

# **Structures and functions of proteins that utilize and modify Wall Teichoic Acid**

**Dissertation**

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Cengiz Koç

aus Hanau / Hessen

Tübingen

2016

**Tag der mündlichen Qualifikation: Oktober/November 2016**

**Dekan: Prof. Dr. Wolfgang Rosenstiel**

**1. Berichterstatter: Prof. Dr. Thilo Stehle**

**2. Berichterstatter: PD Dr. Frank Essmann**

# 1 Contents

I	ABBREVIATIONS.....	C
II	Summary (english).....	E
III	Summary (german) .....	F
IV	List of publications .....	H
V	Personal contribution .....	I
2	INTRODUCTION .....	1
2.1	Gram-positive Bacteria .....	1
2.1.1	<i>Staphylococci</i> .....	2
2.1.1.1	$\beta$ -hemolysis and Coagulase-negative <i>Staphylococci</i> .....	2
2.1.1.2	MRSA & MSSA .....	4
2.1.2	Actinobacteria.....	5
2.1.3	Anaerobic Bacilli .....	5
2.2	The Gram-Positive Cell Envelope and the Cell Wall .....	7
2.2.1	Glycosyltransferases and biofilm.....	9
2.2.2	Autolysins and Penicillin binding proteins (PBPs).....	9
2.3	Wall teichoic acid (WTA) & lipoteichoic acid (LTA).....	12
2.3.1	WTA <i>de novo</i> biosynthesis .....	13
2.3.2	TarM/TarS glycosylation .....	14
2.3.3	Immune evasion via WTA alanylation .....	15
2.4	Immunogenic relevance of cell surface molecules.....	17
2.5	Mobile genetic elements (MGE).....	19
2.6	Bacteriophages .....	21
2.6.1	Lytic cycle and lysogenic cycle .....	22
2.6.2	The order <i>Caudovirales</i> .....	23
2.6.2.1	Siphoviridae.....	23
2.6.2.2	Myoviridae.....	24
2.6.3	<i>S. aureus</i> Pathogenicity Islands (SaPIs) .....	26
2.6.4	Siphoviridial architecture .....	27
2.6.4.1	The siphoviridial tail end polymorphism .....	27
2.6.4.2	Receptor binding proteins .....	29
2.6.4.3	Teichoic acids as receptors for bacterial viruses .....	30
2.7	Aims of Thesis .....	31
3	RESULTS AND DISCUSSION .....	32
3.1	TarM .....	32
3.1.1	Isolation and crystallization.....	32

3.1.2	Diffraction data collection and model refinement .....	32
3.1.3	Overall structure and domain organization of TarM .....	32
3.1.4	Structural and mutational analysis of the active site.....	33
3.1.5	DUF1975 .....	34
3.2	$\phi$ 11 Receptor Binding Protein Gp45 .....	36
3.2.1	Characterization of Gp45 .....	36
3.2.2	HHPRD analyses of baseplate components .....	36
3.2.3	Isolation and crystallization.....	37
3.2.4	Diffraction data collection and phasing .....	37
3.2.5	NCS and phase improvement.....	38
3.2.6	Gp45 structure (Rbp) .....	39
3.2.6.1	Stem domain .....	39
3.2.6.2	Shoulder domain .....	40
3.2.6.3	Head domain .....	41
3.2.7	Negative staining electron microscopy of the $\phi$ 11 baseplate .....	42
4	FUTURE RESEARCH.....	43
4.1	TarM .....	43
4.2	Gp45 .....	43
5	REFERENCES .....	46
6	APPENDIX .....	60
6.1	Gp45 crystal data and phasing statistics tables .....	60
6.2	Publications .....	65

## I ABBREVIATIONS

Amp	Ampicillin
ACP	Acyl-carrier protein
asu	Asymmetric unit
CAMP	Cationic anti-microbial peptide
CD	Circular dichroism
CoNS	Coagulase-negative <i>Staphylococci</i>
cpm	Capsid morphogenesis (gene)
DLS	Dynamic light scattering
DUF	Domain of unknown function
EM	Electron microscopy
EOP	Efficiency of plaquing
G+C	Guanine+cytosine content of DNA
Gal	Galactose
Glc	Glucose
GlcNAc	N-acetylglucosamine
Gp	Gene product
GroP	Glycerolphosphate
GT	Glycosyltransferase
hgt	Horizontal gene transfer
HMW	High-molecular weight
IMAC	Immobilized metal affinity chromatography
LMW	Low-molecular weight
LPS	Lipopolysaccharide
LTA	Lipoteichoic acid
ManNAc	N-acetylmannosamine
MES	2-(N-Morpholino)ethanesulfonic acid
MGE	Mobile genetic elements
MIRAS	Multiple Isomorphous Replacement with Anomalous Scattering
MOPS	3-(N-Morpholino)propanesulfonic acid
MRSA	Methicillin-resistant <i>S. aureus</i>
MSSA	Methicillin-sensitive <i>S. aureus</i>
MurNAc	N-acetyl muramic acid
NAG	N-acetyl glucosamine
NAM	N-acetyl muramic acid
NCS	Non-crystallographic symmetry
NMR	Nuclear magnetic resonance
OGP	n-octyl- $\beta$ -D-glucopyranoside
ORF	Open reading frame
NTP	Nucleotide-triphosphate
PAMP	Pathogen-associated molecular pattern
PBP	Penicillin-binding protein
PG	Peptidoglycan
PIA	Polysaccharide intercellular adhesin
PNAG	Poly- $\beta$ -1,6-glucosamine
PS/A	Polysaccharide/adhesin
PRR	Pathogen recognition receptors
RboP	Ribitolphosphate
RBP	Receptor-binding protein
rmsd	root mean square deviation

---

## Abbreviations

---

SCC	Staphylococcal cassette chromosome
SIRAS	Single isomorphous replacement with anomalous scattering
TLR	Toll-like receptor
SaPI	<i>S. aureus</i> pathogenicity islands
SEC	Size exclusion chromatography
SEM	Scanning electron microscope
SP	Scaffolding protein
Tag	Teichoic acid glycerol
Tar	Teichoic acid ribitol
UDP	Uridine 5'-diphosphate
vgt	Vertical gene transfer
VRE	Vancomycin-resistant <i>Enterococci</i>
VRSA	Vancomycin-resistant <i>S. aureus</i>
WTA	Wall teichoic acid

## II Summary (english)

By utilizing x-ray crystallography and electron microscopy (EM), this thesis describes the structure determination of two novel proteinaceous macromolecules together with their characterization by microbiological infection assays. These two objects of study revolve around an abundant cell surface molecule, wall teichoic acid (WTA), of *S. aureus* and its variants. These two structures are for one TarM, a glycosyltransferase involved in the final steps of the in-vivo synthesis pathway of WTA, and secondly Gp45, a receptor binding protein located in the baseplate of siphovirus  $\phi 11$ , a molecule that recognizes WTA by an unknown mechanism facilitating phage adsorption. Both recently discovered proteins, TarM and Gp45, were recombinantly overproduced in *E. coli*, isolated and crystallized. Since no homologous structure was available to facilitate the usage of the measured structure factor amplitudes for molecular replacement calculations, an appropriate structure solution protocol was sought. TarM crystals were treated with iodide ions and Gp45 crystals were likewise soaked in Tantalbromide solution in order to introduce anomalous scatterers into the protein crystal lattice. Consecutively, applying a combination of single and multiple isomorphous replacement with anomalous scattering (SIRAS and MIRAS, respectively) could retrieve the respective structure factor phase information for the protein atoms. A final solvent flattening routine enabled the structure solution of TarM. Gp45 could be interpreted in a model building process only after a more complex electron density modification routine, combining histogram matching, solvent flattening and averaging protocols. These unbiased models were refined against high-resolution datasets of 2.2 Ångstroms each. The position of Gp45 in the bacteriophage baseplate was elucidated by using negative staining EM on  $\phi 11$  particles and by superposing this x-ray structure model on the 3D reconstruction image derived therefrom. Semi-quantitative infection assays involving WTA deficient in glycosylation using TarM/TarS knockouts or mutants in strain *S. aureus* RN4220 or by blocking the cell surface specifically with recombinant Gp45 were carried out. Additionally, phages blocked with antibodies against Gp45/Gp54 were used to investigate dependencies governing impact and interaction of TarM and Gp45 with WTA in- and outside the cell.

*S. aureus* is responsible for many fatalities in clinical environments caused by infections. Treatments cannot only become tedious but are also a difficult economic factor. The fast resistance developing nature of this germ, of which many details are

not well understood, is a major obstacle to overcome lethal infections and therefore requires scientific research. The cell envelope represents a major research field and is in some cases affiliated with the emergence of resistance development. This dissertation describes a structure-based research utilizing in-vitro and in-silico laboratory methods, covering a tiny portion of this wide topic by presenting the structures of TarM and Gp45, and is to be seen as a contribution to the understanding of the complicated processes in and around the cell surface of a potent Gram-positive pathogen. The general goal of these interdisciplinary endeavours is to find sensitive treatments by implementing the vast data coming from cell envelope studies into drug design targeted against Gram-positive bacteria, specifically targeting crucial switches of the metabolism or components of the architecture itself.

### III Summary (german)

In dieser Arbeit wird die Strukturlösung zweier neuartiger proteinogener Makromoleküle mittels Röntgen-Kristallographie und Elektronenmikroskopie (EM) zusammen mit einer charakterisierenden Untersuchung durch Infektionsassays beschrieben. Diese zwei Studienobjekte drehen sich thematisch um ein häufig vorkommendes Zelloberflächenmolekül, nämlich Wandteichonsäure (WTS), aus *S. aureus* und seiner Varianten. Bei diesen zwei jüngst entdeckten Strukturen handelt es sich zum Einen um TarM, einer Glycyltransferase, die an den finalen Stufen der *in vivo* Biosynthese von WTS beteiligt ist, und zum Anderen um Gp45, dem Rezeptorbindenden Protein aus der Grundplatte von Siphovirus  $\phi 11$ , einem Molekül, das WTS nach einem unbekanntem Mechanismus erkennt und zur Adsorption des Phagen befähigt. Beide Proteine TarM und Gp45 wurden rekombinant hergestellt, isoliert und kristallisiert. Da keine homologen Strukturen vorlagen um fuer molekulare Ersatz Berechnungen die gemessenen Strukturfaktor Amplituden einsetzen zu können, wurde ein passendes ab initio Struktur Lösungs Protokoll gesucht. TarM Kristalle wurden mit Iodid-Ionen behandelt und Gp45 wurde gleichermassen in Tantal-Bromid aufgeschwämmt, mit der Absicht einen anomalen Streuer in das Proteinkristallgitter einzufügen. Schliesslich konnten daraus durch die Anwendung einer Kombination aus SIRAS und MIRAS die Strukturfaktor Phasen fuer die jeweiligen Proteinatome ermittelt werden. Eine abschliessende Solventabflachungsroutine konnte die Struktur von TarM schliesslich lösen. Andererseits konnte Gp45 einer strukturinterpretierenden Modellbau Prozedur erst ausgesetzt werden, nachdem eine komplexe Elektronendichte



Modifikations Routine, nämlich eine Kombination aus Histogramm Angleichung, Solventabflachung und Mittelung durchgeführt wurde. Diese unverfälschten Modelle wurden gegen jeweils hoch auflösende 2.2 Angström Datensätze verfeinert. Die faktische Position von Gp45 in der Phagen Grundplatte wurde durch Überlagerung dieser Röntgenkristallstruktur mit der Frost-EM 3D-Rekonstruktion präparierter Phagen Partikel ermittelt. Semi-quantitative Versuchsreihen zur Infektivität wurden mit TarM/TarS ausgeschalteten Stämmen, die damit defizient in der Glykosylierung ihrer WTS waren, durchgeführt oder indem Zelloberflächen spezifisch mit rekombinantem Gp45 geblockt wurden. Zusätzlich wurden Phagen mit Antikörpern, die gegen Gp45/54 gerichtet sind, geblockt, um den Einfluss und die Interaktionen in- und ausserhalb der Zelle von TarM und Gp45 zu sondieren.

*S. aureus* ist verantwortlich für viele Tote in klinischen Umgebungen, die durch Ansteckung verursacht sind. Nicht nur, dass die Behandlungen zäh verlaufen können, sie sind auch wirtschaftlich problematisch. Die schnelle Resistenzentwicklungsfähigkeit, über deren Details fast nichts bekannt ist, dieses Keimes ist ein Haupthindernis um den Todesursachen beizukommen. Diese Dissertation beschreibt eine im Schwerpunkt struktur-basierte Forschungstätigkeit, die in-vitro und in-silico Labor-Methoden nutzt, um einen winzigen Teil eines breiten Themenfeldes abzudecken, indem die molekularen Strukturen von TarM und Gp45 präsentiert werden, und es ist beabsichtigt Verwendung als Bereicherung zum allgemeinen Verständnis der komplizierten Prozesse in und um die Zelloberfläche eines starken Gram-positiven Pathogens zu finden. Das allgemeine Ziel dieser interdisziplinären Unternehmungen ist es, sensible Behandlungen zu entwickeln, die durch die Implementierung grosser Datenmengen aus Zelloberflächen-Studien im Bereich Drogen-Design hervorgehen und gegen Gram-positive Bakterien gerichtet sind, insbesondere spezifisch an metabolischen Schaltern oder der Architektur selbst wirken.

## IV List of publications

### Structural and Enzymatic Analysis of TarM Glycosyltransferase from *Staphylococcus aureus* Reveals an Oligomeric Protein Specific for the Glycosylation of Wall Teichoic Acid

Received for publication, October 20, 2014, and in revised form, February 10, 2015 Published, JBC Papers in Press, February 19, 2015, DOI 10.1074/jbc.M114.619924

Cengiz Koç, David Gerlach, Sebastian Beck, Andreas Peschel, Guoqing Xia, and Thilo Stehle

Anionic glycopolymers known as wall teichoic acids (WTAs) functionalize the peptidoglycan layers of many Gram-positive bacteria. WTAs play central roles in many fundamental aspects of bacterial physiology, and they are important determinants of pathogenesis and antibiotic resistance. A number of enzymes that glycosylate WTA in *Staphylococcus aureus* have recently been identified. Among these is the glycosyltransferase TarM, a component of the WTA de novo biosynthesis pathway. TarM performs the synthesis of  $\alpha$ -O-N-acetylglycosylated poly-5-phosphoribitol in the WTA structure. We have solved the crystal structure of TarM at 2.4 Å resolution, and we have also determined a structure of the enzyme in complex with its substrate UDP-GlcNAc at 2.8 Å resolution. The protein assembles into a propeller-like homotrimer in which each blade contains a GT-B-type glycosyltransferase domain with a typical Rossmann fold. The enzymatic reaction retains the stereochemistry of the anomeric center of the transferred GlcNAc moiety on the polyribitol backbone. TarM assembles into a trimer using a novel trimerization domain, here termed the HUB domain. Structure-guided mutagenesis experiments of TarM identify residues critical for enzyme activity, assign a putative role for the HUB in TarM function, and allow us to propose a likely reaction mechanism.

### An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*

Received: 08 December 2015 Accepted: 28 April 2016 Published online: 23 May 2016 *Scientific Reports* 6, Article number: 26455 (2016) doi:10.1038/srep26455

Xuehua Li, Cengiz Koç, Petra Kühner, York-Dieter Stierhof, Bernhard Krismer, Mark C. Enright, José R. Penadés, Christiane Wolz, Thilo Stehle, Christian Cambillau, Andreas Peschel, Guoqing Xia

Despite the importance of phages in driving horizontal gene transfer (HGT) among pathogenic bacteria, the underlying molecular mechanisms mediating phage adsorption to *S. aureus* are still unclear. Phage  $\phi$ 11 is a siphovirus with a high transducing efficiency. Here, we show that the tail protein Gp45 localized within the  $\phi$ 11 baseplate. Phage  $\phi$ 11 was efficiently neutralized by anti-Gp45 serum, and its adsorption to host cells was inhibited by recombinant Gp45 in a dose-dependent manner. Flow cytometry analysis demonstrated that biotin-labelled Gp45 efficiently stained the wild-type *S. aureus* cell but not the double knockout mutant  $\Delta tarM/S$ , which lacks both  $\alpha$ - and  $\beta$ -O-GlcNAc residues on its wall teichoic acids (WTAs). Additionally, adsorption assays indicate that GlcNAc residues on WTAs and O-acetyl groups at the 6-position of muramic acid residues in peptidoglycan are essential components of the  $\phi$ 11 receptor. The elucidation of Gp45-involved molecular interactions not only broadens our understanding of siphovirus-mediated HGT, but also lays the groundwork for the development of sensitive affinity-based diagnostics and therapeutics for *S. aureus* infection.

### Structure of the host-recognition device of *Staphylococcus aureus* phage $\phi$ 11

Received: 09 December 2015 Accepted: 17 May 2016 Published online: 10 June 2016 *Scientific Reports* 6, Article number: 27581 (2016) doi:10.1038/srep27581

Cengiz Koç, Guoqing Xia, Petra Kühner, Silvia Spinelli, Alain Roussel, Christian Cambillau, Thilo Stehle

Phages play key roles in the pathogenicity and adaptation of the human pathogen *Staphylococcus aureus*. However, little is known about the molecular recognition events that mediate phage adsorption to the surface of *S. aureus*. The lysogenic siphophage  $\phi$ 11 infects *S. aureus* SA113. It was shown previously that  $\phi$ 11 requires  $\alpha$ - or  $\beta$ -N-acetylglucosamine (GlcNAc) moieties on cell wall teichoic acid (WTA) for adsorption. Gp45 was identified as the receptor binding protein (RBP) involved in this process and GlcNAc residues on WTA were found to be the key component of the  $\phi$ 11 receptor. Here we report the crystal structure of the RBP of  $\phi$ 11, which assembles into a large, multidomain homotrimer. Each monomer contains a five-bladed propeller domain with a cavity that could accommodate a GlcNAc moiety. An electron microscopy reconstruction of the  $\phi$ 11 host adhesion component, the baseplate, reveals that six RBP trimers are assembled around the baseplate core. The Gp45 and baseplate structures provide insights into the overall organization and molecular recognition process of the phage  $\phi$ 11 tail. This assembly is conserved among most glycan-recognizing *Siphoviridae*, and the RBP orientation would allow host adhesion and infection without an activation step.

**V Personal contribution**

The author of this thesis was responsible for solving the crystal structures, interpreting the data and writing the manuscripts for the publications entitled “Structural and Enzymatic Analysis of TarM Glycosyltransferase from *Staphylococcus aureus* Reveals an Oligomeric Protein Specific for the Glycosylation of Wall Teichoic Acid” and “Structure of the host-recognition device of *Staphylococcus aureus* phage  $\phi$ 11”. The author of this thesis performed experiments incorporated in the manuscript entitled “An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*”.

## 2 INTRODUCTION

The outbreak of incurable diseases in confined sterile places such as clinics and hospitals, and the accompanying lack on possibilities to identify treatments to even faster developing community-acquired infections was and is still one of the catalysts for a drive in bio-technological development. In the last 30 years research on the cellular envelope could be intensified due to the establishment of possibilities given by the modernization of methods applicable in laboratories on a standardized basis. Genetic tools allow for quick characterization and identification, and together with online databases give the means to analyze freshly occurring pathogens on the fly. However, the ultimate aim is rather to find solutions in real-time. Essentially, germs of the *staphylococcus* genus, which represent the main entity of this thesis, infect patients in care and prevent their recuperation<sup>1</sup>. Neither the origination of their resistances nor their reported evolution are fully understood. The death toll of clinically-derived diseases is hence high and rising<sup>2,3</sup>. The approach by structural research is an highly advanced analytical instrument for the molecular level and represents a first step for drug design facilitating counter-strategies to pandemics. This chapter provides an overview of the cell envelope of Gram-positive bacteria, of which quite a few apart from some *Staphylococcus* strains are potent pathogens and therefore medically of relevance. The target of this thesis is nested amid the canopy of this broad topic.

### 2.1 Gram-positive Bacteria

The phylum *Firmicutes* combines rod- (bacilli<sup>4</sup>) and sphere-shaped (cocci<sup>5</sup>) bacteria with a cell wall and represents a major branch of Gram-positive bacteria<sup>1</sup>. As a major distinction to Gram-negative bacteria, the feature of Gram-positive bacteria's ability to adsorb and hold Gram-stain<sup>6</sup> stands out, with only *Mycoplasma*<sup>7</sup> being an exception. The most prominent Gram-positive genera are *Clostridium*<sup>8</sup>, *Lactobacillus*<sup>9</sup>, *Bacillus*<sup>4</sup>, *Mycoplasma*<sup>7</sup> and *Streptomyces*<sup>10</sup>. This staining feature is possible due to the lack of an outer membrane, which in contrast Gram-negative bacteria are surrounded by, their cell wall being part of the periplasmic space. As a consequence, the latter exhibit stronger physical and antibiotic resistances when compared to the former. One interesting feature of Gram-positive bacteria are their teichoic acids<sup>11-13</sup> in their cell walls, where Gram-negative bacteria have Lipopolysaccharides<sup>14</sup> (LPS) instead. Taxonomic

division is also possible by the content of guanine+cytosine bases in DNA (G+C-ratio)<sup>15</sup> of the bacterial chromosome. A low ratio, i.e. between 25 to 40 mol% is found for the most common *Firmicutes*, which share certain characteristics, whereas a high G+C content (more than 60 mol%) indicates actinomycetes<sup>16</sup>.

### 2.1.1 *Staphylococci*

The genus *Staphylococcus* is a member of the *Bacillales* order. It is non-spore-forming with a low G+C-ratio and consists of ca. 50 species<sup>17</sup>. These species belong to the family of *Micrococcaceae* and are catalase<sup>18</sup> producers. The most interesting species are skin-colonizing *Staphylococcus epidermidis*<sup>19</sup> and the characteristically yellow and spherical *Staphylococcus aureus* (Figure 1). *Staphylococci* can be found in dust, soil, mucous membranes, on skin, i.e. particularly all places with a certain degree of moisture<sup>20</sup>.

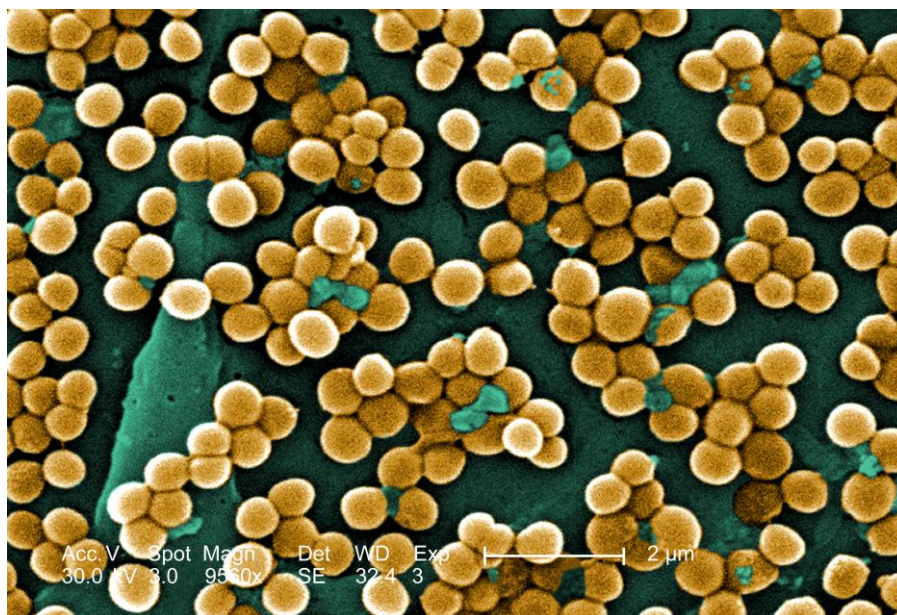


Figure 1 Scanning electron microscope (SEM) micrograph of agglutinated *S. aureus* (source: [phil.cdc.gov](http://phil.cdc.gov), permission for reproduction courtesy of CDC/Janice Haney Carr; Jeff Hageman, MHS)

#### 2.1.1.1 $\beta$ -hemolysis and Coagulase-negative *Staphylococci*

Besides certain *Listeria*- and *Streptococcus*-species<sup>21,22</sup>, the also pathogenic *Staphylococcus aureus* has the ability to clot blood. A test differentiating among these species is performed on so-called blood agar<sup>23</sup>. Plated  $\beta$ -hemolyzing species are indicated by the yellow and opaque lawn around bacterial colonies generated from lysed hemoglobins of red blood

cells (see Figure 2). In comparison,  $\alpha$ -hemolysis is present when red blood cells are partially broken down, releasing the greenish biliverdin, an intermediary byproduct of hemoglobin catabolism. Lastly,  $\gamma$ -hemolysis denotes the absence of any hemolysis reaction. On a side note, the catalase activity exhibited by *Staphylococcus* spp. is used to discriminate between the genera *Staphylococcus* and *Streptococcus*<sup>24</sup>. Other than *S. aureus*, coagulase-negative *Staphylococci* (CoNS) lack blood-clotting exoenzymes such as hemolysins and virulence factors but have other means of pathogenesis<sup>25</sup>. Although generally non-pathogenic, they are related to nosocomial infections in patients with implants<sup>26-28</sup>. These usually harmless skin commensals may cause hospital-acquired infections including central nervous system shunt infection<sup>29</sup>, prosthetic valve endocarditis<sup>30-32</sup>, urinary tract infection<sup>33-35</sup> or endophthalmitis<sup>36</sup>. The special arrangement of extracellular polysaccharides<sup>37</sup> (PS/A<sup>38</sup> capsular polysaccharide/adhesin<sup>39</sup>), also referred to as 'slime' or 'biofilm'<sup>40</sup>, is one major reason behind this, and enables adhesion of bacteria to synthetic devices.

Chronically, the presence or absence of coagulase was given as a reference to distinguish between coagulase-negative and -positive *Staphylococci*. The coagulase assay in tandem with the thermonuclease reaction and the novobiocin susceptibility resistance test however turned out to be more useful tools when differentiating up to the species level<sup>41-43</sup>.



Figure 2  $\beta$ -hemolysing *S. aureus* on blood agar (source: [phil.cdc.gov](http://phil.cdc.gov), permission for reproduction courtesy of CDC/Melissa Dankel; James Gathany, MHS).

#### 2.1.1.2 MRSA & MSSA

Methicillin-resistant *Staphylococcus aureus* (MRSA) is the direct progeny of penicillin-resistant strains<sup>44-47</sup>. Historically, infected patients were treated with penicillin<sup>48</sup>. An inflationary usage of penicillin led to decrease of susceptibility through the acquisition of  $\beta$ -lactamase. Methicillin, a more potent  $\beta$ -lactam, was used instead of penicillin since it was not affected by  $\beta$ -lactamase at that time<sup>49-52</sup>. Eventually,  $\beta$ -lactam-resistant MRSA emerged. The mechanism of *S. aureus* defying methicillin cell toxicity is not yet fully understood. In direct contrast to that stand the methicillin-susceptible *Staphylococcus aureus* (MSSA)<sup>53</sup>, which, from a medical point of view are less harmless than the MRSA, due to their treatability<sup>54</sup>. MRSA is the most widespread community-acquired infection. It is one of the major causes of nosocomial pneumonia, surgical site infection and blood stream infection. Due to the difficulty of its treatment, it remains a severe threat to health care systems worldwide. Antibiotics resistance in *S. aureus* is conveyed by a vast number of plasmids most likely obtained by horizontal gene transfer (HGT).

### 2.1.2 Actinobacteria

A major phylum of Gram-positive bacteria belongs to the superfamily of actinobacteria<sup>55-59</sup>. More than half of the known species of this 'high G+C' (around 60 mol%) branch are assumed to be inhabitants of human mucous membranes. Some are capable to dissolve even sturdy biomaterial like cellulose<sup>60</sup> or chitin<sup>61</sup>. Actinomycosis is a severe inflammatory illness caused by the more pathogenic species and primarily affect cattle<sup>62</sup> but can also cause infectious epidemics in humans<sup>63</sup>.

*Mycobacterium* has its name derived from the mycolic acid<sup>64</sup> present on its surface, making it highly hydrophobic and facilitating characterization by the so-called "acid-fast"<sup>65</sup> staining technique. The slow-growing and non-sporulating *M. tuberculosis* and *M. leprae* are agents responsible for tuberculosis and leprosy, respectively. *M. tuberculosis* captivates the host's macrophages to cause severe damage to peripheral tissue by repeated invasion of macrophages, which lead to malignancies of the host's immune system. Tubercles, formed by phagocytizing macrophages<sup>66,67</sup> in lungs, cause lesions and lead to the characteristic symptom of tuberculosis. Eventually the parasite escapes the host through the aerial tract to infect other hosts.

*Streptomyces*<sup>68</sup> are mainly soil bacteria with the highest count (ca. 500) of species among *Firmicutes*. Under favourable conditions, they grow sporophores from the hyphal tip of their spore germ tubes<sup>69,70</sup>. This multinuclear mycelium<sup>71</sup> makes them phenotypically reminiscent of fungi. Several species are symbionts and convey their hosts resilience against diverse pathogens, which makes *Streptomyces* important producers of essential antibiotics<sup>72-76</sup>. Examples of such secondary metabolites<sup>77,78</sup> are streptomycin (from *Streptomyces griseus*), chloramphenicol, nystatin (fungi-antagonist), erythromycin, tetracyclines and others.

### 2.1.3 Anaerobic Bacilli

Bacilli usually denote rod-shaped bacteria. They are mostly aerobic or facultatively anaerobic spore-formers and inhabit soil and water. The majority in the family of *Bacillaceae* are non-pathogenic. However, exceptions include *B. anthracis*<sup>79</sup>, *B. cereus*<sup>80</sup> and *B. thuringiensis*<sup>81</sup>. Moreover, genes from the latter



and from *B. amyloliquefaciens* are especially applied in the field of genetic modifications of plants because of their ability to target infected cells<sup>82-85</sup>.

*B. subtilis* is a well-studied non-pathogenic species with a low G+C-ratio. It was first discovered in the 19<sup>th</sup> century in hay in two forms<sup>86</sup>. Firstly, heat-resistant endospores, which grow vegetatively, are resilient to harsh environmental conditions. Secondly, a predominant heat-prone form being the actual bacterium. *B. subtilis* is the producer of bacitracin. Its cylindrical shape is allegedly a consequence of the specially arranged “circumferential” PG layer and the actin homolog MreB it harbours<sup>87-89</sup>.

Clostridium is one exemplary object for sporulating Gram-positive rods, which thrive under anaerobic conditions. In the medical field two species are of special interest: the causative agent of botulism, *C. botulism*, and the causer of tetanus, *C. tetanus*<sup>90</sup>. Examples for other species of interest are *C. perfringens* and *C. difficile*<sup>91</sup>, which are wide-spread causers of food poisoning.

Other obligate anaerobes include *Propionibacterium*, species of which cause acne, and *Bifidobacterium* which is also commonly being used to treat people with loss of intestinal flora, for example as a result of radiation- or chemotherapy and carcinoms<sup>92</sup>. A facultative anaerobic genus, especially of interest in dairy production and generally in food production, is *Lactobacillus* spp<sup>93-95</sup>.

## 2.2 The Gram-Positive Cell Envelope and the Cell Wall

The chemical composition of the cell envelope consists of conserved macromolecules of proteoglycan nature and has glycosylated cell-surface proteins tethered to the bacterial rim<sup>96</sup>. As a major part of the cell envelope, the *S. aureus* cell wall is composed of different classes of glycosylic structures<sup>97</sup>, including glycosylated proteins, biofilm<sup>98</sup> Polysaccharide Intercellular Adhesin (PIA or PS/A), capsular polysaccharides, cell wall peptidoglycan (PG), lipoteichoic acid (LTA) and wall teichoic acid (WTA)<sup>99,100</sup> (see Figure 3). Moreover, specialized Gram-positive bacteria bear surface layer protein (*B. anthracis*)<sup>101</sup> or mycolic acid (*Mycobacterium*)<sup>102</sup>, for example.

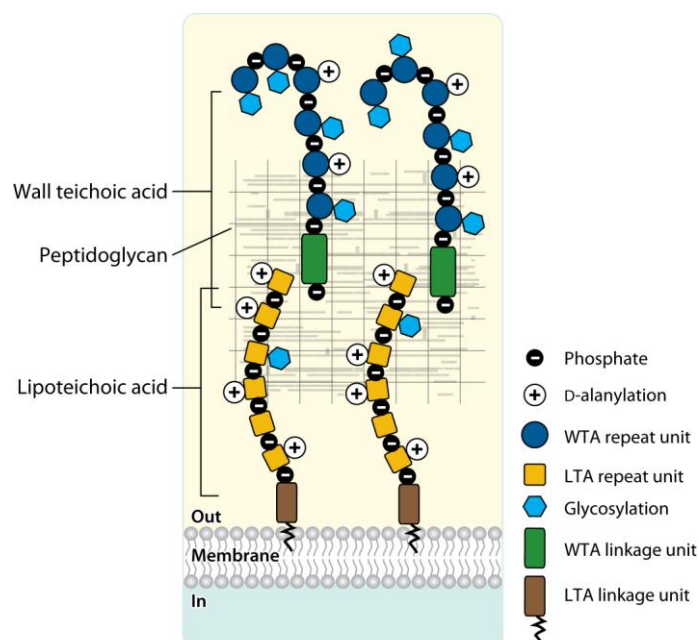
The integrity of the cell is maintained by the cell wall, a peptidoglycan scaffold. With its multigenous layer the tightly meshed cell wall is able to convey means of protection from and interaction with the environment. It serves as a protective mesh against cell adversaries<sup>103</sup>. Furthermore it is flexible in a manner, which allows cell division, conveys mechanical protection, balances osmotic pressure and permits interchanging of macromolecules with the environment<sup>17</sup>. Making up 90% of the dry weight of a cell, the cell wall of up to 80 nm is able to counter balance the Turgor pressure of 10-25 atm. Flagelli, which reach beyond the cell wall, are also ubiquitous components for any mobile cell.

The chemical composition of the cell wall mainly consists of N-Acetyl muramic acid (MurNAc) and N-Acetyl glucosamine (GlcNAc),  $\beta$  (1-4)-linked by glycosidic bonds. The interconnection of these singular glycan strands occurs through stem-peptides and pentaglycine bridges (*Staphylococcus*) via the MurNAc carboxylate moiety to form a vast network. According to a calculated model derived from an x-ray structure of a fragment, they are believed to form pseudo-helical strands<sup>104</sup>. Other studies have shown the cell wall to be compartmentalized in a low density inner wall zone, a region of up to 20 nm depending on the genera, and a high density outer wall zone, which generally represents the PG layer and its attachments<sup>87,88,104</sup>.

Lipoteichoic acids (LTA) are attached to the cytoplasmic membrane and almost predominantly consist of poly-glycerolphosphate (GroP) in the bacterial kingdom<sup>105</sup>. Free osmosis through the cell wall is a common process, whereas

Gram-negative bacteria have specialized pore-proteins, called porins, and lipopolysaccharides (LPS), consisting of lipid A, core and O-Antigen, instead of LTA. Gram-negative organisms are surrounded by a secondary outer membrane, whereas the Gram-positive bacterial PG layer represents the margin of the cell and is merely covered by capsular polysaccharides and proteins. The murein layer of Gram-negative bacteria is thin, circumferential and located in the periplasmaticum. Molecularly, PGs of both types are very similar apart from minor variations. Furthermore PG functions as an exoskeleton and is a determining factor of cellular shape. The study of the cell envelope is essential for the understanding of the maintenance and development of virulence<sup>88,103,106</sup>.

Especially WTA is reported to be involved in infections in wounds of skin, flesh and blood, e.g. olfactory tissue in animal models<sup>107-110</sup>. It plays a defensive role by protecting the cell from human lysozyme degradation. Acetylation of MurNAc by OatA on approximately every second C6 in conjunction with WTA on approximately every ninth MurNAc-C6 are presumably responsible for this<sup>111-113</sup>. It is involved in biofilm output control, which is necessary for colonization and it functions as cell receptor. The involvement of WTA in processes related to cell proliferation is crucial, although unclear.



**AR** Brown S, et al. 2013.  
Annu. Rev. Microbiol. 67:313–36

Figure 3 Schematic representation of WTA and LTA in the Gram-positive cell envelope (source: given)

below picture, permission for reproduction courtesy of Annual review of microbiology by ANNUAL REVIEWS. Reproduced with permission of ANNUAL REVIEWS in the format Thesis/Dissertation via Copyright Clearance Center.)

### 2.2.1 Glycosyltransferases and biofilm

The synthesis of each glycosidic bond in an organism is done by its own invariant glycosyltransferase (GT). Overall, GT structures can roughly be divided into two big families: GT-A and GT-B. Connecting characteristics for both include a predominant Rossmann fold, which assumes an evolutionary relationship for both GTs. These two superfamilies subsists of numerous subclasses, which define substrate and stereo specificities and targets, conservations in active sites, *in-vivo* enzyme locations, reaction classes and kinetics. An extensive collection of these data is accessible in the Carbohydrate-Active enZymes Database ([www.cazy.org](http://www.cazy.org))<sup>114,115</sup>.

An example for a GT is to be found in PIA production. PIA is referred to a polycistronic *locus* in biofilm-active *Staphylococci* consisting of the operon *icaABCDR*, encoding a GT and other auxiliary proteins. It is related to the exopolysaccharide production of CoNS on artificial prosthetic devices and represents a crucial virulence factor in pathogenesis. This slime mainly consists of poly- $\beta$ -1,6-glucosamine (PNAG) and has an approximate molecular weight of around 30,000 kDa. The deacetylated form of PNAG, like capsular polysaccharide conjugates<sup>116</sup>, was also reported to be useful in vaccination<sup>117</sup>.

After deacetylation and sparse substitution with amino acids causing the generation of positive charging of the linear molecule, it becomes electrostatically attracted to the staphylococcal cell surface due to the paramount negative charge caused by the presence of WTA<sup>118</sup>. This slime production is activated upon infection by an unknown mechanism. PIA presumably assists in intercellular as well as cell-host adhesion and facilitates cell proliferation<sup>98,119</sup>.

### 2.2.2 Autolysins and Penicillin binding proteins (PBPs)

PG hydrolyzing enzymes are referred to as autolysins and are essential for PG regeneration<sup>120,121</sup>. Carboxypeptidases, endopeptidases, N-

acetylmuramoyl-L-alanine amidases and lytic transglycosylases are all counted to this enzyme class. Whereas carboxy- and endopeptidases cleave peptidic bonds of the stem-peptide, the lytic transglycosylases<sup>122,123</sup> cleave the MurNAc- $\beta$ -(1,4)-GlcNAc glycosidic bond by releasing a peri-cyclic  $\alpha$ -MurNAc-intermediary in distinction to lysozyme (amidase)<sup>124,125</sup>.

The expression of the major autolysin is under control of the *atl*-operon. The cleavage of the gene product results in two autonomous enzymes, N-acetylmuramoyl-L-alanine amidase and  $\beta$ -N-acetyl-glucosaminidase. The amidase is hydrolyzing the peptidoglycan bond between L-alanine of the stem-peptide and MurNAc of the glycan strand, the glucosaminidase is dissolving the glycosidic  $\beta$ (1 $\rightarrow$ 4) bond between GlcNAc and MurNAc (glycosidase)<sup>124</sup>. The regulatory importance of autolysin capability to degrade the cell wall becomes a crucial factor during cell wall growth<sup>126</sup>, separation of daughter cells during cell division<sup>127</sup> and sporulation<sup>128,129</sup>. Pore enlargement and release of turnover products are other features of *atl* activity in cell wall regulation<sup>130</sup>.

Class A- and B-PBPs are high-molecular weight (HMW) membrane-anchored, multimodular transpeptidases, which form the peptidic bond between the stem-peptide and the "cross-bridge", the connector between two adjacent peptidoglycan strands. Except for this transpeptidase domain, most class A-PBPs also harbour a transglycosylase domain. On the other hand, low-molecular weight (LMW) PBPs are well-characterized endopeptidases and are related to transpeptidases by catalyzing the same reaction, but in the reverse direction. Other LMW PBPs are carboxypeptidases, and together with endopeptidases these belong to class C of PBPs. Class C-PBPs have two typical sequence motifs in common, identical to those of HMW-PBPs.

PBPs<sup>131</sup> have acquired their denomination due to their sensitivity to  $\beta$ -lactam compounds (ampicillin, carbenicillin, penicillin, etc.), which covalently bind to the serine of the conserved SXXK sequence motif of the penicillin-binding/endopeptidase or penicillin-binding/transpeptidase domain. This domain harbors a D-Ala-D-Ala binding site, an intermediary of all classes of PBP activities. The irreversible blocking of the active site by  $\beta$ -lactams via

stable protein-acyl-complexes blocks consequently cell wall biosynthesis at a crucial step<sup>132-134</sup>.

Of the few *S. aureus* PBPs, the most prominent is PBP2a. It belongs to PBP-class B1 of the HMW transpeptidases and is encoded by *mecA* in the staphylococcal cassette chromosome *mec* (SCC*mec*) on transposon TN552<sup>135</sup> under regulatory control of MecI and MecR1. The regulatory system and the functional arrangement of its gene products are almost identical to those of the *lac* operon<sup>136</sup>. Induced overexpression of PBP2a does not change expression levels of other PBPs in the cell, but since it has a lower sensitivity to  $\beta$ -lactams, it allows for replacement of the enzymes of the murein layer biosynthesis.

PBP2a shares little homology with any PBPs from *E. coli* or *B. subtilis*, but it is similar to subclass B1 *Enterococcus faecium* PBP5<sup>137</sup>. However, striking differences are not only found in the tertiary structure of the DD-transpeptidase-domain, but PBP2a is also very unique in its active site and its N-terminal domain<sup>138</sup>.

## 2.3 Wall teichoic acid (WTA) & lipoteichoic acid (LTA)

WTAs of various species and different strains are composed of variable chemical units<sup>58,100,139,140</sup>. Some of these structures are related in different genera of Gram-positive bacteria, which is indicative of a common ancestry, e.g. the linkage unit<sup>141,142</sup> GroP<sub>2</sub>-3-ManNAc-β(1,4)-GlcNAc-1-P to peptidoglycan-MurNAc-C6 and up to 60 repetitive units covalently connected by phosphodiester. Predominant repeating units of the WTA main chain are either 1,3-GroP or 1,5-D-ribitolphosphate (RboP) (Fig 4). Different *B. subtilis* strains have GroP, RboP or even 6-Gal-α(1,2)-GroP. The genus *Staphylococcus* can have either GroP or RboP as WTA<sup>139</sup>.

LTA molecules are covalently attached to the plasma membrane. The predominant LTA type I of low-G+C Gram-positive bacteria consists of GroP repeating units linked to the cell outer membrane by the glycolipid anchor Glc(β1,6)Glc(β1,3)-(gentiobiosyl)diacyl-Gro via C6 of the non-reducing sugar. Other LTA classes are defined as follows: type II LTA has GalGal-GroP and type III has Gal-GroP repetitive units. Rare LTAs, those which do not fall under the categories described above, can be further classified according to chemical nature and degree of GroP-substitutions, glycolipid compositions and chain lengths<sup>99,143-149</sup>.

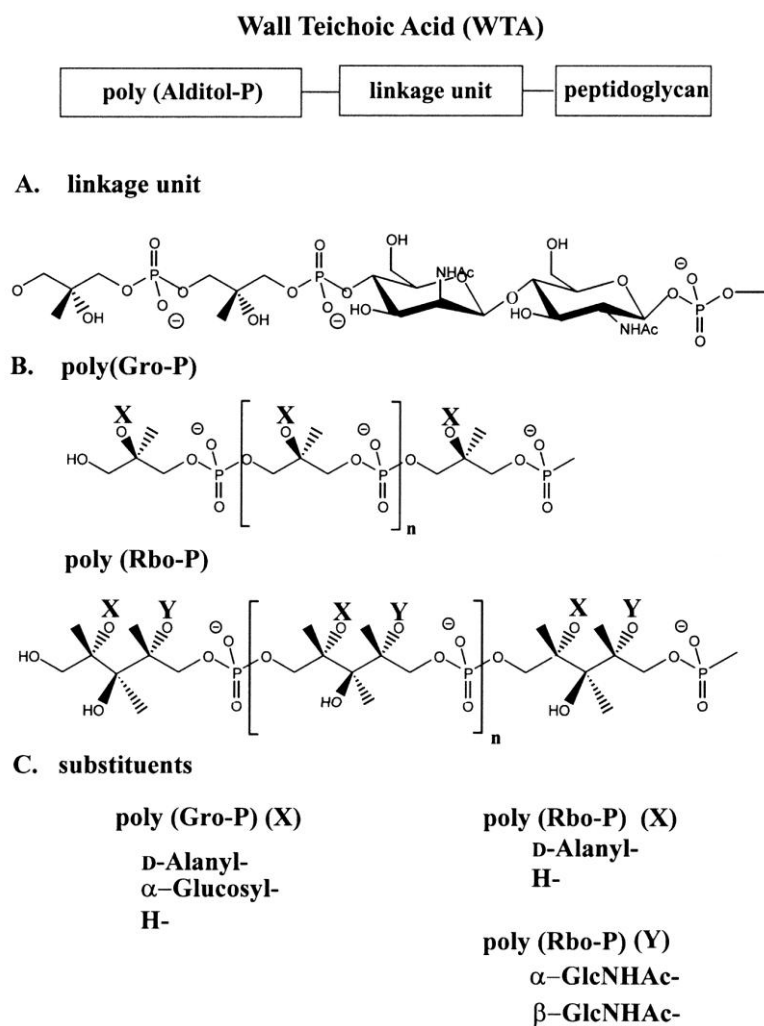


Figure 4 chemical structures of the linkage unit (A) and *B. subtilis* 168 GroP-type (B top) and *S. aureus* RboP-type (B bottom) WTA with their decorations (C) (source: Neuhaus, Francis C. et al.; *A Continuum of Anionic Charge: Structures and Functions of D-Alanyl-Teichoic Acids in Gram-Positive Bacteria*, *Microbiol. Mol. Biol. Rev.* December 2003, vol. 67, no. 4, 686-723; copyright permission courtesy of Microbiology and molecular biology reviews: MMBR by American Society for Microbiology. Reproduced with permission of AMERICAN SOCIETY FOR MICROBIOLOGY in the format Thesis/Dissertation via Copyright Clearance Center.).

### 2.3.1 WTA *de novo* biosynthesis

Biosynthesis of RboP-type WTA in *S. aureus* is carried out intracellularly by a cluster of enzymes belonging to the teichoic acid ribitol (Tar) synthesizing pathway, whereas the related cluster of enzymes for GroP-synthesis are denoted teichoic acid glycerol (Tag), i.e. LTA and WTA are generally synthesized by two independent pathways. The synthesis of the linkage unit disaccharide and attachment onto the lipid carrier undecaprenyl monophosphate, embedded in the inner leaflet of the cell membrane, is



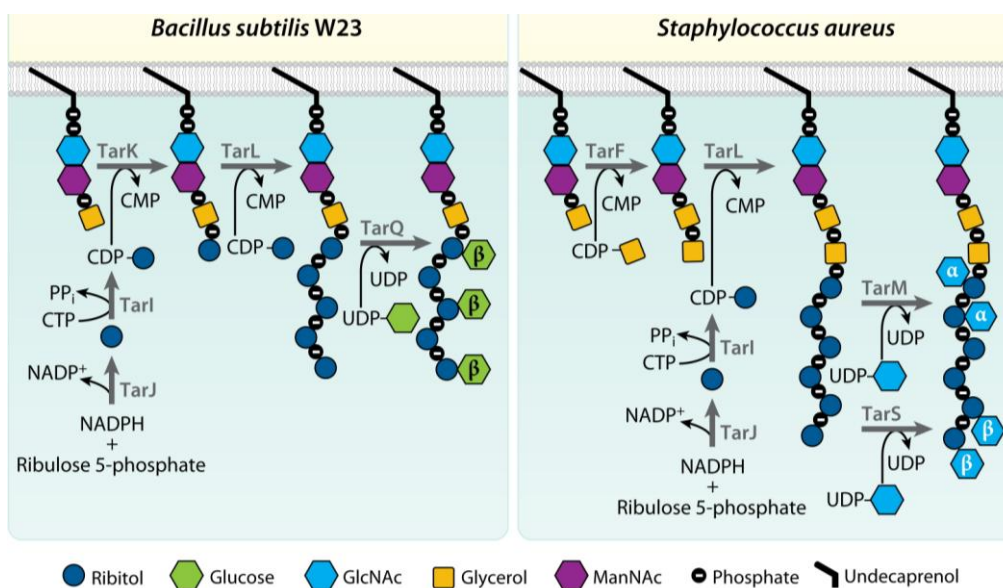
carried out by TarA, TarB (GroP-primase), TarD and TarO, which represent the general steps of WTA-biosynthesis. After this step, WTA-synthesis becomes strain- or species-dependent, culminating in the vast variety of main chain compositions and decorations.

In concert with TarI, a cytidyltransferase, and TarJ, an alcohol dehydrogenase, which generate the CDP-RboP supply, TarL (both primase and polymerase in *S. aureus*) adds additional GroP to the linkage unit and elongates the RboP backbone utilizing CDP-RboP. On a side note, some *B. subtilis* strains with a RboP-genotype have a separate primase (TarK), which putatively adds one primer RboP to the linkage unit, and a polymerase (TarL) for chain elongation (Fig. 5). After completion of biosynthesis, RboP is flipped to the outer leaflet of the plasma membrane via the ABC type transporter TarG/TarH and is covalently attached to the MurNAc moiety of the peptidoglycan<sup>150-155</sup>.

### 2.3.2 TarM/TarS glycosylation

RboP-carrying strains may be  $\alpha$ -,  $\beta$ -glycosylated at the RboP-C4'-position or D-alaninylated at the C2'-position. Other strains may have glycosylated GroP. For example, in *E. faecalis*, glycosylated GroP are reported to carry an additional D-alanine on their sugars<sup>156</sup>. Repeating units comprising GroP-glycosyl-P or GlcNAcP are also known (*Staphylococcus lactis*)<sup>157</sup>. ArabitolP in *Agromyces cerinus*<sup>158</sup> or ErythritolP in *Glycomyces tenuis*<sup>159</sup> are other known unusual repeating units.

A SaRN4220-mutant with a transposon in a gene of unknown function was discovered to have WTA with an altered phenotype. Concomitantly, it was found to be resistant to phage infection. A genetic analysis to find the cause for this effect resulted in identifying an open reading frame (ORF) transcribing a GT that was later termed *tarM*<sup>160</sup>. On the opposite side, TarS is a GT-A type according to sequence-derived predictions and is specifically decorating WTA with  $\beta$ -GlcNAc<sup>161</sup> (Fig. 5).




 Brown S, et al. 2013.  
Annu. Rev. Microbiol. 67:313–36

Figure 5 Elongation and decoration of of *B. subtilis* W23 GroP-type (left panel) and *S. aureus* RboP-type (right panel) WTA (source: given below picture, permission for reproduction courtesy of Annual review of microbiology by ANNUAL REVIEWS. Reproduced with permission of ANNUAL REVIEWS in the format Thesis/Dissertation via Copyright Clearance Center.).

By now only a handful of teichoic acid biogenesis affiliated protein structures have been solved for either poly-(GroP)-type or complex-type WTA, e.g. TagF from *B. subtilis*<sup>162</sup>, *Streptococcus* TarI<sup>163</sup> (pdb 2vsh), MnaA from *B. anthracis* (pdb 3beo)<sup>164</sup>, putative WTA ligase from *Streptococcus*<sup>165</sup> or lipoteichoic acid polymerase LtaS<sup>166</sup>.

### 2.3.3 Immune evasion via WTA alanylation

After the biosynthesis of WTA and LTA has been completed and both have been brought to their respective locations in the cell envelope, the sparse D-alanylation at the 2'-position of RboP commences. The gene product DltABCD is the responsible enzyme complex for the esterification. At first, DltA esterifies D-alanine with AMP and transfers the activated substrate to DltC. Sequentially, the pantothenate cofactor of DltC forms a thioester with D-alanine, concurrently releasing AMP. From here on, the membrane-anchored DltB and membrane-spanning DltD transfer the substrate from DltC to WTA or LTA by an unknown mechanism. According to studies, DltD is thought to hydrolyze the DltC-D-alanine-bond and facilitate transfer of D-alanine. The

predicted thioesterase domain of DltD is in accordance with the research performed on hydrolysis of D-alanine on acyl-carrier-protein (ACP). Even less is known about DltB, hence it was suggested that DltB must somehow be involved in a translocation process either as D-alanine-DltC or alone<sup>118,167</sup>.

In order to counter the host's native immune system, Gram-positive bacteria have a sensory system, the *aps-xsr* operon, which upregulates the genes necessary for D-alanylation, lysylation (*mprF*) of phosphatidylglycerol of the cytoplasmic membrane and a transporter system. With the aid of this machinery the bacteria are capable of evading CAMP or lysozyme action by electrostatic repulsion through the increased positive charge on their surfaces. Alternatively, the bacterial response specifically to linear CAMPs are proteases, e.g. V8 or aureolysin in *S. aureus*.

## 2.4 Immunogenic relevance of cell surface molecules

The recognition of the pathogen by the invaded host (mammal, plant, insect) is essential for the host's defense. For a successful countermeasure several mechanisms are employed. The distinction of the innate and the adaptive immune system of the host is important since the specialized defense triggers a response reaction according to site of infection and the tools at hand at that location. Cell surface molecules of the invader play a crucial role in this recognition process, and the type of immune reaction is dependent on the specific macromolecule<sup>168</sup>.

The invasion of bacteria in mucosal membranes or the subepithelial cells triggers two kinds of innate immunogenic responses carried out by neutrophils. Neutrophils (leukocytes) like granulocytes engulf whole bacteria and dissolve them in their vacuoles and process bacterial components to the adaptive immune system<sup>169</sup>. Bacteria are degraded via toxic oxygen intermediates by myeloperoxidase with NADPH in the oxygen-dependent mechanism<sup>170,171</sup>. The oxygen-independent mechanism as a second route triggers the release of cationic antimicrobial peptides (CAMPs)<sup>172</sup> upon sensing cell envelope glycopolymers via Toll-like-receptors (TLR)<sup>173,174</sup> on dendritic cells and subsequent NF- $\kappa$ B translocation.

CAMPs are microbicidal positively charged small molecules able to perforate the invader's cytosolic membrane, and they are released as a response to invasion of the host cell. They are part of the innate immune system of vertebrates, fungi, plants, insects (e.g. cecropins) and bacteria<sup>175</sup>. Prevalent in neutrophils of human tissue, e.g. mucous membranes and skin are defensins<sup>176</sup>, a major class of human CAMPs. Other examples include cathelicidins (e.g. LL-37)<sup>177</sup>, which are produced in keratinocytes and polymorphonuclear neutrophils and thrombicidins. In low concentrations, however, CAMPs are able to attract leukocytes via a process called chemotaxis<sup>178</sup>. CAMPs complement the transient pathogen response system.

Pathogen-associated molecular patterns (PAMPs) are microbe-related invariant macromolecules of polysaccharide nature and are readily recognized by pathogen recognition receptors (PRRs). PAMPs are exposed on microbial cell-

surfaces of Gram-positive as well as Gram-negative bacteria. For example, lipid A of LPS is recognized and transduces a proinflammatory signal in the immune response to activate against infection of Gram-negative bacteria<sup>179</sup>. An analogous reaction is triggered when LTA is recognized, a signal referring to the infection of Gram-positive bacteria. One PAMP or one bacterium can be an agonist to several PRRs. Examples for PAMPs include PG, lipoarabinomannan, glycerophosphoinositol, LTA, dsRNA, CpG DNA, flagellin, and others. PAMPs are expressed constitutively and represent conserved structures of germ-lines. These are distinct to the tightly regulated virulence factors, a response for evading the host defense system.

PRRs are highly specialized molecules of macrophages, neutrophils and dendritic cells recognizing PAMPs. They are variable membrane-spanning or membrane-anchored (e.g. CD14) protein-classes of the innate immune system<sup>180</sup>. Some PRRs are either highly functionalized, i.e. N- or O-glycosylated, cysteine-rich, etc. or might comprise several conserved domain features, e.g. C-type lectin domains, epidermal growth factors, fibronectin types. Prominent representatives include scavenger receptors, integrins, lectins, Toll-like-receptors and others<sup>181,182</sup>. One example for plant PRRs are PAMP-sensing LysM-domains. Utilization of PRRs is a first step in opsonization, phagocytosis, apoptosis, complement activation pathway, coagulation cascades and proinflammatory signalling. However, highly virulent pathogens like *Yersinia pestis*, *Salmonella enterica*, *Legionella pneumophila*, *Bordetella pertussis* are known to use the type III and type IV secretion systems to transmit effector molecules into macrophages in order to bypass host responses. The discrimination between pathogens and non-pathogenic commensal microflora is suppressed with anti-inflammatory cytokines by an unknown mechanism<sup>183-186</sup>.

WTA is able to activate the complement cascade, leading to opsonophagocytosis. In the case of WTA the mannose-binding lectin (MBL) pathway is activated via WTA-sugar decorations in general, however for the classical pathway and generally for antibody-recognition exclusively  $\beta$ -GlcNAc is of relevance.

## 2.5 Mobile genetic elements (MGE)

As the name suggests, MGEs are mobile DNA-segments in prokaryotes. These gene clusters can either be inserted into a commensal plasmid or into the chromosomal genome. MGEs exhibit features that make them discernable from the specimen's core chromosomal DNA. For example, transposases *per se* or site-specific recombinases are crucial factors by which they can be recognized. Deviating G+C-ratios is another such indicator. Adjacent virulence factors or resistance genes are also very common. It is being widely assumed that such adaptability and pathogenesis markers were inherited by horizontal gene transfer (hgt), although the mobility could have been lost in due course<sup>187-190</sup>. All MGEs, such as plasmids, temperate phages, chromosomal cassettes, pathogenicity islands or transposons are also passed on freely by vertical gene transfer (vgt).

Vgt describes the passing on of genetic information by means of replication to the progeny, hgt is encompassing all other modes of transfer of genetic information to related recipients. Three general mechanisms of hgt are known. Firstly, transformation describes the adsorption of genetic material from the environment by competent cell organisms when responsible systems for nutrient uptake are switched on. Secondly, transduction is a more specific mechanism. It occurs when temperate phages are transferring DNA upon entering the infectious state during the lytic cycle. Thirdly, conjugation is a complex mechanism by which pili are formed between a donor and an acceptor cell through which DNA is channelled. Conjugation is exclusively exploited by prokaryotes. This mechanism utilizes so-called integrative conjugative elements, which include transposases, integrases and other genes essential for conjugation. These *loci* can be as large as chromosomal DNA (megaplasmids) with independent origins of replication. Integrated conjugative elements are able to drag host chromosomal elements or other accessory genes during hgt. As an example, the occurrence of vancomycin-resistant *S. aureus* (VRSA) strains originates from vancomycin-resistant *Enterococcus* (VRE) through conjugation of the mobile genetic element TN1546. This transposon harbours the *van* operon and is activated upon vancomycin sensing. Gene products *vanA* and *vanH* are synthesizing a D-Ala-D-Lac intermediary with lower affinity to vancomycin,

whereas VanX is a dipeptidase specific for D-Ala-D-Ala targets. Other gene products of the *van* operon fulfill putative regulatory functions. Non-homologous recombination of accessory genes is an interesting phenomenon, as these incidents are currently believed to be the driving force in acquiring antibiotics resistances, making *S. aureus* and related pathogens harmful entities.

## 2.6 Bacteriophages

The major autonomous components of a bacteriophage consist of a head, often a tail, and a tail end. The head is a proteinaceous capsid, which represents the repository of the viral DNA encoding the enzyme machinery necessary for the phage's initial preservation after invasion. The head contains the viral DNA and can either be of icosahedral symmetry or filamentous and. The tail is a hollow tube-like fiber-protein complex that is able to channel the DNA from the head to the tail tip. The tail end is generally a multimeric protein complex and has its function in attachment to the host's cell surface and permits penetration of the cell barrier for injection of the viral DNA into the cell. Some bacteriophages have additional tail fibers. The dimensions of a full bacteriophage can be up to 200x100 nm.

Taxonomic classification of bacteriophages is based on morphology, gene type, genome size and infectivity. Bacteriophages are primarily classified according to their family (tailed phages), on a second level according to order (overall shape, infection type, genome-size) and finally according to genus (specific host range, replication machinery). The genus is the most relevant classification, since it describes the least diversification. Phages can be directly determined according to their functions and mechanisms.

Phages of the *Caudovirales* order were originally classified according to their reaction to specific polyclonal antibodies. Hence, they were divided into eleven serogroups. With the technical development of EM-techniques, morphological differences were subsumed inside these classical serogroups. Another way to classify bacteriophages is by the lysogenic mechanism they adopt, as described below. This general classification roughly divides phages into sets of contractile tails (*Myoviridae*), non-contractile tails (*Siphoviridae*) and without tails (*Podoviridae*) (Fig. 6).

The parasitic nature of phages is determined by the intracellular multiplication of its own genetic code by exploiting the host's replicative biosystem. By an unknown mechanism, the phage adsorbs to the cell surface with its baseplate or tail fibers. Both baseplate and tail fibers are important determinants of host tropism. The adsorption by tail fibers is generic and weak, whereas the



attachment by the baseplate is specific and utilizes diverse cell surface components as could be shown for a series of Gram-positive bacteria, e.g. *Lactobacilli*<sup>191</sup>, *Lactococci*<sup>192</sup>, *Streptococci*<sup>193</sup>. In the case of *Staphylococcus* spp and *Bacillus* spp, it could be shown that the presence of teichoic acids is crucial for virus adsorption<sup>194,195</sup>. The cell barrier is overcome by the concerted action of each domain, ultimately leading to the injection of the viral genome and leaving the protein husk outside the cell. This mechanism is in stark contrast to the invasion mode of eukaryotic viruses, which enter cells through endocytosis. After injection of the phage DNA, the fate of the invaded host is determined either by lytic or lysogenic infection and the invasion takes its course.

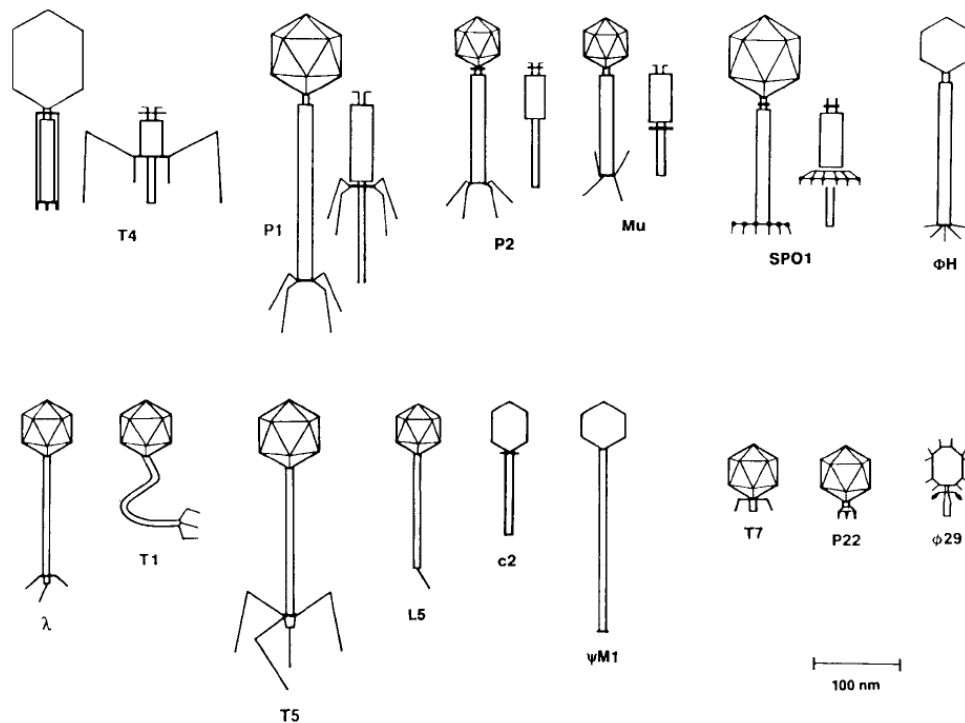


Figure 6 Schematic overview of the most frequent genera in the Caudovirales order of tailed phages. Myovirus: T4, P1, P2, Mu, SPO1, PhiH; Siphovirus:  $\lambda$ , T1, T5, L5, c2,  $\Psi$ M1; Podovirus: T7, P22,  $\phi$ 29 (source: Maniloff, Jack et al.; *Taxonomy of bacterial viruses: establishment of tailed virus genera and the other Caudovirales*, H. Arch. Virol. October 1998, vol. 143, no. 10, 2051-2063; copyright permission courtesy of Archives of Virology by Springer Link. Reproduction of figure with permission of Springer Link via Copyright Clearance Center.).

### 2.6.1 Lytic cycle and lysogenic cycle

The goal of the lytic cycle is to multiply and eventually destroy the host. In order to achieve this, phages hijack the transcription machinery and produce viral mRNA in the eclipic phase, meanwhile preventing the host's gene

transcription and translation. In due process the host's chromosome is degraded. After phage assembly is completed in the late phase of the lytic cycle, lysis proteins are expressed until the host cell bursts and phages are released.

The lysogenic infection is in many cases characterized by the incorporation of the viral DNA into the host's chromosome by site-specific recombination and sustenance in a dormant state, in which the transcription is blocked by a phage-encoded and phage-specific repressor. In this state the phage is referred to as prophage and the cell as lysogen. The replication system of the host is not hindered, and the viral genetic code is passed on together with the bacterial genome to its progeny. However, upon an induction event the prophage is activated and enters the lytic mode. Induction parameters are mainly in adverse to the host's thriving. Lytic mode is generally activated by the host's exposure to mutagenics, UV-light, ionic radiation and desiccation. Responsible for the induction is the protease RecA, which cleaves the viral repressor and enables the transcription of the prophage genome.

### 2.6.2 The order *Caudovirales*

Caudovirales comprise three major families of characteristically tailed phages: *Myo*-, *Sipho*- and *Podoviridae*. *Siphoviridial* dsDNA is comprised of 50 kb with a non-contractile tail. *Myoviridae* also have dsDNA of up to 280 kb, but a contractile tail instead. *Podoviridae* have short tails and dsDNA of around 40 kb. Before DNA is packaged, the procapsid is built by scaffolding proteins and a major capsid protein during assembly. After concatamerization of DNA, it is channeled into the procapsid, an ATP-dependant step carried out by the terminase-complex consisting of the small TerS and the large TerL proteins. The tail and the baseplate are attached to complete the virion.

#### 2.6.2.1 *Siphoviridae*

One example for a *Siphovirus* are the  $\lambda$ -like phages. They utilize a flexible DNA ejection system. After the cytoplasmic membrane has been punctured, the phage opens its cap, which is located between the head and the tail, in order to tunnel the DNA into the cell. From here the viral DNA circularizes by ligation at its *cos* (cohesive end site) sites by exploiting a host ligase. The first proteins to be expressed at the I-promoter site are the regulators

*cro* and *cII*. Generally, elevated cellular levels of the regulatory gene product *cro* are liable to promote lysis. *Cro* itself is dependant on environment variables that control its production. If the amount of *cro* exceeds that of *cII*, the phage enters its lytic pathway, if vice versa, the latent stage is enabled. In the latent stage an integrase is expressed, and the phage DNA recombines at chromosomal *attB*-sites regulated by *cl*. Either by degradation of *cl* and subsequent upregulation of *xis* and *int* or by entering the lytic pathway right from the beginning of infection, the circular virion is replicated. In the late stage, DNA is concatamerized, packaged, and finally the cell is being lysed.

### 2.6.2.2 Myoviridae

The *Myovirus* phage  $\mu$  has a contractile tail and six fibers at its baseplate, which allow for host tropism. The modification of tail fibers occurs either by inverting the direction of genetic transcription or by mutagenesis. After binding of the phage to the cell surface, a puncturing device cuts through the outer membrane and cell wall and fuses with the inner membrane. The outer tail tube contracts, pushing the dsDNA through the inner membrane. Once inside the host, the co-invading viral protein N holds the linear DNA circularized by keeping the ends together. This diminishes host nuclease activity on the bacterial 3'-ends of the phage DNA. Before going into replication or latency mode, the DNA is integrated into the bacterial chromosome as a transposon element by MuA and MuB. The early genes MuA and MuB are responsible for DNA insertion and target site selection. MuA is a Mg-ion-binding recombinase of the DDE family and MuB is the target DNA activator<sup>196</sup>. After the integration process is concluded, expression of repressors Ner facilitates latency by preventing early genes from expression or Repc controls further viral transposition yielding in up to 100 new copies of the viral transposon with the assistance of MuA and MuB. After this event, the expression of late genes leads to the formation of the structural phage proteins, DNA packaging and eventual lysis.

Coliphage P1 is another temperate bacteriophage of the *myoviridial* family. It has an icosahedral head and stores a dsDNA of around 90 kb size. It is distinct because of its different state of the virion in the lysogenic state. An

*attP*-site is lacking, therefore the linear DNA is homologously recombined at its terminally redundant sequences using the viral Cre-*loxP* system, a specific recombinase and its recombinase site. The result is a prophage plasmid, which during cell division replicates via its Holliday junction and promotes segregation to daughter cells by so-called binary fission<sup>197-199</sup>.

### 2.6.3 *S. aureus* Pathogenicity Islands (SaPIs)

Pathogenicity islands represent gene *loci* associated with specialized virulence factors, resistance genes and enterotoxin genes, e.g. toxic shock syndrome (*tst*). Of the identified SaPIs, best known are SaPI1 and SaPIbov1. SaPIs are stably integrated into the *S. aureus* host genome at *attC*-sites with their own autonomous replicons with the whole cassettes ranging from 14-27 kbs.

SaPIs have the ability to utilize structural proteins and enzymes of phages of the *Caudovirales* order to package their DNA into transducing particles. The generation of transducing particles occurs during the lytic phase of the phage upon induction, e.g. SOS-response, superinfection of a non-lysogen SaPI by a helper phage or concomitant infection of SaPI and bacteriophage. The general act of piracy, describing the exploitation of host genomic products to suit the needs of the invading virus, is reversed. The host becomes the “pirate” and the virion a mere “helper”. The helper phages are commonly found to be recruited from the family of *Siphoviridae*, and each SaPI has a specific helper phage range. The formation of SaPI-particles follows a mechanism similar to the general assembly pathway of *Caudovirales*, although one major difference of most SaPI-particles to common phages is the smaller size its head. Responsible for this size redistribution are additional scaffolding proteins (SPs). Most SaPI gene-clusters share the so called capsid morphogenesis genes (*cpm*), which assist in the formation of the small SaPI-procapsid by replacing SPs of the general virion assembly. However, exempt from *cpm*-assisted assembly are SaPIbov2 and SaPIbov5.

All of these phage-related chromosomal islands contain common regulatory genes, among them genes for excision (*xis*) and integrases (*int*), which make them highly mobile intra- and intercellularly. In the absence of phages, SaPIs circularize at redundant termini before integrating back into the host chromosome. The SaPI is not excised and replicated, lest helper phage repressors act on *stl*. When *stl* is not repressing *str*, the process of excision turns on. The viability of SaPIs depend on this regulation, since an excised state during cell proliferation would lead to their loss. Eventually, SaPI replication is carried out at the *ori* site via Rep, a replication initiator with

helicase activity, and leads to concatamer formation.

Subsequently, DNA packaging of SaPI-particles occurs analogously to the DNA packaging in phages. The *pac*-site of SaPI-DNA is recognized by SaPI-encoded small subunit TerS of the terminase-complex and translocated by the phage-encoded large subunit TerL inside the procapsid. When SaPI-particles are invading new hosts alone, they integrate into the new chromosome at specific sites. However, when invasion is accompanied by phages, they replicate immediately to form new particles.

## 2.6.4 Siphovirial architecture

### 2.6.4.1 The siphovirial tail end polymorphism

Parts of the baseplate complex of a *Siphovirus* are the distal tail protein (Dit), the tail associated lysin (Tal), the tape measure protein (Tmp) and the receptor binding protein (Rbp). In the structure of the *lactococcus*-infecting siphovirial bacteriophage TP901-1, 18 trimers in a total of 54 RBPs are building the baseplate protein lower (BppL) unit, which is connected to the core by 6 molecules of baseplate protein upper (BppU) in a six-fold symmetrical arrangement around the baseplate core. Intermolecular contacts between the oligomeric proteins in several copies each builds up a complex of 1.8 MDa (Fig 7).

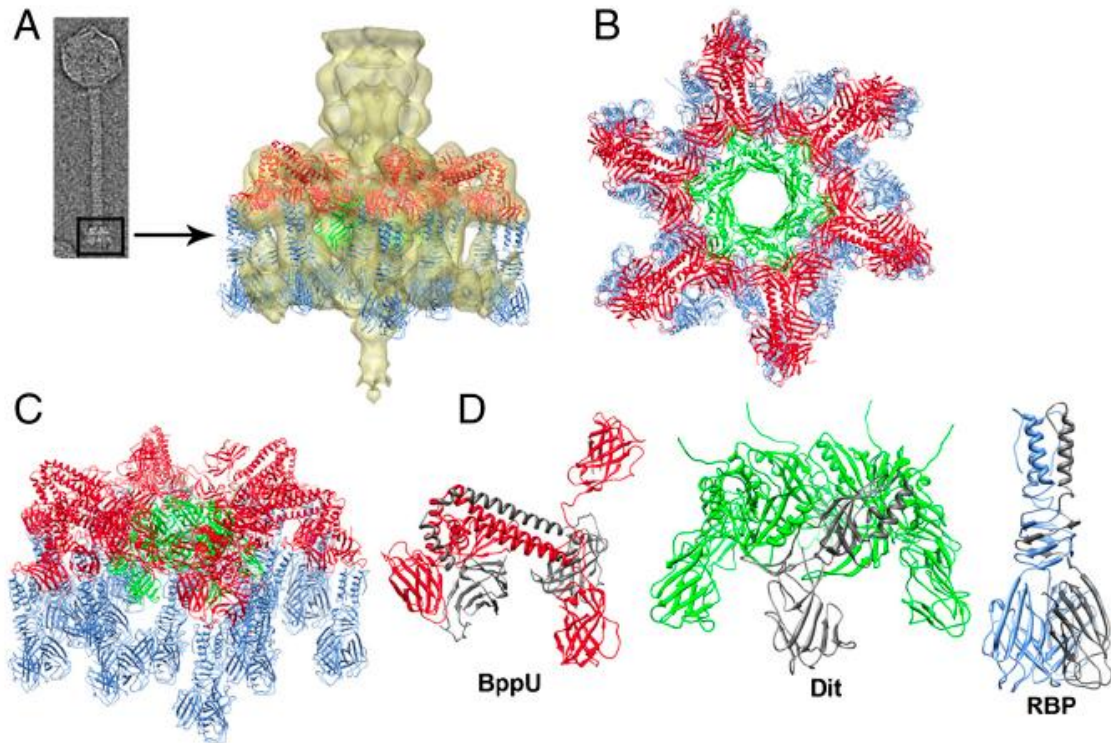


Figure 7 The TP901-1 baseplate architecture and its components. Negative-stained electron micrograph of TP901-1 Siphovirus (A left) with EM reconstruction image of the baseplate with fitted x-ray structure models (A right), baseplate complex viewed from top (B), baseplate complex in diagonal perspective (C), baseplate complex components BppU trimer, Dit hexamer and RBP trimer (D). (source: Veesler, D. et al.; Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism, PNAS April 2012, vol.109, no.23, 8954–8958; courtesy of National Academy of Sciences.)

The baseplate of *lactococcal* p2 is another example for *siphoviridial* diversity. Here, 6 RBPs are located laterally on the distal tail end in a fashion comparable to the TP901-1. For the p2 baseplate, two distinct conformations are found, active and resting, which are switched on and off by the presence or absence of calcium ions. Apart from that, p2 is rather comparable to TP901-1, especially because of the presence of a baseplate. The genetic locus encoding the proteins for the baseplate is well conserved among *Lactococcus spp.* infecting *Siphoviridae*. The prophage gene cluster harboring structural segments, head and tail for instance, is located amid open reading frames, which are also responsible for lysogeny, replication, packaging and lysis. Conventionally, the baseplate is related to the tail, since not every siphophage carries a baseplate in contrast to *Myoviridae*, which in general do have baseplates.

The *B. subtilis Siphovirus SPP1* has a peculiar trunk-like tail tip that is bridged via a hinge to the tail cap. The tail tip is attached to the tube cap distal to the head enabling flexibility independent of tail tube orientation and serves as an adhesion tool, recognizing *B. subtilis* YueB-ectodomain. This tail tip is larger than the usual tail spike protein of other phages, and another difference is the loss of it after adhesion, leading to conformational rearrangements of the tube cap, which facilitates the opening of the tube channel for tunneling viral DNA into the host.

Another baseplate structure is to be found in *Siphovirus* coliphage T5, a well-studied model phage of Gram-negative enterobacteria. Besides the presence of a tail tip like in SPP1, this phage also utilizes tail fibers for cell surface recognition.

#### 2.6.4.2 Receptor binding proteins

Tail ends of phages of the *Siphoviridae* family are structurally multivariant determinants of selectivity. These proteinaceous molecules, which are harboured in the tube distal end define host tropism and invasiveness. For lactococcal phages, slight alterations of a given Rbp elicit shifts of selectivity to closely related species from a narrow evolutionary scope. Relatively small structural modifications of the Rbp lead to the receptor recognition shifting from dairy phage group P335 to group 936<sup>200,201</sup>, both of which are distinct with respect to DNA homology and phage morphology<sup>202</sup>, as is seen for Rbp structures of Tuc2009, TP901-1, p2 and bIL170, for instance (Fig 8). The availability of a cell surface receptor can lead to natural adaptation of Rbps to recognize additional cell surface molecules as receptors, as is reported for phage  $\lambda$  and its recognition of OmpF apart from its cognate receptor LamB<sup>203</sup>. Other examples for such occurrences are also common for other phage families<sup>204-207</sup>.



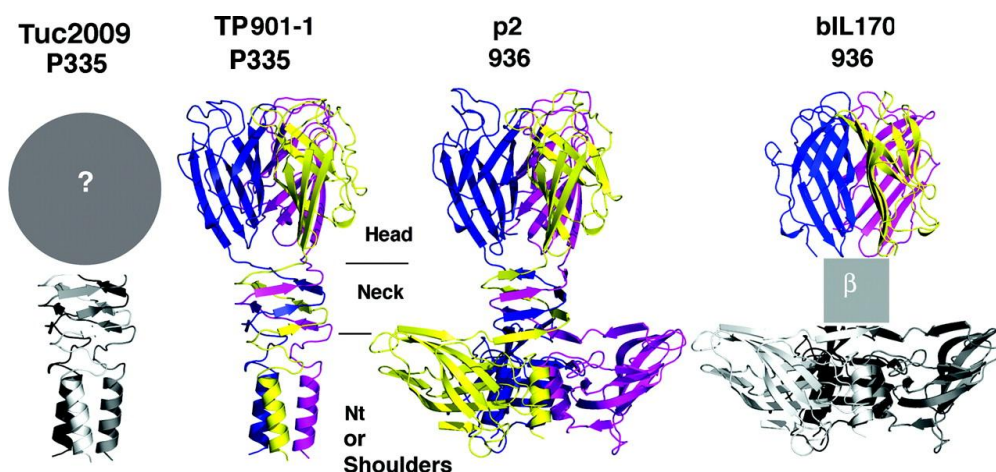


Figure 8 Structural comparison of four lactococcal Rbps out of two lactococcal groups (source: Ricagno, Stefano et al.; Crystal Structure of the Receptor-Binding Protein Head Domain from *Lactococcus lactis* Phage bIL170, *J. Virol.* September 2006, vol. 80, no. 18, 9331-9335, permission for reproduction courtesy of Journal of Virology by American Society for Microbiology. Reproduced with permission of AMERICAN SOCIETY FOR MICROBIOLOGY in the format Thesis/Dissertation via Copyright Clearance Center.).

#### 2.6.4.3 Teichoic acids as receptors for bacterial viruses

For invasion to happen, an irreversible recognition and successive adhesion step needs to take place on the cell surface. Formerly, LTA was believed to be the infection receptor of phages. However, phages of *Listeria* spp. and allegedly *L. lactis* are reported to recognize specific side-chains of teichoic acids as receptors<sup>208,209</sup>. *Listeria* phages temperate A118 and virulent P35 evidently recognize sugar moieties (rhamnose, GlcNAc) on poly-RboP-C2 or -C4. Whereas for A118 siphophage rhamnose is crucial, P35 requires both, GlcNAc and rhamnose, for adhesion. For lactococcal phages of groups c2, 936 and P335 it is assumed that they might be binding sugars or even glycerol units of GroP-type WTA<sup>200</sup>. For several staphylococcal phages, recognition of GlcNAc regardless of their anomeric configuration has been reported<sup>195</sup>.

## 2.7 Aims of Thesis

The general intention of this thesis is to describe and discuss recently identified WTA-affiliated proteins, and to highlight structural similarities and differences to other archived proteins of interest. The molecular analysis of these structures could assist in understanding WTA better. The characterization of novel domains or features specifically could identify molecular patterns to broaden our yet small knowledge of WTA needed for the development of treatments and vaccines against widely spread WTA bearing germs. Structure analysis of proteins associated with WTA biosynthesis are rare and the contribution through this thesis could serve as a puzzle piece to generate models of *in vivo* interdependencies of WTA.

Concretely, the structure solution and biochemical characterization of TarM of the biosynthesis pathway of RboP-type WTA and Gp45, the receptor binding protein of *S. aureus* infecting phages, which recognizes sugar moieties on WTA are the subject matters of this thesis.

## 3 RESULTS AND DISCUSSION

### 3.1 TarM

#### 3.1.1 Isolation and crystallization

TarM was produced in *E. coli* and lysed by ultrasonication. After affinity chromatography and size exclusion chromatography (SEC), the fused thioredoxin protein was removed and TarM isolated.

Crystallization conditions were screened in a sparse setup, and a successful initial hit was later refined to obtain bigger crystals. For a binary complex crystal, purified TarM was incubated with UDP-GlcNAc (15 mM, 1 h, 4 °C) and a new crystallization condition was screened and refined. Crystals were vitrified in liquid nitrogen. For anomalous dispersion experiments, wild-type crystals were treated with KI (30 mM, 30 ', 16 °C), backsoaked in crystallization condition and vitrified<sup>210</sup>. (See publication in appendix for details)

#### 3.1.2 Diffraction data collection and model refinement

Data were collected on a PILATUS 2 M hybrid pixel detector using synchrotron beam line X06DA at the Swiss Light Source (SLS) super-bending magnet (2.9 tesla). The XDS package was used for data reduction. The crystals belong to spacegroup P6<sub>3</sub>22 with cell parameters of a = b = 123.7 Å and c = 223.3 Å. One TarM monomer (58 kDa, 493 residues) comprises the asymmetric unit (asu), which has a solvent content of 74%. Data collection of anomalous data on the beamline was performed by irradiating the crystal in predefined wedges at low electron dosage.

Alternating cycles of COOT model building and REFMAC5 or PHENIX refinement revealed additional residues, which were included in the refinement until convergence had been achieved. The final model includes residues 1–493 (PDB ID 4wac, 4wad)<sup>210</sup>. (Data collection and refinement statistics can be found in the publication in the appendix).

#### 3.1.3 Overall structure and domain organization of TarM

TarM is the first enzyme structure in the biogenesis pathway of poly-(RboP)-type WTA to be reported. TarM is an  $\alpha$ -O-N-acetylglycosyltransferase specific for RboP and a member of the RboP biosynthesis pathway, also found in other

related species and strains. A structural study of this enzyme identifies a glycosyltransferase of type B (GT-B). TarM is a homotrimeric molecule (174 kDa) with a broadened oblate shape according to combinatorial biophysical characterization via dynamic light scattering (DLS), SEC and bioinformatic calculations (PISA-server<sup>211</sup>). The C-termini protrude from the molecular center, which is constructed by a threefold symmetry axis, so that the C-termini surround a large hollow center. Each TarM monomer consists of one domain of unknown function (DUF) and one GT-module. This DUF1975 is inserted into the GT. Three DUFs stack their 10-stranded antiparallel  $\beta$ -sheets against each other and thus exclusively form the assembly interface. This stacking allows for a 30 Å wide opening at the exterior of the central hub domain, the inner side on the other hand is closed tight. The interface of this domain buries a total surface area of 777 Å<sup>2</sup> (PISA-server).

The GT-module has the typical GT-B fold of anomer-retaining enzymes as is evident from the bipartite architecture separated by a flexible linkage of ca. 10 residues. Crystal contacts are prevalent in the N-terminal protein part, however these are becoming drastically less in the C-terminal protein parts, the latter subsuming extraordinarily high B-factor values. A DALI query of this second major domain returned typical GT-4 class homologs, e.g. MshA (Z-score 33.1, pdb 3c4q) and BshA (Z-score 32.9, pdb 3mbo). The unexpectedly overall low Z-scores are presumably a result of the conformational flexibility of the GT-module due to a flexible 10-residue linker. Each GT-B crystal structure is frozen in a slightly different relative orientation of N-termini against C-termini, exacerbating a complete superposition of 2 GT-Bs. Known acceptor substrates for GT-4 class enzymes range from 1-L-Ins-1-P (inositolphosphate) to LPS and to S-layer glycoproteins, and many others. The C-terminus is overall less well defined by electron density than the remainder of the protein, but the electron density for UDP- $\alpha$ -GlcNAc is nevertheless unambiguous.

### 3.1.4 Structural and mutational analysis of the active site

The sugar-transfer in GT-4 class glycosyltransferases is typically carried out at the interface between the N-terminus and the C-terminus. Binding of UDP-GlcNAc does not lead to any remarkable structural rearrangements (r.m.s.d. value of 0.81 Å for 493 aligned residues). According to the structural model,

UDP-GlcNAc is located in a cleft formed by five loops. The anomeric center carbon C1 lies in close proximity to the carboxylate function of Glu-403. Arg-326 and Lys-331 side chains are stabilizing the phosphate moieties. The uridyl unit is held in position by several prominent loops. The ribose is anchored to the Glu-411 side chain. Specificity for uracil is generated through several polar and hydrophobic interactions, namely Tyr-382, Thr-383, and Pro-386 and the side chain of Ile-324 appears relevant here, too. In summary, the UDP-GlcNAc binding pocket is well conserved with other GT-Bs.

For a mutational analysis of the UDP-GlcNAc binding site, RN4220 was complemented with various *tarM* variants to be used as hosts for phage  $\phi$ 11. EOP (efficiency of plaquing) utilizes a link between  $\alpha$ -O- and  $\beta$ -O-GlcNAc glycosylated WTA and bacteriophage adsorption to *S. aureus* hosts. A high EOP corresponds to a high WTA glycosylation, whereas inactive complements lead to a low degree of glycosylation.

This setup has been probed for active site variants E403A, K331S, R326S, and H249A. The integrities of the respective protein structures were verified via circular dichroism (CD) and DLS. Glu403 and Lys331 were found to be crucial for catalysis, whereas the latter two residues, Arg326 and His249, appear to be of lesser importance due to residual plaquing capacity.

### 3.1.5 DUF1975

A BLAST sequence search of protein databases finds DUF1975 only in TarM homologs of other Gram-positive bacteria. As TarM-mediated WTA glycosylation is thought to constitute a general pathway in Gram-positive bacteria with RboP-WTA, it seems likely that the DUF1975 fulfills a similar role in these related organisms. One noteworthy feature is the site DUF1975 is inserted. Here, the homologs without the insertion show a small region of undefined structure. Incidentally, this point of insertion also happens to be the dimerization site for MshA and BshA. The direct distance of two neighbouring active sites in the TarM trimer is approximately 72 Å, which corresponds to about eight or nine ribitol-units of an extended chain. Thus, a single TarM trimer could simultaneously glycosylate the same poly-RboP substrate at different locations. It is not currently known which RboP units in the long

polyribitolphosphate chain are glycosylated, but it is likely that the glycosyltransferases acting on WTA have a mechanism that enables them to move along the polyribitol chain and selectively glycosylate specific units.

Single amino acid substitutions (K136S, N138Q, N180W), a double (V159Y/C164R) and a triple (V159Y/C164R/K136S) mutation near the trimer interface were studied in EOP tests. We observed a substantial decrease in TarM EOP for K136S as well as the double and the triple mutant. However, where, under *in vitro* and EOP conditions, E403A was not capable to glycosylate WTA, the case for the triple mutant was different. Interestingly, under *in vitro* conditions the triple mutant shows a stronger enzymatic capability than in the EOP assay. Unfortunately, a conclusive explanation for this phenomenon is not possible with the data at hand. Expanding on this, the CD spectra of the double and the triple mutants show an additional shoulder around 205 nm, indicating a noteworthy structural alteration.

The streptococcal GT-B GtfA (pdb 4pqq) has also a DUF1975, and that is also inserted into the GT module at a similar location<sup>212</sup>. Furthermore, GtfA DUF1975 is the only structure in the DALI database with any significant structural homology to the TarM domain (Z-score 11.6, sequence identity 16%). However, GtfA is clearly monomeric, and the enzyme also does not act on WTA. DUF1975 from GtfA and TarM are very much the same module structurally, however the mode of action and the specificities for both enzymes are diverging dramatically. Here, it is almost impossible to deduce from one DUF1975 the function of the other one.

## 3.2 $\phi$ 11 Receptor Binding Protein Gp45

### 3.2.1 Characterization of Gp45

Gp45 could be identified by bioinformatic analysis. The protein shares identity of about 45 % with ORF636 of PhiSLT, an annotated cell adhesion molecule. Moreover, antibodies were raised against recombinantly purified Gp45, and an immunogold-labeling was performed. Transmission electron microscopy verified the location of Gp45 and Gp54, a suspected BppU homolog, to be within the tail tube tip. The antisera exhibited a dose-dependent effect when blocking the tail tip with antisera against both Gp45 and Gp54. Neutralization of cells with recombinant Gp45 showed a comparable effect for both variations of experimental setups, plating efficiency of preincubated cells and flow cytometry analysis with biotin-labeled recombinant protein. Mutants deficient in WTA-biosynthesis were also tested, and a direct correlation could be verified between the  $\alpha$ - or  $\beta$ -O-glycosylation capacity of the cells and the Rbp binding ability. It was established that both 6-O acetylation<sup>213</sup> of PG-MurNAc or WTA integrity in general need to be fully functional for phage adsorption to occur. Furthermore, YueB homologues in *S. aureus* denoted Pip (phage infection protein) were ruled out as membrane-protein receptors (see publication in appendix for details)<sup>214</sup>.

### 3.2.2 HHPRED analyses of baseplate components

HHPRED analyses of  $\phi$ 11 baseplate components Gp43, Gp44 and Gp54 revealed similarities with components of the lactococcal phage TP901-1 baseplate. Gp43 is supposed to exist as a hexamer and represents the  $\phi$ 11 distal tail protein (SPP1 Dit) ring and Gp44 as a trimer forming the tail-associated lysin (Tal) N-terminus and extension, while Gp54 N-terminus (the functional equivalent of BppU N-terminus) may form a second ring<sup>215,216</sup>. Furthermore, the N-terminal segment and the first helical bundle of  $\phi$ 11 RBP are structurally homologous to the N-terminal part of the phage TP901-1 or Tuc2009 Rbp trimer, a segment that connects BppU to BppL. The TP901-1 Dit hexamer was fitted together with the BppU N-terminus into the electron density map. The ring of the Dit had appropriate dimensions to fit the map above the Rbps. In contrast, the structure equivalent to BppU could not be

fitted unequivocally as the internal density is not defined sufficiently.

It is not unusual that Dit and Tal could have been conserved through evolution, a phenomenon shared by other components such as the capsids MCP, the connector, as well as the tail MTP. Only the periphery of Dit (its C-terminal domain) and the Tal extension (e.g. a C-terminal fiber) could have been adapted to specific phage infection-style requirements. In contrast,  $\phi$ 11 Gp45, the Rbp, does not exhibit analogy with other phage Rbps, in particular with those from lactococcal phages that also bind to saccharidic receptors (See publication in appendix for details)<sup>214</sup>.

### 3.2.3 Isolation and crystallization

Overproduced Gp45 was isolated according to standard procedures in a purification process involving two steps Immobilized metal affinity chromatography (IMAC) and SEC. The final concentration was adjusted using a concentration-cell.

A first round of sparse matrix condition screening had been performed using a pipetting robot to pipet 10x96 commercial conditions for protein concentrations adjusted to several values between 5 and 15 mg/mL. The successful crystallization conditions were refined by varying pH, additive and cryoprotectant amounts in small ranges. The refinement in 24-well plates using the hanging-drop variation gave bigger crystals (ca. 500x200x50  $\mu$ M). Derivatives were prepared by soaking native crystals in crystallization solution with 2 mM Ta<sub>6</sub>Br<sub>12</sub> for up to 2 weeks before backsoaking briefly in crystallization condition and vitrification in liquid nitrogen<sup>216</sup>. (See publication in appendix for details)

### 3.2.4 Diffraction data collection and phasing

Data for native and derivative crystals were collected at the Swiss Light Source (SLS) on beamline X06DA (PXIII) using a PILATUS 2M hybrid pixel detector. The inverse beam diffraction data collection method was applied for derivative crystals. For the determination of peak, inflection, high-remote and low-remote wavelengths from fluorescence spectra, the program CHOOCH was adjusted to the absorption edges of Ta-L-II and Ta-L-III.

Collected data were processed with the XDS-package. Structure



determination was performed by a combination of SIRAS and non-crystallographic symmetry (NCS) averaging: Anomalous data were processed according to MAD, SAD, MIRAS and SIRAS protocols using SHARP/autoSHARP. Initial heavy atom coordinates and B-factors found with SHELXDE were reedited with the SHARP-module Sushi and refined until electron density maps showed good contrast. The outcome was then used to further refine the atomic substructure and calculate electron density maps for the protein. Resulting electron density maps were inspected for characteristic features of a protein. The outcomes of the various phasing protocols were compared and the map derived from the SIRAS protocol was taken for further improvement. The tables in the appendix section document the most relevant phasing results.

### 3.2.5 NCS and phase improvement

A threefold NCS was elicited from the self-rotation function via `polarrfn` (`ccp4`), giving a strong signal for rotation in reciprocal space for Eulerian angles ( $\alpha=357.4$ ,  $\beta=63.6$ ,  $\gamma=110.6$ ) corresponding to polar angles ( $\theta=37.4$ ,  $\phi=33.4$ ,  $\kappa=120.1$ ). Applying `GETAX` (`ccp4`) using the above set of coordinates, the orientation of the major rotation axis and all of its locations in the asu could be obtained in another set of translation vectors. These results were combined in four rotation-translation matrices, which could be used for improving the map further.

Due to multiple domains the large dimensions of the protein made it necessary to split the molecule for averaging purposes into four domains: 2 different subdomains for the stem, the shoulder domain and the head domain. Four different masks were generated to include at least these 4 structurally different regions. To verify the accuracy of the unrefined matrices in real space, a testwise generation of correlation maps with the combined NCS-operators using `maprot` (`ccp4`) on the phasing solution was tried. By comparing the unbiased map with the correlation maps the NCS-operators could be assigned to the related domains. The refining of these NCS-matrices with `IMP` and a testwise averaging with `AVE` converged the incomplete input electron density to above 80 % of correlation after averaging, indicating a good outcome for the averaging process. The increase of electron density and

the increase of percentage correlation after each averaging round lead to visually interpretable maps, confirming the correctness of the results.

For combined density improvement, each subdomain was integrated with their respective NCS-matrices into a DM script. Starting at 5.8 Å, 80 cycles of consecutive solvent flattening, NCS-averaging, histogram matching and phase extension to a final resolution of 2.9 Å resulted in a final map that was clearly distinguishable from the unbiased calculated map. Main chains and side chains became interpretable so that a large portion of the protein model could be docked. Initial refinements were carried out with REFMAC5 and PHENIX after each step of COOT model building. Finally, refinement was pursued with autoBUSTER leading to  $R_{\text{free}}/R_{\text{work}}$  values of 21.1 % and 17.5 %.

### 3.2.6 Gp45 structure (Rbp)

Although the electron density was generally of good quality, a few loops of the propeller domains (amino-acids 140-440) have very weak electron density, explaining the persistence of a few outliers in the Ramachandran plot. The Rbp of  $\phi 11$  assembles into an elongated homotrimer, with overall dimensions of approximately 160×120×100 Å. The structure consists of the N-terminal stem region, the central shoulder region and the C-terminal head region. A total buried surface of ~35,000 Å<sup>2</sup> is found for the trimer according to PISA ([http://www.ebi.ac.uk/msd-srv/prot\\_int/cgi-bin/piserver](http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver)). Most of the buried surface area is concentrated in the stem and C-terminal regions, while the propeller domains engage in few intermolecular contacts (See publication in appendix for details)<sup>216</sup>.

#### 3.2.6.1 Stem domain

Overall, the stem is composed of 4 triple-helical bundles, connected either by  $\beta$ -hairpins or undefined regions. Between the first two bundles, an Fe<sup>3+</sup> is positioned on the 3-fold axis of the first helix bundle and octahedrally coordinated by the side chains of His42 and His50 from each of the three monomers. The HisN<sub>ε2</sub>-Fe distances range from 2.19 Å to 2.32 Å. This His-Fe-octahedron is surrounded by Glu46, Arg43 and Gln54 from all chains. Similar Fe<sup>3+</sup> binding geometries and bond lengths are known of viral proteins, i.e. membrane-piercing spike proteins of phages P2 (PDB code

3qr7) and  $\phi 92$  (PDB code 3pqh), as well as in the receptor-binding domain of the long tail fiber of phage T4 (PDB code 2xgf).  $\phi 11$  Rbp could act in a similar fashion as it was proposed for phage P2,  $\phi 92$  and T4 spike structures by enforcing the protein to pierce the cell wall.

While the first and second triple helices are collinear, a hinge is introduced into the structure, which forms a sharp angle of  $\sim 30^\circ$  between the second and the third bundle. This angle allows for contacts between stem units and between the stem domain and the shoulder domain. The hinge geometry is such that the sequences of the three helices of bundles two and three remain in phase. Bundles  $\alpha 3$  are followed by extended hairpin-structures, and helix  $\alpha 4$  completes the stem (See publication in appendix for details)<sup>216</sup>.

### 3.2.6.2 Shoulder domain

The stem is followed by five-bladed  $\beta$ -propeller modules of the shoulder domain, which encompass residues 142–439 and form the midsection of the protein, and it is linked via a short helix (residues 425–432) to the C-terminal domain. Between the third helical bundle and the head domain of  $\phi 11$  Rbp is the shoulder, which contains three five-bladed propellers in an area that is  $\sim 100$  Å wide and  $\sim 40$  Å thick. Each interface between two propeller domains buries a surface area of only  $457$  Å<sup>2</sup> from solvent, and much of this surface is buried due to a helix-helix contact at the center of the trimer axis. The plane of the propeller is not perpendicular to the 3-fold axis, but is tilted upwards by an angle of  $\sim 30^\circ$ . This tilt improves access to the lower face of the propeller, and this might be linked to the function of Rbp in interacting with ligands.

Sets of four anti-parallel  $\beta$ -strands form each blade, and the N-terminal  $\beta$ -strand closes the fold by forming the final blade (blade 5) with the three C-terminal  $\beta$ -strands of the domain. Most similar folds are shared by enzymes that mediate degradation of carbohydrates, e.g. glutamine cyclotransferase from *Zymomonas mobilis* (Z=16.1, r.m.s.d.= 3.1 Å, pdb 3nol) or possess such endosialidases of coliphages K1F and phi92. Putative Rbps with  $\beta$ -propellers include a five-bladed propeller of the RBP-P2 head domain of phage PRD1, a *Tectiviridae* member infecting Gram-negative bacteria (PDB

code 1n7u) and the C-terminal domain gp131C of the *Pseudomonas* myophage PhiKZ, forming a seven-bladed  $\beta$ -propeller domain (PDB code 4gbf).

Using the *Zymomonas mobilis*  $\beta$ -propeller structure to superpose onto the Rbp propeller domain, the active site of the enzyme revealed a pocket into which a GlcNAc molecule could be fitted. Three water molecules are indicating hydroxyl groups of a modelled GlcNAc molecule. The cavity is lined with polar residues (Gln165, Thr211, Gln330) and apolar residues Met164 and Met329, which could accommodate the modelled GlcNAc molecule (See publication in appendix for details)<sup>216</sup>.

### 3.2.6.3 Head domain

The head domain contains two structurally similar domains, which are each formed by three five-stranded anti-parallel  $\beta$ -sheets. The two C-terminal tower domains (residues 440–541 and 542–636) form a structure of dimensions  $60 \times 50 \times 50$  Å. These two domains are very similar in structure (C $\alpha$  r.m.s.d.=1.7 Å). A DALI search performed with these domains returned the highest Z-score (Z=6.1; r.m.s.d.=3.2Å) for uracil-DNA glycosylase inhibitor, a small all- $\beta$  monomeric protein (PDB code 2ugi). The two structures essentially share the same anti-parallel  $\beta$ -sheet but differ in their oligomeric state and their surrounding structural features. The identity (9%) is not high enough to assign possible functions to the C-terminal domains.

### 3.2.7 Negative staining electron microscopy of the $\phi$ 11 baseplate

512 images of  $\phi$ 11 were collected, and 778 baseplate particles were boxed. The final map has a resolution of 23 Å and accommodates six  $\phi$ 11 Rbp trimers in lateral position, when the hinge angle between the second and the third helix bundles are increased to a value of  $\sim 90^\circ$ . The correlation is 0.845 with 95.5% of the atoms inside the map, calculated for a Rbp orientation fit with the head domain inclined towards the bottom of the baseplate. A “heads up” conformation results in a correlation of 0.826 with 74.9% of the atoms inside the map. However, it has to be taken into account that the “heads-down” conformation is not a completely non-flexible state. The angle of the stem might in fact vary from the minimum observed in the X-ray structure to larger values when the phage scans the host’s surface by moving the Rbps around the calculated average position for adhesion to the specific receptor. The remainder of the Rbp structure was left unaltered, and the modified trimers fit well in a peripheral region of the map that could accommodate the triangular shape of the shoulder domain.

## 4 FUTURE RESEARCH

### 4.1 TarM

TarM was found to be a GT module in a novel fold. A highly resolved ternary structure with the acceptor substrate would facilitate discerning a mechanism for the glycosylation of WTA and a probable involvement of the DUF1975 in this enzymatic action. It could be shown that one RboP molecule fits in the TarM structure (Simon Huber et al., Data not shown). A TarM structure of around 3.5 Å could pinpoint a density for one RboP-unit on the bottom side of the DUF1975 domain opposite to the sugar-binding cleft. This spot is located inside the cradle which is put up by the C-terminal lobes. However, a better resolved structure, probably with more than one RboP-unit, could allow for a more detailed analysis.

DUF1975 domain functionality was characterized as an oligomerization interface and proposed to be a domain able to present substrate molecules to the activated sugar ready to be transferred. According to enzyme specialization this domain could be able to bind conserved structures. Two substrates specific for DUF1975 are known: RboP-type WTA in the case of TarM and PsrP, a surface adhesin, in the case of GtfA. For a systematic comparison more DUF1975-variants are crucial to derive any sensible deductions from a bioinformatical approach. Although in the case of WTA there is generally much information, the role of WTA is not clarified and does not allow for appointing a role in the environment of the bacterial envelope to which it clearly belongs. However, the only role for DUF1975 we can report at that time is confined to the assemblation capacity of three GT modules and a putative site for binding of a RboP acceptor chain. Conserved portions of this domain in the vicinity of the identified active site region definitely indicates an acceptor site here.

### 4.2 Gp45

A novel fold of three distinct domains, stem shoulder and head, was found in the case of the Gp45 structure. Unexpectedly, the stem contains a bound iron. The function of this iron is unknown as its location differs from irons found in other spike proteins. It would be interesting to see if the stem structure integrity falls apart, when the iron is removed from the coordination center. This could either be achieved by using chelating agents on an isolated solute form or by using

culture media devoid by iron but rather using a non-toxic heavy metal of similar size. Both setups could be conducted by biophysical methods in a way that allows for transient processes to be inspected. The binding motif and other overall structural parameters could be studied structurally when inserting an alternative coordination center and compared to wild-type values.

The invariable TP901-1 BppU protein, seems to be present in parts in Rbp (Gp45) or in the subsequent protein (Gp54). This finding suggests an evolutionary interchange even between remote phages with different hosts. Since the role of Gp54 is not documented, it is difficult to predict any functions for it. Presumably, a subdomain of phage  $\phi$ 11 Gp54 might accommodate the Rbp N-terminus, comparable to TP901-1 BppU. Another isolation attempt at an *in-vivo* preformed or *in-vitro* produced complex between Gp45 and Gp54 could be taken into account. An appropriate SEC-column (e.g. Sephacryl) allowing for separation ranges around 1 MDa would be of crucial material demands in a protein purification procedure of the anticipated high molecular weight Gp45-Gp54 complex.

The molecular orientations as hinted by the electron microscopy low-resolution structure gives and the five-bladed propeller domain suggests roles for the Gp45 subdomains as exposed to the host cell surface. Although, a well-defined cavity in the size of a monosaccharide was identified, a much larger crevice surrounding this cavity could harbour a bigger molecule, of which GlcNAc might be only part. Although soaking and cocrystallization experiments with GlcNAc were not successful, modelling suggests a reasonable location for the GlcNAc binding site in the shoulder domain. However, further structural data are necessary to develop this hypothesis.

A large volume of the EM map remains to account for the Gp54 and for the Tal. When attaching the trimeric Tal N-terminal domain below the Dit hexamer, the three carbohydrate binding modules (2wao) identified by HHpred project in the direction of the tail tip. These three bulky modules should fill the electron density map in between the six RBP trimers. The functional investigation of the singular domains could be carried out analogously to the procedure as outlined before in the case of TP901-1 BppL, by connecting camelid antibody avidity and

enzymatic analysis. Since multiple nanobodies recognize several epitopes of a given molecule, a structural analysis of nanobody-complexes in combination with competitive enzyme assays would help in the characterization attempt.



## 5 REFERENCES

- 1 Baron, S. *Medical microbiology*. 4th edn, (University of Texas Medical Branch at Galveston, 1996).
- 2 Velez, R. & Sloand, E. Combating antibiotic resistance, mitigating future threats and ongoing initiatives. *J Clin Nurs* **25**, 1886-1889, doi:10.1111/jocn.13246 (2016).
- 3 Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T* **40**, 277-283 (2015).
- 4 Berkowitz, F. E. The gram-positive bacilli: a review of the microbiology, clinical aspects, and antimicrobial susceptibilities of a heterogeneous group of bacteria. *Pediatr Infect Dis J* **13**, 1126-1138 (1994).
- 5 Wells, C. L. & Wilkins, T. D. in *Medical Microbiology* (ed S. Baron) (1996).
- 6 Mittwer, T., Bartholomew, J. W. & Kallman, B. J. The mechanism of the gram reaction. II. The function of iodine in the gram stain. *Stain Technol* **25**, 169-179 (1950).
- 7 Adler, H. E. Mycoplasma, the cause of chronic respiratory disease. *Annals of the New York Academy of Sciences* **79**, 703-712 (1960).
- 8 Clostridium difficile and chronic bowel disease. *Lancet* **1**, 402-403 (1980).
- 9 Fornachon, J. C., Douglas, H. C. & Vaughn, R. H. The pH Requirements of Some Heterofermentative Species of Lactobacillus. *Journal of bacteriology* **40**, 649-655 (1940).
- 10 Hopwood, D. A., Wildermuth, H. & Palmer, H. M. Mutants of Streptomyces coelicolor defective in sporulation. *J Gen Microbiol* **61**, 397-408, doi:10.1099/00221287-61-3-397 (1970).
- 11 Burger, M. M. & Glaser, L. The Synthesis of Teichoic Acids. I. Polyglycerophosphate. *The Journal of biological chemistry* **239**, 3168-3177 (1964).
- 12 Baddiley, J. Teichoic Acids and the Bacterial Cell Wall. *Endeavour* **23**, 33-37 (1964).
- 13 Archibald, A. R., Armstrong, J. J., Baddiley, J. & Hay, J. B. Teichoic acids and the structure of bacterial walls. *Nature* **191**, 570-572 (1961).
- 14 Rowley, D., Ali, W. & Jenkin, C. R. A reaction between fresh serum and lipopolysaccharides of gram-negative bacteria. *Br J Exp Pathol* **39**, 90-98 (1958).
- 15 Adams, R. L. & Eason, R. DNA methylation and G+C rich DNA. *Nature* **312**, 407-408 (1984).
- 16 Dams, E. *et al.* Structure of 5S rRNA in actinomycetes and relatives and evolution of eubacteria. *Journal of molecular evolution* **25**, 255-260 (1987).
- 17 Zapun, A., Vernet, T. & Pinho, M. G. The different shapes of cocci. *FEMS microbiology reviews* **32**, 345-360, doi:10.1111/j.1574-6976.2007.00098.x (2008).
- 18 Marie, A. L. & Parak, F. Factors affecting the growth and the catalase synthesis in Micrococcus luteus cells. *Hoppe Seylers Z Physiol Chem* **361**, 603-606 (1980).
- 19 Huebner, J. & Goldmann, D. A. Coagulase-negative staphylococci: role as pathogens. *Annu Rev Med* **50**, 223-236, doi:10.1146/annurev.med.50.1.223 (1999).
- 20 Shen, Y. E., Kurup, V. P. & Fink, J. N. Circulating antibodies against thermophilic actinomycetes in farmers and mushroom workers. *J Hyg Epidemiol Microbiol Immunol* **35**, 309-316 (1991).

- 21 Janetschke, P. [Studies of poultry blood (hen, turkey, goose, duck) for *Listeria agglutinins*]. *Arch Exp Veterinarmed* **22**, 171-179 (1968).
- 22 Clarke, D. J. & Knowles, C. J. The effect of haematin and catalase on *Streptococcus faecalis* var. *zymogenes* growing on glycerol. *J Gen Microbiol* **121**, 339-347, doi:10.1099/00221287-121-2-339 (1980).
- 23 Lorian, V. & Popoola, B. Pneumococci producing beta hemolysis on agar. *Appl Microbiol* **24**, 44-47 (1972).
- 24 Davis, T. E., Fuller, D. D. & Aeschleman, E. C. Rapid, direct identification of *Staphylococcus aureus* and *Streptococcus pneumoniae* from blood cultures using commercial immunologic kits and modified conventional tests. *Diagn Microbiol Infect Dis* **15**, 295-300 (1992).
- 25 Ma, X. X., Wang, E. H., Liu, Y. & Luo, E. J. Antibiotic susceptibility of coagulase-negative staphylococci (CoNS): emergence of teicoplanin-non-susceptible CoNS strains with inducible resistance to vancomycin. *J Med Microbiol* **60**, 1661-1668, doi:10.1099/jmm.0.034066-0 (2011).
- 26 Gastmeier, P. *et al.* [Surgical site infections in hospitals and outpatient settings. Results of the German nosocomial infection surveillance system (KISS)]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **47**, 339-344, doi:10.1007/s00103-004-0805-8 (2004).
- 27 Gastmeier, P. *et al.* [Mortality in German intensive care units: dying from or with a nosocomial infection?]. *Anesthesiol Intensivmed Notfallmed Schmerzther* **40**, 267-272, doi:10.1055/s-2005-861354 (2005).
- 28 Stone, S. P. *et al.* The ORION statement: guidelines for transparent reporting of outbreak reports and intervention studies of nosocomial infection. *Lancet Infect Dis* **7**, 282-288, doi:10.1016/S1473-3099(07)70082-8 (2007).
- 29 Zhang, B. *et al.* Pharmacokinetics of intravenous moxifloxacin in the cerebrospinal fluid of a patient with central nervous system shunt infection. *Diagn Microbiol Infect Dis* **84**, 249-251, doi:10.1016/j.diagmicrobio.2015.10.001 (2016).
- 30 Amat-Santos, I. J. *et al.* Prosthetic valve endocarditis after transcatheter valve replacement: a systematic review. *JACC Cardiovasc Interv* **8**, 334-346, doi:10.1016/j.jcin.2014.09.013 (2015).
- 31 Natsheh, A. *et al.* Prosthetic valve endocarditis due to *Streptococcus pneumoniae*. *Springerplus* **3**, 375, doi:10.1186/2193-1801-3-375 (2014).
- 32 Grubitzsch, H. *et al.* Outcome after surgery for prosthetic valve endocarditis and the impact of preoperative treatment. *J Thorac Cardiovasc Surg* **148**, 2052-2059, doi:10.1016/j.jtcvs.2014.05.025 (2014).
- 33 Brown, P. D. Ciprofloxacin for the management of urinary tract infection. *Womens Health (Lond Engl)* **2**, 509-516, doi:10.2217/17455057.2.4.509 (2006).
- 34 Muder, R. R. *et al.* Isolation of *Staphylococcus aureus* from the urinary tract: association of isolation with symptomatic urinary tract infection and subsequent staphylococcal bacteremia. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **42**, 46-50, doi:10.1086/498518 (2006).
- 35 Suman, E., Jose, J., Varghese, S. & Kotian, M. S. Study of biofilm production in *Escherichia coli* causing urinary tract infection. *Indian J Med Microbiol* **25**, 305-306 (2007).

- 36 Anijeet, D. Endophthalmitis after cataract surgery. *Ophthalmology* **117**, 853-853 e851; author reply 853-855, doi:10.1016/j.ophtha.2010.01.006 (2010).
- 37 Vuong, C. & Otto, M. The biofilm exopolysaccharide polysaccharide intercellular adhesin--a molecular and biochemical approach. *Methods in molecular biology* **431**, 97-106 (2008).
- 38 Lin, H., Lin, Z., Meng, C., Huang, J. & Guo, Y. Preparation and immunogenicity of capsular polysaccharide-surface adhesin A (PsaA) conjugate of *Streptococcus pneumoniae*. *Immunobiology* **215**, 545-550, doi:10.1016/j.imbio.2009.08.008 (2010).
- 39 O'Riordan, K. & Lee, J. C. Staphylococcus aureus capsular polysaccharides. *Clinical microbiology reviews* **17**, 218-234 (2004).
- 40 Kogan, G., Sadovskaya, I., Chaignon, P., Chokr, A. & Jabbouri, S. Biofilms of clinical strains of Staphylococcus that do not contain polysaccharide intercellular adhesin. *FEMS Microbiol Lett* **255**, 11-16, doi:10.1111/j.1574-6968.2005.00043.x (2006).
- 41 Suarez, M. J., Arias, M. L. & del Mar Gamboa, M. [Staphylococcus aureus enterotoxin A detection using the polymerase chain reaction (PCR) and its correlation with coagulase and thermonuclease tests]. *Arch Latinoam Nutr* **58**, 59-63 (2008).
- 42 Xia, S. R. [Thermonuclease assay--a rapid method for screening Staphylococcus aureus food-poisoning]. *Zhonghua Yu Fang Yi Xue Za Zhi* **22**, 164-166 (1988).
- 43 Leighton, P. M. & White, M. Rapid determination of novobiocin susceptibility for the identification of Staphylococcus saprophyticus. *Diagn Microbiol Infect Dis* **1**, 261-264 (1983).
- 44 Akerele, J. O., Obasuyi, O., Onyeagwara, N. & Ottih, I. Methicillin-Resistant Staphylococcus aureus (MRSA): An Emerging Phenomenon among non-Hospitalized Otorhinolaryngological Patients in Benin City, Nigeria. *West Afr J Med* **33**, 276-279 (2014).
- 45 Imaura, M. *et al.* Initial dosing regimen of vancomycin to achieve early therapeutic plasma concentration in critically ill patients with MRSA infection based on APACHE II score. *Eur J Drug Metab Pharmacokinet* **41**, 211-218, doi:10.1007/s13318-014-0246-1 (2016).
- 46 Kerschner, H., Harrison, E. M., Hartl, R., Holmes, M. A. & Apfalter, P. First report of mecC MRSA in human samples from Austria: molecular characteristics and clinical data. *New Microbes New Infect* **3**, 4-9, doi:10.1016/j.nmni.2014.11.001 (2015).
- 47 Ng, M. *et al.* Induction of MRSA Biofilm by Low-Dose beta-Lactam Antibiotics: Specificity, Prevalence and Dose-Response Effects. *Dose Response* **12**, 152-161, doi:10.2203/dose-response.13-021.Kaplan (2014).
- 48 Jansson, J. A. A Direct Spectrophotometric Assay for Penicillin Beta-Lactamase (Penicillinase). *Biochimica et biophysica acta* **99**, 171-172 (1965).
- 49 Stewart, G. T. Changes in Sensitivity of Staphylococci to Methicillin. *Br Med J* **1**, 863-866 (1961).
- 50 Douthwaite, A. H., Trafford, J. A., Mc, G. D. & Evans, I. E. Methicillin. *Br Med J* **2**, 6-8 (1961).
- 51 Callaghan, R. P., Cohen, S. J. & Stewart, G. T. Septicaemia due to colonization of Spitz-Holter valves by staphylococci. Five cases treated with methicillin. *Br Med J* **1**, 860-863 (1961).

- 52 Yourassowsky, E. [Treatment of resistant staphylococcal disorders with sodium methicillin]. *Acta Clin Belg* **16**, 259-270 (1961).
- 53 Coia, J. E., Browning, L., Haines, L., Birkbeck, T. H. & Platt, D. J. Comparison of enterotoxins and haemolysins produced by methicillin-resistant (MRSA) and sensitive (MSSA) *Staphylococcus aureus*. *J Med Microbiol* **36**, 164-171, doi:10.1099/00222615-36-3-164 (1992).
- 54 Melzer, M., Eykyn, S. J., Gransden, W. R. & Chinn, S. Is methicillin-resistant *Staphylococcus aureus* more virulent than methicillin-susceptible *S. aureus*? A comparative cohort study of British patients with nosocomial infection and bacteremia. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **37**, 1453-1460, doi:10.1086/379321 (2003).
- 55 Conn, H. J. A Possible Function of Actinomycetes in Soil. *Journal of bacteriology* **1**, 197-207 (1916).
- 56 Waksman, S. A. On the Classification of Actinomycetes. *Journal of bacteriology* **39**, 549-558 (1940).
- 57 Waksman, S. A. & Henrici, A. T. The Nomenclature and Classification of the Actinomycetes. *Journal of bacteriology* **46**, 337-341 (1943).
- 58 Naumova, I. The teichoic acids of actinomycetes. *Microbiol Sci* **5**, 275-279 (1988).
- 59 Streshinskaya, G. M., Shashkov, A. S., Usov, A. I., Evtushenko, L. I. & Naumova, I. B. Cell wall teichoic acids of actinomycetes of three genera of the order actinomycetales. *Biochemistry (Mosc)* **67**, 778-785 (2002).
- 60 Distler, W. & Hagemann, D. [Starch breakdown by oral Streptococci and Actinomycetes]. *Dtsch Zahnarztl Z* **37**, 673-676 (1982).
- 61 Skujins, J., Pukite, A. & McLaren, A. D. Chitinase of *Streptomyces* sp: purification and properties. *Enzymologia* **39**, 353-370 (1970).
- 62 Norton, J. H. Tuberculosis-like granulomas in cattle caused by Actinomycetes. *Aust Vet J* **52**, 455-457 (1976).
- 63 Mackinnon, J. E. & Artagaveytiaallende, R. C. The main species of pathogenic aerobic actinomycetes causing mycetomas. *Trans R Soc Trop Med Hyg* **50**, 31-40 (1956).
- 64 Buscher-Daubenbuchel, L., Gillissen, G., Harth, H. & Turba, F. [Demonstration of mycolic acid and of phthionic acid in *Mycobacterium tuberculosis*]. *Zentralbl Bakteriol Orig* **167**, 298-306 (1956).
- 65 Gordon, R. E. The Classification of Acid-Fast Bacteria. *Journal of bacteriology* **34**, 617-630 (1937).
- 66 Berthrong, M. The macrophage in tuberculosis. *Bibl Tuberc* **26**, 1-27 (1970).
- 67 Kaufmann, S. H. & Flesch, I. E. The role of T cell-macrophage interactions in tuberculosis. *Springer Semin Immunopathol* **10**, 337-358 (1988).
- 68 Waksman, S. A. Species concept among the actinomycetes with special reference to the genus *Streptomyces*. *Bacteriol Rev* **21**, 1-29 (1957).
- 69 Kalakutskii, L. V., Nikitina, N. I. & Artamonova, O. I. [Spore germination in actinomycetes]. *Mikrobiologiya* **38**, 834-841 (1969).
- 70 Kalakoutskii, L. V. & Agre, N. S. Endospores of actinomycetes: dormancy and germination. *Soc Appl Bacteriol Symp Ser* **2**, 179-195 (1973).

- 71 Skinner, F. A. A method for distinguishing between viable spores and mycelial fragments of actinomycetes in soils. *J Gen Microbiol* **5**, 159-166, doi:10.1099/00221287-5-1-159 (1951).
- 72 Burkholder, P. R. Studies on the antibiotic activity of actinomycetes. *Journal of bacteriology* **52**, 503 (1946).
- 73 Stapley, E. O. *et al.* Cephamycins, a new family of beta-lactam antibiotics. I. Production by actinomycetes, including *Streptomyces lactamdurans* sp. n. *Antimicrobial agents and chemotherapy* **2**, 122-131 (1972).
- 74 Abou-Zeid, A. Z. & Abd El Hamid, M. M. Actinomycetes producing antibiotics. *Zentralbl Bakteriell Parasitenkd Infektionskr Hyg* **130**, 541-555 (1975).
- 75 Hotta, K., Takahashi, A., Okami, Y. & Umezawa, H. Relationship between antibiotic resistance and antibiotic productivity in actinomycetes which produce aminoglycoside antibiotics. *J Antibiot (Tokyo)* **36**, 1789-1791 (1983).
- 76 Procopio, R. E., Silva, I. R., Martins, M. K., Azevedo, J. L. & Araujo, J. M. Antibiotics produced by *Streptomyces*. *Braz J Infect Dis* **16**, 466-471, doi:10.1016/j.bjid.2012.08.014 (2012).
- 77 Horinouchi, S. & Beppu, T. Autoregulatory factors of secondary metabolism and morphogenesis in actinomycetes. *Crit Rev Biotechnol* **10**, 191-204, doi:10.3109/07388559009038207 (1990).
- 78 Cundliffe, E. Antibiotic production by actinomycetes: the Janus faces of regulation. *J Ind Microbiol Biotechnol* **33**, 500-506, doi:10.1007/s10295-006-0083-6 (2006).
- 79 Barker, A. E. On Some Points regarding the Distribution of *Bacillus Anthracis* in the Human Skin in Malignant Pustule. *Med Chir Trans* **69**, 127-134 121 (1886).
- 80 Callegan, M. C., Jett, B. D., Hancock, L. E. & Gilmore, M. S. Role of hemolysin BL in the pathogenesis of extraintestinal *Bacillus cereus* infection assessed in an endophthalmitis model. *Infection and immunity* **67**, 3357-3366 (1999).
- 81 Peng, D. *et al.* A novel metalloproteinase virulence factor is involved in *Bacillus thuringiensis* pathogenesis in nematodes and insects. *Environ Microbiol* **18**, 846-862, doi:10.1111/1462-2920.13069 (2016).
- 82 Gao, C. H. *et al.* Antibacterial and antilarval compounds from marine gorgonian-associated bacterium *Bacillus amyloliquefaciens* SCSIO 00856. *J Antibiot (Tokyo)* **63**, 191-193, doi:10.1038/ja.2010.7 (2010).
- 83 Raza, W., Ling, N., Yang, L., Huang, Q. & Shen, Q. Response of tomato wilt pathogen *Ralstonia solanacearum* to the volatile organic compounds produced by a biocontrol strain *Bacillus amyloliquefaciens* SQR-9. *Sci Rep* **6**, 24856, doi:10.1038/srep24856 (2016).
- 84 Vidal-Quist, J. C., Rogers, H. J., Mahenthiralingam, E. & Berry, C. *Bacillus thuringiensis* colonises plant roots in a phylogeny-dependent manner. *FEMS microbiology ecology* **86**, 474-489, doi:10.1111/1574-6941.12175 (2013).
- 85 Zhou, Y., Choi, Y. L., Sun, M. & Yu, Z. Novel roles of *Bacillus thuringiensis* to control plant diseases. *Appl Microbiol Biotechnol* **80**, 563-572, doi:10.1007/s00253-008-1610-3 (2008).
- 86 Rettger, L. F. On the Germination of the Spores of *B. Subtilis* and *B. Megatherium*. *Public Health Pap Rep* **28**, 541-542 (1902).

- 87 Beeby, M., Gumbart, J. C., Roux, B. & Jensen, G. J. Architecture and assembly of the Gram-positive cell wall. *Molecular microbiology* **88**, 664-672, doi:10.1111/mmi.12203 (2013).
- 88 Vollmer, W., Blanot, D. & de Pedro, M. A. Peptidoglycan structure and architecture. *FEMS microbiology reviews* **32**, 149-167, doi:10.1111/j.1574-6976.2007.00094.x (2008).
- 89 Strahl, H., Burmann, F. & Hamoen, L. W. The actin homologue MreB organizes the bacterial cell membrane. *Nature communications* **5**, 3442, doi:10.1038/ncomms4442 (2014).
- 90 Simpson, L. L. *Botulinum neurotoxin and tetanus toxin*. (Academic Press, 1989).
- 91 Scholars Portal Books: Legislative Library of Ontario. 1 online resource (s.n., S.I., 2007).
- 92 Hoffman, R. M. *Bacterial therapy of cancer : methods and protocols*. (Humana Press, 2016).
- 93 Brown, C. D. *Fermentation of milk by Lactobacillus bifidus*, (1970).
- 94 National Agricultural Library. (s.n., S.I.).
- 95 National Agricultural Library. (s.n., S.I.).
- 96 Schaffer, C. & Messner, P. The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together. *Microbiology* **151**, 643-651, doi:10.1099/mic.0.27749-0 (2005).
- 97 Costerton, J. W., Irvin, R. T. & Cheng, K. J. The bacterial glycocalyx in nature and disease. *Annual review of microbiology* **35**, 299-324, doi:10.1146/annurev.mi.35.100181.001503 (1981).
- 98 Gerke, C., Kraft, A., Sussmuth, R., Schweitzer, O. & Gotz, F. Characterization of the N-acetylglucosaminyltransferase activity involved in the biosynthesis of the Staphylococcus epidermidis polysaccharide intercellular adhesin. *The Journal of biological chemistry* **273**, 18586-18593 (1998).
- 99 Weidenmaier, C. & Peschel, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature reviews. Microbiology* **6**, 276-287, doi:10.1038/nrmicro1861 (2008).
- 100 Endl, J., Seidl, H. P., Fiedler, F. & Schleifer, K. H. Chemical composition and structure of cell wall teichoic acids of staphylococci. *Arch Microbiol* **135**, 215-223 (1983).
- 101 Kern, J. *et al.* Structure of surface layer homology (SLH) domains from Bacillus anthracis surface array protein. *The Journal of biological chemistry* **286**, 26042-26049, doi:10.1074/jbc.M111.248070 (2011).
- 102 Toubiana, R., Berlan, J., Sato, H. & Strain, M. Three types of mycolic acid from Mycobacterium tuberculosis Brevanne: implications for structure-function relationships in pathogenesis. *Journal of bacteriology* **139**, 205-211 (1979).
- 103 Sieradzki, K. & Tomasz, A. Alterations of cell wall structure and metabolism accompany reduced susceptibility to vancomycin in an isogenic series of clinical isolates of Staphylococcus aureus. *Journal of bacteriology* **185**, 7103-7110 (2003).
- 104 Meroueh, S. O. *et al.* Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 4404-4409, doi:10.1073/pnas.0510182103 (2006).
- 105 Reichmann, N. T. & Grundling, A. Location, synthesis and function of glycolipids and

- polyglycerolphosphate lipoteichoic acid in Gram-positive bacteria of the phylum Firmicutes. *FEMS Microbiol Lett* **319**, 97-105, doi:10.1111/j.1574-6968.2011.02260.x (2011).
- 106 Clements, A. *et al.* The major surface-associated saccharides of *Klebsiella pneumoniae* contribute to host cell association. *PLoS one* **3**, e3817, doi:10.1371/journal.pone.0003817 (2008).
- 107 Weidenmaier, C. *et al.* Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. *Nature medicine* **10**, 243-245, doi:10.1038/nm991 (2004).
- 108 Koprivnjak, T., Weidenmaier, C., Peschel, A. & Weiss, J. P. Wall teichoic acid deficiency in *Staphylococcus aureus* confers selective resistance to mammalian group IIA phospholipase A(2) and human beta-defensin 3. *Infection and immunity* **76**, 2169-2176, doi:10.1128/IAI.01705-07 (2008).
- 109 Weidenmaier, C., McLoughlin, R. M. & Lee, J. C. The zwitterionic cell wall teichoic acid of *Staphylococcus aureus* provokes skin abscesses in mice by a novel CD4+ T-cell-dependent mechanism. *PLoS one* **5**, e13227, doi:10.1371/journal.pone.0013227 (2010).
- 110 Baur, S. *et al.* A nasal epithelial receptor for *Staphylococcus aureus* WTA governs adhesion to epithelial cells and modulates nasal colonization. *PLoS pathogens* **10**, e1004089, doi:10.1371/journal.ppat.1004089 (2014).
- 111 Herbert, S. *et al.* Molecular basis of resistance to muramidase and cationic antimicrobial peptide activity of lysozyme in staphylococci. *PLoS pathogens* **3**, e102, doi:10.1371/journal.ppat.0030102 (2007).
- 112 Bera, A. *et al.* Influence of wall teichoic acid on lysozyme resistance in *Staphylococcus aureus*. *Journal of bacteriology* **189**, 280-283, doi:10.1128/JB.01221-06 (2007).
- 113 Bera, A., Herbert, S., Jakob, A., Vollmer, W. & Gotz, F. Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of *Staphylococcus aureus*. *Molecular microbiology* **55**, 778-787, doi:10.1111/j.1365-2958.2004.04446.x (2005).
- 114 Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research* **42**, D490-495, doi:10.1093/nar/gkt1178 (2014).
- 115 Park, B. H., Karpinets, T. V., Syed, M. H., Leuze, M. R. & Uberbacher, E. C. CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology* **20**, 1574-1584, doi:10.1093/glycob/cwq106 (2010).
- 116 Robbins, J. B., Schneerson, R., Xie, G., Hanson, L. A. & Miller, M. A. Capsular polysaccharide vaccine for Group B *Neisseria meningitidis*, *Escherichia coli* K1, and *Pasteurella haemolytica* A2. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 17871-17875, doi:10.1073/pnas.1114489108 (2011).
- 117 Gening, M. L. *et al.* Synthetic  $\beta$ -(1 $\rightarrow$ 6)-linked N-acetylated and nonacetylated oligoglucosamines used to produce conjugate vaccines for bacterial pathogens. *Infection and immunity* **78**, 764-772, doi:10.1128/IAI.01093-09 (2010).
- 118 Neuhaus, F. C. & Baddiley, J. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**, 686-723 (2003).

- 119 O'Gara, J. P. *ica* and beyond: biofilm mechanisms and regulation in *Staphylococcus epidermidis* and *Staphylococcus aureus*. *FEMS Microbiol Lett* **270**, 179-188, doi:10.1111/j.1574-6968.2007.00688.x (2007).
- 120 Johnson, J. W., Fisher, J. F. & Mobashery, S. Bacterial cell-wall recycling. *Annals of the New York Academy of Sciences* **1277**, 54-75, doi:10.1111/j.1749-6632.2012.06813.x (2013).
- 121 Yamaguchi, T. *et al.* Inhibitors for Bacterial Cell-Wall Recycling. *ACS Med Chem Lett* **3**, 238-242, doi:10.1021/ml2002746 (2012).
- 122 Scheurwater, E., Reid, C. W. & Clarke, A. J. Lytic transglycosylases: bacterial space-making autolysins. *Int J Biochem Cell Biol* **40**, 586-591, doi:10.1016/j.biocel.2007.03.018 (2008).
- 123 Suvorov, M., Lee, M., Heseck, D., Boggess, B. & Mobashery, S. Lytic transglycosylase MltB of *Escherichia coli* and its role in recycling of peptidoglycan strands of bacterial cell wall. *Journal of the American Chemical Society* **130**, 11878-11879, doi:10.1021/ja805482b (2008).
- 124 Vollmer, W., Joris, B., Charlier, P. & Foster, S. Bacterial peptidoglycan (murein) hydrolases. *FEMS microbiology reviews* **32**, 259-286, doi:10.1111/j.1574-6976.2007.00099.x (2008).
- 125 Frirdich, E. & Gaynor, E. C. Peptidoglycan hydrolases, bacterial shape, and pathogenesis. *Current opinion in microbiology* **16**, 767-778, doi:10.1016/j.mib.2013.09.005 (2013).
- 126 Doyle, R. J. & Koch, A. L. The functions of autolysins in the growth and division of *Bacillus subtilis*. *Critical reviews in microbiology* **15**, 169-222, doi:10.3109/10408418709104457 (1987).
- 127 Smith, T. J. & Foster, S. J. Characterization of the involvement of two compensatory autolysins in mother cell lysis during sporulation of *Bacillus subtilis* 168. *Journal of bacteriology* **177**, 3855-3862 (1995).
- 128 Rogers, H. J. & Taylor, C. Autolysins and shape change in *rodA* mutants of *Bacillus subtilis*. *Journal of bacteriology* **135**, 1032-1042 (1978).
- 129 Chastanet, A. & Losick, R. Engulfment during sporulation in *Bacillus subtilis* is governed by a multi-protein complex containing tandemly acting autolysins. *Molecular microbiology* **64**, 139-152, doi:10.1111/j.1365-2958.2007.05652.x (2007).
- 130 Wheeler, R. *et al.* Bacterial Cell Enlargement Requires Control of Cell Wall Stiffness Mediated by Peptidoglycan Hydrolases. *mBio* **6**, e00660, doi:10.1128/mBio.00660-15 (2015).
- 131 Zapun, A., Contreras-Martel, C. & Vernet, T. Penicillin-binding proteins and beta-lactam resistance. *FEMS microbiology reviews* **32**, 361-385, doi:10.1111/j.1574-6976.2007.00095.x (2008).
- 132 Lopez, R., Ronda, C. & Garcia, E. Autolysins are directly involved in the bactericidal effect caused by penicillin in wild type and in tolerant pneumococci. *FEMS Microbiol Lett* **54**, 317-322 (1990).
- 133 Isturiz, R., Metcalf, J. A. & Root, R. K. Enhanced killing of penicillin-treated gram-positive cocci by human granulocytes: role of bacterial autolysins, catalase, and granulocyte oxidative pathways. *Yale J Biol Med* **58**, 133-143 (1985).
- 134 Lim, D. & Strynadka, N. C. Structural basis for the beta lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus*. *Nat Struct Biol* **9**, 870-876,



- doi:10.1038/nsb858 (2002).
- 135 Rowland, S. J. & Dyke, K. G. Characterization of the staphylococcal beta-lactamase transposon Tn552. *The EMBO journal* **8**, 2761-2773 (1989).
- 136 Clarke, S. R. & Dyke, K. G. Studies of the operator region of the *Staphylococcus aureus* beta-lactamase operon. *J Antimicrob Chemother* **47**, 377-389 (2001).
- 137 Rice, L. B. *et al.* Enterococcus faecium low-affinity pbp5 is a transferable determinant. *Antimicrobial agents and chemotherapy* **49**, 5007-5012, doi:10.1128/AAC.49.12.5007-5012.2005 (2005).
- 138 Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A. & Charlier, P. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS microbiology reviews* **32**, 234-258, doi:10.1111/j.1574-6976.2008.00105.x (2008).
- 139 Baddiley, J. Bacterial cell walls and membranes. Discovery of the teichoic acids. *Bioessays* **10**, 207-210, doi:10.1002/bies.950100607 (1989).
- 140 Yokoyama, K., Mizuguchi, H., Araki, Y., Kaya, S. & Ito, E. Biosynthesis of linkage units for teichoic acids in gram-positive bacteria: distribution of related enzymes and their specificities for UDP-sugars and lipid-linked intermediates. *Journal of bacteriology* **171**, 940-946 (1989).
- 141 Araki, Y. & Ito, E. Linkage units in cell walls of gram-positive bacteria. *Critical reviews in microbiology* **17**, 121-135, doi:10.3109/10408418909105745 (1989).
- 142 Kojima, N., Araki, Y. & Ito, E. Structure of the linkage units between ribitol teichoic acids and peptidoglycan. *Journal of bacteriology* **161**, 299-306 (1985).
- 143 Coley, J., Duckworth, M. & Baddiley, J. The occurrence of lipoteichoic acids in the membranes of gram-positive bacteria. *J Gen Microbiol* **73**, 587-591, doi:10.1099/00221287-73-3-587 (1972).
- 144 Coley, J., Duckworth, M. & Baddiley, J. Extraction and purification of lipoteichoic acids from Gram-positive bacteria. *Carbohydr Res* **40**, 41-52 (1975).
- 145 Fischer, W. Physiology of lipoteichoic acids in bacteria. *Adv Microb Physiol* **29**, 233-302 (1988).
- 146 Schurek, J. & Fischer, W. Distribution analyses of chain substituents of lipoteichoic acids by chemical degradation. *Eur J Biochem* **186**, 649-655 (1989).
- 147 Fischer, W., Mannsfeld, T. & Hagen, G. On the basic structure of poly(glycerophosphate) lipoteichoic acids. *Biochem Cell Biol* **68**, 33-43 (1990).
- 148 Ruhland, G. J. & Fiedler, F. Occurrence and structure of lipoteichoic acids in the genus *Staphylococcus*. *Arch Microbiol* **154**, 375-379 (1990).
- 149 Sutcliffe, I. C. & Shaw, N. Atypical lipoteichoic acids of gram-positive bacteria. *Journal of bacteriology* **173**, 7065-7069 (1991).
- 150 Xia, G. & Peschel, A. Toward the pathway of *S. aureus* WTA biosynthesis. *Chemistry & biology* **15**, 95-96, doi:10.1016/j.chembiol.2008.02.005 (2008).
- 151 Ginsberg, C., Zhang, Y. H., Yuan, Y. & Walker, S. In vitro reconstitution of two essential steps in wall teichoic acid biosynthesis. *ACS Chem Biol* **1**, 25-28, doi:10.1021/cb0500041 (2006).
- 152 Brown, S., Zhang, Y. H. & Walker, S. A revised pathway proposed for *Staphylococcus aureus* wall teichoic acid biosynthesis based on in vitro reconstitution of the intracellular steps. *Chemistry & biology* **15**, 12-21, doi:10.1016/j.chembiol.2007.11.011 (2008).

- 153 Meredith, T. C., Swoboda, J. G. & Walker, S. Late-stage polyribitol phosphate wall teichoic acid biosynthesis in *Staphylococcus aureus*. *Journal of bacteriology* **190**, 3046-3056, doi:10.1128/JB.01880-07 (2008).
- 154 Swoboda, J. G., Campbell, J., Meredith, T. C. & Walker, S. Wall teichoic acid function, biosynthesis, and inhibition. *ChemBiochem* **11**, 35-45, doi:10.1002/cbic.200900557 (2010).
- 155 Brown, S., Santa Maria, J. P., Jr. & Walker, S. Wall teichoic acids of gram-positive bacteria. *Annual review of microbiology* **67**, 313-336, doi:10.1146/annurev-micro-092412-155620 (2013).
- 156 Wicken, A. J. & Baddiley, J. Structure of intracellular teichoic acids from group D streptococci. *The Biochemical journal* **87**, 54-62 (1963).
- 157 Archibald, A. R., Baddiley, J., Heckels, J. E. & Heptinstall, S. Further studies on the glycerol teichoic acid of walls of *Staphylococcus lactis* I3. Location of the phosphodiester groups and their susceptibility to hydrolysis with alkali. *The Biochemical journal* **125**, 353-359 (1971).
- 158 Shashkov, A. S., Streshinskaya, G. M., Gnilozub, V. A., Evtushenko, L. I. & Naumova, I. B. Poly(arabitol phosphate) teichoic acid in the cell wall of *Agromyces cerinus* subsp. *cerinus* VKM Ac-1340T. *FEBS letters* **371**, 163-166 (1995).
- 159 Potekhina, N. V., Tul'skaya, E. M., Naumova, I. B., Shashkov, A. S. & Evtushenko, L. I. Erythritolteichoic acid in the cell wall of *Glycomyces tenuis* VKM Ac-1250. *Eur J Biochem* **218**, 371-375 (1993).
- 160 Xia, G. *et al.* Glycosylation of wall teichoic acid in *Staphylococcus aureus* by TarM. *The Journal of biological chemistry* **285**, 13405-13415, doi:10.1074/jbc.M109.096172 (2010).
- 161 Brown, S. *et al.* Methicillin resistance in *Staphylococcus aureus* requires glycosylated wall teichoic acids. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18909-18914, doi:10.1073/pnas.1209126109 (2012).
- 162 Lovering, A. L. *et al.* Structure of the bacterial teichoic acid polymerase TagF provides insights into membrane association and catalysis. *Nature structural & molecular biology* **17**, 582-589, doi:10.1038/nsmb.1819 (2010).
- 163 Baur, S., Marles-Wright, J., Buckenmaier, S., Lewis, R. J. & Vollmer, W. Synthesis of CDP-activated ribitol for teichoic acid precursors in *Streptococcus pneumoniae*. *Journal of bacteriology* **191**, 1200-1210, doi:10.1128/JB.01120-08 (2009).
- 164 Velloso, L. M., Bhaskaran, S. S., Schuch, R., Fischetti, V. A. & Stebbins, C. E. A structural basis for the allosteric regulation of non-hydrolysing UDP-GlcNAc 2-epimerases. *EMBO reports* **9**, 199-205, doi:10.1038/sj.embor.7401154 (2008).
- 165 Kawai, Y. *et al.* A widespread family of bacterial cell wall assembly proteins. *The EMBO journal* **30**, 4931-4941, doi:10.1038/emboj.2011.358 (2011).
- 166 Campeotto, I. *et al.* Structural and mechanistic insight into the *Listeria monocytogenes* two-enzyme lipoteichoic acid synthesis system. *The Journal of biological chemistry* **289**, 28054-28069, doi:10.1074/jbc.M114.590570 (2014).
- 167 Kiriukhin, M. Y. & Neuhaus, F. C. D-alanylation of lipoteichoic acid: role of the D-alanyl carrier protein in acylation. *Journal of bacteriology* **183**, 2051-2058, doi:10.1128/JB.183.6.2051-2058.2001 (2001).
- 168 Comstock, L. E. & Kasper, D. L. Bacterial glycans: key mediators of diverse host

- immune responses. *Cell* **126**, 847-850, doi:10.1016/j.cell.2006.08.021 (2006).
- 169 van Kessel, K. P., Bestebroer, J. & van Strijp, J. A. Neutrophil-Mediated Phagocytosis of *Staphylococcus aureus*. *Front Immunol* **5**, 467, doi:10.3389/fimmu.2014.00467 (2014).
- 170 Kowanko, I. C., Ferrante, A., Clemente, G. & Kumaratilake, L. M. Tumor necrosis factor primes neutrophils to kill *Staphylococcus aureus* by an oxygen-dependent mechanism and *Plasmodium falciparum* by an oxygen-independent mechanism. *Infection and immunity* **64**, 3435-3437 (1996).
- 171 Hampton, M. B., Kettle, A. J. & Winterbourn, C. C. Involvement of superoxide and myeloperoxidase in oxygen-dependent killing of *Staphylococcus aureus* by neutrophils. *Infection and immunity* **64**, 3512-3517 (1996).
- 172 Peschel, A. & Sahl, H. G. The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature reviews. Microbiology* **4**, 529-536, doi:10.1038/nrmicro1441 (2006).
- 173 Medzhitov, R. Toll-like receptors and innate immunity. *Nat Rev Immunol* **1**, 135-145, doi:10.1038/35100529 (2001).
- 174 Akhmatova, N. K., Egorova, N. B., Kurbatova, E. A. & Akhmatov, E. A. Activation of innate immunity by bacterial ligands of toll-like receptors. *Front Immunol* **5**, 89, doi:10.3389/fimmu.2014.00089 (2014).
- 175 Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389-395, doi:10.1038/415389a (2002).
- 176 Dua, H. S., Otri, A. M., Hopkinson, A. & Mohammed, I. In vitro studies on the antimicrobial peptide human beta-defensin 9 (HBD9): signalling pathways and pathogen-related response (an American Ophthalmological Society thesis). *Trans Am Ophthalmol Soc* **112**, 50-73 (2014).
- 177 Hollands, A. *et al.* A bacterial pathogen co-opts host plasmin to resist killing by cathelicidin antimicrobial peptides. *The Journal of biological chemistry* **287**, 40891-40897, doi:10.1074/jbc.M112.404582 (2012).
- 178 Agerberth, B. *et al.* The human antimicrobial and chemotactic peptides LL-37 and alpha-defensins are expressed by specific lymphocyte and monocyte populations. *Blood* **96**, 3086-3093 (2000).
- 179 Williamson, S. I. *et al.* LPS regulation of the immune response: separate mechanisms for murine B cell activation by lipid A (direct) and polysaccharide (macrophage-dependent) derived from *Bacteroides* LPS. *Journal of immunology* **133**, 2294-2300 (1984).
- 180 Yauch, L. E., Mansour, M. K., Shoham, S., Rottman, J. B. & Levitz, S. M. Involvement of CD14, toll-like receptors 2 and 4, and MyD88 in the host response to the fungal pathogen *Cryptococcus neoformans* in vivo. *Infection and immunity* **72**, 5373-5382, doi:10.1128/IAI.72.9.5373-5382.2004 (2004).
- 181 Pluddemann, A., Mukhopadhyay, S. & Gordon, S. The interaction of macrophage receptors with bacterial ligands. *Expert Rev Mol Med* **8**, 1-25, doi:10.1017/S1462399406000159 (2006).
- 182 Bell, J. K. *et al.* Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol* **24**, 528-533 (2003).
- 183 Loyi, T., Kumar, H., Nandi, S. & Patra, M. K. Expression of pathogen recognition receptors and pro-inflammatory cytokine transcripts in clinical and sub-clinical endometritis cows. *Anim Biotechnol* **26**, 194-200, doi:10.1080/10495398.2014.987389

- (2015).
- 184 Pulendran, B. Variagation of the immune response with dendritic cells and pathogen recognition receptors. *Journal of immunology* **174**, 2457-2465 (2005).
- 185 Wilsson, A., Lind, S., Ohman, L., Nilsson-Augustinsson, A. & Lundqvist-Setterud, H. Apoptotic neutrophils containing *Staphylococcus epidermidis* stimulate macrophages to release the proinflammatory cytokines tumor necrosis factor- $\alpha$  and interleukin-6. *FEMS Immunol Med Microbiol* **53**, 126-135, doi:10.1111/j.1574-695X.2008.00412.x (2008).
- 186 Vance, R. E., Isberg, R. R. & Portnoy, D. A. Patterns of pathogenesis: discrimination of pathogenic and nonpathogenic microbes by the innate immune system. *Cell Host Microbe* **6**, 10-21, doi:10.1016/j.chom.2009.06.007 (2009).
- 187 Koonin, E. V. & Wolf, Y. I. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic acids research* **36**, 6688-6719, doi:10.1093/nar/gkn668 (2008).
- 188 Christie, G. E. & Dokland, T. Pirates of the Caudovirales. *Virology* **434**, 210-221, doi:10.1016/j.virol.2012.10.028 (2012).
- 189 Gogarten, J. P., Doolittle, W. F. & Lawrence, J. G. Prokaryotic evolution in light of gene transfer. *Mol Biol Evol* **19**, 2226-2238 (2002).
- 190 Novick, R. P., Christie, G. E. & Penades, J. R. The phage-related chromosomal islands of Gram-positive bacteria. *Nature reviews. Microbiology* **8**, 541-551, doi:10.1038/nrmicro2393 (2010).
- 191 Raisanen, L., Schubert, K., Jaakonsaari, T. & Alatosava, T. Characterization of lipoteichoic acids as *Lactobacillus delbrueckii* phage receptor components. *Journal of bacteriology* **186**, 5529-5532, doi:10.1128/JB.186.16.5529-5532.2004 (2004).
- 192 McCabe, O. *et al.* The targeted recognition of *Lactococcus lactis* phages to their polysaccharide receptors. *Molecular microbiology*, doi:10.1111/mmi.12978 (2015).
- 193 Quiberoni, A., Stiefel, J. I. & Reinheimer, J. A. Characterization of phage receptors in *Streptococcus thermophilus* using purified cell walls obtained by a simple protocol. *Journal of applied microbiology* **89**, 1059-1065 (2000).
- 194 Baptista, C., Santos, M. A. & Sao-Jose, C. Phage SPP1 reversible adsorption to *Bacillus subtilis* cell wall teichoic acids accelerates virus recognition of membrane receptor YueB. *Journal of bacteriology* **190**, 4989-4996, doi:10.1128/JB.00349-08 (2008).
- 195 Xia, G. *et al.* Wall teichoic Acid-dependent adsorption of staphylococcal siphovirus and myovirus. *Journal of bacteriology* **193**, 4006-4009, doi:10.1128/JB.01412-10 (2011).
- 196 Baker, T. A., Mizuuchi, M. & Mizuuchi, K. MuB protein allosterically activates strand transfer by the transposase of phage Mu. *Cell* **65**, 1003-1013 (1991).
- 197 Fu, T. J., Tse-Dinh, Y. C. & Seeman, N. C. Holliday junction crossover topology. *Journal of molecular biology* **236**, 91-105, doi:10.1006/jmbi.1994.1121 (1994).
- 198 Abremski, K., Wierzbicki, A., Frommer, B. & Hoess, R. H. Bacteriophage P1 Cre-loxP site-specific recombination. Site-specific DNA topoisomerase activity of the Cre recombination protein. *The Journal of biological chemistry* **261**, 391-396 (1986).
- 199 Nash, H. A. LambdaattB-attP, a lambda derivative containing both sites involved in integrative recombination. *Virology* **57**, 207-216 (1974).
- 200 Spinelli, S. *et al.* Modular structure of the receptor binding proteins of *Lactococcus lactis*

- phages. The RBP structure of the temperate phage TP901-1. *The Journal of biological chemistry* **281**, 14256-14262 (2006).
- 201 Ricagno, S. *et al.* Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bIL170. *Journal of virology* **80**, 9331-9335 (2006).
- 202 Deveau, H., Labrie, S. J., Chopin, M. C. & Moineau, S. Biodiversity and classification of lactococcal phages. *Appl Environ Microbiol* **72**, 4338-4346, doi:10.1128/AEM.02517-05 (2006).
- 203 Meyer, J. R. *et al.* Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**, 428-432, doi:10.1126/science.1214449 (2012).
- 204 Pepin, K. M., Domsic, J. & McKenna, R. Genomic evolution in a virus under specific selection for host recognition. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **8**, 825-834, doi:10.1016/j.meegid.2008.08.008 (2008).
- 205 Paterson, S. *et al.* Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275-278, doi:10.1038/nature08798 (2010).
- 206 Qimron, U., Marintcheva, B., Tabor, S. & Richardson, C. C. Genomewide screens for *Escherichia coli* genes affecting growth of T7 bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19039-19044, doi:10.1073/pnas.0609428103 (2006).
- 207 Ravin, V., Raisanen, L. & Alatosava, T. A conserved C-terminal region in Gp71 of the small isometric-head phage LL-H and ORF474 of the prolate-head phage JCL1032 is implicated in specificity of adsorption of phage to its host, *Lactobacillus delbrueckii*. *Journal of bacteriology* **184**, 2455-2459 (2002).
- 208 Biemann, R. *et al.* Receptor binding proteins of *Listeria monocytogenes* bacteriophages A118 and P35 recognize serovar-specific teichoic acids. *Virology* **477**, 110-118, doi:10.1016/j.virol.2014.12.035 (2015).
- 209 Tremblay, D. M. *et al.* Receptor-binding protein of *Lactococcus lactis* phages: identification and characterization of the saccharide receptor-binding site. *Journal of bacteriology* **188**, 2400-2410, doi:10.1128/JB.188.7.2400-2410.2006 (2006).
- 210 Koc, C. *et al.* Structural and enzymatic analysis of TarM glycosyltransferase from *Staphylococcus aureus* reveals an oligomeric protein specific for the glycosylation of wall teichoic acid. *The Journal of biological chemistry* **290**, 9874-9885, doi:10.1074/jbc.M114.619924 (2015).
- 211 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774-797, doi:10.1016/j.jmb.2007.05.022 (2007).
- 212 Shi, W. W. *et al.* Structure of a Novel O-Linked N-Acetyl-d-glucosamine (O-GlcNAc) Transferase, GtfA, Reveals Insights into the Glycosylation of Pneumococcal Serine-rich Repeat Adhesins. *The Journal of biological chemistry* **289**, 20898-20907, doi:10.1074/jbc.M114.581934 (2014).
- 213 Moynihan, P. J. & Clarke, A. J. O-Acetylated peptidoglycan: controlling the activity of bacterial autolysins and lytic enzymes of innate immune systems. *Int J Biochem Cell Biol* **43**, 1655-1659, doi:10.1016/j.biocel.2011.08.007 (2011).
- 214 Li, X. *et al.* An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*. *Sci Rep* **6**, 26455, doi:10.1038/srep26455 (2016).
- 215 Veesler, D. *et al.* Structure of the phage TP901-1 1.8 MDa baseplate suggests an

- alternative host adhesion mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8954-8958, doi:10.1073/pnas.1200966109 (2012).
- 216 Koc, C. *et al.* Structure of the host-recognition device of Staphylococcus aureus phage varphi11. *Sci Rep* **6**, 27581, doi:10.1038/srep27581 (2016).

## 6 APPENDIX

### 6.1 Gp45 crystal data and phasing statistics tables

DATA COLLECTION	crystal 1	SIRAS reference (crystal 2)	Native (crystal 3)	crystal 4
wavelength (Å)	1.25472	1.0000	1.06754	1.25511
absorption position	peak	-	-	inflection
space group	P1	P1	P1	P1
cell dimensions	87.65, 89.60,	86.96, 89.24,	87.06, 89.01,	87.62, 89.56,
a, b, c (Å)	93.73, 92.7,	93.50, 92.9,	93.26 93.0,	93.72, 92.7,
a, b, $\gamma$ (°)	105.7, 117.9	105.3, 117.6	105.2, 117.6	105.7, 117.9
resolution limits (Å) <sup>a</sup>	50-3.30 (3.37-3.30)	50-3.30 (3.38-3.30)	50-2.20 (2.25-2.20)	50-3.30 (3.39-3.30)
R-meas (%) <sup>b</sup>	12.5 (100.4)	8.9 (31.3)	7.1 (72.6)	9.8 (75.2)
CC <sup>1/2</sup>	99.9 (87.1)	99.6 (94.9)	99.9 (77.6)	99.8 (77.0)
no. unique reflections	71,266 (5,052)	34,939 (2,510)	116,568 (8,510)	68, 928 (4,733)
mean (I)/ $\sigma$ d(I)	20.4 (2.9)	16.7 (4.8)	14.9 (2.0)	13.6 (2.0)
completeness (%)	99.3 (95.0)	98.6 (95.7)	97.8 (96.9)	96.4 (89.5)
multiplicity	13.7 (12.0)	3.6 (3.4)	3.6 (3.4)	3.6 (3.3)
B-wilson (Å <sup>2</sup> )	74.5	35.6	44.9	71.8

Table 1: Data collection statistics 1 (used for structure solution)

DATA COLLECTION	crystal 5	crystal 6	crystal 7	crystal 8
wavelength (Å)	1.25472	1.25511	1.25473	1.25511
absorption position	peak	inflection	peak	inflection
space group	P1	P1	P1	P1
cell dimensions	87.67, 89.52,	87.65, 89.51,	87.57, 89.58,	87.56, 89.60,
a, b, c (Å)	93.51, 92.8,	93.50, 92.8,	93.29, 92.6,	93.04 92.7,
a, b, $\gamma$ (°)	105.6, 117.8	105.6, 117.8	105.4, 117.7	105.3, 117.7
resolution limits (Å) <sup>a</sup>	50-3.10 (3.17-3.10)	50-3.10 (3.38-3.10)	50-3.68 (3.77-3.68)	50-3.66 (3.74-3.66)
R-meas (%) <sup>b</sup>	12.4 (96.3)	10.3 (72.9)	15.1 (122.9)	11.1 (71.6)
CC <sup>1/2</sup>	99.8 (74.4)	99.8 (72.5)	99.7 (51.8)	99.8 (74.4)
no. unique reflections	83,616 (5,741)	82,956 (5,741)	49, 368 (2,970)	50,039 (3,175)
mean (I)/ $\sigma$ d(I)	13.8 (2.1)	11.8 (2.0)	11.1 (1.1)	12.2 (2.1)
completeness (%)	96.8 (89.5)	96.4 (90.3)	95.8 (77.3)	95.5 (82.8)
multiplicity	7.1 (6.3)	3.6 (3.3)	6.6 (2.8)	3.5 (3.2)
B-wilson (Å <sup>2</sup> )	69.1	67.5	93.6	84.0

Table 2: Data collection statistics 2 (Ta-LIII absorption edge)

Appendix

<b>DATA COLLECTION</b>	crystal 9	crystal 9	crystal 9	crystal 9	crystal 9
wavelength (Å)	1.24231	1.24233	1.25476	1.25516	1.26759
absorption position	High remote	High remote chi-40	peak	inflection	Low remote
space group	P1	P1	P1	P1	P1
cell dimensions a, b, c (Å)	87.66, 89.69, 94.18,	87.75, 89.77, 94.18,	87.83, 89.90, 94.29,	87.75, 89.81, 94.26,	87.88, 89.95, 94.31,
a, b, $\gamma$ (°)	92.5, 105.7, 118.0	92.5, 105.7, 118.1	92.5, 105.7, 118.1	92.5, 105.7, 118.1	92.5, 105.7, 118.2
resolution limits (Å) <sup>a</sup>	50-3.20 (3.27-3.20)	50-3.30 (3.38-3.30)	50-3.35 (3.43-3.35)	50-3.30 (3.38-3.30)	50-3.45 (3.53-3.45)
R-meas (%) <sup>b</sup>	7.9 (56.4)	8.8 (176.7)	9.3 (78.5)	8.9 (132.6)	7.7 (112.5)
CC <sup>1/2</sup>	99.7 (75.6)	99.7 (66.3)	99.7 (68.6)	99.7 (68.8)	99.8 (71.8)
Nr. unique reflections	68,621 (5,281)	59,705 (4,635)	57,630 (4,651)	61,192 (4,878)	51,574 (4,230)
mean (I)/ $\sigma$ d(I)	12.8 (2.0)	12.1 (1.2)	10.6 (1.7)	11.7 (1.3)	13.9 (1.3)
completeness (%)	86.9 (90.2)	82.8 (85.8)	83.4 (90.2)	84.7 (90.0)	81.5 (90.3)
multiplicity	1.8 (1.8)	1.8 (1.9)	1.8 (1.9)	1.8 (1.9)	1.8 (1.9)
B-wilson (Å <sup>2</sup> )	47.9	39.1	31.7	42.2	21.7

*Table 3: Data collection statistics 3 (Ta-LIII absorption edge)*

<b>DATA COLLECTION</b>	crystal 10	crystal 10	crystal 10	crystal 10
wavelength (Å)	1.10333	1.11325	1.11375	1.12333
absorption position	High remote	peak	inflection	Low remote
space group	P1	P1	P1	P1
cell dimensions a, b, c (Å)	88.27, 90.28, 94.87, 92.1,	88.38, 90.36, 94.93, 92.1,	88.47, 90.43, 94.96, 92.1,	88.53, 90.48, 94.99, 92.1,
a, b, $\gamma$ (°)	106.2, 118.1	106.2, 118.2	106.2, 118.2	106.1, 118.3
resolution limits (Å) <sup>a</sup>	50-3.50 (3.58-3.50)	50-3.60 (3.68-3.60)	50-3.60 (3.68-3.60)	50-3.50 (3.58-3.50)
R-meas (%) <sup>b</sup>	9.4 (93.9)	9.6 (55.6)	10.0 (59.4)	10.7 (170.3)
CC <sup>1/2</sup>	99.8 (68.1)	99.8 (83.5)	99.8 (83.8)	99.8 (43.7)
no. unique reflections	57,267 (5,281)	53,067 (3,482)	53,129 (3,483)	57,240 (3,826)
mean (I)/ $\sigma$ d(I)	8.8 (2.0)	8.7 (2.0)	8.6 (1.8)	8.2 (0.7)



<b>DATA COLLECTION</b>	crystal 10	crystal 10	crystal 10	crystal 10
completeness (%)	93.2 (84.8)	93.7 (82.3)	93.7 (81.9)	92.7 (83.4)
multiplicity	1.8 (1.9)	1.8 (1.9)	1.8 (1.9)	1.8 (1.7)
B-wilson ( $\text{\AA}^2$ )	91.1	92.6	95.1	101.5

Table 4: Data collection statistics 4 (Ta-LII absorption edge)

Site	Occupancy	B-factor ( $\text{\AA}^2$ )
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 1	0.21	162.2
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 2	0.20	90.1
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 3	0.36	33.6
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 4	0.18	143.2
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 5	0.29	187.3
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 6	0.12	190.3
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 7	0.39	152.1
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 8	0.20	138.2
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 9	0.10	185.7
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 10	0.05	64.1
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 11	0.28	109.1
Spherical Cluster Ta <sub>6</sub> Br <sub>12</sub> :Ta 12	0.20	114.3

Table 5: Refinement parameters of atomic substructure

BIN	ISOMORPHOUS						ANOMALOUS			
	D <sub>min</sub>	D <sub>max</sub>	N <sub>acen</sub>	N <sub>cen</sub>	LOC <sub>acen</sub>	LOC <sub>cen</sub>	N <sub>acen</sub>	N <sub>cen</sub>	LOC <sub>acen</sub>	LOC <sub>cen</sub>
1	48.79	14.16	431	0	1.031	0.000	0	0	0.000	0.000
2	14.16	10.23	733	0	1.002	0.000	0	0	0.000	0.000
3	10.23	8.41	931	0	1.002	0.000	0	0	0.000	0.000
4	8.41	7.31	1096	0	1.001	0.000	0	0	0.000	0.000
5	7.31	6.56	1244	0	0.996	0.000	0	0	0.000	0.000
6	6.56	5.99	1376	0	0.997	0.000	0	0	0.000	0.000
7	5.99	5.55	1492	0	0.996	0.000	0	0	0.000	0.000
8	5.55	5.20	1600	0	0.998	0.000	0	0	0.000	0.000
9	5.20	4.91	1705	0	0.996	0.000	0	0	0.000	0.000
10	4.91	4.66	1807	0	0.996	0.000	0	0	0.000	0.000
11	4.66	4.44	1895	0	0.997	0.000	0	0	0.000	0.000
12	4.44	4.25	1965	0	0.997	0.000	0	0	0.000	0.000
13	4.25	4.09	2052	0	0.997	0.000	0	0	0.000	0.000
14	4.09	3.94	2133	0	0.993	0.000	0	0	0.000	0.000
15	3.94	3.81	2224	0	0.992	0.000	0	0	0.000	0.000
16	3.81	3.69	2265	0	0.986	0.000	0	0	0.000	0.000

Appendix

17	3.69	3.58	2349	0	0.987	0.000	0	0	0.000	0.000
18	3.58	3.48	2438	0	0.977	0.000	0	0	0.000	0.000
19	3.48	3.39	2488	0	0.972	0.000	0	0	0.000	0.000
20	3.39	3.30	2519	0	0.959	0.000	0	0	0.000	0.000
Overall			34743		0.990					

*Table 6: Lack of closure for the reference native dataset determined by SIRAS protocol divided by total variance for crystal 2*

ISOMORPHOUS							ANOMALOUS			
BIN	D <sub>min</sub>	D <sub>max</sub>	N <sub>acen</sub>	N <sub>cen</sub>	LOC <sub>acen</sub> <sup>c</sup>	LOC <sub>cen</sub> <sup>c</sup>	N <sub>acen</sub>	N <sub>cen</sub>	LOC <sub>acen</sub> <sup>c</sup>	LOC <sub>cen</sub> <sup>c</sup>
1	48.79	14.16	436	0	2.966	0.000	431	0	1.605	0.000
2	14.16	10.23	735	0	1.247	0.000	735	0	1.101	0.000
3	10.23	8.41	935	0	1.091	0.000	927	0	1.039	0.000
4	8.41	7.31	1111	0	0.989	0.000	1102	0	1.020	0.000
5	7.31	6.56	1253	0	0.966	0.000	1248	0	1.102	0.000
6	6.56	5.99	1387	0	0.985	0.000	1382	0	1.011	0.000
7	5.99	5.55	1499	0	1.020	0.000	1489	0	0.994	0.000
8	5.55	5.20	1608	0	1.037	0.000	1604	0	0.981	0.000
9	5.20	4.91	1717	0	1.011	0.000	1712	0	0.999	0.000
10	4.91	4.66	1818	0	1.005	0.000	1813	0	1.074	0.000
11	4.66	4.44	1912	0	1.041	0.000	1906	0	1.100	0.000
12	4.44	4.25	1986	0	1.050	0.000	1984	0	1.200	0.000
13	4.25	4.09	2078	0	1.014	0.000	2076	0	1.182	0.000
14	4.09	3.94	2150	0	0.993	0.000	2148	0	1.119	0.000
15	3.94	3.81	2255	0	1.015	0.000	2252	0	1.085	0.000
16	3.81	3.69	2288	0	0.966	0.000	2287	0	1.018	0.000
17	3.69	3.58	2384	0	0.985	0.000	2382	0	0.975	0.000
18	3.58	3.48	2464	0	0.898	0.000	2462	0	0.881	0.000
19	3.48	3.39	2528	0	0.864	0.000	2528	0	0.849	0.000
20	3.39	3.30	2556	0	0.807	0.000	2553	0	0.811	0.000
Overall			35100	0	1.032	0.000	35021	0	1.033	0.000

*Table 7: Lack of closure (LOC) for the derivative crystal determined by SIRAS protocol for crystal 1*

ISOMORPHOUS							ANOMALOUS			
BIN	D <sub>min</sub>	D <sub>max</sub>	N <sub>acen</sub>	PP <sub>acen</sub>	RC <sub>acen</sub> <sup>d</sup>	PP/RC <sub>cen</sub>	N <sub>acen</sub>	PP <sub>acen</sub> <sup>f</sup>	RC <sub>acen</sub> <sup>e</sup>	PP/RC <sub>cen</sub>
1	48.79	14.16	430	1.183	0.608	0.000	431	5.116	0.292	0.000
2	14.16	10.23	733	1.949	0.523	0.000	735	6.759	0.232	0.000
3	10.23	8.41	925	1.723	0.566	0.000	927	5.747	0.269	0.000
4	8.41	7.31	1092	1.600	0.662	0.000	1102	4.909	0.306	0.000
5	7.31	6.56	1243	1.278	0.759	0.000	1248	3.177	0.412	0.000
6	6.56	5.99	1372	0.930	0.849	0.000	1382	2.999	0.432	0.000

Appendix

7	5.99	5.55	1488	0.656	0.892	0.000	1489	2.291	0.518	0.000
8	5.55	5.20	1597	0.411	0.932	0.000	1604	1.651	0.665	0.000
9	5.20	4.91	1702	0.249	0.969	0.000	1712	1.089	0.787	0.000
10	4.91	4.66	1803	0.155	0.991	0.000	1813	0.623	0.897	0.000
11	4.66	4.44	1893	0.089	0.978	0.000	1906	0.390	0.956	0.000
12	4.44	4.25	1963	0.044	0.980	0.000	1984	0.163	0.988	0.000
13	4.25	4.09	2049	0.013	0.969	0.000	2076	0.043	0.999	0.000
14	4.09	3.94	2133	0.022	0.971	0.000	2148	0.058	0.999	0.000
15	3.94	3.81	2223	0.045	0.959	0.000	2252	0.099	0.997	0.000
16	3.81	3.69	2264	0.068	0.954	0.000	2287	0.122	0.996	0.000
17	3.69	3.58	2348	0.082	0.936	0.000	2382	0.135	0.996	0.000
18	3.58	3.48	2436	0.103	0.940	0.000	2462	0.147	0.997	0.000
19	3.48	3.39	2488	0.114	0.929	0.000	2528	0.144	0.998	0.000
20	3.39	3.30	2517	0.126	0.945	0.000	2553	0.143	0.998	0.000
Overall			34699	0.693	0.834	0.000	35021	1.402	0.744	0.000

Figure 2 Phasing power (PP) and R-Cullis (RC) for the derivative crystal determined by SIRAS protocol for crystal I

a: values in parentheses correspond to the highest resolution shell.

b:  $R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n_{hkl}}{n_{hkl}-1}} \sum_i^{n_{hkl}} |I_i(hkl) - \langle I_i(hkl) \rangle|}{\sum_{hkl} \sum_i^{n_{hkl}} I_i(hkl)}$ , with  $\langle I_i(hkl) \rangle = \frac{\sum_i^{n_{hkl}} I_i(hkl)}{n_{hkl}}$  and  $I_i(hkl)$  the intensity of a unique reflection.

c:  $LOC = \sum_{hkl} | |F_P(\text{obs}) + F_H(\text{calc})| - F_{PH}(\text{obs}) |$

d:  $R_{Cullis\ iso} = \frac{\sum_{\phi} P(\phi) |LOC|}{\sum_{hkl} |F_{PH}(\text{obs}) - F_P(\text{obs})|}$

e:  $R_{Cullis\ ano} = \frac{\sum_{hkl} \sum_{\phi} P(\phi) |\Delta_{ano}(\text{obs}) - \Delta_{ano}(\text{calc})|}{\sum_{hkl} |\Delta_{ano}(\text{obs})|}$

f:  $PP = \frac{|F_H|}{\sigma(|\sum_{\phi} P(\phi) |LOC||)}$

## 6.2 Publications

# Structural and Enzymatic Analysis of TarM Glycosyltransferase from *Staphylococcus aureus* Reveals an Oligomeric Protein Specific for the Glycosylation of Wall Teichoic Acid\*

Received for publication, October 20, 2014, and in revised form, February 10, 2015. Published, JBC Papers in Press, February 19, 2015, DOI 10.1074/jbc.M114.619924

Cengiz Koç<sup>‡</sup>, David Gerlach<sup>§</sup>, Sebastian Beck<sup>§</sup>, Andreas Peschel<sup>§¶</sup>, Guoqing Xia<sup>§||1</sup>, and Thilo Stehle<sup>‡¶\*\*2</sup>

From the <sup>‡</sup>Interfaculty Institute of Biochemistry, University of Tübingen, 72076 Tübingen, Germany, <sup>§</sup>Interfaculty Institute of Microbiology and Infection Medicine, Cellular and Molecular Microbiology Section, University of Tübingen, 72076 Tübingen, Germany, <sup>¶</sup>German Center for Infection Research (DZIF), Partner site Tübingen, 72076 Tübingen, Germany, <sup>||</sup>Faculty of Medical and Human Sciences, Stopford Building, Institute of Inflammation and Repair, The University of Manchester, Oxford Road, Manchester, M13 9PT, United Kingdom, and <sup>\*\*</sup>Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

**Background:** TarM catalyzes the addition of  $\alpha$ -GlcNAc to 4'-polyribitol-phosphate of wall teichoic acid (WTA) in *S. aureus*.

**Results:** Structural analysis shows that TarM is a homotrimeric propeller-like glycosyltransferase.

**Conclusion:** Enzyme processivity is linked to a novel domain that generates the trimer.

**Significance:** Our structure-function analysis helps define the biosynthetic pathway leading to WTA glycosylation in *S. aureus*.

Anionic glycopolymers known as wall teichoic acids (WTAs) functionalize the peptidoglycan layers of many Gram-positive bacteria. WTAs play central roles in many fundamental aspects of bacterial physiology, and they are important determinants of pathogenesis and antibiotic resistance. A number of enzymes that glycosylate WTA in *Staphylococcus aureus* have recently been identified. Among these is the glycosyltransferase TarM, a component of the WTA *de novo* biosynthesis pathway. TarM performs the synthesis of  $\alpha$ -O-N-acetylglycosylated poly-5'-phosphoribitol in the WTA structure. We have solved the crystal structure of TarM at 2.4 Å resolution, and we have also determined a structure of the enzyme in complex with its substrate UDP-GlcNAc at 2.8 Å resolution. The protein assembles into a propeller-like homotrimer in which each blade contains a GT-B-type glycosyltransferase domain with a typical Rossmann fold. The enzymatic reaction retains the stereochemistry of the anomeric center of the transferred GlcNAc-moiety on the polyribitol backbone. TarM assembles into a trimer using a novel trimerization domain, here termed the HUB domain. Structure-guided mutagenesis experiments of TarM identify residues critical for enzyme activity, assign a putative role for the HUB in TarM function, and allow us to propose a likely reaction mechanism.

*Staphylococcus aureus* is a leading cause of nosocomial pneumonia, surgical site infections, and blood stream infections.

\* This work was supported by Collaborative Research Grant SFB766 (to A. P., G. X., and T. S.) and TRR34 (to T. S. and A. P.) from the German Research Foundation and by the German Center for Infection Research (DZIF; to G. X. and A. P.).

The atomic coordinates and structure factors (codes 4WAC and 4WAD) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

<sup>1</sup> To whom correspondence may be addressed: Faculty of Medical and Human Sciences, Stopford Bld., Institute of Inflammation and Repair, The University of Manchester, Oxford Rd., Manchester, M13 9PT, UK. Tel.: 44-161-27-51903; E-mail: guoqing.xia@manchester.ac.uk.

<sup>2</sup> To whom correspondence may be addressed: Interfaculty Institute of Biochemistry, University of Tübingen, Hoppe-Seyler-Strasse 4, 72076 Tübingen, Germany. Tel.: 49-7071-29-73043; Fax: 49-7071-29-5565; E-mail: thilo.stehle@uni-tuebingen.de.

The bacterium remains a severe threat to human health, in part due to the continued emergence of strains that are resistant to existing antibiotics (1). To survive, *S. aureus* relies heavily on virulence and adaptability to its environment. The *S. aureus* cell envelope structure is highly complex, and this complexity is central to the survival and adaptability of the organism. Major components of the cell envelope are glycosylated structures (2, 3), including glycoproteins, polysaccharide intracellular adhesin, capsular polysaccharides, peptidoglycan, lipoteichoic acid, and wall teichoic acid (WTA).<sup>3</sup> The unique ability of methicillin-resistant *S. aureus* (MRSA) to develop resistance to  $\beta$ -lactams as well as other antibiotics (4) is in part due to the structure and composition of specific cell wall components (5–7). The role of WTAs in these processes is complex and not well understood at the molecular level. WTAs serve to protect the cell from degradation through lysozyme (8) or from the action of cationic antimicrobial oligopeptides (9). However, WTAs also assist in staphylococcal adhesion and colonization (10, 11). Furthermore, they play a critical role in cell division and biofilm formation (12). The chemical structure of WTA varies substantially among Gram-positive bacteria (13), and this variability represents one strategy that allows these organisms to adapt to the environment or react to host defense systems (3).

Most of the *S. aureus* strains produce poly-ribitol-phosphate (RboP)-type WTA, which is composed of ~40 RboP units that are connected by 1,5-phosphodiester bonds. Some of the C4 hydroxyl groups of the WTA RboP unit are either substituted with  $\alpha$ -O- or  $\beta$ -O-GlcNAc, whereas the C2 hydroxyls sometimes carry a D-alanine (Fig. 1).

Biosynthesis of WTA in *S. aureus* is carried out by a cluster of enzymes belonging to the teichoic acid ribitol (Tar) synthesizing pathway, many of which have only been recently character-

<sup>3</sup> The abbreviations used are: WTA, wall teichoic acid; Tar, teichoic acid ribitol; RboP, ribitol-phosphate; dlt, D-alanyl-lipoteichoic acid; DUF, domain of unknown function; EOP, efficiency of plaquing; DLS, dynamic light scattering; Gt-N, glycosyltransferase N-terminal lobe; Gt-C, glycosyltransferase C-terminal lobe; r.m.s.d., root mean square deviation; MRSA, methicillin-resistant *S. aureus*; GT, glycosyltransferase.

ized. The polyribitol backbone is covalently attached to the *N*-acetylmuramic acid moiety of the peptidoglycan via a disaccharide (ManNAc $\rightarrow\beta$ (1,4)-GlcNAc-1-P) linkage unit followed by two units of glycerol phosphates (14) as shown in Fig. 1. In concert with TarA, TarB, TarI, TarJ, and TarL, the main chain is synthesized on the lipid carrier undecaprenyl monophosphate (C55P), which is embedded in the inner leaflet of the cell membrane. After the completion of glycosylation, the main chain of WTA is flipped to the outer leaflet of the plasma membrane via the ABC-type transporter TarG/TarH (13, 15).

The regulated addition of alanines at the ribitol 2'-position by the *D*-alanyltransferase, one gene product of the *dltABCD* (16) gene cluster, as a final modification counterbalances the predominant negative charge of the linking phosphate groups and results in WTA becoming zwitterionic. The evolution of host-pathogen interaction is thought to have led to the increase of positive charges in the bacterial cell wall to circumvent the action of cationic antimicrobial peptides (17, 18).

The enzymes TarM and TarS decorate the WTA backbone with  $\alpha$ -GlcNAc and  $\beta$ -GlcNAc, respectively (6, 19). The  $\beta$ -GlcNAcylation of RboP is critical for the resistance of *S. aureus* (MRSA) to  $\beta$ -lactams (6). Furthermore  $\beta$ -GlcNAc residues on WTA are recognized by the mannose-binding lectin, leading to complement activation pathway of the human innate immune system as well as by antibodies in the adaptive immune system (20). The role of the  $\alpha$ -GlcNAcylation is not yet known.

To define the mechanism of RboP glycosylation, we have performed a structure-function analysis of the glycosyltransferase TarM, a 171-kDa protein. Sequence analysis and database research predicted one domain of TarM to belong to the GT-B superfamily of glycosyltransferases (21), whereas the second domain was assigned DUF1975 (domain of unknown function) according to the Pfam database (22). The crystal structure of TarM reveals a propeller-like trimer, with the three GT-B

domains arranged as blades around a central hub formed by the three DUF1975 domains. Accordingly, we suggest the name HUB for DUF1975. The structure analysis of TarM bound to its substrate UDP-GlcNAc identifies the active site, defines essential contacts with this ligand, and suggests a plausible reaction mechanism. As TarM is the first known enzyme structure in the biogenesis pathway of poly-RboP WTA, our work sheds light on an essential aspect of *S. aureus* glycosylation and provides an initial framework for investigating parameters that dictate glycosylation of WTAs in bacteria.

## EXPERIMENTAL PROCEDURES

**Strains and Media**—*S. aureus* strains were cultured in BM media (1% (w/v) Tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 0.1% (w/v) glucose, 0.1% (w/v) K<sub>2</sub>HPO<sub>4</sub>, pH 7.2). *Escherichia coli* strains were cultivated in LB media (1% (w/v) casein hydrolysate peptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 1% (w/v) glucose, 1‰ (w/v) K<sub>2</sub>HPO<sub>4</sub>, pH 7.2).

**Cloning and Expression of tarM and Mutant tarM**—Wild-type *tarM* (*SACOL 1043*) was subcloned as reported previously (19). QuikChange (Stratagene) was used to introduce point mutations into the glycosyltransferase active site in either pRB474-*tarM* or pBAD-TOPO-102/202-*tarM* (EcoRI/BamHI, Amp or Kan) as template. pRB474 shuttle vectors containing *tarM* variants (wt or mutant) were transformed into *S. aureus* RN4220 mutant  $\Delta tarM\Delta tarS$  for determining the efficiency of plaquing (EOP). Thus *tarM* and *tarM* mutants were fused to a hexahistidine tag at the N terminus and subcloned into the pBAD vector for recombinant expression in *E. coli* strain Top10 (Table 1).

**Expression**—Single colonies of *E. coli* transformants containing *tarM* variants were grown on antibiotics containing LB agar (1.5% (w/v) agar-agar in LB-medium). They were inoculated into 2 ml of LB medium and grown overnight at 37 °C. For large scale protein production, bacterial culture was induced at the log phase ( $A_{600} \approx 0.5-1.0$ ) with *L*-arabinose at a final concentration of 0.001% (w/v) at 20 °C for 12–20 h before harvesting by centrifugation at 7,900  $\times g$  for 13 min. After washing once with buffer A (10 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA), the cells were resuspended with buffer B (100 mM triethanolamine, pH 8.5, 500 mM LiCl, 5 mM EDTA, 1 mM DTT) for storage at –80 °C or for purification as described below.

**Purification**—Cells were lysed by ultrasonication (Digital Sonifer, Branson). After centrifugation at 38,000  $\times g$  for 55 min, the supernatant containing recombinant TarM was collected and used as the crude TarM preparation after dialyzing against

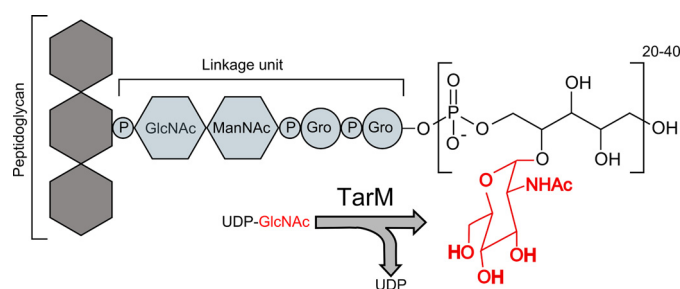


FIGURE 1. Schematic representation of the reaction catalyzed by TarM in WTA biogenesis. Gro, glycerol; ManNAc, *N*-acetyl-*D*-mannosamine.

TABLE 1  
Design of TarM mutations

Mutation	Location	Rationale
H249A	Gt-N	Removes contact with M18, possible indirect impact on catalytic activity
R326S	Gt-C	Removes contact with UDP-GlcNAc phosphate
K331S	Gt-C	Removes contact with UDP-GlcNAc phosphate, removes possible contact with Glu-403
E403A	Gt-C	Removes carboxylic function, removes possible catalytic activity
E411A	Gt-C	Removes interaction with UDP ribose
K136S	HUB	Disrupts intra-HUB contacts, possible other function
N138Q	HUB	Disrupts intra-HUB contacts
N180W	HUB	Disrupts intra-HUB contacts
V159Y/C164R	HUB	Disrupts intra-HUB contacts, possible other function
K136S/V159Y/C164R	HUB	Disrupts intra-HUB contacts, possible other function

## Structure of TarM

buffer C (50 mM triethanolamine, pH 8.5, 500 mM LiCl, 25 mM imidazole, 1 mM DTT). The crude preparation was applied onto a His-Trap-FF nickel-chelate affinity column (GE Healthcare, 5 ml), and the column was subsequently washed with 30 column volumes of buffer C followed by another washing step with 10 column volumes of buffer C containing 10 mM imidazole. Pure TarM was eluted using a gradient ranging from 36 to 400 mM imidazole in buffer C. The pure sample was then concentrated to 1 mg/ml (Sartorius vivaspin20 PES, 50,000 molecular weight cut-off), dialyzed against buffer D (50 mM triethanolamine, pH 8.5, 250 mM LiCl, 5 mM EDTA, 1 mM DTT), and subjected to 2 successive treatments with enterokinase (Ekmx, 20 °C, 15–20 h, 1.5 units/mg; Life Technologies). Aggregated proteins and excess enterokinase were removed by gel filtration (Superdex200). A final concentration step yielded highly monodisperse pure protein that was then used for crystallization and biophysical characterization. Purity and homogeneity of TarM were assessed by SDS-PAGE as well as dynamic light scattering (DLS).

**Structure Determination**—Initial small crystals grew as trap-ezoid-shaped plates (diameter  $\approx$  50  $\mu$ m) in 600-nl drops containing 300 nl of TarM protein solution in buffer D and 300 nl of crystallizing buffer E (100 mM imidazole, pH 8.0, 200 mM Ca(OAc)<sub>2</sub>, 20%(w/v) PEG-1000) in 96-well plates using the sitting-drop vapor diffusion method. Refinement yielded the final crystallization condition (100 mM imidazole, pH 7.7–8.2, 18–21% (w/v) PEG-1000, 0.1–0.25 M Ca(OAc)<sub>2</sub>) which produced large, single crystals (diameter  $\approx$  250  $\mu$ m). After soaking crystals in cryoprotection solution (buffer D:buffer E = 1:1, 10% (v/v) (4s)-2-methyl-2,4-pentanediol) they were directly frozen in liquid nitrogen. Data were collected on a PILATUS 2 M hybrid pixel detector using synchrotron beam line X06DA at the Swiss Light Source (SLS) super-bending magnet (2.9 tesla), and they were processed with the XDS package (23). The crystals belong to spacegroup P6<sub>3</sub>22 with cell parameters of  $a = b = 123.7$  Å and  $c = 223.3$  Å (Table 2). They contain one TarM monomer in their asymmetric unit, with a solvent content of 74%. For the determination of phases, crystals were soaked in crystallization condition containing the anomalous scatterer iodide (buffer D:buffer E = 1:1, 400 mM KI). After soaking them for 30 min, the crystals were back-soaked in cryoprotection solution and frozen in liquid nitrogen. The autoSHARP routine protocol (24) for SIRAS (single isomorphous replacement of anomalous scatterer) was used in conjunction with heavy atom detection implemented in SHELXDE (25) for initial phase calculation, and autoSHARP-implemented density modification package (DM) (26) was employed to perform solvent flattening for phase improvement. ARP/wARP (27) was then used to trace the first 300 residues of the protein. Alternating cycles of COOT (28) model building and REFMAC5 (29, 30) or PHENIX (31) refinement subsequently revealed additional residues, which were included in the refinement until convergence had been achieved. The final model includes residues 1–493 (PDB ID 4WAC). Data collection and refinement statistics are given in Table 2.

To solve the structure of the ligand-bound complex, purified TarM was preincubated with UDP-GlcNAc (15 mM, buffer D, 1 h, 4 °C) and then subjected to high throughput crystallization screening using a robot. Diffracting crystals were obtained with

crystallization solution buffer F (100 mM Tris-HCl, pH 7.0, 200 mM MgCl<sub>2</sub>, 10% (w/v) PEG-8000). Although slightly smaller than the crystals obtained with unliganded TarM, the complex crystals shared a similar morphology. The crystals were soaked in the new crystallization condition containing a higher amount of cocrystallant (buffer D:buffer F = 1:1, 50 mM UDP-GlcNAc) and subsequently transferred to cryoprotectant-containing solution (buffer D: buffer F = 1:1, 50 mM UDP-GlcNAc, 10%(v/v) (4s)-2-methyl-2,4-pentanediol) before freezing them in liquid nitrogen. Data were collected on a PILATUS 2 M detector at beamline X06DA of the Swiss Light Source. Data processing using XDS (23) yielded the same space group as the crystals of unbound TarM, with slightly altered cell parameters of  $a = b = 122$  Å and  $c = 212$  Å. The unbiased whole native structure solution was used as a molecular replacement (30, 32) input model for phasing the new data. After one refinement run of the phased structure model, the UDP- $\alpha$ -GlcNAc-moiety was clearly visible in the unbiased electron density maps, and thus the ligand was incorporated into the model using the refrac library (33) in COOT (28). TLS refinement utilizing REFMAC5 and PHENIX yielded the final model for the binary complex (PDB ID 4WAD). Data collection and refinement statistics are given in Table 2. Figures were generated with PyMOL (34).

**Dynamic Light Scattering**—DLS measurements were performed on a Nano Zetasizer (Malvern) with purified TarM samples at 1 mg/ml in buffer D or buffer G (10 mM Na<sub>2</sub>HPO<sub>4</sub>, 0.01 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.5, 200 mM NaF). Data were recorded and evaluated using Zeta Software (Malvern).

**Circular Dichroism**—CD measurements were performed on a JASCO J-720 spectropolarimeter with purified TarM samples at  $\sim$ 0.5 mg/ml in buffer G. A path length of 0.1 cm was used, and the samples were scanned at a speed of 50 nm/min. Data were recorded and evaluated using the software Spectra Manager (Jasco).

**Plaquing Efficiency of  $\phi$ 11**—To analyze the *in vivo* activity of TarM and its variants, the plaquing efficiency of bacteriophage  $\phi$ 11 was determined by plating  $\phi$ 11 on *S. aureus* mutant strain RN4220 $\Delta$ tarM $\Delta$ tarS complemented with empty plasmid (pRB474), a plasmid encoding wt TarM (pRB474-tarM), or plasmids encoding TarM variants (see Table 1). To determine the plaquing efficiency, 100  $\mu$ l of  $\phi$ 11 lysate with  $\sim$ 1000 plaque-forming units (pfu) was mixed well with 100  $\mu$ l of bacteria culture containing  $\sim$ 4  $\times$  10<sup>7</sup> colony forming units. After incubation at 25 °C for 10 min, the infection mixture was mixed well with 5 ml of soft agar and then poured onto BM plates containing 10  $\mu$ g/ml chloramphenicol. The plates stood at 37 °C overnight (16–24 h) and following up, the pfu was enumerated. The plaquing efficiency of  $\phi$ 11 on tarM-complemented RN4220- $\Delta$ tarM $\Delta$ tarS was set to 100%.

**WTA Glycosyltransferase Activity Assay**—The colorimetric assay was prepared according to Mulder's procedure (35) with slight modifications. 1.5  $\mu$ g of recombinant TarM variants (in 20 mM Tris, pH 8.0, 10 mM MgCl<sub>2</sub>) were incubated with UDP-GlcNAc (2 mM) and non-glycosylated WTA (25  $\mu$ M) that was isolated from RN4220  $\Delta$ tarM $\Delta$ tarS (36) and a reaction mixture consisting of phosphoenolpyruvate and NADH (0.2 mM each). The release of UDP by TarM was

**TABLE 2**  
Data and refinement statistics

Data set	native	complex	anomalous
Space Group	P6 <sub>3</sub> 22	P6 <sub>3</sub> 22	P6 <sub>3</sub> 22
a, b, c (Å)	124.75, 124.75, 223.25	124.02, 124.02, 217.03	124.77, 124.77, 225.49
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution (Å)	48.62-2.40 (2.54-2.40) <sup>a</sup>	48.13-2.80 (2.98-2.80)	50.0-3.3 (3.38-3.3)
unique reflections	40,716 (6,350)	24,904 (3,899)	29,958 (2,158)
Completeness (%)	99.6 (98.4)	99.4 (98.6)	99.8 (97.7)
Redundancy	9.79 (9.53)	12.89 (13.08)	42.60 (39.84)
I/σI	14.33 (2.16)	12.24 (1.56)	19.20 (2.48)
R <sub>meas</sub> (%) <sup>b</sup>	12.8 (147.2)	24.1 (184.4)	30.0 (221.8)
CC(1/2) (%)	99.6 (62.0)	99.5 (53.5)	99.8 (66.7)
Wilson B (Å <sup>2</sup> )	54.5	57.7	64.2
Anomalous phasing power	-	-	0.207 (0.031)
Figure of merit			
acentric/centric reflections	-	-	0.141/0.176 (0.019/0.025)
After solvent flattening <sup>c</sup>	-	-	0.85
Refinement statistics			
Resolution (Å)	48.62-2.40 (2.49-2.40)	48.13-2.80 (2.96-2.80)	-
R <sub>work</sub> (%) <sup>d</sup>	21.70	25.20	-
R <sub>free</sub> (%) <sup>e</sup>	25.14	29.09	-
Number of atoms/asu	4194	4168	-
Protein	4079	4074	-
Solvent/ligands	115	94	-
rmsd <sup>f</sup>			
Bonds (Å)	0.0119	0.004	-
Angles (°)	1.326	0.799	-
Average B-factor (Å <sup>2</sup> )	67.56	70.58	-
Protein	67.98	70.85	-
Solvent/ligands	47.58	45.88	-
Ramachandran <sup>g</sup>			
Favoured (%)	95.2	89.7	-
Outlier (%)	0.2	1.6	-

a: values in parentheses correspond to the highest resolution shell.

$$b: R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n_{hkl}}{n_{hkl} - 1}} \sum_i^{n_{hkl}} |I_i(hkl) - \langle I_i(hkl) \rangle|}{\sum_{hkl} \sum_i^{n_{hkl}} I_i(hkl)}, \text{ with } \langle I_i(hkl) \rangle = \frac{\sum_i^{n_{hkl}} I_i(hkl)}{n_{hkl}} \text{ and } I_i(hkl) \text{ the intensity of a unique reflection.}$$

reflection.

c: mean value as determined by Solomon (ccp4).

$$d: R_{work} = \frac{\sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}}, \text{ with } F_{hkl}^{obs} \text{ and } F_{hkl}^{calc} \text{ being the structure-factor amplitudes for unique reflections and model-derived calculations, respectively.}$$

e: R<sub>free</sub> represents 5% of excluded test data, calculated according to the equation for R<sub>work</sub>.

f: root mean square deviation

g: determined by Molprobit

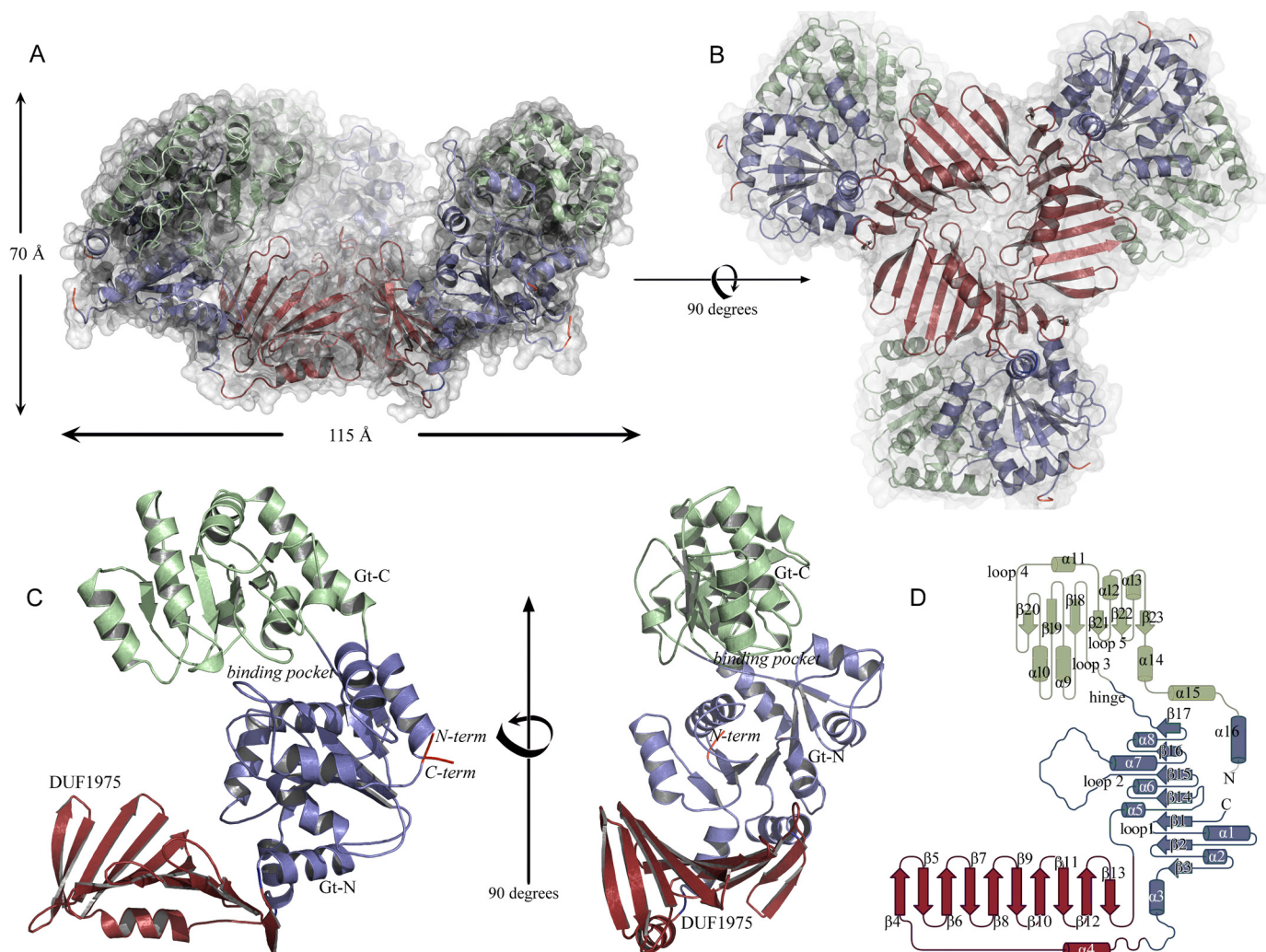
assayed through the coupled conversion of NADH to NAD<sup>+</sup> (340 nm, 40 min, 25 °C) by pyruvate kinase and lactate dehydrogenase (2 units each), leading to the decrease of absorbance.

## RESULTS

Overall Structure and Domain Organization of TarM—TarM assembles into a symmetric, propeller-like homotrimer, with three blades projecting from the central hub (Fig. 2, A and



## Structure of TarM



**FIGURE 2. Overall structure and domain organization of TarM.** *A* and *B*, ribbon representation of the TarM homotrimer viewed from two orthogonal angles. The molecular surface of the trimer is shown as a semitransparent surface. *C*, ribbon drawing of a TarM monomer, viewed from two orthogonal angles. *D*, topology drawing using topdraw of one TarM monomer. The Gt-N, Gt-C, and HUB domains are colored in blue, green, and red, respectively, in all panels.

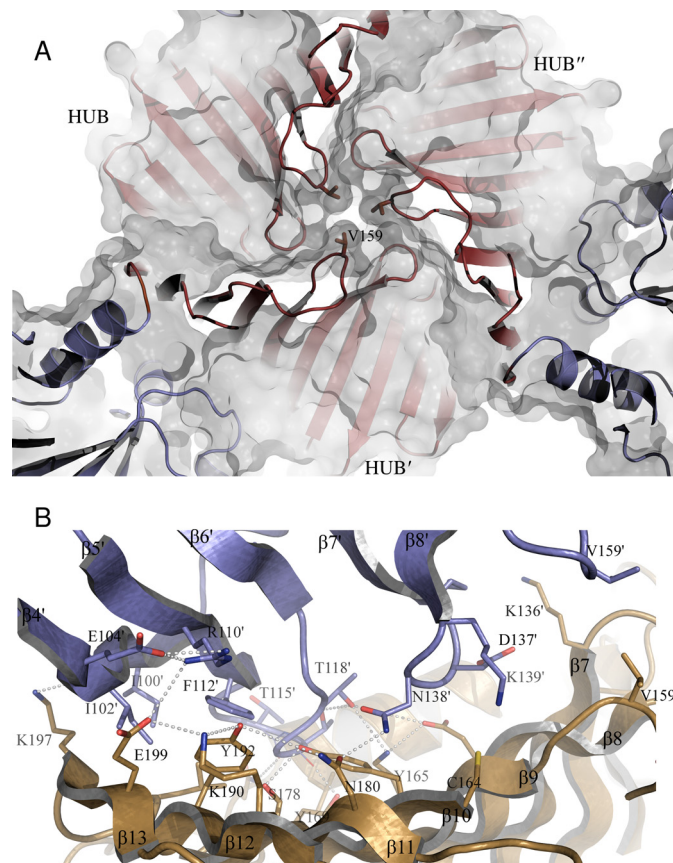
*B*). The three blades project at angles of  $\sim 120$  degrees from the hub (Fig. 2*A*), giving the propeller a cradle-like appearance, with a large cavity at its center. Each TarM monomer can be divided into two regions (Fig. 2, *C* and *D*); the glycosyltransferase (GT) domain forms the blade, which can be further subdivided into an N-terminal domain (Gt-N; residues 1–80, 202–309) and a C-terminal domain (Gt-C; residues 310–493). The trimer is assembled by three copies of a domain (residues 81–201) that was originally annotated as a domain of unknown function (DUF1975) and that is inserted into Gt-N. This domain features a 10-stranded antiparallel  $\beta$ -sheet composed of strands  $\beta 4$  through  $\beta 13$ , with one face of the sheet covered by a single  $\alpha$ -helix ( $\alpha 4$ ). Given its function in TarM trimerization, we refer to this domain as the HUB domain. The Gt-N and HUB domains are well ordered and exhibit low overall temperature factors (B-factors). In contrast, large portions of Gt-C display higher mobility and elevated B-factors, probably as a result of the larger surface-exposed area of this domain, the paucity of its interactions in the crystal lattice, and its flexible linkage to Gt-N. The Gt-N domain is positioned atop the HUB, whereas Gt-C projects away from this assembly at an angle of  $\sim 40$

degrees. Each TarM monomer, therefore, has a bent, hook-like conformation, giving rise to the cradle-like structure of the trimer (Fig. 2).

Together, subdomains Gt-N and Gt-C form the glycosyltransferase unit, referred to as GT-B. Gt-N consists of a parallel seven-stranded  $\beta$ -sheet (strands  $\beta 1$ – $\beta 3$  and  $\beta 14$ – $\beta 17$ ) connected by eight  $\alpha$ -helices (helices  $\alpha 1$ – $\alpha 3$ ,  $\alpha 5$ – $\alpha 8$  and  $\alpha 16$ ), whereas Gt-C contains a central six-stranded parallel  $\beta$ -sheet (strands  $\beta 18$ – $\beta 23$ ) and seven flanking  $\alpha$ -helices (helices  $\alpha 9$ – $\alpha 15$ ) (Fig. 2*D*). The Gt-N and Gt-C domains are linked by  $\sim 10$  solvent-exposed residues that connect strands  $\beta 17$  and  $\beta 18$ . A DALI (37) query of the Gt-N/Gt-C-unit returned several hits for structural homologs in the GT-B glycosyltransferase superfamily. The glycosyltransferases MshA from *Corynebacterium glutamicum* (Z-score 33.1, r.m.s.d. 2.9 Å, 334 aligned residues, 18% sequence identity, PDB ID 3C4Q (38)) and BshA from *Bacillus anthracis* (Z-score 32.9, r.m.s.d. 3.2 Å, 336 aligned residues, sequence identity 18%, PDB 3MBO (39)) yielded the highest scores, directly followed by the streptococcal enzyme Gtfa (Z-score 32.8, r.m.s.d. 2.8 Å, 361 aligned residues, sequence identity 23%, PDB 4PQG (40)). Other homolo-

**TABLE 3**  
Alignment of conserved GT4 residues

Structure	Conserved residues appear in columns with the numbering according to the deposited PDB files: 4WAD, TarM; 4PQG, GtfA; 3MBO, BshA; 3C4Q, MshA; 3OKA, PimB'; 3OY7, B736L.																			
	Residue code																			
TarM	Gly-16	Gly-17	Met-18	His-249	Val-250	Val-324	Ile-324	Ser-325	Arg-326	Pro-329	Lys-331	Gly-356	Glu-401	Gly-403	Gly-404	Gly-406	Leu-407	Glu-411	Ala-412	
GtfA	Ser-14	Ser-15	His-242	Ala-243	Asp-309	Ala-326	Ser-327	Arg-328	Tyr-397	Thr-402	Glu-404	Gly-405	Gly-407	Leu-408	Glu-412	Ala-413				
BshA	Gly-14	Gly-15	Ser-16	His-120	Asp-176	Ile-204	Ser-205	Asn-206	Lys-211	Gly-235	Lys-235	Gly-282	Ser-283	Gly-285	Glu-290	Ala-291				
MshA	Gly-22	Gly-23	Met-24	His-133	Asp-198	Thr-134	Asp-231	Pro-234	Lys-236	Gly-264	Pro-296	Val-309	Glu-316	Ser-317	Gly-319	Glu-324	Ala-325			
PimB'	Gly-19	Gly-20	His-118	His-118	Asp-173	Asp-173	Ser-205	Arg-206	Pro-209	Lys-211	Gly-236	Phe-276	Arg-281	Glu-290	Gly-291	Glu-294	Glu-298	Ala-299		
B736L	Gly-15		Asp-115	Val-117	Asp-165	Asn-191	Arg-192	Ala-195	Lys-197	Val-267	Ile-277	Ser-282	Glu-284	Gly-285	Gly-287	Leu-288	Glu-292	Gly-293		



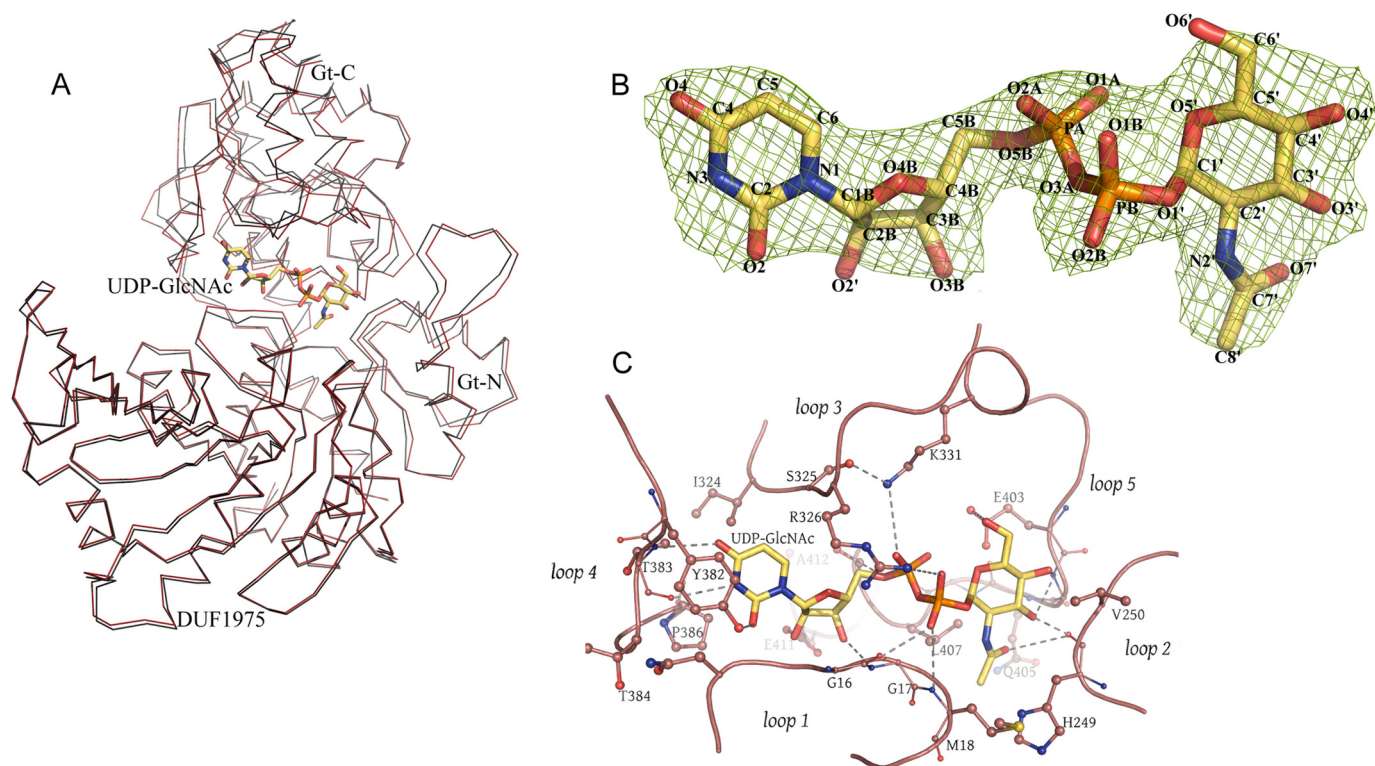
**FIGURE 3. Close-up view of the HUB domain-generated trimer interface.** *A*, view along the trimer axis showing three crystallographically related valines (Val-159) closing the HUB on the side opposite to the HUB maw. The color code is the same as in Fig. 2. *B*, close-up view showing the interface between two monomers. Three crystallographically related lysines (Lys-136) point with their side chains into the cavity. Other interfacial residues of relevance, determined by PISA, are represented as *sticks*. The TarM trimer interface is composed of 20 residues (amino acids 98, 100–104, 110, 112–119, 136–139, and 159) in one monomer (*orange*) and 28 residues (amino acids 85, 87, 88, 91, 136, 141, 142, 144, 154, 156–159, 164, 165, 167, 169, 174, 175, 178, 180, 190, 192–194, and 197–199) in the other monomer (*blue*). Residues contributing to the interface are concentrated in loops  $\beta 4'$ – $\beta 8'$  and  $\beta 9$ – $\beta 13$ .

gous proteins include PimB' (Z-score of 30.6, r.m.s.d. of 3.0 Å, 331 aligned residues, 14% sequence identity, PDB 3OKA (41)) and chlorovirus NY-2A gene product B736L (Z-score of 28.7, r.m.s.d. of 3.4 Å, 333 aligned residues, 14% sequence identity, PDB 3OY7 (42)). Table 3 lists the conservation of the binding pocket.

*The Oligomeric State of TarM*—The asymmetric unit of the crystals contains one TarM monomer (58 kDa, 493 residues) that assembles into the trimeric structure shown in Fig. 2 through a crystallographic three-fold symmetry operator. Trimer contacts exclusively involve the HUB domains, which form a funnel-like arrangement that is  $\sim 30$  Å wide at one end. The other end of the funnel is almost closed as a result of three closely approaching Val-159 side chains (Fig. 3A). The trimer interface includes hydrogen bonds and salt bridges as well as hydrophobic interactions (Fig. 3B), and it buries a total surface area of 777 Å<sup>2</sup> with a solvation free energy gain of  $-9$  kcal/M at each monomer-monomer contact.

It is of course possible that the observed propeller-like trimer is a crystallization artifact. However, the PISA server (43),

## Structure of TarM



**FIGURE 4. The ligand-binding site of one TarM chain.** *A*, superposition of free and UDP-GlcNAc-bound structures in ribbon representation. The UDP-GlcNAc-bound TarM structure is shown in red, and unbound TarM is depicted in black. *B*,  $2F_o - F_c$  map at 1.0  $\sigma$  contour-level of UDP-GlcNAc in the active site of TarM, seen from the face-on side. *C*, representation of the bound sugar surrounded by five prominent and well conserved loops (loop-1 (residues 9–18), loop-2 (residues 248–251), loop-3 (residues 324–332), loop-4 (residues 380–386), and loop-5 (residues 401–406)). Dashed lines indicate putative contacts with distances ranging from 2.2 to 4.5 Å. A distance cutoff of 4.5 Å was used to show hydrogen bonds and salt bridges.

which evaluates the physiologic relevance of crystallographic interfaces, classifies this interface as significant for complexation ( $css = 1$ ), in contrast to all other contacts of TarM subunits in the crystals. To examine whether the trimer also exists in solution, we performed size-exclusion chromatography and dynamic light scattering with purified TarM. Both experiments provide evidence for a trimeric state of the enzyme in solution. TarM elutes as a single peak in gel filtration, with a hydrodynamic diameter of  $\sim 14$  nm corresponding to a molecular mass of  $\sim 300$  kDa. Although this value is higher than that calculated for the trimer (174 kDa), the protein deviates significantly from a globular shape (Fig. 2A) and would, therefore, be expected to elute at a higher apparent molecular weight. A calculated Perrin-factor of  $\sim 1.5$  suggests a molecular shape deviant from a spherical protein (44). In accordance with this, the molecular shape derived from the structural data is reminiscent of an oblate rather than a sphere.

**Architecture of the Active Site**—Glycosyltransferases of the GT-B class typically bind their substrates at the interface between Gt-N and Gt-C. To characterize the ligand binding site of TarM, we solved the structure of the enzyme bound to its substrate UDP-GlcNAc through incubation of soluble TarM with UDP-GlcNAc and subsequent cocrystallization of the complex. The overall structures of unbound and UDP-GlcNAc-bound TarM are highly similar (r.m.s.d. value of 0.81 Å for 493 aligned residues, Fig. 4A), and thus binding of UDP-GlcNAc does not lead to any larger structural rearrangements. Unambiguous electron density in the active site cleft

allowed us to build the UDP-GlcNAc substrate and assign contacts (Fig. 4B). UDP-GlcNAc is located in a cleft formed by five loops (loops 1–5, Fig. 4C).

The GlcNAc moiety rests in a shallow pocket formed by residues in loops 1, 2, and 5, with one face of the sugar ring buried and the other exposed to solvent. The *N*-acetyl group faces into a small, hydrophobic pocket formed by Met-18 (loop 1), His-249 (loop 2), and Leu-407 (loop 5). The carbonyl oxygen as well as the C3 and C4 hydroxyl groups form hydrogen bonds to loop residues, whereas the anomeric center carbon C1 lies in close proximity to the carboxylate function of Glu-403. The negative charge of the pyrophosphate unit is negated by salt bridges to the Arg-326 and Lys-331 side chains. One of the phosphates also forms contacts with the backbone amide of Gly-17, whereas the other is hydrogen-bonded to Ser-408. Finally, the uridyl unit lies in a narrow pocket that is lined by residues from loops 1 and 3 and closed at the rear end by loop 4. The ribose faces toward the Glu-411 side chain. Specificity for uracil is generated through several polar and hydrophobic interactions with loop 4 residues Tyr-382, Thr-383, and Pro-386 as well as the side chain of Ile-324 within loop 3.

**EOP**—EOP was utilized to assess *in vivo* functionality of several TarM variants. Enzymatic activity of wild-type and mutant TarM was assayed using an established semiquantitative method that is based on a link between glycosylated WTA and bacteriophage adsorption to *S. aureus* hosts (45, 46). Recently we showed sugar residues on WTAs served as the receptor of siphophage such as  $\phi 11$ . The laboratory strain RN4220 lacks all

resistance mechanisms; hence, phage plaquing efficiency on strains derived from RN4220 indicates the abundance of GlcNAc residues on WTA, reflecting the *in vivo* activity of a WTA glycosyltransferase (19, 45). We also showed that a double mutant RN4220  $\Delta tarM\Delta tarS$ , which lacks both  $\alpha$ -O- and  $\beta$ -O-GlcNAc, did not only produce any GlcNAc on WTA but was resistant to  $\phi 11$  infection (6). In this study we complemented this mutant with various *tarM* variants, and the resulting complemented strains were used as hosts for plating  $\phi 11$  (Table 1). The efficiency of  $\phi 11$  plaquing reflects the level of WTA glycosylation in those *tarM* variant complemented strains. Cells expressing wild-type TarM, therefore, show the highest EOP, whereas cells lacking TarM activity do not show any plaque forming capacity. Falsifying concentration effects were ruled out by performing EOP experiments for each mutant at threshold titers from the same freshly prepared phage cultures.

**Structure-guided Mutagenesis of Active Site Residues**—To obtain insight into the catalytic mechanism and assess the validity of the observed interactions, several of the amino acids that lie in close proximity to the bound UDP-GlcNAc were mutated, and the enzymatic activities of the mutated proteins were analyzed in each case (“Experimental Procedures”). We specifically generated mutants E403A, K331S, R326S, and H249A, all of which probe interactions with substrate (Table 1). To confirm that the mutated proteins are still folded, each protein was purified and subjected to circular dichroism (CD) spectroscopy experiments and DLS analysis (Fig. 5, B and D). These data show that all mutants are structurally intact and have secondary structure elements that are indistinguishable from those of WT TarM.

Residue Glu-403 is clearly among the most important residues for catalysis. Its mutation to alanine essentially renders TarM inactive as it was not able to produce almost any observable spots on the bacterial lawn and generated no detectable output in the EOP-measurement (Fig. 5). The Glu-403 carboxyl group is thus essential for catalysis. Likewise, the mutation of Lys-331 to serine diminished all transferase activity in the EOP measurement. This mutation was aimed at removing a contact with the pyrophosphate group of UDP-GlcNAc as well as removing a potentially stabilizing interaction with Glu-403, as Lys-331 lies in close proximity to Glu-403, and the two residues could form a salt bridge during catalysis. Our results show that Lys-331 plays an essential role in substrate binding and/or catalysis. Residue Glu-411, which lies near the ribose of UDP, is also highly conserved. Its mutation to alanine also leads to severely reduced enzymatic activity, probably because the Glu-411 side chain is an integral part of the UDP-GlcNAc-binding site. Mutations of Arg-326 to serine and His-249 to alanine led to 20 and 30%, respectively, remaining WT activity. This suggests that both residues are important contact points that are, however, not essential for the reaction to proceed.

After the phospholysis reaction, the activated GlcNAc oriented on Glu-403 has its anomeric carbon pointed to the gap between Gt-N and Gt-C, where the activated acceptor (poly-ribitol-phosphate) must be located for the chemical reaction of glycosylation to occur. Unfortunately we lack a structure of

TarM bound to WTA fragments, which would shed light on the exact structure of the sugar-transfer transition state.

**Physiologic Role of the HUB and the Trimer**—To obtain insight into the putative function of the novel HUB domain, we selected a small number of residues in this domain for site-directed mutagenesis (Table 1). The rationale of these experiments was to subtly alter HUB regions mediating trimerization and to test the impact of these mutants on enzyme turnover efficiency. We generated single amino acid substitutions (K136S, N138Q, N180W) as well as a double (V159Y/C164R) and a triple (V159Y/C164R/K136S) mutation near the trimer interface. We observed a substantial decrease in TarM EOP (Fig. 5A) for K136S as well as the double and the triple mutant. We next selected the triple mutant and tested, alongside the wild type and E403A, its enzyme activity under *in vitro* conditions. Although E403A substantially decreased the EOP outcome and the enzymatic activity in the same order (Fig. 5E), indicating fully impaired glycosylation of WTA, we could not observe a comparable outcome for the triple mutant. Thus, the triple mutant produces different results *in vitro* and in the EOP assay. To rationalize this, we hypothesize that the triple mutant may lead to a subtle alteration or destabilization of the TarM trimer structure. Such a subtle change might not affect the catalytic activity of the enzyme in solution, but it might elicit a more severe effect in a physiologic setting. In support of this hypothesis, the CD spectra of the double and the triple mutants (Fig. 5B) show an additional shoulder around 205 nm, indicating a small alteration of secondary structure elements in the HUB domain. According to DLS and size exclusion chromatography analysis of recombinant TarM-variants, the variations in molecular dimensions are at best marginal (Fig. 5), indicating that the putative alteration is small.

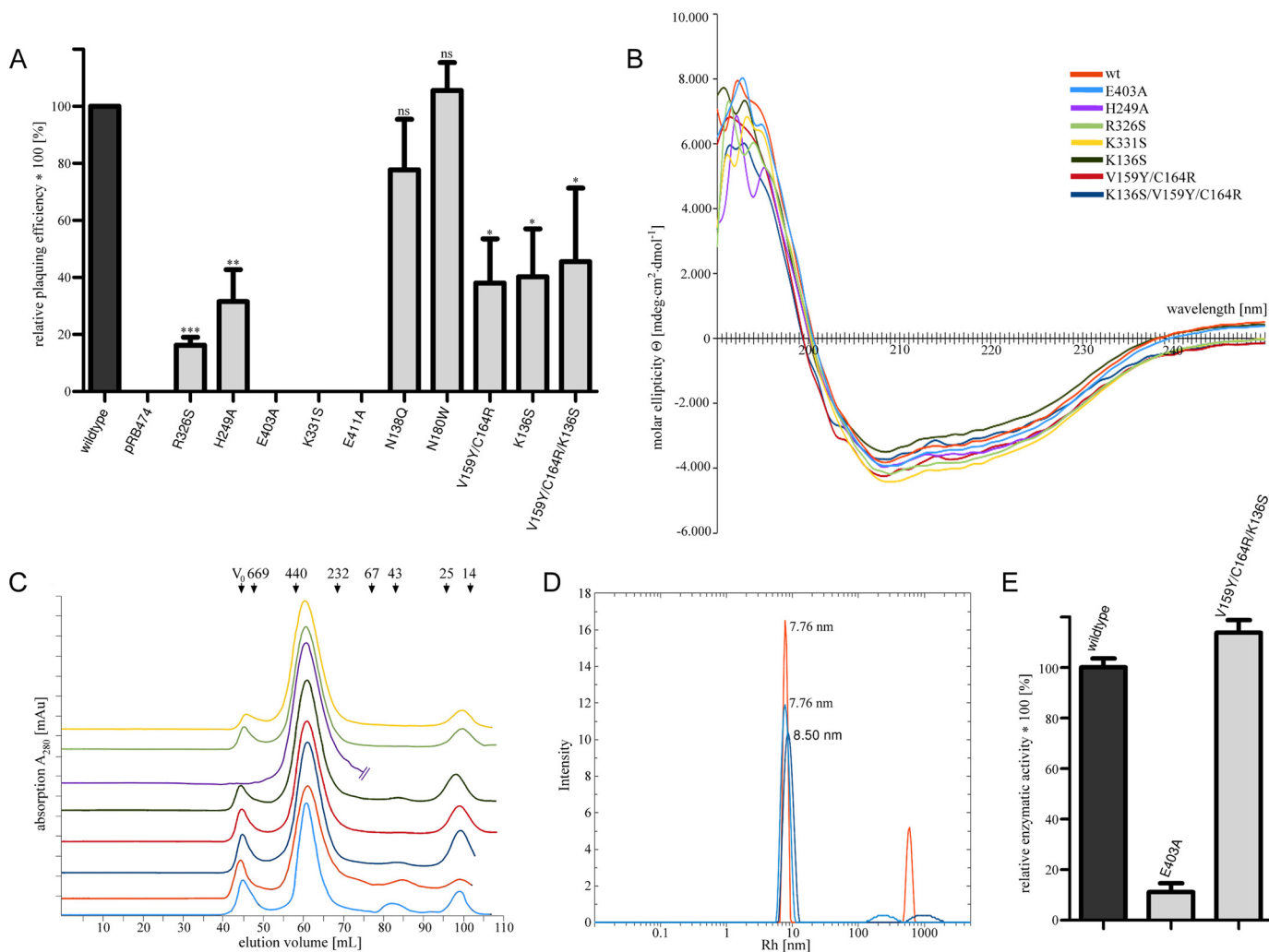
## DISCUSSION

We have determined the first structure of an enzyme in the biogenesis pathway of poly-RboP WTA, and we have characterized the ligand binding site of this enzyme. Our work sheds light on an essential aspect of *S. aureus* glycosylation and can be used as a template for understanding similar reactions in related organisms.

Glycosyltransferases can be classified into two groups that either retain the stereochemistry of the donor anomeric bond ( $\alpha \rightarrow \alpha$ ) or that invert this bond during the transfer reaction ( $\alpha \rightarrow \beta$ ). A common feature of GT-4 class enzymes is that they retain glycosyltransferases, and combined with previous biochemical data our structural analysis suggests that TarM is also a retaining glycosyltransferase that employs an  $S_N1$ -like mechanism in accordance with the widely acknowledged mechanism for a typical GT-4 class enzyme. The most salient structural features are shared by TarM and closely related GT-4 class enzymes MshA and BshA, and these latter enzymes can, therefore, serve as a useful basis for comparison.

The reaction mechanism for this class of enzymes has been established for MshA (38) and others (47–49). The acceptor substrates of GT-4 enzymes range from small molecules such as inositol phosphate to lipopolysaccharides and to S-layer glycoproteins. Although the resolution of UDP-GlcNAc-bound TarM is only 2.8 Å and although the Gt-C domain is more

## Structure of TarM



**FIGURE 5. Plating assay & CD-spectra of active site variants.** *A*, EOP assay for functional analysis of TarM point mutations. The histogram shows the effects of key amino acids of TarM on phage susceptibility, as probed by complementation of RN4220  $\Delta tarMtarS$  with TarM carrying specific point mutations. Approximately 1000 pfu of phage 11 were mixed with 100  $\mu$ l of bacteria suspension of optical density 0.4. After a brief incubation, soft agar was applied, and mixture was poured onto agar plates followed by overnight incubation at 37  $^{\circ}$ C. pfu was counted, and EOP of  $\Delta tarMtarS$  complemented with wild-type *tarM* was designated as 100%. Mutant-*tarM* complements are indicated in relation as the mean of four experiments  $\pm$  S.D. Statistically significant differences of mutant TarM from wild-type TarM complementation were calculated by paired two-tailed Student's *t* test: *ns*, not significant,  $p > 0.05$ ; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ . *B*, overlaid CD spectra of recombinant TarM active site mutants H249A, R326S, K331S, and E403A and HUB domain mutants K136S, V159Y/C164R, and K136S/V159Y/C164R. Recombinant enzymes were purified according to the same protocol used for the wild-type enzyme (see "Experimental Procedures"). CD measurements were performed at concentrations ranging from 0.5 to 0.8 mg/ml in buffer G. *C*, overlay of size exclusion chromatography (SD200 16/60, GE Healthcare) elution profiles from the last step of recombinant protein purification with buffer D at 4  $^{\circ}$ C. All TarM variants elute in a volume range with an average peak point corresponding to a molecular size estimated to be 300 kDa according to the size calibration proteins shown on top. Color coding is the same as in Fig. 5B. *D*, overlaid DLS spectra of recombinant wild type, E403A, and K136S/V159Y/C164R TarM at concentrations of 0.5–1 mg/ml performed at 20  $^{\circ}$ C in buffer D. The calculated radius for a spherical protein model of 170 kDa is 5.35 nm. Color coding is the same as in Fig. 5B. *E*, relative *in vitro* activity of TarM and selected TarM variants. Activity of wild-type TarM was set to 100%. Values are given as the mean of three experiments ( $n = 3$ )  $\pm$  S.D. The reactions were carried out in the presence of 2 mM UDP-GlcNAc and 25  $\mu$ M WTA at room temperature. The reaction was followed via a coupled enzymatic assay with non-saturating amounts of TarM variants.

mobile and less well defined by electron density than the remainder of the protein, the electron density for UDP- $\alpha$ -GlcNAc is nevertheless unambiguous and allows placement of the ligand into the structure in the conformation shown in Fig. 4.

A critical difference to other GT-4 class enzymes such as MshA and BshA is that TarM has a HUB domain that is inserted into the Gt-N domain between helices  $\alpha$ 4 and  $\alpha$ 5 (Fig. 2D) and that folds into a long antiparallel  $\beta$ -sheet. The point of insertion of the HUB domain into Gt-N also happens to be the dimerization site for MshA and BshA. The HUB domain gives rise to a unique trimeric, propeller-like assembly of three gly-

cosyltransferase domains. Given the proximity of the three-fold symmetry axis to the active sites, the HUB-generated trimer may also participate in interactions with WTA and assist with catalysis. Our mutational analysis clearly implicates the HUB domain in this process. It is interesting that a BLAST (50) sequence search of protein databases only finds HUB-like sequences in TarM homologs of other Gram-positive bacteria (NCBI# WP\_029331270.1, identity 53%, similarity 71%; NCBI# WP\_014124998.1, identity 43%, similarity 65%; NCBI# WP\_025702814.1, identity 27%, similarity 50%; NCBI# WP\_003756742.1, identity 33%, similarity 55%). To analyze the level of conservation of residues in TarM and its homologs, we generated a

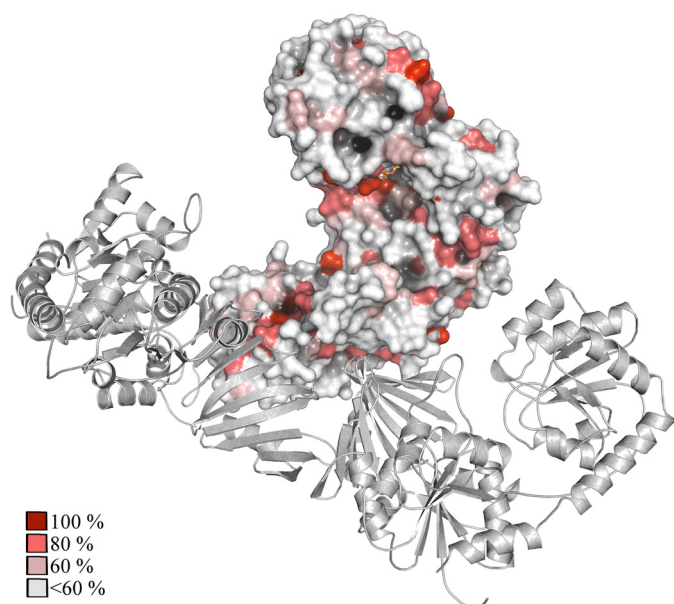


FIGURE 6. **Conservation of TarM.** Conserved residues were identified via multiple alignment (ClustalW) of TarM with glycosyl transferase family 1 (*Exiguobacterium oxidotolerans*) (NCBI# WP\_029331270.1), hypothetical protein (*Paenibacillus forsythiae*) (NCBI# WP\_025702814.1), putative glycosyltransferase (*Tetragenococcus halophilus*) (NCBI# WP\_014124998.1), glycosyl transferase (*Listeria grayi*) (NCBI# WP\_003756742.1). Conserved residues were highlighted by coloring the TarM model surface according to the following scheme: 100% conserved (dark red), 80% conserved (medium red), 60% conserved (light red), no significant conservation (white).

sequence alignment (not shown) and colored the TarM surface according to the level of conservation (Fig. 6). As expected, the active site region and the UDP-GlcNAc binding site are rather conserved (red in Fig. 6). Interestingly, surface-exposed portions of the HUB domain that lie adjacent to the active site region are also well conserved, and because there is no obvious structural reason for this conservation, we predict that these regions might be involved in the binding of the second substrate, the RboP acceptor chain. As TarM-mediated WTA glycosylation is thought to constitute a general pathway in Gram-positive bacteria with RboP-WTA (19), it seems likely that the HUB domain acts similarly in these related organisms. Our mutagenesis results indirectly suggest a role for the HUB in WTA glycosylation, although the exact mechanism remains to be elucidated.

A recent structural analysis of the streptococcal glycosyltransferase GtfA has identified a novel domain that is very similar in structure to the HUB domain and that is also inserted into a glycosyltransferase subdomain at a similar location (40). The GtfA domain is the only structure in the DALI database with any significant structural homology to the TarM HUB (Z-score 11.6, r.m.s.d. 2.1 Å, 103 aligned residues, sequence identity 16%, PDB 4PQG). However, GtfA is clearly monomeric, and the enzyme also does not act on WTA. A structural alignment shows that the novel GtfA domain is unlikely to form a similar trimeric arrangement due to an insertion sequence (PVDNK) that extends the turn connecting strands  $\beta 9$  and  $\beta 10$  (Fig. 7). This loop is much shorter in TarM, allowing trimer formation, and the tip of the loop moreover carries Val-159, which makes direct contacts to the two other Val-159 residues in the trimer and thus stabilizes the trimeric arrangement. Con-

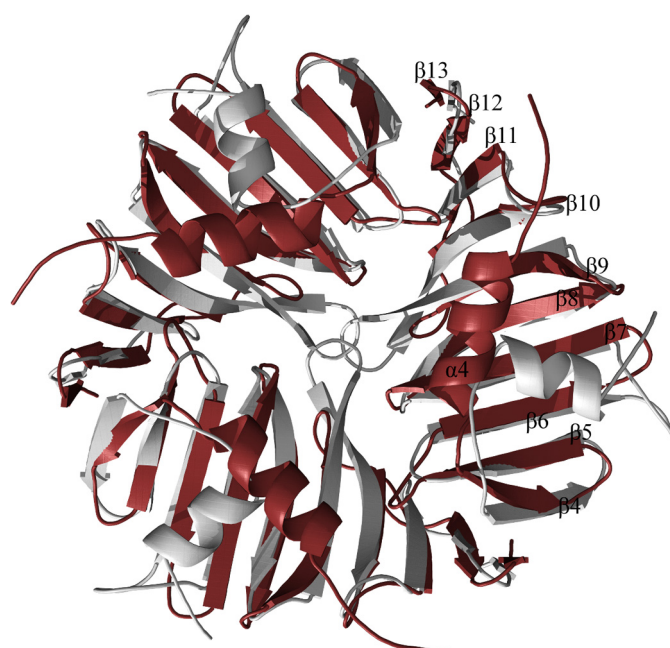


FIGURE 7. **Comparison of the HUB domain trimer of TarM with the monomeric DUF1975 domain of SpGtfA.** Superposition of the trimeric TarM HUB (red) with three monomeric DUF1975 domains of GtfA (silver) (r.m.s.d. 2.2 Å, aligned residues 116). It is evident that a trimeric arrangement of the GtfA domains would lead to clashes.

sistent with this, a mutation of Val-159 that would disrupt the trimer affects the ability of TarM to process WTA.

It is possible that WTA-GTs such as TarM, and its relatives have evolved the HUB domain to assemble into trimers and thereby facilitate the glycosylation of complex glycopolymers. The direct distance of two neighboring active sites in the TarM trimer is 72 Å, which corresponds to about eight or nine ribitol units in an extended chain. Thus, a single TarM trimer could simultaneously glycosylate the same poly-RboP substrate at different locations. It is not currently known which RboP units in the long polyribitol-phosphate chain are glycosylated, but it is likely that the glycosyltransferases acting on WTA have a mechanism that enables them to move along the polyribitol chain and selectively glycosylate specific units. The result of such a glycosylation pattern is for example relevant for the selectivity of pattern recognition receptors. Future studies of TarM in complex with WTA components should help reveal the molecular details of this process.

By now only a handful of teichoic acid biogenesis-affiliated protein structures have been solved for either polyglycerol phosphate-type or complex-type WTA, e.g. TagF from *Bacillus subtilis* (51), *Streptococcus* TarI (spr1149, PDB 2VSH) (52), MnaA (BA5590) from *B. anthracis* (PDB 3BEO) (53), putative WTA ligase from *Streptococcus* (54), or lipoteichoic acid polymerase LtaS (55). TarM is the first enzyme structure in the biogenesis pathway of poly(RboP)-type WTA to be reported.

*Acknowledgments*—X-ray data were collected at beam line X06DA of the Swiss Light Source (Villigen, Switzerland), and we are grateful to the beam line staff, particularly Vincent Olieric and Tomizaki Takashi, for assistance. We also thank Nicolas Binder and Simon Huber for the purification of recombinant TarM variants used for CD spectroscopy.

## REFERENCES

- Baron, S. (1996) *Medical Microbiology*, 4th Ed., University of Texas Medical Branch at Galveston, Galveston, TX
- Schäffer, C., and Messner, P. (2005) The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together. *Microbiology* **151**, 643–651
- Weidenmaier, C., and Peschel, A. (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat. Rev. Microbiol.* **6**, 276–287
- Leonard, P. G., Golemi-Kotra, D., and Stock, A. M. (2013) Phosphorylation-dependent conformational changes and domain rearrangements in *Staphylococcus aureus* VraR activation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8525–8530
- Maki, H., Yamaguchi, T., and Murakami, K. (1994) Cloning and characterization of a gene affecting the methicillin resistance level and the autolysis rate in *Staphylococcus aureus*. *J. Bacteriol.* **176**, 4993–5000
- Brown, S., Xia, G., Luhachack, L. G., Campbell, J., Meredith, T. C., Chen, C., Winstel, V., Gekeler, C., Irazoqui, J. E., Peschel, A., and Walker, S. (2012) Methicillin resistance in *Staphylococcus aureus* requires glycosylated wall teichoic acids. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18909–18914
- Qamar, A., and Golemi-Kotra, D. (2012) Dual roles of FmtA in *Staphylococcus aureus* cell wall biosynthesis and autolysis. *Antimicrob. Agents Chemother.* **56**, 3797–3805
- Bera, A., Biswas, R., Herbert, S., Kulauzovic, E., Weidenmaier, C., Peschel, A., and Götz, F. (2007) Influence of wall teichoic acid on lysozyme resistance in *Staphylococcus aureus*. *J. Bacteriol.* **189**, 280–283
- Kristian, S. A., Lauth, X., Nizet, V., Goetz, F., Neumeister, B., Peschel, A., and Landmann, R. (2003) Alanylation of teichoic acids protects *Staphylococcus aureus* against Toll-like receptor 2-dependent host defense in a mouse tissue cage infection model. *J. Infect. Dis.* **188**, 414–423
- Weidenmaier, C., Kokai-Kun, J. F., Kristian, S. A., Chanturiya, T., Kalbacher, H., Gross, M., Nicholson, G., Neumeister, B., Mond, J. J., and Peschel, A. (2004) Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. *Nat. Med.* **10**, 243–245
- Weidenmaier, C., Peschel, A., Xiong, Y. Q., Kristian, S. A., Dietz, K., Yeaman, M. R., and Bayer, A. S. (2005) Lack of wall teichoic acids in *Staphylococcus aureus* leads to reduced interactions with endothelial cells and to attenuated virulence in a rabbit model of endocarditis. *J. Infect. Dis.* **191**, 1771–1777
- Atilano, M. L., Pereira, P. M., Yates, J., Reed, P., Veiga, H., Pinho, M. G., and Filipe, S. R. (2010) Teichoic acids are temporal and spatial regulators of peptidoglycan cross-linking in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18991–18996
- Brown, S., Santa Maria, J. P., Jr., and Walker, S. (2013) Wall teichoic acids of gram-positive bacteria. *Annu. Rev. Microbiol.* **67**, 313–336
- Kojima, N., Araki, Y., and Ito, E. (1985) Structure of the linkage units between ribitol teichoic acids and peptidoglycan. *J. Bacteriol.* **161**, 299–306
- Brown, S., Zhang, Y. H., and Walker, S. (2008) A revised pathway proposed for *Staphylococcus aureus* wall teichoic acid biosynthesis based on *in vitro* reconstitution of the intracellular steps. *Chem. Biol.* **15**, 12–21
- Weidenmaier, C., Peschel, A., Kempf, V. A., Lucindo, N., Yeaman, M. R., and Bayer, A. S. (2005) DltABCD- and MprF-mediated cell envelope modifications of *Staphylococcus aureus* confer resistance to platelet microbicidal proteins and contribute to virulence in a rabbit endocarditis model. *Infect. Immun.* **73**, 8033–8038
- Peschel, A., and Sahl, H. G. (2006) The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat. Rev. Microbiol.* **4**, 529–536
- Li, M., Lai, Y., Villaruz, A. E., Cha, D. J., Sturdevant, D. E., and Otto, M. (2007) Gram-positive three-component antimicrobial peptide-sensing system. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9469–9474
- Xia, G., Maier, L., Sanchez-Carballo, P., Li, M., Otto, M., Holst, O., and Peschel, A. (2010) Glycosylation of wall teichoic acid in *Staphylococcus aureus* by TarM. *J. Biol. Chem.* **285**, 13405–13415
- Kurokawa, K., Jung, D. J., An, J. H., Fuchs, K., Jeon, Y. J., Kim, N. H., Li, X., Tateishi, K., Park, J. A., Xia, G., Matsushita, M., Takahashi, K., Park, H. J., Peschel, A., and Lee, B. L. (2013) Glycoepitopes of staphylococcal wall teichoic acid govern complement-mediated opsonophagocytosis via human serum antibody and mannose-binding lectin. *J. Biol. Chem.* **288**, 30956–30968
- Breton, C., Fournel-Gigleux, S., and Palcic, M. M. (2012) Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* **22**, 540–549
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heeger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222–D230
- Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
- Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007) Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230
- Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772–1779
- Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972
- Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008) Automated macromolecular model building for x-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367
- Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
- Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25
- Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184–2195
- DeLano, W. L. (2010) *The PyMOL Molecular Graphics System*, Version 1.3r1, Schrodinger, LLC, New York
- Mulder, G. J., and van Doorn, A. B. (1975) A rapid NAD<sup>+</sup>-linked assay for microsomal uridine diphosphate glucuronyltransferase of rat liver and some observations on substrate specificity of the enzyme. *Biochem. J.* **151**, 131–140
- Winstel, V., Sanchez-Carballo, P., Holst, O., Xia, G., and Peschel, A. (2014) Biosynthesis of the unique wall teichoic acid of *Staphylococcus aureus* lineage ST395. *mBio* **5**, e00869
- Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549
- Vetting, M. W., Frantom, P. A., and Blanchard, J. S. (2008) Structural and enzymatic analysis of MshA from *Corynebacterium glutamicum*: substrate-assisted catalysis. *J. Biol. Chem.* **283**, 15834–15844
- Parsonage, D., Newton, G. L., Holder, R. C., Wallace, B. D., Paige, C., Hamilton, C. J., Dos Santos, P. C., Redinbo, M. R., Reid, S. D., and Clairborne, A. (2010) Characterization of the N-acetyl- $\alpha$ -D-glucosaminyl l-malate synthase and deacetylase functions for bacillithiol biosynthesis in *Bacillus anthracis*. *Biochemistry* **49**, 8398–8414
- Shi, W. W., Jiang, Y. L., Zhu, F., Yang, Y. H., Shao, Q. Y., Yang, H. B., Ren, Y. M., Wu, H., Chen, Y., and Zhou, C. Z. (2014) Structure of a novel O-linked N-acetyl-D-glucosamine (O-GlcNAc) transferase, GtfA, reveals insights into the glycosylation of pneumococcal serine-rich repeat adhesins. *J. Biol. Chem.* **289**, 20898–20907
- Batt, S. M., Jabeen, T., Mishra, A. K., Veerapen, N., Krumbach, K., Egg-

- eling, L., Besra, G. S., and Fütterer, K. (2010) Acceptor substrate discrimination in phosphatidyl-myo-inositol mannoside synthesis: structural and mutational analysis of mannosyltransferase *Corynebacterium glutamicum* PimB'. *J. Biol. Chem.* **285**, 37741–37752
42. Xiang, Y., Baxa, U., Zhang, Y., Steven, A. C., Lewis, G. L., Van Etten, J. L., and Rossmann, M. G. (2010) Crystal structure of a virus-encoded putative glycosyltransferase. *J. Virol.* **84**, 12265–12273
  43. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
  44. Wright, A. K., Duncan, R. C., and Beekman, K. A. (1973) A numerical inversion of the Perrin equations for rotational diffusion constants for ellipsoids of revolution by iterative techniques. *Biophys. J.* **13**, 795–803
  45. Xia, G., Corrigan, R. M., Winstel, V., Goerke, C., Gründling, A., and Pechel, A. (2011) Wall teichoic acid-dependent adsorption of staphylococcal siphovirus and myovirus. *J. Bacteriol.* **193**, 4006–4009
  46. Wendlinger, G., Loessner, M. J., and Scherer, S. (1996) Bacteriophage receptors on *Listeria monocytogenes* cells are the *N*-acetylglucosamine and rhamnose substituents of teichoic acids or the peptidoglycan itself. *Microbiology* **142**, 985–992
  47. Greenfield, L. K., Richards, M. R., Vinogradov, E., Wakarchuk, W. W., Lowary, T. L., and Whitfield, C. (2012) Domain organization of the polymerizing mannosyltransferases involved in synthesis of the *Escherichia coli* O8 and O9a lipopolysaccharide O-antigens. *J. Biol. Chem.* **287**, 38135–38149
  48. Steiner, K., Hagelueken, G., Messner, P., Schäffer, C., and Naismith, J. H. (2010) Structural basis of substrate binding in WsaF, a rhamnosyltransferase from *Geobacillus stearothermophilus*. *J. Mol. Biol.* **397**, 436–447
  49. Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2006) Insights into the synthesis of lipopolysaccharide and antibiotics through the structures of two retaining glycosyltransferases from family GT4. *Chem. Biol.* **13**, 1143–1152
  50. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402
  51. Lovering, A. L., Lin, L. Y., Sewell, E. W., Spreter, T., Brown, E. D., and Strynadka, N. C. (2010) Structure of the bacterial teichoic acid polymerase TagF provides insights into membrane association and catalysis. *Nat. Struct. Mol. Biol.* **17**, 582–589
  52. Baur, S., Marles-Wright, J., Buckenmaier, S., Lewis, R. J., and Vollmer, W. (2009) Synthesis of CDP-activated ribitol for teichoic acid precursors in *Streptococcus pneumoniae*. *J. Bacteriol.* **191**, 1200–1210
  53. Velloso, L. M., Bhaskaran, S. S., Schuch, R., Fischetti, V. A., and Stebbins, C. E. (2008) A structural basis for the allosteric regulation of non-hydrolysing UDP-GlcNAc 2-epimerases. *EMBO Rep.* **9**, 199–205
  54. Kawai, Y., Marles-Wright, J., Cleverley, R. M., Emmins, R., Ishikawa, S., Kuwano, M., Heinz, N., Bui, N. K., Hoyland, C. N., Ogasawara, N., Lewis, R. J., Vollmer, W., Daniel, R. A., and Errington, J. (2011) A widespread family of bacterial cell wall assembly proteins. *EMBO J.* **30**, 4931–4941
  55. Campeotto, I., Percy, M. G., MacDonald, J. T., Förster, A., Freemont, P. S., and Gründling, A. (2014) Structural and mechanistic insight into the *Listeria monocytogenes* two-enzyme lipoteichoic acid synthesis system. *J. Biol. Chem.* **289**, 28054–28069



**Structural and Enzymatic Analysis of TarM Glycosyltransferase from *Staphylococcus aureus* Reveals an Oligomeric Protein Specific for the Glycosylation of Wall Teichoic Acid**

Cengiz Koç, David Gerlach, Sebastian Beck, Andreas Peschel, Guoqing Xia and Thilo Stehle

*J. Biol. Chem.* 2015, 290:9874-9885.

doi: 10.1074/jbc.M114.619924 originally published online February 19, 2015

---

Access the most updated version of this article at doi: [10.1074/jbc.M114.619924](https://doi.org/10.1074/jbc.M114.619924)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 53 references, 27 of which can be accessed free at <http://www.jbc.org/content/290/15/9874.full.html#ref-list-1>

# SCIENTIFIC REPORTS



OPEN

## An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*

Received: 08 December 2015  
Accepted: 28 April 2016  
Published: 23 May 2016

Xuehua Li<sup>1</sup>, Cengiz Koç<sup>2</sup>, Petra Kühner<sup>1</sup>, York-Dieter Stierhof<sup>3</sup>, Bernhard Krismer<sup>1</sup>, Mark C. Enright<sup>4</sup>, José R. Penadés<sup>5</sup>, Christiane Wolz<sup>1</sup>, Thilo Stehle<sup>2,6,7</sup>, Christian Cambillau<sup>8</sup>, Andreas Peschel<sup>1,7</sup> & Guoqing Xia<sup>1,7,9</sup>

Despite the importance of phages in driving horizontal gene transfer (HGT) among pathogenic bacteria, the underlying molecular mechanisms mediating phage adsorption to *S. aureus* are still unclear. Phage  $\phi 11$  is a siphovirus with a high transducing efficiency. Here, we show that the tail protein Gp45 localized within the  $\phi 11$  baseplate. Phage  $\phi 11$  was efficiently neutralized by anti-Gp45 serum, and its adsorption to host cells was inhibited by recombinant Gp45 in a dose-dependent manner. Flow cytometry analysis demonstrated that biotin-labelled Gp45 efficiently stained the wild-type *S. aureus* cell but not the double knockout mutant  $\Delta tarM/S$ , which lacks both  $\alpha$ - and  $\beta$ -O-GlcNAc residues on its wall teichoic acids (WTAs). Additionally, adsorption assays indicate that GlcNAc residues on WTAs and O-acetyl groups at the 6-position of muramic acid residues in peptidoglycan are essential components of the  $\phi 11$  receptor. The elucidation of Gp45-involved molecular interactions not only broadens our understanding of siphovirus-mediated HGT, but also lays the groundwork for the development of sensitive affinity-based diagnostics and therapeutics for *S. aureus* infection.

Recently, there has been a renewed interest in phage-bacteria interactions because phages have not only profound influence on the biology of bacterial pathogens<sup>1,2</sup> but also promising applications in the detection of pathogens, the biocontrol of bacterial food contamination<sup>3</sup>, and the treatment of bacterial infections<sup>4</sup>.

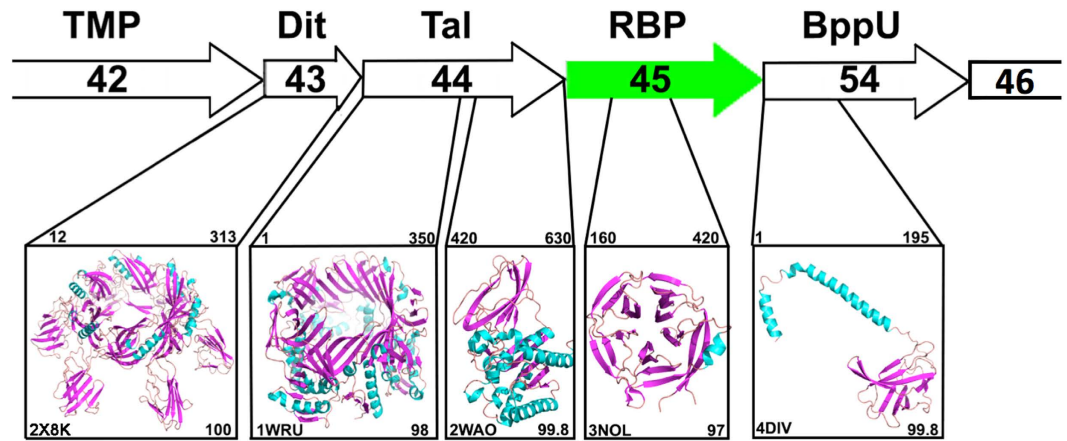
Phages infecting gram-positive bacteria need to adsorb and penetrate a cell envelope with a thick peptidoglycan meshwork. The mechanism of phage adsorption and genome translocation across the gram-positive cell envelope remains largely unknown for many phages, with the exception of a few dairy phages infecting *Lactobacillus*, *Lactococcus*, or *Streptococcus* spp. Genome comparison of several dairy phages with different host ranges enabled the identification of their receptor binding proteins (RBPs), which are essential for phage adsorption and virulence. The first RBP recognizing a gram-positive cell envelope was identified from phage Dt1 infecting *Streptococcus thermophilus*<sup>5</sup>. Recently, the structures of RBPs from several lactococcal phages were solved<sup>6</sup>. These RBPs are generally homotrimeric and are composed of three modular structures, which encompass the N-terminal shoulder domain for connection to the virion, a  $\beta$ -helical linker or the neck domain, and the C-terminal head domain bearing the receptor binding site for host recognition<sup>6</sup>.

Wall teichoic acids (WTAs) are phosphate-rich anionic glycopolymers covalently linked to the peptidoglycan in gram-positive bacteria. The two common types of WTA are either poly-1,3 glycerol-phosphate (GroP) or poly-1,5 ribitol-phosphate (RboP). The main chains of both types of WTAs can be further substituted with sugar residues and alanyl groups<sup>7</sup>. Previous studies on *Bacillus* phage SPP1 revealed that adsorption of this phage to its

<sup>1</sup>Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, 72076 Tübingen, Germany.

<sup>2</sup>Interfaculty Institute of Biochemistry, University of Tübingen, 72076, Tübingen, Germany. <sup>3</sup>Center for Plant Molecular Biology, University of Tübingen, 72076, Tübingen, Germany. <sup>4</sup>School of Healthcare Sciences, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, United Kingdom. <sup>5</sup>Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom.

<sup>6</sup>Vanderbilt University, School of Medicine, Nashville, TN 37232, USA. <sup>7</sup>German Center for Infection Research (DZIF), partner site Tübingen, Germany. <sup>8</sup>Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique, UMR 6098, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France. <sup>9</sup>Institute of Inflammation & Repair, Faculty of Medical and Human Sciences, University of Manchester, Oxford Road, Manchester, M13 9PT, United Kingdom. Correspondence and requests for materials should be addressed to G.X. (email: guoqing.xia@manchester.ac.uk)



**Figure 1.** HHpred analysis of the four structural gene products following the *tmp* (*gp42*) of  $\phi 11$ . The genes *gp42*, *gp43*, *gp44*, *gp45*, *gp54* and *gp46* are represented by arrows. The structural homologues of these tail proteins are presented in the boxes beneath the corresponding genes. The PDB identifiers and ribbon structures ( $\alpha$ -helices in blue,  $\beta$ -strands in violet) are shown for the structural homologues. The starting and ending amino acid residues of the regions, which could well align with these structural homologues are indicated above the boxes. The similarity probability (%) returned by HHpred is indicated to the right of the PDB identifier. The PDB entries shown here include 2  $\times$  8K, Bacillus phage SPP1 baseplate Dit protein; 1WRU, Tail associated lysin (Tal) of bacteriophage MU; 2WAO, carbohydrate esterase of *Clostridium thermocellum*; 3NOL, glutaminy cyclase of *Zymomonas mobilis*; and 4DIV, BppU of Lactococcus phage TP901-1.

host cell initially depends on the reversible binding to WTAs, which accelerates the subsequent irreversible binding to membrane receptor YueB<sup>8</sup>. Interestingly, incubation of the purified SPP1 virions with recombinant YueB leads to phage DNA release *in vitro*<sup>9</sup>, indicating that the binding of this protein is the trigger for DNA injection.

*Staphylococcus aureus* is a gram-positive pathogen that causes not only superficial skin infections but also severe, deep tissue infections such as endocarditis, osteomyelitis, septic arthritis, and bacteraemia. It is very well known that phages or mainly siphoviruses play vital roles in the virulence, adaptation, and evolution of *S. aureus*<sup>1,2</sup>. However, it remains unclear how siphoviruses recognize *S. aureus* and what ligand-receptor interactions mediate phage adsorption to the cell surface of *S. aureus*.

Among all *S. aureus* phages,  $\phi 11$  is probably one of the best-studied siphoviruses due to its high transducing efficiency and broad application in transducing genetic markers among *S. aureus* strains. Recently, there has been a growing interest in studying the function of  $\phi 11$  as a helper phage mediating the horizontal gene transfer (HGT) of *S. aureus* pathogenicity islands (SaPIs)<sup>10</sup>. We have shown that staphylococcal siphoviruses use  $\alpha$ -O-GlcNAc modified WTA as a receptor<sup>11</sup> and that WTA structures govern phage-mediated horizontal transfer of SaPIs among major bacterial pathogens<sup>12</sup>. Although many structural proteins of  $\phi 11$  have been reported<sup>13,14</sup>, its receptor binding protein (RBP) has yet to be identified. Here we report the identification and characterization of the  $\phi 11$  RBP and the major components of its receptor in the cell wall of *S. aureus*. These data not only provide novel insight into phage-host recognition at the staphylococcal cell surface, but also establish a molecular basis to develop novel diagnostics and therapeutic treatments of *S. aureus* infection.

## Results

**Sequence analysis of the putative baseplate proteins of  $\phi 11$ .** In staphylococcal siphovirus genomes, the genes coding for tail proteins are usually located downstream of the gene of the tape measure protein (TMP) and upstream of the lysis module<sup>2,15</sup>. Among the genes localized between *tmp* (*gp42*) and the lysis module, *gp43*, *gp44*, *gp45* and *gp54* (Fig. 1) were previously shown to be essential for phage  $\phi 11$  infectivity<sup>13,16</sup>. Of note, *gp54* was not initially annotated in the genome of  $\phi 11$ <sup>17</sup>, but it was later identified as an open reading frame localized between *gp45* and *gp46*<sup>13</sup>. To advance an understanding of the putative functions of the proteins encoded by these four essential genes, HHpred<sup>18</sup> (Homology detection and structure prediction by HMM-HMM comparison) analysis was carried out for each protein in addition to BlastP analysis at NCBI (<http://goo.gl/DE9BkO>).

The HHpred analysis identified Gp43 with 100% probability as a distal tail protein (Dit) because it is similar to the Dit protein (PDB 2  $\times$  8K) in the baseplate of the siphophage SPP1, which infects *Bacillus subtilis*<sup>19</sup> (Fig. 1). The N-terminal regions of Dit proteins form a hexameric ring and are very conserved among phages<sup>20</sup>, although their C-terminal peripheral domains may differ considerably<sup>21</sup>.

BlastP search revealed that Gp44 possesses an endopeptidase domain at its N-terminus (1–350 residues) and a SGNH/GDSL hydrolase domain at its C-terminus (400–633 residues). Of note, the SGNH hydrolase represents a diverse family of lipases and esterases, but the enzyme activity of Gp44 is yet to be characterized experimentally. Further sequence analysis by HHpred revealed that the N-terminal domain of Gp44 aligns well with the tail associated lysin (Tal) of bacteriophage MU (PDB 1WRU)<sup>22</sup>, and its C-terminal domain exhibits striking similarity to a carbohydrate esterase (PDB 2WAO) from *Clostridium thermocellum* (Fig. 1). Tal proteins are structurally

similar to Gp27, a baseplate component of the puncturing device of phage T4<sup>23</sup>. Notably, the gene *tal* is always localized directly downstream of the gene *dit* in siphophage genomes. In the  $\phi$ 11 tail module, *gp44* exists directly downstream of *gp43* (*dit*). Hence, both sequence homology and conserved genome localization suggest that *gp44* encodes a Tal protein. Recently, it was shown that phage mutants deficient in Gp43 (Dit), or Gp44 (Tal) were defective in tails, suggesting that these two baseplate proteins are required for tail formation<sup>16</sup>. Furthermore, it was shown that the tail protein Gp49 possesses peptidoglycan hydrolase activity but is dispensable for  $\phi$ 11 infectivity<sup>16,24</sup>. These facts suggest that  $\phi$ 11 may have two virion-associated peptidoglycan hydrolases, Tal and Gp49, but the activity of Tal needs to be verified by further experiments.

BlastP search with Gp45 as a query returned a hit of ORF636, which shares 44% identity with Gp45 and is localized at the tail tip of phage  $\phi$ SLT, a serogroup A phage of *S. aureus*. Of note, the tail protein ORF636 was characterized as an adhesion protein essential for  $\phi$ SLT adsorption and infectivity<sup>25</sup>. HHpred analysis revealed that the central part of Gp45, covering amino acid residues 160–420, shares high similarity with 5-bladed propeller proteins (Fig. 1), for example the glutamyl cyclase of *Zymomonas mobilis* (PDB 3NOL). The segment upstream was predicted to be  $\alpha$ -helical by Jpred<sup>26</sup>, while the segment downstream was predicted to form  $\beta$ -strands.

Just downstream of *gp45*, *gp54* most likely encodes an upper baseplate protein (BppU)<sup>27</sup>. The N-terminus of Gp54 (amino-acids 1–195) displays high similarity to a large part of the BppU<sup>27</sup>, which attaches the RBP to the central baseplate core in lactococcal phage TP901-1 (Fig. 1). In TP901-1, BppU assembles as a trimer. Its N-terminus (amino-acids 1–120) is a stand-alone domain, while amino-acids 121–193 assemble as a triple  $\alpha$ -helix bundle. This structure is followed by a trimeric all-beta domain (~100 residues), to which the N-terminus of RBP is plugged in<sup>27</sup>. Thus, the C-terminus of BppU and the RBP exhibit strong shape complementarity in phage TP901-1. However, the C-terminus of Gp54 possesses a domain of unknown function, which includes ~400 amino acid residues, and is much larger than that of BppU in TP901-1.

Taken together, the HHpred analyses revealed that Gp43 (Dit), Gp44 (Tal), Gp45 (ORF636-like protein), and Gp54 (BppU) very likely constitute the baseplate of  $\phi$ 11. Moreover, the central part of the  $\phi$ 11 baseplate gathering Dit, Tal, and the N-terminus of BppU, forming the dsDNA passage, is similar to that of other phages<sup>20</sup>, whereas the role of the tail proteins, Gp45 and Gp54, most likely located at the periphery of the baseplate is elucidated below.

**Localization of Gp45 and Gp54 at the baseplate of  $\phi$ 11.** Baseplate proteins or tail fibre proteins play critical roles in phage adsorption, the first step of phage replication cycle<sup>6</sup>. Previously it was shown that the two putative tail proteins Gp45 and Gp54 were essential for phage infectivity<sup>13</sup>. To demonstrate that both Gp54 and Gp45 are localized at the tail tip, both Gp54 and Gp45 antisera were raised and used for immunogold labelling of  $\phi$ 11. Electron micrographs of negatively stained phage samples indicate that Gp45 and Gp54 are clearly localized at the tail baseplate of  $\phi$ 11 (Fig. 2a,b), whereas immunogold labelling of mutant phages deficient in *gp45* or *gp54* resulted in negligible background labelling (Fig. 2c).

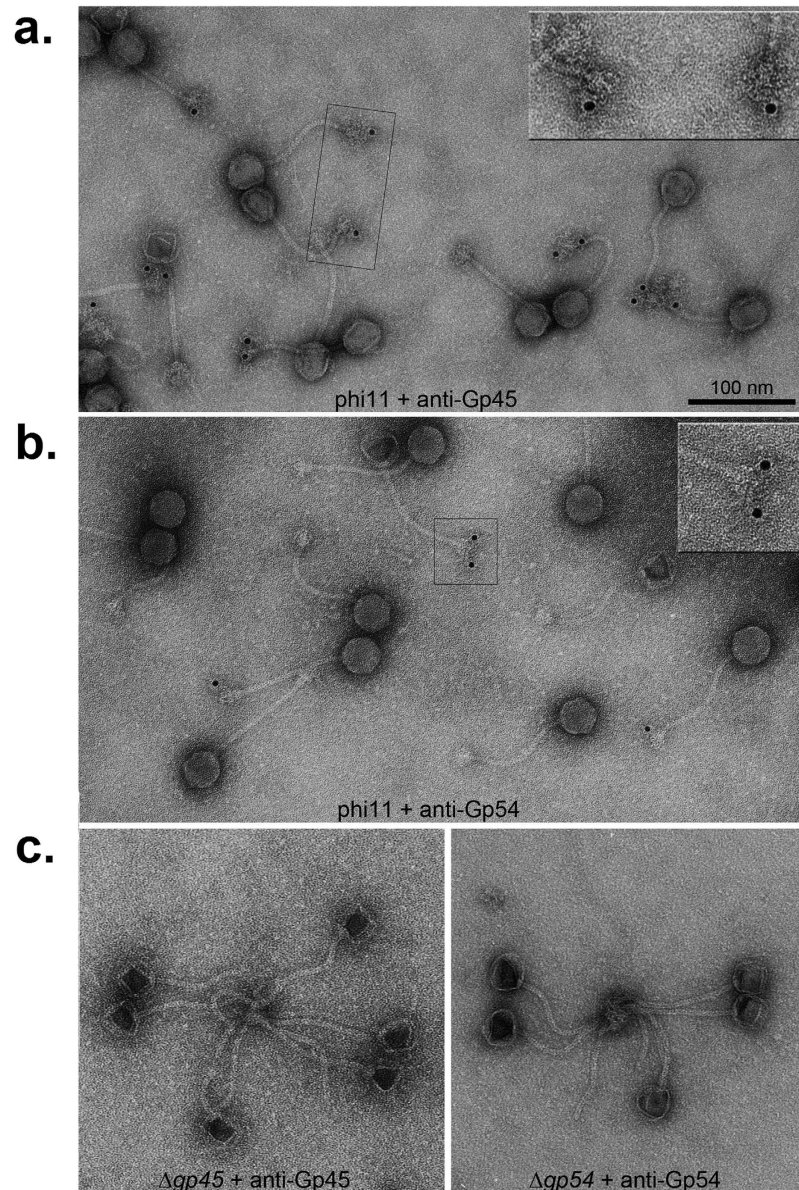
**Neutralization of  $\phi$ 11 infection with anti-Gp45 or anti-Gp54 serum.** As both Gp45 and Gp54 are baseplate proteins, their roles in phage adsorption and infection were analysed. Phage  $\phi$ 11 virions were pre-incubated with increasing concentrations of antisera before plating on the host. Notably, pre-immune sera exhibited hardly any inhibitory effects on phage plating efficiency (data not shown), whereas both anti-Gp45 and anti-Gp54 serum decreases the plating efficiency of  $\phi$ 11 in a dose-dependent manner (Fig. 3a,b), which clearly suggests that these sera can specifically neutralize  $\phi$ 11 infectivity. It is most likely that masking of Gp45 or Gp54 with antiserum prevents their access to the phage receptor in the cell wall, hence blocks the phage adsorption and leads to neutralization of  $\phi$ 11.

**Gp45 binds to the cell wall with  $\alpha$ - or  $\beta$ -O-GlcNAc modified WTAs.** To investigate the molecular interaction of  $\phi$ 11 with its cognate receptor on the host cell surface, recombinant Gp45 was expressed and purified (Supplementary Fig. S1). Pre-incubation of host cells with increasing concentrations of recombinant Gp45 led to dose-dependent inhibition of  $\phi$ 11 adsorption (Fig. 4a).

We recently demonstrated that *S. aureus* siphoviruses use  $\alpha$ -O-GlcNAc modified WTAs as their adsorption receptor<sup>11</sup>. To examine whether Gp45 binds to WTAs, *S. aureus* wild-type strain RN4220 and mutants with altered WTAs were stained with biotin-labelled recombinant Gp45 and subsequently analysed by flow cytometry. In contrast to the well-stained wild-type *S. aureus* with glycosylated WTA, the mutants  $\Delta tarM/S$ , which lacks  $\alpha$ -O- and  $\beta$ -O-GlcNAc residues on WTA, or  $\Delta tagO$ , which is deficient in WTA, demonstrated drastically decreased background staining (Fig. 4b). These results indicate that Gp45 binds to the cell wall with  $\alpha$ - or  $\beta$ -O-GlcNAc modified WTAs. Unfortunately, recombinant Gp54 purified from *E. coli* was found to be susceptible to degradation and was therefore not suitable for flow cytometry analysis.

**The major components of the  $\phi$ 11 receptor in the cell wall of *S. aureus*.** Previous studies have shown that the entire cell wall of *S. aureus* could inactivate *S. aureus* phages, while the isolated WTAs could not<sup>28,29</sup>. Additionally, treating the cell wall preparations with either muramidase or amidase or using deacetylated cell walls destroyed the phage inactivation capacity of these preparations<sup>30,31</sup>. These observations suggested that peptidoglycan may participate in phage adsorption directly or indirectly by providing rigid support for WTAs. These data prompted us to re-examine phage adsorption with an extended set of *S. aureus* cell wall mutants. In particular, we aimed to investigate how phages interact with WTAs, and how peptidoglycan structures affect their adsorption.

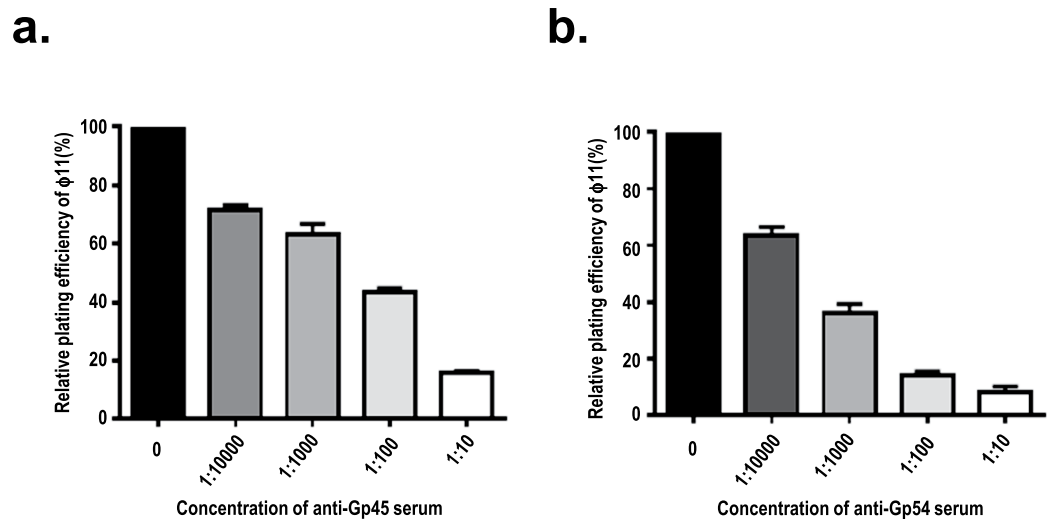
Adsorption assays were carried out using isogenic mutants with altered WTAs as hosts. As shown in Fig. 5a,  $\phi$ 11 virions were able to adsorb to either the  $\Delta tarM$  mutant with only  $\beta$ -GlcNAc residues on WTA or the  $\Delta tarS$



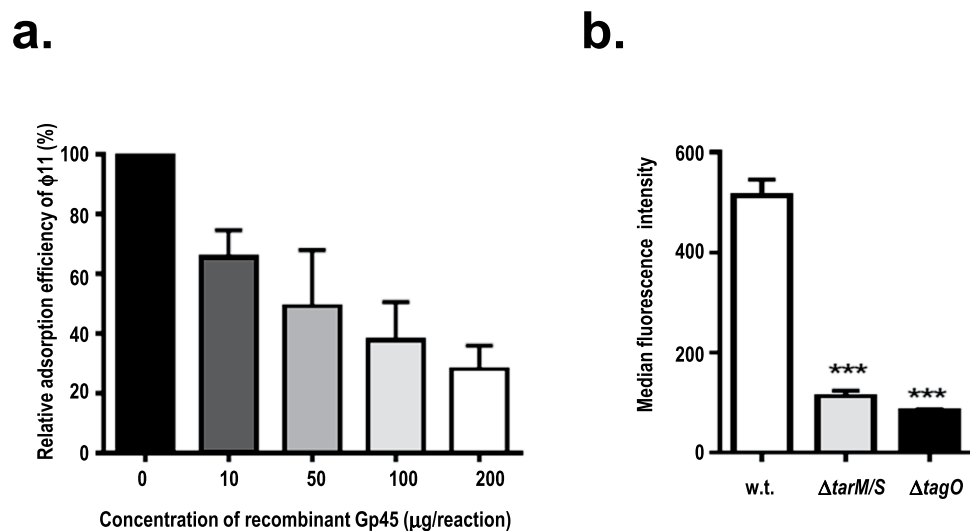
**Figure 2. Immunogold labelling of tail proteins Gp45 and Gp54.** (a,b) Transmission electron microscopy (TEM) images of negatively stained  $\phi 11$  after immunogold labelling with anti-Gp45 serum (a) and anti-Gp54 serum (b), respectively. (c) TEM images of mutant phages  $\Delta gp45$  and  $\Delta gp54$ . (Left), mutant phage  $\Delta gp45$  labelled with anti-Gp45 serum; (Right), mutant phage  $\Delta gp54$  labelled with anti-Gp54 serum. Insets show enlarged views of the boxed areas.

mutant with only  $\alpha$ -GlcNAc residues on WTA with efficiency comparable to that of wild-type cells. In contrast, phage adsorption was significantly impaired when the  $\Delta tagO$  mutant, devoid of WTAs, or the double mutant  $\Delta tarM/S$ , deficient in both  $\alpha$ - and  $\beta$ -GlcNAc residues on WTAs, were used as hosts (Fig. 5a). Consistent with these findings, the cell wall preparation from the wild-type strain dose-dependently inactivated phage, with full inactivation reached at a concentration of 240 nmol phosphate per reaction, whereas the cell wall preparation from double mutant  $\Delta tarM/S$  exhibited significantly less inhibitory effect on plating efficiency at a similar concentration (Fig. 5b). Collectively, these observations demonstrate that GlcNAc residues on WTAs are essential for phage adsorption regardless of their anomeric configurations.

*S. aureus* cell wall preparations with deacetylated peptidoglycan fail to inactivate phage 52A<sup>31</sup>, which is also a serogroup B phage like  $\phi 11$ . To examine if peptidoglycan acetylation is involved in  $\phi 11$  adsorption, the *oatA* mutant<sup>32</sup> deficient in 6-O acetylation of muramic acid residues in peptidoglycan was used as a host for the adsorption assay. As shown in Fig. 5c, phage adsorption efficiency decreased to 50% when compared to the wild-type adsorption, suggesting that peptidoglycan acetylation favours  $\phi 11$  adsorption.



**Figure 3. Neutralization of  $\phi 11$  infection with rabbit anti-Gp45 or anti-Gp54 serum.** Data represent means  $\pm$  standard deviations (SD,  $n = 3$ ). Inhibition of  $\phi 11$  plating efficiency with anti-Gp45 (a) or anti-Gp54 serum (b) were shown respectively.

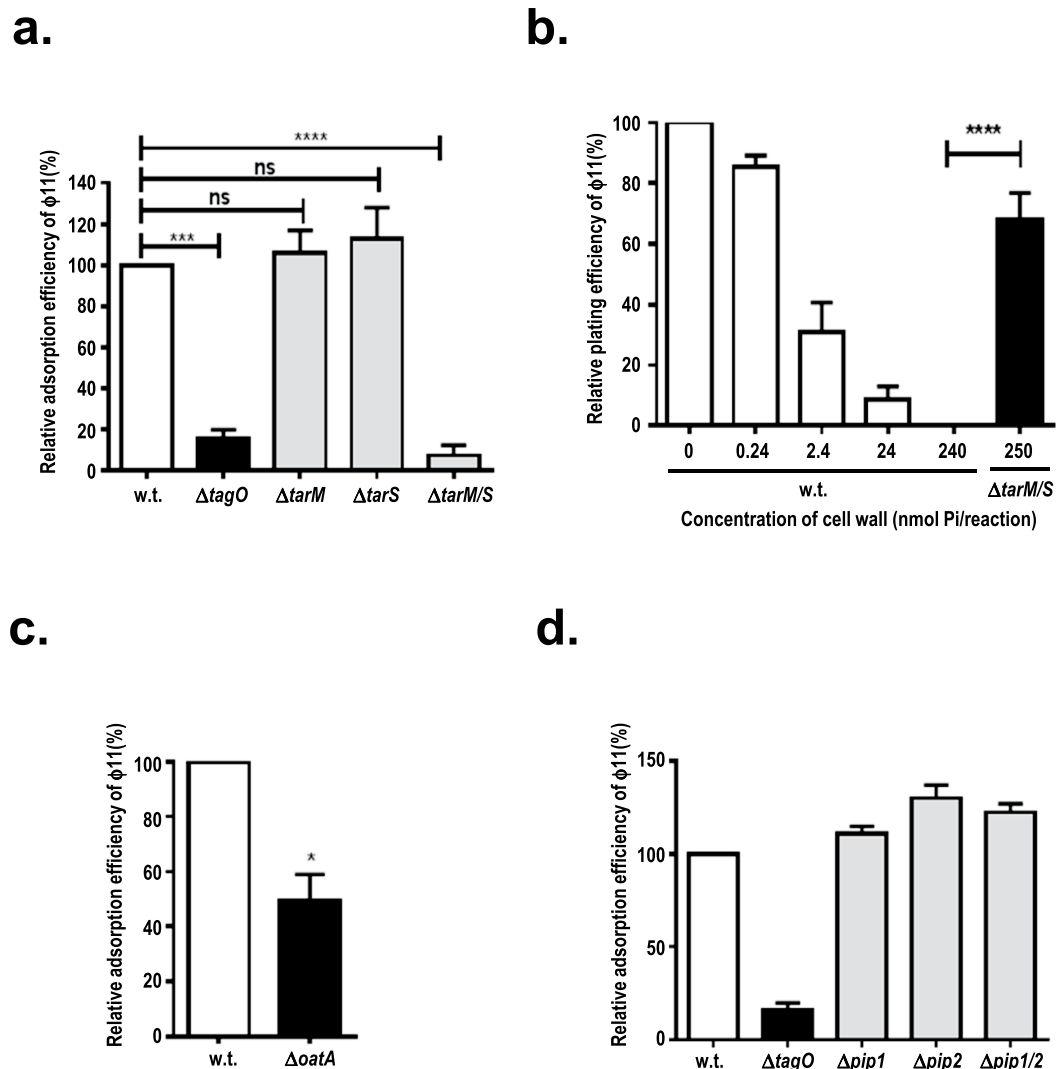


**Figure 4. The RBP (receptor binding protein) activity of Gp45.** (a) Dose-dependent inhibition of  $\phi 11$  adsorption with recombinant Gp45. (b) Flow cytometry analysis of *S. aureus* wild-type cells and mutant cells stained with biotin-labelled Gp45. Wild-type (w.t., white bar),  $\Delta tarM/S$  mutant (grey bar),  $\Delta tagO$  mutant (black bar). Values are given as means  $\pm$  standard deviations (SD,  $n = 3$ ). Statistical significant differences calculated by one way ANOVA with Bonferroni's post-test are indicated: \*\*\* $p < 0.001$ .

**Pip homologues in *S. aureus* do not play a role in  $\phi 11$  adsorption.** Some phages require a membrane-embedded protein receptor for irreversible binding before the translocation of the phage genome into the host cell<sup>8</sup>. Previous studies on phage-resistant mutants derived from *L. lactis* identified the phage infection protein (Pip) as the membrane receptor for lactococcal phage  $c2^{33}$ . YueB, the Pip homologue in *B. subtilis*, was also identified as the membrane receptor for siphophage SPP1<sup>34</sup>. Using the amino acid sequence of YueB or Pip as a probe, two homologues with conserved membrane topology and 40% similarity to YueB were identified from the *S. aureus* genome and designated as Pip1 (SAV2643) and Pip2 (SAV0283), respectively. To determine if these two membrane proteins are involved in  $\phi 11$  adsorption, knockout mutants deficient in *pip1*, *pip2*, or both were generated. Interestingly,  $\phi 11$  plates well on *pip* mutants, and no decrease in  $\phi 11$  adsorption efficiency was observed when these mutants were used as a host (Fig. 5d), suggesting that Pip homologues in *S. aureus* are not involved in phage  $\phi 11$  adsorption.

## Discussion

Research on *S. aureus* phages has a very long history that can be traced back to the early studies of bacteriophages. Since the discovery of bacteriophages, many *S. aureus* phages have been isolated, and these were classified into



**Figure 5. Efficiency of  $\phi 11$  adsorption to *S. aureus* cell wall mutants.** (a) Efficiency of  $\phi 11$  adsorption to mutants with altered WTAs. The *S. aureus* wild-type strain (w.t., white bar) and mutants with altered WTAs were used as host. Mutants  $\Delta tagO$ ,  $\Delta tarM$ ,  $\Delta tarS$ , and  $\Delta tarM/S$  are indicated. (b) Dose-dependent inhibition of  $\phi 11$  plating efficiency with wild-type cell wall but not the cell wall from mutant  $\Delta tarM/S$  deficient in WTA GlcNAc residues. Wild-type cell wall concentrations used in this experiment range from 0 to 240 nmol Pi/reaction (white bars), and the concentration of mutant cell wall  $\Delta tarM/S$  used in this experiment is 250 nmol Pi/reaction, which is indicated with black bar. (c) Efficiency of  $\phi 11$  adsorption to  $\Delta oatA$  mutant deficient in peptidoglycan acetylation. wild-type strain (w.t., white bar);  $\Delta oatA$  mutant (black bar). (d) Efficiency of  $\phi 11$  adsorption to *pip* mutants. wild-type strain (w.t., white bar),  $\Delta tagO$  mutant (black bar), and *pip* mutants (grey bars) are indicated. Values are given as means  $\pm$  standard deviations (SD,  $n = 3$ ). Statistical significant differences calculated by one way ANOVA with Bonferroni's post-test (Fig. 5a,b) or by the two-tailed Student's paired t-test (Fig. 5c) are indicated: not significant (ns); \* $p < 0.05$ ; \*\* $p < 0.001$ ; and \*\*\*\* $p < 0.0001$ .

three families and a few major serogroups<sup>2</sup>. Before molecular techniques became available, *S. aureus* phages had been widely used for typing *S. aureus*. It was known for a long time that many *S. aureus* phages carry virulence genes and are required for *S. aureus* virulence and adaptation<sup>2</sup>. Despite comprehensive studies on phage genomes<sup>15</sup> and the role of *S. aureus* phages in horizontal transfer of resistance and virulence genes among clones and species<sup>12</sup>, the molecular interactions mediating phage adsorption to the staphylococcal cell surface remain poorly understood.

The mechanism underlying *S. aureus* phage adsorption has often been assumed to be similar to that of phages infecting gram-negative bacteria. However, as gram-positive bacteria have a very different cell wall structure compared to that of gram-negative bacteria, phages infecting gram-positive bacteria may employ adsorption mechanisms different from those infecting gram-negative bacteria. Accounting for over 50% of the cell wall mass, WTAs are the most abundant surface molecules in the cell wall of bacteria belonging to the order *Bacillales*, which includes genera such as *Bacillus*, *Listeria* and *Staphylococcus*<sup>7,35</sup>. Hence, it is most likely that phages infecting bacteria of these genera need to interact with WTAs for successful adsorption.

In this study, we demonstrated that GlcNAc residues on WTAs are essential for  $\phi 11$  adsorption regardless of their anomeric configurations. We also found that 6-O-acetylation of muramic acid residues in peptidoglycan is involved in  $\phi 11$  adsorption. We showed that Gp45 and Gp54 are two baseplate proteins critical for  $\phi 11$  infection, as both antisera can neutralize  $\phi 11$  infection dose-dependently. Recombinant Gp45 inhibits  $\phi 11$  adsorption in a dose-dependent manner and binds to glycosylated WTAs, demonstrating that Gp45 is the RBP of  $\phi 11$ . Unfortunately, recombinant Gp54 purified from *E. coli* was not stable and hence unsuitable for cell wall binding studies, and its functions could not be tested.

Staphylococcal pathogenicity islands (SaPIs) have an intimate relationship with temperate staphylococcal phages. Phages can induce the SaPI cycle, which allows the SaPIs to be efficiently encapsidated into special small phage heads commensurate with their size<sup>10</sup>. Previous mutational analyses of the genes present in the morphogenesis cluster of  $\phi 11$  demonstrated that the Gp45 was essential for both the phage infectivity and transduction of its cognate SaPI<sup>13</sup>. Of note,  $\Delta gp54$  seemed to lose its baseplates and failed to plate on *S. aureus*. Surprisingly,  $\Delta gp54$  was still able to transduce SaPIs, although with a 100-fold reduction in transduction efficiency when compared with wild-type  $\phi 11$ . These results highlight that Gp45 is essential for the recognition process, while the presence of the Gp54 significantly increases the binding affinity between the phage and its receptor. As the N-terminus of Gp54 was predicted to be similar to that of BppU, which maintains the attachment of RBP to the baseplate core in TP901-1, it is tempting to speculate that Gp54 plays an important role in anchoring RBP in the baseplate.

Previously, a tail protein ORF636 from a serogroup A phage phiSLT was characterized as an adhesion protein that binds to poly-glycerolphosphate (GroP) chain of lipoteichoic acids (LTAs)<sup>25</sup>. Notably, ORF636 shares high homology with Gp45 (62% similarity) and the ORF636 sequence exists in all known serogroup A phages infecting *S. aureus*. However, it was shown that all tested serogroup A phages can still form plaques on a *S. aureus* mutant deficient in LTAs, but not on a mutant deficient in WTAs<sup>11</sup>, suggesting that WTAs but not LTAs are required for *S. aureus* phage infection. The tight binding of glycerol and glycerolphosphate for the RBPs suggested that LTAs could act as receptors for lactococcal phages<sup>36</sup>, however the structure of LTAs is well conserved and thought to be too simple to explain the different host specificities of various lactococcal phages. Recently, by mutational analysis, it was demonstrated that cell wall polysaccharide (CWPS) is the host cell surface receptor of tested lactococcus phages of different groups and that differences between the CWPS structures play a crucial role in determining phage host range<sup>37</sup>.

It is noteworthy that many phages need a protein receptor for adsorption, for example, Fhu A, OmpA, OmpC, LamB for *E. coli* phages, YueB for *Bacillus* phages and Pip for *Lactococcus* phage<sup>38</sup>. Interestingly, all these protein receptors are non-essential and many of them were identified by transposon mutagenesis. However, by screening a mutant library of *S. aureus* we were unable to isolate  $\phi 11$ -resistant mutants, which carry transposon insertions in genes encoding membrane proteins<sup>39</sup>. It is now generally acknowledged that carbohydrate recognizing phages possess a broad baseplate structure with multiple receptor binding sites. Conversely, phages with stubby ends or tail fibres, including the lactococcal c2 phages and the *Bacillus* phage SPP1, may recognize protein receptors on the cell surface<sup>40</sup>. The crystal structure of Gp45 was solved and it was found that Gp45 forms six trimers in the baseplates of  $\phi 11$  and that each monomer of Gp45 contains a five-bladed propeller domain with a cavity that could accommodate a GlcNAc moiety (Koc *et al.*, unpublished data). Hence, the presence of 18 receptor binding sites in the baseplate of  $\phi 11$  suggests that its receptors are saccharides but not proteins.

Accounting for over 50% of the cell wall mass, WTAs are considered to be the most abundant surface molecules in *S. aureus* and have been implicated in various critical processes and interactions such as staphylococcal cell division, biofilm formation,  $\beta$ -lactam resistance, and staphylococcal pathogenesis<sup>41,42</sup>. Due to the in-homogeneity of WTA, its analysis has proven to be very challenging. Unlike research carried out on DNA, RNA or protein, methods available for studying WTA function are very limited. Despite technical limitations, a few WTA-interacting proteins such as FmtA<sup>43</sup>, WTA antibody, MBL<sup>44</sup>, and SREC-1<sup>45</sup> have recently been identified. Here, we report Gp45 as a new WTA-interacting protein. Our results may eventually provide new tools for labelling and detecting the subdomain structures in the cell wall of *S. aureus*. Additionally, this study establishes a solid basis for the development of sensitive affinity-based infection diagnostics<sup>46</sup> and therapeutics for MRSA infection.

## Materials and Methods

**Bacterial strains and growth conditions.** *S. aureus* strains used in these studies are listed in Table 1. Bacteria were grown at 37 °C in BM broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.1% K<sub>2</sub>HPO<sub>4</sub>, 0.1% glucose) under agitation.

**Construction of *S. aureus* mutants.** The deletion mutants  $\Delta pip1$ ,  $\Delta pip2$ , and  $\Delta pip1/2$  were constructed by allelic exchange. For knock-out plasmid construction, the primers listed in Table S1 in the Supplementary Information were used. For deletion of *pip1*, flanking regions were amplified with primer pairs pip1-F1-up/pip1-F1-dn and pip1-F2-up/pip1-F2-dn. Purified PCR products were digested with Sall/NheI and NheI/EcoRI respectively, and subsequently ligated into the Sall/EcoRI digested knockout vector pBASE6<sup>47</sup>. The resulting plasmid was used for allelic exchange<sup>48</sup>. For the construction of the *pip2* deletion mutant, a similar approach was pursued. The flanking regions of *pip2* were amplified with primer pairs pip2-F1-up/pip2-F1-dn and pip2-F2-up/pip2-F2-dn, digested with XbaI and ligated. Afterwards this marker-less knockout cassette was subcloned into pKOR-1, and the resulting plasmid was used for mutant construction via allelic exchange<sup>48</sup>.

**Overexpression and purification of the recombinant Gp45 and Gp54.** Both *gp45* and *gp54* were amplified by PCR from *S. aureus* strain SA113, which is a  $\phi 11$  lysogen. The primers used for the PCR reaction are



Bacterial strain	Description	Reference
BL21	<i>E. coli</i> BL21, host of inducible recombinant protein expression	Invitrogen
RN4220	<i>S. aureus</i> strain deficient in restriction, capsule, or prophage.	51
$\Delta tagO$	RN4220, $\Delta tagO$	11
$\Delta tarM$	RN4220, $\Delta tarM$	52
$\Delta tarS$	RN4220, $\Delta tarS$	52
$\Delta tarM/S$	RN4220, $\Delta tarM$ , $\Delta tarS$	52
$\Delta pip1$	RN4220, $\Delta pip1$	This study
$\Delta pip2$	RN4220, $\Delta pip2$	This study
$\Delta pip1/2$	RN4220, $\Delta pip1$ , $\Delta pip2$	This study
SA113	Derivative of <i>S. aureus</i> strain NCTC8325 harboring prophages $\phi 11$ , $\phi 12$ , and $\phi 13$	53
$\Delta oatA$	SA113, $\Delta oatA$	32

**Table 1. Bacterial strains used in this study.**

listed in Table S1 in the Supplementary Information. The amplified *gp45* or *gp54* genes were subcloned into the expression vector pET28a between the *NheI* and *XhoI* sites. The resulting plasmids were transformed into *E. coli* BL21 for overexpression of Gp45 or Gp54. Both proteins were fused to a hexa-histidine-tag at the N-terminus to facilitate purification. After IPTG induction of the host cells, recombinant Gp45 was extracted and purified according to the procedure described previously<sup>49</sup>. Briefly, cells were lysed via ultrasonication (Digital Sonifier, Branson). After centrifugation at  $38,000 \times g$  for 55 min, cell debris was removed, and the supernatant containing recombinant Gp45 protein was loaded on a 5 mL Ni-NTA-column (GE Healthcare). Fractions containing Gp45 were pooled and concentrated to 1 mg/mL using Vivaspin 20 centrifugal concentrators with a molecular size cut-off of 50,000 (Sartorius, Göttingen, Germany). The concentrated sample was then loaded on a size-exclusion chromatography column SD200 pre-equilibrated with SEC-buffer containing 25 mM HEPES, 150 mM NaCl, 1 mM DTT. Fractions containing Gp45 were pooled and concentrated as pure Gp45 preparations. The purity and folding of the recombinant Gp45 were assessed with SDS-PAGE, Circular dichroism (CD) spectroscopy and dynamic light scattering (DLS). Gp54 was purified by the same procedure as for Gp45.

**Preparation of cell wall from *S. aureus* strains.** The cell wall was extracted according to the procedure described previously<sup>50</sup>. Briefly, *S. aureus* overnight cultures were harvested by centrifugation at  $5000 \times g$  for 10 minutes. The cells were washed with 20 mM  $NH_4Ac$  buffer (pH 4.8) and re-suspended in the same buffer. After disruption in a cell disrupter (Euler, Frankfurt am Main, Germany), the cell lysates were centrifuged at  $5000 \times g$  to remove the intact cells. The supernatant was collected as a crude extract of cell wall and mixed well with 5 mM  $MgSO_4$ , 40 U/mL DNase and 80 U/mL RNase at final concentrations before overnight incubation at 37 °C. Next, to remove any cell membrane contamination, SDS was added to a final concentration of 2%, followed by ultra-sonication for 15 min. After heating at 65 °C for one hour, the cell wall preparations were washed six times with 20 mM  $NH_4Ac$  buffer by centrifugation at  $12,000 \times g$ . Finally, the cell wall preparations were re-suspended in distilled water and quantified by measuring the amount of inorganic phosphate using the QuantiChrom™ Phosphate Assay Kit (BioAssay Systems, USA) as described previously<sup>39</sup>.

**Bacteriophage experiments.** Using the double layer soft agar method,  $\phi 11$  was propagated with the indicator strain, *S. aureus* strain RN4220, as a host.

Phage plating efficiencies were determined to investigate the effects of Gp45, Gp54 anti-sera and cell wall preparations on the inactivation of  $\phi 11$ . In brief, 100  $\mu L$  of  $\phi 11$  ( $3 \times 10^6$  PFU/mL) was mixed with 100  $\mu L$  of cell wall preparations or antisera of certain concentrations and incubated at 37 °C for 10 min. Samples pre-incubated without any cell wall preparations or sera served as controls. Next, the mixtures were diluted before plating on the indicator strain (*S. aureus* strain RN4220) using double agar overlay methods. After overnight incubation at 37 °C, the plaques were enumerated. The efficiency of plating was calculated relative to that of plating of  $\phi 11$  pre-incubated without any sera or cell wall preparations.

Adsorption assays were performed according to the procedure described previously<sup>11</sup>. Briefly, 200  $\mu L$  of *S. aureus* wild-type or mutant cells containing  $8 \times 10^7$  CFU were mixed with 100  $\mu L$  of  $\phi 11$  containing  $3 \times 10^5$  PFU and incubated at 37 °C for 15 min. The bound phages were separated from the free phages by centrifugation at  $13,000 \times g$  for 5 min. Adsorption was calculated by determining the number of PFU of the unbound phage in the supernatant and subtracting it from the total number of input PFU. Adsorption efficiency was expressed relative to the adsorption of wild-type strain RN4220. Each adsorption assay was repeated at least three times. To study the inhibition of adsorption by Gp45, cells were pre-incubated with the purified recombinant Gp45 of indicated concentrations for 15 min before adding phages to the host cells.

**Purification of  $\phi 11$  and electron microscopy methods.** Phage  $\phi 11$  lysate was centrifuged at  $73000 \times g$ , 4 °C for two hours (Beckman Optima XL-80K). The resulting pellet was re-suspended in 500  $\mu L$  of TMN buffer

containing 10 mM Tris-HCl, pH 7.5, 10 mM MgSO<sub>4</sub>, 500 mM NaCl. The sample was then mixed well with 55% CsCl in TMN-buffer to give a final concentration of 42% CsCl and subjected to ultracentrifugation at 245,000 × g, 15°C for 20 hours (Beckman). The visible phage band on the CsCl gradient was collected and sequentially dialyzed for two hours each in a D-Tube Dialyzer Mini (Novagen®, Merck Millipore, Darmstadt, Germany) against decreasing concentrations of NaCl in TMN buffer (10 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 4 M NaCl) until the NaCl concentration after each round of dialysis was at 4 M, 2 M, 1 M and 10 mM NaCl, respectively.

For immunogold labelling, purified phage samples were adsorbed to glow discharged, pioloform and carbon-coated grids. The grids were then blocked with 0.2% gelatin in phosphate-buffered saline for 10 min followed by incubation with rabbit anti-Gp45 or rabbit anti-Gp54 serum, which were diluted in blocking buffer at 1:20 and 1:100, respectively. Polyclonal rabbit antisera were raised against purified recombinant Gp45 or Gp54 using a custom antibody service, Speedy 28-Day polyclonal program from Eurogentec (Brussels, Belgium). After blocking at room temperature for 60 min, the grids were washed six times with blocking buffer for a total time of 15 min before incubation with goat anti-rabbit IgG coupled with 12 nm gold colloids (Dianova, Hamburg), which was diluted with blocking buffer at 1:30. After incubation at room temperature for 60 min, the grids were washed three times with blocking buffer for 10 min and three times with phosphate-buffered saline for 10 min, followed by washing four times with double-distilled water for 2 min. Finally, the grids were negatively stained with 1% (w/v) aqueous uranyl acetate before examination with a JEM-1400Plus transmission electron microscope (JEOL, Japan)

**Flow cytometry analysis.** Flow cytometry was carried out to evaluate the binding of recombinant Gp45 to the *S. aureus* cell surface. Purified recombinant Gp45 was labelled with biotin using the EZ-Link™ NHS-Biotin kit (Thermo Fisher Scientific). Biotin-labelled Gp45 was then incubated with *S. aureus* wild-type or mutant cells for 30 min with shaking at room temperature. Cells were washed and stained with strep-Alu488 (Invitrogen) for one hour at 4°C. Finally, cells were fixed for flow cytometry analysis.

**Statistical analysis.** Results are expressed as the means ± standard deviations from at least three independent experiments. Statistical analysis was performed using GraphPad Prism (GraphPad Software, Inc., La Jolla, USA, Version 5.04). Statistically significant differences were calculated with two-tailed Student's t-test or one-way ANOVA with Bonferroni's post-test as indicated.

## References

- Lindsay, J. A. Staphylococcus aureus genomics and the impact of horizontal gene transfer. *Int J Med Microbiol* **304**, 103–109, doi: 10.1016/j.ijmm.2013.11.010 (2014).
- Xia, G. & Wolz, C. Phages of Staphylococcus aureus and their impact on host evolution. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases* **21**, 593–601, doi: 10.1016/j.meegid.2013.04.022 (2014).
- Hagens, S. & Loessner, M. J. Bacteriophage for biocontrol of foodborne pathogens: calculations and considerations. *Current pharmaceutical biotechnology* **11**, 58–68 (2010).
- Fischetti, V. A., Nelson, D. & Schuch, R. Reinventing phage therapy: are the parts greater than the sum? *Nature biotechnology* **24**, 1508–1511, doi: 10.1038/nbt1206-1508 (2006).
- Duplessis, M. & Moineau, S. Identification of a genetic determinant responsible for host specificity in Streptococcus thermophilus bacteriophages. *Molecular microbiology* **41**, 325–336 (2001).
- Spinelli, S., Veessler, D., Bebeacua, C. & Cambillau, C. Structures and host-adhesion mechanisms of lactococcal siphophages. *Frontiers in microbiology* **5**, 3, doi: 10.3389/fmicb.2014.00003 (2014).
- Xia, G., Kohler, T. & Peschel, A. The wall teichoic acid and lipoteichoic acid polymers of Staphylococcus aureus. *International journal of medical microbiology: IJMM* **300**, 148–154, doi: 10.1016/j.ijmm.2009.10.001 (2010).
- Baptista, C., Santos, M. A. & Sao-Jose, C. Phage SPP1 reversible adsorption to Bacillus subtilis cell wall teichoic acids accelerates virus recognition of membrane receptor YueB. *Journal of bacteriology* **190**, 4989–4996, doi: 10.1128/jb.00349-08 (2008).
- Sao-Jose, C. *et al.* The ectodomain of the viral receptor YueB forms a fiber that triggers ejection of bacteriophage SPP1 DNA. *J Biol Chem* **281**, 11464–11470, doi: 10.1074/jbc.M513625200 (2006).
- Penades, J. R. & Christie, G. E. The phage-inducible chromosomal islands: a family of highly evolved molecular parasites. *Annual Review of Virology* **2**, 181–201, doi: 10.1146/annurev-virology-031413-085446 (2015).
- Xia, G. *et al.* Wall teichoic Acid-dependent adsorption of staphylococcal siphovirus and myovirus. *Journal of bacteriology* **193**, 4006–4009, doi: 10.1128/jb.01412-10 (2011).
- Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nature communications* **4**, 2345, doi: 10.1038/ncomms3345 (2013).
- Tormo, M. A. *et al.* Staphylococcus aureus pathogenicity island DNA is packaged in particles composed of phage proteins. *Journal of bacteriology* **190**, 2434–2440, doi: 10.1128/jb.01349-07 (2008).
- Tallent, S. M., Langston, T. B., Moran, R. G. & Christie, G. E. Transducing particles of Staphylococcus aureus pathogenicity island SaPI1 are comprised of helper phage-encoded proteins. *Journal of bacteriology* **189**, 7520–7524, doi: 10.1128/jb.00738-07 (2007).
- Kwan, T., Liu, J., DuBow, M., Gros, P. & Pelletier, J. The complete genomes and proteomes of 27 Staphylococcus aureus bacteriophages. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 5174–5179, doi: 10.1073/pnas.0501140102 (2005).
- Quiles-Puchalt, N., Martinez-Rubio, R., Ram, G., Lasa, I. & Penades, J. R. Unravelling bacteriophage varphi11 requirements for packaging and transfer of mobile genetic elements in Staphylococcus aureus. *Molecular microbiology* **91**, 423–437, doi: 10.1111/mmi.12445 (2014).
- Iandolo, J. J. *et al.* Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of Staphylococcus aureus 8325. *Gene* **289**, 109–118 (2002).
- Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244–248, doi: 10.1093/nar/gki408 (2005).
- Veessler, D. *et al.* Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): a baseplate hub paradigm in gram-positive infecting phages. *J Biol Chem* **285**, 36666–36673, doi: 10.1074/jbc.M110.157529 (2010).
- Veessler, D. & Cambillau, C. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiology and molecular biology reviews: MMBR* **75**, 423–433, first page of table of contents, doi: 10.1128/mmb.00014-11 (2011).
- Flayhan, A. *et al.* Crystal Structure of pb9, the Distal Tail Protein of Bacteriophage T5: a Conserved Structural Motif among All Siphophages. *J Virol* **88**, 820–828, doi: 10.1128/JVI.02135-13 (2014).

22. Kondou, Y. *et al.* Structure of the central hub of bacteriophage Mu baseplate determined by X-ray crystallography of gp44. *Journal of molecular biology* **352**, 976–985, doi: 10.1016/j.jmb.2005.07.044 (2005).
23. Kanamaru, S. *et al.* Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553–557 (2002).
24. Rodriguez-Rubio, L. *et al.* The peptidoglycan hydrolase of Staphylococcus aureus bacteriophage 11 plays a structural role in the viral particle. *Applied and environmental microbiology* **79**, 6187–6190, doi: 10.1128/aem.01388-13 (2013).
25. Kaneko, J., Narita-Yamada, S., Wakabayashi, Y. & Kamio, Y. Identification of ORF636 in phage phiSLT carrying Pantone-Valentine leukocidin genes, acting as an adhesion protein for a poly(glycerophosphate) chain of lipoteichoic acid on the cell surface of Staphylococcus aureus. *Journal of bacteriology* **191**, 4674–4680, doi: 10.1128/jb.01793-08 (2009).
26. Cole, C., Barber, J. D. & Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197–201, doi: 10.1093/nar/gkn238 (2008).
27. Veesler, D. *et al.* Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8954–8958, doi: 10.1073/pnas.1200966109 (2012).
28. Coyette, J. & Ghuyssen, J. M. Structure of the cell wall of Staphylococcus aureus, strain Copenhagen. IX. Teichoic acid and phage adsorption. *Biochemistry* **7**, 2385–2389 (1968).
29. Lindberg, A. A. Bacteriophage receptors. *Annual review of microbiology* **27**, 205–241, doi: 10.1146/annurev.mi.27.100173.001225 (1973).
30. Murayama, Y., Kotani, S. & Kato, K. Solubilization of phage receptor substances from cell walls of Staphylococcus aureus (strain Copenhagen) by cell wall lytic enzymes. *Biken journal* **11**, 269–291 (1968).
31. Shaw, D. R. & Chatterjee, A. N. O-Acetyl groups as a component of the bacteriophage receptor on Staphylococcus aureus cell walls. *Journal of bacteriology* **108**, 584–585 (1971).
32. Bera, A., Herbert, S., Jakob, A., Vollmer, W. & Gotz, F. Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of Staphylococcus aureus. *Molecular microbiology* **55**, 778–787, doi: 10.1111/j.1365-2958.2004.04446.x (2005).
33. Geller, B. L., Ivey, R. G., Trempy, J. E. & Hettlinger-Smith, B. Cloning of a chromosomal gene required for phage infection of Lactococcus lactis subsp. lactis C2. *J Bacteriol* **175**, 5510–5519 (1993).
34. Sao-Jose, C., Baptista, C. & Santos, M. A. Bacillus subtilis operon encoding a membrane receptor for bacteriophage SPP1. *Journal of bacteriology* **186**, 8337–8346, doi: 10.1128/jb.186.24.8337-8346.2004 (2004).
35. Brown, S., Santa Maria, J. P., Jr. & Walker, S. Wall teichoic acids of gram-positive bacteria. *Annual review of microbiology* **67**, 313–336, doi: 10.1146/annurev-micro-092412-155620 (2013).
36. Spinelli, S. *et al.* Modular structure of the receptor binding proteins of Lactococcus lactis phages. The RBP structure of the temperate phage TP901-1. *J Biol Chem* **281**, 14256–14262, doi: 10.1074/jbc.M600666200 (2006).
37. Ainsworth, S. *et al.* Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. *mBio* **5**, e00880–00814, doi: 10.1128/mBio.00880-14 (2014).
38. Bertozzi Silva, J., Storms, Z. & Sauvageau, D. Host receptors for bacteriophage adsorption. *FEMS microbiology letters* **363**, doi: 10.1093/femsle/fnw002 (2016).
39. Xia, G. *et al.* Glycosylation of wall teichoic acid in Staphylococcus aureus by TarM. *J Biol Chem* **285**, 13405–13415, doi: 10.1074/jbc.M109.096172 (2010).
40. Mahony, J., McDonnell, B., Casey, E. & van Sinderen, D. Phage-Host Interactions of Cheese-Making Lactic Acid Bacteria. *Annual review of food science and technology*, doi: 10.1146/annurev-food-041715-033322 (2016).
41. Winstel, V., Xia, G. & Peschel, A. Pathways and roles of wall teichoic acid glycosylation in Staphylococcus aureus. *International journal of medical microbiology: IJMM* **304**, 215–221, doi: 10.1016/j.ijmm.2013.10.009 (2014).
42. Misawa, Y. *et al.* Staphylococcus aureus Colonization of the Mouse Gastrointestinal Tract Is Modulated by Wall Teichoic Acid, Capsule, and Surface Proteins. *Plos pathogens* **11**, e1005061, doi: 10.1371/journal.ppat.1005061 (2015).
43. Qamar, A. & Golemi-Kotra, D. Dual roles of FmtA in Staphylococcus aureus cell wall biosynthesis and autolysis. *Antimicrobial agents and chemotherapy* **56**, 3797–3805, doi: 10.1128/aac.00187-12 (2012).
44. Kurokawa, K. *et al.* Glycoepitopes of staphylococcal wall teichoic acid govern complement-mediated opsonophagocytosis via human serum antibody and mannose-binding lectin. *J Biol Chem* **288**, 30956–30968, doi: 10.1074/jbc.M113.509893 (2013).
45. Baur, S. *et al.* A nasal epithelial receptor for Staphylococcus aureus WTA governs adhesion to epithelial cells and modulates nasal colonization. *Plos pathogens* **10**, e1004089, doi: 10.1371/journal.ppat.1004089 (2014).
46. Schmelcher, M. & Loessner, M. J. Application of bacteriophages for detection of foodborne pathogens. *Bacteriophage* **4**, e28137, doi: 10.4161/bact.28137 (2014).
47. Geiger, T. *et al.* The stringent response of Staphylococcus aureus and its impact on survival after phagocytosis through the induction of intracellular PSMs expression. *Plos pathogens* **8**, e1003016, doi: 10.1371/journal.ppat.1003016 (2012).
48. Bae, T. & Schneewind, O. Allelic replacement in Staphylococcus aureus with inducible counter-selection. *Plasmid* **55**, 58–63, doi: 10.1016/j.plasmid.2005.05.005 (2006).
49. Koc, C. *et al.* Structural and enzymatic analysis of TarM glycosyltransferase from Staphylococcus aureus reveals an oligomeric protein specific for the glycosylation of wall teichoic acid. *J Biol Chem* **290**, 9874–9885, doi: 10.1074/jbc.M114.619924 (2015).
50. Weidenmaier, C. *et al.* Role of teichoic acids in Staphylococcus aureus nasal colonization, a major risk factor in nosocomial infections. *Nature medicine* **10**, 243–245, doi: 10.1038/nm991 (2004).
51. Kreiswirth, B. N. *et al.* The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* **305**, 709–712 (1983).
52. Brown, S. *et al.* Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18909–18914, doi: 10.1073/pnas.1209126109 (2012).
53. Iordanescu, S. & Surdeanu, M. Two restriction and modification systems in Staphylococcus aureus NCTC8325. *Journal of general microbiology* **96**, 277–281, doi: 10.1099/00221287-96-2-277 (1976).

## Acknowledgements

This work was supported by SFB766 to T.S., A.P. and G.X. from the German Research Foundation (DFG).

## Author Contributions

X.L. and G.X. designed this study; X.L., C.K., P.K., Y.S. and B.K. performed the experiments; X.L., Y.S., M.E., J.P., C.W., T.S., C.C., A.P. and G.X. analysed the data; C.C., M.E. and G.X. wrote the manuscript. Every author reviewed the manuscript prior to submission.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Li, X. *et al.* An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*. *Sci. Rep.* **6**, 26455; doi: 10.1038/srep26455 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

# An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*

## Supplementary information:

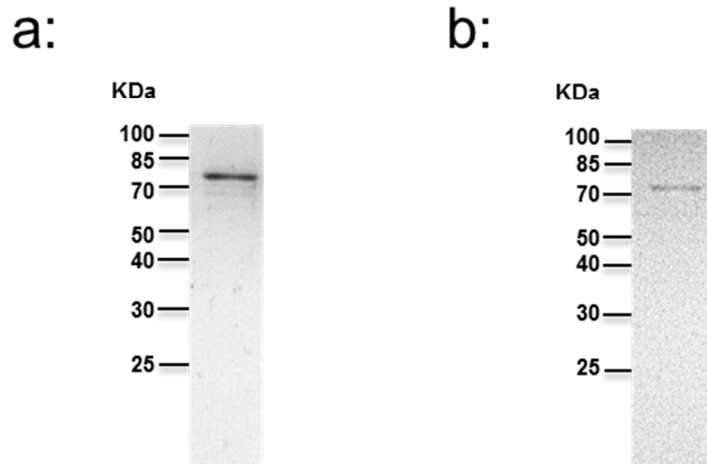
Xuehua Li<sup>1</sup>, Cengiz Koç<sup>2</sup>, Petra Kühner<sup>1</sup>, York-Dieter Stierhof<sup>3</sup>, Bernhard Krismer<sup>1</sup>, Mark C Enright<sup>4</sup>, José R Penadés<sup>5</sup>, Christiane Wolz<sup>1</sup>, Thilo Stehle<sup>2,6,8</sup>, Christian Cambillau<sup>7</sup>, Andreas Peschel<sup>1,8</sup>, Guoqing Xia<sup>1,8,9,\*</sup>

1. Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, 72076 Tübingen Germany,
2. Interfaculty Institute of Biochemistry, University of Tübingen, 72076 Tübingen, Germany
3. Center for Plant Molecular Biology, University of Tübingen, 72076 Tübingen, Germany
4. School of Healthcare Sciences, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, United Kingdom
5. Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
6. Vanderbilt University, School of Medicine, Nashville, TN 37232, USA
7. Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique, UMR 6098, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France
8. German Center for Infection Research (DZIF), partner site Tübingen, Germany
9. Institute of Inflammation & Repair, Faculty of Medical and Human Sciences, University of Manchester, Oxford Road, Manchester, M13 9PT, United Kingdom

\* Correspondence to Dr. Guoqing Xia, e-mail: [guoqing.xia@manchester.ac.uk](mailto:guoqing.xia@manchester.ac.uk)


**Supplementary Table S1: Primers used in this study.**

Primer	Sequence
gp45-up (NheI)	CACGGCTAGCATGAGTAATAAACTAATTACAGATTTAAGTAG
gp45-dn (XhoI)	CAGTCTCGAGTTATTCAACCACCTTTCCTTCGAAT
gp54-up	CACGGCTAGCATGGTAGTAGATAATTTTTCGAAAGATGA
gp54-dn	CAGTCTCGAGTTATTTCCACCCATAATTTTAGTTATAGT
pip1-F1-up	CTAGGATCTGTGTCGACTAATGATGTTG
pip1-F1-dn	GGAAGCTAGCTAGTCGATTGATACCTATC
pip1-F2-up	CCAAAGTAACGCTAGCGAATAAAAATT
pip1-F2-dn	CAGAATTCACATTTAAAAATACTTCAGCG
pip2-F1-up	GGGACAAGTTTGTACAAAAAAGCAGGCTCGCGCCGGTGTCTTTATTC ACTTTG
pip2-F1-dn	GTCATCTAGAACTTTGGGATCCTTTTTCTGATTGATCTCC
pip2-F2-up	GTCATCTAGAGTATGTTCTTAATGATTGCTGCATTGGGTC
pip2-F2-dn	GGGACCACTTTGTACAAGAAAGCTGGGTGAAACTGTCACGTAGTTCA GTTAATTCAGC



**Supplementary Figure S1:** Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the purified Gp45 (a) and Gp54 (b) preparation from *E. coli*. The gel was stained with Coomassie brilliant blue R-250.

# SCIENTIFIC REPORTS



OPEN

## Structure of the host-recognition device of *Staphylococcus aureus* phage $\phi$ 11

Received: 09 December 2015

Accepted: 17 May 2016

Published: 10 June 2016

Cengiz Koç<sup>1</sup>, Guoqing Xia<sup>2,3,4</sup>, Petra Kühner<sup>2</sup>, Silvia Spinelli<sup>5,6</sup>, Alain Roussel<sup>5,6</sup>, Christian Cambillau<sup>5,6</sup> & Thilo Stehle<sup>1,3,7</sup>

Phages play key roles in the pathogenicity and adaptation of the human pathogen *Staphylococcus aureus*. However, little is known about the molecular recognition events that mediate phage adsorption to the surface of *S. aureus*. The lysogenic siphophage  $\phi$ 11 infects *S. aureus* SA113. It was shown previously that  $\phi$ 11 requires  $\alpha$ - or  $\beta$ -N-acetylglucosamine (GlcNAc) moieties on cell wall teichoic acid (WTA) for adsorption. Gp45 was identified as the receptor binding protein (RBP) involved in this process and GlcNAc residues on WTA were found to be the key component of the  $\phi$ 11 receptor. Here we report the crystal structure of the RBP of  $\phi$ 11, which assembles into a large, multidomain homotrimer. Each monomer contains a five-bladed propeller domain with a cavity that could accommodate a GlcNAc moiety. An electron microscopy reconstruction of the  $\phi$ 11 host adhesion component, the baseplate, reveals that six RBP trimers are assembled around the baseplate core. The Gp45 and baseplate structures provide insights into the overall organization and molecular recognition process of the phage  $\phi$ 11 tail. This assembly is conserved among most glycan-recognizing *Siphoviridae*, and the RBP orientation would allow host adhesion and infection without an activation step.

*Staphylococcus aureus* is a Gram-positive bacterium that causes a wide range of infections. It is a leading cause of bacteremia, infective endocarditis, as well as osteoarticular, skin and soft tissue, pleuropulmonary and device related infections<sup>1</sup>. Methicillin-resistant *S. aureus* (MRSA) remains a severe global problem threatening the health care system as resistance restricts treatment options to a few drugs of last resort.

All *S. aureus* genomes sequenced to date contain one or several prophages<sup>2,3</sup>. Most *S. aureus* phages can be integrated into the bacterial chromosome or exist as extra-chromosomal elements. It is known that many of these phages encode a large variety of *S. aureus* virulence or fitness factors that allow the bacterium to escape the host immune system. Among all mobile genetic elements in *S. aureus*, phages are probably most efficient in mediating horizontal gene transfer of virulence or resistance genes between strains, and across species or even genus. Therefore, phages play important roles in staphylococcal pathogenicity and adaptation of *S. aureus* to different hostile environments<sup>2,3</sup>.

The large number of staphylococcal phages sequenced to date display an extensive mosaicism in their gene structure, which is a consequence of gene shuffling among different phages that can infect staphylococcal species. The resulting mosaic gene organization is consistent with a modular evolution involving exchanges of genome modules by horizontal transfer and genetic recombination. The genetic exchanges of modules can involve single genes, protein domains, groups of genes, or even functional modules<sup>3</sup>.

Although phages are the most abundant and diversified biological entity on earth, each phage can only infect a limited number of bacterial strains. This specific phage-host interaction is determined, in part, by the protein recognition device located at the tip of the phage tail, which engages a receptor at the bacterial cell surface. Since

<sup>1</sup>Interfaculty Institute of Biochemistry, University of Tübingen, 72076 Tübingen, Germany. <sup>2</sup>Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, 72076 Tübingen, Germany. <sup>3</sup>German Center for Infection Research (DZIF), partner site Tübingen, Germany. <sup>4</sup>Institute of Inflammation and Repair, Faculty of Medical and Human Sciences, University of Manchester, Manchester, United Kingdom. <sup>5</sup>Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, UMR 7257 13288 Marseille Cedex 09, France. <sup>6</sup>Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique, UMR 6098, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France. <sup>7</sup>Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. Correspondence and requests for materials should be addressed to C.C. (email: ccambillau@gmail.com) or T.S. (email: thilo.stehle@uni-tuebingen.de)

bacterial cell wall polysaccharides or glycopolymers project from the cell surface and are thus easily accessible, they are the most common molecules targeted by bacteriophages<sup>4–6</sup>.

The cell wall of *S. aureus* typically contains poly-ribitol phosphate type wall teichoic acid (WTA), which is modified with D-alanine and N-acetyl-glucosamine (GlcNAc). *S. aureus*  $\phi$ 11 is often used as model to study horizontal gene transfer of virulence genes<sup>7</sup>. Recently it was shown that  $\phi$ 11 requires GlcNAc residues on WTA for adsorption<sup>8</sup>. Gp45 of  $\phi$ 11 was identified and characterized as the receptor-binding protein (RBP) of  $\phi$ 11<sup>9</sup>. Furthermore, it was shown that  $\phi$ 11 was unable to bind to the cell wall in the absence of WTA–GlcNAc, identifying glycosylated WTA as the receptor.

Phages adopt a two-fold strategy for host adhesion. They first deploy adhesion modules on fibers or on capsid or tail that recognize the host's cell wall glycan structures in a reversible way: this allows cell wall scanning in search for the final, specific receptor, to which they bind irreversibly<sup>10–12</sup>. This final receptor can be a protein, generally membrane embedded<sup>12,13</sup>, or a cell wall polysaccharide as observed in the case of lactococcal phages<sup>4–6</sup>. Attachment to membrane protein often requires a unique and strong attachment of the phage's tail tip, as observed in phage T5<sup>14,15</sup>. In contrast, the loose affinity observed between saccharides and proteins requires the presence of several attachment sites provided by a multimeric RBP carrying device, the baseplate<sup>16–19</sup>. In order to provide a foundation for understanding the initial recognition mechanism of phage  $\phi$ 11 and its receptor in the cell wall of *S. aureus*, we embarked on structural analyses of the Gp45 and baseplate of  $\phi$ 11 using X-ray crystallography and electron microscopy. The RBP structure reveals a trimer with a complex fold that can be divided, from N- to C-terminus, into a “stem”, a “platform” and a “tower”. The stem is formed by a long, severely bent triple  $\alpha$ -helical coiled coil that features three interruptions: the first and third interruptions are both  $\beta$ -hairpin structures and the second is a short disordered region. A putative “hinge”-like feature is located between the second and the third interruption. The stem is followed by a “platform” of three  $\beta$ -propellers, and the protein terminates with a “tower” formed by a repetitive all- $\beta$  domain. Platform and tower are interconnected by a fifth short triple helix buried inside of the protein on the molecule's longest three-fold axis. An unusual iron is located at the C-terminal end of the first coiled coil and may play a role in mediating flexibility or conformational rearrangements within the helical domain. Six copies of the trimer assemble around the baseplate core. This hexameric organization is commonly observed in lactococcal *Siphoviridae*<sup>17,18</sup>, and it is compatible with host adhesion and infection in the absence of an activation step.

## Results

**Structure determination.** The *gp45* gene was cloned into the pET28 vector (Novagen) for overexpression in *E. coli* as described elsewhere<sup>9</sup>. Briefly, the protein was expressed with an N-terminal hexa-histidine-tag, and purified by nickel-affinity chromatography and size exclusion chromatography as a trimer. Structure determination was performed with a Ta<sub>6</sub>Br<sub>12</sub> derivative using single isomorphous replacement with anomalous scattering (SIRAS) and exploiting the non-crystallographic symmetry (NCS) present in the crystals. Initial refinement with PHENIX<sup>20</sup> was followed by several runs with autoBUSTER<sup>21</sup>, and alternating refinement and model building<sup>22</sup> cycles resulted in excellent R<sub>free</sub> and R<sub>work</sub> values of 21.1% and 17.5%, respectively, for the final model (Table 1). Although the map is generally of good quality, a few loops of the propeller domains have very weak electron density, explaining the persistence of a small number of outliers in the Ramachandran plot (0.5%). Of the remaining residues, 94.5% are located in regions of preferred conformation and 5% in regions that are classified as allowed.

**Overall structure of  $\phi$ 11 RBP.** The RBP of  $\phi$ 11 assembles into an elongated homotrimer, with overall dimensions of approximately 160 × 120 × 100 Å (Fig. 1A). The structure can be divided into an N-terminal “stem” region that forms a triple-helical bundle (Figs 1B and S1), a central “platform” region composed of three  $\beta$ -propeller domains (Figs 1D and S1) and a C-terminal “tower” region (Figs 1E and S1). Overall, the stem contains three non-helical interruptions. The first of these occurs between residues 46 and 67 and contains a bound iron as well as a  $\beta$ -hairpin that faces away from the bundle axis (Figs 1C and 2). The second and the third interruptions are located between residues 81 and 107, which introduce a sharp kink into the stem (the “hinge”) and thus break the shared three-fold symmetry of  $\alpha$ 1 and  $\alpha$ 2 (Figs 1 and 3). Helix  $\alpha$ 3 (residues 88–97), a short triple helical coil located in the hinge, has an independent rotation axis not aligned to the remainders of the molecule (Fig. 3). This helical bundle is followed by five-bladed  $\beta$ -propeller modules of the platform, which encompass residues 142–439 and form the midsection of the protein (Fig. 1). This “platform” is linked via a short helix (residues 425–432) to the C-terminal “tower”. The latter contains two structurally similar domains (residues 440–541 and 542–636), which are each formed by three five-stranded anti-parallel  $\beta$ -sheets, one from each monomer, that are covered on their surface-exposed side by loops and one short  $\alpha$ -helix each (Fig. 1).

Each  $\phi$ 11 RBP monomer forms extensive contacts with the two others monomers in the trimer. For each contact, 2 × 5,800 Å<sup>2</sup> (11,600 Å<sup>2</sup>) are buried in the interaction as calculated by PISA<sup>23</sup>. This results in a total buried surface of ~35,000 Å<sup>2</sup> for the trimer. Most of the buried surface area is concentrated in the stem and C-terminal regions, while the propeller domains engage in few intermolecular contacts (Fig. S2).

**The stem structure.** The stem comprises three separate triple-helical bundles, which are composed of helices  $\alpha$ 1,  $\alpha$ 2 and  $\alpha$ 3/ $\alpha$ 4, respectively (Fig. 1B). The helices pack tightly together in each of the bundles, and almost every residue of each monomer is in contact with a residue of one of the two other monomers (Fig. S2) through central hydrophobic contacts or lateral hydrogen or ionic bonds. The trimeric ensemble comprising the extended N-terminus and helix  $\alpha$ 1 can be superposed onto the 30 first residues of phage TP901-1 RBP (PDB code 3U6X) with an r.m.s.d. value of 1.8 Å for 90 C $\alpha$  atoms (Fig. 1B). The remaining helical bundles most closely resemble those found in phage TP901-1 Baseplate protein Upper (BppU) and other viral trimeric helix bundles<sup>18,19</sup>.

A strong electron density feature suggesting the presence of a metal ion was observed at the junction between the first two bundles. Using an RBP crystal and extended X-ray absorption fine structure (EXAFS) spectroscopy,

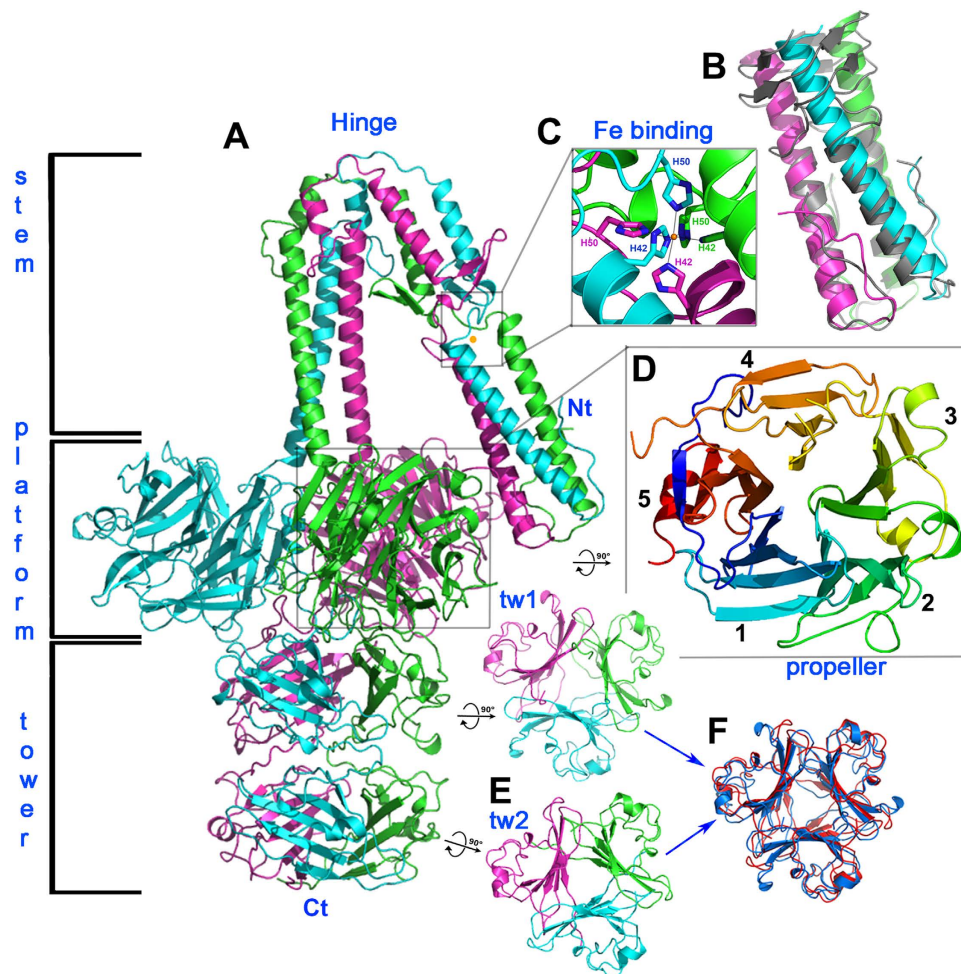


	Derivative Ta6Br12	Native
Data collection		
Space group	P1	P1
Cell dimensions		
a, b, c (Å)	87.65, 89.60, 93.73	87.06, 89.08, 93.3
$\alpha, \beta, \gamma$ (°)	92.7, 105.7, 117.9	93.0, 105.2, 117.6
Resolution (Å)	50–3.30 (3.37–3.30)	44.4–2.20 (2.38–2.20)
R <sub>meas</sub>	12.5 (100.4)	6.6 (57.5)
CC (1/2) (%)	99.9 (87.1)	99.8 (83.5)
I/ $\sigma$ I	20.4 (2.9)	12.6 (2.4)
Completeness (%)	99.3 (95.0)	97.1 (97.1)
Redundancy	13.7 (12.0)	3.2 (3.2)
Phasing		
Sites	12	
Anomalous phasing power	1.4 (0.143)	
Figure of merit - acentric	0.26 (0.05)	
Refinement		
Resolution (Å)		44.4–2.20 (2.26–2.20)
No. reflections		116,564 (8497)
R <sub>work</sub> /R <sub>free</sub> (%)		17.5/21.1 (21.7/25.1)
No. atoms		
Protein		15324
Ligand/ion		1
Water		1446
B-factors		
Protein		51.1
Ligand/ion		33.6
Water		54.2
R.m.s deviations		
Bond lengths (Å)		0.010
Bond angles (°)		1.12

**Table 1. X-ray data collection, phasing and refinement statistics of Gp45.**

the identity of this ion was determined to be iron, which probably exists in its oxidized form Fe<sup>3+</sup> (analyzed at SOLEIL beamline PX1) (See Fig. S3). The Fe<sup>3+</sup> ion is positioned along the 3-fold axis of the first helix bundle and coordinated by the side chains of His42 and His50 from each of the three monomers (Fig. 1C). This gives rise to a near-perfect octahedral coordination, in which the HisN<sub>ε</sub><sub>2</sub>-Fe distances range from 2.19 Å to 2.32 Å. His42<sub>X</sub> is in close hydrogen-bond distance to an acid/base pair, which forms a second shell around the His-Fe-octahedron. The side-chain functional groups of Glu46<sub>X</sub> and Arg43<sub>Z</sub> allow for an arrangement in which the deprotonated glutamate-carboxylate is oriented to His42<sub>X</sub>-N<sub>ε</sub><sub>1</sub> (2.66–2.81 Å), forcing His42<sub>X</sub>-N<sub>ε</sub><sub>2</sub> to point towards the Fe-center. A similar tautomerization effect might occur for the diametrically opposed His50<sub>Z</sub>, as it is in close distance to Gln54<sub>Z</sub> (3.24–3.27 Å), forcing His50<sub>Z</sub>-N<sub>ε</sub><sub>2</sub> to coordinate to Fe<sup>3+</sup>, resulting in an intertwined chelate-complex comprised of all three protein chains (see Fig. 2). It is noteworthy that the His42-N<sub>ε</sub><sub>2</sub>-Fe<sup>3+</sup> distances are comparable (2.295 ± 0.025 Å) but significantly longer than the His50-N<sub>ε</sub><sub>2</sub>-Fe<sup>3+</sup> distances (2.21 ± 0.02 Å). Such a Fe<sup>3+</sup> binding geometry has previously been observed in the membrane-piercing spike proteins of phages P2 (PDB code 3QR7) and φ92 (PDB code 3PQH)<sup>24</sup>, as well as in the receptor-binding domain of the long tail fiber of phage T4 (PDB code 2XGF)<sup>25</sup>. In the three reported cases, the N<sub>ε</sub><sub>2</sub>-Fe distances were 2.20 ± 0.01, 2.23 ± 0.01 and 2.31 ± 0.06, respectively, compared to an average of 2.25 ± 0.07 Å for the distances observed in φ11 RBP. The distance values in φ11 RBP are also close to the averages of those reported for N-Fe<sup>3+</sup> bonds in average-resolution and high-resolution protein structures deposited in the PDB, 2.25 ± 0.15 Å and 2.16 ± 0.13 Å, respectively<sup>26</sup>. However, all of these values are larger than those observed for distances between a heme Fe<sup>3+</sup> and the N<sub>ε</sub><sub>2</sub> of histidines coordinating it axially in myoglobin (2.00–2.11 Å)<sup>27</sup>. Indeed, iron ions often absorb in visible light wavelength ranges, giving rise to a red color for hemes and a brownish color for Fe-S clusters. However, the φ11 RBP and the related phage proteins discussed above are all colourless in solution<sup>24</sup>. It is worth noting that the Fe<sup>3+</sup> binding regions in the phage P2, φ92 and T4 spike structures involve histidines within a His-X-His motif at the apex of an intertwined triple β-helix. It has been proposed that this ion binding structure might strengthen the puncturing device of phages that pierce the cell wall<sup>24</sup>. In φ11 RBP, the His42-X<sub>7</sub>-His50 motif lies at a junction between two helical bundles. We therefore suggest that it serves a different role, perhaps by helping to stabilize the bundles that undergo a sharp turn at the hinge. If the Fe lock would not be in place, the structure of the α1/α2 segment would likely not be maintained as a rigid unit.

While the first and second helical bundles are collinear, the hinge introduces a sharp angle of ~30° between the second and the third bundle. This angle is the smallest possible since the second and third bundles are in contact at

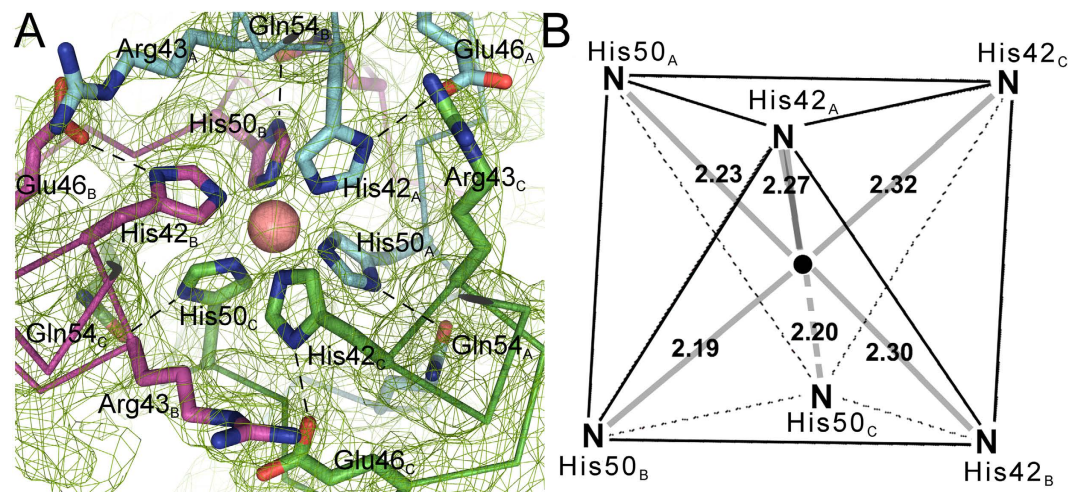


**Figure 1. The  $\phi 11$  RBP overall structure and domains.** (A) The structure of the complete RBP trimer is represented as a ribbon, with monomers coloured blue, green and violet. (B) The trimeric N-terminus of the  $\phi 11$  RBP has been superposed onto the 30 first amino acids of the trimeric TP901-1 RBP (r.m.s.d. of 1.8 Å). (C) Schematic representation of the  $\text{Fe}^{3+}$  binding motif. (D) The 5-bladed propeller domain rainbow coloured from dark blue (N-terminus) to red (C-terminus). The components of blade 5 are the N-terminal and C-terminal  $\beta$ -strands of the domain. (E) The two C-terminal trimeric domains, each assembling into three 4-stranded  $\beta$ -sheets. (F) Superposition of the two C-terminal domains. The r.m.s.d. between these domains after superimposition is 1.7 Å (blue and red structures).

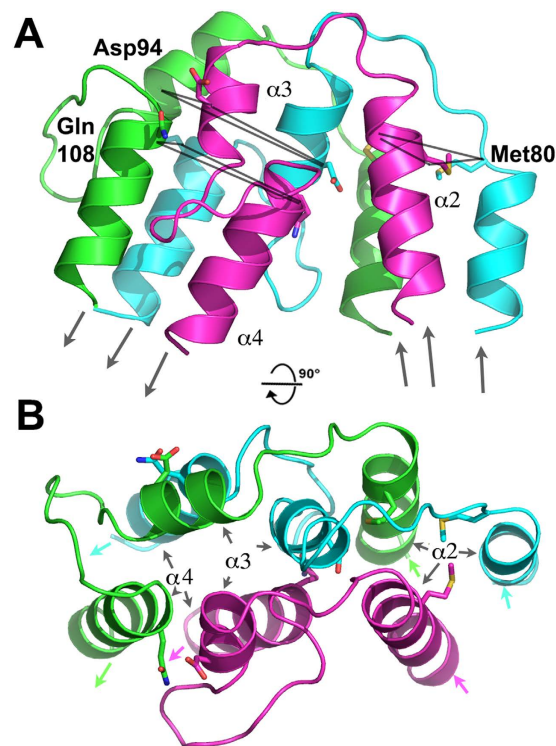
positions 61–64 and 116–121, while the first bundle contacts the propeller domain at position 219 (Figs 1 and S1). The hinge geometry is such that the sequences of the three helices of bundles two and three remain in phase (Fig. 3). The second bundle terminates with Met80, and the following sequences in the three monomers adopt a coil structure that abuts helices  $\alpha 3$  where the three sequences are already in phase (see Asp94, Fig. 3). Helices  $\alpha 3$  are followed by extended hairpin-structures and helices  $\alpha 4$  forming the final bundle.

**The five-bladed propeller platform and the two C-terminal tower domains.** The C-terminal end of the third helical bundle abuts the three five-bladed propeller domains that form the platform of  $\phi 11$  RBP (Fig. 1D). The three propellers are all equidistant to each other and to the molecule's main NCS-axis. This whole platform domain occupies a space that is  $\sim 100$  Å wide and  $\sim 40$  Å thick. Contacts between the three propeller domains are sparse, as each interface between two propeller domains buries a surface area of only  $457$  Å<sup>2</sup> from solvent, and much of this surface is buried due to a helix-helix contact at the center of the trimer axis (Fig. S2). The plane of the propeller is not perpendicular to the 3-fold axis, but is tilted upwards (as represented in Fig. 1) by an angle of  $\sim 30^\circ$ . This tilt improves access to the lower face of the propeller, and this might be linked to the function of RBP in interacting with ligands (see below).

As in other propeller structures, sets of four anti-parallel  $\beta$ -strands form each blade, and the N-terminal  $\beta$ -strand closes the fold by forming the final blade (blade 5) with the three C-terminal  $\beta$ -strands of the domain (Fig. 1D). A DALI search<sup>28</sup> with the  $\phi 11$  RBP propeller returned many significant hits above a Z-score of 15, and with r.m.s.d. values ranging from 3.1 to 4.0 Å. Most of the identified proteins are enzymes that mediate the degradation of carbohydrates. The highest score ( $Z = 16.1$ , r.m.s.d. = 3.1 Å), however, was obtained for the

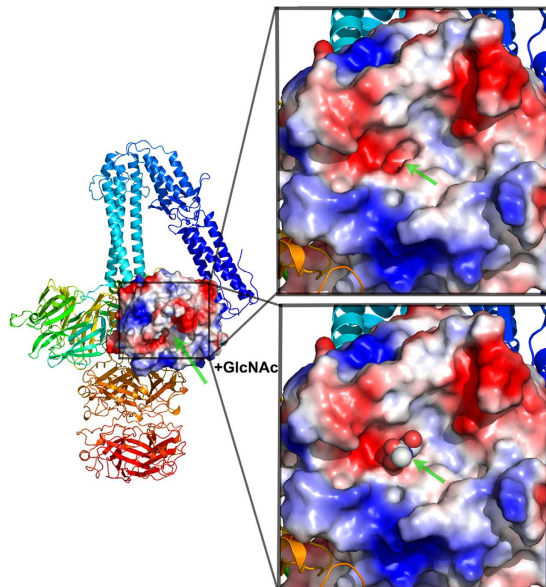


**Figure 2. The Fe<sup>3+</sup> binding motif.** (A) Final 2F<sub>o</sub>-F<sub>c</sub> electron density map of the Fe<sup>3+</sup> binding motif contoured at 2.0  $\sigma$ . The salmon-coloured sphere represents the octahedrally coordinated iron center in the histidine-rich stem region of helix-bundle  $\alpha$ 1. A second shell of residues, responsible for maintaining the active tautomeric states of the histidines, are represented with their side chains as sticks. Dashed lines indicate hydrogen-bonds and polypeptide chains are coloured blue for chain A, magenta for chain B and green for chain C. (B) Schematic representation of the octahedral motif, with the distances between His-N $\epsilon$ 2 atoms from residues His 42/His 50 and the Fe<sup>3+</sup> ion.



**Figure 3. The hinge domain.** (A) Ribbon representation of the trimeric hinge domain, encompassing amino acids Met80 to Gln108. Equivalent residues in the three monomers have been joined, and the sequence phase is conserved at this position. The direction of each chain is indicated by arrows. (B) 90° rotated view.

enzyme glutamine cyclotransferase from *Zymomonas mobilis* (PDB code 3NOL)<sup>29</sup>. To our knowledge, only two other examples of  $\beta$ -propellers in putative RBPs have been reported. For one, a distorted five-bladed propeller has been identified as the head domain of the RBP-P2 protein of phage PRD1, a *Tectiviridae* member infecting Gram-negative bacteria (PDB code 1N7U)<sup>30</sup>. The second example is the C-terminal domain gp131C of the *Pseudomonas* myophage PhiKZ, forming a seven-bladed  $\beta$ -propeller domain (PDB code 4GBF)<sup>31</sup>, and its position



**Figure 4. The platform domain.** The complete structure as a ribbon view with the surface of the five-bladed  $\beta$ -propeller domain coloured according to its electrostatic surface potential. **Inset above:** close-up view of the domain evidencing a deep cavity at its center. **Inset below:** model of a GlcNAc molecule (the main receptor's component) docked into the cavity.

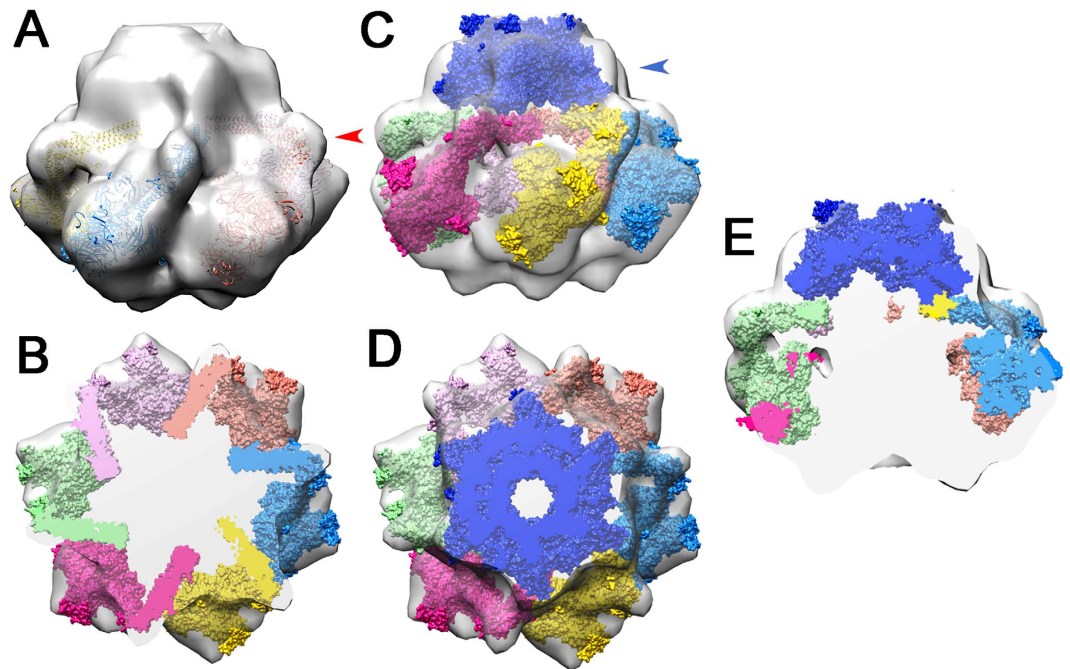
at the periphery of the baseplate has led to speculation that the propeller might act as the receptor-binding domain or as a cell-degrading enzyme. Neither of these hypotheses, however, have been confirmed experimentally.

When the *Zymomonas mobilis* glutamine cyclotransferase structure<sup>29</sup> was superimposed onto the RBP propeller domain, the active site of the enzyme overlaid a deep crevice located within the RBP lower face (Fig. 4). Modelling indicates that a cavity in this crevice has the correct size to accommodate a GlcNAc molecule, the cell wall teichoic acid (WTA) component specifically recognized by phage  $\phi 11$ <sup>9</sup>. Three water molecules occupy this cavity (Fig. S4A), which can be nicely replaced with hydroxyl groups of a modelled GlcNAc molecule<sup>8</sup>. The cavity is lined with polar residues (Gln165, Thr211, Gln330), which could serve to establish hydrogen bonds with the modelled GlcNAc molecule (Fig. S4B). Apolar residues Met164 and Met329 complete the walls of the cavity.

The two C-terminal tower domains form a structure of dimensions  $60 \times 50 \times 50 \text{ \AA}$ . These two domains are very similar in structure, which is confirmed by a superposition that yields a low r.m.s.d. value of  $1.7 \text{ \AA}$  for their C $\alpha$  atoms (Fig. 1E,F). A DALI search<sup>28</sup> performed with these domains returned only lower Z-scores, with the highest of these ( $Z = 6.1$ ; r.m.s.d. =  $3.2 \text{ \AA}$ ) for uracil-DNA glycosylase inhibitor, a small all- $\beta$  monomeric protein (PDB code 2UGI). The two structures essentially share the same anti-parallel  $\beta$ -sheet but differ in their oligomeric state and their surrounding structural features, and the identity (9%) is not high enough to assign possible functions to the C-terminal domains.

**Negative staining electron microscopy structure of the  $\phi 11$  baseplate.** To define the topology of the  $\phi 11$  baseplate and allow location of the RBP, we conducted electron microscopy analysis using negative staining of the virion. This approach has been successful in other cases<sup>17,18,32,33</sup>. We collected 512 images of the phage, and boxed 778 baseplate particles (see experimental procedures section). The final map has a resolution of  $23 \text{ \AA}$  (determined using the 0.5 FSC criterion) and allowed us to unambiguously place six  $\phi 11$  RBP trimers (Fig. 5A,B). To optimize this fit, we modified the hinge angle between the second and the third helix bundles from a value of  $\sim 30^\circ$  to  $\sim 90^\circ$ . The correlation is 0.845 with 95.5% of the atoms inside the map, calculated for a RBP orientation fit with the tower domain inclined towards the bottom of the baseplate. Compared to this, an orientation in which the tower domain would be “heads up”, reminiscent of the resting state of p2 baseplate<sup>17</sup>, only resulted in a correlation of 0.826 with 74.9% of the atoms inside the map. However, it has to be taken into account that the “heads-down” conformation is not a completely non-flexible state. The angle of the stem might in fact vary from the minimum observed in the X-ray structure to larger values when the phage scans the host's surface by moving the RBPs around the calculated average position for adhesion to the specific receptor. Such movements have been observed for several phages, such as phage T7<sup>10,11</sup>.

The remainder of the RBP structure was left unaltered, and the modified trimers fit well in a peripheral region of the map that could accommodate the triangular shape of the platform domain's platform. In order to explain the remaining density of the baseplate, we performed HHPRED<sup>34</sup> analyses of  $\phi 11$  proteins Gp43, Gp44 and Gp54, which are the most likely candidates for baseplate components<sup>9</sup>. This analysis revealed similarities with components of the lactococcal phage TP901-1 baseplate<sup>9</sup>, suggesting that the central part of the  $\phi 11$  baseplate is organized similarly to that of TP901-1<sup>18</sup>. Based on this analysis, Gp43 is predicted to exist as a hexamer and form the distal tail protein (Dit) ring and Gp44 as a trimer forming the tail-associated lysin (Tal) N-terminus and extension, while Gp54 N-terminus (the functional equivalent of BppU N-terminus) may form a second



**Figure 5. Electron microscopy negative-staining map of the RBP baseplate.** (A) The electron microscopy map is displayed with the six RBPs, represented as ribbons, fitted manually into the map using Chimera<sup>71</sup>. The red arrow indicates the hinge that was modified to provide a good fit. (B) A section of the baseplate with the six RBPs cut as shown in A. (C) View of Dit and the six RBPs fitted in the baseplate. The blue arrow indicates the Dit structure. (D) A section of the baseplate with Dit cut as shown in C. (E) Vertical slicing of the baseplate showing the cut volumes of Dit and the RBPs as well as the empty volume putatively occupied by the Tal and Gp54 proteins.

ring. Furthermore, the N-terminal segment and the first helical bundle of  $\phi 11$  RBP are structurally homologous to the N-terminal part of the phage TP901-1 RBP trimer, a structural domain that anchors the RBP into the BppU C-terminus<sup>18</sup>. We therefore also attempted to fit the phage TP901-1 Dit hexamer together with the BppU N-terminus (amino acids 1–160) into our electron density map<sup>18</sup>. The ring of the Dit had appropriate dimensions to fit the map above the RBPs (Fig. 5C,D). In contrast, the structure equivalent to BppU could not be fitted unequivocally as the internal density is not defined sufficiently. A large volume of the EM map remains to account for the Gp54 and for the Tal (Fig. 5E). When attaching the trimeric Tal N-terminal domain below the Dit hexamer, the three carbohydrate binding modules (2WAO) identified by HHpred project in the direction of the tail tip. These three bulky modules should fill the electron density map in between the six RBP trimers.

## Discussion

We have solved the crystal structure of  $\phi 11$  RBP and located this protein in the tail spike of the assembled phage using electron microscopy. Our analysis defines the domain organization of RBP, which can be divided into a ring region, a platform domain and a tower-like C-terminal structure composed of two nearly identical domains. Interestingly, the stem displays a severely bent, hook-like conformation that may undergo a conformational change as the protein can only be fitted into the electron density of the tail spike in a less bent arrangement. Unexpectedly, the stem also contains a bound iron. The function of this iron is unknown as its location differs from irons found in other spike proteins. The platform region is formed by three propeller domains and likely harbours the binding site for the substrate GlcNAc. Although soaking and cocrystallization experiments with GlcNAc were not successful, modelling suggests a reasonable location for the GlcNAc binding site in the platform region. Of note, the propeller fold was identified in the endosialidases of several phages. These enzymes cleave polysialic acid at the surface of their host in order to obtain access to the capsular cell wall. For example, coliphages K1F and phi92 possess such endosialidases, which also exhibit trimeric propeller domains attached to a stem<sup>35,36</sup>.

Comparison with the phage TP901-1 tail spike assembly allows us to also assign a putative location of the Gp43 and Gp54 proteins of  $\phi 11$ . Gp43 likely forms the hexameric Dit, while the N-terminus of Gp54 resembles the first 160 residues of TP901-1 BppU<sup>18</sup>. The remainder of the electron density is likely occupied by the rest of the large Gp54 and by Gp44, the Tal protein. Interestingly, the N-terminal folds of Dit and Tal are found in a wide range of phages<sup>37</sup>, including *Myoviridae* infecting Gram-negative bacteria (T4 or Mu<sup>38,39</sup>), *Siphoviridae* from Gram-negative (T5<sup>40</sup>) or Gram-positive (SPP1<sup>41,42</sup>) bacteria, lactococcal phages<sup>17,33,43,44</sup>, or even mycobacteria (Araucaria<sup>45</sup>). It is worth noting that the Tal protein is also found in the type VI secretion system machinery<sup>46</sup>. This observation suggests that the block formed by Dit and Tal could have been conserved through evolution, a phenomenon shared by other components such as the capsids MCP<sup>47,48</sup>, the connector<sup>49</sup>, as well as the tail MTP<sup>50</sup>.

Only the periphery of Dit (its C-terminal domain) and the Tal extension (e.g. a C-terminal fiber) could have been adapted to specific phage infection-style requirements<sup>51</sup>. In contrast,  $\phi 11$  Gp45, the RBP, does not exhibit analogy with other phage RBPs, in particular with those from lactococcal phages that also bind to saccharidic receptors. Lactococcal phages p2<sup>52</sup>, TP901-1<sup>53</sup>, Tuc2009<sup>19</sup>, bIL170<sup>54</sup> and 1358<sup>4</sup> all possess a trimeric receptor recognition head sharing a *bona fide* or a modified jelly-roll motif. The rest of their RBPs share common motifs in the neck or in the N-terminal domain (or stem). In Gp45, only the first 30 amino-acids of the stem resemble those of phages TP901-1 or Tuc2009.

Although the phage TP901-1 BppU protein does not seem to share such an extensive evolution coverage, the presence of a large part of it either in the RBP (Gp45) or in the subsequent protein (Gp54) was quite surprising. This finding suggests that phages might not capture only widespread elements in the protein domains repertoire, but also less diffused components, even between remote phages with different hosts. In the present case, the role of Gp54 is not documented and difficult to predict. We think it likely that in phage  $\phi 11$  Gp54, a large C-terminal domain might also accommodate the RBP N-terminus, but with different structural features compared to TP901-1 BppU, because of its much increased size. The electron microscopy low-resolution structure gives hints of the putative receptor binding sites, located below the five-bladed propeller domain. This arrangement allows for a correct orientation to capture the GlcNAc of the cell wall teichoic acids. Although, a well-defined cavity in the size of a monosaccharide exists, a much larger crevice surrounds this cavity, suggesting that other WTA components might complement the interaction. However, further structural data are necessary to develop this hypothesis.

## Experimental Procedures

**Overexpression, purification and crystallization of Gp45.** Gp45 was produced and purified as described elsewhere<sup>9</sup>. Briefly, after induction with IPTG the protein was purified to homogeneity using nickel-affinity chromatography and size exclusion chromatography. The purified protein carries a hexa-histidine tag at its N-terminus. Two similar crystallization solutions (0.1 M bicine/Trizma base pH 8.5, 10% w/v PEG 8000, 20% v/v ethylene glycol, 0.12 M monosaccharide-mix<sup>55</sup>) yielded initial crystals ( $50 \times 20 \times 5 \mu\text{m}$ ) of triangular shape that grew to bouquets at 16 °C over 1 week. Reproduction of the crystals in 5  $\mu\text{L}$  hanging drops lead to bigger crystals ( $500 \times 200 \times 50 \mu\text{m}$ ) that were used for X-ray structural analysis.

**Phasing, construction, refinement.** Data for native and derivative crystals were collected at the Swiss Light Source (SLS) on beamline X06DA (PXIII) using a PILATUS 2M hybrid pixel detector. For the determination of peak, inflection, high-remote and low-remote wavelengths from fluorescence spectra, the program CHOOCH was adjusted to the absorption edges of Ta-L-II (1.11325 Å peak) and Ta-L-III (1.25476 Å peak). Data were processed with the XDS package<sup>56</sup>. The crystals belong to spacegroup P1 and have unit-cell dimensions  $a = 87.06 \text{ \AA}$ ,  $b = 89.01 \text{ \AA}$ ,  $c = 93.26 \text{ \AA}$ ,  $\alpha = 93.0^\circ$ ,  $\beta = 105.2^\circ$  and  $\gamma = 117.6^\circ$ . A Ta<sub>6</sub>Br<sub>12</sub> derivative ( $a = 87.65 \text{ \AA}$ ,  $b = 89.60 \text{ \AA}$ ,  $c = 93.73 \text{ \AA}$ ,  $\alpha = 92.7^\circ$ ,  $\beta = 105.7^\circ$  and  $\gamma = 117.9^\circ$ ) was prepared by soaking native crystals in crystallization solution + 2 mM Ta<sub>6</sub>Br<sub>12</sub> for up to 2 weeks before backsoaking in crystallization condition and vitrification in liquid nitrogen. Anomalous data were processed according to MAD, SAD, MIRAS and SIRAS protocols using SHARP/autoSHARP<sup>57</sup>. Initial heavy atom coordinates and B-factors found with SHELXDE<sup>58</sup> were reedited with the SHARP-module Sushi and were refined until electron density maps showed good contrast. The outcomes of the various phasing protocols were compared, and the map derived from the SIRAS protocol was selected for further improvement.

A threefold NCS was elicited from the self-rotation function via polarrfn (ccp4<sup>59</sup>), giving a strong signal for rotation in reciprocal space for eulerian angles ( $\alpha = 357.4$ ,  $\beta = 63.6$ ,  $\gamma = 110.6$ ) corresponding to polar angles ( $\theta = 37.4$ ,  $\phi = 33.4$ ,  $\kappa = 120.1$ ). Using this self-rotation solution, GETAX<sup>60</sup> was able to find a set of translation vectors for the asymmetric unit in real space.

Due to the size of the multi-domain protein, it was split in various parts for further processing: two for the stem (before and after the 'hinge'), the platform domain and the C-terminal tower domain. Molecular masks<sup>61</sup> were created for each part and by generating correlation maps of them separately, a set of NCS operators could be assigned to each of them. The NCS-matrices were refined with IMP and averaging with AVE<sup>62–64</sup> converged the respective domains to about 80–90% of correlation. Each subdomain was integrated with respective NCS-matrices into a DM script for a combined density modification<sup>65,66</sup>. Starting at 5.8 Å, 80 cycles of consecutive solvent flattening, NCS-averaging, histogram matching and phase extension to a final resolution of 2.9 Å resulted in an interpretable map, which was clearly distinguishable from the unbiased calculated map. Initial refinement was carried out with REFMAC5<sup>59,67</sup> and PHENIX<sup>20</sup>, and after each step model building was done in COOT<sup>22</sup>. The final rounds of refinement were performed with autoBUSTER<sup>21</sup>, leading to Rfree/Rwork values of 21.1 and 17.5% (Table 1). A portion of the electron density map is shown in Fig. 3. Structural images were generated using pymol<sup>68</sup>.

**Negative staining electron microscopy.** Phages were purified as previously described<sup>9</sup>. Purified  $\phi 11$  phage (5  $\mu\text{L}$ ,  $10^9$  pfu) was applied to glow-discharged carbon-coated grids and left to adsorb for one min. Sample excess was blotted off and the grids were stained with 10  $\mu\text{L}$  of 1% uranyl acetate for 30 sec. Micrographs (512) were recorded on a 2Kx 2K FEI Eagle CCD camera using a Tecnai Spirit electron microscope operated at 120 kV and a magnification of 48,500 (resulting in a pixel size of 4.83 Å/pixel) (Fig. S5A). The three-dimensional reconstruction was produced using a single particle procedure and the XMIPP software package<sup>69</sup>. Particles defined around the baseplate (778) were manually picked and subjected to maximum likelihood (ML) classification and alignment implemented in Xmipp<sup>70</sup> imposing a 6-fold symmetry. The initial volume was determined using a random sample consensus (RANSAC) approach<sup>69</sup> with 5 2D classes. The resolution of the final volume was estimated at 23 Å using the Fourier Shell Correlation (FSC) 0.50 criterion (Fig. S5B).

**Molecular fitting and structure visualization.** Molecular graphics and analyses were performed with the UCSF Chimera package (Resource for Biocomputing, Visualization, and Informatics at UC-San Francisco)<sup>71</sup>. The model/EM map fitting was performed by the option “fit in map” of the “volume” register. The Dit fitting resulted in a correlation coefficient of 0.85 with 94% of the atoms in the map volume. The correlation coefficient calculated for six RBPs, with an orientation of the tower domains inclined towards the bottom of the baseplate, is 0.845 with 95.5% of the atoms inside the map.

**Data deposition.** X-ray structures and structure factors have been deposited with the Protein Data Bank (PDB, www.rcsb.org) under accession code 5EFV. The EM map of the baseplate reconstruction has been deposited at the Electron Microscopy Data Bank (EMDB, emdatabank.org).

## References

- Tong, S. Y., Davis, J. S., Eichenberger, E., Holland, T. L. & Fowler, V. G. Jr. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* **28**, 603–661, doi: 10.1128/cmr.00134-14 (2015).
- Lindsay, J. A. Staphylococcus aureus genomics and the impact of horizontal gene transfer. *Int J Med Microbiol* **304**, 103–109, doi: 10.1016/j.ijmm.2013.11.010 (2014).
- Xia, G. & Wolz, C. Phages of Staphylococcus aureus and their impact on host evolution. *Infect Genet Evol* **21**, 593–601, doi: 10.1016/j.meegid.2013.04.022 (2014).
- McCabe, O. *et al.* The targeted recognition of Lactococcus lactis phages to their polysaccharide receptors. *Mol Microbiol* **96**, 875–886, doi: 10.1111/mmi.12978 (2015).
- Chapot-Chartier, M. P. *et al.* Cell surface of Lactococcus lactis is covered by a protective polysaccharide pellicle. *J Biol Chem* **285**, 10464–10471, doi: 10.1074/jbc.M109.082958 (2010).
- Ainsworth, S. *et al.* Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. *mBio* **5**, e00880–00814, doi: 10.1128/mBio.00880-14 (2014).
- Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nat Commun* **4**, 2345, doi: 10.1038/ncomms3345 (2013).
- Xia, G. *et al.* Wall teichoic Acid-dependent adsorption of staphylococcal siphovirus and myovirus. *J Bacteriol* **193**, 4006–4009, doi: 10.1128/JB.01412-10 (2011).
- Li, X. *et al.* An essential role for the baseplate protein Gp45 in phage adsorption to Staphylococcus aureus. *Sci Rep* **6**, 26455, doi: 10.1038/srep26455 (2016).
- Hu, B., Margolin, W., Molineux, I. J. & Liu, J. The bacteriophage T7 virion undergoes extensive structural remodeling during infection. *Science* **339**, 576–579, doi: 10.1126/science.1231887 (2013).
- Gonzalez-Garcia, V. A. *et al.* Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J Biol Chem* **290**, 10038–10044, doi: 10.1074/jbc.M114.614222 (2015).
- Parent, K. N. *et al.* OmpA and OmpC are critical host factors for bacteriophage Sf6 entry in Shigella. *Mol Microbiol* **92**, 47–60, doi: 10.1111/mmi.12536 (2014).
- Plancon, L. *et al.* Characterization of a high-affinity complex between the bacterial outer membrane protein FhuA and the phage T5 protein pb5. *J Mol Biol* **318**, 557–569, doi: 10.1016/S0022-2836(02)00089-X (2002).
- Flayhan, A., Wien, F., Paternostre, M., Boulanger, P. & Breyton, C. New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its Escherichia coli receptor FhuA. *Biochimie* **94**, 1982–1989, doi: 10.1016/j.biochi.2012.05.021 (2012).
- Breyton, C. *et al.* Assessing the conformational changes of pb5, the receptor-binding protein of phage T5, upon binding to its Escherichia coli receptor FhuA. *J Biol Chem* **288**, 30763–30772, doi: 10.1074/jbc.M113.501536 (2013).
- Kostyuchenko, V. A. *et al.* Three-dimensional structure of bacteriophage T4 baseplate. *Nat Struct Biol* **10**, 688–693, doi: 10.1038/nsb970 (2003).
- Sciara, G. *et al.* Structure of lactococcal phage p2 baseplate and its mechanism of activation. *Proc Natl Acad Sci* **107**, 6852–6857, doi: 10.1073/pnas.1000232107 (2010).
- Veesler, D. *et al.* Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc Natl Acad Sci* **109**, 8954–8958, doi: 10.1073/pnas.1200966109 (2012).
- Legrand, P. *et al.* The Atomic Structure of the Phage Tuc2009 Baseplate Tripod Suggests that Host Recognition Involves Two Different Carbohydrate Binding Modules. *mBio* **7**, e01781–01715, doi: 10.1128/mBio.01781-15 (2016).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**, 213–221, doi: 10.1107/S0907444909052925 (2010).
- Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D* **60**, 2210–2221, doi: 10.1107/S0907444904016427 (2004).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D* **66**, 486–501, doi: 10.1107/S0907444910007493 (2010).
- Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797, doi: 10.1016/j.jmb.2007.05.022 (2007).
- Browning, C., Shneider, M. M., Bowman, V. D., Schwarzer, D. & Leiman, P. G. Phage pierces the host cell membrane with the iron-loaded spike. *Structure* **20**, 326–339, doi: 10.1016/j.str.2011.12.009 (2012).
- Bartual, S. G. *et al.* Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci* **107**, 20287–20292, doi: 10.1073/pnas.1011218107 (2010).
- Zheng, H., Chruszcz, M., Lasota, P., Lebioda, L. & Minor, W. Data mining of metal ion environments present in protein structures. *J Inorg Biochem* **102**, 1765–1776, doi: 10.1016/j.jinorgbio.2008.05.006 (2008).
- Lima, F. A. *et al.* Probing the electronic and geometric structure of ferric and ferrous myoglobins in physiological solutions by Fe K-edge absorption spectroscopy. *PCCP Phys Chem Chem Phys* **16**, 1617–1631, doi: 10.1039/c3cp53683a (2014).
- Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DALI-Lite v.3. *Bioinformatics* **24**, 2780–2781, doi: 10.1093/bioinformatics/btn507 (2008).
- Carrillo, D. R. *et al.* Kinetic and structural characterization of bacterial glutaminyl cyclases from Zymomonas mobilis and Myxococcus xanthus. *Biol Chem* **391**, 1419–1428, doi: 10.1515/BC.2010.130 (2010).
- Xu, L., Benson, S. D., Butcher, S. J., Bamford, D. H. & Burnett, R. M. The receptor binding protein P2 of PRD1, a virus targeting antibiotic-resistant bacteria, has a novel fold suggesting multiple functions. *Structure* **11**, 309–322, doi: 10.1016/S0969-2126(03)00023-6 (2003).
- Sycheva, L. V. *et al.* Crystal structure and location of gp131 in the bacteriophage phiKZ virion. *Virology* **434**, 257–264, doi: 10.1016/j.virol.2012.09.001 (2012).
- Spinelli, S. *et al.* Cryo-electron microscopy structure of lactococcal siphophage 1358 virion. *J Virol* **88**, 8900–8910, doi: 10.1128/JVI.01040-14 (2014).

33. Spinelli, S., Veesler, D., Bebeacua, C. & Cambillau, C. Structures and host-adhesion mechanisms of lactococcal siphophages. *Front Microbiol* **5**, 3, doi: 10.3389/fmicb.2014.00003 (2014).
34. Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244–248, doi: 10.1093/nar/gki408 (2005).
35. Stummeyer, K., Dickmanns, A., Muhlenhoff, M., Gerardy-Schahn, R. & Ficner, R. Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nature Struct Biol* **12**, 90–96, doi: 10.1038/nsmb874 (2005).
36. Schwarzer, D. *et al.* Structure and biochemical characterization of bacteriophage phi92 endosialidase. *Virology* **477**, 133–143, doi: 10.1016/j.virol.2014.11.002 (2015).
37. Veesler, D. & Cambillau, C. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol Mol Biol Rev* **75**, 423–433, first page of table of contents, doi: 10.1128/MMBR.00014-11 (2011).
38. Arisaka, F., Kanamaru, S., Leiman, P. & Rossmann, M. G. The tail lysozyme complex of bacteriophage T4. *Int J Biochem Cell Biol* **35**, 16–21, doi: 10.1016/S1357-2725(02)00098-5 (2003).
39. Kanamaru, S. *et al.* Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553–557, doi: 10.1038/415553a (2002).
40. Flayhan, A. *et al.* Crystal Structure of pb9, the Distal Tail Protein of Bacteriophage T5: a Conserved Structural Motif among All Siphophages. *J Virol* **88**, 820–828, doi: 10.1128/JVI.02135-13 (2014).
41. Veesler, D. *et al.* Crystal structure of bacteriophage SPP1 Distal Tail Protein (gp19.1), a baseplate hub paradigm in Gram-positive infecting phages. *J Biol Chem* **285**, 36666–36673, doi: 10.1074/jbc.M110.157529 (2010).
42. Goulet, A. *et al.* The opening of the SPP1 bacteriophage tail, a prevalent mechanism in Gram-positive-infecting siphophages. *J Biol Chem* **286**, 25397–25405, doi: 10.1074/jbc.M111.243360 (2011).
43. Bebeacua, C. *et al.* Visualizing a Complete Siphoviridae Member by Single-Particle Electron Microscopy: the Structure of Lactococcal Phage TP901-1. *J Virol* **87**, 1061–1068, doi: 10.1128/JVI.02836-12 (2013).
44. Campanacci, V. *et al.* Solution and electron microscopy characterization of lactococcal phage baseplates expressed in *Escherichia coli*. *J Struct Biol* **172**, 75–84, doi: 10.1016/j.jsb.2010.02.007 (2010).
45. Sassi, M., Bebeacua, C., Drancourt, M. & Cambillau, C. The First Structure of a Mycobacteriophage, the Mycobacterium abscessus subsp. bolletii Phage Araucaria. *J Virol* **87**, 8099–8109, doi: 10.1128/JVI.0120913 (2013).
46. Leiman, P. G. *et al.* Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci* **106**, 4154–4159, doi: 10.1073/pnas.0813360106 (2009).
47. Fokine, A. *et al.* Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci* **102**, 7163–7168, doi: 10.1073/pnas.0502164102 (2005).
48. Benson, S. D., Bamford, J. K., Bamford, D. H. & Burnett, R. M. Does common Architecture reveal a viral lineage spanning all three domains of life? *Mol Cell* **16**, 673–685, doi: 10.1016/j.molcel.2004.11.016 (2004).
49. Cardarelli, L. *et al.* The crystal structure of bacteriophage HK97 gp6: defining a large family of head-tail connector proteins. *J Mol Biol* **395**, 754–768, doi: 10.1016/j.jmb.2009.10.067 (2010).
50. Pell, L. G., Kanelis, V., Donaldson, L. W., Howell, P. L. & Davidson, A. R. The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proc Natl Acad Sci* **106**, 4160–4165, doi: 10.1073/pnas.0900044106 (2009).
51. Dieterle, M. E. *et al.* Exposing the secrets of two well-known *Lactobacillus casei* phages, J-1 and PL-1, by genomic and structural analysis. *Appl Environ Microbiol* **80**, 7107–7121, doi: 10.1128/AEM.02771-14 (2014).
52. Spinelli, S. *et al.* Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses. *Nat Struct Mol Biol* **13**, 85–89, doi: 10.1038/nsmb1029 (2006).
53. Spinelli, S. *et al.* Modular structure of the receptor binding proteins of *Lactococcus lactis* phages. The RBP structure of the temperate phage TP901-1. *J Biol Chem* **281**, 14256–14262, doi: 10.1074/jbc.M600666200 (2006).
54. Ricagno, S. *et al.* Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bLL170. *J Virol* **80**, 9331–9335, doi: 10.1128/JVI.0116006 (2006).
55. Gorrec, F. The MORPHEUS protein crystallization screen. *J Appl Crystallogr* **42**, 1035–1042, doi: 10.1107/S0021889809042022 (2009).
56. Kabsch, W. XDS. *Acta Crystallogr D* **66**, 125–132, doi: 10.1107/S0907444909047337 (2010).
57. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol Biol* **364**, 215–230, doi: 10.1385/1-59745-266-1:215 (2007).
58. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr D* **58**, 1772–1779, doi: 10.1107/S0907444902011678 (2002).
59. CCP4, C. C. P. N. The CCP4 suite: programs for crystallography. *Acta Crystallogr D* **50**, 760–766, doi: 10.1107/S0907444994003112 (1994).
60. Vonrhein, C. & Schulz, G. E. Locating proper non-crystallographic symmetry in low-resolution electron-density maps with the program GETAX. *Acta Crystallogr D* **55**, 225–229, doi: 10.1107/S0907444998007914 (1999).
61. Kleywegt, G. J. & Jones, T. A. Software for handling macromolecular envelopes. *Acta Crystallogr D* **55**, 941–944, doi: 10.1107/S0907444999001031 (1999).
62. Kleywegt, G. J. Validation of protein models from C $\alpha$  coordinates alone. *J Mol Biol* **273**, 371–376, doi: 10.1006/jmbi.1997.1309 (1997).
63. Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. In *International Tables for Crystallography* Vol. F (ed Arnold, E. & Rossmann, M. G.) Ch. 17.1, 353–367, doi: 10.1107/97809553602060000111 (Kluwer Academic Publishers, 2001).
64. Kleywegt, G. J. & Read, R. J. Not your average density. *Structure* **5**, 1557–1569, doi: 10.1016/S0969-2126(97)00305-5 (1997).
65. Terwilliger, T. C. Statistical density modification with non-crystallographic symmetry. *Acta Crystallogr D* **58**, 2082–2086, doi: 10.1107/S0907444902016360 (2002).
66. Wang, B. C. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol* **115**, 90–112, doi: 10.1016/0076-6879(85)15009-3 (1985).
67. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* **67**, 355–367, doi: 10.1107/S0907444911001314 (2011).
68. Schrödinger, L. L. C. The PyMOL molecular graphics system. *PyMOL*, New York, USA URL <https://www.pymol.org/> (2016).
69. Sorzano, C. O. *et al.* XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148**, 194–204, doi: 10.1016/j.jsb.2004.06.006 (2004).
70. Scheres, S. H., Nunez-Ramirez, R., Sorzano, C. O., Carazo, J. M. & Marabini, R. Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc* **3**, 977–990, doi: 10.1038/nprot.2008.62 (2008).
71. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612, doi: 10.1002/jcc.20084 (2004).

## Acknowledgements

We thank the Resource for Biocomputing, Visualization, and Informatics at UC-San Francisco for Chimera distribution (supported by NIGMS P41-GM103311). We acknowledge the Paul Scherrer Institut, Villigen, Switzerland for provision of synchrotron radiation beamtime at beamline PXIII (x06da) of the SLS and would like to thank its beamline support team for assistance. This work was supported in part by the French Infrastructure



for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01. TS, GX and CK were supported by Collaborative Research Centers 766 and TR34 from the Deutsche Forschungsgemeinschaft.

### Author Contributions

C.K. solved crystal structure, interpreted data and wrote manuscript. G.X. supervised experiments and wrote manuscript. P.K. produced and purified the phage particles. S.S. collected EM data. A.R. supervised experiments. C.C. and T.S. supervised the work and wrote the manuscript. Every author reviewed the manuscript prior to submission.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Koç, C. *et al.* Structure of the host-recognition device of *Staphylococcus aureus* phage  $\phi$ 11. *Sci. Rep.* **6**, 27581; doi: 10.1038/srep27581 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

# 1 **Structure of the host-recognition device of *Staphylococcus aureus* phage $\phi$ 11**

2

## 3 **Supplementary information**

4

5 Cengiz Koç<sup>1</sup>, Guoqing Xia<sup>2, 3, 4</sup>, Petra Kühner<sup>2</sup>, Silvia Spinelli<sup>5, 6</sup>, Alain Roussel<sup>5, 6</sup>, Christian  
6 Cambillau<sup>5, 6</sup> \* and Thilo Stehle<sup>1, 3, 7</sup> \*

7

8 1. Interfaculty Institute of Biochemistry, University of Tübingen, 72076 Tübingen, Germany

9 2. Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, 72076  
10 Tübingen, Germany

11 3. German Center for Infection Research (DZIF), partner site Tübingen, Germany

12 4. Institute of Inflammation & Repair, Faculty of Medical and Human Sciences, University of  
13 Manchester, Manchester, United Kingdom

14 5. Architecture et Fonction des Macromolécules Biologiques, UMR 7257 CNRS, 13288 Marseille  
15 Cedex 09, France.

16 6. Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche  
17 Scientifique, UMR 6098, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France

18 7. Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee,  
19 USA

20

21 \*Correspondence to:

22 Thilo Stehle, E-mail: [thilo.stehle@uni-tuebingen.de](mailto:thilo.stehle@uni-tuebingen.de)

23 or Christian Cambillau, E-mail: [ccambillau@gmail.com](mailto:ccambillau@gmail.com)

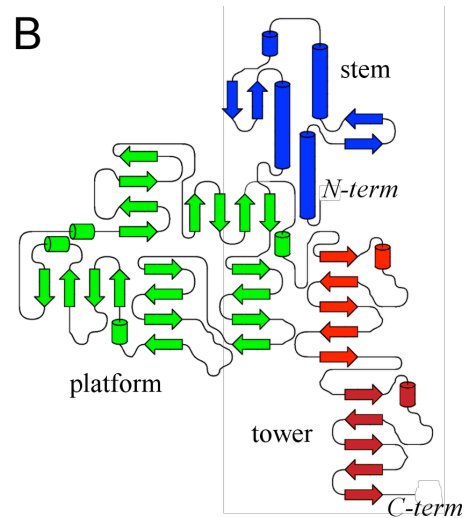
24

25

A

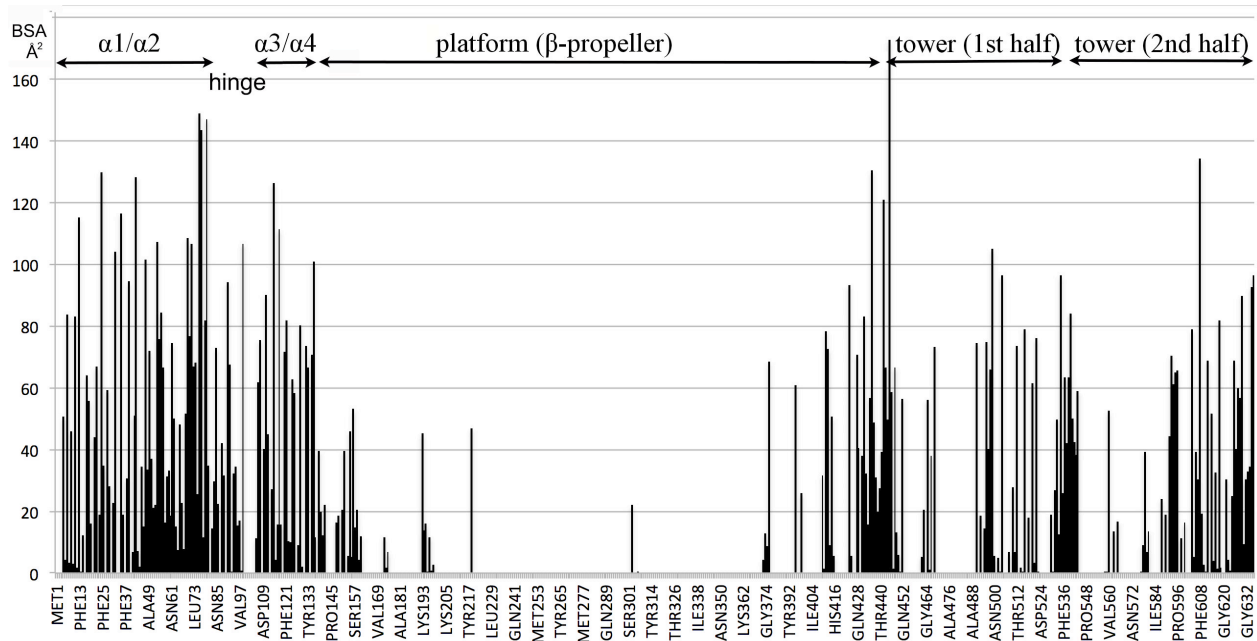
50 MSNKLI TDLSRVFDYRYVDENEYNFKLI SDMLTDFNFSLEY **HRNKEVFAH**  
 100 **DGEQIKYEHLNVTSN**VSDFLTYLNGRFSNM **VLGHNGD**GINEVKDARV **DNT**  
 150 **GYGHK**TLQDRLYHDYSTLDVFTKKVEKAVDEHYKEYRATEYRFEPKEQEP  
 200 EFITDLSPYTNAV **MQ**SFWVDPRTKIIYMTQARPGNHMLSR LKPNQQFID  
 250 RLLVKNGGGHGT HNAYRYIDGELWIYS AVLD SNKNNKFVRPQYRTGEITYG  
 300 NEMQDVMPNIFNDRYTSAIYNPVENLMIFRREYKPTERQLKNSLNFVEVR  
 350 SADDIDKGD KVLVYQMDIPMEYTSDTQP **MQ**GITYDAGILYWTGDSNTAN  
 400 PNYLQGF DIKTKELLFKRRIDIGV **NNNFKGD**FQEAEGLDMYDLETGRK  
 450 ALLIGVTIGPGNNRHHHSIYSIGQRGVNQFLKNIAPQV **SMTDSGGRRVKPLP**  
 500 **IQNPA**YLS DITEVGHYYIYTQDTQNALDFPLPKAFRDAGWFLDVLPGHYN  
 550 **GALRQ**VLTRNSTGRNMLKFERVIDIFNKKNGAWNFCPQNAGYWEHIPKS  
 600 **ITKLS**DLKIVGLDFYITTEESNRFTDFPKDFKGIAGWILEVKSNTPGNTT  
 636 **QVLR**RNNFP **SAHQ**FLVRNFGTGGVGKWSLFEKGKVV E

B



25  
 26 **Supplementary Figure S1. Primary and secondary structure of one Gp45 protomer.** A/  
 27 Colour-coded sequence of Gp45 monomer: “Stem” in blue, “platform” in green, “tower1” in  
 28 light-red and “tower2” in dark-red. Highlighted in coloured boxes are special features: the iron-  
 29 binding-site in dark red, the interruptions in the stem in black, the putative GlcNAc binding motif  
 30 in pink and the disordered loop in yellow. **B/** Topology of one Gp45 protomer partitioned in a  
 31 colour-code according to Figure S1A.

32  
 33  
 34  
 35  
 36



37

38 **Supplementary Figure S2. Buried surface area along the Gp45 chain.** The water accessible

39 surface area of one monomer, buried by the two other monomers, has been calculated by PISA <sup>12</sup>

40 for each residue. Note the extensive interaction of the trimers along the stem (amino acids 1-140),

41 while practically no interaction is observed between the β-propeller domains.

42

43

44

45

46

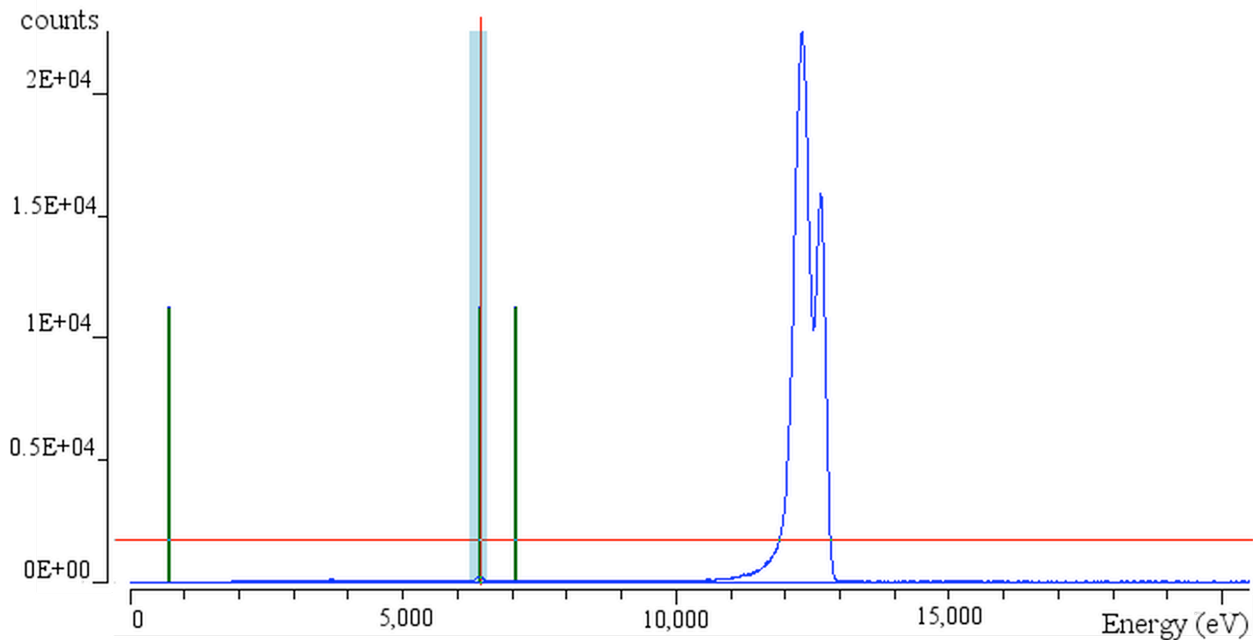
47

48

49

50

51



52  
53 **Supplementary Figure S3. EXAFS spectrum of  $\phi 11$  RBP crystal.** X-ray fluorescence spectra  
54 obtained upon illumination of a  $\phi 11$  RBP crystal by 12.65 keV incident X-rays. The only  
55 detectable fluorescent line emission corresponds to the iron K- $\alpha$  energy at 6.40 keV (highlighted  
56 in light blue). The large peak on the right corresponds to the diffused scattering of incident X-  
57 rays.

58

59

60

61

62

63

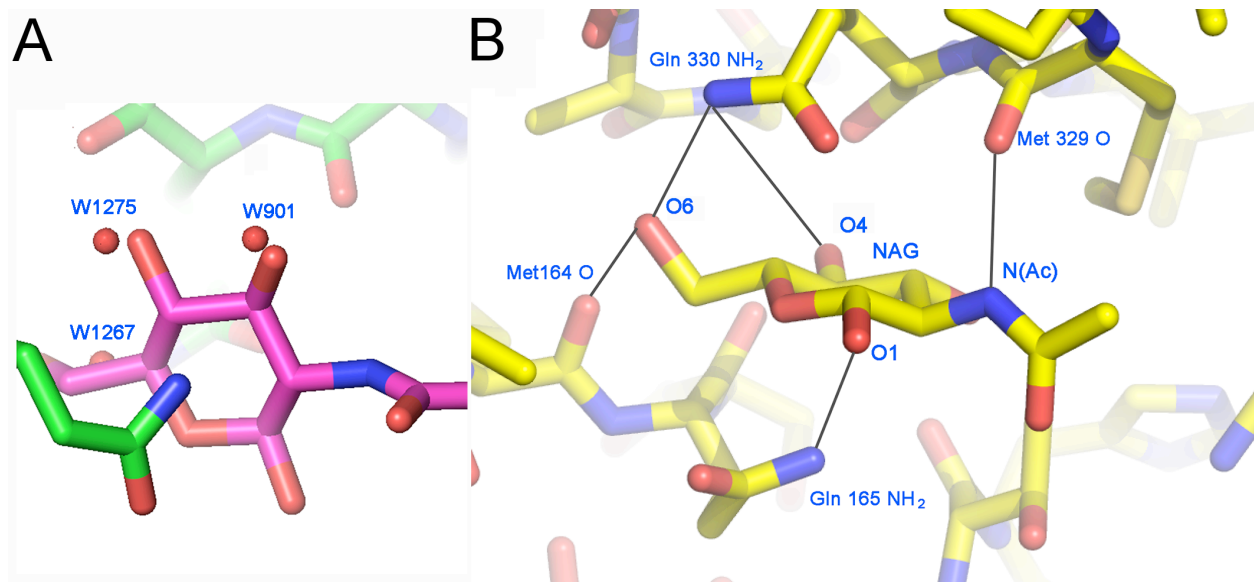
64

65

66

67

68



69

70 **Supplementary Figure S4. Putative GlcNAc binding-site.** **A/** Hydroxyl groups O3, O4 and O6

71 of a GlcNAc substrate can be overlaid with W901, W1275 and W1267, respectively, defining a

72 specific binding mode for a glucose residue. **B/** The overlay shown in A facilitates the modelling

73 of GlcNAc, a major moiety of staphylococcal peptidoglycan, in a preformed cavity. The peptide-

74 oxygen of M164 and the sidechain of Q165 of blade 2 are in appropriate proximity of O6 and O1,

75 respectively, allowing for hydrogen bonds. The second half of GlcNAc is accommodated by

76 M329 and Q330 of blade 5, allowing for hydrogen bridges to the nitrogen of the N-acetyl moiety,

77 O4 and O6, respectively. This arrangement keeps GlcNAc stably in position.

78

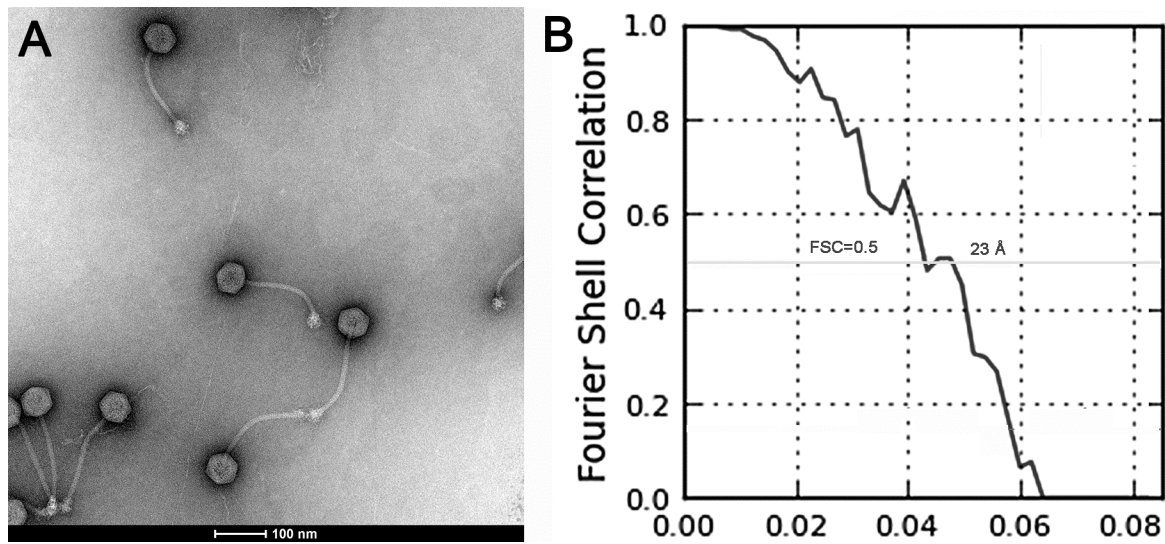
79

80

81

82

83



84  
85 **Supplementary Figure S5. Electron microscopy of phage  $\phi$ 11.** **A/** An EM picture of the whole  
86 phage  $\phi$ 11. **B/** The Fourier shell correlation curve plotted against 1/resolution. The FSC 0.5 cut-  
87 off yields a resolution of 23 Å.

88  
89  
90  
91