1 and HGFP-SAC2. M_r's in kDa of reference proteins are indicated in the left margin.

Interaction of HGFP-SAC, HGFP-SAC1 and HGFP-SAC2 with native *L. acidophilus* and *L. casei* cell wall fragments. In our previous work we demonstrated the capacity of HGFP-SAC to bind to *L. acidophilus* cells from which the S-layer had been stripped. The binding conditions were further optimized using purified CWF's. A-specific binding of *E. coli* proteins was minimized by the inclusion of NaCl in the binding buffer (50 mM Tris-HCl, pH7.5). Up to 250 mM NaCl could be used without affecting the binding of HGFP-SAC. We chose 50 mM Tris-HCl, pH 7.5, 150 mM NaCl as standard incubation buffer (SIB) for all further binding experiments. Figure 2 shows that HGFP-SAC and HGFP-SAC1 bound quite well to *L. acidophilus* CWF's, but HGFP-SAC2 did not bind. At NaCl concentrations higher than 150 mM binding of HGFP-SAC1 was severely reduced, while

at NaCl concentrations higher than 250 mM binding of HGFP-SAC was reduced (not shown). None of the peptides bound to purified *L. casei* CWF's, as was previously shown for HGFP-SAC (not shown). To further analyze the interaction of HGFP-SAC and HGFP-SAC1 with native CWF's we bound them to CWF's in SIB and re-extracted both peptides with buffers containing increasing concentrations of NaCl. We found that HGFP-SAC could be removed from the CWF's with NaCl concentrations above 250 mM, while for HGFP-SAC1 concentrations higher than 150 mM were enough to remove the peptide (figure 3). A small fraction of both HGFP-SAC and HGFP-SAC1 remained bound even after extraction with 2 M NaCl. The binding of HGFP-SAC and HGFP-SAC1 was not affected by the presence of monosaccharides (galactose, glucose or N-acetylglucosamine) or non-ionic detergents (results not shown).

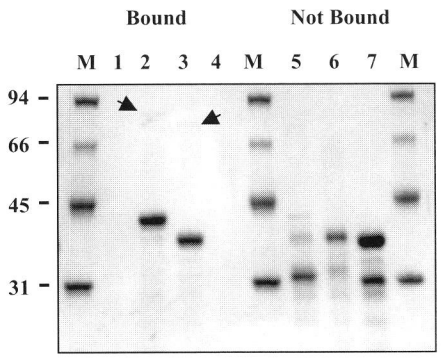


Figure 2. SDS-PAGE of HGFP-SAC (lane 2 and 5), HGFP-SAC1 (lane 3 and 6) and HGFP-SAC2 (lane 4 and 7) interacting with purified native *L. acidophilus* CWF's. lane 1, native CWF's only; lane 2 to 4, CWF-bound material; lanes 5 to 7, remaining non-bound material. M: low Mr reference proteins in kDa. Arrows indicate dimers of HGFP-SAC and HGFP-SAC1.

Binding of HGFP-SAC and HGFP-SAC1 to *L. acidophilus* HF- and phenol-treated CWF's. *L. acidophilus* CWF's were subjected to HF extraction to determine whether SAC binds to the peptidoglycan, the main constituent of bacterial cell walls or to PG-associated cell wall components. HF treatment of CWF's resulted in the complete absence of binding by HGFP-SAC and HGFP-SAC1 (figure 4A). Phenol treatment of CWF's, which extracts LTA from cell wall fragments of various Gram-positive bacteria (Fischer *et al.*, 1983), did not affect binding of either HGFP-SAC or HGFP-SAC1 (figure 4B).

Chemical composition of native, HF- and phenol-extracted *L. acidophilus* cell wall fragments and material extracted by HF treatment. Treatment of native *L. acidophilus* CWF's with HF resulted in a loss in mass of about 16%. Only a small fraction of the solubilized material could be recovered by ethanol precipitation. Subsequently, the monosaccharide composition and phosphorus content was determined for several CWF samples. The monosaccharide composition as determined by of native and HF-extracted CWF's showed the presence of glucose and galactose as well as an unidentified

carbohydrate. The amount of galactose in the HF-extracted CWF's was about one third of that of native CWF's. The amounts of glucose and unidentified sugar were not changed, compared to native CWF's. Analysis of the phosphorus content revealed large differences. Native CWF's contained 1.12 µmol of phosphorus per mg dry weight while in the HF-CWF's only 0.01 µmol was detected. The phosphorus was recovered in the non-precipitable fraction of the HF extract. Phenol treatment of CWF's resulted in a lowering of the phosphorus content from 1.12 to 0.71 µmol/mg.

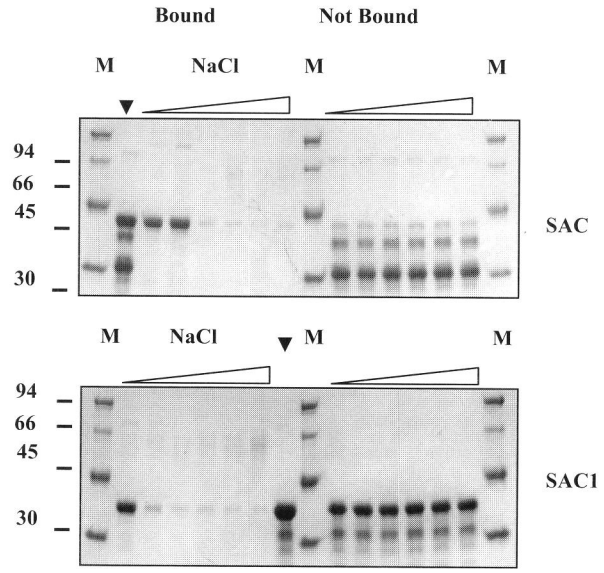


Figure 3. Interaction of HGFP-SAC (top panel), HGFP-SAC1 (bottom) with native *L. acidophilus* CWF's and re-extraction using NaCl. Peptides were bound to CWF's in SIB and re-extracted with 50 mM Tris-HCl, pH7.5 containing 0.15, 0.25, 0.5, 0.75, 1.0 and 2.0 M NaCl (increasing from left to right). M_r's of reference proteins (lanes indicated with M) are given in the left margin. The black triangles indicate purified HGFP-SAC and HGFP-SAC1. The arrows indicate the position of dimers of HGFP-SAC and HGFP-SAC1.

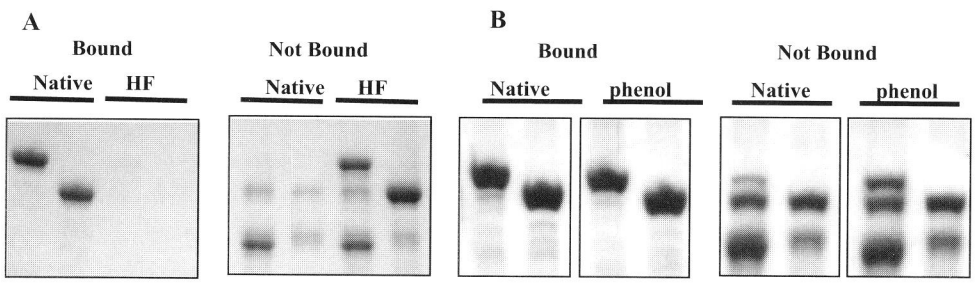


Figure 4. SDS-PAGE of HGFP-SAC and HGFP-SAC1 interacting with (A) *L. acidophilus* native and HF-CWF's or (B) native and phenol-CWF's. (A) native and HF-CWF's, (B) native and phenol-CWF's. In each panel the left lane shows GFP-SAC and the right lane GFP-SAC1.

Discussion

In the present study we determined the properties of the cell wall binding domain, SAC of the *L. acidophilus* S-protein and it's constituent repeats SAC1 and SAC2. Analysis of binding of HGFP-SAC, HGFP-SAC1 and HGFP-SAC2 to native CWF's revealed that equal amounts of HGFP-SAC and HGFP-SAC1 bound to CWF's, while HGFP-SAC2 did not bind at all, implying that SAC1 is necessary and sufficient for binding *in vitro*. Re-extraction experiments showed that the binding of SAC with native CWF's through electrostatic interaction is stronger than for SAC1 since more NaCl is needed to remove SAC. These data suggest a co-operative role for SAC2 in cell wall binding.

Our analyses also revealed that CWF-bound SAC and SAC1 formed dimers, stable even after boiling in 1% SDS, suggesting that the association between the monomers is quite strong. Apparently, SAC1 is necessary and sufficient for dimerization of the cell wall binding domain. Dimerization of SAC is probably needed to enhance effective interaction with the cell wall. It has been suggested that oligomeric SLH structures represent the apparent functional unit *in vivo* rather than the monomer (Engelhardt & Peters, 1998). SLH domain-containing molecules were indeed found to form dimers (and trimers) that resist boiling in 1% SDS (Brechtel & Bahl, 1999). Dimerization of SAC and SAC1 was observed for bound peptides but not for the fusion peptides without CWF or for purified HGFP. An explanation, as to why dimers are only observed in CWF-bound preparations and not in solution, could be that only upon binding monomers come in sufficiently close contact. Another possibility is that a conformational change in SAC or SAC1 during the binding process promotes dimer formation.

To identify the cell wall receptor for SAC, chemical extraction experiments of native *L. acidophilus* CWF's were performed. HF extraction of CWF removes PG-associated polymers like lipoteichoic, teichoic acids (LTA or TA) and polysaccharides (PS) by hydrolyzing phosphodiester bonds. LTA and TA, polymers of glycerol-phosphate or ribitol-phosphate, are degraded into phosphate, glycerol and monosaccharides by HF, while PS's attached to PG by phosphodiester bonds are removed intact (Ekwunife *et al.*, 1991; Fischer *et al.*, 1983). The SAC/SAC1 receptor was removed by this treatment, which indicates that a PG-associated polymer, and not PG itself, is responsible for binding of SAC and SAC1.

The finding that hardly any HF-extracted material could be recovered as ethanol precipitable material indicated that the PG-associated polymer was degraded. This suggests that acid-labile compounds such as LTA or TA rather than PS or teichuronic

acid, which are all removed by HF treatment, are involved in the binding process (Fischer *et al.*, 1983). The presence of teichoic acids in native CWF's and their subsequent removal by HF was confirmed by a high level of phosphorus in native CWF's and its absence in HF-CWF's. To discriminate between LTA and TA, we treated native CWF's with hot aqueous phenol, an effective method to specifically extract LTA from CWF's (Fischer *et al.*, 1983). Binding of SAC and SAC1 was not influenced by this treatment, although it resulted in the removal of about one third of cell wall associated phosphorus. Thus we conclude that PG-associated TA is the cell wall receptor for SAC.

The SLH domain-containing S-protein of *B. anthracis* binds to a pyruvylated (thus anionic) PG-associated polysaccharide, while the SLH-containing SbsB protein of *B. stearothermophilus* binds to a different PG-associated anionic polymer (Mesnage *et al.*, 2000; Sára *et al.*, 1998). The *B. stearothermophilus* S-layer proteins SbsA and SbsC, which do not possess SLH domains, bind to teichuronic acid (Sára, 2001). It appears that anionic PG-associated cell wall polymers like teichuronic acids, teichoic acids and polysaccharides are important S-protein receptors in many Gram-positive bacteria.

Carbohydrate analysis of native and HF-CWF's revealed that the galactose but not the glucose or unidentified sugar content of HF-treated CWF's was lowered by one third the level of native CWF's. This could suggest that the material removed by HF treatment (i.e. TA) contains galactose. A possible role for galactose in binding of SAC is not unexpected considering that *Clostridium difficile* toxin A, which showed the strongest homology with SAC, binds a galactose-containing oligosaccharide (Chaves-Olarte *et al.*, 1997). Our finding that galactose had no effect on binding of SAC to CWF's seems not at variance with this conclusion since carbohydrate recognition often requires a specific structure of the carbohydrate ligand that is absent in a monosaccharide (Weis, 1997).

LTA was shown to be important for the surface association of cell wall lytic enzymes and invasion-mediating proteins (Jonquieres *et al.*, 1999; Yother & White, 1994). These proteins belong to the family of choline-binding proteins and possess repetitive domains responsible for attachment. This well-studied category of proteins include the streptococcal proteins LytA and PspA that specifically bind the choline moiety of LTA (Garcia *et al.*, 1998). Thus it seems that not the LTA backbone but substitutions such as choline could confer protein binding specificity. Considering that TA's may be substituted with glycosyl groups (for example galactosyl, glucosyl or N-acetylglucosamine) (Fischer, 1990; Ward, 1981; Yasui & Yoda, 1997) these groups could be responsible for the specific binding of SAC.

Among animal lectins, plant lectins or lectins of bacterial origin little if any primary sequence identity exists, but some striking similarities can be noted. Most lectins consist exclusively of beta-sheets and possess conserved aromatic residues important for ligand binding. They also tend to form dimers or tetramers (Barondes *et al.*, 1994). Since SAC shares all these properties we suggest that SAC constitutes a lectin-like domain. Also the SLH domain and the CWBD's of SbsA and SbsC were suggested to be lectin-like (Engelhardt & Peters, 1998; Jarosch *et al.*, 2000). SLH domains show little resemblance to the SAC domain except that they consist of similarly sized repeats. The primary amino acid sequences of the SAC and SLH domains do not show any homology and SAC has an all-beta predicted secondary structure compared to a predicted helix-loop-helix motif for SLH. Interestingly, SLH domain-containing proteins showed strongly diminished cell wall binding properties and increased proteolytic sensitivity when one or two copies of the SLH domain were C-terminally deleted, indicating that three SLH repeats probably represent the functional domain (Brechtel & Bahl, 1999; Mesnage *et al.*, 1999a).

Multiple repeats are also required for binding of the streptococcal proteins PspA and LytA, but not for the lactococcal autolysin AcmA. AcmA possesses three repeats of which only one is needed for cell wall binding (Buist, 1997; Yother & White, 1994). Our studies show that the C-terminal repeat of SAC can be deleted without compromising cell wall binding capacity or proteolytic resistance. This indicates that SAC1 is both a structural and a functional unit.

In conclusion, we have demonstrated that the SAC domain, found in the S-layer protein of *L. acidophilus*, in S-proteins of related species and in other surface associated proteins of *Lactobacillus*, represents a novel type of CWBD. It consists of a tandemly repeated sequence, the C-terminal repeat of which can be deleted without compromising binding capacity. This repeat could provide specificity to the cell wall interaction of SAC without possessing a direct binding capacity itself. We further suggest that SAC is a lectin-like domain interacting with a cell wall-associated teichoic acid substituted with galactosyl residues.

Acknowledgement. We thank Jos Seegers for carefully reading of the manuscript.

Chapter 4

Insertion mutagenesis of the S-layer protein crystallization domain of *Lactobacillus acidophilus* ATCC 4356 and surface display of the c-myc epitope

Egbert Smit, Dennis Jager, Beatriz Martinez, Frans J. Tielen and Peter H. Pouwels

Submitted *J. Mol. Biol.*

Summary

The structure of the crystallization domain, SAN, of the S_A-protein of *Lactobacillus acidophilus* ATCC 4356 was analyzed by insertion mutagenesis. Mutant S_A-protein synthesized in *E. coli* with 7-13 amino acid insertions near the N-terminus or within regions of sequence variation in SAN (amino acid position 7, 45, 114, 125, 193) could form crystalline sheets, whereas insertions in conserved regions or in regions with predicted secondary structure elements (positions 30, 66, 88 and 156) destroyed this capacity. An insertion in the cell wall binding domain (position 345) did not affect crystallization. FACscan analysis of *L. acidophilus* synthesizing three crystallizing and one non-crystallizing S_A-protein c-myc (19 amino acids) insertion mutant was performed with c-myc antibodies. Fluorescence was most pronounced for insertions at positions 125 and 156, less for position 45 and severely reduced for position 7. Immunofluorescence microscopy revealed a fluorescent ring in a fraction of the bacterial population, suggesting that these bacteria synthesized mutant S_A-protein only. The finding that the chromosomal *slp*A gene was replaced by the mutant allele in a subset of the population corroborates this conclusion. The capacity of S_A-protein of *L. acidophilus* to present epitopes, up to ~ 19 amino acids in length, at the bacterial surface and the exchange of the wild type by a mutant allele *in vivo,* makes the system, in principle, suitable for application as an oral delivery vehicle.

Introduction

Surface layers or S-layers have been found in up to 400 different species of Eu- and Archaebacteria. They consist of one species of (glyco-)protein, the S-protein, which assembles into characteristic two-dimensional crystalline layers at the cell surface. This assembly is an entropy-driven process during which individual S-protein monomers form multiple interactions with each other and with the under-lying cell envelope (Beveridge, 1994; Sleytr & Messner, 1983). Why certain bacteria possess an S-layer is not always known but they have been shown to function as molecular sieve, scaffold for extracellular enzymes, protective coat or virulence factor (Egelseer *et al.*, 1995; Noonan & Trust, 1997; Sára & Sleytr, 1987a).

S-protein structure-function relationships remain poorly understood due to the lack of suitable methods to determine S-protein structure at the atomic level. Certainly the best understood aspect of S-protein structure is the association of S-proteins with the cell wall. Several cell wall associating domains have been described for S-proteins from *Bacillus*, *Corynebacterium*, *Lactobacillus* and *Thermoanaerobacterium*. Recent studies have shown that many S-proteins contain one or more copies of the so-called Surface Layer Homology or SLH domain for cell surface attachment (Egelseer *et al.*, 1998; Lemaire *et al.*, 1998; Mesnage *et al.*, 1999b; Olabarria, 1996). In contrast, much less is known about the role of structural components or domains in S-protein crystal formation but a few studies have shown that distinct crystallization domains exist in some S-proteins, possibly corresponding to the morphological domains observed in EM studies (Baumeister *et al.*, 1989; Engelhardt & Peters, 1998; Jarosch *et al.*, 2001; Mesnage *et al.*, 2000; Smit *et al.*, 2001).

Several species of the genus *Lactobacillus* possess an S-layer. The S-layers of *Lactobacillus acidophilus* and related species are composed of a single S-protein (S_A-protein) of around 45 kDa (Boot *et al.*, 1996). The S_A-protein of *L. acidophilus* ATCC 4356 and CbsA of *L. crispatus* JCM 5810 are the best studied with respect to structure and function (Boot *et al.*, 1993; Sillanpää *et al.*, 2000; Smit *et al.*, 2001). The function of the S-layer of these organisms is unknown but lactobacillar S-layers may be important for bacterial adhesion to intestinal epithelial cells and extracellular matrix (ECM) components (Schneitz *et al.*, 1993; Toba *et al.*, 1995).

S_A-protein of *L. acidophilus* shows considerable similarity to the putative product of *slpB* encoded by a silent S-protein gene of *L. acidophilus,* to S-proteins from *L. helveticus* and *L. crispatus*, and to a haem-agglutinating protein (HAP50) from *L. acidophilus*, but not to

other S-proteins (Boot *et al.*, 1996; Boot *et al.*, 1995). Analysis of the primary amino acid sequences of these proteins showed that homology between these proteins is highest in the C-terminal one third of the proteins (77-99% identity) and lower in the N-terminal two third (30-72% identity). While the N-terminal region shows a preponderance of hydrophobic amino acids, the C-terminal region is mainly composed of hydrophilic residues, a large fraction of which consists of basic amino acids, rendering this region highly positively charged.

Recently, we demonstrated that the N-terminal and C-terminal parts of S_A-protein constitute different structural and functional domains. The N-terminal part of S_A-protein (amino acids 1-290) constitutes the crystallization domain, SAN, and is able to form S-layer crystals with lattice parameters similar to those of crystals formed by intact S_A-protein. The C-terminal part of S_A-protein (amino acids 291-412), SAC, serves to attach the S-layer to the cell wall (Smit *et al.*, 2001).

S_A-protein is among the smallest S-proteins known, making S_A-protein an ideal candidate for the study of structure-function relationship of S-proteins. Insertion mutagenesis is a method widely used to determine structure-function relationships of proteins (Bingle *et al.*, 1997; Norton *et al.*, 1998; van Geest & Lolkema, 2000; Wong & Hancock, 2000). In the present study we employed insertion mutagenesis and proteolytic treatment to gain further insight into the structural organization of the *L. acidophilus* S-layer protein and its crystallization domain SAN in particular. Our results suggest a structural model for S_A-protein in which SAN is comprised of two subdomains of around 15 kDa each flanking a loop region that is exposed to the environment. To present further evidence for the model DNA sequences encoding the two hypothetical subdomains were expressed in *E. coli* and the properties of the resulting polypeptides were determined. In addition, we show that functional S_A-protein mutants can be used for the efficient surface exposure of a selected epitope making the system, in theory, suitable for application as an oral vaccination vehicle.

Materials and Methods

Bacterial strains and growth conditions. *L. acidophilus* ATCC 4356 and *L. casei* ATCC 393 were cultivated anaerobically in MRS broth (Difco) at 37°C. *Lactococcus (Lc.) lactis* MG1614 was cultivated aerobically in GM17 medium at 30°C. *Escherichia coli* M15 [pREP4] (Qiagen) and *E. coli* DH5α (Phabagen, The Netherlands) were cultivated aerobically in Luria Bertani (LB) broth at 37°C. When appropriate, media were

supplemented with 1,5% agar, 5 µg/ml (*Lc. lactis*) or 7.5 µg/ml (*L. casei* and *L. acidophilus*) chloramphenicol, 100 µg/ml ampicillin and/or 25 µg/ml kanamycin (*E. coli*).

Linker insertion mutagenesis of the *L. acidophilus slp*A gene. A pET5a (Promega) derived vector, pT*slp*A10 (previously named pTA10), containing a *Bam*HI-*Hin*dIII PCR fragment encoding mature S_A-protein (Smit *et al.*, 2001) was used for insertion mutagenesis. Double stranded DNA linkers were inserted in unique restriction sites or in sites occurring twice in the *slpA* gene (Figure 1a). Linkers consisted of *Nco*I, *Not*I and *Xho*I (NcNoX) sites flanked by restriction site-specific 5′ and 3′ sticky ends. Ligation mixtures were introduced in *E. coli* DH5α and clones harboring pT*slp*A10 vector with the desired linker insert, were identified by colony PCR and analysed by DNA sequencing. The resulting plasmids are indicated with pT*slp*A followed by the mutant number as indicated in Table 1 (pT*slp*A11, pT*slp*A12 etc).

Functional analysis of S_A-protein insertion mutants in *E. coli*. Mutant S_A-protein genes were cloned in pQE30ΔXN as *Bam*HI-*Hin*dIII fragments introducing an N-terminal six histidine tag (plasmids are indicated with pH*slp*A followed by mutant number). Mutant S_A-protein gene expression and metal affinity purification were carried out as previously described (Smit *et al.*, 2001). Purified mutant S_A-protein (500 µg/ml) was dialysed exhaustively against 50 mM Tris-HCl, pH 7.5 and checked for precipitation, which is indicative of S_A-protein crystal formation (Smit *et al.*, 2001). Soluble and precipitated S-protein fractions were separated by centrifugation and analysed by SDS-PAGE. Precipitated material was also analysed by electron microscopy (EM) (Smit *et al.*, 2001). In three mutant plasmids pH*slp*A11, pH*slp*A12 and pH*slp*A13 the NcNoX linker was removed by *Nco*I and *Xho*I digestion and replaced by a *c-myc* epitope linker (*Nco*I-*c-myc*-*Xho*I). The mutant proteins were produced and analysed as described for the other mutants.

Construction of plasmid pH*slp*A9c. An additional insertion mutant was constructed by PCR using vector pBK1 (Boot *et al.*, 1993) as template and primers 5′SLPA1 (5′ GCG CGA ATT CAG ATC TAT CGT GGT AAG TAA TAG GAC GTG 3′) and CMYCRE (5′CAG CGA ATT CCT CGA GGT TTA AAT CTT CTT CTG AAA TTA ACT TTT GTT CTG CGT TAA TAG TAG TAG CAG CGC 3′) yielding 5′*slp*A9c. This PCR product was introduced in vector pT*slp*A16-3 (containing *Sal*I, *Bam*HI and *Xho*I sites, after amino acid 7) as *Bgl*II-*Xho*I fragment yielding pT5′*slp*A9c. For purification and functional analysis of the mutant S_A-protein in *E. coli* the *slp* region of pT5′*slp*A9c was amplified using primers CEAMYC1 (5′GGG GGG ATC CGG TAC CGC TAC TAC TAT TAA CGC AGA AC 3′) and CEA2 (5′ CCC CGG ATC CAA GCT TAT CGA AGT ATC AGA AGA TCC TAT T 3′) and the *Bam*HI-*Bst*EII fragment was

transferred to pHs*lp*A10 yielding pHs*lp*A9c. Functional analysis of the purified mutant protein was performed as described above.

Re-introduction of 5' *slpA* expression signals in mutants *slpA*11c, *slpA*12c and *slpA*13c. A fragment including the 5' expression signals of the *slpA* gene (nt -190 to + 150 relative to the ATG), flanked by *Bgl*II (5') and *Hin*dIII (3') sites, was amplified from plasmid pBK1 (Boot *et al.*, 1993) by PCR using primers 5'SLPA1 (5' GCG CGA ATT CAG ATC TAT CGT GGT AAG TAA TAG GAC GTG 3') and 5'SLPA2 (5' GGG GAA GCT TCA GTA GTG CTA CCA GCA GCA G 3'). The PCR product was inserted in pGEM-T and after sequence confirmation excised with *Eco*RI and *Hin*dIII and inserted in pUC19 to give p5'*slpA*. Mutant S_A-protein cassettes isolated from pT*slpA*11c, pT*slpA*12c, and pT*slpA*13c as *Bst*EII or *Pst*I-*Hin*dIII fragments were cloned in p5'*slpA* to yield p5'*slpA*11c, p5'*slpA*12c and p5'*slpA*13c. Clones containing the complete *slp* cassettes were identified by colony PCR and restriction analysis.

Construction of expression vectors for mutant S_A-protein genes in *L. casei* and *L. acidophilus*. To achieve the synthesis of mutant S_A-proteins S_{A9c}, S_{A11c}, S_{A12c} and S_{A13c} in *L. casei* and *L. acidophilus*, the multi-host range vector pLP601-T was used. This vector is derived from pNZ124 (Platteeuw *et al.*, 1994) and contains the inducible promoter of the *L. amylovorus* α-amylase gene controlling expression of the *E. coli* β-glucuronidase (*gusA*) gene (Pouwels *et al.*, 2001). The complete *slp* cassettes were isolated from pT5'*slpA*9c, p5'*slpA*11c, p5'*slpA*12c and p5'*slpA*13c as *Bgl*II-*Hin*dIII fragments and cloned in *Bam*HI/*Hin*dIII digested pLP601-T (which removes *gusA*) to yield vectors pLP*slpA*9c, pLP*slpA*11c, pLP*slpA*12c and pLP*slpA*13c. Ligation mixtures were introduced in *L. lactis* MG1614 by electroporation and correct clones were selected by colony PCR. Highly purified plasmid DNA isolates were prepared by cesium chloride (CsCl) gradient centrifugation, digested with *Not*I to remove the *ldh* terminator, ligated and introduced in *L. casei* ATCC 393. Colony PCR was used to confirm the absence of the terminator sequence. These constructions are shown in figure 1b. The terminator-less vectors were prepared by CsCl gradient centrifugation and introduced in *L. acidophilus*.

Expression analysis of mutant S_A-protein genes in *L. casei* and *L. acidophilus*. *L. casei* and *L. acidophilus* transformants harboring pLP*slpA*9c, pLP*slpA*11c, pLP*slpA*12c and pLP*slpA*13c and untransformed *L. casei* ATCC 393 and *L. acidophilus* ATCC 4356 were inoculated 1:100 from ON cultures into 50 ml MRS medium and incubated for 16 h at 37°C. Cells were collected by centrifugation (25 min, 3,000 x g, 4°C) and proteins in the culture supernatant were precipitated with TCA. The cell pellet was washed once with physiological salt and subjected to 1 M and 5 M LiCl extraction to remove surface

associated proteins as described previously (Smit *et al.*, 2001). The 1 M and 5 M LiCl extracts were dialyzed against distilled water at 4°C.

Exposition of c-myc at the S-layer surface of *L. acidophilus* transformants. Exposition of the c-myc epitope on the S-layer surface by *L. acidophilus* was determined by flow cytometry and immunofluorescence microscopy. Transformants containing pLP*slpA*9c, pLP*slpA*11c, pLP*slpA*12c and pLP*slpA*13c and wildtype *L. acidophilus* cells were harvested at the end of the log phase and washed once with phosphate buffered saline (PBS) supplemented with 1% (w/v) fetal calf serum (FCS). After centrifugation the cell pellet was resuspended in PBS-FCS supplemented with 1% FCS containing anti-S_A-protein or anti-c-myc (1:5000 and 1:50 diluted, respectively) antibodies. After incubation with fluorescein isothiocyanate (FITC)-labeled goat-anti-mouse antibodies, cells were analyzed with a FACSCalibur flow cytometer (Becton Dickinson, San Jose, CA) as described (Pouwels *et al.*, 2001), or inspected with an immunofluorescence microscope.

PCR detection of chromosomal *slp* recombinants. Genomic DNA was isolated from transformants containing expression vectors pLP*slpA*9c, pLP*slpA*11c, pLP*slpA*12c and pLP*slpA*13c. PCR was performed according to standard procedures using primer A9 (forward, *slp* promoter-specific: 5'-CTTGCTATTTCTTGAAGAG-3', position –228 to –210 relative to the *slpA* start codon) and primer CMYCINT (reverse, *c-myc* –insert-specific: 5'-GCGTTTAAATCTTCTTCTGAA-3') and 15 ng of chromosomal DNA as template. To avoid misinterpretation of the results due to a Polymerase Halt-mediated Linkage Of Primers or PHLOP event, PCR was also performed using purified chromosomal DNA obtained from wild type *L. acidophilus* which was mixed with purified pLP*slpA*12c DNA at physiological relevant molar ratios (chrDNA : plDNA ratio's of 1:0.3, 1:3, 1:30 and 1:300).

DNA manipulation, protein analysis. DNA manipulations and protein analysis were performed following standard procedures (Laemmli, 1970; Sambrook *et al.*, 1989). Plasmid DNA was isolated from *Lactococcus/Lactobacillus* according to published procedures (Kok *et al.*, 1984; Posno, 1991). Genomic DNA was isolated from lactobacilli using cetyltrimethylammonium bromide (CTAB) (Towner, 1991). Transformation of *Lc. lactis*, *L. casei* and *L. acidophilus* was performed following published procedures (Kok *et al.*, 1984; Posno, 1991; Walker *et al.*, 1996). DNA fragments obtained after digestion of lambda DNA with *Pst*I (fragment sizes: 11509, 5080, 4649, 4505, 2840, 2577, 2454, 2443, 2140, 1980, 1700, 1159, 1092, 805, 516, 467, 448, 339, 265, 247 and 210 bp) was used as reference marker. Proteolytic treatment of S_A-protein was performed as previously described (Smit *et al.*, 2001).

Results

Insertion of peptide linkers in the S_A-protein of *L. acidophilus*. Our linker insertion mutagenesis strategy yielded ten different *slpA* insertion mutants *slpA*11, *slpA*12, *slpA*13, *slpA*14, *slpA*15, *slpA*16, *slpA*18, *slpA*19, *slpA*20 and *slpA*21, possessing linker insertions (NcNoX; 7 or 8 amino acids) after amino acid 45, 125, 156, 177, 30, 66, 88, 114, 193 and 349, respectively, of the mature S_A-protein. One additional mutant, *slpA*9c, contained a *c-myc* epitope insertion at amino acid 7 instead of the NcNoX linker insertion (figure 1a). The mutant genes and un-modified *slpA* (*slpA*10), were transferred to expression vector pQE30ΔXN and the corresponding proteins S_{HA9c}, S_{HA11}, S_{HA12}, S_{HA13}, S_{HA14}, S_{HA15}, S_{HA16}, S_{HA18}, S_{HA19}, S_{HA20}, S_{HA21} and S_{HA10} were synthesized in *E. coli* and purified by metal-affinity chromatography (Table 1). All mutant genes were expressed at high levels in *E. coli*. SDS-PAGE analysis revealed that the gene products had the expected molecular mass (Mr) of about 46 kDa, slightly larger than wild type S_A-protein (results not shown). For mutant S_{HA14} a truncated peptide of about 25 kDa was observed, the result of a frame-shift mutation already observed during sequencing. Mutant S_{HA9c} showed a Mr of around 47 kDa due to a larger insert, the c-myc epitope.

Functional analysis of S_A-protein mutants synthesized in *E. coli*. Centrifugation of dialyzed mutant S_A-proteins allowed discrimination between crystallized and soluble S_A-protein. Mutant proteins S_{HA9c}, S_{HA11}, S_{HA12}, S_{HA15}, S_{HA16} and S_{HA20} and the wild type protein S_{HA10} formed a precipitate, indicative of S_A-protein crystallization (Smit *et al.*, 2001), mutant S_{HA21} formed less precipitate than the previous mutants, while mutants S_{HA13}, S_{HA18}, S_{HA19} and the truncated S_{HA14} did not show any precipitate formation. SDS-PAGE analysis of the soluble and insoluble fractions confirmed that a substantial part of S-protein was present in the pellet fractions of S_{HA9c}, S_{HA10}, S_{HA11}, S_{HA12}, S_{HA15}, S_{HA16}, and S_{HA20}. For all other mutants, S_A-protein was found only in the supernatant fraction (Figure 2). Electron microscopical analysis of the precipitates formed by wild type (His-tagged) S_A-protein (S_{HA10}) and the precipitating mutants showed crystalline sheets possessing oblique symmetry, similar to that of S_A-protein (data not shown).

Three insertion mutants, *slpA*11, *slpA*12, and *slpA*13, were selected for replacement of the NcNoX linker by a *c-myc* epitope linker. This yielded mutant proteins S_{HA11c}, S_{HA12c} and S_{HA13c} with an insert of 19 amino acids and an Mr of around 47 kDa (Table 1). Introduction of the larger insert did not alter the properties of the mutant proteins. S_{HA11c} and S_{HA12c} could still form S-layer crystals whereas mutant S_{HA13c} was unable to form crystals as determined by EM (data not shown).

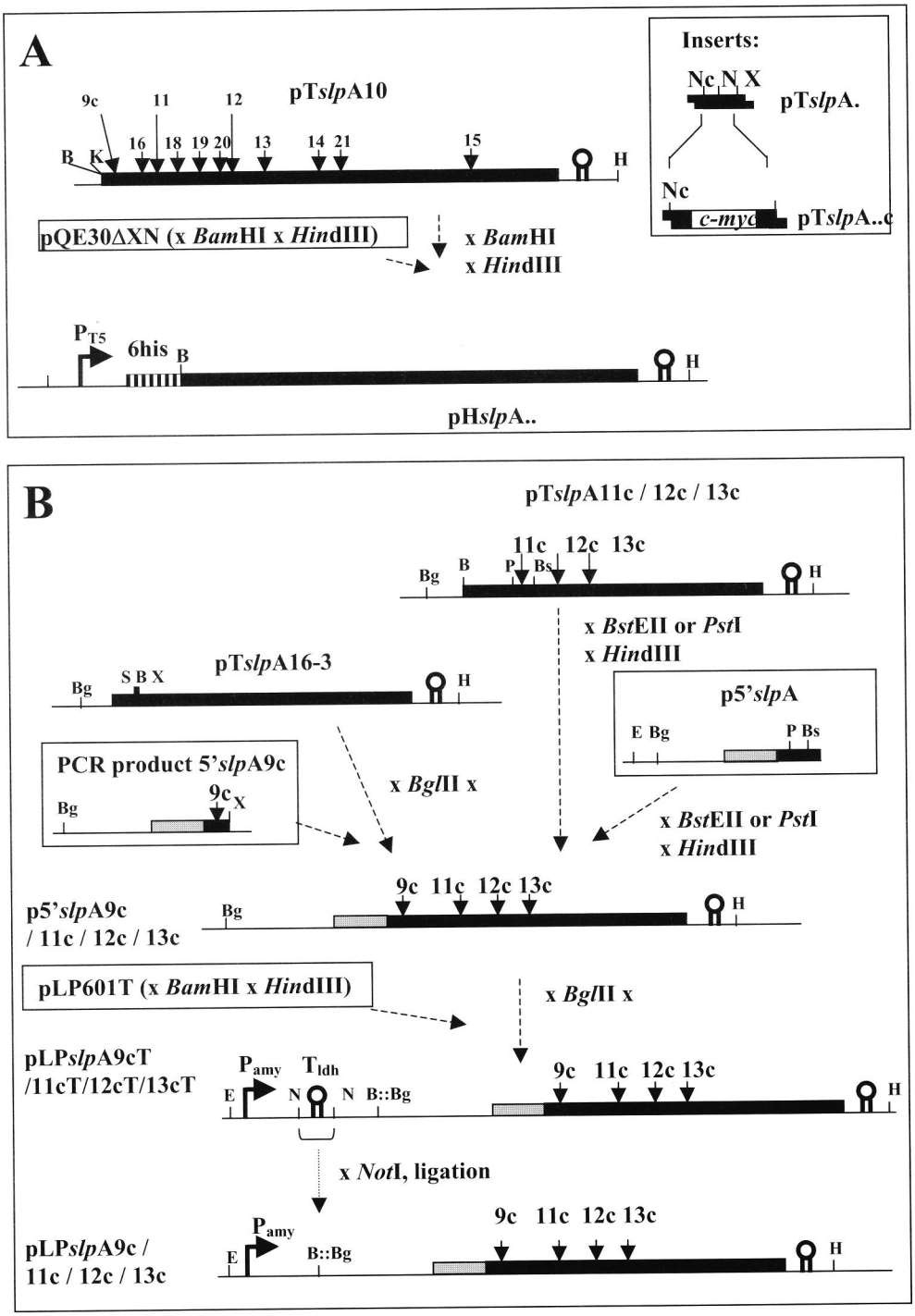

A
pTslpA10
9c 11 12
B K 16 18 19 20 13 14 21 15 H
pQE30ΔXN (x BamHI x HindIII)
x BamHI
x HindIII
Inserts:
Nc N X
pTslpA.
Nc
c-myc
pTslpA..c
P T5
6his B
H
pHslpA..
B
pTslpA11c / 12c / 13c
11c 12c 13c
Bg B P Bs H
x BstEII or PstI
x HindIII
pTslpA16-3
Bg S B X H
p5'slpA
E Bg P Bs
PCR product 5'slpA9c
Bg 9c X
x BglII x
x BstEII or PstI
x HindIII
p5'slpA9c
/ 11c / 12c / 13c
Bg 9c 11c 12c 13c H
pLP601T (x BamHI x HindIII)
x BglII x
pLPslpA9cT
/11cT/12cT/13cT
P amy T ldh
E N N B::Bg
9c 11c 12c 13c H
x NotI, ligation
pLPslpA9c /
11c / 12c / 13c
P amy
E B::Bg
9c 11c 12c 13c H

Figure 1. (Previous page). (A) Linker insertion mutagenesis and *E. coli* expression vector constructions. Sites chosen for linker insertion are indicated by solid vertical arrows. For details on linker insertion positions the reader is referred to table 1. (B) *Lactobacillus* vector constructions. Promoter sequences are indicated by an arrow pointing to the right and a stem-loop indicates the presence of a transcription terminator sequence. Black bar indicates the *slpA* sequence encoding mature S_A-protein and gray bars the signal sequence. Restriction sites are abbreviated as follows: B, *Bam*HI; K, *Kpn*I; P, *Pst*I; Bs, *Bst*EII; H, *Hin*dIII, Nc, *Nco*I; N, *Not*I; X, *Xho*I, S, *Sal*I; Bg, *Bgl*II; E, *Eco*RI.

Table 1. Properties of S_A-protein insertion mutants.

Mutant	Site/type modification	Position[1]	AA composition[2]	Size (AA)	Crystals
10	6 Histag [3]	0	- R G S H_6 G S G T – ATTIN	13	+++
11	*Bst*EII	45	AIAGN - A M A A A R G N – LTGTI	8	+++
12	*Age*I	125	VKFTG - A M A A A R A G – TNSDN	8	+++
13	*Mun*I	156	TNVSI - T M A A A R G I – ANVYA	8	-
14[4]	*Spe*I	177	YDVTS - A M A A A R A S – GATVT	"8"	-
15	*Cla*I	349	NAANI - V A M A A A R G – DGTKR	8	+++
16	*Pst*I	30	PSVSA - S M A A A R G A – VAANT	8	-
18	*Alw*NI	67	LKADT - M A A A R N T – ENATI	7	-
19	*Pst*I	88	AELAA - S M A A A R G A – GVAYT	8	-
20	*Sty*I	114	KTVTL - A M A A A R G L – GSANS	8	+++
21	*Sty*I	193	NADNQ - A M A A A R G Q – VNVAN	8	+
9c	c-myc	7	TTINA – E Q K L I S E E D L N L E - SSSAI	13	+++
11c	c-myc	45	- A M V N E Q K L I S E E D L N A R G N -	19	+++
12c	c-myc	125	- A M V N E Q K L I S E E D L N A R A G -	19	+++
13c	c-myc	156	- T M V N E Q K L I S E E D L N A R G I -	19	-

(1) Amino acid (AA) positions relative to the start of the mature S_A-protein; (2) Original S_A-protein sequences are underlined; (3) This His-tag is present at the N-terminus of all mutants listed; (4) Insertion of 23 bp instead of 24 bp resulted in a frame-shift mutation and truncation of the protein.

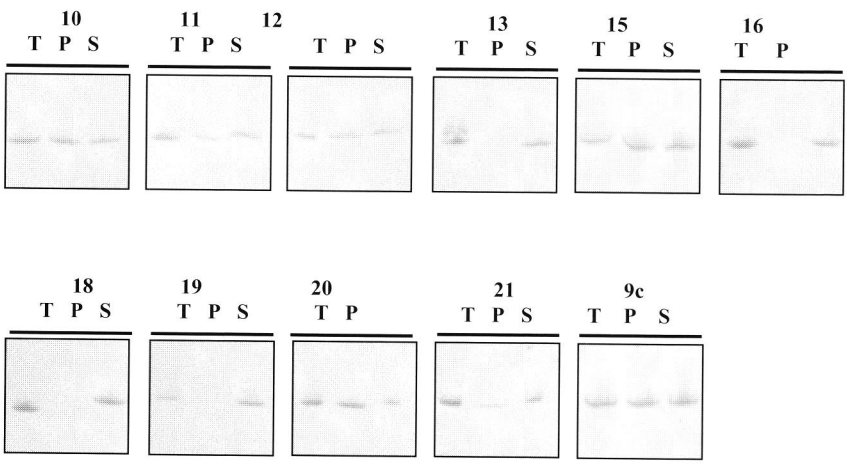


Figure 2. SDS-PAGE analysis of S_A-protein mutants purified from *E. coli*. Total (T), pellet (P) and supernatant (S) fractions obtained by dialysis and centrifugation. The numbers above each panel represent mutant numbers.

Construction of expression vectors for *Lactobacillus* and synthesis of S_A-protein with a c-myc insertion in *L. casei* and *L. acidophilus*. To determine whether the c-myc epitope in mutants S_{HA9c}, S_{HA11c}, S_{HA12c} and S_{HA13c} is buried in the S_A-protein interior or is exposed on the S-layer surface, mutant S_A-proteins were co-expressed with wild type S_A-protein in *L. acidophilus*. For the expression of mutant S-protein genes in *Lactobacillus* new vectors were constructed in which the *slpA*9c, *slpA*11c, *slpA*12c and *slpA*13c genes were combined with the original *slpA* expression signals (the 5' leader sequence for mRNA stabilization, ribosome binding site (RBS), start codon and secretion signal) (Figure 1b). The complete *slp* cassettes were then transferred to the *E. coli-Lactobacillus* shuttle vector pLP601-T, introduced in *Lc. lactis* MG1614 (has a higher transformation efficiency than *L. casei* ATCC 393, so it is more suitable for the introduction of ligation mixtures). After isolation and identification of vectors containing the different *slp* cassettes these were introduced into *L. casei* ATCC 393.

Analysis of mutant *slp* gene expression revealed that in both *Lc. lactis* and *L. casei* all transformants produced mutant S_A-protein irrespective of whether the *ldh* terminator sequence between the promoter and mutant *slp* gene was present or not (results not shown). Mutant S_A-protein production in *L. casei* under inducing conditions from vectors without the terminator was only marginally increased compared to that from vectors with the transcription terminator. The constitutive production of mutant S_A-protein in *Lc. lactis* and *L. casei* under non-inducing conditions from vectors with the *ldh* transcription terminator, suggests the presence of a previously undetected promoter sequence in the 5' region of the *slp* fragments that is active in both *Lc. lactis* and *L. casei*. In further experiments we used this promoter activity for constitutive mutant S_A-protein production in *L. casei* and *L. acidophilus*. After confirmation of the synthesis and secretion by *L. casei* of the mutant S_A-proteins, of their correct relative molecular mass and the presence of the c-myc epitope by SDS-PAGE and Western blot analysis (data not shown), CsCl purified vector DNA (without *ldh* terminator) obtained from *L. casei* was introduced in *L. acidophilus* ATCC 4356. *L. acidophilus* transformants produced less mutant S_A-protein compared to *L. casei*. From *L. acidophilus* transformants producing S_{A9c}, S_{A11c}, S_{A12c} and S_{A13c} mutant protein as well as the wild type strain S_A-protein could be completely removed by 5 M LiCl extraction, indicating that the bacteria contain a mosaic S-layer comprised of wild type and mutant protein. Mutant proteins were not detected in the culture medium. Using SDS-PAGE and Western blot analysis c-myc-containing S_A-protein mutants were detected in the 5 M LiCl extract of the four *L. acidophilus* transformants.

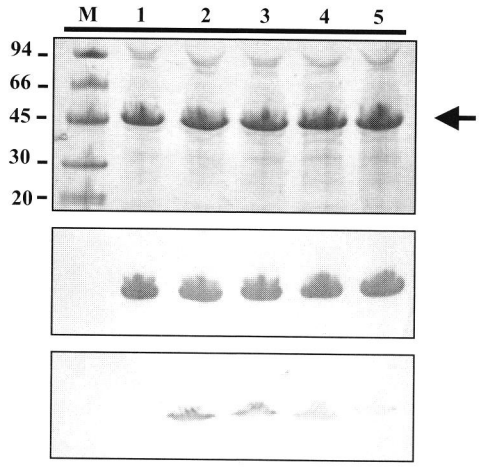


Figure 3. SDS-PAGE and Western blot analysis of surface protein extracts obtained from wild type *L. acidophilus* (lane 1) and transformants producing mutant S_A-proteins S_{A9c}, S_{A11c}, S_{A12c} and S_{A13c} (lanes 2 to 5). The top panel shows a Coomassie stained SDS-PAGE gel, the middle panel a Western blot analysis with anti-S_A-protein and the lower panel a Western blot analysis with anti-c-myc. The arrow indicates S_A-protein (43 kDa) and the Mr's (kDa) of the reference proteins are given in the left margin.

Wild type S_A-protein did not, confirming the presence of the c-myc epitope (results not shown). It was also observed that mutants S_{A9c} and S_{A11c} were produced at equal levels, while the S_{A12c} production level was lower and that of S_{A13c} even more so (Figure 3).

SDS-PAGE analysis of soluble and precipitated S_A-protein fractions obtained by dialysis and centrifugation of the 5 M LiCl extracts revealed that all four S_A-protein mutants associated with both pellet and supernatant fractions, a behavior typically observed for fully functional wild type S_A-protein. We did not observe a preferential association of the non-crystallizing mutant S_{A13c} with the dialysate supernatant fraction although this was observed for the S_{A13c}-protein purified from *E. coli* (data not shown).

Exposition of the c-myc epitope at the cell surface of *L. acidophilus*. Surface presentation of the c-myc epitope was analyzed by flow cytometry and immunofluorescence microscopy. FACScan analysis with c-myc antibodies yielded a signal for each of the *L. acidophilus* transformants although important differences were observed. The detected fluorescent signal increased in the following order: S_{A9c}, S_{A11c}, S_{A13c} S_{A12c} (Figure 4). Interestingly, the cytometry signals were not proportional to the mutant protein production levels, which were lowest for the mutants with the highest fluorescent signals (S_{A12c} and S_{A13c}). The flow cytometry experiments also showed, especially for mutants S_{A12c} and S_{A13c}, a large heterogeneity in fluorescence intensity. Using immunofluorescence microscopy two types of fluorescent cells were observed, one with diffuse fluorescent spots randomly distributed on the cells and another with an intense fluorescent ring completely surrounding the cell (data not shown).

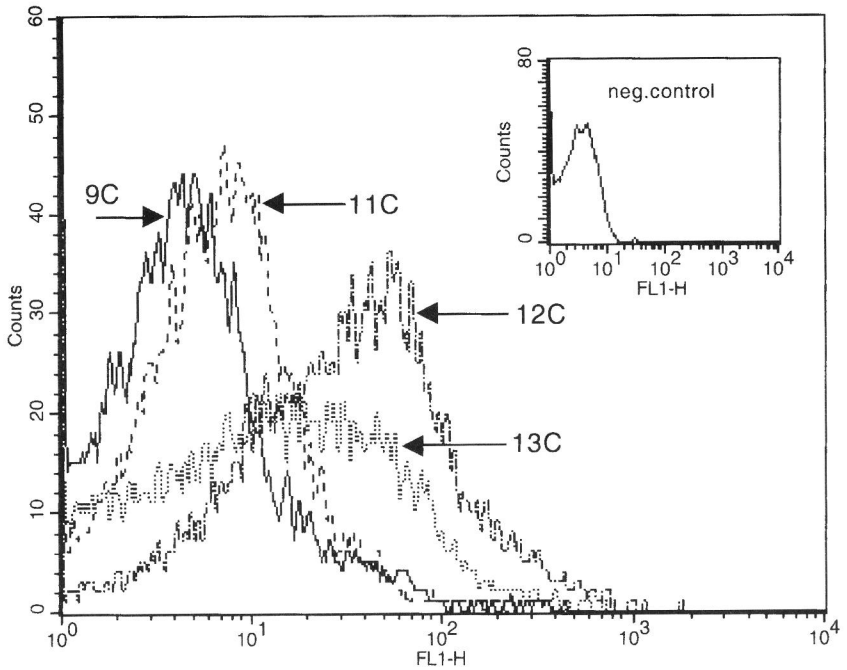


Figure 4. Detection of c-myc at the cell surface of *L. acidophilus* transformants using flow cytometry. Cells were coated with mouse monoclonal anti-c-myc antibodies followed by fluorescent labeling with goat-anti-mouse FITC-conjugated antibodies. Fluorograms of *L. acidophilus* producing S_{A9c}, S_{A11c}, S_{A12c} and S_{A13c} are indicated with 9c, 11c 12c and 13c, respectively. The inset shows the negative control (wild type *L. acidophilus*). The y-axis shows counts x 10^3. Fluorescence intensity as detected by the FL1 detector is shown on the x-axis on a logarithmic scale.

Detection of recombination at the chromosomal *slpA* locus. The S-layer of cells possessing a surrounding fluorescent ring possibly consists solely of mutant S_A-protein since a similar intense ring was observed using wild type cells and anti-S_A-protein serum (Boot, 1996). To determine whether the chromosomal copy *slpA* copy, normally downstream of the *slp* promoter, had been replaced in these cells by the mutant *slpA* gene carried by the expression vector, a PCR strategy was applied. Using genomic DNA preparations isolated from *L. acidophilus* transformants harboring pLP*slp*A9c, pLP*slp*A11c, pLP*slp*A12c, and pLP*slp*A13c we were able to amplify recombination-specific fragments for all four transformants with expected sizes, indicating that the chromosomal *slpA* gene had been replaced by the mutant allele (figure 5). In a parallel control experiment, using wild type *L. acidophilus* genomic DNA mixed with purified pLP*slp*A12c as template, we did not find any PCR product (figure 5).

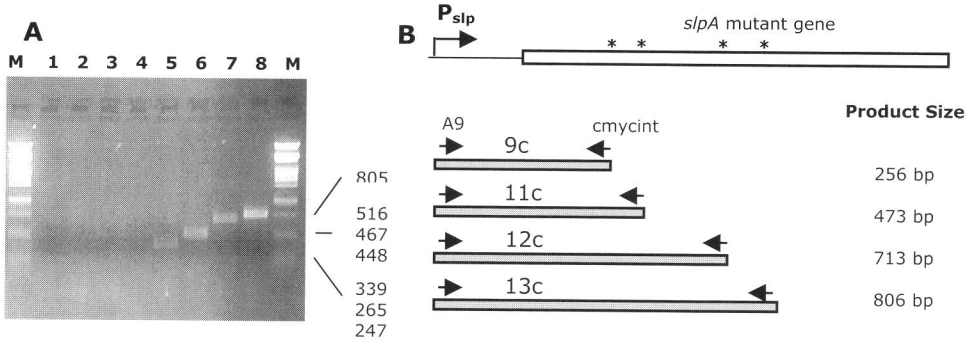


Figure 5. (A) Detection of mutant *slpA* chromosomal integration by PCR analysis using a chromosomal integration-specific primer set (A9 forward primer, *slp* promoter-specific and cmycint reverse primer, c-myc insert-specific). Chromosomal DNA template was isolated from *L. acidophilus* transformants harboring expression vectors pLP*slp*A9c, pLP*slp*A11c, pLP*slp*A12c, and pLP*slp*A13c and the PCR products obtained are shown in lanes 5 to 8. Lanes 1 to 4 show controls in which *L. acidophilus* wild type chromosomal DNA was mixed with purified pLP*slpA*12c DNA at molar ratio's of 1:0.3, 1:3, 1:30 and 1:300, respectively. Lane M: DNA reference marker (lambda DNA digested with *PstI*). (B) Schematic overview of the PCR strategy. *SlpA* mutant genes are represented by a white bar (position of the *c-myc* insertions indicated by *). Gray bars represent the integration-specific PCR products of which the calculated theoretical sizes are given on the right. The *slpA* promoter is indicated with P_{slp}.

Discussion

Traditionally, S-layers are studied by electron microscopical techniques and as a result much is known about their ultrastructure. Little is known, however, about S-protein structure-function relationships. In recent years, with the advent of versatile recombinant DNA technologies, valuable new clues to the structural organization of S-proteins have been obtained. Methods like deletion analysis, sub-cloning of domains, linker mutagenesis and "cysteine-scanning" mutagenesis have been used for this purpose (Jarosch *et al.*, 2001; Mesnage *et al.*, 1999b; Smit *et al.*, 2001).

In the present study we applied linker insertion to clarify the structural organization of the crystallization domain, SAN, of the *L. acidophilus* S_A-protein. Our analysis comprised insertion of seven to thirteen amino-acid residues at eleven positions randomly distributed throughout the S_A-protein, ten of which were located in SAN. All mutant proteins contained an additional N-terminal His-tag (13 amino acids), which did not interfere with S_A-protein crystallization as was previously confirmed by electron microscopy.

Introduction of insertions at positions 30, 66, 88, and 156 of S_A-protein (S_{HA13}, S_{HA16}, S_{HA18} and S_{HA19}) resulted in complete abolishment of the capacity to form crystals *in vitro*.

Mutants with insertions after amino acids 7, 45, 125, 114 (S_{HA9c}, S_{HA11}, S_{HA12}, and S_{HA20}) formed crystals similar to those formed by the His-tagged wild type protein (Smit *et al.*, 2001), while an insertion at position 193 (S_{HA21}) resulted in a partially functional mutant S_A-protein. An insertion in the cell wall binding domain (S_{HA15}) had no effect on assembly, as expected.

For the interpretation of these results multiple S-protein sequence alignments and predicted S-protein sequence motifs (Sillanpää *et al.*, 2000; Chapter 5) were used. The data obtained from these analyses showed that the SAN domain of S_A-protein consists of regions with high variability in composition and length (gap regions) alternating with regions that show considerably less variation. It was proposed that the regions of higher conservation are important for domain structure-function (responsible either for intra- or intermolecular interactions) (Sillanpää *et al.*, 2000; Smit *et al.*, 2001). Disruption of these regions may result in loss of domain function, while regions of variable length and composition are more likely to represent protein surface regions or loops (Miyazawa & Jernigan, 2000) and may accept the insertion of additional amino acid residues. Based on the *Lactobacillus* S-protein and S_A-protein-PrtY alignments presented in chapter 5 we concluded that the conserved N-terminal motifs are important for S-protein monomer integrity, whereas the various loop regions could be involved in protein-protein interaction.

Our present results indicate that, generally, insertions in regions that were well aligned in S_A-protein alignments or contained predicted secondary structure elements were not allowed. This in contrast to insertions in variable, non-conserved regions without any predicted secondary structure elements, which were accepted without disrupting S-layer formation. We highlight several mutated regions.

First the N-terminal region of the S_A-protein. The N-terminus itself is flexible and does not directly play a role in S_A-protein crystallization since it accepts an extension of 13 amino acids without loss of domain function (Smit *et al.*, 2001). Insertion of additional amino acids at position 7 (S_{HA9c}) was probably accepted because of the proximity to the S_A-protein N-terminus. If the insertion was made further away from the N-terminus (amino acid 30; S_{HA16}) then protein function was compromised, possibly due to improper folding of the region or to interruption of protein-protein interactions between S_A-protein monomers. The region seems to be important for S_A-protein crystallization.

The second region in which mutations 12 and 20 were made (positions 125 and 114, respectively), is interesting because of its location in the center of the SAN domain where

insertions of up to 19 amino acids did not affect S_A-protein crystallization. This central region is one of the largest variable region observed in alignments and is the only region in SAN containing sites that are accessible to trypsin and chymotrypsin. Finally, insertion mutant 18 (position 67) showed an important atypical behavior i.e. it produced a non-functional S-protein despite its location in a large loop region. This may signify the importance of some loop regions in protein-protein interaction of SAN.

Flow cytometric and immunofluorescence analysis of mutant S_A-protein-producing *L. acidophilus* showed that the regions near amino acids 125 and 156 (S_{A12c} and S_{A13c}) were strongly reacting with anti-c-myc antibodies, whereas those near amino acids 9 and 45 (S_{A9c} or S_{A11c}) were not. From these results we conclude that the former regions are facing the environment, while the N-terminal region is poorly accessible to antibodies and thus is either buried within SAN or is facing the S-layer pore or cell wall.

Interestingly, mutant S_{A13c} , although non-functional in *in vitro* analyses, forms an integral part of the S-layer *in vivo*, since the mutant protein could be extracted from *L.* acidophilus transformants by 5M LiCl. Moreover, a strong positive signal was observed in the FACS analysis implying that c-myc in S_{A13c} is located at the bacterial surface. Since the S-layer of the transformants is comprised of wild type and mutant S_A-protein, we conclude that S_{A13c} can functionally interact with wild type S_A-protein.

The difference in distribution of the flow cytometry signals, most clearly detectable for *L. acidophilus* transformants producing S_{A12c} or S_{A13c}, suggested a difference in mutant protein production levels within the population of bacteria which was confirmed by immunofluorescence microscopy. As to why some cells produce much more of the mutant protein than other could be explained by the occurrence of an integration event resulting in placement of the mutant *slp* cassette directly downstream of the *slp* promoter. PCR experiments with an integration-specific primer set indeed showed this was the case. We did not determine whether the whole expression plasmid integrated at the *slp* locus or that the wild type and mutant slp cassettes were exchanged only. In the case of a single cross-over event the chromosome contains the complete vector sequence with the mutant *slp* cassette under control of the strong *slp* promoter and the wild type copy under control of the unidentified, but weaker, promoter resulting in the production of a mixed S_A-protein population. The amount of mutant S_A-protein will be higher than that of the wild type because of the stronger promoter directing transcription of the mutant gene and this could explain why such a strong fluorescent signal was detected by immunofluorescence microscopy. This signal could also be generated by an S-layer exclusively consisting of mutant S-protein (when wild type *slpA* has been replaced by a

double cross-over event), but immunofluorescence cannot discriminate between these two situations.

Finally, we have demonstrated the possibility to surface-expose epitopes on the S-layer of *L. acidophilus*. Based on this finding we suggest that the presentation of antigens on the surface of lactobacilli using the S-layer may be useful for the development of oral or nasal vaccines. Lactobacilli possess several additional properties that make them highly suitable for such applications, i.e. their GRAS (generally regarded as safe) status, their immuno-modulating properties and their capacity to evoke mucosal and systemic immune responses against associated antigens (Maassen *et al.*, 1999; Shaw *et al.*, 2000). The observation that mutant S-layer genes spontaneously integrate into the chromosome resulting in a high mutant S_A-protein production, further adds to the application potential of genetically engineered *Lactobacillus* S-proteins.

Acknowledgements. We thank Ton Muijsers for performing the N-terminal sequencing and Alain Brisson for providing EM facilities.

Chapter 5

The crystallization domain of the *Lactobacillus acidophilus* S-layer protein: Evidence for protein-protein interaction of two subdomains

Egbert Smit and Peter H. Pouwels

Summary

The S-layer protein, S_A, consists of a ~15 kDa cell wall binding domain (SAC, 123 amino acids) and a 30 kDa crystallization domain (SAN, 290 amino acids). We have previously shown that SAN forms two-dimensional crystals identical to those formed by the S_A-protein. We have also shown that SAN contains a central region, exposed on the surface of the native S-layer, which could be used to insert a 19 amino acid epitope. In the present paper we describe a more detailed structural analysis of the SAN domain. The central surface-exposed region showed sensitivity to trypsin and chymotrypsin as was determined by N-terminal sequencing of several 14-18 kDa tryptic and chymotryptic peptides. Based on these results we constructed and functionally analyzed a set of N- and C-terminal truncated SAN peptides. Truncation of SAN by up to 177 C-terminal amino acids yielded a stable peptide. Only 113 N-terminal amino acids could be removed to produce a stable peptide. Removal of more residues resulted in unstable peptides that could not be purified. None of the purified peptides was able to form crystalline arrays but in Western blot analysis peptides SAN2 (131), SAN4 (141), SAN6 (177 C-terminal residues removed), SAN7 (113 N-terminal residues removed) and SAN were able to form aggregation products that were stable in 1% SDS even at 100°C. Dimeric products were observed for all peptides, while a trimeric form was also observed for SAN. Further Western blot analysis with S-layer-surface-specific antibodies showed that the surface-exposed epitopes of SAN are mainly located between residues 149 and 159, which coincides with the region of the S_A-protein in which inserted epitopes were surface exposed. These data led us to conclude that S_A-protein consists of two separate subdomains, one of around 113 amino acids (N-domain) and one of 139 to 177 amino acids (C-domain) that are connected by a surface-exposed region. We show for the first time that an S-protein contains functional subdomains responsible for protein-protein interaction.

Introduction

Surface layer or S-layer proteins occur as the most abundant protein at the cell surface of many Gram-positive Eubacteria and also of the Archaea. They are almost universally present in Archaea as an important shape-determining cell envelope component, and in many Eubacteria, where they can serve as molecular sieve, scaffold for extracellular enzymes or can play a role in bacterial adhesion and pathogenesis (Baumeister & Lembcke, 1992; Egelseer *et al.*, 1995; Noonan & Trust, 1997; Sára & Sleytr, 1987a). S-layers usually consist of only one type of protein or glyco-protein monomer arranged in a two-dimensional crystalline monolayer possessing either oblique (p1, p2), square (p4) or hexagonal (p3 or p6) lattice symmetry (Sára & Sleytr, 2000).

Structural studies employing electron-crystallography in combination with computer-aided image analysis have provided morphological information about many S-layers but have not yielded structural information at the atomic level such as can be obtained with X-ray crystallography or NMR. This is due to intrinsic properties of S-layer proteins such as low solubility in aqueous solution and a preference for the formation of two-dimensional but not three-dimensional crystals. As a consequence the structure-function relationship of S-layer proteins and the structural principles underlying S-layer formation are poorly understood.

In recent years many S-protein encoding genes have been cloned and sequenced which allowed extensive analysis and comparison of different S-protein sequences. Strikingly though, S-protein primary structures proved to be very different and few homologous regions were found. The first, and up to the present only, domain found to be conserved among S-proteins from different species is the Surface-Layer Homology (SLH) domain (Lupas, 1996a; Lupas *et al.*, 1994). SLH domains are found in many S-proteins but also in non-S-layer cell surface proteins. This domain has been characterized in detail and was recently shown to anchor S-proteins to peptidoglycan- (PG) associated receptors (Mesnage *et al.*, 2000; Sára, 2001). Other S-proteins such as those of *Lactobacillus acidophilus* and related species share a homologous cell wall binding domain (CWBD) with several lactobacillar extracellular proteinases (Siezen, 1999; Smit *et al.*, 2001; Yamamoto *et al.*, 1999). The S-proteins SbsA and SbsC of *Bacillus stearothermophilus* possess yet another type of CWBD, but this domain and the CWBD from *Lactobacillus* seem to occur in one or a few species only.

In contrast to CWBD's, S-protein domains necessary for crystallization have not been investigated very intensively. Such domains responsible for crystal formation could be

detected in a number of S-proteins and some have been cloned and analyzed. Extensive sequence comparisons have not revealed any conserved motifs in these domains (Jarosch *et al.*, 2000; Kuen *et al.*, 1997) Interestingly, even crystallization domains from S-proteins of closely related species did not yield such motifs. An exception are the S-proteins from *L. acidophilus* and related species since in these proteins several conserved motifs have been found (Sillanpää *et al.*, 2000; Smit *et al.*, 2001). Whether such motifs are important in determining the structure of a single monomer or for S-layer formation, or both, is not known.

Our previous studies have shown that the S-layer protein, S_A, of *L. acidophilus* ATCC 4356 contains a cell wall binding domain (SAC) and a separate crystallization domain (SAN) (Smit *et al.*, 2001). The present work extends the structural analysis of the SAN domain. Here we provide evidence for dimerisation of two subdomains from the *L. acidophilus* S_A-protein. We present a structural model for the SAN domain and for the formation of S-protein/SAN oligomers by dimerisation of subdomains.

Materials and Methods

Bacterial strains and growth conditions. *Escherichia coli* M15 [pREP4] (Qiagen) and *E. coli* DH5α (Phabagen, The Netherlands) were cultivated aerobically in Luria Bertani (LB) medium at 37°C. For overproduction of peptides *E. coli* M15 [pREP4] harboring the appropriate expression vector was cultivated in LB medium at 37°C to an optical density at 695 nm (A_{695}) of 0.6. After cooling of the culture to room temperature (RT) IPTG was added to a final concentration of 1 mM and incubation was continued overnight. When appropriate, media were supplemented with 1,5% agar, 100 µg/ml ampicillin and/or 25 µg/ml kanamycin.

DNA and protein techniques. Plasmid DNA was isolated using Tip-20 or Tip-100 columns (Qiagen) according to the instructions of the supplier. DNA restriction, ligation and introduction in *E. coli* were performed according to standard procedures (Sambrook *et al.*, 1989). DNA fragments were amplified by PCR using SuperTaq DNA polymerase and products were cloned in pGEM-T (Promega). All cloned fragments were sequenced using universal primers. S_A-protein was digested with trypsin and chymotrypsin as described previously (Smit *et al.*, 2001). Peptides were separated by SDS-PAGE and blotted to Immobilon-P membranes (AP Biotech, Sweden) in CAPS/NaOH buffer and those in the range of 14 to 18 kDa were excised from the membrane and subjected to N-terminal sequencing. SDS-PAGE and Western blotting were performed according to

standard procedures using the BioRad Miniprotean™ system (Laemmli, 1970; Sambrook *et al.*, 1989).

Construction of expression vectors. DNA fragments encoding three C-terminally and three N-terminally truncated SAN peptides were amplified by PCR using pBK1 (Boot *et al.*, 1993) as template. Primer CEA1 (5'-GGG GGG ATC CGG TAC CGC TAC TAC TAT TAA CGC AAG TTC –3', forward) was used in combination with primer SANA1 (5'-CCC CGA ATT CAA GCT TTT ATT AGT ATA CGT TTG CAA TTG AAA CAT TAG-3', reverse) to amplify peptide SAN2, SANA2 (5'-CCC CAA GCT TTT ATT ATG AAG CAA CAC CGT TTT GGT C-3', reverse) for peptide SAN4 and with SANA3 (5'-CCC CAA GCT TTT ATT AAC CAA GGG TAA CAG TCT TAC C-3', reverse) for peptide SAN6. Fragments encoding three N-terminally truncated peptides were amplified using primer and SANB1 (5'-GGG GGG ATC CGG TAC CAA AGT TAA GTT AGA CCA AAA CGG TG-3', forward) for peptide SAN3, SANB2 (5'-GGG GGG ATC CGG TAC CCT TAC TAA TGT TTC AAT TGC AAA CG-3', forward) for peptide SAN5 and SANB3 (5'-GGG GGG ATC CGG TAC CTC AGC TAA CTC AAA TGT AAA ATT-3', forward) for peptide SAN7 in combination with primer SAN1 (5'-CCC CGA ATT CAA GCT TTT ATT AAA TTC TCT TGC TTA GCT GGG CTA C-3', reverse, reverse) and pBK1 as template. PCR products were cloned in pGEM-T (Promega) and sequenced. Fragments encoding SAN2, SAN3, SAN4, SAN5, SAN6 and SAN7 were excised as *Bam*HI-*Hin*dIII fragments and introduced in *Bam*HI/*Hin*dIII-digested vector pQE30ΔXN (Smit *et al.*, 2001) introducing an N-terminal six histidine tag. This strategy yielded expression vectors pHSAN2 to pHSAN7.

Purification and functional analysis of truncated SAN peptides. C-terminally truncated peptides SAN2, SAN4, and SAN6, and N-terminally truncated peptides SAN3, SAN5 and SAN7 were synthesized in *E. coli* M15 [pREP4] harboring the appropriate vector and purified in the presence of 6 M guanidinium hydrochloride (GHCl) as previously described for SAN (Smit *et al.*, 2001). Functional analysis consisted of dialysis of purified peptides (500 µg/ml) against several changes of 50 mM Tris-HCl, pH 7.5 at 4°C followed by SDS-PAGE and Western blot analysis.

Purification of an *L. acidophilus* S-layer surface-specific antibody fraction. *L. acidophilus* cells were inoculated 1:100 into 10 ml pre-warmed MRS and cultured till an OD_{695} of 0.35 was reached. Cells were collected by centrifugation (3000 x g, 15 min, 4°C) and washed three times with phosphate-buffered saline (PBS) (the last wash containing 0,01 % NaN_3). 40 µl polyclonal anti-S_A-protein antibodies was mixed with 960 µl PBS (0,01 % NaN_3) and the diluted antibody solution was used to resuspended the *L. acidophilus* cell pellet. The suspension was then incubated at 4°C for 1 hour. After

incubation the cells were pelleted by centrifugation (3000 x g, 15 min, 4°C) and washed with PBS. Cell-bound antibodies were extracted by resuspending the cells in 500 µl 0.1 M glycine, pH 2.5. The suspension we centrifuged directly (3.000 x g, 15 min, 4°C), the supernatant collected and centrifuged again (10.000 x g, 2 min, room temperature). After the second centrifugation the supernatant was directly neutralized by the addition of 200 µl 0.5 M Na_2HPO_4, pH 7. The S-layer surface-specific (SS) antibodies purified in this manner were stored in 100 µl at – 20 °C and used for western blots 1:500 diluted (compared to 1:20.000 for the polyclonal serum).

Protein primary structure analysis. Alignments of protein sequences were generated with DNAMAN 4.1 (Lynnon Biosoft). The deduced amino acid sequences of following S-protein genes were used: *L. acidophilus* ATCC4356 *slpA* and *slpB* (accession number: X71412 and X89376); *L. helveticus* CNRZ 892 *slpH* (X91199); *L. crispatus* LMG 9749 *slpTA* and *slpTB* (B. Martinez, personal communication); *L. crispatus* LMG 12003 *slpNA* and *slpNB* (AF253043/AF253044); *L. crispatus* JCM 5810 *cbsA* and *cbsB* proteins (AF001313/AF079365); *L. crispatus* M247 s-layer protein gene (AJ007839); *L. crispatus* NCFB5 S-layer protein gene (L. Morelli, unpublished). Protein sequence alignments were submitted to the PredictProtein server (http:\\www.embl-heidelberg.de/predictprotein) for secondary structure, accessible surface area (ASA), coiled coil and signal sequence cleavage site prediction (Lupas, 1996b; Nielsen *et al.*, 1997; Rost, 1996).

Results

Identification of a major protease-sensitive region in the SAN domain

We previously reported that trypsin and chymotrypsin digestion of S_A-protein yielded a range of degradation products, varying in size from 30 to 14 kDa. These fragments were shown to represent intact SAN and degradation products of SAN (Smit *et al.*, 2001). We now subjected the smallest fragments (14-18 kDa) to N-terminal sequencing. The excised peptide bands yielded in most cases mixtures of two sequences. The N-termini of the tryptic peptides were either ATTIN (S_A-protein residues 1 to 5) or VKLDQ (residues 140 to 144), while those of chymotryptic peptides were either ATTIN (residues 1 to 5) or AINTT (residues 160-164). Apparently, only sites in the central region of SAN are susceptible to proteolytic attack despite the presence of many other potential cleavage sites in SAN.

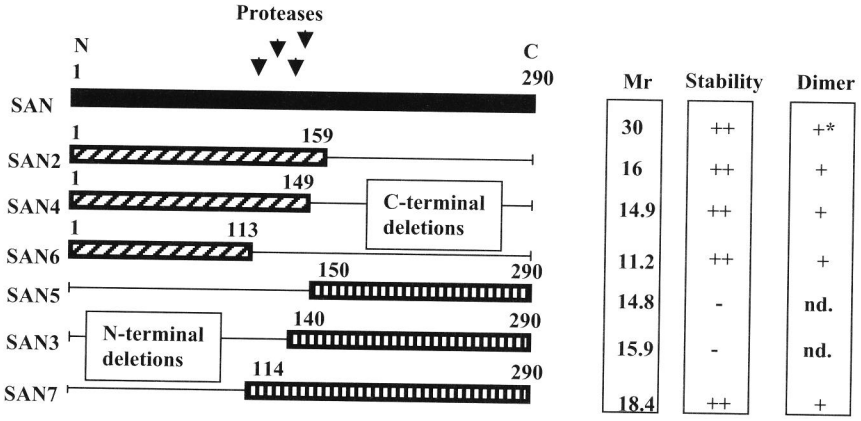


Figure 1. Properties of SAN and SAN truncated peptides. The position and size of the truncated peptides are indicated as well as their Mr (kDa), resistance to *E. coli* proteases and ability to form dimers. *: for SAN trimers were also observed. All peptides possess an N-terminal 6 histidine tag which is not shown. Nd: not determined.

Production and purification of truncated SAN peptides. We designed a series of six different truncated SAN peptides (figure 1) and produced them in *E. coli*. Figure 2 shows the *E. coli* expression levels and the yields obtained for the C-terminally truncated peptides SAN2, SAN4 and SAN6 and the N-terminally truncated peptides SAN5, SAN6 and SAN7. SAN2, SAN4, SAN6 and SAN7 were produced well and could be easily purified while SAN3 and SAN5 were produced at very low levels and could not be purified in large amounts. Purified SAN3 and SAN5 also proved highly sensitive to proteolysis upon storage.

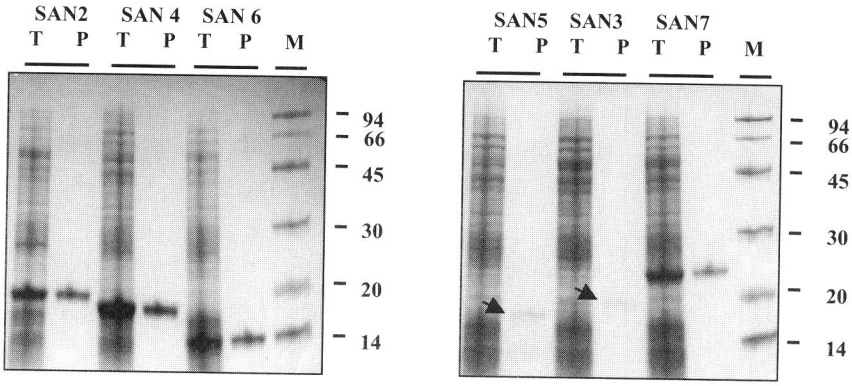


Figure 2. Production of N-and C-terminally truncated SAN peptides in *E. coli* and purification from lysates. Truncated peptide numbers are indicated. t: total cell lysate, p: purified preparation. The arrows indicate the positions of purified SAN5 and SAN3. M: reference proteins, Mr's indicated in the right margins in kDa.

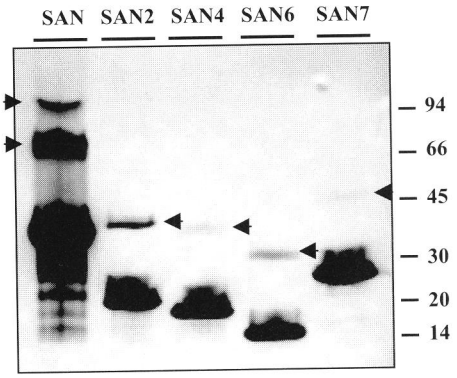


Figure 3. Western blot detection of dimeric and oligomeric aggregates of SAN and truncated SAN peptides. Truncated peptide numbers are indicated and SAN by S. The arrows indicate dimeric and trimeric peptide forms. The Mr's of the reference proteins are indicated in the right margin in kDa.

Detection of aggregated forms of SAN and truncated peptides. Purified wild type S_A-protein was previously shown to precipitate during removal of the denaturing agent used for extraction (5 M LiCl; Smit *et al.*, 2001). This precipitate was shown to consist of crystalline sheets and stacked sheets by electron microscopical (EM) analysis. Since we did not observe precipitate formation of any of the truncated SAN peptides (they are unable to form high molecular weight assembly products)we did not perform EM analysis. For this we the method described below to determine whether the truncated peptides had retained the capacity to form low molecular weight assembly products. In the Western blot analysis with polyclonal anti-S_A antibodies we detected several lower mobilities bands for SAN, SAN2 and SAN7, suggestive of the presence of aggregated forms of these peptides. Further analysis with monoclonal anti-RGS-His6 antibodies showed that such bands could also be detected for SAN, SAN2, SAN4, SAN6 and SAN7 (figure 3), but not for SAN3 and SAN5 (not shown). For SAN at least two additional bands were observed with estimated Mr's of around 70 kDa and 100 kDa, for SAN2 one of around 40 kDa, for SAN4 one of 38 kDa, for SAN6 one of 32 kDa and for SAN7 one of 48 kDa. The absence of such bands for SAN3 and SAN5 could be due to the small amount of peptide that was used in SDS-PAGE analysis.

Detection of surface located regions in the S-protein. To determine the presence and location of S-layer surface-specific epitopes in the different peptides we separated them by SDS-PAGE and subjected them to Western blot analysis with anti-S_A-protein (SP) antibodies or S_A-protein surface-specific (SS) antibodies (figure 4). Both antibody preparations bound equally well to SAN and SAN2, while binding of SAN4, SAN5, SAN3 and SAN7 by SS antibodies was clearly reduced compared to the binding by SP antibodies. SAN6, although present in similar amounts as SAN2 or SAN, bound no SS antibodies at all.

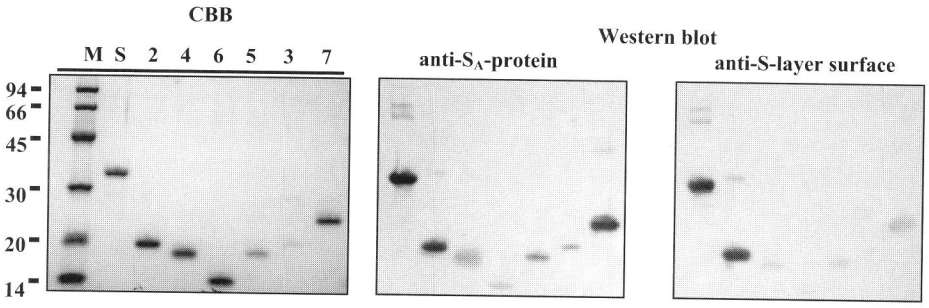


Figure 4. SDS-PAGE and Western blot analysis of purified SAN and truncated SAN peptides. The left panel shows the coomassie (CBB) stained gel. The middle and right panel show Western blot analyses with unpurified (SP) and S-layer surface-specific antibodies, respectively. Truncated peptide are indicated by numbers and SAN by S. M: reference proteins, the Mr's are indicated in the left margin in kDa.

Analysis of SAN primary and secondary structure. Secondary structure prediction of four aligned *Lactobacillus* deduced S-protein amino acid sequences has previously shown that the SAN domain almost exclusively consists of β-strands and only two α-helices, predicted in the C-terminal half (Smit *et al.*, 2001). In addition, several high-identity regions were observed representing conserved S-protein motifs.

Figure 5 shows an extended alignment of SAN with the corresponding regions from ten other related S-proteins. Regions in which conserved amino acid residues are clustered could be easily identified and these clusters mostly corresponded to the S-protein motifs defined previously (Sillanpää *et al.*, 2000; Smit *et al.*, 2001). Table 1 lists the motifs as we find them in the present alignment. Although within motif 4 there are sections of relatively low homology (such as amino acids 147 to 168 of SAN), we have included it in the list since all eleven sequences are well-aligned in this region and possesses a well-predicted secondary structure (see below).

Motif	Sequence (AA)	Position (AA)
1	VDVTPSVS	21-28
2	GNLTGTISASYNGKTYTANL	43-62
3	AGVAYTVTNDVSFNFGSENAGKTVTL	87-113
4	TLKVKLD -> INSKYFAAQ	137-212
5	KAALKDQ -> TVASVSKRI	230-290

Table 1. Conserved amino acid motifs observed in the SAN domain of *L. acidophilus* S_A-protein and the corresponding domains from related S-proteins. Motifs are numbered according to their position in the SAN sequence. See also figure 5: black underlined residues.

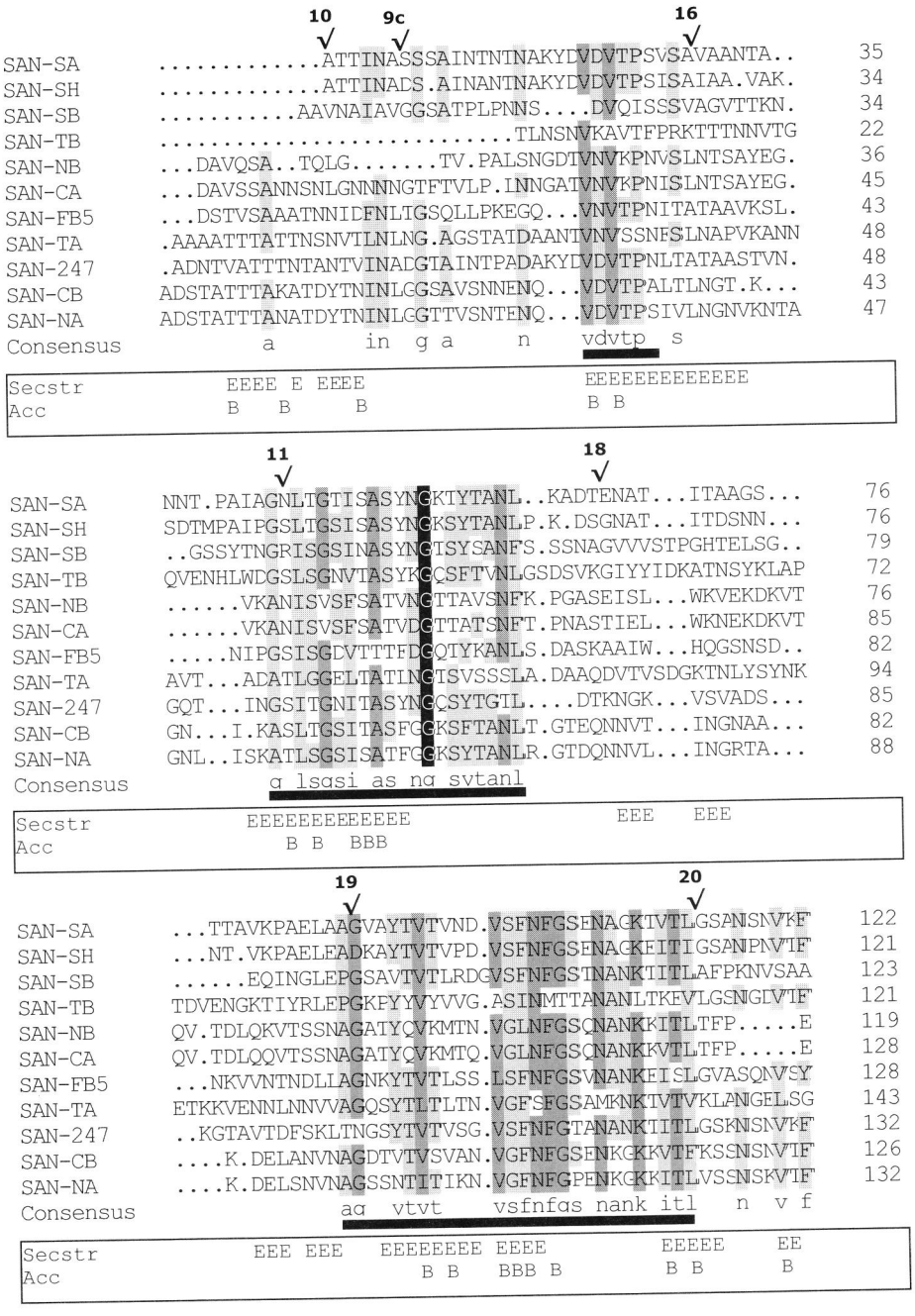


Figure 5 (continued next page).

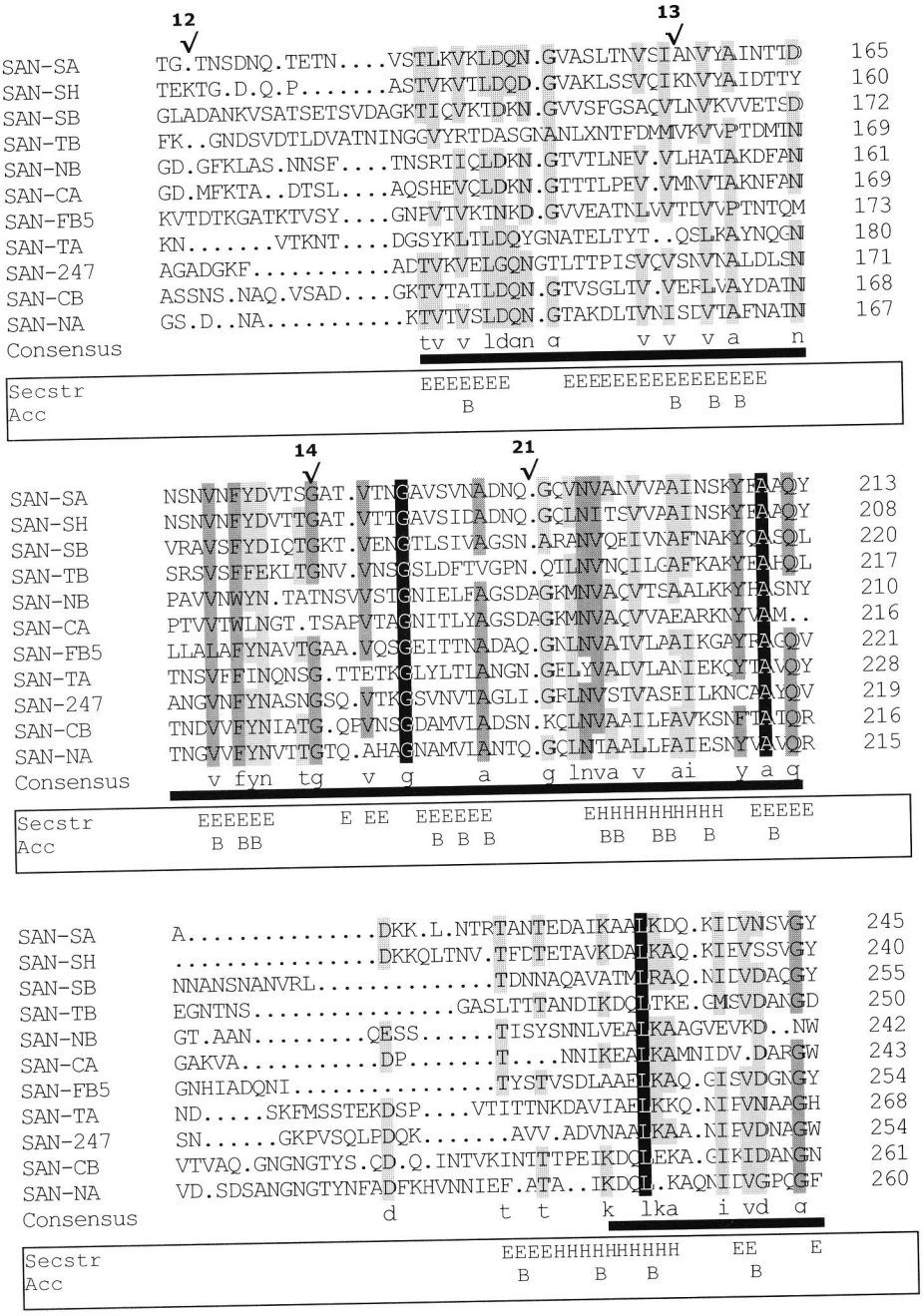


Figure 5 (continued next page).

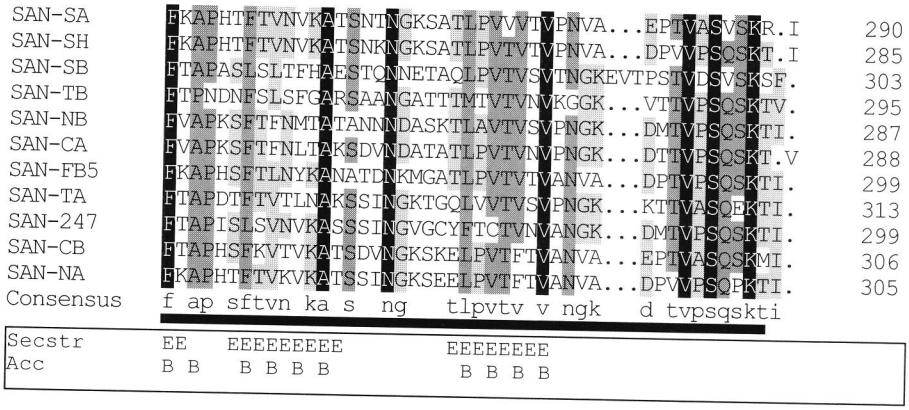


Figure 5 (Previous pages). Alignment, secondary structure and accessible surface area prediction of the deduced amino acid sequences of SAN and corresponding regions of ten other *Lactobacillus* S-proteins. SA/SB: *L. acidophilus* ATCC 4356 S_A-protein/S_B-protein; SH: *L. helveticus* CNRZ 892 S_H-protein; TA/TB: *L. crispatus* LMG 9749 S_{TA}-protein/S_{TB}-protein; NA/NB: *L. crispatus* LMG 12003 S_{NA}-protein/S_{NB}-protein; CA/CB : *L. crispatus* JCM 5810 CbsA/CbsB proteins; 247: *L. crispatus* M247 S-layer protein; FB5: *L. crispatus* NCFB5 S-layer protein. For proteins with suffix A and B: A indicates expressed gene and B non-expressed gene. Identical residues are shown in black, >75% conserved residues in dark gray and 75% - 50% conserved residues in light gray shading. Consensus sequence (> 50% conserved) is given below the alignment in lower case character. Below the consensus the alignment based predicted secondary structure is indicated. E: β-strand and H: α-helix. Below the secondary structure prediction the predicted solvent accessibility (Acc) per amino acid residue is given. B indicates a buried residue. Insertion mutant positions are indicated with mutant number and '√' (from table 1, chapter 4). Conserved S-protein motifs are underlined by a black bar (see also table 1).

Secondary structure and per residue accessible surface area (ASA; see next paragraph) were predicted based on the new alignment. The secondary structure analysis showed sixteen β-strands (of more than three residues) and two α-helices. The predicted β-strands and α-helices are indicated in the alignment and were found to largely coincide with well-aligned and conserved regions. The region containing the two α-helices (amino acids 198 to 235 of SAN) was analyzed for the presence of a coiled-coil forming sequence pattern (a known motif responsible for dimerization and oligomerization of proteins, (Lupas, 1996b)) but such a structure was not observed.

ASA analysis based on a set of related protein sequences can give a reliable prediction, which amino acids are likely to be buried in the protein interior. Analyzing the present alignment showed that overall buried residues were associated with the predicted β-strands and α-helices. Most of the buried residues corresponded to well-conserved hydrophobic residues and some polar amino acid residues. Especially the last two β-strands and segments of several others also show an alternating pattern of exposed and buried amino acids suggestive of amphipathic β-strands, which are likely to form a beta-sheet.

In our previous work we observed a high level of homology between the C-terminal regions of the S-proteins and those of several extracellular proteinases (Smit *et al.*, 2001). Recently, the sequence of a new small extracellular proteinase, PrtY, was described, which not only possesses a C-terminal region homologous to SAC (Yamamoto *et al.*, 1999), but also an N-terminal region that showed strong homology (35% identical residues) to the SAN domain. Alignment of the N-terminal sequences of SAN and the PrtY proteinase showed several regions of high homology that correspond to the conserved motifs observed in the S-protein alignment, but which did not include the active site residues of the proteinase (6 figure; table 1).

Figure 6. Alignment of deduced amino acid sequences of SAN and the N-terminal region of the *L. helveticus* CP790 extracellular proteinase PrtY. Identical residues are shown in gray shading. PrtY putative active site residues serine (S) and aspartic acid (D) are indicated by arrows.

Discussion

In the present work we analyzed the products that are formed by proteolytic treatment of S_A-protein to obtain structural information of the crystallization domain, SAN. In addition, we determined the presence of surface-exposed regions and the subdomain organization of SAN, by the production and characterization of N- and C-terminally truncated peptides. The design of the truncated peptides was primarily based on S-protein sequence comparisons and results from our SAN insertion mutagenesis study (present work; Smit *et al.*, 2001). Functional analysis of the purified peptides consisted

of the formation of crystalline structures or oligomeric forms after removal of the chaotropic agent.

N-terminal sequencing of 14-18 kDa peptides obtained after mild tryptic and chymotryptic digestion of S_A-protein revealed that SAN contains protease-accessible peptide bonds in the central region (residues K_{139} and Y_{159}). These residues are located quite close to a large variable region just upstream, including residues 114 to 136, that was previously identified in *Lactobacillus* S-layer protein sequence alignments. In chapter 4 we report that small inserts (up to 19 amino acid residues) in this region do not affect S_A-protein crystal formation *in vitro* and are S-layer-surface exposed *in vivo*. Based on these data and our new observations we conclude that the S_A-protein central region, encompassing at least residues 114 to 159 is a flexible and surface-exposed region not directly involved in either S-protein monomer structural integrity or S-protein-S-protein interaction. This conclusion is in agreement with surface exposure of epitopes inserted at positions 125 and 156 (mutant 12 and 13, chapter 4).

Based on the assumption that the central region of S_A-protein is surface-exposed we designed, produced and purified several truncated SAN peptides to determine whether SAN can be divided into subdomains. During the production in *E. coli* and purification from cell lysates large differences in stability were observed between these peptides. Deletion of residues 114 to 290, 150 to 290 or 160 to 290 resulted in stable peptides (SAN6, SAN4 and SAN2, respectively) that could be produced in large amounts and purified easily.
Peptides with deletions of residues 1 to 149 and 1 to 139 (SAN5, and SAN3), however, were produced at low levels and purified with low yields, in contrast to a peptide with a deletion of residues 1 to 113 only, which was produced and purified in large amounts (SAN7). Purified SAN5 and SAN3 proved highly unstable upon storage and during analyses, which was not the case for SAN7. Thus the regions covering amino acid residues 1 to 113 and 114 to 290 could represent two structured subdomains of SAN.

The issue whether the truncated SAN peptides are functional entities was analyzed by measuring the ability of the peptides to form oligomeric or possibly multimeric forms. S_A-protein and SAN formed multimers (i.e. two-dimensional crystals) in the absence of chaotropic agents. Such multimers could be easily identified by their insolubility in water and their structure by EM analysis (Smit *et al.*, 2001). In this study we observed dimeric and oligomeric forms for S_A-protein and SAN after SDS-PAGE and Western blot analysis. Thus S_A-protein and SAN remain associated even after boiling in sample buffer (1% SDS). For the peptides SAN2, SAN4, SAN6 and SAN7 we also detected dimeric forms, but

no higher Mr forms, after similar treatment. Thus we conclude that the four peptides all represent functional SAN fragments. The observation that boiling in 1 % SDS does not completely disrupt the interactions between S_A-protein or SAN monomers as well as dimers of truncated peptides SAN2, SAN4, SAN6 and SAN7 indicates that strong interactions exist between monomers and also between these peptides.

The results of SDS-PAGE and Western blot analysis, in combination with our observations concerning the stability of truncated peptides produced in *E. coli*, indicate that the N-terminal 113 amino acid residues and the C-terminal 177 amino acids constitute structural and functional subunits and thus represent SAN subdomains (N- and C-subdomain, respectively). The C-terminal boundary of the N subdomain probably lies not too far upstream from residue 113, since the best conserved SAN motif is located directly upstream. The N-terminal boundary of the C domain is at present less well defined, but lies between residues 114 and 140, since the transition from stable to unstable peptide occurred upon deletion of these residues. The N- and C subdomains with probable Mr's of $\sim$ 11 kDa and $\sim$ 18 kDa, respectively, could correspond to the two morphological domains observed in our previous EM studies (Smit *et al.*, 2001). Since we found that the truncated peptides could form dimers, we conclude that the N subdomain and C subdomain both contain intact protein-protein interaction site(s).

We determined the distribution of surface-exposed residues in SAN and in the truncated peptides using anti-S_A-protein antibodies and purified S-layer surface-specific antibodies. Binding of the two antibody preparations was identical for SAN and SAN2, but binding of SAN4, SAN6, SAN5, SAN3 and SAN7 by the S-layer surface-specific antibodies was significantly reduced compared to binding by the non-purified antibodies. The observation that SAN2 bound surface-specific antibodies as well as SAN, shows that major surface-exposed epitopes are included in the SAN2 peptide. Since deletion of residues 149-159 resulted in a significant reduction of antibody binding, we conclude that major surface exposed epitopes are located in this region. This is in agreement with the results of proteolysis experiments (this chapter and chapter 2) and surface exposure of an epitope inserted at position 156 (mutant 13, chapter 4). Interestingly, SAN3 and SAN7, which have an overlap with the SAN2 peptide that includes the 149-159 region, did not show high reactivity. This leads us to conclude that the epitopes are either conformational or discontinuous and require residues upstream of position 114 for antibody recognition.

We also re-analyzed the sequences of SAN and those from corresponding regions of ten other S-proteins for conserved sequence motifs and secondary structure elements. Since we included eleven instead of four sequences with varying levels of sequence identity,

the results of the present analysis will be more accurate than before (Rost & Sander, 1994). In our present analysis we excluded both the N-terminal signal sequence and the highly conserved C-terminal sequences (SAC) since these are of great influence on the results, especially in the border regions. Generally, we observed that conserved sequences were included in predicted secondary structure elements. In the present analysis one additional short sequence motif (residues 21 to 30), included in a predicted β-strand, was identified in the N-terminal region. The N-terminal regions of all SAN domains up to this first motif are highly variable. The data presented here confirm an extend those observed by Sillanpää *et al.* (2000) and Smit *et al.* (2001). The S-protein motifs represent the best conserved sequences, but conservation also extends beyond these regions. At present we define five conserved regions and four highly variable gap regions (figure 5). Since the best conserved residues in an alignment usually constitute amino acids that are crucial for proper protein folding and function (Ladunga & Smith, 1997), we suggest that these motifs play an important role in the structural integrity of the SAN subdomains. The variable regions most likely correspond to surface exposed sequences. Prediction of the residue accessible surface area (ASA) of SAN indeed showed that most of the amino acids classified as buried correspond to well-conserved hydrophobic residues included in the conserved regions. This further supports our suggestion that these conserved residues may be involved in packing of structure elements in the subdomain interior and may be important for structural integrity.

The homologous sequences of the SAN domain and the PrtY proteinase include all the S-protein sequence motifs previously described, while the regions that include the active site residues of the proteinase are markedly different from the corresponding SAN sequences. The presence of S-protein motifs in PrtY suggests that this proteinase possesses a similar structure as the S-protein. Since the proteinase does not form a crystalline structure (N. Yamamoto, personal communication), the motifs must be important for S-protein and proteinase structural integrity rather than monomer-monomer interactions, in full agreement with the conclusion based on our secondary structure analysis of S-proteins. Which region(s) in S-protein differing from those in the proteinase are essential for the capacity to form oligomers remains to be determined. The observation that introduction of an insertion at position 67 in the large loop region in the SAN N-terminal subdomain (residues 64 and 86) resulted in a non-aggregating S-protein (Chapter 4) might suggest that this region plays role in monomer-monomer interaction.

In conclusion, we here provide new insights into the structure and properties of the S_A-protein crystallization domain and thereby into the S_A-protein crystallization process. We

show that SAN forms dimers and trimers as stable assembly products. We further show that SAN consists of two separate subdomains linked by an S-layer surface-exposed region. Since these subdomains are able to form strongly associated dimers, we suggest that the first step in S_A-protein crystallization, the formation of oligomeric assembly products (Sleytr *et al.*, 1996), is initiated by dimerization of the subdomains.

Acknowledgements.
We thank Ton Muijsers for the N-terminal sequencing of peptides. We thank L. Morelli and B. Martinez for kindly providing the sequences of some unpublished S-protein genes.

Chapter 6

The *Lactobacillus acidophilus* S-layer protein assembles in an ordered array on a gold surface. Implications for the development of a biosensor coating

Egbert Smit, Edwin Stigter, Heidi Diettrich, Stefan Past, Peter Pouwels

Summary

In this chapter we describe some *in vitro* adhesion properties of the surface layer (S_A) protein of *Lactobacillus acidophilus*. We used surface plasmon resonance (SPR) to determine the adsorption characteristics of S_A-protein to a solid gold substrate. When S_A-protein was allowed to interact for a few minutes, the adsorbed material could be removed by treatment with 1M NaCl or 25mM NaOH. After incubation for 3 hours adsorption of S_A-protein was resistant to treatment with these agents. The efficient adsorption of S_A-protein to the moderately hydrophobic gold surface suggests that interaction takes place via the hydrophobic outer face. The observation that immobilized S_A-protein does not react with surface-specific antibodies but strongly reacts with polyclonal antibodies corroborates this conclusion. From our data we conclude that S_A-protein is bound in a highly ordered fashion. Chemical modification of carboxyl and amino groups in S_A-protein did not affect the adsorption properties indicating that these groups are not directly involved in S_A-protein adsorption. Biotin and fumonisin B1 could be exposed at the surface of the adsorbed S_A-protein layer via carboxyl or amino groups. An antibody-based inhibition assay allowed us to determine fumonisin B1 in concentrations as low as 20 ng/ml. The potential of S_A-protein biocoatings in biosensor technology is discussed.

Introduction

Surface layers or S-layers are found as a major cell surface component in many bacteria. They consist of a single protein or glycoprotein subunit forming two-dimensional crystalline arrays. For most Archaea the S-layer is an essential shape-determining and -maintaining cell envelope component. Eubacterial S-layers have been reported to function as molecular sieve or scaffold for extracellular enzymes (Egelseer *et al.*, 1995; Sára & Sleytr, 1987a). They may also have a role in adhesion to epithelial surfaces or extracellular matrix components (Garduno *et al.*, 2000; Toba *et al.*, 1995).

Electron microscopical (EM) studies have shown that the outer surface of S-layers tends to be smooth and that the inner surface has a corrugated appearance. Also with respect to physico-chemical properties S-layers are anisotropic structures and generally have a relatively hydrophobic outer surface and a more hydrophilic inner surface. Many S-layer proteins are easily extracted by treatment with chaotropic agents or high salt and can be reassembled *in vitro* into arrays identical to those present on the bacteria (Sleytr *et al.*, 1999). Reassembly can be performed in solution or at interfaces like lipid monolayers, an air-water interface, on liposomes or on solid substrates such as mica, gold or glass. S-proteins assemble on a substrate with a specific orientation dependent on the physico-chemical properties of the substrate. Due to the repetitive structure of S-layers they have a very regular distribution of functional groups that can be used for chemical modification and coupling of for example antibody fragments, enzymes or antigens (Pum, 1993; Sára, 1993; Sleytr *et al.*, 1999). S-protein assembly products have been used in several industrial applications such as isoporous S-layer Ultrafiltration Membranes (SUMs) and as support matrix in biosensors and immuno-assays (Sleytr & Beveridge, 1999).

Our studies focus on the structure and properties of the S-layer of *Lactobacillus acidophilus*. *L. acidophilus* is a constituent of many probiotic food products. Due to its surface-location the S-layer could play a role in adhesion to and persistence at mucosal surfaces, processes which are thought to be of great importance for the exertion of health-promoting properties of a probiotic bacterium (Dunne *et al.*, 1999). As a result there is great interest to increase understanding of *Lactobacillus* cell surface properties that mediate adhesion and persistence.

Our previous studies have shown that the S-protein (S_A) of *L. acidophilus* can be assembled at an air-water interface and on lipid monolayers (Smit *et al.*, 2001). Assembly at the hydrophobic air-water interface resulted in an S-layer outer face-to-air orientation suggesting that the outer face is more hydrophobic than the inner face. The

orientation of assembly at lipid layers varied depending on the lipid head group. On negatively charged lipids S_A-protein assembly occurred with an inner face-to-lipid orientation. This latter finding was explained by the surplus of positively charged amino acid residues in the cell wall binding domain of the S_A-protein which is thought to bind anionic cell wall-associated polymers (Smit & Pouwels, 2002 (chapter 3)).

Surface Plasmon Resonance (SPR) is a technique that allows the measurement of the interaction of biological molecules with or at surfaces in real-time. In the present work we used SPR to analyze the adsorption of the *L. acidophilus* surface layer protein to gold surfaces and the characteristics of the adsorbed protein layer. Since there is great interest in stable, reusable and easily obtainable sensor coatings in the field of biosensor research, we studied the possibilities of application of gold-adsorbed S_A-protein as a coating (Brown, 1997; Woodbury *et al.*, 1998).

Materials and Methods

Chemicals and reagents. 1-Ethyl-3-(3-dimethylamino-propyl)carbodiimide hydrochloric acid (EDC) was obtained from Fluka (Switzerland). 2-(N-morpholino) ethanesulfonic acid (MES), biotin, biotinamidocaproic acid 3-sulfo-N-hydroxy-succinimide ester sodium salt (NHS-biotin), N-hydroxysuccinimide (NHS), NHS-long chain (LC), 4-(2-hydroxyethy)-1-piperazine ethanesulfonic acid (HEPES), lithium chloride (LiCl), Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), guanidine hydrochloride (GHCl), mouse anti-biotin antibody (clone BN-34) and fumonisin B1 were obtained from Sigma. Methyl-N-succimimidyl adipate (MSA) and EZ-Link Biotin-LC-Hydrazide were obtained from Pierce (Rockford, IL). Polyclonal S_A-protein antiserum was obtained as described (Boot *et al.*, 1993), monoclonal anti-biotin (clone B4) was purchased from Sigma and anti-fumonisin B1 (clone P2D5) was a kind gift from Dr. C. Maragos (FDA, USA).

Cultivation of bacteria and isolation of surface layer protein. The surface layer protein, S_A-protein, used in this study was isolated from *Lactobacillus acidophilus* bacteria (American Type Culture Collection (ATCC) 4356) as previously described (Smit *et al.*, 2001). Briefly, bacteria were grown anaerobically in MRS medium (Difco) to an optical density (measured at 695 nm) of 0.6 and harvested by centrifugation (20.000 x g, 20 min, 4°C). S_A-protein was extracted from bacteria with 5 M LiCl after washing with 0.9% NaCl and 1 M LiCl. The 5 M LiCl extract contained highly purified S_A-protein (2.5 mg/ml) that could be stored indefinitely in 5 M LiCl at room temperature (RT). LiCl was

removed from the S_A-protein solution using a Pall Gelman (Ann Arbor, MI) Microsep, low protein binding, ultrafiltration unit (molecular weight cut-off (MWCO) 30.000 Da) and replaced by 50 mM Tris-HCl, pH 7.5. This S_A-protein preparation could be stored for a few days at 4°C. Prior to use the solution was diluted in MilliQ to 50 µg/ml.

SPR equipment and data analysis. SPR analyses were performed with a BIAcore 1000 apparatus and standard gold chips (SIA kit, BIAcore AB, Sweden). Experimental data was analyzed using BIAcore software. All buffers were prepared with MilliQ water and filter-sterilized. Solutions were prepared with buffer or filter-sterilized MilliQ. HEPES-buffered saline (HBS) was used as standard buffer.

S_A-protein adsorption to gold-coated sensor chips. Before use, sensor chips were ethanol vapor-cleaned (EVC), fixed to a polystyrene support and then mounted in the BIAcore apparatus according to the instructions of the supplier. S_A-protein adsorption to the chips was determined during a six minute injection of a 50 µg/ml S_A-protein solution. After injection, the chips were regenerated with 0.5 M NaCl, 1 M NaCl and 25 mM NaOH (twice each) and the amount of S_A-protein that remained adsorbed was assessed after each regeneration step. In so-called incubation experiments, S_A-protein was allowed to adsorb to sensor chip surfaces for 3 h at RT prior to mounting of the chip in the BIAcore apparatus. Before regeneration, the amount of S_A-protein adsorbed was determined. Regeneration was performed with 0.5 M NaCl, 1 M NaCl, 25 mM NaOH and 6 M GHCl. Bovine serum albumin (BSA) (50 µg/ml) was used as a negative control in a similar incubation experiment. Flow rate was set at 10 µl/min in all experiments.

Binding of specific antibodies to un-modified S_A-protein-coated chips. S-layer-surface-specific antibodies were extracted from the polyclonal S_A-protein antiserum according to the method described by Martinez *et al.* (2000) (Martinez *et al.*, 2000). Chip surfaces were regenerated twice with 25 mM NaOH and twice with 6 M GHCl before antibody binding. The recognition of S_A-protein immobilized on EVC gold chips by both antibody preparations (10 µg/ml in HBS) was studied during 7 min injections at 10 µl/min. 6 M GHCl was used to regenerate after binding of the antibodies.

Chemical modification of S_A-protein. Carboxyl groups of the S_A-protein were activated by mixing 500 µl S_A-protein solution (100 µg/ml, 10 mM MES, pH 6) with 100 µl 200 mM EDC/50 mM NHS (10 mM MES, pH 6) and incubation for 10 min at RT under continuous stirring. Biotin was coupled to carboxyl-activated S_A-protein by addition of 400 µl EZ-Link Biotin-LC-hydrazide (2 mg/ml, 50 mM borate buffer) and incubation for 30 min at RT. Coupling of biotin to amino groups was carried out by mixing 500 µl (2

mg/ml, 50 mM borate buffer) NHS-LC-biotin with 500 µl of S_A-protein solution (100 µg/ml) and incubation for 30 min under continuous stirring at RT. The mycotoxin fumonisin B1 was coupled to S_A-protein carboxyl groups as follows. 600 µl of an activated S_A-protein solution was mixed with 400 µl fumonisin B1 solution (0.5 mg/ml, 50 mM borate buffer) and incubated for 10 min at RT under continuous stirring. Prepared S_A-protein material was used directly for immobilization on gold chips by incubation as described above. A chip coated with un-modified S_A-protein served as a negative control.

Inhibition assays with chemically modified S_A-protein chips. Inhibition assays were performed with COOH- or NH_2-biotinylated or fumonisin B1-modified S_A-protein chips. Biotinylated, fumonisin-modified and un-modified S_A-protein sensor chips were regenerated with 25 mM NaOH and 6 M GHCl for 1 min each prior to anti-biotin antibody injection. Before injection monoclonal anti-biotin antibodies (10 µg/ml in HBS) or anti-fumonisin B1 antibodies (10 µg/ml in HBS) were mixed with increasing concentrations (1 mg/ml to 0.3 ng/ml) of free biotin or fumonisin B1 (100 µg/ml to 0.1 ng/ml). Chips were exposed to antibody-biotin or antibody-fumonisin B1 mixtures by seven minute injections at 10 µl/min. Surfaces were regenerated with 6 M GHCl. Assays were performed in triplicate.

AFM measurements. To visualize S_A-protein immobilized on chip surfaces AFM was performed with a Nanoscope III (Digital instruments, Santa Barbara, CA) as described by Rinia *et al.* (1999). S_A-protein-coated gold slides and untreated slides (~1 cm x 1 cm) were prepared as described above. Prior to microscopy, slides were washed with MilliQ or 10 mM NaOH followed by MilliQ and fixed to a metal support.

Results

Adsorption of S_A-protein to gold surfaces. Adsorption of S_A-protein to ethanol vapor-cleaned EVC gold surfaces was first analyzed by exposing a BIAcore-mounted and regenerated chip to an S_A-protein solution. Figure 1 shows a sensorgram of a typical injection experiment. The baseline response of the chip prior to injection was subtracted from the relative response to obtain the normalized values given in the figure 1. During the course of the injection (6 min) a rapid initial increase of the response was observed which reached a level of about 5,000 RU after about one minute. In the remaining 5 min only a slight increase to a plateau level of 5,600 RU was observed. Following injection a slow dissociation of material from the surface was observed and after regeneration of the chip surface with 0.5 M NaCl (figure 1, regeneration 1 and 2), 1 M NaCl (regeneration 3

and 4) and 25 mM NaOH (regeneration 5) none of the adsorbed material remained bound.

Adsorption experiments in which S_A-protein was incubated with the EVC gold surface for 3 h prior to mounting in the BIAcore apparatus yielded quite different results compared to injection experiments. Figure 2 shows a sensorgram of a typical incubation experiment. After incubation and mounting in the BIAcore a high response (5,800 RU) was initially measured, which was not affected by treatment with 0.5 or 1 M NaCl (results not shown). Regeneration with NaOH and GHCl resulted in a final response level of 2,200 RU (figure 2). Some variation was observed between flow cells (± 400 RU) and between chips, but final responses after regeneration were never below 1,800 RU's. The adsorbed S_A-protein layer proved very stable over repetitive (up to 30) NaOH/GHCl washing cycles, indicating strong adsorption to the gold surface (results not shown). BSA, used as a control, was also adsorbed but was completely removed by regeneration with NaOH or GHCl (results not shown).

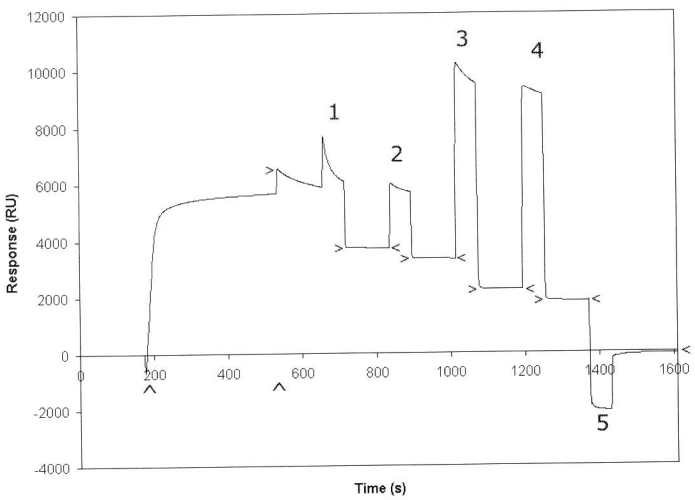


Figure 1. S_A-protein adsorption to an EVC gold surface upon injection. The absolute response is indicated in response units (RU). Start and end of the S_A-protein injection are indicated with '^'. Regenerations are indicated: (1) and (2) 0.5 M NaCl, (3) and (4) 1 M NaCl and (5) 25 mM NaOH. The response level reached after each regeneration as well as start and end levels are indicated with '>' and '<'. Sharp up or down peaks of the curve associated with regeneration steps represent changes in refractive index caused by the change in buffer or solution. The chip baseline response was 11,765 RU.

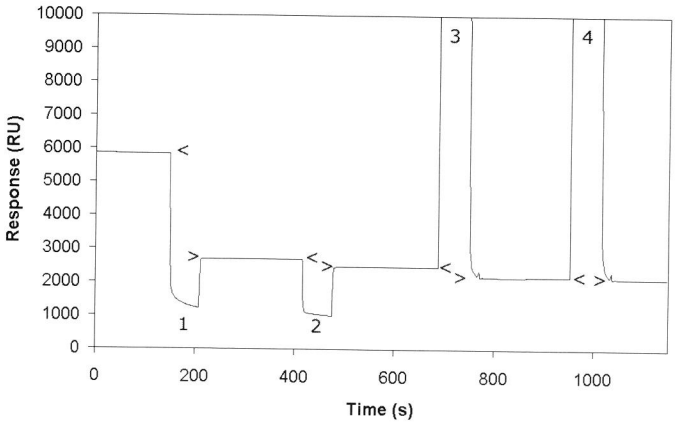


Figure 2. S_A-protein adsorption to an EVC gold surface by incubation. The absolute response is indicated units (RU). Regeneration was done with 25 mM NaOH (1 and 2) and 6 M GHCl (3 and 4). The response level reached after each regeneration as well as start and end levels are indicated between '>' and '<'. The baseline response of the chip before incubation was ~ 9,200.

Characteristics of the immobilized S_A-protein layer. Using anti-S_A-protein polyclonal antibodies and anti-S-layer-surface antibodies we determined whether S_A-protein adsorbed to the gold surface with a specific orientation. While polyclonal antibodies bound well to the chip surface, purified anti-S-layer-surface polyclonal antibodies hardly showed any binding (figure 3). Binding of the non-purified polyclonal antibodies, when injection was stopped after 7 min, was still increasing indicating that there were additional epitopes available for antibody binding. A response level of about 75 RU's was reached after injection which corresponds to one antibody molecule for every one hundred S_A-protein molecules. Repetitive adsorption and regeneration with GHCl of the S_A-protein chip showed good stability of the antibody response (not shown).

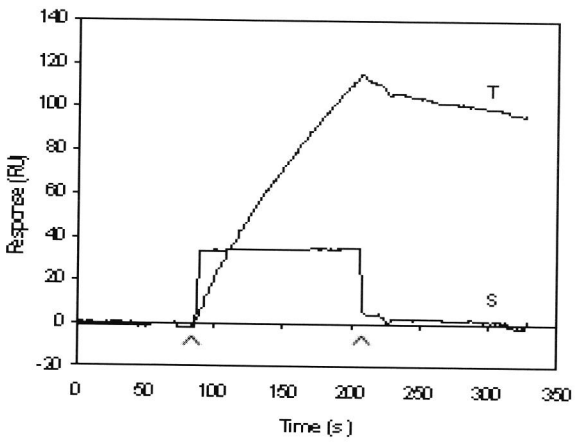


Figure 3. Binding of a polyclonal S_A-protein antiserum (T) and a purified, surface-specific, fraction (S) to an S_A-protein coated and NaOH/GHCl regenerated sensor surface. Normalized responses are indicated in RU. '^' indicate start and end of the antibody injection.

We also used AFM in an attempt to visualize S_A-protein immobilized on the gold chips and to determine whether the protein was bound in a crystalline form. Since the gold surface itself was found to be quite rough, we were not able to visualize S_A-protein crystals. We did however observe a distinct 'smoothing' of the surface of the S_A-protein-coated surface compared to an untreated surface. Regeneration with 25 mM NaOH resulted in an additional smoothing. This result can be explained by assuming that the S_A-protein that is removed forms one or more stacked layers of S_A-protein superimposed upon that bound to gold.

The effect of chemical modification of S_A-protein on immobilization. Having determined the conditions for permanent and stable immobilization of S_A-protein on EVC gold, we analyzed the effect of chemical modification of S_A-protein on immobilization. Carboxyl or amino groups of S_A-protein were modified prior to immobilization but these modifications did not affect the capacity of S_A-protein to be permanently immobilized on EVC gold.

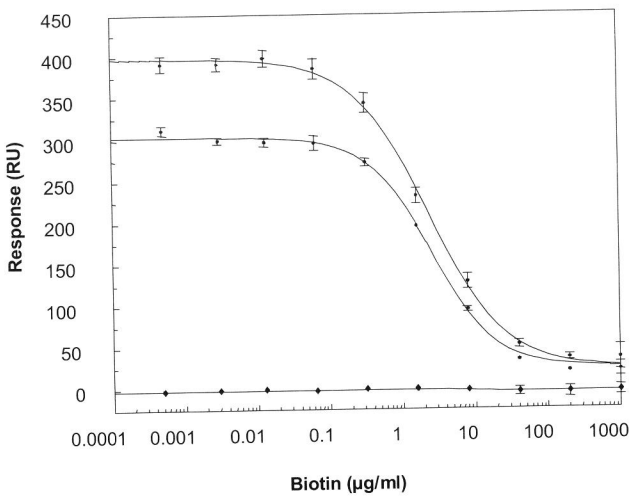


Figure 4. Inhibition assays performed with anti-biotin monoclonal antibody and NH_2-biotinylated (top curve), carboxyl- biotinylated (middle) or unmodified (bottom) S_A-protein sensor coatings. Sensor surfaces were regenerated with NaOH/GHCl before use.

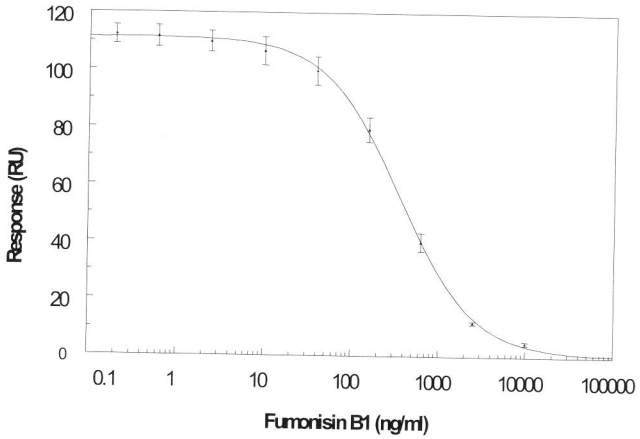


Figure 5. Inhibition assays performed with anti-fumonisin B1 monoclonal antibody and a S_A-protein sensor coating with fumonisin B1-coupled carboxyl groups. Sensor surfaces were regenerated with NaOH/GHCl before use.

Biotinylation of these groups also did not influence this capacity (figure 4). Using carboxyl- or amino-biotinylated-S_A-protein chips we analyzed the accessibility of the biotin on the chip surface using anti-biotin. Figure 4 shows the results of inhibition assays with these chips. Biotin was accessible for antibody binding on both types of modified S_A-protein chips. The response obtained for the amino-biotinylated chip was about 400 RU compared to 300 RU for the carboxyl-biotinylated chip. No response was obtained for an un-modified S_A-protein chip. At a concentration of about 2 µg biotin per ml the response of the amino-modified S_A-protein coating was reduced by 50 % (IC_{50} value). A biotin concentration of 0.3 µg/ml however, already resulted in a significant reduction of the SPR response. Figure 5 shows the result of an inhibition assay performed with a fumonisin B1-modified S_A-protein coating. The maximum response obtained was approximately 110 RU, much lower than observed for biotinylated S_A-protein, but the sensitivity of the measurements was considerably higher. The IC_{50} value for fumonisin B1 was around 300 ng/ml. and already at a fumonisin B1 concentration as low as 40 ng/ml a significant reduction of the SPR response was measured.

Discussion

S-layers, as the outermost layer of the cell envelope of many bacteria, are important for the interaction of bacteria with their environment. Especially S-layers of pathogenic bacteria seem to mediate cell adhesion as well as immune evasion (Dworkin & Blaser, 1997; Garduno *et al.*, 2000; Kotiranta *et al.*, 2000). The cell surface properties of

probiotic lactobacilli are the focus of many recent studies since it is thought that these properties play an important role in bacterial adhesion and persistence at mucosal surfaces, prerequisites for the effectiveness of a probiotic (Dunne *et al.*, 1999; Kailasapathy & Chin, 2000). Lactobacilli from the *L. acidophilus* A homology group from which a number of well-known probiotic bacteria originate, possess an S-layer. Although a role for these abundant cell surface proteins in adhesion seems obvious, only a few studies have addressed this issue (Greene & Klaenhammer, 1994; Schneitz *et al.*, 1993; Toba *et al.*, 1995). *In vitro* studies of the interaction of S-proteins with solid substrates could help to yield insight in the interactions of S-layered bacteria with mucosal surfaces *in vivo*.

In this chapter we describe our studies on the interaction of the *L. acidophilus* S-layer protein as determined by SPR. SPR allows to study the interaction between molecules and with different substrates in real-time. The substrate chosen for the present study was an ethanol vapor-cleaned (EVC) gold surface, which has moderately hydrophobic surface characteristics (water contact angle > 45°).

We analyzed the interaction of S_A-protein with EVC gold in two ways. First the process of adsorption over a short period of time was determined. When an S_A-protein solution was passed over the gold surface (injection experiments) for a few minutes a steep increase of the adsorption curve was measured resulting in an almost saturated gold surface after about one minute. A final response of 5,500 RU above background was reached corresponding to 5-6 ng bound protein per mm^2 (1000 RU ~ 1ng/mm^2, BIAcore manual; Silin & Plant, 1997). Since an amount to 2 - 3 ng/mm^2 is needed to completely cover the available gold surface with a densely packed/crystalline monolayer of S_A-protein, a higher response may be the result of stacking of S_A-protein molecules. Regeneration with NaCl and NaOH, respectively, resulted in complete removal of all adsorbed material showing that S_A-protein binds only weakly to the substrate over the course of a few minutes.

When S_A-protein was allowed to interact with the gold surface for a longer period of time the results were altogether different. Exposing the gold surface to S_A-protein for 3 hours resulted in a similarly high response as observed in the injection experiments, but now regeneration with NaCl, and even NaOH or GHCl was only partly effective in removing adsorbed S_A-protein. About 2.2 ng/mm^2 remained bound after regeneration with all three agents, an amount that corresponds closely to the theoretical value for an ordered monolayer.

In previous *in vitro* studies we found that S_A-protein binds to anionic substrates via it's cell wall binding domain, while it interacts with a hydrophobic surface via it's outer crystallization domain. The efficient adsorption of S_A-protein to the moderately hydrophobic gold surface suggests that interaction takes place via the hydrophobic crystallization domain. Binding of a polyclonal anti-S_A-protein antiserum and a purified, S-layer surface-specific antiserum to the adsorbed S_A-protein layer, provided strong indications that S_A-protein adsorbs with a specific orientation. Unpurified antibodies efficiently recognized S_A-protein epitopes, while purified, surface-specific antibodies did not bind to gold-adsorbed S_A-protein. These results reinforce the conclusion that S_A-protein is oriented to the gold substrate with its outer face. Since no reactivity of bound S_A-protein with S-layer surface-specific antiserum could be detected, all S_A-protein molecules are oriented in the same way in the adsorbed layer, strongly suggesting that the layer possesses a crystalline structure. Whether the layer has crystalline properties cannot be concluded from these experiments.

Orientation of S_A-protein with its outer face to the gold surface means that the C-terminal cell wall binding domain of the S-layer protein is exposed to the environment. To determine whether carboxyl or amino groups are involved in the stable immobilization to a gold substrate and in monomer-monomer interaction, we chemically modified these groups in S_A-protein prior to the immobilization phase. Coupling of biotin to amino and carboxyl groups and fumonisin B1 to carboxyl groups did not interfere with stable adsorption. This finding suggests that the modified groups are not directly involved in either the S_A-protein-gold or monomer-monomer interaction. The high responses that were observed with biotin- and fumonisin B1-specific antibodies indicate that many carboxyl and amino groups are exposed on the S_A-protein covered surface. This corroborates our previous conclusion that the C-terminal region of the S_A-protein contains many amino and carboxyl groups.

Finally, our data suggests that S_A-protein is ideally suited for application as a generic biosensor coating. S_A-protein can be easily isolated and stored indefinitely, when dissolved in for example 6 M GHCl or 5 M LiCl. Moreover, S_A-protein coating is carried out in a single step obviating the need for complex and time-consuming procedures. The coated layer proved highly stable over repeated regeneration with chaotropic agents. Furthermore, modification of S_A-protein is easily performed and has no effect on the coating of gold. Our inhibition assays also showed the versatility and applicability of S_A-protein coating. The sensitivity of fumonisin B1 detection (~40 ng/ml) and IC_{50} in this assay compares favorably with ELISA-based assays. Elissalde *et al.* reported ELISA IC_{50} values of 300 to 670 ng/ml for fumonisin B1 depending on the monoclonal antibody used

(Elissalde *et al.*, 1995). Commercial ELISA-based fumonisin B1 assays (for example from R-Biopharm, Darmstadt, Germany) show a sensitivity the same order of magnitude (maximum sensitivity ~10 ng/ml).

Chapter 7

Summary and concluding remarks

1. Molecular biological research on S-layers

Although S-layers were discovered over forty years ago (Houwink, 1953), up to ten years ago research centered mainly around the biochemistry of S-proteins and electron microscopical studies on S-layer ultrastructure. One of the main factors hampering molecular biological research in areas such as S-protein gene expression, regulation, secretion and structure has been the lack of S-protein sequence information. Since the end of the eighties however, recombinant DNA techniques have been used to clone many S-protein genes. Consequently, more profound studies on the molecular biology of S-layer proteins were possible. The one area that still lags behind is structural research on S-layer proteins, resulting in a poor understanding of the structural-function relationships of S-layer proteins. This thesis aims to provide a deeper understanding of the structure-function relationship of *Lactobacillus* S-layer proteins. The *L. acidophilus* S_A-protein is taken as a representative member of a group of S-proteins from lactobacilli related to *L. acidophilus*.

In a recent review it was stated that application of NMR, cryo-electron microscopy and X-ray crystallographic techniques in combination with recombinant DNA techniques should allow researchers to determine the three-dimensional structure of S-layers and S-layer proteins (Engelhardt & Peters, 1998). Our work on the domain organization of the S_A-protein has generally proceeded along these lines.

2. S_A-protein domains

Alignment of sequences from several related proteins is always very informative since well-conserved amino acids usually represent either structurally or functionally (or both) important residues. Also the analysis of the hydrophilicity / hydrophobicity profile of a primary sequence, the search for regions or positions with structural flexibility (presence of for example glycine residues) or constraints (proline residues) and alignment-based secondary structural predictions yield relevant structural information. These analyses cannot be used for most S-layer proteins since for few S-proteins enough related sequences are available for comparison. The S-proteins of *L. acidophilus* and related species however have a unique position among S-proteins because extensive amino acid analyses can be performed (Chapter 2 and 5). The most striking feature of these related S-proteins is that all possess a highly identical C-terminal region (~120 amino acids). In contrast, the N-terminal regions have a much lower level of identity, but do possess several clearly conserved motifs. Whether these motifs are important for S-protein

structural integrity or protein-protein interaction of the individual monomers cannot be deduced from primary sequence data alone.

The division of the S_A-protein into two distinct regions was confirmed by proteolysis experiments. The C-terminal region (SAC: **S_A**-protein **C**-terminal region) proved very sensitive to proteolysis while the N-terminal region (SAN: **S_A**-protein **N**-terminal region) was much more resistant. Based on the sequence analysis and proteolysis data we cloned the N- and C-terminal regions (amino acids 1-290 and 291-413, respectively) and characterized their properties separately. The purified N-terminal region showed crystallizing properties similar to S_A-protein, whereas the C-terminal domain did not such behavior and was found to be soluble under native conditions.

2.1. The crystallization domain. We successfully obtained and compared the two-dimensional reconstructions of both wild type S_A-protein and the SAN domain crystals. It is the first study comparing the two-dimensional structure of a wild type S-protein with that of a truncated form. Since the structural parameters of the crystals formed by the two peptides were identical and since the SAN structure can form independently from the SAC region we named SAN the S_A-protein crystallization domain.

Although these results should be further extended with three-dimensional reconstructions of the crystal structures of both peptides, we suggest this strategy could already be applied to other S-proteins. Several groups did report the existence of distinct crystallization domains in other S-layer proteins such as the SbsA, SbsB, SbsC, Sap and Eag S-proteins. However, the boundaries of these domains were always deduced after removal of conserved cell wall binding regions such as SLH domains, but the crystallization domains were never analyzed separately (Sára *et al.*, 1998; Jarosch *et al.*, 2001; Mesnage *et al.*, 1997).

The next important question was whether SAN could be subdivided into smaller, but still structurally stable and functional, units. Clues that smaller functional subunits (or 'subdomains') did exist came from insertion mutagenesis of the SAN domain. Insertion mutagenesis was carried out with two objectives. The first was to obtain further structural information about the SAN domain and the second was to use mutant S_A-protein to present epitopes at the cell surface *in vivo* (the latter will be discussed below).

As was shown by a group working on the *Caulobacter crescentus* S-layer protein, insertion mutagenesis can be a powerful tool for obtaining functional information about specific S-protein regions (Bingle *et al.*, 1997). Unfortunately, this group only determined

the effect of insertions on the secretion and cell wall attachment of the S-protein and did not extend their work to crystallization properties of the protein. We applied this strategy specifically to gain more insight in the functional characteristics of specific SAN domain regions (chapter 4). In general, we found that insertions in or close to one of the detected SAN motifs resulted in crystallization-compromised mutants, while insertions in the intermittent regions were usually accepted. Especially, the central region of the SAN domain accepted insertions at various positions and of various sizes. This suggested the presence of a flexible loop region and subsequent analysis of *in vivo* expressed insertion mutants showed that the inserted amino acids are exposed on the S-layer surface. The surface exposed flexible region is located more or less in the center of the SAN domain and we thought it could be a loop separating the upstream region from the downstream region. This question we aimed to answer with the deletion analysis described below.

Using available proteolysis data and the results obtained from the insertion mutagenesis of the SAN domain we designed, purified and analyzed a set of SAN N- and C-terminally deleted peptides (chapter 5). Neither of the peptides formed large assembly products so that EM could not be performed and so functional analysis consisted of SDS-PAGE followed by western analysis. We used this simple method as a functional assay because we previously observed that dimeric and trimeric assembly products of S_A-protein and SAN could be detected in this way. The observation of these assembly products even after boiling in 1 % SDS indicates that the interactions between monomers are strong and suggests that hydrophobic rather that electrostatic interaction plays an important role in the association process (SDS interferes with protein-protein interactions and protein structure primarily by disrupting electrostatic interactions (Creighton, 1993)). From the fact that we could detect dimeric forms of four of the peptides we concluded that they are functional units. Furthermore, because they are functional units this must mean that they also fold into a native structure. Two of the three N-terminally truncated peptides proved to be unstable when produced in *E. coli* (sensitive to *E. coli* proteases) indicating that these were probably not folded properly. Finally, we conclude that SAN consists of two functional subunits or subdomains, residues 1 to ~113 and ~114 to 290 respectively, which are connected by a surface exposed loop roughly consisting of residues 114 to 140. Further refinement of the deletion analysis is needed to better define these boundaries, but these results nevertheless represent the first evidence of dimerization or oligomerization of defined S-protein domains or subdomains.

Further support for the importance of hydrophobic interactions for S-protein assembly came from the work described in chapter 6. There we describe the capacity of purified S_A-protein to irreversibly bind to a moderately hydrophobic gold surface as determined

by SPR. We obtained results strongly suggesting that the S_A-protein assembles on the gold surface into a highly ordered, most likely crystalline layer, after prolonged (90 min) incubation with the gold surface. Chemical modification of either carboxyl or amino groups prior to incubation did not affect this assembly. From these observations we conclude that hydrophobic forces play a significant role in both the S_A-protein-gold interactions and S_A-protein-S_A-protein interactions, but that electrostatic forces and hydrogen bridging cannot be excluded. More extensive chemical modification studies are needed and could be combined with time-course experiments, which would give insight into the process of S-protein assembly and the role of specific S-protein functional groups.

2.2. The cell wall binding domain. As mentioned above, the S_A-protein contained a highly conserved C-terminal region (SAC) that was found to be highly protease-sensitive (Chapter 2 and 3). Hydrophilicity analysis of the S_A-protein showed that the C-terminal region is very hydrophilic and can be clearly distinguished from the N-terminal region due to the presence of a large number of basic amino acid residues (mostly lysine and some arginine). Cloning of this region was successful as a fusion with GFP and we used the fusion protein in cell wall binding assays. The surplus of positive charges in SAC was an indication that it could have a role in binding the bacterial cell wall which usually have an anionic character. Using either whole cells from which the S-layer had been stripped or purified cell wall fragments, we showed that SAC indeed interacts very well with the *L. acidophilus* cell wall and is the S_A-protein cell wall binding domain. The interaction of SAC with the cell wall occurs primarily via electrostatic interaction since bound GFP-SAC was easily removed with NaCl. In addition, SAC also bound to the *L. helveticus* and *L. crispatus* cell wall indicating that the S-proteins of these bacteria also attach to the cell wall via their C-terminal domain.

Interestingly, further sequence analysis of SAC revealed that it actually consists of a tandemly repeated sequence (SAC1 and SAC2) and alignment revealed a core region with several conserved basic and aromatic residues (chapter 2 and 3). This conserved region possessed a striking similarity to repetitive carbohydrate-binding domains of various other bacterial proteins, such as clostridial toxins and extracellular glycosyltransferases. These types of amino acids are also conserved in carbohydrate-binding proteins, or lectins, of both prokaryotic and eukaryotic origin, where they serve to correctly coordinate the carbohydrate in the binding pocket via hydrophobic interaction with the planar side of the sugar molecule and by hydrogen bonding with the hydroxyl side chains. Analysis of the cell wall binding properties of GFP-SAC1 and GFP-SAC2 fusions was determined and to our surprise showed that SAC1 bound as well as SAC,

SAC2 did not bind at all. This was surprising given that SAC1 and SAC2 are possess the same conserved amino acids. Thus the difference in cell wall binding of the repeats must lie in differences in primary structure probably resulting in a different structure of SAC2. We have observed that SAC1 is more easily extracted from cell wall fragments with NaCl than SAC indicating that SAC2 at least has a role in strengthening the electrostatic interaction with the cell wall.

Taken together these experimental data and sequence characteristics suggest that the SAC domain interacts with an anionic cell wall component that may also contain specific carbohydrate residues. To test this we extracted cell wall associated polymers from cell wall fragments with hydrofluoric acid, a method that yields essentially pure peptidoglycan. Binding of SAC was completely abolished by this treatment, which coincided with the loss of all cell wall-associated phosphorus and in a lowering of the galactose content. Specific extraction of cell wall-associated lipoteichoic acid or LTA (a glycerolphosphate or ribitolphosphate polymer non-covalently associated with the cell wall) by hot phenol treatment lowered the phosphorus content, but did not affect the binding of SAC. These results indicate the role of a phosphorus-containing cell wall polymer in the binding of SAC. Since specific extraction of LTA from cell wall fragments did not affect binding of SAC covalently-associated cell wall teichoic acid (TA) is the most likely cell wall receptor. Free glucose galactose or N-acetyl glucosamine did not affect the SAC cell wall interaction but this does not mean that carbohydrates are not involved. Carbohydrate-binding can be very specific for a certain conformation of the sugar ligand, which is not found in the free form. Thus the involvement of carbohydrates in SAC cell wall binding cannot be excluded (especially with the above-mentioned sequence similarities in mind). One way to further analyze the interaction between SAC and anionic cell wall polymers could be to obtain purified polymers. This may be done by digesting the peptidoglycan with cell wall lytic enzymes followed by anion-exchange chromatography purification of the polymers. An important additional advantage of this method is that acid-induced modifications of the polymer or of the peptidoglycan are avoided.

Finally, our data combined allows us to construct a structural model of the S_A-protein and how individual monomers interact with neighboring monomers to form the macromolecular S-layer lattice (figure 1).

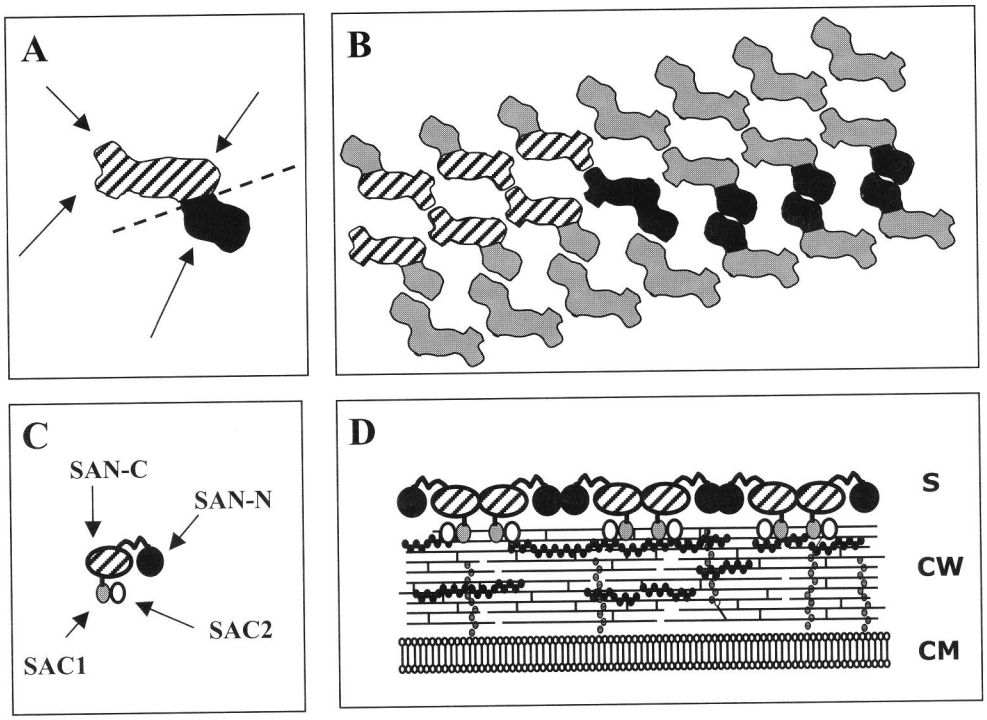


Figure 1. Schematic overview of the domain and subdomain organization of the *L. acidophilus* S_A-protein. (A) shows a SAN monomer with the putative N- (SAN-N, black) and C-terminal (SAN-C, shaded) subdomains and the possible contact points with neighboring subunits (arrows). The putative subdomain boundary is indicated with a dashed line. (B) shows the organization of a small section of a SAN two-dimensional crystal. A SAN monomer is indicated in gray and SAN-N and SAN-C as in (A). (C) shows the domain/subdomain organization of a SA-protein monomer in cross-section. (D) shows the putative structural organization of the *L. acidophilus* cell envelop in cross-section and how the S_A-protein monomers are organized at the cell surface in an S-layer lattice. The different domains are indicated as in (C). The cytoplasmic membrane (CM), peptidoglycan (PG, horizontal lines with vertical connections), teichoic acid (black chains) and lipoteichoic acid (gray chains) are indicated.

3. S-layers, cell surface enzymes and enzymatic protein assemblies

As mentioned in chapter 5 the *L. helveticus* cell-surface proteinase PrtY, similar in size to *Lactobacillus* S-layer proteins, possessed the same conserved sequence motifs detected in S_A and related S-proteins. This observation supports our previous conclusion that the motifs are important for protein structural integrity and not protein-protein interaction, since the PrtY protein did not show crystallizing properties. It further suggests that the S_A-protein may be evolutionary related to this extracellular enzyme. The conserved active site residues of the proteinase are not found at corresponding position in S_A-protein for which, not surprisingly, we were not able to detect proteolytic activity (as determined by milk overlay assays). Nothing is known about the abundance of the PrtY

protein at the *L. helveticus* cell surface but one could speculate that the *L. acidophilus* S_A-protein is a highly abundant cell-surface proteinase that at some stage during evolution lost its enzymatic property. One could as easily suppose however, that the PrtY proteinase is an S-protein that at some stage during evolution acquired a catalytic property.

The existence of an S-protein with an enzymatic function was recently reported for an S-layer proteins from *Clostridium difficile.* This bacterium has a double S-layer structure and amidase activity was detected for the outermost S-layer component (Cerquetti *et al.*, 2000). This is the first report describing an S-layer protein with an enzymatic function and one could speculate that other bacteria with a double S-layer structure (for example *Brevibacillus brevis*) could represent a similar case, although enzymatic studies have not been published. One could wonder whether the outer S-layer is a highly abundant cell-surface enzyme using the inner S-layer as scaffolding to present itself to the environment. That S-layers can function as attachment sites for enzymes was demonstrated for the S-layer of *Bacillus stearothermophilus* DSM 2358, which bound a high-molecular weight amylase (Egelseer *et al.*, 1995). Other S-layer-carrying bacteria, including *L. acidophilus,* possess additional cell surface proteins besides the S-layer protein and these may use the S-layer as an attachment site. Carrying out interaction studies for several S-layer systems may show that S-layers have an important general function in attachment of other cell surface proteins. As such S-layer may mediate adhesion by binding adhesins or in acquiring nutrients by binding extracellular enzymes or enzymatic complexes.

With respect to S-proteins and extracellular enzymatic complexes cellulosomes must be mentioned. Cellulosomes are multi-subunit enzymatic complexes digesting exogenous cellulose and are found in clostridia and related bacteria. Many cellulosomal subunits have been identified and shown to possess the same cell wall anchoring domains as many S-proteins, i.e. S-layer Homology (SLH) domains. This establishes yet another connection between S-proteins and other cell surface-associated molecules.

4. Application potential of *Lactobacillus* S-layer proteins

Part of the work described in chapters 4 and 6 was aimed at exploring the possibilities for biotechnological application of *Lactobacillus* S-proteins. The applications that we focused on were (1) the use of live lactobacilli for presentation of epitopes or antigens at the cell

surface using the S-layer and (2) the use of purified S_A-protein as a surface coating in biosensor technology.

As GRAS organisms lactobacilli have been the subject of applied research for many years. One of the main interests today is the application of lactobacilli as probiotics for man and animal. Lactobacilli possess health-promoting properties that beneficially affect the microbial balance in the gastrointestinal tract (GIT), can survive passage through the GIT and many strains are able to colonize the gastrointestinal environment. These traits are of special interest to researchers that study the potential of lactobacilli as mucosal vaccine delivery vehicles.

Presentation of antigens from bacterial pathogens to the mucosal immune system can elicit both mucosal and systemic immune responses. This is the way in which the body defends itself against infections via mucosal surfaces. Traditionally, immunizations are carried out with either attenuated non-virulent, or killed pathogens, but the risk of the subject getting ill, due to pathogens reverted to a virulant form or to incomplete killing, remains. The use of harmless or even beneficial bacteria for the presentation of antigens derived from these pathogens may eliminate the risks of using attenuated or killed pathogens. Such vaccines also eliminate the need for invasive administration methods, which is another advantage. Furthermore, lactobacilli have also been shown to possess immuno-modulating properties that may stimulate the response to an antigen (Matsuzaki & Chin, 2000).

Antigens may be produced in lactobacilli in cytoplasmic, cell-surface anchored or secreted form and application of the bacteria to mucosal sites has shown that both mucosal and systemic immune response can be elicited (Pouwels *et al.*, 2001; Shaw *et al.*, 2000; Grangette *et al.*, 2001). The use of the S-layer of a *Lactobacillus* provides an alternative for the already available cell-surface anchored presentation of antigens. The advantage of using an S-layer is the high number of protein subunits that is present at the cell surface ($\sim 10^5$ molecules per cell) which exceeds the number of molecules that can be anchored using other methods such as covalent anchoring with the PrtP cell wall anchor (max 10^4 molecules per cell). Thus the use of "S-layer presentation" could allow the administration of a higher antigen dose which in turn may have a strong effect on the potency of the immune response. Our S_A-protein insertion mutagenesis has yielded mutants that were not affected by the insertion of 20 amino acid epitopes and thus will most likely form an S-layer when produced *in vivo*. Production of several mutants in *L. acidophilus* showed that an epitope insertion in the central loop region of the SAN domain

was exposed at the surface of the S-layer, while an insertion at the extreme N-terminus was buried in the S-layer.

Finally, the results described in chapter 6 show that the S_A-protein can be irreversibly and stably immobilized on a commonly used gold biosensor surface. The intrinsic property of S_A-protein to form assembly products, consisting of multiple subunits that interact with both the gold surface and each other, is the basis of this phenomenon. Furthermore, standard chemical modifications can be used without affecting immobilization. We demonstrate that a sensitive biosensor for the detection of fungal toxins (mycotoxins) can be constructed quickly and easily without the need for complex multi-step sensor constructions.

5. Conclusions and future prospects

As mentioned earlier, the integration of traditional electron microscopical (EM) techniques with the new possibilities offered by recombinant DNA technology should provide new insights into the structure-function relationship of bacterial S-layer proteins. In our work we have indeed shown the potential of this integration. We have performed studies that may provide new impulses to the field of S-layer structural research. The following results have using novel combinations of existing methodologies:

(1) comparative EM studies of wild type S_A-protein with a truncated form, resulting in the definition of a distinct crystallization domain and distinct cell wall binding domain,
(2) insertion mutagenesis of the crystallization domain providing important evidence for the existence of subdomains, while also yielding several soluble mutant S_A-proteins,
(3) N- and C-terminal deletion mutagenesis confirming the existence of dimerizing subdomains, the building blocks of the crystallization domain.

These three methodologies have provided insights to a level of structural and functional detail not reached previously for S-proteins, and may help in devising ways to solve the atomic structure of an S-protein or S-protein domain by NMR or X-ray crystallography.

Since the actual realization of these methodologies may be some way off yet, more advanced or even new EM techniques may make the 'dissection' of S-proteins into smaller and smaller sections unnecessary, thus avoiding the risk that structures of such section do not represent the structure in the native S-layer. Such advances could lie in the development of more sensitive technologies and powerful image processing software,

but also in the discovery of methods for obtaining proper three-dimensional crystals of S-layer proteins.

Already today, mutagenesis strategies such as the construction of N- and C-terminal fusions could be used for structural studies in combination with existing EM techniques. For example, the N-and C-termini could be localized by constructing fusions and rendering them visible in two- and three-dimensional reconstructions. Small N-terminal extensions such as purification tags are probably too small to be visible in a reconstruction (an N-terminal six histidine-tag was not visible in the SAN two-dimensional structure). Other larger fusion partners could probably be visualized in this way as long as they do not interfere with crystal formation.

References

Adlerberth, I., Ahrne, S., Johansson, M. L., Molin, G., Hanson, L. A. & Wold, A. E. (1996). A mannose-specific adherence mechanism in *Lactobacillus plantarum* conferring binding to the human colonic cell line HT-29. *Appl. Environ. Microbiol.* **62**, 2244-2251

Alander, M., Satokari, R., Korpela, R., Saxelin, M., Vilpponen-Salmela, T., Mattila-Sandholm, T. & von Wright, A. (1999). Persistence of colonization of human colonic mucosa by a probiotic strain, *Lactobacillus rhamnosus* GG, after oral consumption. *Appl. Environ. Microbiol.* **65**, 351-354.

Awram, P. & Smit, J. (1998). The *Caulobacter crescentus* paracrystalline S-layer protein is secreted by an ABC transporter (type I) secretion apparatus. *J. Bacteriol.* **180**, 3062-3069.

Axelsson, L. (1998). Lactic acid bacteria: Classification and Physiology. In *Lactic acid bacteria. Microbiology and functional aspects* (Wright, S. S. a. A. v., ed.), pp. 1-72. Marcel Dekker Inc., New York.

Barondes, S. H., Cooper, D. N., Gitt, M. A. & Leffler, H. (1994). Galectins. Structure and function of a large family of animal lectins. *J. Biol. Chem.* **269**, 20807-20810.

Baumeister, W., Engelhardt, H. (1987). Three-dimensional structure of bacterial surface layers. In: *Electron Microscopy of proteins*. Harris, J.R., Horne, R.W. (eds.). Academic Press, London, Vol 6, pp. 110-154.

Baumeister, W. & Lembcke, G. (1992). Structural features of Archaebacterial cell envelopes. *J. Bioenerg. and Biomem.* **24**(6), 567-575.

Baumeister, W., Wildhaber, I. & Phipps, B. M. (1989). Principles of organization in eubacterial and archaebacterial surface proteins. *Can. J. Microbiol.* **35**, 215-227.

Beachey, E. H. (1981). Bacterial adherence: Adhesin-receptor interactions mediating the attachment of bacteria to mucosal surfaces. *J. Infect. Dis.* **143**, 325-345.

Beveridge, T. J. (1994). Bacterial S-layers. *Curr. Opin. Struct. Biol.* **4**, 204-212.

Beveridge, T. J. & Graham, L. L. (1991). Surface layers of bacteria. *Microbiol. Rev.* **55**, 684-705.

Bingle, W. H., Nomellini, J. F. & Smit, J. (1997). Linker mutagenesis of the *Caulobacter crescentus* S-layer protein: Toward a definition of an N-terminal anchoring region and a C-terminal secretion signal and the potential for heterologous protein secretion. *J. Bacteriol.* **179**, 601-611.

Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.

Blaser, M. J., Smith, P. F., Repine, J. E. & Joiner, K. A. (1988). Pathogenesis of *Campylobacter fetus* infections. *J. Clin. Invest.* **81**, 1434-1444.

Boonaert, C. J. & Rouxhet, P. G. (2000). Surface of lactic acid bacteria: relationships between chemical composition and physicochemical properties. *Appl. Environ. Microbiol.* **66**, 2548-2554.

Boot, H. J. (1996). The surface layer protein genes of *Lactobacillus acidophilus.* In: *PhD.-Thesis.* University of Amsterdam, The Netherlands. Chapter 8, pp. 115-121.

Boot, H. J., Kolen, C. P. A. M., Pot, B., Kersters, K. & Pouwels, P. H. (1996). The presence of two S-layer protein encoding genes is conserved among species related to *Lactobacillus acidophilus*. *Microbiol.* **142**, 2375-2384.

Boot, H. J., Kolen, C. P. A. M. & Pouwels, P. H. (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.

Boot, H. J., Kolen, C. P. A. M., van Noort, J. M. & Pouwels, P. H. (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. & Pouwels, P. H. (1996). Expression, secretion and antigenic variation of bacterial S-layer proteins. *Mol Microbiol*. **28**, 1209-1213

Brechtel, E. & Bahl, H. (1999). In *Thermoanaerobacterium thermosulfurigenes* EM1 S-layer homology domains do not attach to peptidoglycan. *J. Bacteriol.* **181**, 5017-5023.

Brechtel, E., Matuschek, M., Hellberg, A., Egelseer, E. M., Schmid, R. & Bahl, H. (1999). Cell wall of *Thermoanaerobacterium thermosulfurigenes* EM1: isolation of its components and attachment of the xylanase XynA. *Arch Microbiol* **171**, 159-165.

Breitwieser, A., Küpcü, S., Howorka, S., Weigert, S., Langer, C., Hoofmann-Sommergruber, K., Sleytr, U. B. & Sára, M. (1996). 2-D protein crystals as an immobilization matrix for producing reaction zones in dipstick-style immunoassays. *BioTechniques* **21**, 918-925.

Brown, S. (1997). Metal-recognition by repeating polypeptides. *Nat. Biotechnol.* **15**, 269-272.

Buist, G. (1997). The *N*-acetylmuramidase of *Lactococcus lactis* binds the cell wall by means of repeated motifs located in the C-terminal domain. In: *PhD Thesis*. University of Groningen, The Netherlands. Chapter 3, pp. 55-76.

Calabi, E., Ward, S., Wren, B., Paxton, T., Panico, M., Morris, H., Dell, A., Dougan, G. & Fairweather, N. (2001). Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol. Microbiol.* **40**, 1187-1199.

Callegari, M. L., Riboli, B., Sanders, J. W., Cocconcelli, P. S., Kok, J., Venema, G. & Morelli, L. (1998). The S-layer gene of *Lactobacillus helveticus* CNRZ 892: cloning, sequence and heterologous expression. *Microbiol.* 144, 719-726.

Cerning, J. (1990). Exocellular polysaccharides produced by lactic acid bacteria. *FEMS Microbiol. Rev.* **87**, 113-130.

Cerquetti, M., Molinari, A., Sebastianelli, A., Diociaiuti, M., Petruzzelli, R., Capo, C. & Mastrantonio, P. (2000). Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microb. Pathog.* **28**, 363-372.

Chami, M., Bayan, N., Peyret, J. L., Gulikkrzywicki, T., Leblon, G. & Shechter, E. (1997). The S-layer protein of *Corynebacterium glutamicum* is anchored to the cell wall by its C-terminal hydrophobic domain. *Mol. Microbiol.* **23**, 483-492.

Chassy, B. M. & Flickinger, J. L. (1987). Transformation of *Lactobacillus casei* by electroporation. *FEMS Microbiol. Let.* **44**, 173-177.

Chaves-Olarte, E., Weidmann, M., Von Eichel-Streiber, C. & Thelestam, M. (1997). Toxins A and B from *Clostridium difficile* differ with respect to enzymatic potencies, cellular substrate specificities and surface binding to cultured cells. *J. Clin. Invest.* **100**, 1734-1741.

Chu, S., Cavaignac, S., Feutrier, J., Phipps, B. M., Kostrzynska, M., Kay, W. W. & Trust, T. J. (1991). Structure of the tetragonal surface virulence array protein and gene of *Aeromonas salmonicida*. *J. Biol. Chem.* **266**, 15258-15265.

Collins, M. D., Rodrigues, U., Ash, C., Aguirre, M., Farrow, J. A. E., Martinezmurcia, A., Phillipis, B. A., Williams, A. M. & Wallbanks, S. (1991). Phylogenetic analysis of the genus *Lactobacillus* and related lactic acid bacteria as determined by reverse transcriptase sequencing of 16S RNA. *FEMS Microbiol. Lett.* **77**, 5-12.

Creighton, T. E. (1993). Proteins. Structures and Molecular Properties, Second edition. W.H. Freeman and Company, New York.

Cunningham, M. W. (2000). Pathogenesis of group A streptococcal infections. *Clin. Microbiol. Rev.* **13**, 470-511.

de Vos, W. M. (1999). Gene expression systems for lactic acid bacteria. *Curr. Opin. Microbiol.* **3**, 289-295.

Deblaere, R. (1995). Patent application EP9500147.

Delcour, J., Ferain, T., Deghorain, M., Palumbo, E. & Hols, P. (1999). The biosynthesis and functionality of the cell wall of lactic acid bacteria. *Antonie van Leeuwenhoek* **76**, 159-184.

Demel, R. A. (1994). Monomolecular layers in the study of biomembranes. *Subcell. Biochem.* **23**, 83-120.

Dooley, J. S. G., Engelhardt, H., Baumeister, W., Kay, W.W., Trust, T.J. (1989). Three-dimensional structure of an open form of the surface layer from fish pathogen *Aeromonas salmonicida*. *J. Bacteriol.* **171**, 190-197.

Dooley, J. S. G., McCubbin, W. D., Kay, C. M. & Trust, T. J. (1988). Isolation and biochemical characterization of the S-layer protein from a pathogenic *Aeromonas hydrophila* strain. *J. Bacteriol.* **170**, 2631-2638.

Dunne, C., Murphy, L., Flynn, S., O'Mahony, L., O'Halloran, S., Feeney, M., Morrissey, D., Thornton, G., Fitzgerald, G., Daly, C., Kiely, B., Quigley, E. M. M., O'Sullivan, G. C., Shanahan, F. & Collins, J. K. (1999). Probiotics: from myth to reality. Demonstration of functionality in animal models of disease and in human clinical trials. *Antonie van Leeuwenhoek* **76**, 279-292.

Dworkin, J. & Blaser, J. (1996). Generation of *Campylobacter fetus* S-layer protein diversity utilizes a single promoter on an invertible DNA segment. *Mol. Micriobiol.* **19**, 1241-1253.

Dworkin, J. & Blaser, M. J. (1997). Nested DNA inversion as a paradigm of programmed gene rearrangement. *Proc. Nat. Acad. Sci. USA* **94**, 985-990.

Egelseer, E., Schocher, I., Sára, M. & Sleytr, U. B. (1995). The S-layer from *Bacillus stearothermophilus* DSM 2358 functions as an adhesion site for a high-molecular-weight amylase. *J. Bacteriol.* **177**, 1444-1451.

Egelseer, E. M., Danhorn, T., Pleschberger, M., Hotzy, C., Sleytr, U. B. & Sára, M. (2001). Characterization of an S-layer glycoprotein produced in the course of S-layer variation of *Bacillus stearothermophilus* ATCC 12980 and sequencing and cloning of the *SbsD* gene encoding the protein moiety. *Arch. Microbiol.* **177**, 70-80.

Egelseer, E. M., Leitner, K., Jarosch, M., Hotzy, C., Zayni, S., Sleytr, U. B. & Sára, M. (1998). The S-layer proteins of two *Bacillus stearothermophilus* wild-type strains are bound via their N-terminal region to a secondary cell wall polymer of identical chemical composition. *J. Bacteriol.* **180**, 1488-1495.

Egelseer, E. M., Schocher, I., Sleytr, U. B. & Sára, M. (1996). Evidence that an N-terminal S-layer protein fragment triggers the release of a cell-associated high-molecular-weight amylase in *Bacillus stearothermophilus* ATCC 12980. *J. Bacteriol.* **178**, 5602-5609.

Ekwunife, F. S., Singh, J., Taylor, J. k. G. & Doyle, R. J. (1991). Isolation and purification of a cell wall polysaccharide of *Bacillus anthracis* (Sterne). *FEMS Microbiol. Lett.* **82**, 257-262.

Elissalde, M. H., Kamps-Holtzapple, C., Beier, R. C., Plattner, R. D., Rowe, L. D. & Stanker, L. H. (1995). Development of an improved monoclonal antibody-based ELISA for fumonisin B1-3 and the use of molecular modeling to explain observed detection limits. *Food Agr. Imm.* **7**, 109-122.

Engelhardt, H. & Peters, J. (1998). Structural research on surface layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *J. Struct. Biol.* **124**, 276-302.

Fischer, W. (1990). Bacterial phosphoglycolipids and lipoteichoic acids. In: *Handbook of lipid research*. D. J. Hanahan (ed). Plenum Press, New York, pp. 123-217.

Fischer, W., Koch, H. U. & Haas, R. (1983). Improved preparation of lipoteichoic acids. *Eur. J. Biochem.* **133**, 523-530.

Fuller, R. (1989). A review: probiotics in man and animals. *J. Appl. Bacteriol.* **66**, 365-378.

Garcia, J. L., Sanchez-Beato, A. R., Medrano, F. J. & Lopez, R. (1998). Versatility of Choline-binding domain. *Microb. Drug Resist.* **4**, 25-36.

Garduno, R. A., Moore, A. R., Olivier, G., Lizama, A. L., Garduno, E. & Kay, W. W. (2000). Host cell invasion and intracellular residence by *Aeromonas salmonicida*: role of the S-layer. *Can. J. Microbiol.* **46**, 660-668.

Gilbert, C., Atlan, D., Blanc, B., Portailer, R., Germond, J. E., Lapierre, L. & Mollet, B. (1996). A new cell surface proteinase: sequencing and analysis of the prtB gene from *Lactobacillus delbruekii* subsp. *bulgaricus*. *J. Bacteriol.* **178**, 3059-3065.

Goldin, B. R. & Gorbach, S. L. (1992). Probiotics for humans. In: *Probiotics. The scientific basis*. Fuller, R. (ed.). Chapman and Hall, London, Vol. 1, pp. 355-376.

Granato, D., Perotti, F., Masserey, I., Rouvet, M., Golliard, M., Servin, A. & Brassart, D. (1999). Cell surface-associated lipoteichoic acid acts as an adhesion factor for attachment of *Lactobacillus johnsonii* La1 to human enterocyte-like caco-2 cells. *Appl Environ Microbiol* **65**, 1071-1077.

Grangette, C., Muller-Alouf, H., Goudercourt, D., Geoffroy, M. C., Turneer, M. & Mercenier, A. (2001). Mucosal immune responses and protection against tetanus toxin after intranasal immunization with recombinant *Lactobacillus plantarum*. *Infect Immun* **69**, 1547-1553.

Greene, J. D. & Klaenhammer, T. R. (1994). Factors involved in adherence of lactobacilli to human Caco-2 cells. *Appl. Environ. Microbiol.* **60**, 4487-4494.

Hammes, W. P. & Vogel, R. F. (1995). The Genus *Lactobacillus*. In: *The Genera of Lactic Acid Bacteria.* Holzapfel, B. J. B. W. a. W. H., (eds.), Chapman and Hall, London, pp. 19-54.

Hastie, A. T. & Brinton, C. C. (1979). Specific interations of the tetragonally arrayed protein layer of *Bacillus sphaericus* with its peptidoglycan sacculus. *J. Bacteriol.* **138,** 1010-1021.

Hasty, D. L., Ofek, I., Courtney, H. S. & Doyle, R. J. (1992). Multiple adhesins of streptococci. *Infect. Imm.* **60**, 2147-2152.

Houwink, A.L. (1953). A macromolecular mono-layer in the cell wall of *Spirillum* species. *Biochim. Biophys. Acta* **10**, 360-366.

Jacobsen, C. N., Rosenfeldt Nielsen, V., Hayford, A. E., Moller, P. L., Michaelsen, K. F., Paerregaard, A., Sandstrom, B., Tvede, M. & Jakobsen, M. (1999). Screening of probiotic activities of forty-seven strains of *Lactobacillus* spp. by in vitro techniques and evaluation of the colonization ability of five selected strains in humans. *Appl. Environ. Microbiol.* **65**, 4949-4956.

Jarosch, M., Egelseer, E. M., Huber, C., Moll, D., Mattanovich, D., Sleytr, U. B. & Sára, M. (2001). Analysis of the structure-function relationship of the S-layer protein SbsC of *Bacillus stearothermophilus* ATCC 12980 by producing truncated forms. *Microbiology* **147**, 1353-1363.

Jarosch, M., Egelseer, E. M., Mattanovich, D., Sleytr, U. B. & Sára, M. (2000). S-layer gene *sbsC* of *Bacillus stearothermophilus* ATCC 12980: molecular characterization and heterologous expression in Escherichia coli. *Microbiology* **146**, 273-281.

Johnson, J. L., Phelps, C. F., Cummins, C. S., London, J. & Gasser, F. (1980). Taxonomy of the *Lactobacillus acidophilus* group. *Int. J. Syst. Bact.* **30**, 53-68.

Jonquieres, R., Bierne, H., Fiedler, F., Gounon, P. & Cossart, P. (1999). Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of gram-positive bacteria. *Mol. Microbiol.* **34**, 902-914.

Kahala, M. & Palva, A. (1999). The expression signals of the *Lactobacillus brevis slpA* gene direct efficient heterologous protein production in lactic acid bacteria. *Appl. Microbiol. Biotechnol.* **51**, 71-78.

Kailasapathy, K. & Chin, J. (2000). Survival and therapeutic potential of probiotic organisms with reference to *Lactobacillus acidophilus* and *Bifidobacterium* spp. *Immunol. Cell Biol.* **78**, 80-88.

Kandler, O. & Weiss, N. (1986). Regular, Nonsporing Gram-positive Rods. In: *Bergey's manual of Systematic bacteriology*. Vol. 2, Section 14, pp. 1208-1234.

Kay, W. W. & Trust, T. J. (1991). Form and functions of the regular surface array (S-layer) of *Aeromonas salmonicida*. *Experientia* **47**, 412-414.

Kok, J., Van der Vossen, J. & Venema, G. (1984). Construction of plasmid cloning vectors for Lactic streptococci which also replicates in *Bacillus subtilis* and *Escherichia coli*. *Appl. Environ. Microbiol.* **48**, 726-731.

Kotiranta, A., Lounatmaa, K. & Haapasalo, M. (2000). Epidemiology and pathogenesis of *Bacillus cereus* infections. *Microbes Infect.* **2**, 189-198.

Kuen, B., Koch, A., Asenbauer, E., Sára, M. & Lubitz, W. (1997). Molecular characterization of the *Bacillus stearothermophilus* PV72 S-layer gene *sbsB* induced by oxidative stress. *J. Bacteriol.* **179**, 1664-1670.

Kuen, B. & Lubitz, W. (1996). Analysis of S-layer proteins and genes. In: *Crystalline bacterial cell surface proteins*. U.B. Sleytr, P. M., D. Pum, M. Sára, (eds.), pp. 77-102. R.G. Landes Company, Austin.

Kuipers, O. P., Deruyter, P. G. G. A., Kleerebezem, M. & Devos, W. M. (1997). Controlled overproduction of proteins by lactic acid bacteria. *Trends Biotechnol.* **15**, 135-140.

Kullen, M. J. & Klaenhammer, T. R. (2000). Genetic modification of intestinal lactobacilli and bifidobacteria. *Curr. Issues Mol. Biol.* **2**, 41-50.

Ladunga, I. & Smith, R. F. (1997). Amino acid substitutions preserve protein folding by conserving steric and hydrophobic properties. *Prot. Eng.* **10**, 187-196.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Leer, R. J., H. Christiaens, W. Verstraete, L.Peters, M. Posno and P.H. Pouwels. (1993). Gene disruption in *Lactobacillus plantarum* strain 80 by site-specific recombination: isolation of a mutant strain deficient in conjugated bile salt hydrolase activity. *Mol. Gen. Genet.* **239**, 269-272.

Lemaire, M., Miras, I., Gounon, P. & Beguin, P. (1998). Identification of a region responsible for binding to the cell wall within the S-layer protein of *Clostridium thermocellum*. *Microbiology* **144**, 211-217.

Lortal, S., VanHeijenoort, J., Gruber, K. & Sleytr, U. B. (1992). S-layer of *Lactobacillus helveticus* ATCC 12046: isolation, chemical characterization and re-formation after extraction with lithium chloride. *J. Gen. Microbiol.* **138**, 611-618.

Lowry, O. H., Rosebrough, N.J., Farr, A.L., Randall, R.J. (1951). Protein measurement with the folin phenol reagents. *J. Biol. Chem.* **193**, 265-275.

Lubitz, W., Witte, A., Eko, F. O., Kamal, M., Jechlinger, W., Brand, E., Marchart, J., Haidinger, W., Huter, V., Felnerova, D., Stralis-Alves, N., Lechleitner, S., Melzer, H., Szostak, M. P., Resch, S., Mader, H., Kuen, B., Mayr, B., Mayrhofer, P., Geretschlager, R., Haslberger, A. & Hensel, A. (1999). Extended recombinant bacterial ghost system. *J. Biotechnol.* **73**, 261-273.

Luchansky, J. B., Muriana, P. M. & Klaenhammer, T. R. (1988). Application of electroporation for transfer of plasmid DNA to *Lactobacillus, Lactococcus, Leuconostoc, Listeria, Pediococcus, Bacillus, Staphylococcus, Enterococcus* and *Propionibacterium*. *Mol. Microbiol.* **2**, 637-646.

Lupas, A. (1996a). A circular permutation event in the evolution of the SLH domain? *Mol. Microbiol.* **20**, 897-898.

Lupas, A. (1996b). Prediction and analysis of coiled-coil structures. *Meth. Enzymol.* **266**, 513-525.

Lupas, A., Engelhardt, H., Peters, J., Santarius, U., Volker, S. & Baumeister, W. (1994). Domain structure of the *Acetogenium kivui* surface layer revealed by electron crystallography and sequence analysis. *J. Bacteriol.* **176**, 1224-1233.

Maassen, C. B., Laman, J. D., den Bak-Glashouwer, M. J., Tielen, F. J., van Holten-Neelen, J. C., Hoogteijling, L., Antonissen, C., Leer, R. J., Pouwels, P. H., Boersma, W. J. & Shaw, D. M. (1999). Instruments for oral disease-intervention strategies: recombinant *Lactobacillus casei* expressing tetanus toxin fragment C for vaccination or myelin proteins for oral tolerance induction in multiple sclerosis. *Vaccine* **17**(17), 2117-2128.

Martinez, B., Sillanpää, J., Smit, E., Korhonen, T. K. & Pouwels, P. H. (2000). Expression of *cbsA* encoding the collagen-binding S-protein of *Lactobacillus crispatus* JCM 5810 in *Lactobacillus casei* ATCC 393^{T}. *J. Bacteriol.* **182**, 6857-6861.

Märyä-Mäkinen, A. & Bigret, M. (1998). Industrial use and production of lactic acid bacteria. In: *Lactic acid bacteria. Microbiology and functional aspects*. A. von Wright, S. Salminen (eds.) Second edition, revised and expanded, pp. 73-102. Marcel Dekker Inc., New York.

Masuda, K. (1992). Heterogeneity of S-layer proteins of *Lactobacillus acidophilus* strains. *Microbiol. Immunol.* **36**, 297-301.

Masuda, K. & Kawata, T. (1983). Distribution and chemical characterization of regular arrays in the cell walls of strains of the genus *Lactobacillus*. *FEMS Microbiol. Lett.* **20**, 145-150.

Matsuzaki, T. & Chin, J. (2000). Modulating immune responses with probiotic bacteria. *Immunol. Cell. Biol.* **78**, 67-73.

Matuschek, M., Sahm, K., Zibat, A. & Bahl, H. (1996). Characterization of genes from *Thermoanaerobacterium thermosulfurigenes* EM1 that encode two glycosylhydrolases with conserved S-layer-like domains. *Mol. Gen. Genet.* **252**, 493-496.

Medaglini, D., Oggioni, M. R. & Pozzi, G. (1998). Vaginal immunization with recombinant gram-positive bacteria. *Am. J. Reprod. Immunol.* **39**, 199-208.

Mercenier, A., Muller-Alouf, H. & Grangette, C. (2000). Lactic acid bacteria as live vaccines. *Curr. Issues Mol. Biol.* **2**, 17-25.

Mesnage, S., Fontaine, T., Mignot, T., Delepierre, M., Mock, M. & Fouet, A. (2000). Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *Embo J.* **19**, 4473-4484.

Mesnage, S., Tosi-Couture, E. & Fouet, A. (1999a). Production and cell surface anchoring of functional fusions between the SLH motifs of the *Bacillus anthracis* S-layer proteins and the *Bacillus subtilis* levansucrase. *Mol. Microbiol.* **31**, 927-936.

Mesnage, S., Tosi-Couture, E., Mock, M. & Fouet, A. (1999b). The S-layer homology domain as a means for anchoring heterologous proteins on the cell surface of *Bacillus anthracis*. *J. Appl. Microbiol.* **87**, 256-260.

Mesnage, S., Tosicouture, E., Mock, M., Gounon, P. & Fouet, A. (1997). Molecular characterization of the *Bacillus anthracis* main S-layer component: Evidence that it is the major cell-associated antigen. *Mol. Microbiol.* **23**, 1147-1155.

Millsap, K. W., Reid, G., Vandermei, H. C. & Busscher, H. J. (1997). Cluster analysis of genotypically characterized *Lactobacillus* species based on physicochemical cell surface properties and their relationship with adhesion to hexadecane. *Can. J. Microbiol.* **43**, 284-291.

Miyazawa, S. & Jernigan, R. L. (2000). Identifying sequence-structure pairs undetected by sequence alignments. *Prot. Eng.* **13**, 459-475.

Navarre, W. W. & Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* **63**, 174-229.

Neeser, J. R., Granato, D., Rouvet, M., Servin, A., Teneberg, S. & Karlsson, K. A. (2000). *Lactobacillus johnsonii* La1 shares carbohydrate-binding specificities with several enteropathogenic bacteria. *Glycobiol.* **10**(11), 1193-1199.

Nielsen, H., Engelbrecht, J., Brunak, S. & Von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Prot. Eng.* **10**, 1-6.

Noonan, B. & Trust, T. J. (1997). The synthesis, secretion and role in virulence of the paracrystalline surface protein layers of *Aeromonas salmonicida* and *Aeromonas hydrophila*. *FEMS Microbiol. Lett.* **154**, 1-7.

Norton, D. D., Dwyer, D. S. & Muggeridge, M. I. (1998). Use of a neural network secondary structure prediction to define targets for mutagenesis of herpes simplex virus glycoprotein B. *Virus Res.* **55**, 37-48.

Olabarria, G., Carrascosa, J., L., Pedro, M., A., de, Berenguer, J. (1996). A conserved motif in S-layer proteins is involved in peptidoglycan binding in *Thermus thermophilus*. *J. Bacteriol.* **178**, 4765-4772.

Ouwehand, A. C. (1998). Antimicrobial Components from Lactic Acid Bacteria. In: *Lactic acid bacteria. Microbiology and functional aspects*. A. von Wright, S. Salminen (eds.) Second edition, revised and expanded, pp. 139-159. Marcel Dekker Inc., New York.

Pederson, J. A., Mileski, G. J., Weimer, B. C. & Steele, J. L. (1999). Genetic characterization of a cell envelope-associated proteinase from *Lactobacillus helveticus* CNRZ32. *J. Bacteriol.* **181**, 4592-4597.

Peters, J., Peters, M., Lottspeich, F. & Baumeister, W. (1989). S-layer protein gene of *Acetogenium kivui*: cloning and expression in *Escherichia coli* and determination of the nucleotide sequence. *J. Bacteriol.* **171**, 6307-6315.

Platteeuw, C., Simons, G. De Vos, W.M. (1994). Use of the *Escherichia coli* Beta-Glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria. *Appl. Environ. Microbiol.* **60**, 587-593.

Posno, M., R.J. Leer, N. van Lujk, M.J.F. van Giezen, P.T.H.M. Heuvelmans, B.C. Lokman, and P.H. Pouwels. (1991). Incompatibility of *Lactobacillus* vectors with replicons derived from small cryptic *Lactobacillus* plasmids and segregational instability of the introduced vectors. *Appl. Environ. Microbiol.* **57**, 1822-1828.

Pouwels, P. H., Leer, R. J., Shaw, M., Heijne den Bak-Glashouwer, M. J., Tielen, F. D., Smit, E., Martinez, B., Jore, J. & Conway, P. L. (1998). Lactic acid bacteria as antigen delivery vehicles for oral immunization purposes. *Int. J. Food Microbiol.* **41**, 155-167.

Pouwels, P. H., Vriesema, A., Martinez, B., Tielen, F. J., Seegers, J. F. M. L., Leer, R. J., Jore, J. & Smit, E. (2001). Lactobacilli as vehicles for targeting antigens to mucosal tissues by surface exposition of antigens. *Meth. Enzymol.* **336**, 369-389.

Pugsley, A. P. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiological Reviews* **57**, 50-108.

Pum, D. & Sleytr, U. B. (1996). Molecular nanotechnology and biomimetics with S-layers. In: *Crystalline bacterial cell surface proteins*. U.B. Sleytr, P. M., D. Pum, M. Sára, (eds.), pp. 175-209. R.G. Landes Company, Austin.

Pum, D., Sleytr., U.B. (1993). Molecular nanotechnology with S-layers. In: *Advances in paracrystalline bacterial surface layers*. Beveridge, T., J, and Koval, S.F., (eds.), Vol. 252, pp. 205 - 218, Plenum Press, New York.

Rinia, H. A., Demel. R.A., Van der Eerden, J. P. J. M. & De Kruijff, B. (1999). Blistering of Langmuir-Blodgett bilayers containing anionic phospholipids as observed by Atomic Force Microscopy. *Biophys. J.* **77**, 1683-1693.

Roos, S. & Jonsson, H. (2002). A high-molecular-mass cell-surface protein from *Lactobacillus reuteri* 1063 adheres to mucus components. *Microbiology* **148**, 433-442.

Rost, B. (1996). PHD: Predicting one-dimensional protein structure by profile based neural networks. *Meth. Enzymol.* **266**, 525-539.

Rost, B. & Sander, C. (1993). Prediction of protein structure at better than 70% accuracy. *J. Mol. Biol.* **232**, 584-599.

Russell, W. M. & Klaenhammer, T. R. (2001). Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Appl. Environ. Microbiol.* **67**, 4361-4364.

Salminen, S., Deighton, M. A., Benno, Y. & Gorbach, S. L. (1998). Lactic acid bacteria in health and disease. In: *Lactic acid bacteria. Microbiology and functional aspects*. A. von Wright, S. Salminen (eds.) Second edition, revised and expanded, pp. 211-253. Marcel Dekker Inc., New York.

Salminen, S., Isolauri, E. & E., S. (1996). Clinical uses of probiotics for stabilizing the gut mucosal barrier: successful strains and future challenges. *Antonie van Leeuwenhoek* **70**, 347-358.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanders, M. E. & Klaenhammer, T. R. (2001). Invited review: the scientific basis of *Lactobacillus acidophilus* NCFM functionality as a probiotic. *J. Dairy Sci.* **84**, 319-331.

Sára, M., Egelseer, E. M., Dekitsch, C., Sleytr, U. B. (1998). Identification of two binding domains, one for peptidoglycan and another for a secondary cell wall polymer, on the N-terminal part of the S-layer protein SbsB from Bacillus stearothermophilus PV72/p2. *J. Bacteriol.* **180**, 6780-6783.

Sára, M., Kupcu, S., Weiner, C., Weigert, S. and Sleytr, U.B. (1993). S-layer as immobilization and affinity matrices. In: *Advances in paracrystalline bacterial surface layers*. Beveridge, T., J, and Koval, S.F., (eds.), Vol. 252, pp. 195-204, Plenum Press, New York.

Sára, M. (2001). Conserved anchoring mechanisms between crystalline cell surface S-layer proteins and secondary cell wall polymers in Gram-positive bacteria? *Trends Microbiol.* **9**, 47-50.

Sára, M. & Sleytr, U. B. (1987a). Molecular sieving through S-layers of *Bacillus stearothermopilus* strains. *J. Bacteriol.* **169**, 4092-4098.

Sára, M. & Sleytr, U. B. (1987b). Production and characteristics of ultrafiltration membranes with uniform pores from two-dimensional arrays of proteins. *J. Membr. Sci.* **33**, 27-49.

Sára, M. & Sleytr, U. B. (2000). S-Layer proteins. *J. Bacteriol.* **182**, 859-868.

Savijoki, K., Kahala, M. & Palva, A. (1997). High level heterologous protein production in *Lactococcus* and *Lactobacillus* using a new secretion system based on the *Lactobacillus brevis* S-layer signals. *Gene* **186**, 255-262.

Saxton, W. O. & Baumeister, W. (1986). Principles of organization in S layers. *J. Mol. Biol.* **187**, 251-253.

Schneitz, C., Nuotio, L. & Lounatma, K. (1993). Adhesion of *Lactobacillus acidophilus* to avian intestinal epithelial cells mediated by the crystalline bacterial cell surface layer (S-layer). *J. Appl. Bacteriol.* **74**, 290-294.

Schuster, K. C., Pink, T., Mayer, H. F., Hampel, W. A. & Sára, M. (1997). Oxygen-triggered synchronized variant formation of the S-layer carrying *Bacillus stearothermophilus* PV72 during continuous cultivation. *J. Biotechnol.* **54**, 15-28.

Shaw, D. M., Gaerthe, B., Leer, R. J., Van Der Stap, J. G., Smittenaar, C., Heijne Den Bak-Glashouwer, M., Thole, J. E., Tielen, F. J., Pouwels, P. H. & Havenith, C. E. (2000). Engineering the microflora to vaccinate the mucosa: serum immunoglobulin G responses and activated draining cervical lymph nodes following mucosal application of tetanus toxin fragment C-expressing lactobacilli. *Immunology* **100**, 510-518.

Siezen, R. J. (1999). Multi-domain, cell-envelope proteinases of lactic acid bacteria. *Antonie van Leeuwenhoek* **76**, 139-155.

Silin, V. & Plant A. (1997). Biotechnological applications of surface plasmon resonance. *Tibtech.* **15**, 353-359.

Sillanpää, J., Martinez, B., Antikainen, J., Toba, T., Kalkkinen, N., Tankka, S., Lounatmaa, K., Keränen, J., Höök, M., Westerlund-Wikström, B., Pouwels, P. H. & and Korhonen, T. K. (2000). Characterization of the collagen-binding S-layer protein CbsA of *Lactobacillus crispatus.* *J. Bacteriol.* **182**(22), 6440-6450.

Simon, B., Nomellini, J., Chiou, P., Bingle, W., Thornton, J., Smit, J. & Leong, J. A. (2001). Recombinant vaccines against infectious hematopoietic necrosis virus: production by the *Caulobacter crescentus* S-layer protein secretion system and evaluation in laboratory trials. *Dis. Aquat. Organ.* **44**(1), 17-27.

Sleytr, U. B., Bayley, H., Sára, M., Breitwieser, A., Kupcu, S., Mader, C., Weigert, S., Unger, F. M., Messner, P., Jahn-Schmid, B., Schuster, B., Pum, D., Douglas, K., Clark, N. A., Moore, J. T., Winningham, T. A., Levy, S., Frithsen, I., Pankovc, J., Beale, P., Gillis, H. P., Choutov, D. A. & Martin, K. P. (1997a). Applications of S-layers. *FEMS Microbiol. Rev.* **20**, 151-175.

Sleytr, U. B. & Beveridge, T. J. (1999). Bacterial S-layers. *Trends Microbiol.* **7**, 253-260.

Sleytr, U. B. & Glauert, A. M. (1975). Analysis of regular arrays of subunits on bacterial surfaces; Evidence for a dynamic process of assembly. *J. Ultrastruct. Res.* **50**, 103-116.

Sleytr, U. B. & Messner, P. (1983). Crystalline surface layers on bacteria. *Annu. Rev. Microbiol.* **37**, 311-339.

Sleytr, U. B. & Messner, P. (1988). Crystalline surface layer in procaryotes. *J. Bacteriol.* **170**, 2891-2897.

Sleytr, U. B., Messner, P., Pum, D. & Sára, M. (1996). Occurence, location, ultrastructure and morphogenesis of S-layers. In *Crystalline bacterial cell surface proteins*. U.B. Sleytr, P. M., D. Pum and M. Sára, (eds.), pp. 5-33. R.G. Landes Company, Austin.

Sleytr, U. B., Messner, P., Pum, D. & Sára, M. (1999). Crystalline bacterial cell surface layers (S-layers): From supramolecular cell structure to Biomimetics and Nanotechnology. *Angewandte Chemie* **38**, 1034-1054.

Sleytr, U. B., Pum, D. & Sára, M. (1997b). Advances in S-layer nanotechnology and biomimetics. *Adv. Biophysics* **34**, 71-79.

Smit, E., Jager, D., Martinez, B., Tielen, F. J. & Pouwels, P. H. (2002). Structural and functional analysis of the S-layer protein crystallization domain of *Lactobacillus acidophilus* ATCC 4356: evidence for protein-protein interaction of two subdomains. *Submitted*.

Smit, E., Oling, F., Demel, R., Martinez, B. & Pouwels, P. H. (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**(2), 245-257.

Smit, E. & Pouwels, P. H. (2002). One repeat of the cell wall binding domain is sufficient for the anchoring of the *Lactobacillus acidophilus* surface layer protein. *Submitted*.

Stahl, D. A., Key, R., Flescher, B. & Smit, J. (1992). The phylogeny of marine and freshwater caulobacters reflects their habitat. *J. Bacteriol.* **174**, 2193-2198.

Thomas, S., Austin, J. W., McCubbin, W. D., Kay, C. M. & Trust, T. J. (1992). Roles of structural domains in the morphology and surface anchoring of the tetragonal paracrystalline array of *Aeromonas hydrophila*. *J. Mol. Biol.* **228**, 652-661.

Thompson, S. A., Shedd, O. L., Ray, K. C., Beins, M. H., Jorgensen, J. P. & Blaser, M. J. (1998). *Campylobacter fetus* surface layer proteins are transported by a type I secretion system. *J. Bacteriol.* **180**, 6450-6458.

Toba, T., Virkola, R., Westerlund, B., Björkman, Y., Sillanpää, J., Vartio, T., Kalkkinen, N. & Korhonen, T. K. (1995). A collagen-binding S-layer protein in *Lactobacillus crispatus*. *Appl. Environ. Microbiol.* **61**, 2467-2471.

Towner, P. (1991). Purification of DNA with CTAB. In: *Essential Molecular Biology. A practical Approach.* Brown, T. A., (ed.), pp. 54-55. Oxford University Press, Oxford.

van Geest, M. & Lolkema, J. S. (2000). Membrane topology of the Na(+)/citrate transporter CitS of *Klebsiella pneumoniae* by insertion mutagenesis. *Biochim. Biophys. Acta* **1466**, 328-338.

Vidgrén, G., Palva, I., Pakkanen, R., Lounatmaa, K. & Palva, A. (1992). S-layer protein gene of *Lactobacillus brevis*: cloning by polymerase chain reaction and determination of the nucleotide sequence. *J. Bacteriol.* **174**, 7419-7427.

von Eichel-Streiber, C., Sauerborn, M. & Kuramitsu, H. K. (1992). Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of *Clostridium difficile* toxins and *Streptococcus mutans* glucosyltransferases. *J. Bacteriol.* **174**, 6707-6710.

Walker, D. C., Aoyama, K. & Klaenhammer, T. R. (1996). Electrotransformation of *Lactobacillus acidophilus* group A1. *FEMS Microbiol. Lett.* **138**, 233-237.

Walker, S. G., Smith, S. H. & Smit, J. (1992). Isolation and comparison of the paracrystalline surface layer proteins of freshwater caulobacters. *J. Bacteriol.* **174**, 1783-1792.

Wang, E., Garcia, M. M., Blake, M. S., Pei, Z. & Blaser, M. J. (1993). Shift in S-layer protein expression responsible for antigenic variation in *Campylobacter fetus*. *J. Bacteriol.* **175**, 4979-4984.

Ward, J. B. (1981). Teichoic and teichuronic acids: Biosynthesis, assembly, and location. *Microbiol. Rev.* **45**, 211-

Weis, W. I. (1997). Cell-surface carbohydrate recognition by animal and viral lectins. *Curr. Opin. Struct. Biol.* **7**, 624-30.

Whittaker, C. J., Klier, C. M. & Kolenbrander, P. E. (1996). Mechanisms of adhesion by oral bacteria. *Annu. Rev. Microbiol.* **50**, 513-552.

Wong, K. K. & Hancock, R. E. (2000). Insertion mutagenesis and membrane topology model of the *Pseudomonas aeruginosa* outer membrane protein OprM. *J. Bacteriol.* **182**, 2402-10.

Woodbury, R. G., Wendin, C., Clendenning, J., Melendez, J., Elkind, J., Bartholomew, D., Brown, S. & Furlong, C. E. (1998). Construction of biosensors using a gold-binding polypeptide and a miniature integrated surface plasmon sensor. *Biosens. Bioelectr.* **13**, 1117-1126.

Wren, B. W. (1991). A family of clostridial and streptococcal ligand-binding proteins with conserved C-terminal repeat sequences. *Mol. Microbiol.* **5**, 797-803.

Wright, A. von & Sibakov, M. (1998). Genetic modification of lactic acid bacteria. In: *Lactic acid bacteria. Microbiology and functional aspects*. A. von Wright, S. Salminen (eds.) Second edition, revised and expanded, pp. 161-210. Marcel Dekker Inc., New York.

Yamamoto, N., Shinoda, T. & Takano, T. (1999). Molecular cloning and sequence analysis of a gene encoding an extracellular proteinase from *Lactobacillus helveticus* CP790. *Biosci. Biotechnol. Bioeng.* **64**, 1217-1222.

Yang, D. & Woese, C. R. (1989). Phylogenetic structure of the "leuconostocs": an interesting case of a rapidly evolving organism. *Syst. Appl. Microbiol.* **12**, 145-149.

Yasui, T. & Yoda, K. (1997). Purification and partial characterization of an antigen specific to *Lactobacillus brevis* strains with beer spoilage activity. *FEMS Microbiol. Lett.* **151**, 169-76.

Yother, J. & White, J. M. (1994). Novel surface attachment mechanism of the *Streptococcus pneumoniae* protein PspA. *J. Bacteriol.* **176**, 2976-85.

Nederlandse samenvatting voor niet-ingewijden

Het menselijk lichaam komt dagelijks in contact met vele ziektekiemen uit het ons omringende milieu. Velen daarvan zijn bacteriën, maar daarnaast kunnen b.v. ook virussen en schimmels gevaar voor de gezondheid opleveren. Ziektekiemen moeten het lichaam binnen zien te dringen via b.v. de huid, maar met name de luchtwegen en het maagdarmkanaal zijn belangrijke toegangspoorten voor ziektekiemen. Deze organen worden niet beschermd door een stugge cellaag, zoals de huid, maar slechts door een dunne mucus (slijm)-laag die de cellen eronder bedekt. Daarbij komt ook nog dat longen en b.v. darmen een zeer groot inwendig oppervlak hebben, dat met het milieu in contact staat en wat ziekteverwekkers zeer veel toegangsmogelijkheden biedt.

Het lichaam bezit een belangrijk afweermechanisme dat infectie door ziektekiemen tegengaat, het zogenaamde immuunsysteem. Dit systeem bestaat uit een aantal belangrijke onderdelen die door het lichaam constant worden aangemaakt, waaronder antilichamen en bepaalde celtypes. Komt er een ziektekiem het lichaam binnen, dan wordt deze direct aangevallen en geneutraliseerd. Longen en maagdarmkanaal hebben een groot oppervlak en om te voorkomen dat er ziektekiemen binnenkomen via deze route is er daar een gespecialiseerd immuunsysteem actief, het zogenaamde mucosale immuunsysteem. Mucosaal betekent zoveel als: 'verbonden aan door mucus bedekte oppervlakken'. Naast dit lichaamseigen beschermingsmechanisme kunnen ook bacteriën een rol spelen bij bescherming van mucosale oppervlakken van de longen en het maagdarmkanaal.

Overal in de wereld om ons en ook op en in delen van het menselijk lichaam bevinden zich bacteriën zo ook in b.v. onze longen of het maagdarmkanaal. Deze bacteriën zijn geen ziekteverwekkers en kunnen zelfs zeer gewenste gasten zijn. De zogenaamde melkzuurbacteriën (MZB) behoren tot deze categorie. Ze heten zo omdat deze bacteriën melkzuur maken tijdens hun groei, een eigenschap die wordt gebruikt bij het maken van yoghurt (zuur!!) en kaas uit melk, maar b.v. ook bij de productie van salami. Door de aanwezigheid van melkzuur vinden andere bacteriën het niet prettig om in die producten te groeien, waardoor ze langer houdbaar zijn dan de uitgangsproducten (antibacterieel effect). Bovendien maken MZB naast melkzuur nog verschillende andere stofjes die nog eens extra bijdragen aan dit effect. En aangezien mensen dit soort voedingsproducten veelvuldig eten krijgen zij deze bacteriën ook binnen.

Er is de laatste tien tot twintig jaar zeer veel bekend geworden over de gunstige eigenschappen van (consumptie van) MZB voor de menselijke gezondheid. De

antibacteriële eigenschappen kunnen waarschijnlijk ook zorgen voor de bestrijding van ongewenste bacteriën in ons lichaam. Naast de productie van antibacteriële stoffen kunnen bepaalde MZB ook langere tijd in ons lichaam verblijven, doordat zij goed kunnen hechten aan het mucosale oppervlak. Deze hechting heeft een dubbel effect, namelijk het verhinderen van de hechting van minder gewenste bacteriën, maar geeft ook een langer durende antibacteriële activiteit. Op basis van deze bevindingen zijn aan veel MZB zogenaamde 'probiotische eigenschappen' (eigenschappen gunstig voor de gezondheid van het organisme, oftewel ons lichaam) toegekend. Op de voedselmarkt zijn de laatste jaren een groot aantal producten verschenen met daarin probiotische MZB zoals b.v. de welbekende Yakult en Vifit.

Een belangrijke eigenschap van een probiotische MZB is, zoals beschreven, de hechting aan mucosale oppervlakken. De hechting van een bacterie aan een oppervlak is afhankelijk van de samenstelling van zowel dat oppervlak als van de bacterie.

In ons onderzoek hebben we gewerkt aan de MZB *Lactobacillus acidophilus*. De bacterie heeft een bijzondere celoppervlakte eigenschap. *L. acidophilus* bezit een zeer stevig omhulsel, de celwand, waarop zich nog een laag bevindt, de Surface layer of S-layer (oppervlakte laag). Deze laag zou betrokken kunnen zijn bij hechting aan b.v. het darmoppervlak. Ons onderzoek heeft zich hoofdzakelijk gericht op het bestuderen van de opbouw van deze laag, omdat dat informatie kan opleveren over de rol in hechting van de bacterie aan b.v. darm oppervlaken.

De S-layer is opgebouwd uit een groot aantal identieke bouwstenen, het S_A-eiwit. Eiwitten zijn grote en complexe moleculen, afgeleid van de overeenkomstige genvolgorde op het DNA van de bacterie. Sommige bacteriën hebben b.v. een S-layer bestaande uit vele duizenden S-eiwitten. Deze eiwitten vormen een groot netwerk op het oppervlak van de bacterie. Het netwerk wordt gevormd doordat elk S_A-eiwit graag een aantal andere S_A-eiwit moleculen om zich heen heeft en deze stevig vasthoudt. Dat gebeurt op een hele geordende manier waardoor het netwerk er b.v. uit ziet als een mooi gelegde klinkerstraat op de hele bacterie. Zo'n geordende eiwit structuur heel ook wel een kristallijne structuur.

Onze interesse ging in eerste instantie uit naar het bekijken hoe het S_A-eiwit op het celoppervlak zit en hoe het andere S-eiwit moleculen vasthoudt (Hoofdstuk 2). Wat wij vonden was dat elk S_A-eiwit molecuul eigenlijk uit twee delen (domeinen) bestond die onafhankelijk van elkaar hun werk konden doen. Zo vonden we een deel dat zorgde voor het vasthouden aan het bacterieoppervlak, zodat de S-layer niet zomaar van de bacterie

valt(celwand bindingsdomein). Daarnaast vonden we een ander deel dat zorgde voor het vasthouden van naburige moleculen wat dus zorgt dat het S-layer netwerk kan worden gevormd (kristallisatie domein). Deze losse functionele delen van een eiwit heten ook wel domeinen.

Hoe moet je je dit nu voorstellen? Neem b.v. aan dat een S_A-eiwit molecuul bestaat uit twee kubusvormige delen, een grote kubus bovenop een kleine. De onderkant van de kleine kubus staat op de grond en houdt het geheel daaraan vast. De grote kubus is niet graag alleen en wil graag andere kubussen vasthouden, bijvoorbeeld met elk van de vier zijvlakken één andere kubus. Zo kun je je voorstellen dat, als je een heleboel naast elkaar zet, er een mooi geordende laag ontstaat.

Wat wij vervolgende wilde weten is of we de kubussen (domeinen) verder uit elkaar konden halen om zo meer te weten te komen over hun eigenschappen (hoofdstukken 3 en 5). Uit verder onderzoek bleek inderdaad dat de kleine kubus feitelijk bestond uit twee helften (subdomeinen), waarvan er maar één goed contact leek te maken met de ondergrond. De tweede helft zou nodig kunnen zijn voor het versterken van het contact met de ondergrond ook al kan hij dat contact niet zelf verzorgen. Ook de grote kubus bleek op te delen in twee helften (stel je een verticaal snijvlak door de kubus voor). Beide konden in ieder geval 1 buurmolecuul vasthouden. Mogelijk zijn bij de opdeling de twee andere vlakken doorsneden, waardoor contact aan die zijde met het naastliggende molecuul niet goed is.

Naast het contact van het S-eiwit met zowel het bacterieoppervlak als met zijn buureiwitten, waren we ook geïnteresseerd in de hechtingsmogelijkheden van het S-eiwit aan oppervlakken (Hoofdstuk 4 en 6). Hiertoe moesten we bepalen welk deel van het S-eiwit kristallisatiedomein naar buiten toe is gericht (bovenkant van de kubus). Dit hebben we onderzocht door kleine eiwitonderdelen toe te voegen aan de grote kubus en te bepalen of deze aan de bovenkant terecht kwamen. Bepaalde toevoegingen bleken inderdaad zichtbaar te zijn op de bovenkant. Daarnaast toonde serie andere experimenten aan dat deze bovenkant op kunstmatige oppervlakken goed kan vasthouden. Welk deel van de bovenkant echt belangrijk is, is nog onduidelijk.

Al met al hebben we veel inzicht verkregen in de samenstelling van het S_A-eiwit en hebben we goede aanwijzingen welk deel van het eiwit bij hechting betrokken zou kunnen zijn. En zoals altijd roept een verkregen antwoord steeds weer extra vragen op. Heeft u nog vragen? Ik hoor het wel.

Curriculum vitae

Egbert Smit werd geboren op 26 december 1971 te Eindhoven. Hij verhuisde al snel naar Krommenie, waar een deel van de lagere school tijd werd doorgebracht. Rond zijn negende jaar verhuisde de familie Smit, inmiddels uitgebreid met zusje Renée, naar Coevorden, waar al spoedig broertje Harald het gezin completeerde. Egbert sloot de lagere school in 1985 succesvol af en begon aan de Rijksscholengemeenschap te Coevorden aan het Atheneum, waar hij in mei 1990 zijn diploma ontving. Na zijn afstuderen ging hij Tropische Landbouw studeren aan de Internationale Hogeschool Larenstein in Deventer, waar hij in mei 1991 zijn propedeuse haalde. In september 1991 begon hij met de studie biologie aan de Rijksuniversiteit Groningen. Hij specialiseerde zich in de moleculaire biologie en voerde een tweetal afstudeeronderzoeken uit. De eerste bij de groep Moleculaire Genetica aan de Rijksuniversiteit onder begeleiding van Arjen Nauta, met als onderwerp de analyse van de DNA-bindingseigenschappen van bacteriofaag gecodeerde eiwitten. Zijn tweede afstudeeronderwerp werd uitgevoerd bij TNO Voeding in Zeist in de groep Bioprocessing & Biomonitoring onder begeleiding van Ineke van Geel, met als onderwerp de karakterisatie van extracellulaire glycosyltransferases van *Lactobacillus reuteri*. Volgend op zijn afstuderen in juni 1996 begon hij als onderzoeker in opleiding in de groep Moleculaire Genetica en Gentechnologie (MGG) bij TNO in Rijswijk onder begeleiding van Prof. Peter Pouwels. Zijn promotieonderzoek had als onderwerp de structurele en functionele analyse van het Surface layer eiwit van *Lactobacillus acidophilus*, waarvan de bevindingen in dit proefschrift staan beschreven.

Vanaf maart 2002 werkt hij als post-doc in de groep Humane Biologie van de Universiteit Maastricht aan de ontwikkeling van antibody arrays.

Publications

Nauta, A., Van Sinderen, D., Karsens, H., **Smit, E.**, Venema, G. and Kok, J. (1996) Inducible gene expression mediated by a repressor-operator system isolated from *Lactococcus lactis* bacteriophage r1t. *Mol. Mic.* **19**, 1331-1341

Pouwels, P.H., Leer, R.J., Shaw, M., Heijne den Bak-Glashouwer, M.J., Tielen, F.D., **Smit, E.**, Martinez, B., Jore, J., Conway, P.L. (1998) Lactic acid bacteria as antigen delivery vehicles for oral immunisation purposes. *Int. J. Food Microbiol.* **41**, 155-67

Van Geel-Schutten, G.H., Faber, E.J., **Smit, E.**, Bonting, K., Smith, M.R., Ten Brink, B., Kamerling, J.P., Vliegenthart, J.F., Dijkhuizen, L. (1999) Biochemical and structural characterisation of the glucan and fructan exopolysaccharides synthesised by the *Lactobacillus reuteri* wild-type strain and by mutant strains. *Appl. Environ. Microbiol.* **65**, 3008-3014

Martinez, B., Sillanpaa, J., **Smit, E.**, Korhonen, T.K., Pouwels, P.H. (2000) Expression of cbsA encoding the collagen-binding S-protein of *Lactobacillus crispatus* JCM 5810 in *Lactobacillus casei* ATCC 393(T). *J. Bacteriol.* **182**, 6857-6861.

Pouwels, P.H., Vriesema, A., Martinez, B., Tielen, F.J., Seegers, J.F.M.L., Jore, J., **Smit, E.** (2001) Lactobacilli as vehicles for targeting antigens to mucosal tissues by surface exposition of foreign antigens. *Methods in Enzymology.* **336**, 369-389.

Smit, E., Oling, F., Demel, R., Martinez, B., Pouwels, P.H. (2001) The S-layer protein of *Lactobacillus acidophilus* ATCC 4356: Identification and characterisation of domains responsible for S-protein assembly and cell wall binding. *J. Mol. Biol.* **305**, 245-57.

Smit, E., Jager, D., Tielen F.J., Pouwels, P.H. (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. Submitted *J. Mol. Biol.*

Smit, E., Pouwels P.H. (2002) One repeat of the cell wall binding domain is sufficient for the anchoring of the *Lactobacillus acidophilus* surface layer protein. *J. Bacteriol.* in press

Smit, E., Stigter, E., Diettrich, H., Past, S. and Pouwels, P.H. (2002) The *Lactobacillus acidophilus* S-layer protein assembles in an ordered array on a gold surface. Implications for the development of a biosensor coating. Manuscript in preparation

Dankwoord

En nu dat deel van het proefschrift wat door iedereen als eerste en enige volledig wordt gelezen. Meteen opslaan dat dankwoord: En? Sta ik erbij, sta ik erbij..? Gevolgd door grote teleurstelling, je valt onder het kopje verzamel..... Goed, dat wil niet zeggen dat je niet belangrijk was, maar het is uiteraard onmogelijk om iedereen persoonlijk te bedanken in de ruimte die ik mezelf ter beschikking stel. Als ik me zou laten gaan dan kwamen er nog 120 pagina's bij. Tsja, met wie te beginnen, denk je dan. Je ouders, voor alle ontwikkelingsmogelijkheden en de grote keuzevrijheid die ze je altijd geboden hebben? Of je promotor, voor alle overgedragen kennis? Of moet je echtgenote als eerste, voor alle bijstand in moeilijke momenten en late uurtjes?

Laat ik beginnen bij mijn roots. Lieve ouders, als ik terugdenk begon het allemaal al toen jullie vragen van een vijf jaar oude Egbert, zoals: "Mama, waarom is gras groen? Papa, wat zijn sterren?", serieus gingen beantwoorden. De nieuwsgierigheid naar het onbekende, dat is er nooit meer uitgegaan en het gereedkomen van dit proefschrift is eigenlijk jullie toedoen.

Voordat ik namen ga noemen wil ik de gehele afdeling bedanken voor wat jullie zijn: de beste groep collega's die ik me had kunnen wensen! Jullie waren er voor me op zowel goede als minder goede momenten over de afgelopen jaren en dat heb ik zeer gewaardeerd. Laat ik tenslotte dit zeggen: ook al ben ik uit het directe gezichtsveld verdwenen en is Maastricht niet om de hoek, we zien elkaar op 't Onze Lieve Vrouweplein!

"The foreigners" club: Bea, Bernhard and Eiichi. Actually, I should include myself in this group, since I was named "The foreigner from Groningen" by Rob in the first weeks after I started in Rijswijk. I remember being in Rijswijk with great pleasure (cooking together, having a drink together). Those were great times! Bernhard, the visit of the Dutch/Japanese delegation to Vienna was also very memorable: remember the four liters? I don't remember anything about a newspaper stand either..... Bea, I can't tell you how much I have always appreciated your companionship. You really thought me a lot!

Nu de categorie sport en ontspanning, essentieel buiten het werk. Om te beginnen Iduna, waar je een gezonde balans aantreft tussen ontspannen (biertje?) en inspannen (badminton). Vanaf het moment dat ik bij Iduna ging spelen raakte ik pas goed ingeburgerd in Utrecht, een stad naar mijn hart. Idunesen Michiel, Colinda, Rene,

Marieke, Annemiek, Antoinette, Jeroen en Irene in het bijzonder bedankt voor de nodige prettige maandag- en woensdagavonden en de momenten daarbuiten. Verder wil ik Jasper bedanken voor de ontspannende avondjes snooker, waarin ik na al die tijd nog niet echt beter ben geworden. Ook waren daar Nicole, Ana C., Peter, Vief en Ana L. als collega's, maar ook als vrienden voor gezellige etentjes, drankjes en andersoortige uitjes zo nu en dan.

Irene, jouw adviezen hebben me gestimuleerd het structurele pad op te gaan. Jammer dat we het stadium van structuuroplossing met NMR niet hebben bereikt. Frank, kerel, ook jou ben ik met betrekking tot structurele zaken zeer erkentelijk. Onze elektronmicroscopische samenwerking was zeer vruchtbaar.

En dan mijn paranimfen Jan en Rob. Aan jullie heb ik in mijn labwerk het meest gehad. Jullie adviezen over praktische zaken en discussies over uit te voeren experimenten en resultaten waren onontbeerlijk. Ik waardeer het zeer dat jullie mij tot en met het laatst hebben bijgestaan. Ook de extra handen gedurende de afgelopen jaren, Xavier, Reis en Dennis, bedankt voor jullie inzet.

Met nadruk wil ik Peter bedanken voor de jarenlange wetenschappelijke samenwerking ter 'ontrafeling' van het S_A-eiwit. Ik denk dat ik in een unieke situatie verkeerde met mijn promotor als directe begeleider en ik zal onze lange sessies achter de PC en de uitgebreide discussies niet licht vergeten. Anneke bedankt voor het warme ontvangst in huize Pouwels en goede gesprekken tijdens lunch en diner.

En dan zijn er natuurlijk nog Harald, Aske & Bart, Barry & Ingrid, Jan-Luuk & Marjan en Meinke en voor jullie regelmatige bezoeken, gezamenlijke activiteiten of morele steun.

Jeroen, van met jou 'samenwonen' in het Twijntje, want dat was het bij tijd en wijle, heb ik altijd erg genoten. Vrijdagochtend ontbijten in de stad, een avondborrel op gezette tijden, en vaker wel dan niet samen koken. Als goede vriend zette je me ook weer met beide voeten op aarde wanneer dat nodig was.

Lieve Valérie, jij bent degene die de laatste tijd van het dichtste bij hebt meegemaakt. Ik verbaasde me er vaak over hoe rustig jij altijd bleef als er druk op de ketel kwam. Ik moet zeggen dat het iets is dat ik van je geleerd heb de afgelopen maanden. Door jou waren de laatste loodjes een stuk lichter. Je zult blij zijn met wat rust in de tent nu, met name tot september, want daarna is dat wel weer afgelopen...

ISBN 90-9015872-3