University of Groningen

The proteins non-covalently interact with cell wall in Gram-positive bacteria

Supervised by Prof. Jan.Kok

kexin zhang (S2084775)

7/21/2014

Contents

List of abbrevations

Abbreviation	Full name	Abbreviation	Full name
PGN	peptidoglycan	SCWP	secondary cell wall polymer
TA	teichoic acid	LTA	lipoteichoic acid
PGRP	PGN recognition protein	OM	outer membrane
CHW	cell wall hydrolase	Srt	sortase
iGlu	isoglucose	Dpm	diaminopimelic acid
PBP	penicillin-binding protein	Gro-P	polyglycerol phosphate
Rit-P	poly-ribitol phosphate	Glc-P	poly-glucosyl phophate
Dap	diaminopimelic	CBD	cell wall binding domain
IPG	inter-prong groove	GST	glutathione S-transferase
ChBD	coline binding domain	ChBR	choline binding repeat
hp	hairpin	MTP	muramyl tripeptide
CL	cross-linked	MPP	muramyl pentapeptide
ATP	Adenosine triphosphate	UDP	Uridine diphosphate

Summary

Gram-positive bacteria are encompassed by a phospholipid bilayer and the PGN cell wall that is approximately 200-800 Å thick[1]. This cell wall structure and the negative charge contributed by phospholipids and SCWPs (TA or LTA) distinguishes Gram-positive bacteria from most of eukaryotic cells and makes the bacteria vulnerable to the action ions and other positively charged molecules such as antimicrobial peptides[2]. Some relatively small toxic compounds (20-30 kDa)[3] can easily diffuse through the pores (~20 Å)[4] formed by cross-linked PGN network. The unique composition and distinct negative charge make PGN wall a perfect target for various proteins, including S-layer proteins, autolysins, lysins from bacteriophage as well as PGRPs from eukaryotes.

This essay focuses on the binding/recognition mechanisms of various binding domains. Four binding domains, including SH3b domain from prokaryotes, choline binding domain from phages, protective S-layer proteins from bacteria, and PGRPs from eukaryotes are reviewed in details. To sum up, the binding specificity can be contributed from a series of aspects: stain/species-specific SCWPs, the extra component of SCWPs (choline moiety) and strain/species-specific PGN modifications (unique cross bridge). These binding domains can be applied as effective bactericidal agents. Recombined PGN hydrolase domain with various binding domains can be a new fashion to develop drugs with high specificity and low disruption of commensal microorganisms[14].

Introduction

Bacteria possess a series of characteristics that differentiate them from Archaea and eukaryotes. One of the most predominant features, which has been considered to be unique to bacteria, is the peptidoglycan (PGN) cell wall. This bag-like structure not only provides an external protective shell, but also in a broad range of functions critical to survive, such as cell division, extracellular adhesion, communicating with other cells or environmental factors and preventing cell autolysis from high internal osmotic pressure.

Normally, bacteria are divided into two major sub-groups based on whether they possess a secondary membrane out of the PGN cell wall, the outer membrane (OM)[15]. Gram-negative bacteria are encompassed by an OM, which consists of lipopolysaccharide. An OM is absent in Gram-positive bacteria, which have a thicker PGN cell wall comprised of multiple layers of peptidoglycan interspersed with various secondary cell wall polymers (SCWPs). In both Gram-positive and Gram-negative bacteria the processes of cell growth and division are limited by the PGN wall, or regulated by the surface proteins embedded in the PGN wall[16]. It is well accepted that Gram-positive bacteria have evolved a series of particular mechanisms by which they can immobilize proteins on their PGN cell wall. These mechanisms include either the covalent anchoring to the PGN wall or the non-covalent attachment to either the PGN wall, or to the secondary cell wall polymers, for instance the teichoic acids[17].

The general structure and composition of PGN cell wall are quite conserved. It is conceivable that this general conserved structure and its unique composition make the PGN cell wall an excellent target for various proteins with diversified origins, *e.g.*, the peptidoglycan recognition proteins (PGRPs) from the host immune system (usually eukaryotic cells), the murein hydrolases (CWHs) from bacteriophage, or the endogenous protective proteins like S-layer proteins.

All the proteins interact with the PGN cell wall via either a covalent linkage or through non-covalently attachment. The covalently anchored proteins are linked to the PGN cell wall usually by their conserved C-terminus and transpeptidase, which is called sortase (Srt). This enzyme was first identified in *Staphylococcus aureus* (SrtA), and now its homologs have been identified in numerous Gram-positive bacteria, including *L. monocytogenes*[18], *Bacillus anthracis*[19] and a series of streptococcal species[20-24]. Based on genome sequence analyses, it has been predicted that the number of types of surface proteins anchored by SrtA can vary from 1 or 2 in *Tropberyma whipplei* up to 43 in *L. monocytogenes*[25].

The non-covalently attached proteins interact with PGN wall in several different ways, e.g., some cell wall hydrolases attach to choline-containing teichoic acid (TA) which either anchor to the cytoplasmic membrane or cell wall, some protective proteins like S-layers bind to secondary cell wall polymers (SCWPs), while some of the PGRPs specifically recognize the L-Lys residue on the peptide stem of PGN cell wall. To sum up, all these non-covalently linked proteins show a series of mechanisms to recognize the PGN cell wall specifically.

The main purpose of this essay is to review the current scholarship of the mechanisms of PGN cell wall recognition proteins and their PGN binding motifs. Therefore, the covalently linked proteins and the catalytic properties of non-covalently linked proteins will not be covered. To better understand those binding mechanisms, it is necessary to refresh the structural and metabolic information about PGN wall in Gram-positive bacteria. And the general strategies of binding of protein to PGN wall is discussed in the summary.

The cell wall in Gram-positive bacteria

The architecture and biosynthesis of Peptidoglycan wall

Architecture of PGN wall

Three distinct intracellular compartments can be clearly identified in Gram-positive bacteria: the cytoplasm, cytoplasmic membrane and PGN cell wall[26]. The cell wall is a peptidoglycan macromolecule interspersed with a number of accessory molecules like teichoic acids, polyphosphates, teichuronic acids, or carbohydrates[27]. The PGN wall also functions as a scaffold to anchor capsular polysaccharides[28], cell wall teichoic acids (WTAs)[29] and various proteins either assembled into pili[30], covalently anchored to peptidoglycan[31] or non-covalently attached S-layer structures (Figure.1)[32].

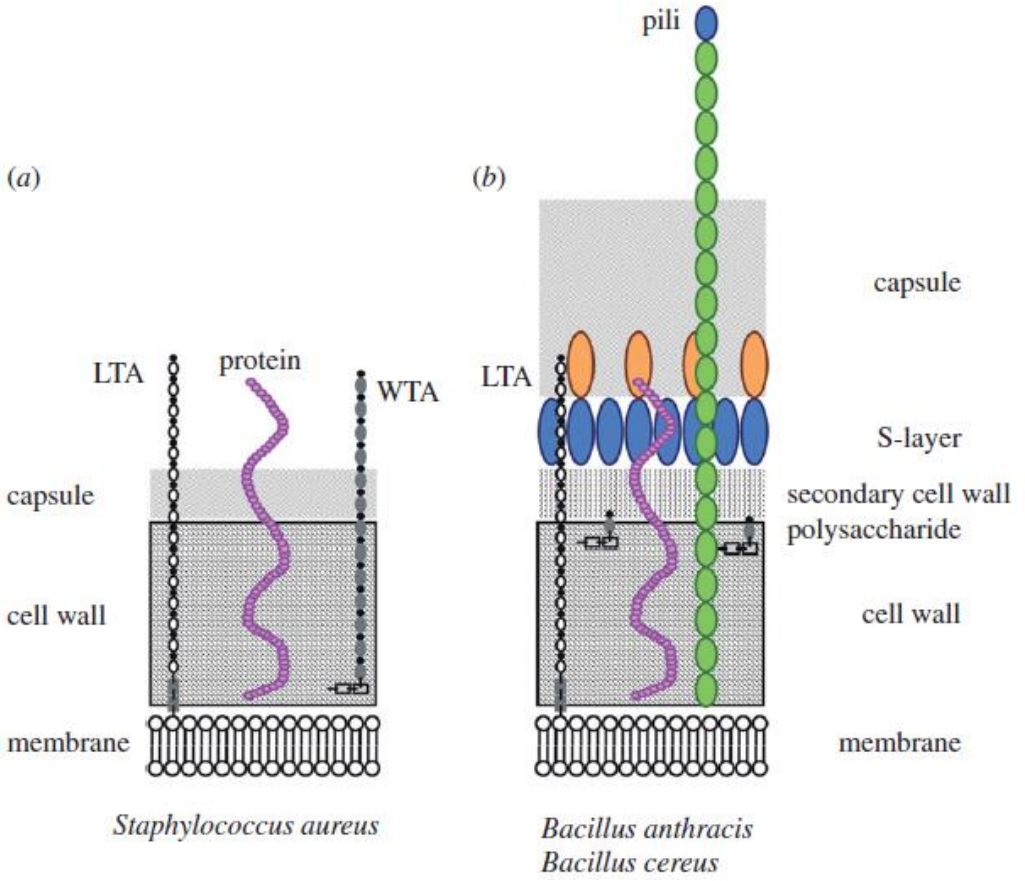


Figure.1 The architecture of the cell envelope in Gram-positive bacteria. (a) Some Gram-positive bacteria, like *Staphylococcus aureus*, use a thick peptidoglycan layer with wall teichoic acids (WTA, or TA), proteins and covered by capsular polysaccharides. The lipoteichoic acids (LTA) are directly linked to the cytoplasmic membrane. (b) Other Gram-positive bacteria, such as *Bacillus anthracis* or *Bacillus cereus* not only have a thick peptidoglycan wall, but also synthetize a secondary cell wall polysaccharide (SCWPs) which encompasses the PGN wall. The SCWPs are covalently linked to peptidoglycan and can serve as attachment sites for the SLH domains of S-layer proteins. (Adapted from Ref.35)

The glycan strands of peptidoglycan consist of repeating N-acetylmuramic acid-(β1-4)-N-acetylglucosamine (MurNAc-GlcNAc)[34] (Figure.2). Their length can vary from 5 to 30 subunits depending on the bacterial species[35, 36]. Usually the D-lactyl moiety of every MurNAc is amide linked to a short pentapeptide, which is called wall peptides[37]. The wall peptides or peptide stems, which are conserved but for a few exceptions, are usually comprised of (L-Ala)-(D-iGlu)-(L-Lys) or (Dpm)-(D-Ala)-(D-Ala). The wall peptides are cross-linked with other wall peptides via a cross-bridge that is variable among species[38-40]. In such a way the glycan strands form a three-dimensional structure that encompasses the whole cell and provides the desired functions.

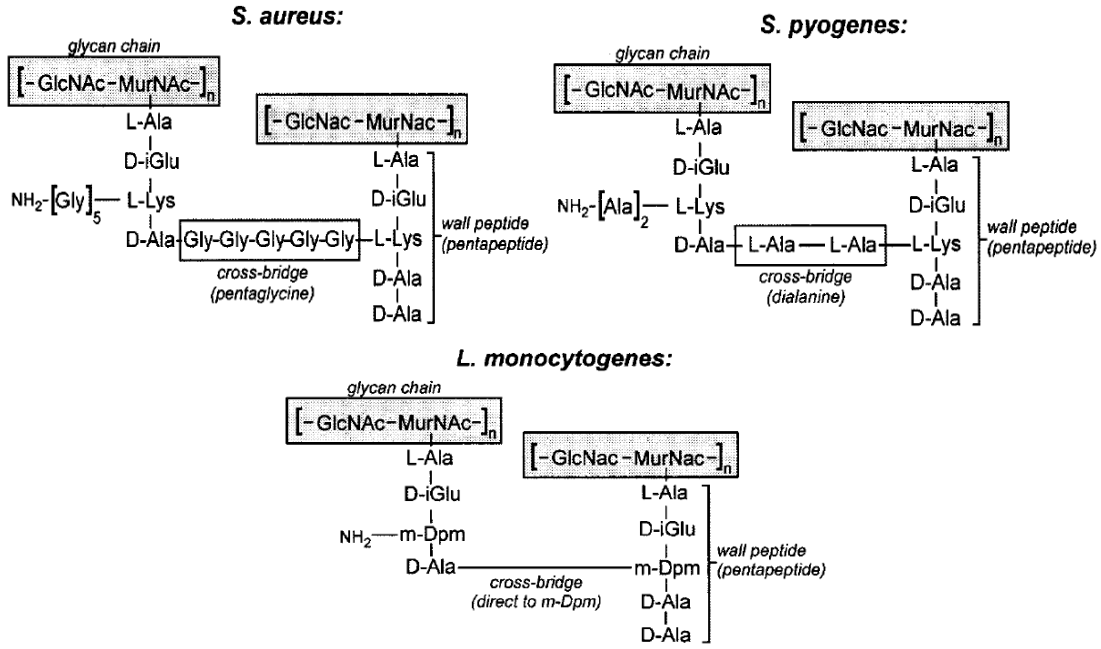


Figure.2 Scheme of the peptidoglycan architectures from *S. aureus*, *S. pyogenes*, and *L. monocytogenes*. These three bacterial species share the same subunit of main glycan strands, which is the repeating disaccharide, MurNAc-GlcNAc. Also, there is little variation in the wall peptide (pentapeptide), with only the L-lys, in *S.aureus* and *S.pyogenes*, is replaced by m-dDpm in that of *L.monocytogenes*. However, the cross-bridge structures connecting neighbouring wall peptides are quite various: the cross-bridge consists of pentaglycine in *S.aureus* and dialanine in *S.pyogenes*. Adjacent wall peptides *in L.monocytogenes* are linked via a direct amide linkage instead of a polypeptide cross-bridge. (Adapted from Ref17)

The biosynthesis pathway of PGN wall

The synthesis of the PGN wall in Gram-positive bacteria can be divided into three distinguishing steps that occur in three distinct compartments: the cytoplasm, the membrane and the peptidoglycan wall[41, 42] (Figure 3). All Gram-positive bacteria share a nearly identical process of cell wall synthesis, thus in this essay *S.aureus* is taken as an example to illustrate cell wall synthesis. The process is initiated by the production of UDP-MurNAc from UDP-GlcNAc and phosphoenolpyruvate in the cytoplasm[43], followed by the consecutive addtion of five amino acids in four steps, the wall peptide, or peptide stem. They are L-Ala, D-isoGlu, L-Lys and D-Ala–D-Ala dipeptide, respectively[44, 45]. Every addtion to the wall peptide consumes one high-energy phosphate, i.e., ATP. The final product of the pathway in the cytoplasm is UDP-MurNAc-pentapeptide, which is also named Park's nucleotide. The Park's nucleotide is anchored to the cytoplasmic membrane via a phosphodiester linkage to form lipid I at the expense of UDP. The full structure of lipid I is C_{55}-PP-MurNAc–LAla–D-isoGlu–L-Lys–D-Ala–D-Ala[46]. A molecule of GlcNAc is linked to MurNAc to generate the lipid II precursor, C_{55}-PP-MurNAc(-L-Ala–DisoGlu–L-Lys–(Gly_5)–D-Ala–D-Ala)-GlcNAc[47, 48]. Lipid II is further modified by the addtion of amino acids onto the ε-amino of lysine[49]. In the case of *S.aureus*, as mentioned above, five glycines are added on the wall peptide, which is called the cross-bridge. However, the compositions and the length of the

corss-bridges vary a lot across species, which makes them a perfect target by which various proteins can discriminate PGN walls. The lipid II precursor is transported across the cytoplasmic membrane and ready as the subunit in peptidoglycan wall assemble. This third stage of PGN wall synthesis, which occurs in the cell wall itself, is the assembly of peptidoglycan macromolecules from their disaccharide subunits. This process is catalyzed by the penicillin-binding proteins (PBPs)[50]. Based on sequence similarities, the PBPs can be sorted into two sub-classes[51]. The class A enzymes catalyze both the ploymerization of peptidoglycan strands from the disaccharide subunits and the transpeptidations between the wall peptides[52]. The first reaction continuously adds the MurNAc to the GlcNAc moiety in another disaccharide subunit, whereas the later reaction links the pentaglycine to the fourth D-alanine in another wall peptide and removes the fifth D-alanine in the C-terminal[53]. Very little is known about class B PBPs (Figure 3). Not all peptidoglycans have the cross-bridge structure. The ratio between cross-linked and free wall peptides may play an important role in shaping the 3-D structure of the whole cell. Peptidoglycan with a low degree of cross-linkage is more sensitive to degradation by cell wall hydrolases[54].

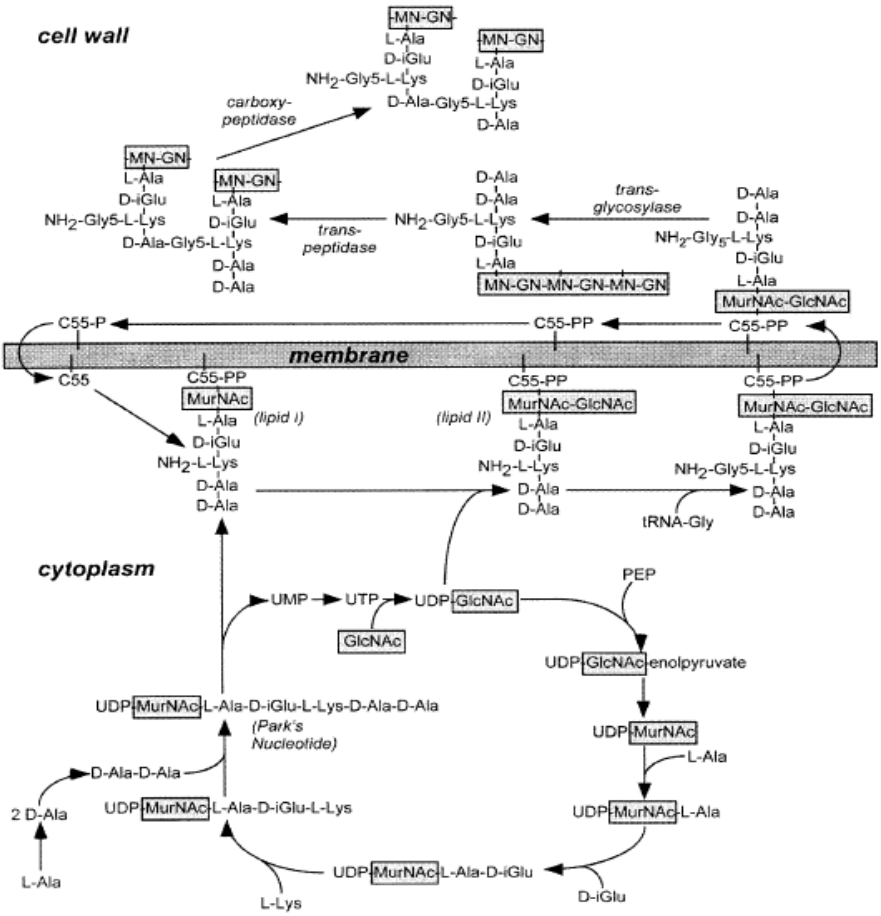


Figure 3 Peptidoglycan cell wall assembly in *S.aureus*. As is mentioned in the text, PGN wall synthesis initiates in the cytoplasm, and leading to the end product UDP-MurNAc-pentapeptides, i.e., Park's nucleotide. The Park's nucleotide is anchored to the cytoplasmic membrane and linked with GlcNAc to generate lipid II precursor. After the addition of pentaglycine to the wall peptide, the lipid II precursor is translocated across the membrane and assembled into peptidoglycan strands by PBPs (Adopted from Ref17).

Accessory components in Gram-positive cell wall

Bacteria usually contain a series of polysaccharides in their cell wall, the major part of which forms the peptidoglycan. As established in the 1980s', Gram-positive bacterial cell wall contains a series of accessory components. These components can be classified into three types based on their structural characteristics: (1) teichoic acids; (2) teichuronic acids, (3) other neutral or acidic polysaccharide modifications. These accessory components are also termed "secondary cell wall polymers" (SCWPs) due to the secondary roles they play in the cell wall. The former two types of SCWPs are also called "classical SCWPs" and the third typfied as a "non-classical SCWP". Here we briefly review the structural and functional features of teichoic acids and teichuronic acids as they are involved in some cell wall targeting or protein binding recognitions[55].

Classical SCWPs: Chemically, teichoic acids are anionic polymers that consist of polyglycerol phosphate (Gro-P), poly-ribitol phosphate (Rit-P), or poly-glucosyl phosphate (Glc-P). Their negative charge can be attributed to either phosphate (in teichoic acids) or carboxyl groups (in teichuronic acids). These subunits, including Gro-P, Rit-P and Glc-P, may connect to each other by amino acid esterification or glycosylation, and covalently link to peptidoglycan[56]. Although the subunits are identical throughout the species, the overall macromolecular structures of these SCWPs are highly diverse[57]. Divers teichoic acid have been documented, whereas their structural features are poorly understood and chemical analyses of this type of SCWP have been documented[58]. Furthermore, the physiological role of the teichoic acids is not completely understood. It is speculated that these anionic cell wall decorations are responsible for capturing divalent cations or forming an electro-barrier to prevent the diffusion of cations. They also seem to serve as binding sites for some proteins (details see below)[59 60].

Non-classical SCWPs: During the investigating of the cell wall in *Bacillaceae*, some novel "non-classical SCWPs" have emerged. Based on their compositions and structural data, these non-classical SCWPs can be classified into three distinct groups[61] (Figure.4).

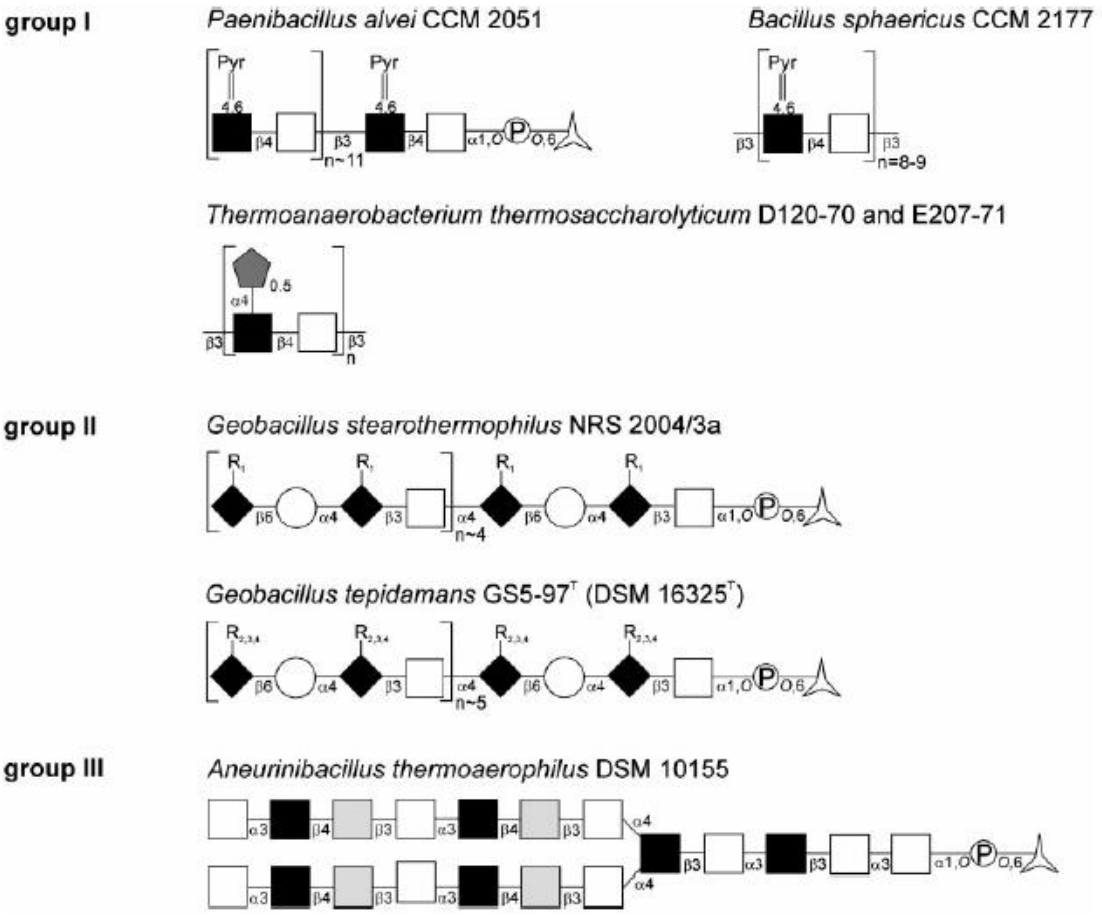


Figure.4 Three distinct groups of non-classical SWCPs. ■, N-acetylmannosamine. ○, D-glucose. □, N-acetylglucosamine N-acetylmuramic acid ◆ 2, 3-di-Nacetylmannosaminuronic acid. ⬟ D-ribofuranose. R1=COOH, R2=$CONH_2$, R3=$CONHCOCH_3$, R4=$CON-(COCH_3)_2$ (Adopted from Ref61).

Group 1. The main glycan chain of this group consists of a repeating dissacharide formed by N-acetylglucosamine and N-acetylmannosamine (Figure.4 group I). These two subunits are linked via alternating β(1→3) and β(1→4) linkages. Either a ribofuranose or a pyruvic acid residue is linked to N-acetylmannosamine, conferring the overall anionic characteristic to the whole glycan chain. The number of repeating dissacharide subunit varies a lot with the species investigated.

Group 2. In the SCWPs of group 2, the subunits are tetrasaccharides comprised of N-acetylglucosamine, D-glucose and two N-acetylmannosaminuronic acids. In each subunit, D-gulcose and N-acetylglucosamine alternatively link to N-acetylmannosaminuronic acid (Figure.4 group II). Usually this subunit repeats six times and ends with a phosphoryl N-acetylmuramic acid. Additional groups are add to every N-acetylmannosaminuronic acid, including COOH, $CONH_2$, $CONHCOCH_3$ or $CON(COCH_3)_2$,

Group 3. The charge-neutral SCWP isolated from *Aneurinibacillus thermoaerophilus* DSM 10155 represent an unusual structure. These non-classical SCWPs consist of N-acetylmannosamine, N-acetylglucosamine, and N-acetylgalactosamine (Figure.4 group III). However, regular repeating subunits have not been observed in this

SCWPs. Only two common features can be summarized: (i) they all have biantennary oligosaccharide structures, and (ii) all the oligosaccharide chains are chemically highly homogeneous.

Differences of peptidoglycan between Gram-positive and Gram-negative bacteria

The composition of peptidoglycan in Gram-positive bacteria is slightly but critically different from that in the Gram-negative bacteria. Figure.5 shows a comparison of the structure between the PGN in Gram-positive and Gram-negative cells. In Gram-positive bacteria, the residue in position three of the wall peptide is L-Lys (Lys-type PGN), whereas in Gram-negative bacteria this position is occupied by m-diaminopimelic acid (Dap-type PGN). Moreover, the wall peptides in Lys-type PGN are usually interconnected by cross-bridges, in which the number and composition of amino acids vary in different species, while the wall peptides in Dap-type PGN are directly cross-linked. These differences result in distinct immune response pathways in insects and mammals[62].

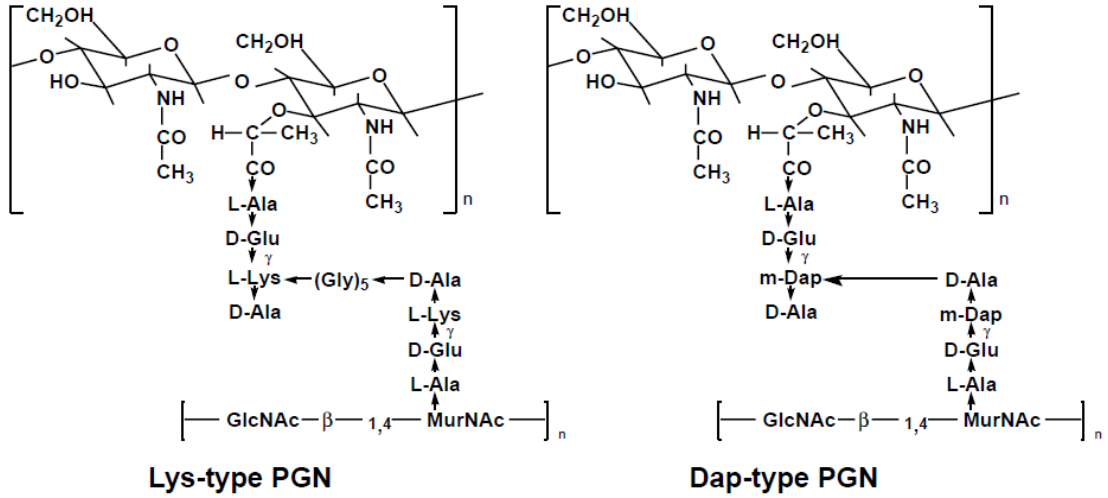


Figure.5 Comparison of the peptidoglycan structure of Gram-positive (Lys-type PGN) and Gram-negative (Dap-type PGN) bacteria. Lys-type PGN: The D-Ala in position 4 of the wall peptides cross-linked with L-Lys in position 3 of another peptide stem via a cross bridge. Dap-type PGN: the D-Ala in position 4 is directly linked with m-Dap in position 3 of another peptide stem (Adopted from Ref137).

Proteins that non-covalently link to cell wall

As detailed above, the bacterial PGN cell wall possesses a unique composition that makes it a perfect target for recognition or attachment of a diversity of proteins. These proteins recognize the PGN structures via various binding motifs. In many cases, for instance in the cell wall hydrolase lysin, their cell wall binding domains (CBDs) are necessary to perform the cytolysis. Numerous studies have been done to elucidate the mechanism of binding and the structures of these binding motifs. Up to date, a series of CBDs have been discovered. See below.

Motif	Example	Ligands	3D-structures	reviewed
GW domain	In1B	SCWPs	Y	N
WxL domain	CscA	unknown	N	N
LysM domain	MurA	PGN	Y	N
choline binding	LytA	choline(TA)	Y	Y
S-layer domain	SLH motif	PGN	Y	Y
SH3b domain	ALE-1	PGN	Y	Y
PGRPs	PGLYRP-3	PGN	Y	Y
PlyPSA CBD	PlyPSA	SCWPs	Y	N

Table.1 Cell wall binding motifs from various proteins. Their ligands and corresponding strains are listed.

The GW domain binds to SCWPs in Gram-positive bacteria and is presented in a tandem form. The binding strength is related to the number of domains[63]. Interestingly, the GW domain is structurally related to the SH3 domain in eukaryotes that is involved in signal transduction and binds to proline-rich sequence[64]. In InlB, a GW domain-containing protein from *L.monocytogenes*, the GW domain consists of seven β-sheet strands, five of which form a barrel construction similar to that in SH3 domain[65].

The C-terminal WxL domain was first predicted in *Lactobacillus plantarum* by using global genome analyses[66] and was then identified in *Enterococcus faecalis*[67]. This domain with a highly conserved Trp-x-Leu signature is involved in *E.faecalis* cell wall localization by directly binding with PGN[68]. However, the biological functions of WxL domain is not completely understood yet. Usually the WxL domain-carrying proteins are exported and only present in low-G+C content Gram-positive bacteria, which may indicate that these proteins play a role in intercellular aggregation[69].

The LysM domain is a cell wall binding domain with approximately 40 amino acid residues first described in muramidase 2 from *E.hirae*[70]. Up to date, more than 1500 proteins containing the LysM domain have been documented, including proteins from Gram-negative bacteria and eukaryotes[71]. The LysM domain is resided to one end of the protein and the number of it is variable. Similar to GW domain, the binding strength is regulated by the number of binding domains[72].

In this essay, four types of binding domains with different origin are described in details. They are the SLH domian in S-layer proteins (bacteria), SH3b domain (bacteria), peptidoglycan recognition proteins (eukaryotes) and choline binding domains (bacteriophage) are described in details in this essay. LysM domain is not included because it has been frequently reviewed.

SH3-like domain in lysostaphin

Lysostaphin is a peptidoglycan hydrolase produced by *Staphylococcus simulans* and *staphylolyticus*, which specifically hydrolyzes the cross-bridge construction in *Staphylococcus aureus* cell wall[73]. ALE-1 is a lysostaphin homologue secreted by *Staphylococcus capitis* EPK1[74]. This enzyme, like other lysostaphins, specifically recognizes PGN of *S.aureus,* has high potential in the treatment of infections caused by antibiotic-resistant *S.aureus*[75]. Previous studies have shown that ALE-1 consists of three main domains: the 13-amino-acids repeat domain in the N-terminus, the 362 amino acids central domain, which contains a zinc ion, and the cell wall binding domain named 92AA in C-terminus[76]. The N-terminal repeat domain is removed during maturation of the preprotein. The central domain possesses glycylglycine endopeptidase activity. 92AA provides high affinity for ALE-1 binding to PGN of *S.aureus*[77]. Interestingly, the CBD in ALE-1, i.e., 92AA, shows significant similarity with eukaryotic Src homology 3 (SH3) domain in tertiary structure[78]. The SH3 domain is involved in protein-protein interactions in signal transduction pathways of eukaryotes and specifically binds to Proline-rich motifs (PXXP)[79]. The 92AA domain belongs to SH3b, the recently discovered bacterial homologs of SH3, which are thought to mediate protein binding to PGN wall[80].

Binding function determination of the SH3b

To confirm if the 92AA domain is responsible for binding to the PGN wall of *S.aureus,* several truncated forms of ALE-1 were constructed (Figure.6 A) and their binding activities are tested by using SDS-treated *S.aureus* 930P cells (Figure.6 B)[81].

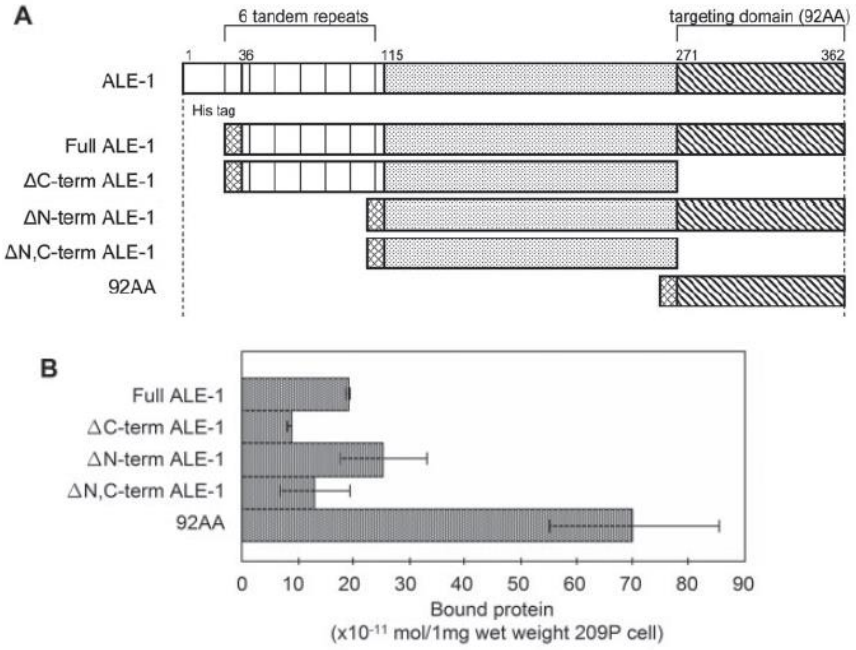


Figure.6 (A) Constructions of natural ALE-1 and its truncated forms. (A) Natural ALE-1 consists of a signal peptide in the N-terminus (residues 1-35), followed by six tandem repeats (residues 36-114), a central part

(residues 115-270) and binding domain 92AA (residues 271-362) in C-terminus. His-tag ALE-1 without signal peptide is used as full ALE-1. Deleting 92AA, six tandem repeats and both resulted in ΔC-term ALE-1, ΔN-term ALE-1 and ΔN, C-term ALE-1, respectively. The single domain, 92AA, is also used to test its binding activity. All these truncated forms contain a His-tag in the N-terminal. (B) The binding abilities of ALE-1 and its truncated forms to SDS-treated *S.aureus* 209P cells (Adapted from Ref81).

According to expectation, the truncated forms of ALE-1 lacking the 92AA domain (ΔC-term ALE-1and ΔN, C-term ALE-1) show a significant decrease in binding activity. Compared to full ALE-1, the ΔN-term ALE-1 shows a similar but higher binding activity. The 92AA domain alone possesses the highest binding activity, which is three times more than that of full ALE-1. It is reasonable to conclude that 92AA provides most of the binding activity of the ALE-1 protein to *S.aureus* 920P cells and cutting down in molecular mass leads to a higher binding capacity (Figure.6).

Determining the binding specificity of the SH3b domain

As mentioned above, the SWCPs, e.g., TA and LTA could be the anchor of various proteins. To test if 92AA domain binds to those SCWPs, *S.aureus* cells were treated with hydrogen fluoride (HF) to block TA binding before incubating with ALE-1. HF treatment breaks the linkage between TA and peptidoglycan. However, the binding process was not influenced by HF-treatment, which excluded TA as the binding ligand. On the other hand, to test the possible interactions between 92AA and PGN wall, *S.aureus* cells were treated by heating, SDS-heating, and trypsinized SDS-heating respectively. These pretreatments increase the linearization of PGN in *S.aureus* cells (Figure.7).

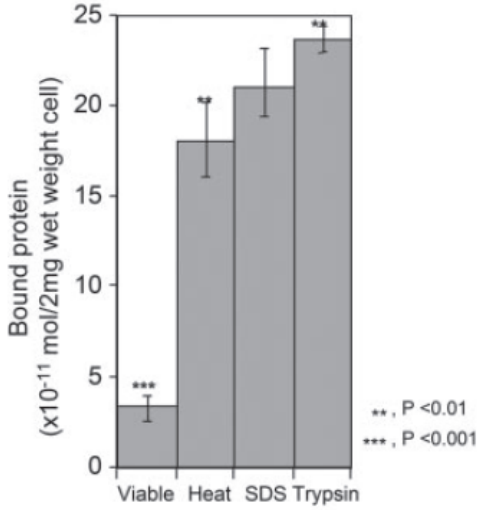


Figure.7 Binding amount of 92AA to *S.aureus* cells with different pretreatments. The cells with bound proteins were treated by SDS and separated by SDS-PAGE. The amount of 92AA was estimated by software NIH image version 1.52 (Adopted from Ref81).

As is shown in Figure.7, the amount of bound protein increases with the increase in a linear fashion, which confirms that binding of 92AA domain to *S.aureus* cells is directly mediated by PGN. Interestingly, previous studies have proved that ALE-1, or lysostaphin, specifically lyses *S.aureus* cells and does not affect lysostaphin producing strains, e.g., *S.simulans*[82]. Experiments proved that the binding ability of 92AA is significantly suppressed when the protein is incubated with PGNs from *Streptococcus mutans*, *Bacillus megaterium*, *Staphylococcus.capitis*, *Micrococcus lysodeikticus* and *Lactobacillus plantarum* species, in which the composition and length of the cross bridge differ from those in *S.aureus*. (See Table 2)

Strains	Cross-bridge Construction
S. mutans	L-Ala-L-Ala
M. lysodeikticus	L-Ala-(D-Glu/Gly)-L-Lys-D-Ala
B. megaterium	NONE
L. plantarum	NONE
S. aureus	Gly-Gly-Gly-Gly-Gly
S. capitis	Gly-Gly-L-Ser-Gly-Gly

Table 2 Construction of the PGN cross-bridge in different bacterial species. The 92AA domain only binds to pentaglycine in *S.aureus*. Binding processes in other strains are significantly suppressed. Of note, the *S.capitis* is an ALE-producing strain.

Furthermore, sequence alignments among ALE-1 and other cell wall binding domains reveal that all *S.aureus*-specific autolysins and *S.aureus*-specific amidases share a highly conserved N-terminal sequence. (See Figure.8)

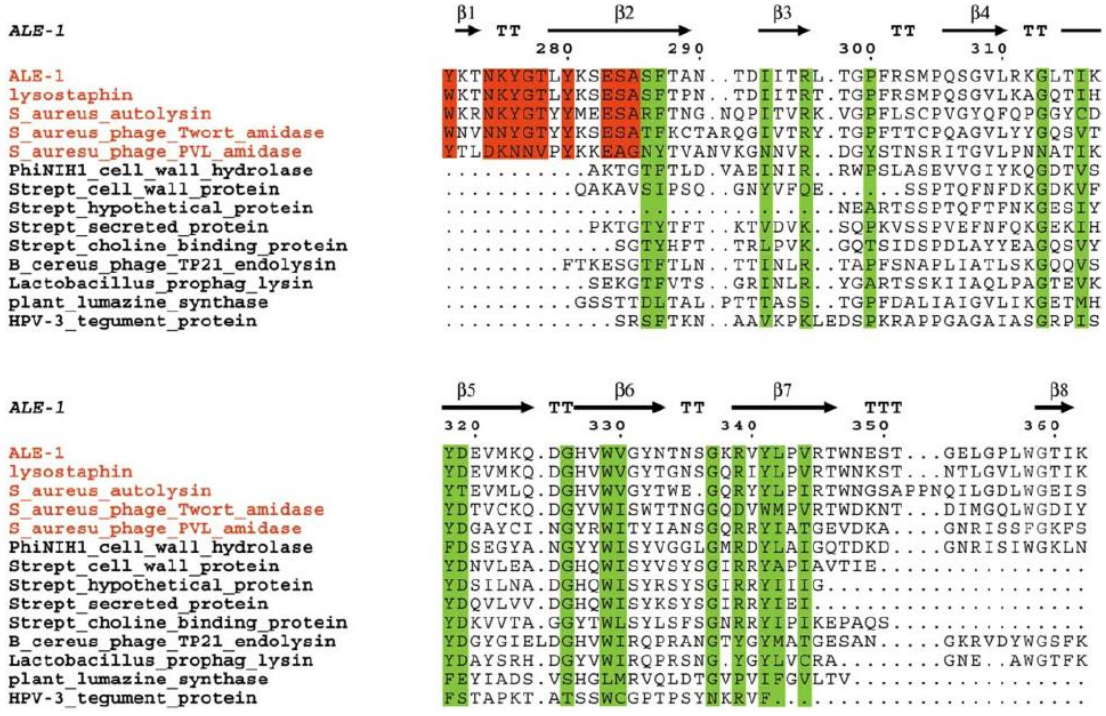


Figure.8 Sequence alignment of cell wall binding domains from different species. Conserved residues in *S.aureus* specific binding domains are depicted by red, the residues conserved in all binding domains are colored by green.

The motifs and its corresponding secondary structures are shown above the sequence (Adapted from Ref81).

Thus, based on the sequence alignments in Figure.8, the conserved residues in the N-terminal of *the S.aureus*-specific binding domains (colored red) seem to be responsible for discriminating the pentaglycine cross bridge from those with other composition.

What needs to be noted is that the ALE-1 producing *S.capitis* contains several genes responsible for cross bridge modification, e.g., *epr*[83] and *lif*[84]. These enzymes alter the composition of the cross bridge and confer resistance to ALE-1[85]. Additionally, FemA and FemB are two peptidyltransferases responsible for extending the cross bridge in *S.aureus* cell wall [86]. FemA inserts the second and third glycines, while FemB adds the fourth and fifth glycines.[87] The absence of FemA or FemB results in a shorter cross-bridge. To test how the composition and length of the cross bridge affects the ALE-1 binding specificity, a series of staphylococcal PGNs with different cross-bridge is used to bind with 92AA and its N-terminal truncated form, the 83-amino-acid domain (See Figure 9).

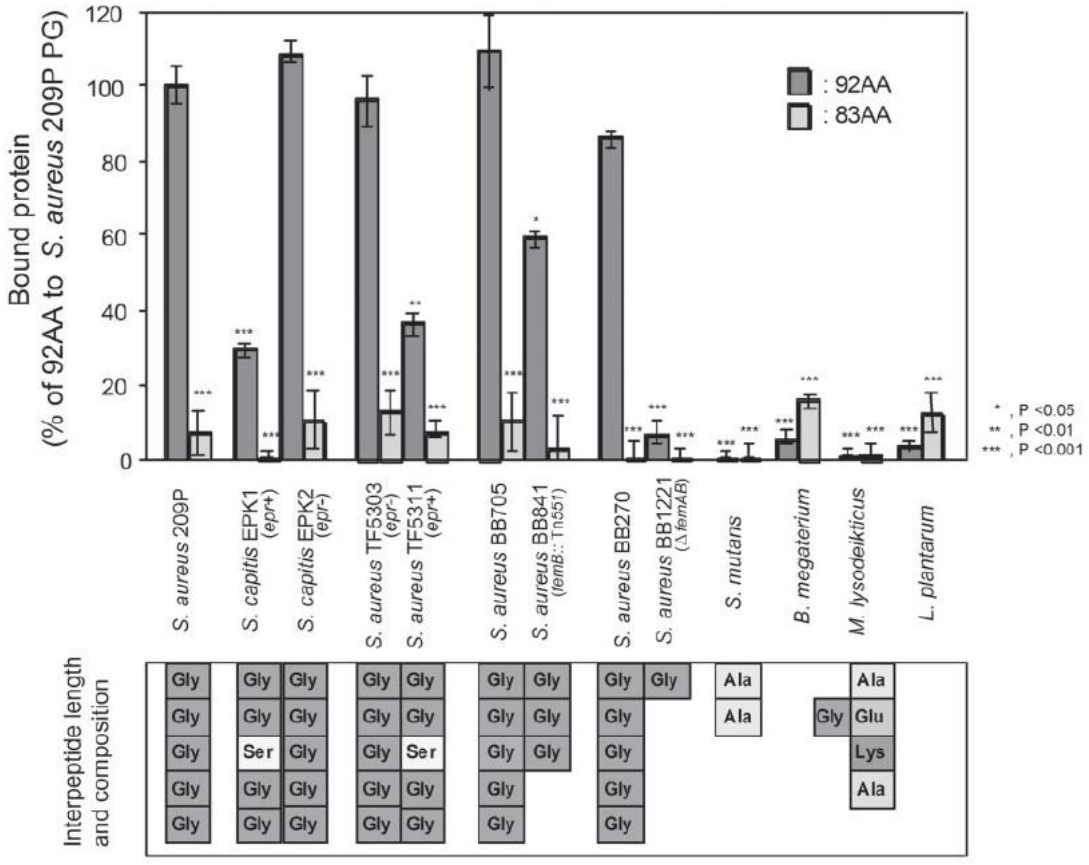


Figure 9 The number of staphylococcal PGNs with different cross-bridge constructions bind with the full ALE-1 binding domain (92AA) and its truncated form 83AA. The constructions of cross bridge are displayed just below the strains name. *S.capitis* EPK2 is an isogenic strain of EPK1 in which *epr* gene is deleted thus the pentaglycine cross bridge is retained. The cross bridge in *S.aureus* BB841 is triglycine due to the mutation of *femB* gene. The *femAB* mutant strain *S.aureus* BB1221 only contain one glycine in its cross bridge. (Adopted from Ref81)

As is shown in Figure 9, the 92AA domain binds very well to the PGNs with pentaglycine cross bridge, i.e., *S.aureus* 209P, *S.aureus* BB270, the *epr* mutant of *S.capitis* EPK2 and *S.aureus* BB705. Once pentaglycine is replaced by Gly_2-Ser-Gly_2, as occurred in *S.capitis* EPK2 and *S.aureus* TF5311, the binding process is significantly inhibited. The length of the cross bridge also plays a critical role in the binding process. The amount of bound 92AA in *ΔfemB S.aureus* BB841 decreases almost half and 92AA barely binds to *ΔfemAB S.aureus* BB1221 cells. The 83AA proteins barely binds to these PGN variants. These findings suggest that the conserved residues in the N-terminal of 92AA domain are the key part in discriminating and binding to PGNs, while the length and composition in pentaglycine cross bridge are also necessary elements in determining the binding with 92AA.

The 3-D structures of 92AA domain in ALE-1

Crystal structures show that the cell wall binding domain in ALE-1 (92AA) is an asymmetric unit comprised of two similar moieties, molecule A and molecule B. the Root-Mean-Square deviation between A and B is only 0.43 Å for all main chain atoms. Thus, molecule B is used for further discussion.

As is shown in Figure 10, 92AA domain contains 8 β-sheet strands and displays a significant similarity with SH3 (PDB code: 1CKA) in tertiary structure[88]. However, their considerable low sequence identity (<20%) may be the indication of their distinguished functions.

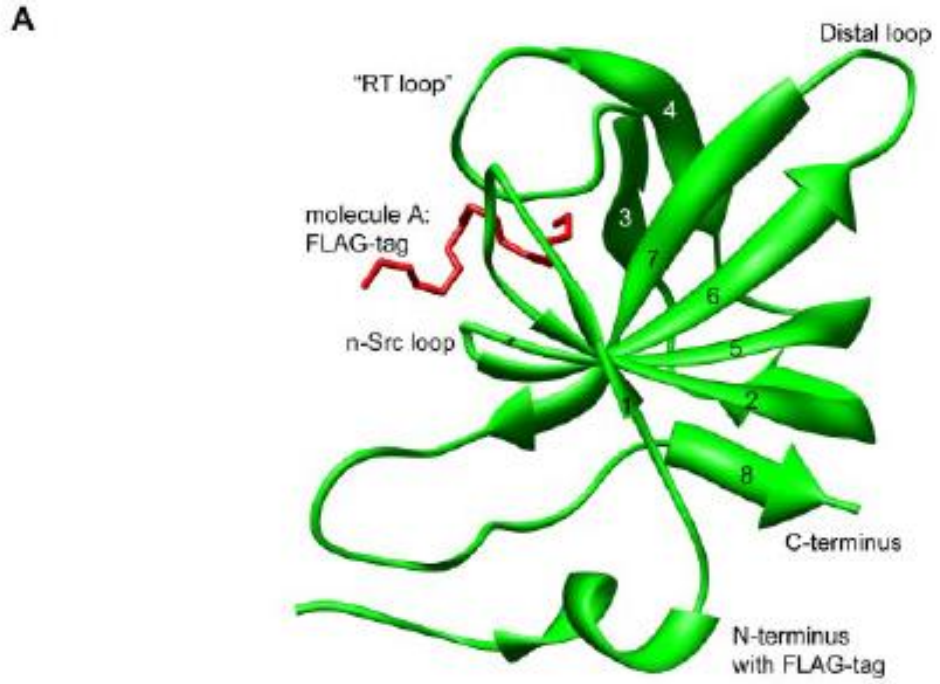


Figure.10 (A)

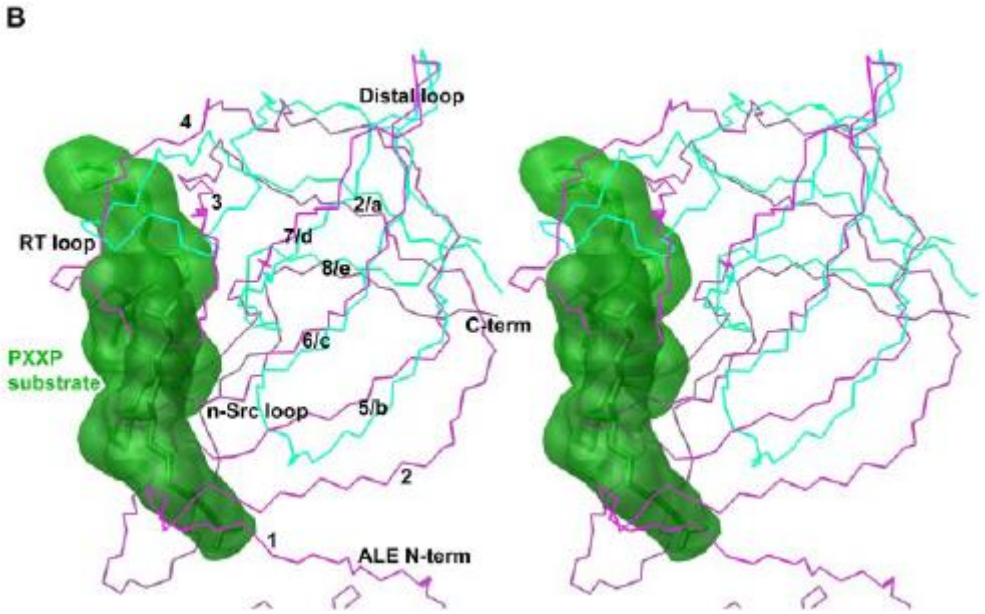


Figure 10 (A) Ribbon diagram of 92AA (molecule B). β-sheet strands are numbered from 1-8. The loops are named after their counterparts in Crk SH3 (details see content). The N-terminus of FLAG-tag from molecule A is shown in red. (B) The superposition of tertiary structures of Crk SH3 (cyan) and 92AA (purple). The β-sheet strands in Crk SH3 are marked by alphabetical order. Every β-sheet strand in Crk SH3 has its counterpart in 92AA of ALE-1. The PXXP substrate of Crk SH3 is shown in stick model with surface rendering green (Adopted from Ref81).

Crk SH3 possesses a β-barrel that consists of five β-sheets (βa to βe). Loops in between are defined as the RT loop (between βa and βb), the n-Src loop (between βb and βc) and the distal loop (between βc and βd). In 92AA, all these five β-sheets and loops are preserved (Figure 10) and three additional β-sheets are found, which are β1, β3 and β4. Compared with Crk SH3, the RT loop in 92AA is much shorter and flanked by β3 and β4 strands. These two strands are not present in Crk SH3. Together with β3 and β4, the RT loop blocks a large area of molecular surface, which is involved in substrate binding by Crk SH3. Furthermore, a tryptophan and proline pair, which are strictly conserved and solvent exposed in SH3[89], are completely buried in 92AA. All these findings support the idea that the biological functions of Crk SH3 and 92AA may be quite distinct.

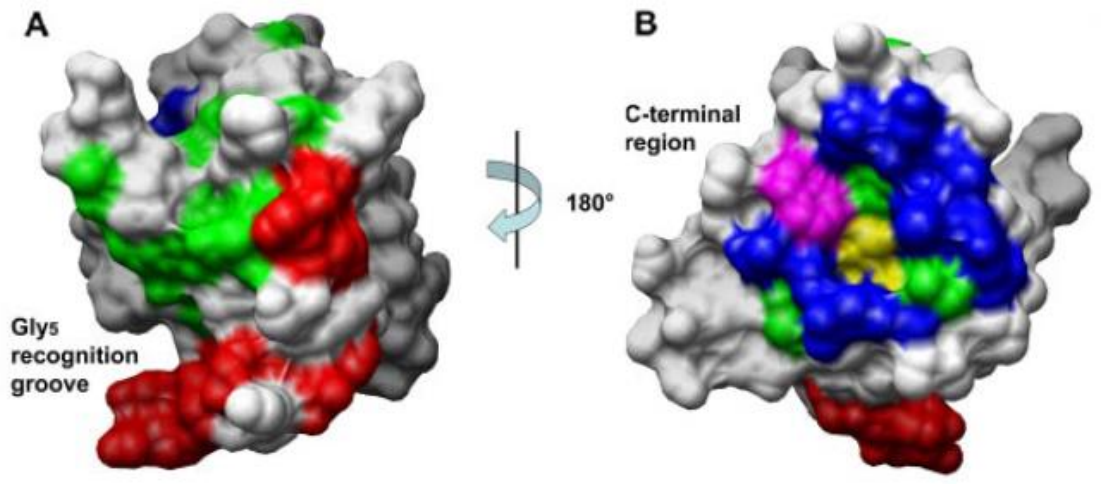


Figure 11 two orientations of the molecular surface of 92AA. The blue residues are those involved in FLAG-tag interactions. The yellow residues are Trp329 and Pro343, which are conserved and solvent exposed in SH3, are completely buried in 92AA. The residues conserved in N-terminal of *S.aureus* specific enzymes are colored by red. These residues form a narrow groove which seems to be responsible for recognition pentaglycine cross bridge in *S.aureus* PGN (Adopted from Ref81).

Figure 11 shows the crystal structure of molecule B[93]. The conserved motif in the N-terminal forms a narrow groove, which may only accommodate pentaglycine. Other cross-bridge construction, even Gly_2-Ser-Gly_2, may disrupt the binding process, as is seen in Figure 9.

Putative binding mode between 92AA and pentaglycine of PGN

To date, the crystal structure of the complex of 92AA bound with pentaglycine has not been obtained yet, and the amino acid residues involved in binding have not been identified. Thus, a putative binding mode was created by a series of the computational methods. Using the Alpha Site Finder function of the MOE program, Hideki Hrakawa et al. [90] identified five putative binding clefts. One of these locates in a narrow groove formed by the β1 and β3 strands, which have been proven to provide functions of discriminating and binding pentaglycines.

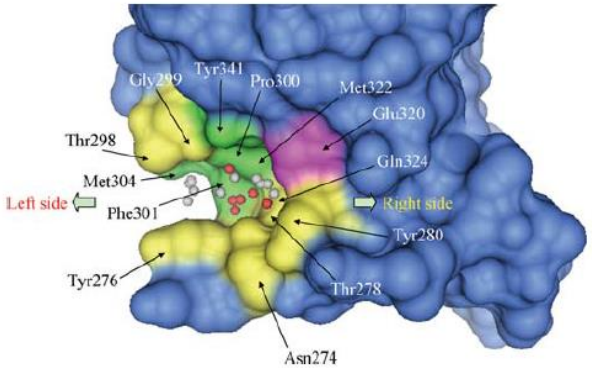


Figure 12 Putative binding groove and its component residues. Hydrophilic and hydrophobic residues are colored

in red and gray. The polar and non-polar residues are depicted in yellow and green. The negatively charged residue Glu320 is marked in magenta. The left and right sides of the groove are also defined (Adapted from Ref90).

Chemical character	Residue positions
polar	Asn274,Tyr276,Thr278,Tyr280,Thr298,Gln324
non-polar	Gly299,Pro300,Phe301,Met304,Met322,Tyr341
negative charged	Glu320

Table 3 the chemical characters of residues in putative binding groove.

As is shown in Figure 12, this narrow groove is comprised of 13 amino acid residues that consist of three types: polar residues, non-polar residues and negatively charged residues. The members of each type have been listed in Table 3. One side of the groove, which is defined by Thr298 and Tyr276, is defined as the left side, the other side is labeled as the right side (See Figure 12).

The 3D structure of pentaglycine is generated by the Edit/Build/Protein commands in the MOE program. Thus, 6561 structures are obtained. Energetically unstable forms (>100kcal/mol) or pentaglycines with nonlinear shape (distance between N-terminus and C-terminus <11Å) are excluded, thus 3070 pentaglycines are ultimately fitted into the putative binding groove.

After energy minimization, 82 pentaglycine-92AA complexes are obtained. Interestingly, both sides of the groove seem to be capable of accommodating pentaglycine. Within these 82 complexes, 21 orient the pentaglycines to the right side, while the other 61 complexes bind pentaglycines in the left side. U_{dock} analysis showed that the average U_{dock} in complexes with pentaglycines bound to the left side is -43.6kcal/mol (S.D. 10.598); the average U_{dock} in complexes bound to the right side is -68.8kcal/mol (S.D. 9.178). This energetic analysis indicates that binding to the right side of the groove leads to an energetically more stable complex. Based on this binding simulation, seven putative residues involved in interactions with these 21 pentaglycines are predicted: Asn274, Tyr276, Thr278, Thr298, Ser303, Glu320 and Tyr 341 (See Figure.13).

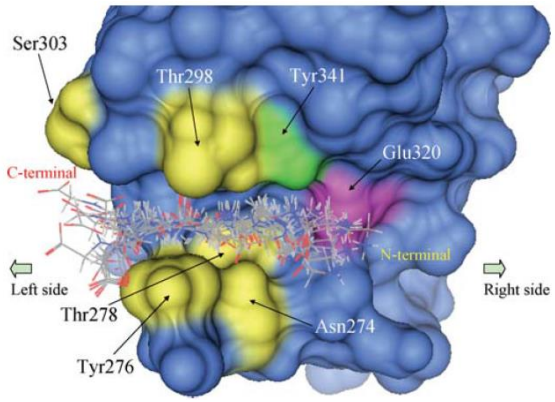


Figure.13 The 21 pentaglycines are modeled to the putative binding groove in 92AA. The pentaglycines are shown in lines. The putative residues interacting with pentaglycines are labelled. The color scheme is the same as in Figure 12 (Adapted from Ref90).

The lowest U_{dock} in these 21 complexes (-83.4 kcal/mol), which is the energetically most stable complex, is used to verify the interactions between 92AA and pentaglycine by applying MD simulation (Molecular Dynamics). Based on MD simulations, five residues are identified directly to interact with the pentaglycine: Asn 274, Tyr276, Thr278, Thr298 and Glu320. All five residues are included in the candidate list given by the binding simulation.

As mentioned previously, the putative binding groove consists of 13 amino acid residues. To estimate if other residues contribute to the binding activity, mutation on each side are introduced by site-directed mutagenesis (See Figure 14). Additionally, four mutations around the groove are also constructed as a controls (N272A, T273A, K275A, S303A).

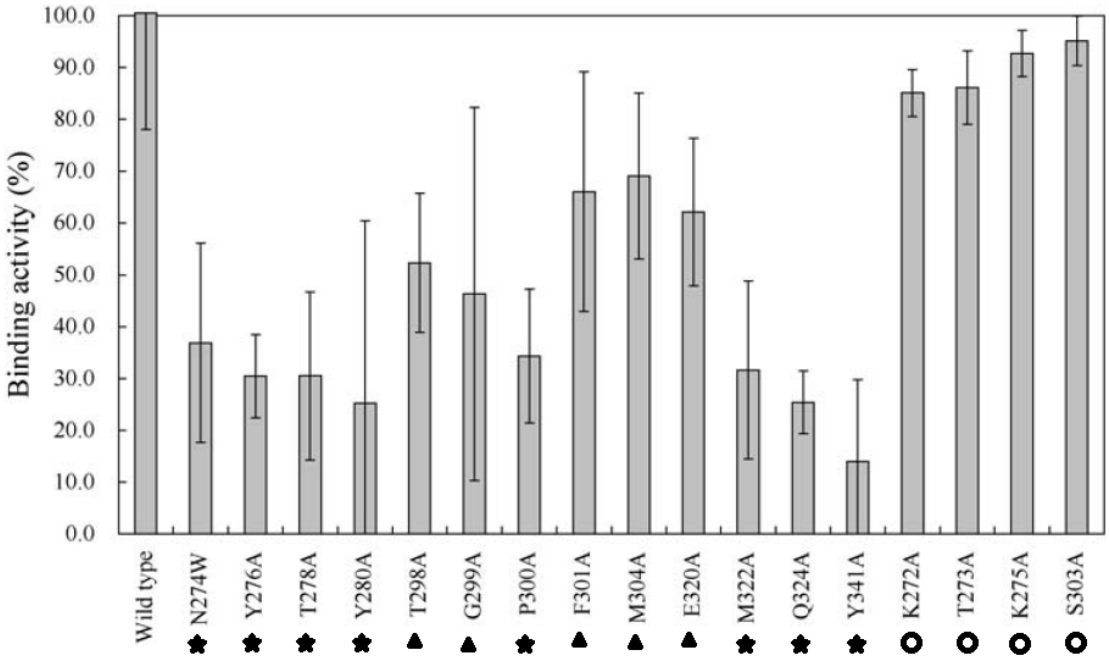


Figure 14 Binding activities of WT ALE-1 and corresponding mutants. The mutations show a dramatic decrease, a moderate decrease and a slight decrease in binding activity are marked by stars, triangles and circles, respectively. Independent experiments were performed and Standard deviations are shown in the error bars. (Adapted from Ref90)

As is shown in Figure 14, the mutations can be classified into three groups based on their impact on the binding activity. Except for four controls (K272A, T273A, K274A and S303A), all the mutants show a suppressed binding activity. MD simulation provides seven residues, six of which are important for binding activity. The only exception is S303. The decrease of binding affinity from other mutations, Y280A, G299A, P300A, F301A, M304A, M322A, Q324A, might be due to changes in the conformation of pentaglycine. These residues may not directly interact with pentaglycine, but they might push the pentaglycine binding in another conformation.

Interestingly, all mutations significantly or moderately decrease the binding affinity orient their side chains toward the binding groove, whereas the side chains from four control residues are not toward but far away from the binding groove. These residues

moderately suppressing binding affinity locate adjacent to the binding groove, which may explain their moderate effect (See Figure.15).

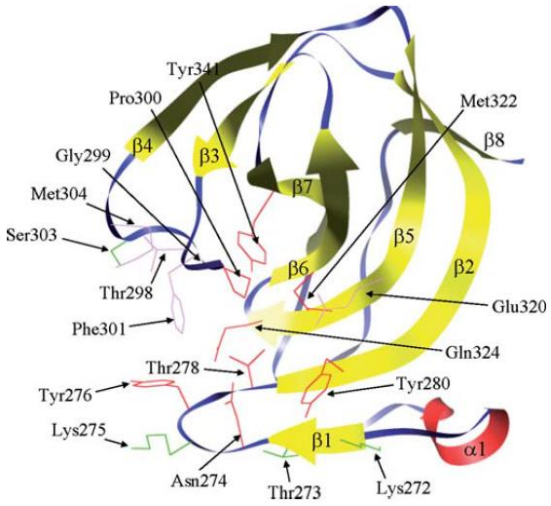


Figure.15 The thirteen residues in the binding cleft of ALE-1 and the orientations of their side chains. The α-helix and β-sheets are labeled and colored red and yellow, respectively. The loops in between are colored blue. The side chains from thirteen binding residues and four control residues are colored red and green, respectively (Adapted from Ref90).

Surface (S)-layer proteins

S-layer proteins are distributed among species of nearly every branch of eubacteria and Archaea[91]. When present, the S-layer proteins usually are the most abundant cellular proteins. They consist of single (glyco-) proteins which are capable of self-assembling into a 2-D crystalline array encompassing the whole cell. Their shapes are either oblique, square or hexagonal. Depending on the species investigated, the S-layer proteins can be comprised from 1-6 identical subunits and the molecular mass of the whole protein varies from 40,000 to 200,000[92]. Most of the S-layer proteins are acidic, and amino acid sequence analyses suggest that their *pKa's* (isoelectric point) range from 4 to 6. Purified isolated S-layer proteins can self-assemble in suspension or form a monolayer at the air-water interface. Although a lot of the structural and genetic information on S-layer proteins is available, the exact physiological role of this outermost envelope components is still in question. On the one hand, these metabolically expensive exo-proteins are ubiquitous in the environment. On the other hand, these S-layer proteins are frequently absent when the S-layer protein carrying organisms are grown under the optimal laboratory conditions. Thus, S-layer proteins must contribute to a selective advantage to the bacteria organism in their natural enviornment. The S-layer is described as a protective coat, an ion trap or an adhesion site for specific exoprotein. In some pathogenic bacteria, S-layer proteins also help the bacterial cells to evade the host immune system. Here, I will explain the binding mechanism and structural features of the S-layer proteins in details.

The binding site for S-layer proteins

S-layer proteins were originally assumed to bind to peptidoglycan, but now it has been demonstrated that the SCWPs serve as their binding sites in Gram-positive bacteria[93]. Moreover, Mesnage *et al.* discovered that, in *B. anthracis* and some other Gram-positive bacteria, the non-covalent attachment of S-layer proteins requires pyruvylation on the SCWPs. *B. anthracis* encodes two S-layer proteins, EA1 and Sap, the genes of which are in the chromosome. Two genes, *csaA* and *csaB*, which are in the upstream region of *ea1* and *sap* clusters, are involved in cell wall metabolism. Studies show that CasB, an enzyme decorating SCWPS with ketal-pyruvate[32, 94], is necessary for binding of EA1 and Sap to the cell wall[95]. The exact position of the ketal-pyruvated modification on the SCWPs is still unknown. Southern blot analysis has shown that S-layer-carrying organisms always possess *csaB* homologs; they also contain pyruvic acid in their cell walls. Pyruvic acid is lost in *csaB*-negative strains. EA1 and Sap only bind to pyruvylated SCWP fractions, they show no affinity for isolated pyruvic acid. This fact indicates that the oligosaccharide moieties of SCWPs also play a critical role in the binding process.

Surface layer homology (SLH) domain

Amino acid sequence analysis among S-layer proteins from different species has revealed a high sequence in the amino terminus of the proteins. Most of the S-layer proteins contain three repeats of a 55-AA sequence in this region, called SLH motif. This motif is necessary for S-layer proteins bind to SCWPs[96, 97]. Here we take the crystal structure of the SLH domains from Sap of *B. anthracis* as an example to illustrate the binding mechanism of S-layer proteins (Figure.16).

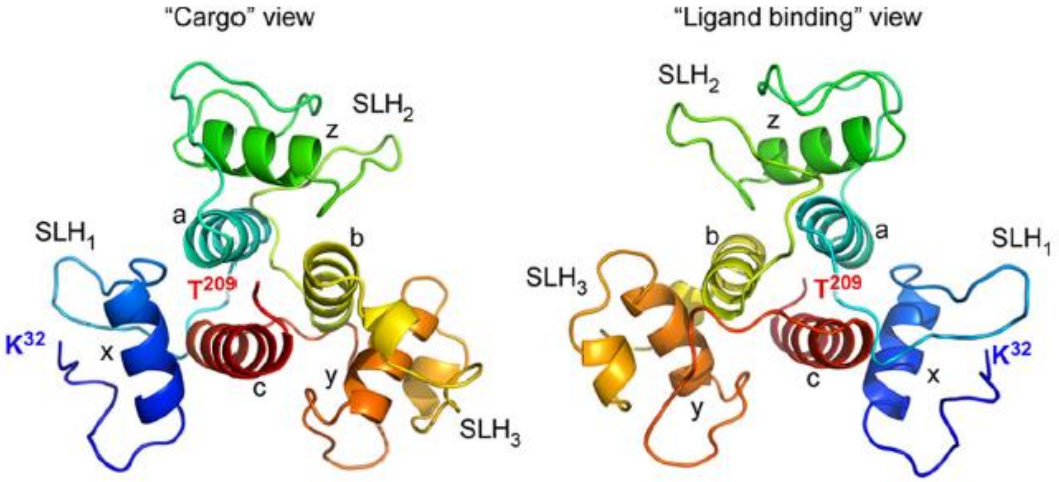


Figure.16 Two views of the Sap_{SLH} domains. Each SLH domain is marked by a different color, and the six main helices are denoted by a, b, c, x, y and z. The left picture is the view from C-terminus. The right picture shows the view from N-terminal, which is the ligand binding surface of the SLH-domain (Adapted from Ref98).

The crystal structure of the SLH domains in Sap was determined by single-wavelength anomalous diffraction (SAD) with Se-Met labeled crystals and refined to 1.8 Å resolution[98]. Three SLH domains form a "three-prong spindle" overall structure, in which every spindle consists of a single SLH domain (Figure.16). Although they possess highly similar 3-D structures **(**Figure.17**)**, the sequence homologies among the three motifs are quite low: SLH_1 versus SLH_2, 26%; SLH_1 versus SLH_3, 39%, and SLH_2 versus SLH_3, 26%. Each SLH domain forms one prong, which contains two loops (A and B in Figure 19) and the beginning part in the "spindle" linker helix (Figure 19). A group of five relatively conserved residues (LTRAE, IDRVS and VTKAE respectively, Figure.18) in each SLH domain occupy the last four residues of loop B and the first residue in the spindle linker helix[99]. Each of the motif of five residues contains a cationic residue: Arg^{72}, Arg^{131} and Lys^{193}, respectively (Figure.18). The ketal-pyruvate modification on the SCWPs introduces a strong negative charge. It is conceivable that the electrostatic interaction between the positively charged residues in the SLH domain and the negatively charged ketal-pyruvated SCWPs promote the non-covalent attachment process.

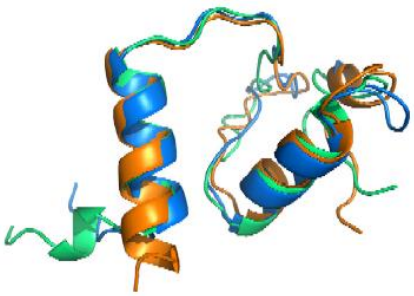


Figure.17 Spatial alignment of SLH1 (blue), SLH2 (green) and SLH3 (orange) (Adapted from Ref98).

Figure.18 Primary sequences of each single SLH domain are aligned via conserved residues (marked in red). Shaded residues are located in inter-prong groove 2, which is the space between the second and the third SLH prong of Sap. Residues of the relatively conserved five amino acids (LTRAE) are underlined (Adapted from Ref112).

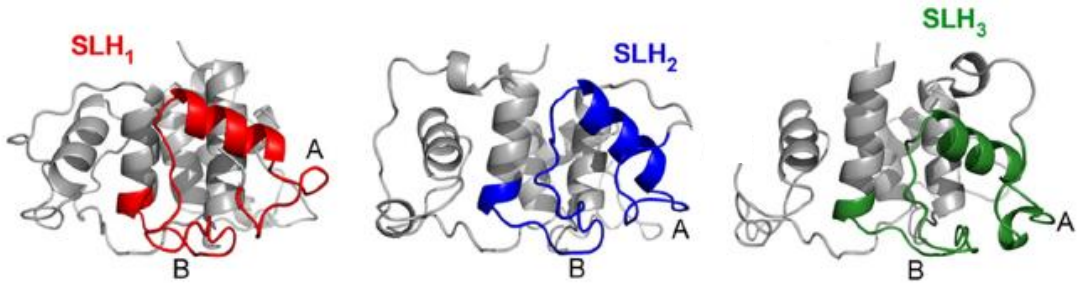


Figure 19 The lateral face of the Sap_{SLH} domains. Each domain is marked in red, blue and green, respectively. The central spindle core formed by the helices are not part of the SLH domain. Each SLH domain contains two loops, which are labeled A and B, respectively (Adapted from Ref98).

Furthermore, the inter-prong grooves (IPGs) formed by adjacent SLH domains promote the interaction between S-layer proteins and SCWPs. Each SLH domain contributes residues to the adjcent IPGs (Figure.20). The three conserved phenylalanine residues, *e.g.*, Phe^{34}, Phe^{95} and Phe^{156}, located at the interface of SLH domains, probably play an important role in ligand binding. In contrast, these three conserved aspartic acid residues (Asp36, Asp97, and Asp156) locate on the bottom of SLH domains, interacting with peptidoglycan.

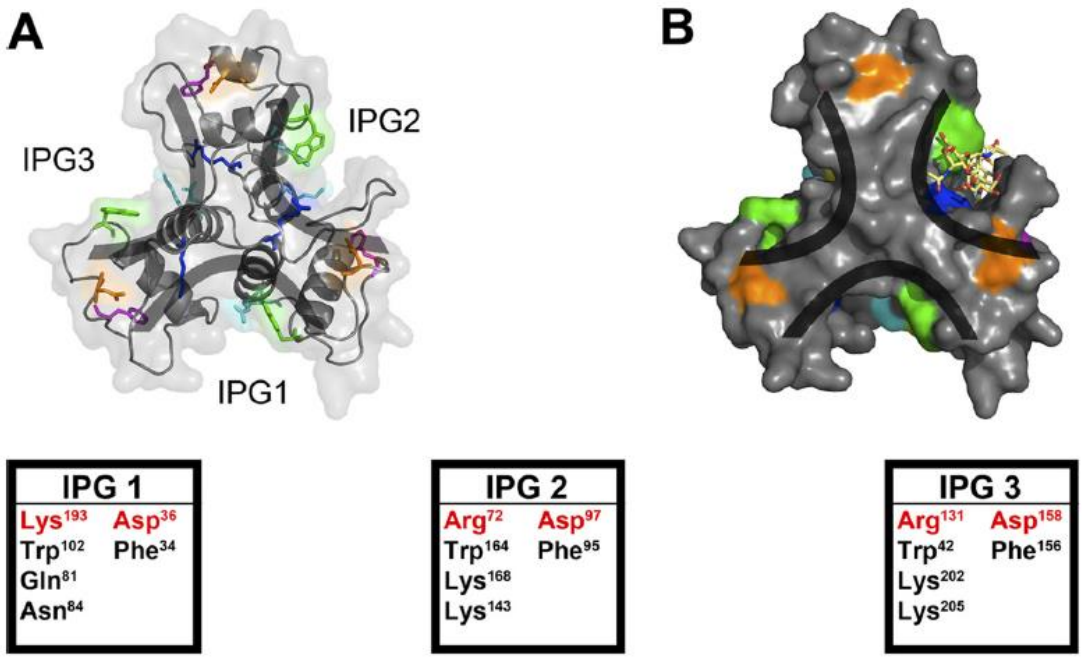


Figure.20 The three-prong spindle structure creates three binding clefts called inter-prong groove (IPG_{1-3}). A, ribbon model of Sap_{SLH}. B, space-filling model of Sap_{SLH}. Residues contributing to each IPG have been listed in the corresponding square. The SCWP molecule was constructed as described in Choudhury et al (Adapted from Ref98).

SLH_{1-3} domains were fused to the C-terminus of glutathione S-transferase (GST). The GST-SLH_{1-3} was incubated with *B. anthracis* vegetative cells from which the S-layer proteins had been stripped. As a measure of SCWPs binding, co-sedimentation of GST-SLH_{1-3} fusion protein was monitored (Figure.21). Clearly, deletion of any individual SLH domain abolishes the binding between the fusion protein and the bacterial cells. All fusion proteins with mutations display a stepwise decrease in cosedimentation, which demonstrates that the charged residues in the IPGs contribute to the binding of S-layer proteins to SCWPs of *B.anthracis.*

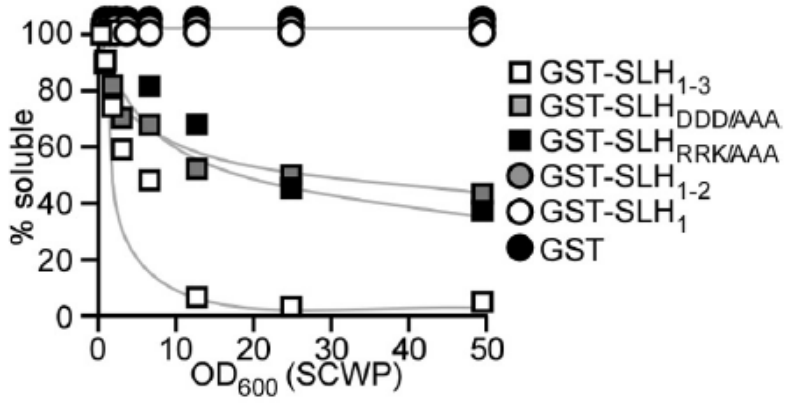


Figure.21 Purified GST-SLH_{1-3} and its variants are analyzed for their capability to co-sediment with vegetative *B.anthracis* cells stripped from their S-layer proteins. The vertical axis represents the percentage of soluble fusion-protein and the horizontal axis represents the amount of SCWP ligand. GST alone is taken as a control; GST-SLH_1 and GST-SLH_{1-2} are variants lacking the second SLH domain or the second plus third SLH domain, GST-$SLH_{DDD/AAA}$ and GST-$SLH_{RRK.AAA}$ are variants carrying mutations in conserved negatively or positively charged residues (Adapted from Ref98).

Choline Binding Proteins in Cell Wall Hydrolases

The choline moiety of TA, an anchor for various proteins

Choline, an important component in eukaryotic cell membranes, can also be found in some prokaryotes[101], such as *S. oralis*[102], *S. mitis*[103], *Clostridium beijerinckii*[104] and *C. acetobutylicum*[105]. Choline was discovered over four decades ago in pneumococcal cells; covalently links to TA and LTA and plays a fundamental biological role in pneumococci. Cells growing without choline show numerous abnormalities, *e.g.*, they are resistant to bacteriophages and form long chains of more than 100 cells[106] (Figure.22). These phenomena suggest that choline metabolism in *S. pneumoniae* is highly associated with phage adhesions and cell wall hydrolysis in cells separation. However, choline is not necessary for growth of other prokaryotes. In *S. pneumoniae*, choline can be specifically recognized by various proteins, including cell wall hydrolases (CWHs) LytA, LytB, LytC and other virulence factors involved cell adhesion and colonization. Thus, choline is considered as a critical ligand that interacts with various proteins. All four pneumococcal cell wall hydrolases possess a tandem of characteristic repeating motif named the choline binding domain (ChBD), which have been identified in a number of proteins, including CspA in *C. beijerinckii*, toxin A and B in *C. difficile*, glycan binding proteins in *S. mutants* and Cpl-1 in pneumococcal bacteriophages. The 3-D structures of ChBD from several proteins, such as Cpl-1, LytA, CbpF and CbpE, have been solved. Here, LytA is taken as an example to elucidate the structural features and choline binding mechanism of ChBDs.

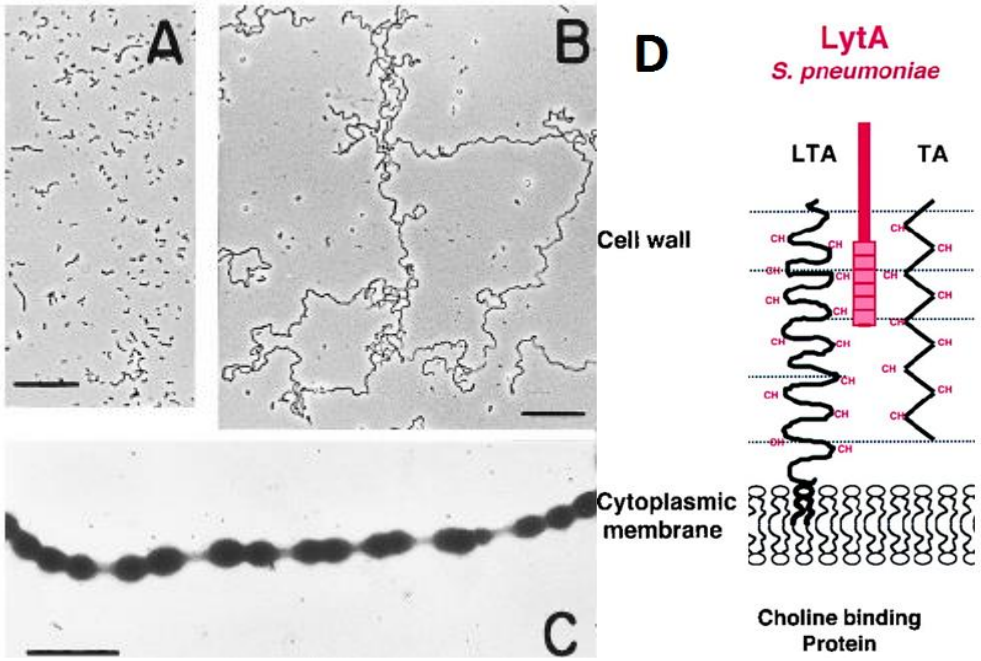


Figure.22 *S. pneumonae* M31B mutant, in which one of the murein hydrolase, gene *LytB*, has been inactivated. The parental strain M31 shows a normal phenotype (A) whereas M31B mutant forms long chains with more than 100 cells (B). (C) Transmission electron microscopy of picture of M13B mutant. The bar in A and B represents 25 μM, that in C represents 2 μM. (D) the putative binding model of LytA. The choline moieties are marked by CH, choline binding domain are colored by pink squares (Adapted from Ref106).

The four CWHs of *S. pneumoniae* and their choline binding domain

As mentioned above, the choline on the surface of the pneumococcal cell wall acts as a ligand of various ChBPs. Four ChBPs have been well characterized to date in *S. pneumococcus.* All of these are cell wall hydrolases (CWHs), i.e., a β-N-acetylmuramidase named LytC, a β-N- acetylglucosaminidase (LytB), the major autolysin LytA and the phosphorylcholine esterase Pce[108]. Figure.23 shows an amino acid sequence/motif comparison of different bacterial and bacteriophage derived CHWs[109].

LytA was the first bacterial autolysin that has been successfully cloned and expressed[110]. The enzyme consists of 318 amino acid residues with a predicted molecular mass of 36,532 (Figure.23). The LytA primary translation product (E-form) has low enzyme activity, while the fully-active mature protein (C-form) is the result of conversion from the E-form in the presence of choline[111]. The LytB gene encodes a 76.4 kDa glucosaminidase with 658 amino acid residues, which includes a 23-AA cleavable signal peptide[112] (Figure.23). LytB is involved in cell separation and is highly expressed during the early exponential phase of growth. LytC is a protein with 501-AA containing a cleavable signal peptide of 33-AA residues (this signal peptide is not included in 501-AA). LytC may function as a lysozyme[113]. The pneumococcal phosphorylcholine esterase Pce contains 627-AA residues, including a 25-AA residues signal peptide. RT-PCR experiments suggest that expression of *pce* is

maximal in the stationary phase of growth, which is quite different from what has been observed in *lytA* and *lytB* (exponential phase).

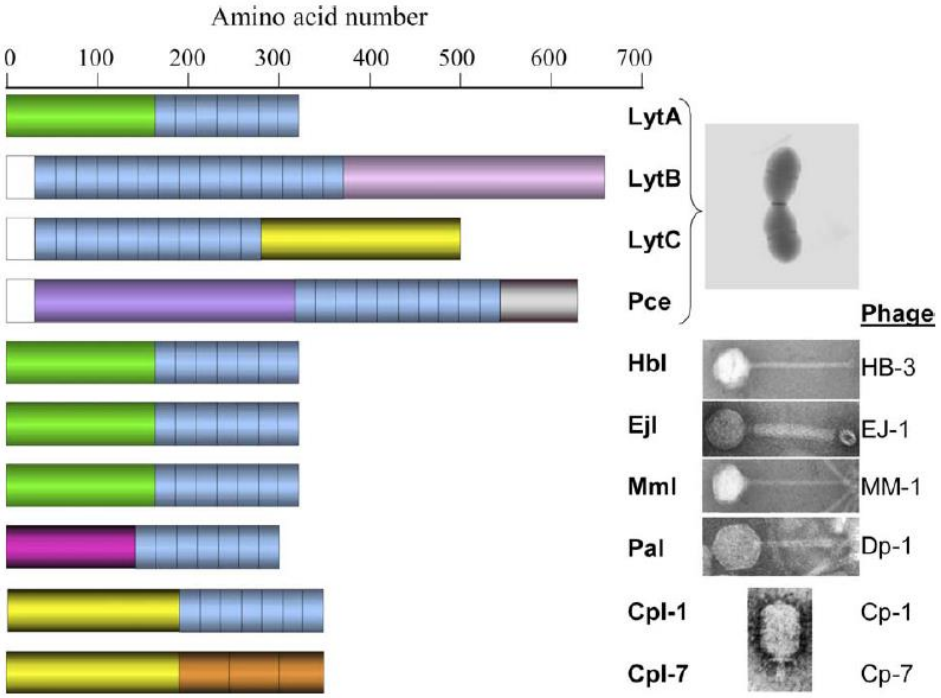


Figure.23 Sequence/motifs constructions among different ChBDs. LytA, LytB, LytC and Pce are CWHs expressed by *S.pneumoniae*, HbI, EjI, MmI, PaI, CpI-1 and CpI-7 are proteins expressed by the bacteriophages. Purple, yellow, pink, violet and green represent the sequence containing the active site in amidase, lysozymes, glucosaminidase, phophorylcholine esterase and cell wall hydrolase, respectively. Signal peptides in LytA, LytB, LytC and Pce are indicated by open rectangles. ChBDs with repeating motifs are marked in blue. The ChBD homolog of CpI-7 is colored by orange. The electron micrographs of *S. pneumoniae* and phages are shown at the right side (Adapted from Ref109).

All four pneumococcal CWHs are comprised of two large domains (Figure.23): the catalytic domain and the choline binding domains (ChBDs). The ChBDs are necessary for cell wall hydrolytic function[114]. Previous studies have shown that both LytA and Cpl-1 are choline-dependent enzymes that share significant C-terminal sequence similarity. The amino acid sequences in their N-termini are distinct. Based on this fact, it is convincible that the choline binding domains (ChBDs) reside on the C-terminus, whereas their active sites locate at N-termini. But the position of ChBDs are vary in CWHs. Figure.23 shows that the ChBDs in LytB and LytC locate in the N-termini of protein. The ChBDs of Pce are near to the C-terminus of the protein and are followed by 85 amino acid residues with unknown function. Further amino acid sequence analyses show that ChBDs are comprised of a number of characteristic repeats with approximately 20-AA residues. These repeats are also called choline binding repeats (ChBRs) and are composed of a conserved motif that can be classified into the cell wall binding repeat family (pfam PF01473). Although more than 1,400 members are documented in PF01473, only few of them display choline binding activity[116]. The number of these ChBRs also varies among proteins with different origins. LytA, LytB, LytC and Pce contain 7, 15, 11 and 10 repeats, respectively. It

has been suggested that a minimum of 4 ChBRs are required for choline binding[117]. It is reasonable to speculate that the binding affinity can be improved by possessing more ChBRs. The amino acid sequence conservation among ChBRs in one CWH enzyme can also be different, e.g., the ChBRs in the C-terminus of LytB are more conserved than other ChBRs. This amino acid sequence variability may explain the flexible specificity of LytB, which binds not only to choline, but also to ethanolamine or phosphorylcholine[118]. Up to date such flexible specificity has not been reported for the other three CWHs of *S. pneumoniae*.

The catalytic activity is independent from choline binding process

The choline moiety only serves as an anchorage of CWHs. It does not contribute to the catalytic activity of the anchored enzyme. Previous studies have shown that pneumococcal LytA can digest murein from Gram-negative bacteria, which contains neither TA nor choline. This finding proves that the catalytic activity of LytA is choline-independent, and free peptidoglycan is a good substrate for LytA[119]. However, choline less peptidoglycan isolated from pneumococci shows a high resistance to LytA[120]. Based on these findings, it seems that the choline moiety is necessary for cell wall hydrolysis as long as the substrate is insoluble, and that TA may function as an inhibitor of CWH activity. The function of the choline might be to unblock the inhibiting effect of TA[121], but there is still few understanding about the exact process of this pathway.

Crystal structures of the LytA monomer

The crystal structure of LytA is taken as an example to illustrate structural features and the choline-binding mechanism of this type of enzyme. As mentioned above, LytA is responsible for cleaving the MurNAC-Ala bond in PGN. The crystal structure of monomeric LytA was determined at 2.6 Å resolution using the multi-wavelength anomalous dispersion method[122]. The LytA monomer shows a height of approximately 60 Å with a diameter of 25 Å. Two different views of this structure clearly show that six hairpin structures are formed by a tandem of ChBRs (Figure.24). These six hairpins have the same length and all are hydrophobic in character. Each of these hairpins consists of two antiparallel β-strands connected by a short loop. With the exception of hairpin 6, each hairpin introduces a 120 ° counter-clockwise rotation in the overall superhelix structure. As is shown in Figure.24 b, three hairpins form a complete triangle upon superimposition and each hairpin climbs up by approximately 10 Å. The singular position of hairpin six indicates that it has a unique function (discussed below).

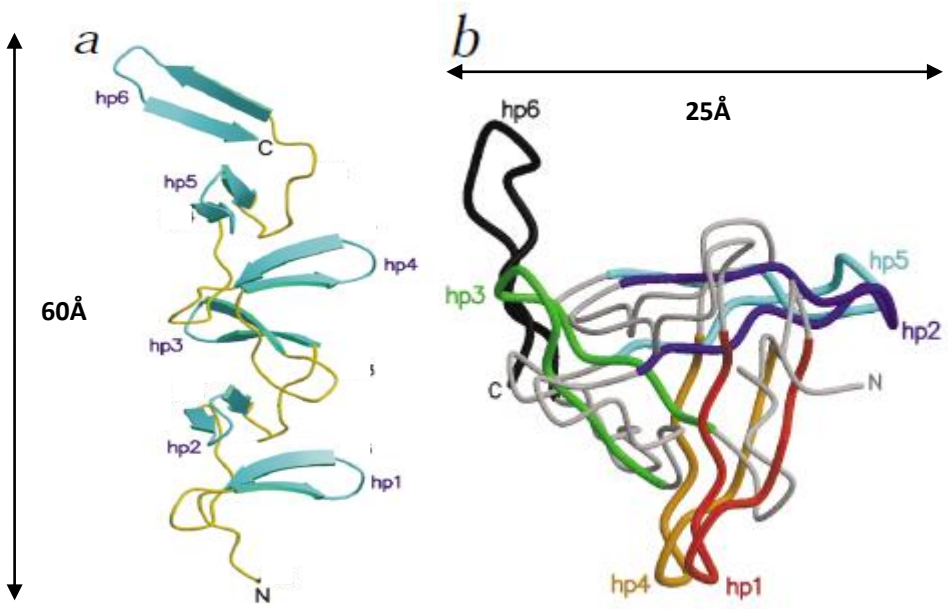
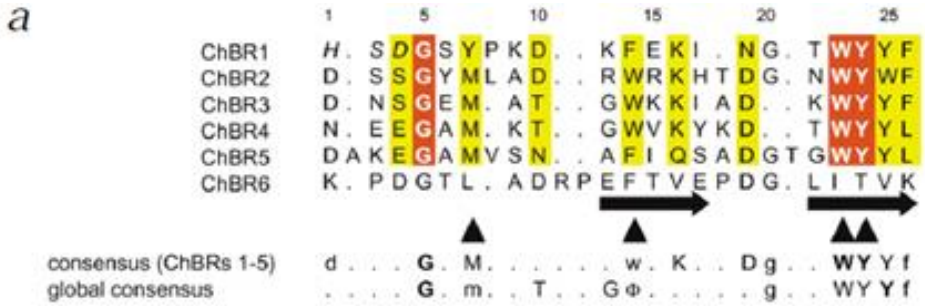


Figure.24 The crystal structure of the LytA monomer. (a) Orientation of LytA monomer from a the lateral view. The six hairpins (hp) are marked cyan and numbered from hp 1 to hp 6, while the loops connecting neighboring hairpins are colored yellow. The N-terminus and C-terminus are labeled by N and C, respectively. (b) The C_α trace of LytA monomer from front view. The hp structures are numbered hp 1 to hp 6. The N-terminus and hp 1 are placed on the top while the C-terminus and hp 6 are put at the bottom (Adapted from Ref122).

Amino acid sequence alignment of the six hairpin structures reveals that hairpin 6 is distinct from the others (Figure.25). The aromatic residues in the second β-strand of hp 1 to hp 5 are well conserved, whereas those in the first β-strand are less conserved. Interestingly, the loop motifs connecting two consecutive hairpins possess an invariant glycine residue at position 5. This residue is considered to be essential for topology. In contrast, hp 6 possesses a quite singular primary sequence and spatial structure (Figure.25). The angle between hp 6 and hp 5 is only 95 °, while none of the conserved residues are present in hp 6. The start of hp 6 contains an insertion of two residues between the conserved amino acid residues (Figure.25 a).



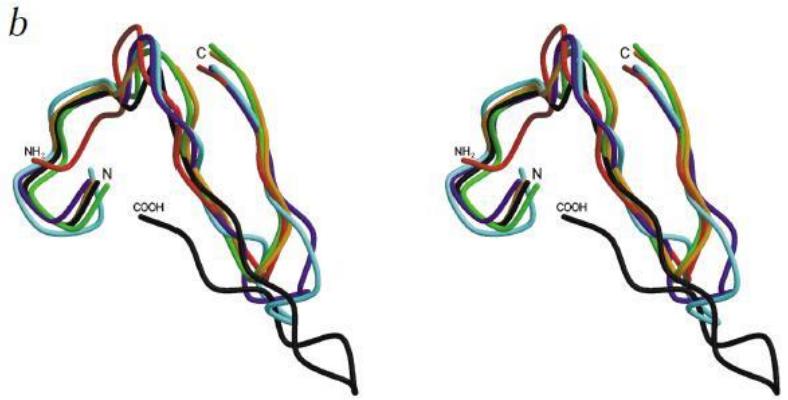


Figure.25 Amino acid sequence and the spatial alignment of the six ChBRs of LytA. (a) Primary sequence alignment of ChBRs. Repeat numbers are listed in the left side, corresponding to hp 1 to hp 6. The residues forming two β-strands are indicated by arrows and residues binding to choline moiety are marked by black triangle. Residues with high conservation are colored red and (100% conservation) or yellow (conservation >50%). The consensus of ChBRs1-5 and the global consensus obtained from comparing >600 cell wall binding motifs of pfam PF01473 are also shown, while small letter indicating >60% conservation, capital letters indicating >80% conservation, bold letters indicating 100% conservation and Φ indicating hydrophobic residues. (b) Superpostion of the six ChBRs in LytA, with the same color scheme as in Figure.24. (Modified from Ref122)

Crystal structure of LytA homodimer and choline binding sites

It was first reported by Usobiaga et al[123] that the LytA forms homodimers in the presence of the choline. Crystal structure (Figure.26**)** shows that two LytA monomers interact with each other in the homodimer via their C-terminal hp 5 and hp 6. This structure is supported by earlier experiments that absence or truncation of the C-terminal in LytA causes a dramatic decrease (>90%) in catalytic activity and dimer dissociation[124]. The homodimer organization is maintained by choline moieties which are deeply buried into the cavities formed by neighboring hairpins. Structural analysis shows that those cavities are predominantly hydrophobic[122] and serve as choline binding sites[125].

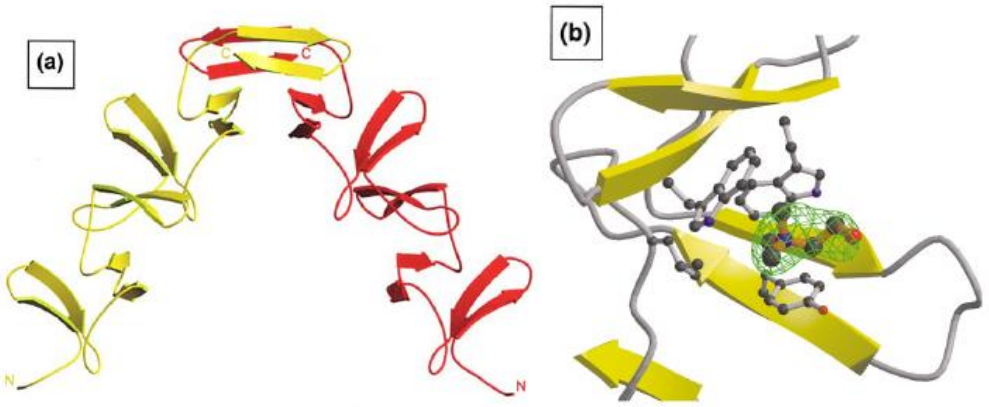


Figure.26 Crystal structures of the LytA homodimer and the choline binding sites. (a) Ribbon diagram of a LytA dimer. Two monomers are colored by red and yellow, respectively. The N-terminal and C-terminal are indicated by letter (N) and (C), respectively. The hp 5 and hp 6 interact with their counterparts in another monomer to form the

LytA dimer. (b) The choline binding site is formed by two neighboring hp strucutres. The β-sheet strands are represented by yellow arrows, while the choline moiety is repersented by ball-and-stick form. The side chains from hydrophobic residues in the binding cavity are shown in ball-and-stick forms (Adapted from Ref 109 and 122).

The homodimer of LytA displays an overall boomerang shape with two 50 Å-long arms with an 85 ° angle (Figure.26 a). It is very likely that the catalytic centers reside at the end of each arm (the N-terminus). The interaction between the monomers is mainly mediated by the aromatic and hydrophobic residues in hp 5 and hp 6. Furthermore, an area of 1,950 $Å^2$ in each monomer is involved in dimerization, which accounts for nearly a quarter of the total surface area in LytA (8,600 $Å^2$).
Four choline binding sites have been found in each LytA monomer, each of which is composed of the interfaces from two consecutive hairpins: ChBS1 (hairpins 1 and 2), ChBS2 (hairpins 2 and 3), ChBS3 (hairpins 3 and 4) and ChB4 (hairpins 4 and 5). The interface between hp 5 and hp 6 is involved in dimerization, thus does not form a choline binding site. In these four binding sites, the choline moiety fits into a cavity with a volume of ~15 $Å^3$. The cavities are predominantly hydrophobic and are formed by three aromatic residues from nearby hairpins and one hydrophobic residue usually from an extending loop connecting two hairpins. The interaction between the electron-rich cavity and the positively charged choline moiety promotes binding affinity. Furthermore, the choline moiety in TA helps shield those hydrophobic residues from solvent molecules and stacks hairpin structures together via hydrophobic interactions. Other Lyt proteins binding with choline analogs also possess similar cavities in their choline binding sites.
Each monomer of LytA forms a left-handed superhelix structure which contains several grooves between the ChBRs. These grooves are filled with polar and charged residues with ~ 10 Å in length and ~7 Å in depth. The proposed function of these grooves is to accommodate TA and to establish hydrogen bonds with N or O atoms on the N-acetylated galactosamine residues of TA. As the number of ChBR varies among different CWH proteins, that may slightly affect the binding specificity.

Binding motifs in peptidoglycan recognition proteins (PGRPs)

The general composition and structure of peptidoglycan strands are well conserved in both Gram-negative and Gram-positive bacteria, which indicates that there may be a general pathway in the host immune system to recognize PGN molecules. Indeed, as the first line of defense against microbes in both vertebrates and invertebrates, the innate immune system recognizes bacterial PGN molecules by several specific recognition proteins, *e.g.*, Toll-like receptor 2 (TLR2), CD14, PGN lytic enzymes and peptidoglycan recognition proteins (PGRPs), which are quite conserved during evolution[126-128]. These proteins specifically interact with certain products from microbial metabolism but not those of the host. To date only PGRP with its PGN ligand has been examined by X-ray crystallography. No structural information is available about how the other PGN recognition proteins interact with their ligands.

Therefore, I will only review the current knowledge on how PGRPs bind with their ligands and describe the mechanism of PGN binding.

Discovery, distribution and functions

The peptidoglycan recognition proteins (PGRPs) were first introduced by *Ashida* in 1996[129]. A 19-kDa protein presents in the cuticle of silkworm was purified (*Bombyx mori*) and shown to not only bind to the PGN in Gram-positive bacteria, but also to activate prophenol oxidase[79]. Subsequently, a mount of homologs of PGRPs were found in mouse, fruit fly (*Drosophila melanogaster*), mosquito (*Anopheles gambiae*) and human genomes. To date, PGRPs have been found in most animals, including insects, echinoderms, molluscs and vertebrates, but not in metazoans or plants[130].
PRGPs are highly conserved from insects to mammals[131] and are critical components of the innate immune system, specifically interacting with PGN from both Gram-positive and Gram-negative bacteria. The PGRP family includes nearly 100 members[80]; all PGRPs contain a conserved PGN-binding domain of approximately 165 amino acids residues. Amino acid sequence analyses indicate that these domains belong to the peptidoglycan-binding type 2 amidase domain, which is the homolog of bacteriophage T7 lysozyme, a Zn^{2+}-dependent amidase that hydrolyzes the PGN wall. Like T7 lysozyme, some of the PGRPs possess hydrolysis activity on the amide bond between L-Ala and MurNAc, whereas most of the PGRPs only bind PGN and have no hydrolysis activity due to the lack of cysteine residues responsible for Zn^{2+} binding[87].
PGRPs in insects can be classified into two groups: 20-kDa extracellular proteins called short PGRPs (PGRP-S), and 90-kDa PGRPs (PGRP-L), which can be either intracellular, extracellular or transmembrane proteins. The 13 PGRP genes in the *D.melanogaster* genome are transcribed into at least 17 PGRPs, seven of which are PGRP-S and contain a signal peptide[132]. The other ten PGRPs have either a signal peptide or a predicted transmembrane domain (transmembrane PGRPs) or neither of them (intracellular PGRPs). The PGRPs in insects usually secrete to immune competent organs, *e.g.,* fat bodies or hemocytes.[133] The secretion process associates with exposure to PGNs or other microbial challenge, indicating that PGRPs might be involved in antimicrobial defense. Indeed, in *Drosophila* the PGRPs activate two distinct immune response pathways depending on the PGN type: the Toll receptor pathway is triggered by Lys-type PGN from Gram-positive bacteria while the Imd pathway is triggered by Dap-type PGN from Gram-negative bacteria. This fact also suggests that the PGRPs are capable of discriminating between Lys-type PGN and DAP-type PGN. Furthermore, in silkworm (*B. mori*) and mealworm (*Tenebrio molitor*) the PGRP-S[134] and PGRP-LE[135] activate the prophenol-oxidase pathway via binding both DAP-and Lys-type peptidoglycan[136]. This pathway produces antimicrobial products covering the infection site, for instance melanin. Figure.27 shows the physiological functions of PGRPs in insects[137].

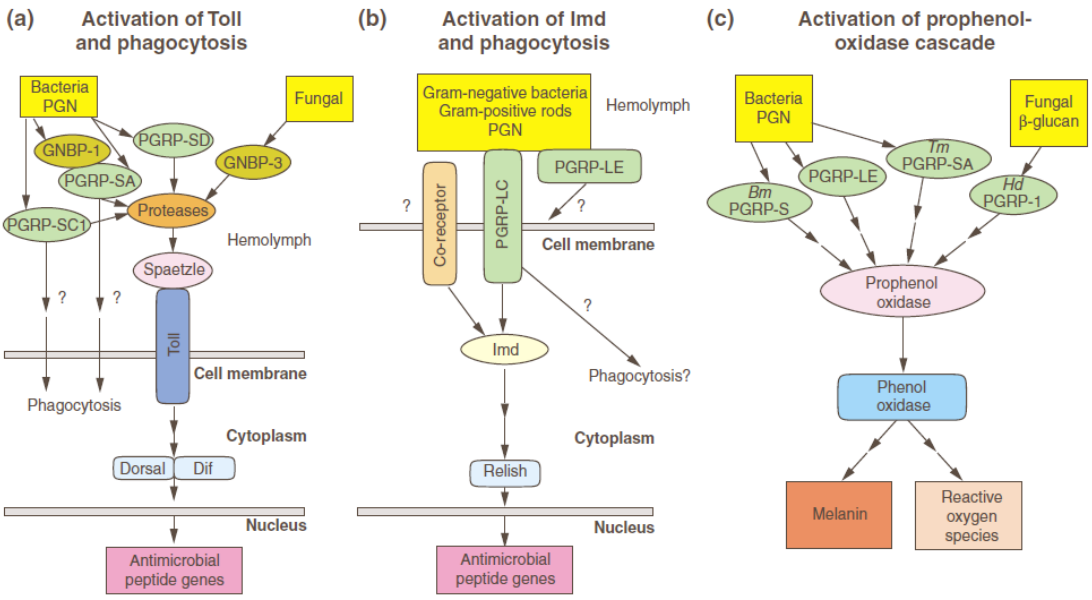


Figure.27 Physiological roles of PGRPs in insects. (a) Three short-type PGRPs, PGRP-SA, PGRP-SD and PGRP-SC1, recognize the Lys-type PG in Gram-positive bacteria and active the protease hydrolyzing Spaetzle, which is an endogenous activator of the Toll-pathway. (b) Recognition and binding to the PGN from Gram-negative or Gram-positive bacteria induce the oligomerization of RP-LC, by which the Imd pathway is activated. This activation process probably requires cooperation with PGRP-LE. (c) The binding of PGN to PGRP in silkworm and mealworm triggers the activation of prophenol oxidase. Multiple steps are marked by multiple arrows, question marks point out unconfirmed functions (Adopted from Ref. 137).

Mammals have four PGRPs, PGRP-S, PGRP-L, PGRP-Iα and PGRP-Iβ (S: short; L: long; I: intermediate); they are expressed in various tissues and organs and have two major functions: to provide amidase and antimicrobial activity. The Human Genome Organization Gene Nomenclature Committee has recently termed these four PGRPs as PGLYRP-1, PGLYRP-2, PGLYRP-3 and PGLYRP-4, respectively. The mammalian PGRPs are all soluble proteins, which means they are either extracellular or intracellular proteins. PGLYRP-2 (PGRP-L) possesses amidase activity and cleaves the lactyl bond between the L-Ala and MurNAc in bacterial PGN[138, 139]. PGLYRP-2 is synthesized in the liver and is continuously secreted into the blood stream to eliminate proinflammatory PGN. Human PGLYRP-1 (PGRP-S) is selectively expressed in bone marrow[140, 141] and has been considered as bactericidal in the presence of Ca^{2+} [142, 143], without which it only shows a bacteriostatic effect[128, 144]. Additionally, mice with a lower expression level of PGRP-S are more susceptible to intraperitoneal infections with Gram-positive bacteria[91]. Human PGLYRP-3 (PGRP-Iα) and PGLYRP-4 (PGRP-Iβ), which can be found in skin epidermis and hair follicles, form either homo- or heterodimers via disulfide linkages. These PGRPs are bactericidal for Gram-positive bacteria and also bacteriostatic against Gram-negative bacteria (Figure.28).

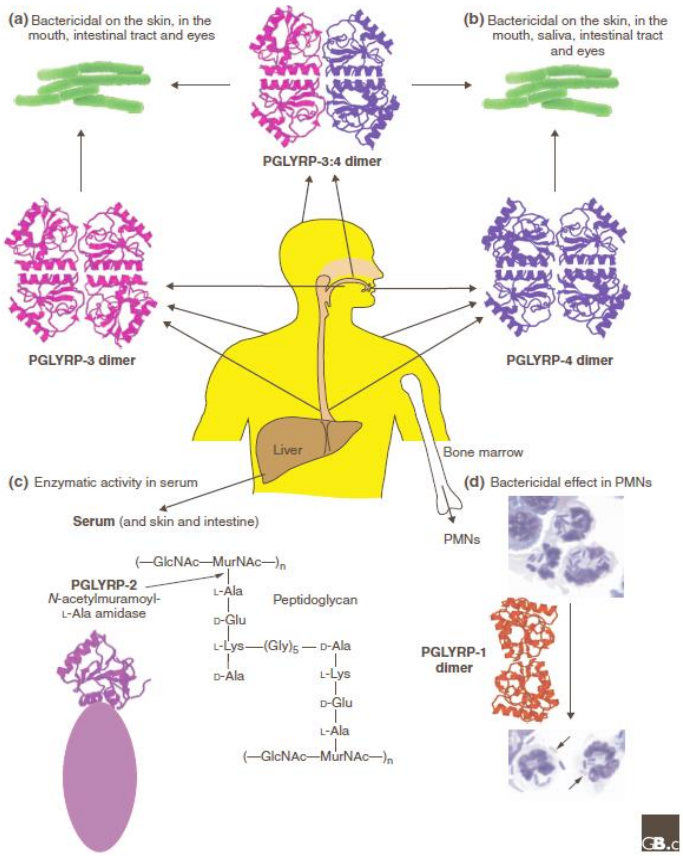


Figure.28 Physiological functions of PGRPs in mammals. The central picture shows the regions where PGLYRP are presented. (a), (b) the homo- and heterodimer of PGLYRP3 and PGLYRP4 possess bactericidal activity and can be found on the skin, in the mouth, intestinal tract and eyes. (c) PGLYRP-2 is continuously expressed in liver and secreted into the blood stream to function as an N-acetylmuramoyl-L-Ala amidase cleaving the bond between MurNAc and L-Ala. (d) PGLYRP-1 is produced in bone marrow and functions as a bactericidal agent (Adopted from Ref. 137).

Overall structures of PGRP-binding domains

All PGRPs contain at least one PGN binding domain, some of which, *e.g.*, in *Drosophila* PGLYRP-3 (PGRP-Iα) and PGLYRP-4 (PGRP-Iβ) and PGRP-LF contain two PGN binding domains[145]. All these PGN binding domain show a fold similar to that in T7 lysozyme, which has been confirmed by the crystal structure of PGRP-LB from *Drosophila*. Six additional crystal structures of PGN binding domains have been determined, including PGRP-Iα C-terminal domain, PGRP-S from human (PGLYRP-1), PGRP-SA, PGRP-LCa, PGRP-LE and PGRP-LCx from *Drosophila.* All of these have a similar fold with five β-sheets (four of which are parallel, and one is antiparallel) and three α-helices in central core. All PGN binding domains contain an N-terminal segment with considerable variability that is absent in T7 lysozyme. This segment is called the PGRP-specific segment and forms a hydrophobic groove with helix α-2. This topologically variable groove is located opposite of the PGN binding site, probably serving as a docking site to interact with different co-effector proteins[146,147]. Moreover, all PGN binding domains contain 1-3 disulfide linkages.

One of these (Cys214-Cys220 in PGLYRP-3) is present in all PGN binding domains except PGRP-LE from *Drosophila*. This disulfide linkage is considered to be necessary for structural integrity of the PGRPs. One of the other two disulfide linkages, Cys194-Cys238 in PGLYRP-3, only occurs in mammalian PGRPs. All PGRP binding domains have a conserved L-shaped PGN binding groove with one deep and one shallow end (Figure 29) [145].

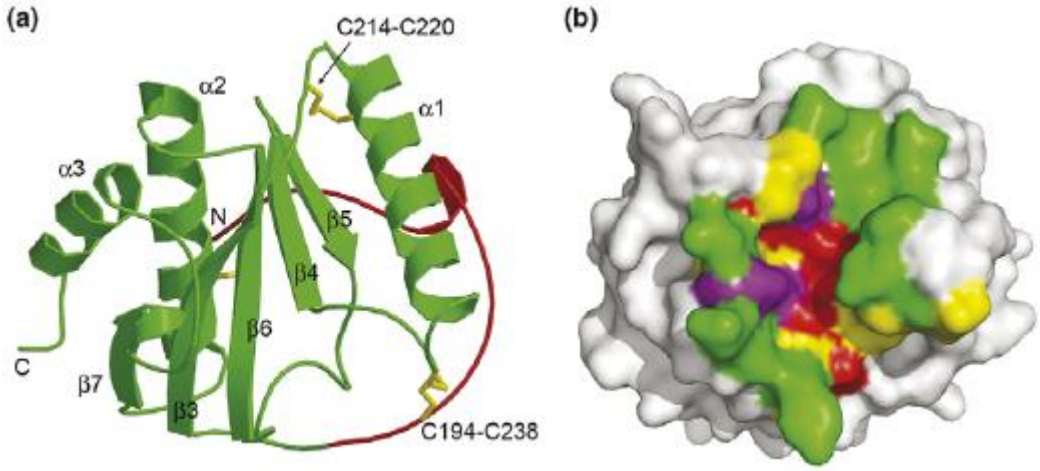


Figure 29 (a) Crystal structures of human PGLYRP-3C. Secondary structural elements (α-helixes and β-sheets) are numbered. The PGRP-specific segment is depicted by red and two disulfide linkages are labeled by C214-C220 and C194-C238, respectively. C-terminus and N-terminus are labeled by C and N, respectively. (b) The binding groove in PGLYRP-3C. The orientation of the molecule is the same as in (a). Surface colors correspond to percentage of conserved residues obtained from aligning 45 PGRPs amino acid sequences: red > 80%, purple 60-80%, yellow 40-60% and green, lower than 40% (adapted from Ref. 13).

Binding mechanisms of PGRPs

As indicated above, PGRPs contain a conserved L-shaped binding groove with one deep and one shallow end. This groove, which is predominately hydrophilic and filled with solvent molecules, binds the MurNAc moiety and the peptide stem of PGN in its deep part while accommodating GlcNAc in the shallow end. Crystallographic analyses performed on PGLYRP-1, the C-terminal of PGLYRP-3, insect PGRP-LB, PGRP-SA and PGRP-LE also show that the minimum binding fragment of Lys-PGN is MTP, MurNAc-L-Ala-D-isoGln-L-Lys, and a muramyl tripeptide. The PGRPs do not bind MurNAc-dipeptide, or the peptide stem without a MurNAc moiety. The high affinity of PGRP for PGN is achieved by burying the MurNAc moiety and peptide stem deeply into the cleft excluding water molecules. Thus, it is conceivable that the relatively well conserved residues (red and purple in Figure.30) interact with the common molecular core and the variable residues around the binding groove (the yellow and green areas in (Figure.30) provide particular specificity of PGRPs to different PGN molecules[145].

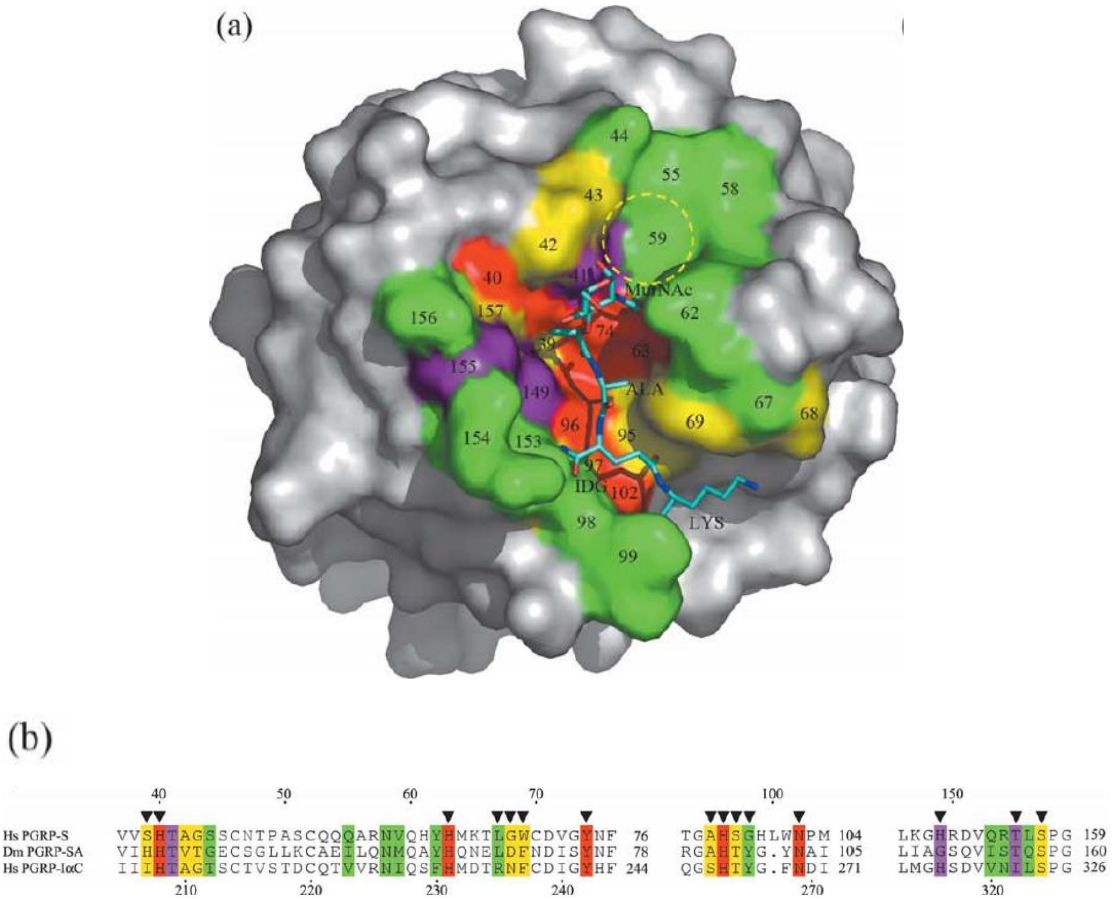


Figure.30 (a) PGN binding groove and putative PGN-interacting residues in human PGYLRP-1 (PGRP-S). Colors scheme is the same as in Figure.14. MTP, the analog of PGN, is shown as stick-and-ball figure. Carbon atoms are marked by cyan while nitrogen and oxygen atoms are depicted by blue and red, respectively. The putative binding site for GlcNAc is marked by a yellow circle. (b) Amino acid sequence alignment of PGN-interacting residues among PGLYRP-1 (PGRP-S), PGLYRP-3 (PGRP-IαC) and *Drosophila* PGRP-SA. Residues are colored according to the color scheme in Figure 29 (adapted from Ref. 147).

Mechanism to recognize PGNs

The structural information mentioned above indicates that the affinity of PGN binding can be modified by structural and conformational difference in PGRPs. Moreover, the capability of discriminating Lys-PGN from Dap-PGN in some PGRPs may be based on these structural differences. In the PGLYPR-3-MTP complex, Asn236 and Phe237 interact with the side chain of L-Lys on the peptide stem via van der Waals forces. These two residues are relatively well conserved in all PGRPs (highlighted in yellow in Figure.30 b and Figure.31). This remarkable structural feature indicates that the variation in these two residues may help the PGRP molecule to discriminate bacterial Lys-type and Dap-type PGNs. Structural alignments among more homologs reveals that these two positions, however, are usually conserved and occupied by Gly and Trp. Moreover, various organisms share conserved binding domains starting with conserved GW residues (GW module).

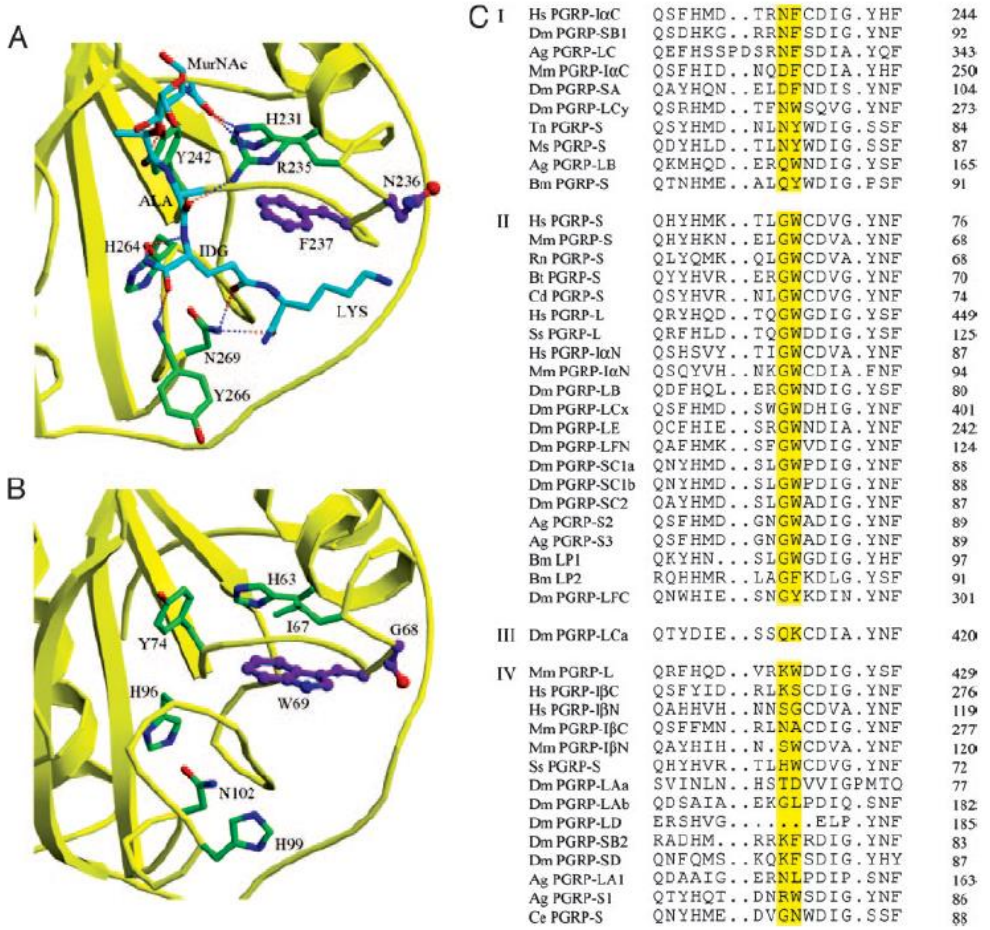


Figure.31 Crystal structure of the MTP bound complex of (A) PGRP-IαC (PGLYRP-3) (B) PGRP-S (PGLYRP-1). (C) Amino acid sequence alignment of specificity determining residues among different PGRPs. MTP is shown as a stick-ball, with carbon, nitrogen, oxygen are colored by cyan, blue and red, respectively. Residues hydrogen bonding with MTP are green. The two pairs of purple residues, N236 and F237 in (A), G68 and W69 in (B), form Van der Waals contacts with the side chain of L-Lys. (C)The PGRPs are classified into four groups, each of which is described in the text. Residues highlighted in yellow are specificity-determining residues corresponding to N236 and F237 in PGLYRP-3. Mammals: Bt, *Bos taurus*; Cd, *Camelus dromedarius*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Ss, *Sus scrofa*. Insects: Ag, *Anopheles gambiae*; Bm, *Bombyx mori*; Ce, *Calpodes ethlius*; Dm, *Drosophila melanogaster*; Ms, *Manduca sexta*; Tn, *Trichoplusia*. (adapted from Ref. 148)

The PGRPs are classified into four distinct groups based on the amino acid sequence alignment of specificity determining residues in PGRPs from insect to human (Figure.31 C). Group I proteins possess Asn-Phe or other homologous residues to discriminate Lys-type and Dap-type PGN. PGRPs from this group show strong prefernence to bind MPP, CL-PGN but not MPP-Dap. Group II PGRPs contain Gly-Trp and show a higher affinity for MPP-Dap than for MPP (no interaction with CL-PGN). Group III, which contains only one member so far (PGRP-LCα), can only bind MPP. The specificity of Group IV proteins is still unknown[148].

To verify the effects of these two residues in determining the specificity of PGRPs, Rongjin Guan *et al.*[148] mutated the Gly and Trp residues in PGLYRP-1 to Asn and Phe, respectively. In contrast to PGLYRP-1, which accommodates MPP 7-fold less affinity

than PGLYRP-3, the double mutated PGLYRP-1 shows a 2-fold better affinity than PGLYRP-3. And similar to PGLYRP-3, the double mutant protein PGLYRP-1 also shows a 25-fold lower affinity to bind with MPP, whereas WT PGLYRP-1 predominately binds to MPP-Dap PGN. Moreover, the double mutated PGLYRP-1 and PGLYRP-3 possess similar affinity to bind with CL-PGN (cross-linked Lys PGN), while the WT PGLYRP-1 shows no interaction with. To sum up, these facts provide an explanation for the ability of PGRPs to discriminate Lys-type from Dap-type PGN.

Protein	MPP	MPP-Dap	TCT	CL-PGN
h-PGRP-IαC	45.0 (±1.0)	NB	20.2 (±0.8)	74.2 (±4.6)
h-PGRP-S	6.3 (±0.4)	47.4 (±6.0)	23.4 (±1.9)	NB
d-PGRP-LCx	NB†	63.5 (±4.7)	19.8 (±0.8)	NB
d-PGRP-LCa	17.8 (±1.5)	NB	8.7 (±1.6)	NB
h-PGRP-S (G68N, W69F)	85.1 (±0.4)	3.6 (±0.9)	45.4 (±2.1)	85.2 (±8.1)
d-PGRP-LCx (G393N, W394F)	23.8 (±1.3)	NB	18.8 (±1.8)	NB
d-PGRP-LCa (Q412N, K413F)	36.3 (±1.7)	72.2 (±14.0)	33.8 (±1.9)	83.0 (±9.6)

Table.4 Binding constants of PGN dervatives to different PGRPs at 2 ℃. *h*- homo species; *d*- *Drosophila*. MPP, monomeric Muramyl pentapeptides. TCT, natural Dap-type PGN fragment. CL-PGN, cross-linked Lys-type PGN (adapted from Ref. 148).

Although the binding of PGN and its derivatives to PGRPs are highly selective, all PGRPs show a significant low affinity (all the K_M >$10^5 M^{-1}$). The enzymes combining high selectivity and low affinity have been observed previously, *i.e.*, the recognition of peptides or MHC ligands by T-cells[149] and carbohydrate- protein interaction[150]. The affinity of binding of PGN to PGRPs can be improved by dimerization or oligomerization of PGRPs, a phenomenon which has been observed in PGRP-LC of *Drosophila* and PGLYRP-3 (or PGLYRP-4) of man. The PGRP-LC in *Drosophila* has three isoforms, named PGRP-LCa, PGRP-LCx and PGRP-LCy. From Figure.31 we can see that each of these three PGRP-Ls contains different specificity determining residues, which indicates that they probably possess different specificities. Human PGLYRP-3 and PGLYRP-4 contain two PGN binding domains, PGRP-IαC / PGRP-IαN and PGRP-IβC / PGRP-IβN, respectively. All of these PGN binding domains contain different specificity-determining residues (Figure.31). By forming homo- or heterodimers, the diversification of the binding specificity can be improved. Furthermore, the host can finely modify the specificity of PGLYRPs by expressing PGLYRP-3 and PGLYRP-4 at different levels (Figure.32).

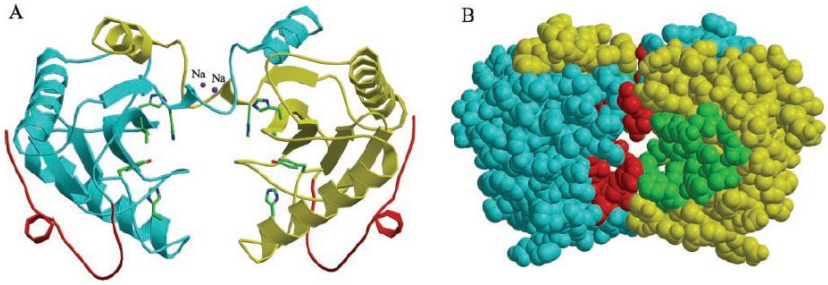


Figure.32 Dimerizaiton of PGLYRP-3 (PGRP-Iα). (A) Ribbon diagram of a PGLYRP-3 homodimer with one domain is shown in blue, the other is in yellow. Na^+ is necessary for the dimerization. The PGRP specific segment

is marked in red, while residues involved in PGN binding are green. (B) Space filling diagram of a PGLYPR-3 homodimer. One domain is blue, the other is yellow. Residues involved in PGN binding are depicted by red and green in each domain (adapted from Ref. 145).

PGRPs are well conserved from insect to human and they might represent one of the most ancient antimicrobal mechanisms. To discriminate various PGN derivatives, the hosts evolved two strategies: one is the variation in specificity determining residues, the other is the recognition of different cross-bridges. In animals, the PGRPs are also involved in some other immune pathways, which is indicated by the fact that PGRPs in bovine kill microorganisms completely lack a cell wall. To sum up, although there is quite some findings about the mechanisms of binding of PGN to PGRPs, the physiological roles of the PGRPs are still less understood.

Discussion

Gram-positive bacteria present numerous of proteins on their surface, many of which are critical for surviving or play an important role in virulence. Four types of surface proteins are currently documented (Figure.33): (1) the proteins contain a hydrophobic transmembrane domain(s), (2) the proteins covalently linked to membrane lipids, which are also called lipoproteins, (3) the proteins covalently linked to PGN wall and (4) the proteins non-covalently attached to PGN wall (or SCWP)[151]. These proteins target to the cell surface via different mechanism. This essay reviewed the current knowledge about the proteins that non-covalently link to PGN wall, and tried to answer the question: how do these proteins with different origins recognize/bind to the PGN wall.

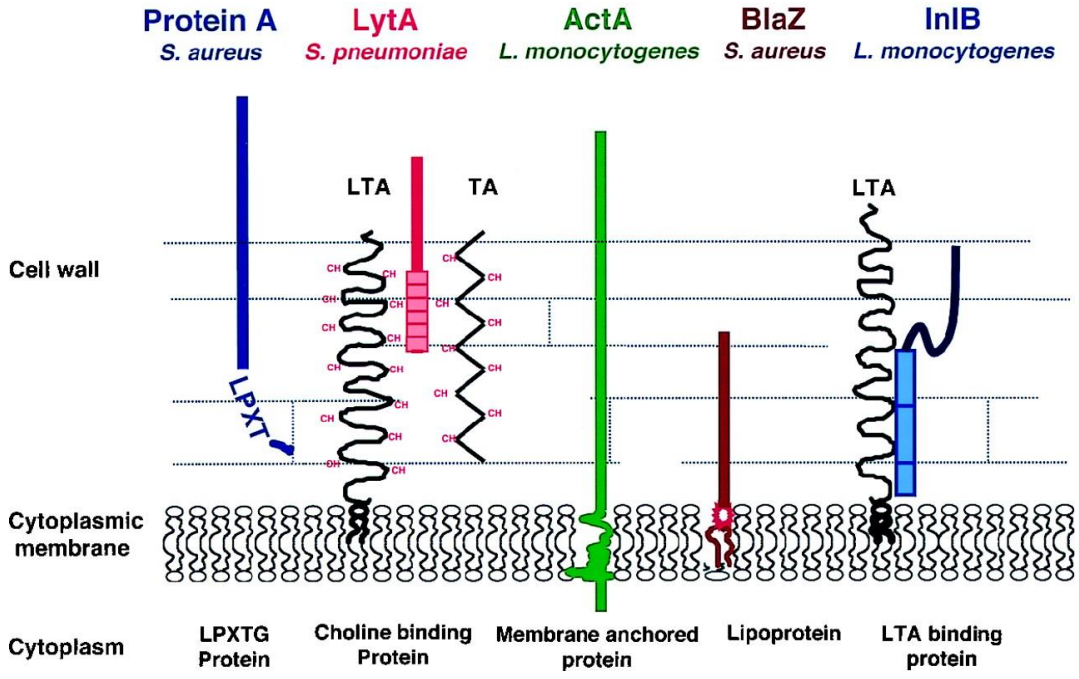


Figure. 33 Major types of surface proteins in Gram-positive bacteria. An example of each type is given. (Adapted from Ref.151)

PGN wall is a dynamic structure and functions as a scaffold where proteins anchor on. This structure expands during cell growth and can be cleaved during cell division. Cell wall structures in different growth stages also represent different properties. For instance, the PGN structure of *B.subtilis* in stationary phase is more cross-linked than that in exponential phase *B.subtilis*[152]. Moreover, modifications or variations of PGN structure are observed throughout species. These modifications can be found in disaccharide backbone, cross bridge region or peptide stem structure, some of which are responsible for antibiotic resistance. For example, the acetyl group in *Streptococcus pneumonia* and *Listeria monocytogenes* are removed, which are responsible for lysozyme resistance[153, 154]. To modify the PGN wall, bacterial cells employ a series of enzymes, which often have one or more binding motif/domains. These binding motifs are necessary for localizing the protein to the cell surface. Beside PGN, there are also other components in cell wall, such as SCWPs, which may function as a ligand for binding various proteins.

Many cell wall binding proteins are highly strain/species specific and usually consist of two main parts: the catalytic domain and the cell wall binding domain[5]. The high specificity and selectivity of these proteins are given by the cell wall binding domain, for instance, the LysM domain and the choline binding domain. Furthermore, a decent catalytic activity requires the strong binding affinity contributed by the binding domain. For example, the *Bacillus anthracis* phage lysin PlyB loses its lytic activity when the GW repeat binding domain is absent[6].

On the other hand, the Gram-positive bacteria have also evolved many ways to avoid lysins or host immune defense. In *Streptococcus pneumoniae*, the positively charged choline moieties are added to the TA; In *Bacillus subtilis*, the head group of the phospholipid is modified[7]; In *Bacillus cereus*, the N-acetyl moieties in the disaccharide backbone of PGN are N-deacetylated, which makes the host cells resistant to human lysozyme[8]. Some Gram-positive bacteria are surrounded by S-layer proteins that are non-covalently linked to SCWPs[9]. The general strategy of protecting the cells from various murein hydrolases is to reduce the negative charge given by phospholipid and SCWPs.

Prokaryotes and eukaryotes use different strategies to recognize PGN structure. Enzymes produced by prokaryotes often target on the species-specific structure. The lysostaphin produced by *S.simulans* is an interesting example, which has a binding domain structurally similar to GW repeat module in prokaryotes and SH3 domain in eukaryotes. GW repeat modules are usually found in amidase and responsible for binding LTA, but lysostaphin specifically recognize the crossbridge structure (pentaglycine) in *S.aureus*. Thus, the *S.simulans* produced lysostaphin specifically lyse *S.aureus* but spare its host cells. However, the PGRPs in eukaryotes have a hydrophobic groove binding to the MurNAc moiety and the peptide stem, which are the basic structures of PGN wall. This binding mechanism can make sure that the immune system in eukaryotes can defense all the PGN structures.

References

1. Schleife.KH, Kandler O. Peptidoglycan types of bacterial cell-walls and their taxonomic implications. *Bacteriol Rev*. 1972;36(4):407-477.

2. Oyston PCF, Fox MA, Richards SJ, Clark GC. Novel peptide therapeutics for treatment of infections. *J Med Microbiol*. 2009;58(8):977-987.

3. Lambert P. Cellular impermeability and uptake of biocides and antibiotics in Gram-positive bacteria and mycobacteria. *J Appl Microbiol*. 2002;92:46S-54S.

4. Demchick P, Koch A. The permeability of the wall fabric of *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol*. 1996;178(3):768-773.

5. Loessner MJ. Bacteriophage endolysins — current state of research and applications. *Curr Opin Microbiol.* 2005;8(4):480-487.

6. Porter CJ, Schuch R, Pelzek AJ, et al. The 1.6 Å crystal structure of the catalytic domain of PlyB, a bacteriophage lysin active against *Bacillus anthracis*. *J Mol Biol*. 2007;366(2):540-550.

7. Roy H. Tuning the properties of the bacterial membrane with aminoacylated phosphatidylglycerol. *IUBMB Life*. 2009;61(10):940-953.

8. Severin A, Tabei K, Tomasz A. The structure of the cell wall peptidoglycan of *Bacillus cereus* RSVF1, a strain closely related to *Bacillus anthracis*. *Microb Drug Resist -Mechan Epidemiol Dis*. 2004;10(2):77-82.

9. Kern J, Ryan C, Faull K, Schneewind O. *Bacillus anthracis* surface-layer proteins assemble by binding to the secondary cell wall polysaccharide in a manner that requires *csaB* and *tagO*. *J Mol Biol*. 2010;401(5):757-775.

10. Perez-Dorado I, Campillo NE, Monterroso B, et al. Elucidation of the molecular recognition of bacterial cell wall by modular pneumococcal phage endolysin CPL-1. *J Biol Chem*. 2007;282(34):24990-24999.

11. Kumar S, Roychowdhury A, Ember B, et al. Selective recognition of synthetic lysine and meso-diaminopimelic acid-type peptidoglycan fragments by human peptidoglycan recognition proteins I alpha and S. *J Biol Chem*. 2005;280(44):37005-37012.

12. Swaminathan C, Brown P, Roychowdhury A, et al. Dual strategies for peptidoglycan discrimination by peptidoglycan recognition proteins (PGRPs). *Proc Natl Acad Sci U S A*. 2006;103(3):684-689.

13. Guan R, Mariuzza RA. Peptidoglycan recognition proteins of the innate immune system. *Trends Microbiol*. 2007;15(3):127-134.

14. Witzenrath M, Schmeck B, Doehn JM, et al. Systemic use of the endolysin Cpl-1 rescues mice with fatal pneumococcal pneumonia. *Crit Care Med*. 2009;37(2):642-649.

15. Errington J. L-form bacteria, cell walls and the origins of life. *Open Biol*. 2013;3:120143.

16. Typas A, Banzhaf M, Gross CA, Vollmer W. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol*. 2012;10(2):123-136.

17. Navarre W, Schneewind O. Surface proteins of Gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev*. 1999;63(1):174-+.

18. Bierne H, Mazmanian S, Trost M, et al. Inactivation of the *srtA* gene in listeria monocytogenes inhibits anchoring of surface proteins and affects virulence. *Mol Microbiol*. 2002;43(4):869-881.

19. Gaspar A, Marraffini L, Glass E, DeBord K, Ton-That H, Schneewind O. *Bacillus anthracis* sortase A (SrtA) anchors LPXTG motif-containing surface proteins to the cell wall envelope. *J Bacteriol*. 2005;187(13):4646-4655.

20. Barnett T, Scott J. Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J Bacteriol*. 2002;184(8):2181-2191.

21. Bolken T, Franke C, Jones K, et al. Inactivation of the *srtA* gene in *Streptococcus gordonii* inhibits cell wall anchoring of surface proteins and decreases in vitro and in vivo adhesion. *Infect Immun*. 2001;69(1):75-80.

22. Kharat A, Tomasz A. Inactivation of the *srtA* gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells in vitro. *Infect Immun*. 2003;71(5):2758-2765.

23. Lee S, Boran T. Roles of sortase in surface expression of the major protein adhesin P1, saliva-induced aggregation and adherence, and cariogenicity of *Streptococcus mutans*. *Infect Immun*. 2003;71(2):676-681.

24. Lalioui L, Pellegrini E, Dramsi S, et al. The SrtA sortase of *Streptococcus agalactiae* is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine. *Infect Immun*. 2005;73(6):3342-3350.

25. Boekhorst J, de Been M, Kleerebezem M, Siezen R. Genome-wide detection and analysis of cell wall-bound proteins with LPXTG-like sorting motifs. *J Bacteriol*. 2005;187(14):4928-4934.

26. Giesbrecht P, Wecke J, Reinicke B. Morphogenesis of cell-wall of staphylococci. *Int Rev Cytol*. 1976;44:225-318.

27. Hancock I. Bacterial cell surface carbohydrates: Structure and assembly. *Biochem Soc Trans*. 1997;25(1):183-187.

28. Ghuysen J. Use of bacteriolytic enzymes in determination of wall structure and their role in cell metabolism. *Bacteriol Rev*. 1968;32(4P2):425-&.

29. Coley J, Archibald A, Baddiley J. Linkage unit joining peptidoglycan to teichoic-acid in *Staphylococcus aureus* H. *FEBS Lett*. 1976;61(2):240-242.

30. Ton-That H, Schneewind O. Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol*. 2003;50(4):1429-1438.

31. Schneewind O, Fowler A, Faull K. Structure of the cell-wall anchor of surface-proteins in *Staphylococcus-aureus*. *Science*. 1995;268(5207):103-106.

32. Mesnage S, Fontaine T, Mignot T, Delepierre M, Mock M, Fouet A. Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J*. 2000;19(17):4473-4484.

33. Schneewind O, Missiakas DM. Protein secretion and surface display in gram-positive bacteria. *Philos Trans R Soc B-Biol Sci*. 2012;367(1592):1123-1139.

34. Ghuysen J, Strominger J. Structure of cell wall of *Staphylococcus aureus*, strain copenhagen .2. separation and structure of disaccharides. *Biochemistry (N Y)*. 1963;2(5):1119-&.

35. Glauner B, Holtje J, Schwarz U. The composition of the murein of *Escherichia-coli*. *J Biol Chem*. 1988;263(21):10088-10095.

36. Henze U, Sidw T, Wecke J, Labischinski H, Bergerbachi B. Influence of *femb* on methicillin resistance and peptidoglycan metabolism in *Staphylococcus-aureus*. *J Bacteriol*. 1993;175(6):1612-1620.

37. Tipper D, Stroming.JL, Ensign J. Structure of cell wall of *Staphylococcus aureus* strain copenhagen .7. mode of action of bacteriolytic peptidase from myxobacter and isolation of intact cell wall polysaccharides. *Biochemistry (N Y)*. 1967;6(3):906-&.

38. Tipper D, Berman M. Structures of cell wall peptidoglycans of *Staphylococcus epidermidis* texas-26 and *Staphylococcus aureus* copenhagen .I. chain length and average sequence of cross-bridge peptides. *Biochemistry (N Y)*. 1969;8(5):2183-&.

39. Tipper D. Structures of cell wall peptidoglycans of *Staphylococcus epidermidis* texas-26 and *Staphylococcus aureus* copenhagen .2. structure of neutral and basic peptides from hydrolysis with myxobacter al-1 peptidase. *Biochemistry (N Y)*. 1969;8(5):2192-&.

40. Tipper D, Stroming.JL. Biosynthesis of peptidoglycan of bacterial cell walls .12. inhibition of cross-linking by penicillins and cephalosporins - studies in *Staphylococcus aureus* in vivo. *J Biol Chem*. 1968;243(11):3169-&.

41. Strominger JL. Penicillin-sensitive enzymatic reactions in bacterial cell wall synthesis. *Harvey Lect*. 1968;64:179-213.

42. Stroming.JL, Izakik, Matsuhas.M Tipper D. Peptidoglycan transpeptidase and D-alanine carboxypeptidase - penicillin-sensitive enzymatic reactions. *Fed Proc*. 1967;26(1):9-&.

43. Gunetile.KG, Anwar R. Biosynthesis of uridine diphospho-N-acetyl muramic acid. *J Biol Chem*. 1966;241(23):5740-&.

44. Ito E, Strominger J. Enzymatic synthesis of peptide in bacterial uridine nucleotides .3. purification + properties of L-lysine-adding enzyme. *J Biol Chem*. 1964;239(1):210-&.

45. Mizuno Y, Yaegashi M, Ito E. Purification and properties of uridine diphosphate N-acetylmuramate - L-alanine ligase. *J Biochem*. 1973;74(3):525-538.

46. Anderson J, Stroming.JL. Isolation and utilization of phospholipid intermediates in cell wall glycopeptide synthesis. *Biochem Biophys Res Commun*. 1965;21(5):516-&.

47. Higashi Y, Siewert G, Stroming.JL. Biosynthesis of peptidoglycan of bacterial cell walls .19. isoprenoid alcohol phosphokinase. *J Biol Chem*. 1970;245(14):3683-&.

48. Higashi Y, Stroming.JL. Biosynthesis of peptidoglycan of bacterial cell walls .21. isolation of free C_{55}-isoprenoid alcohol and of lipid intermediates in peptidoglycan synthesis from *Staphylococcus-aureus*. *J Biol Chem*. 1970;245(14):3697-&.

49. Anderson J, Matsuhas.M, Haskin M, Stroming.JL. Biosynthesis of peptidoglycan of bacterial cell walls .2. phospholipid carriers in reaction sequence. *J Biol Chem*. 1967;242(13):3180-&.

50. Ghuysen J Serine beta-lactamases and penicillin-binding proteins. *Annu Rev Microbiol*. 1991;45:37-67.

51. Goffin C, Ghuysen J. Multimodular penicillin binding proteins: An enigmatic family of orthologs and paralogs. *Microbiol Mol Biol Rev*. 1998;62(4):1079-+.

52. Nakagawa J, Tamaki S, Tomioka S, Matsuhashi M. Functional biosynthesis of cell-wall peptidoglycan by polymorphic bifunctional polypeptides - penicillin-binding protein 1bs of *Escherichia-coli* with activities of transglycosylase and transpeptidase. *J Biol Chem*. 1984;259(22):3937-3946.

53. Tipper D, Stroming.JL. Mechanism of action of penicillins - a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A*. 1965;54(4):1133-&.

54. Stroming.JL, Ghuysen J. Mechanisms of enzymatic bacteriolysis. *Science*. 1967;156(3772):213-&.

55. Schaffer C, Messner P. The structure of secondary cell wall polymers: How Gram-positive bacteria stick their cell walls together. *Microbiology-(UK)*. 2005;151:643-651.

56. Armstrong J, Baddiley J, Buchanan J, Carss B. Nucleotides and the bacterial cell wall. *Nature*. 1958;181(4625):1692-1693.

57. Araki Y, Ito E. Linkage units in cell-walls of Gram-positive bacteria. *Crit Rev Microbiol*. 1989;17(2):121-135.

58. Silhavy TJ, Kahne D, Walker S. The bacterial cell envelope. *Cold Spring Harbor Perspect Biol*. 2010;2(5):a000414.

59. Herbold D, Glaser L. Interaction of N-acetylmuramic acid L-alanine amidase with cell-wall polymers. *J Biol Chem*. 1975;250(18):7231-7238.

60. Holtje J, Tomasz A. Specific recognition of choline residues in cell-wall teichoic-acid by N-acetylmuramyl-L-alanine amidase of pneumococcus. *J Biol Chem*. 1975;250(15):6072-6076.

61. Araki Y, Ito E. Linkage units in cell-walls of Gram-positive bacteria. *Crit Rev Microbiol*. 1989;17(2):121-135.

62. Doyle R, Sonnenfeld E. Properties of the cell-surfaces of pathogenic bacteria. *Int Rev Cytol*. 1989;118:33-92.

63. Jonquieres R, Pizarro-Cerda J, Cossart P. Synergy between the N- and C-terminal domains of InIB for efficient invasion of non-phagocytic cells by listeria monocytogenes. *Mol Microbiol*. 2001;42(4):955-965.

64. Ponting C, Aravind L, Schultz J, Bork P, Koonin E. Eukaryotic signalling domain homologues in Archaea and bacteria. ancient ancestry and horizontal gene transfer. *J Mol Biol*. 1999;289(4):729-745.

65. Marino M, Banerjee M, Jonquieres R, Cossart P, Ghosh P. GW domains of the listeria monocytogenes invasion protein InlB are SH3-like and mediate binding to host ligands. *EMBO J*. 2002;21(21):5623-5634.

66. Kleerebezem M, Boekhorst J, van Kranenburg R, et al. Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci U S A*. 2003;100(4):1990-1995.

67. Paulsen I, Banerjei L, Myers G, et al. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science*. 2003;299(5615):2071-2074.

68. Brinster S, Furlan S, Serror P. C-terminal WxL domain mediates cell wall binding in *Enterococcus faecalis* and other Gram-positive bacteria. *J Bacteriol*. 2007;189(4):1244-1253.

69. Schachtsiek M, Hammes W, Hertel C. Characterization of *Lactobacillus coryniformis* DSM 20001(T) surface protein cpf mediating coaggregation with and aggregation among pathogens. *Appl Environ Microbiol*. 2004;70(12):7078-7085.

70. Joris B, Englebert S, Chu C, et al. Modular design of the *Enterococcus-hirae* muramidase-2 and *Streptococcus faecalis* autolysin. *FEMS Microbiol Lett*. 1992;91(3):257-264.

71. Milohanic E, Jonquieres R, Cossart P, Berche P, Gaillard J. The autolysin ami contributes to the adhesion of *Listeria monocytogenes* to eukaryotic cells via its cell wall anchor. *Mol Microbiol*. 2001;39(5):1212-1224.

72. Steen A, Buist G, Horsburgh G, et al. AcmA of *Lactococcus lactis* is an N-acetylglucosaminidase with an optimal number of LysM domains for proper functioning. *Febs J*. 2005;272(11):2854-2868.

73. Sloan G, Robinson J, Kloos W. Identification of *Staphylococcus staphylolyticus* nrrl B-2628 as a biovar of *Staphylococcus simulans*. *Int J Syst Bacteriol*. 1982;32(2):170-174.

74. Sugai M, Fujiwara T, Ohta K, Komatsuzawa H, Ohara M, Suginaka H. Epr, which encodes glycylglycine endopeptidase resistance, is homologous to *femAB* and affects serine content of peptidoglycan cross bridges in *Staphylococcus capitis* and *Staphylococcus aureus*. *J Bacteriol*. 1997;179(13):4311-4318.

75. Kerr D, Plaut K, Bramley A, et al. Lysostaphin expression in mammary glands confers protection against staphylococcal infection in transgenic mice. *Nat Biotechnol*. 2001;19(1):66-70.

76. RECSEI P, GRUSS A, NOVICK R. Cloning, sequence, and expression of the lysostaphin gene from *Staphylococcus simulans*. *Proc Natl Acad Sci U S A*. 1987;84(5):1127-1131.

77. Baba T, Schneewind O. Target cell specificity of a bacteriocin molecule: A C-terminal signal directs lysostaphin to the cell wall of *Staphylococcus aureus*. *EMBO J*. 1996;15(18):4789-4797.

78. Whisstock J, Lesk A. SH3 domains in prokaryotes. *Trends Biochem Sci*. 1999;24(4):132-133.

79. Pawson T. Protein modules and signaling networks. *Nature*. 1995;373(6515):573-580.

80. Ponting C, Aravind L, Schultz J, Bork P, Koonin E. Eukaryotic signalling domain homologues in archaea and bacteria. ancient ancestry and horizontal gene transfer. *J Mol Biol*. 1999;289(4):729-745.

81. Lu J, Fujiwara T, Komatsuzawa H, Sugai M, Sakon J. Cell wall-targeting domain of glycylglycine endopeptidase distinguishes among peptidoglycan cross-bridges. *J Biol Chem*. 2006;281(1):549-558.

82. Sugai M, Fujiwara T, Akiyama T, et al. Purification and molecular characterization of glycylglycine endopeptidase produced by *Staphylococcus capitis* EPK1. *J Bacteriol*. 1997;179(4):1193-1202.

83. Sugai M, Fujiwara T, Akiyama T, et al. Purification and molecular characterization of glycylglycine endopeptidase produced by *Staphylococcus capitis* EPK1. *J Bacteriol*. 1997;179(4):1193-1202.

84. Thumm G, Gotz F. Studies on prolysostaphin processing and characterization of the lysostaphin immunity factor (lif) of *Staphylococcus simulans* biovar staphylolyticus. *Mol Microbiol*. 1997;23(6):1251-1265.

85. Ehlert K, Tschierske M, Mori C, Schroder W, Berger-Bachi B. Site-specific serine incorporation by *lif* and *epr* into positions 3 and 5 of the staphylococcal peptidoglycan interpeptide bridge. *J Bacteriol*. 2000;182(9):2635-2638.

86. Hegde S, Shrader T. *FemABX* family members are novel nonribosomal peptidyltransferases and important pathogen-specific drug targets. *J Biol Chem*. 2001;276(10):6998-7003.

87. Ehlert K, Schroder W, Labischinski H. Specificities of FemA and FemB for different glycine residues: FemB cannot substitute for FemA in staphylococcal peptidoglycan pentaglycine side chain formation. *J Bacteriol*. 1997;179(23):7573-7576.

88. Holm L, Sander C. Mapping the protein universe. *Science*. 1996;273(5275):595-602.

89. Cesareni G, Panni S, Nardelli G, Castagnoli L. Can we infer peptide recognition specificity mediated by SH3 domains? *FEBS Lett*. 2002;513(1):38-44.

90. Hirakawa H, Akita H, Fujiwara T, Sugai M, Kuhara S. Structural insight into the binding mode between the targeting domain of ALE-1 (92AA) and pentaglycine of peptidoglycan. *Protein Eng Des Sel*. 2009;22(7):385-391.

91. Engelhardt H. Are S-layers exoskeletons? The basic function of protein surface layers revisited. *J Struct Biol*. 2007;160(2):115-124.

92. Navarre W, Schneewind O. Surface proteins of Gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev*. 1999;63(1):174-+.

93. Lupas A, Engelhardt H, Peters J, Santarius U, Volker S, Baumeister W. Domain-structure of the acetogenium-kivui surface-layer revealed by electron crystallography and sequence-analysis. *J Bacteriol*. 1994;176(5):1224-1233.

94. Kern J, Ryan C, Faull K, Schneewind O. *Bacillus anthracis* surface-layer proteins assemble by binding to the secondary cell wall polysaccharide in a manner that requires *csaB* and *tagO*. *J Mol Biol*. 2010;401(5):757-775.

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

96. Mesnage S, Tosi-Couture E, Fouet A. 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*. 1999;31(3):927-936.

97. May A, Pusztahelyi T, Hoffmann N, Fischer R, Bahl H. Mutagenesis of conserved charged amino acids in SLH domains of *Thermoanaerobacterium thermosulfurigenes* EM1 affects attachment to cell wall sacculi. *Arch Microbiol*. 2006;185(4):263-269.

98. Kern J, Wilton R, Zhang R, Binkowski TA, Joachimiak A, Schneewind O. Structure of surface layer homology (SLH) domains from *Bacillus anthracis* surface array protein. *J Biol Chem*. 2011;286(29):26042-26049.

99. Lupas A. A circular permutation event in the evolution of the SLH domain? *Mol Microbiol*. 1996;20(4):897-898.

100. Mesnage S, TosiCouture E, Mock M, Gounon P, Fouet A. Molecular characterization of the *Bacillus anthracis* main S-layer component: Evidence that it is the major cell-associated antigen. *Mol Microbiol*. 1997;23(6):1147-1155.

101. Garcia J, Sanchez-Beato A, Medrano F, Lopez R. Versatility of choline-binding domain. *Microb Drug Resist -Mechan Epidemiol Dis*. 1998;4(1):25-36.

102. Horne D, Tomasz A. Possible role of a choline-containing teichoic-acid in the maintenance of normal-cell shape and physiology in *Streptococcus-oralis*. *J Bacteriol*. 1993;175(6):1717-1722.

103. Gillespie S, Ainscough S, Dickens A, Lewin J. Phosphorylcholine-containing antigens in bacteria from the mouth and respiratory tract. *J Med Microbiol*. 1996;44(1):35-40.

104. SanchezBeato A, Garcia J. Molecular characterization of a family of choline-binding proteins of *Clostridium beijerinckii* NCIB 8052. evolution and gene redundancy in prokaryotic cell. *Gene*. 1996;180(1-2):13-21.

105. Sanchezbeato A, Ronda C, Garcia J. Tracking the evolution of the bacterial choline-binding domain - molecular characterization of the *Clostridium-acetobutylicum* ncib-8052 cspa gene. *J Bacteriol*. 1995;177(4):1098-1103.

106. Lopez R, Gonzalez M, Garcia E, Garcia J, Garcia P. Biological roles of two new murein hydrolases of *Streptococcus pneumoniae* representing examples of module shuffling. *Res Microbiol*. 2000;151(6):437-443.

107. Kolberg J, Hoiby E, Jantzen E. Detection of the phosphorylcholine epitope in streptococci, haemophilus and pathogenic neisseriae by immunoblotting. *Microb Pathog*. 1997;22(6):321-329.

108. Tettelin H, Nelson K, Paulsen I, et al. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science*. 2001;293(5529):498-506.

109. Lopez R, Garcia E. Recent trends on the molecular biology of *Pneumococcal capsules*, lytic enzymes, and bacteriophage. *FEMS Microbiol Rev*. 2004;28(5):553-580.

110. Garcia E, Garcia J, Ronda C, Garcia P, Lopez R. Cloning and expression of the pneumococcal autolysin gene in *Escherichia-coli*. *Mol Gen Genet*. 1985;201(2):225-230.

111. Tomasz A, Westphal M. Abnormal autolytic enzyme in a pneumococcus with altered teichoic acid composition. *Proc Natl Acad Sci U S A*. 1971;68(11):2627-&.

112. Garcia P, Gonzalez P, Garcia E, Lopez R, Garcia J. LytB, a novel pneumococcal murein hydrolase essential for cell separation. *Mol Microbiol*. 1999;31(4):1275-1277.

113. Garcia P, Gonzalez M, Garcia E, Garcia J, Lopez R. The molecular characterization of the first autolytic lysozyme of *Streptococcus pneumoniae* reveals evolutionary mobile domains. *Mol Microbiol*. 1999;33(1):128-138.

114. Tomasz A. Biological consequences of replacement of choline by ethanolamine in cell wall of pneumococcus - chain formation loss of transformability and loss of autolysis. *Proc Natl Acad Sci U S A*. 1968;59(1):86-&.

115. Garcia E, Garcia J, Garcia P, Arraras A, Sanchezpuelles J, Lopez R. Molecular evolution of lytic enzymes of *Streptococcus-pneumoniae* and its bacteriophages. *Proc Natl Acad Sci U S A*. 1988;85(3):914-918.

116. Jedrzejas M. Pneumococcal virulence factors: Structure and function. *Microbiol Mol Biol Rev*. 2001;65(2):187-+.

117. Garcia J, Diaz E, Romero A, Garcia P. Carboxy-terminal deletion analysis of the major pneumococcal autolysin. *J Bacteriol*. 1994;176(13):4066-4072.

118. De las Rivas B, Garcia J, Lopez R, Garcia P. Purification and polar localization of pneumococcal LytB, a putative endo-beta-N-acetylglucosaminidase: The chain-dispersing murein hydrolase. *J Bacteriol*. 2002;184(18):4988-5000.

119. Garciabustos J, Tomasz A. Teichoic acid-containing muropeptides from *Streptococcus-pneumoniae* as substrates for the pneumococcal autolysin. *J Bacteriol*. 1987;169(2):447-453.

120. Yother J, Leopold K, White J, Fischer W. Generation and properties of a *Streptococcus pneumoniae* mutant which does not require choline or analogs for growth. *J Bacteriol*. 1998;180(8):2093-2101.

121. Severin A, Horne D, Tomasz A. Autolysis and cell wall degradation in a choline-independent strain of *Streptococcus pneumoniae*. *Microb Drug Resist -Mechan Epidemiol Dis*. 1997;3(4):391-400.

122. Fernandez-Tornero C, Lopez R, Garcia E, Gimenez-Gallego G, Romero A. A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. *Nat Struct Biol*. 2001;8(12):1020-1024.

123. Usobiaga P, Medrano F, Gasset M, et al. Structural organization of the major autolysin from *Streptococcus pneumoniae*. *J Biol Chem*. 1996;271(12):6832-6838.

124. Varea J, Saiz J, Lopez-Zumel C, et al. Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? *J Biol Chem*. 2000;275(35):26842-26855.

125. Fernandez-Tornero C, Garcia E, Lopez R, Gimenez-Gallego G, Romero A. Two new crystal forms of the choline-binding domain of the major pneumococcal autolysin: Insights into the dynamics of the active homodimer. *J Mol Biol*. 2002;321(1):163-173.

126. Dziarski R, Gupta D. *Staphylococcus aureus* peptidoglycan is a toll-like receptor 2 activator: A reevaluation. *Infect Immun*. 2005;73(8):5212-5216.

127. Dziarski R, Tapping R, Tobias P. Binding of bacterial peptidoglycan to CD14. *J Biol Chem*. 1998;273(15):8680-8690.

128. Dziarski R, Platt K, Gelius E, Steiner H, Gupta D. Defect in neutrophil killing and increased susceptibility to infection with-nonpathogenic Gram-positive bacteria in peptidoglycan recognition protein-S (PGRP-S)-deficient mice. *Blood*. 2003;102(2):689-697.

129. Yoshida H, Kinoshita K, Ashida M. Purification of a peptidoglycan recognition protein from hemolymph of the *silkworm, Bombyx mori*. *J Biol Chem*. 1996;271(23):13854-13860.

130. Dziarski R, Gupta D. The peptidoglycan recognition proteins (PGRPs). *Genome Biol*. 2006;7(8):232.

131. Christophides G, Zdobnov E, Barillas-Mury C, et al. Immunity-related genes and gene families in *Anopheles gambiae*. *Science*. 2002;298(5591):159-165.

132. Werner T, Liu G, Kang D, Ekengren S, Steiner H, Hultmark D. A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*. 2000;97(25):13772-13777.

133. Hoffmann J. The immune response of Drosophila. *Nature*. 2003;426(6962):33-38.

134. Yoshida H, Kinoshita K, Ashida M. Purification of a peptidoglycan recognition protein from hemolymph of the *silkworm, Bombyx mori*. *J Biol Chem*. 1996;271(23):13854-13860.

135. Takehana A, Yano T, Mita S, Kotani A, Oshima Y, Kurata S. Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in drosophila immunity. *EMBO J*. 2004;23(23):4690-4700.

136. Park J, Je B, Piao S, et al. A synthetic peptidoglycan fragment as a competitive inhibitor of the melanization cascade. *J Biol Chem*. 2006;281(12):7747-7755.

137. Dziarski R, Gupta D. The peptidoglycan recognition proteins (PGRPs). *Genome Biol*. 2006;7(8):232.

138. Gelius E, Persson C, Karlsson J, Steiner H. A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine amidase activity. *Biochem Biophys Res Commun*. 2003;306(4):988-994.

139. Wang Z, Li X, Cocklin R, et al. Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase. *J Biol Chem*. 2003;278(49):49044-49052.

140. Medzhitov R, Janeway C. Decoding the patterns of self and nonself by the innate immune system. *Science*. 2002;296(5566):298-300.

141. Schleife.KH, Kandler O. Peptidoglycan types of bacterial cell-walls and their taxonomic implications. *Bacteriol Rev*. 1972;36(4):407-477.

142. Liu C, Gelius E, Liu G, Steiner H, Dziarski R. Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth. *J Biol Chem*. 2000;275(32):24490-24499.

143. Tydell C, Yuan J, Tran P, Selsted M. Bovine peptidoglycan recognition protein-S: Antimicrobial activity, localization, secretion, and binding properties. *J Immunol*. 2006;176(2):1154-1162.

144. Cho J, Fraser I, Fukase K, et al. Human peptidoglycan recognition protein S is an effector of neutrophil-mediated innate immunity. *Blood*. 2005;106(7):2551-2558.

145. Guan R, Malchiodi E, Wang Q, Schuck P, Mariuzza R. Crystal structure of the C-terminal peptidoglycan-binding domain of human peptidoglycan recognition protein I alpha. *J Biol Chem*. 2004;279(30):31873-31882.

146. Kim M, Byun M, Oh B. Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster*. *Nat Immunol*. 2003;4(8):787-793.

147. Guan R, Wang Q, Sundberg E, Mariuzza R. Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 angstrom resolution. *J Mol Biol*. 2005;347(4):683-691.

148. Swaminathan C, Brown P, Roychowdhury A, et al. Dual strategies for peptidoglycan discrimination by peptidoglycan recognition proteins (PGRPs). *Proc Natl Acad Sci U S A*. 2006;103(3):684-689.

149. Kersh E, Shaw A, Allen P. Fidelity of T cell activation through multistep T cell receptor zeta phosphorylation. *Science*. 1998;281(5376):572-575.

150. Swaminathan C, Wais N, Vyas V, et al. Entropically assisted carbohydrate recognition by a natural killer cell-surface receptor. *Chembiochem*. 2004;5(11):1571-1575.

151. Cossart P , and Jonquières R. Sortase, a universal target for therapecutic agents against Gram-positive bacteria? PNAS 2000;97:5013-5015.

152. Atrih A, Bacher G, Allmaier G, Williamson MP, Foster SJ: Analysis of peptidoglycan structure from vegetative cells of *Bacillus subtilis* 168 and role of PBP 5 in peptidoglycan maturation. J Bacteriol 1999; 181: 3956–3966.

153. Vollmer W, Tomasz A: Peptidoglycan Nacetylglucosamine deacetylase, a putative virulence factor in *Streptococcus pneumoniae*. Infect Immun 2002; 70: 7176–7178.

154. Boneca IG, Dussurget O, Cabanes D, Girardin SE: A critical role for peptidoglycan N-deacetylation in *Listeria* evasion from the host innate immune system. Proc Natl Acad Sci USA 2007; 104: 997–1002.