

Biosynthesis of  
soluble capsule precursors in  
*Staphylococcus aureus*

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Hannah Ulm

aus

Saarlouis

Bonn 2015



Angefertigt mit Genehmigung der Mathematisch Naturwissenschaftlichen Fakultät  
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Referent: Prof. Dr. Hans-Georg Sahl
2. Referent: Prof. Dr. Günter Mayer

Tag der Promotion: 13.7.2016

Erscheinungsjahr: 2016



*Meinen Eltern*



## Abstract

Most invasive pathogens produce polysaccharide capsules, which play an essential role in conferring protection against host immune defense. In the case of *Staphylococcus aureus*, most clinical isolates are encapsulated, and inhibition of capsule biosynthesis may offer a valuable strategy for novel anti-infective treatment. Despite their importance for pathogenicity, the biochemistry underlying the assembly of staphylococcal capsular polysaccharides is not fully understood. Three nucleotide-activated sugar precursors, synthesized in the cytoplasm of the cell, are required for *S. aureus* capsule production. In this thesis, the UDP-*N*-acetyl-D-glucosamine 4,6-dehydratases CapD and CapE of *S. aureus* serotype 5, which catalyze the first steps in the synthesis of the soluble capsule precursors UDP-*N*-acetyl-D-fucosamine (UDP-D-FucNAc) and UDP-*N*-acetyl-L-fucosamine (UDP-L-FucNAc), respectively, were purified and characterized. CapD is an integral membrane protein and was obtained for the first time in a purified, active form. Using a new, robust and sensitive capillary electrophoresis-based method for detection of the soluble capsule precursors, kinetic studies for CapD and CapE were performed, and a compound library was screened in search for enzyme inhibitors. Several active compounds were identified and characterized, including suramin (IC<sub>50</sub> at CapE 1.82 μM) and ampicillin (IC<sub>50</sub> at CapD of 40.1 μM). Moreover, the cell wall precursors UDP-*N*-acetylmuramyl-pentapeptide and lipid II appear to function as inhibitors of CapD enzymatic activity, suggesting an integrated mechanism of regulation for cell envelope biosynthesis pathways in *S. aureus*. A further aim of the present study was the elucidation of the pathway underlying UDP-D-FucNAc biosynthesis in *S. aureus*. Using the purified recombinant enzymes CapD and CapN, enzymatic *in vitro* synthesis of the soluble capsule precursor UDP-D-FucNAc was achieved for the first time. Furthermore, the role of the tyrosine kinase complex CapAB in the regulation of capsular polysaccharide production was investigated. The capsule biosynthetic enzyme CapE was identified as regulatory target of the CapAB kinase complex, and it was shown that tyrosine phosphorylation enhances CapE catalytic activity *in vitro*. Besides substantial progress in the understanding of capsule biosynthesis and the underlying regulation, the established *in vitro* systems provide the molecular basis for screening for potential anti-virulence agents.





## Acknowledgements

My heartfelt thanks go to:

- **German-Israeli Foundation for Scientific Research and Development**
- my academic mentors **Prof Hans-Georg Sahl** and **Prof Tanja Schneider**
- my thesis committee members **Prof Günter Mayer** (LIMES-Institute, Chemical Biology and Chemical Genetics), **Prof Thorsten Lang** (LIMES-Institute, Membrane Biochemistry) and **Prof Ulrich Kubitscheck** (Institute of Physical and Theoretical Chemistry, Biophysical Chemistry)
- **Prof Christa E. Müller** (Institute of Pharmacy, Pharmaceutical Chemistry I), **Prof Jean C. Lee** (Brigham and Women's Hospital/Harvard Medical School, Boston, USA) and **Prof Yechiel Shai** (The Weizmann Institute of Science, Rehovot, Israel)
- **Dr Wenjin Li** (Pharmaceutical Chemistry I), who performed capillary electrophoresis analyses
- **Dr Marc Sylvester** (Institute of Biochemistry and Molecular Biology) and my colleague **Michaele Josten**, who performed mass spectrometry analyses
- my colleague **Marvin Rausch**, who performed some of the *in vitro* kinase assays presented in this thesis
- all other members of the Sahl working group, former and present
- **Prof Gabriele Bierbaum** and her group at the Institute of Medical Microbiology, Immunology and Parasitology
- faculty and staff at the Institute for Microbiology and Biotechnology



## Abbreviations and symbols

aa	amino acid(s)
ABC	ATP binding cassette
ADP	adenosine 5'-diphosphate
<i>aqua a.i.</i>	<i>aqua ad iniectabilia</i>
ATCC	American Type Culture Collection
ATP	adenosine 5'-triphosphate
Bis-Tris	2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
BY-kinase	bacterial tyrosine kinase
°C	degree Celsius
C18	alkyl chain(s) of 18 carbon atoms
CapDtr	truncated soluble CapD mutant protein
CAPS	3-(cyclohexylamino)-1-propanesulfonic acid
CE	capillary electrophoresis
Ci	Curie
CP	capsular polysaccharide
CP5/CP8	<i>Staphylococcus aureus</i> serotype 5/8 capsular polysaccharide
C <sub>55</sub> P(P)	undecaprenyl-(pyro)phosphate
Da	Dalton
DDM	<i>n</i> -dodecyl-β-D-maltoside
DE3	λ prophage carrying bacteriophage T7 RNA polymerase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates
dTDP	deoxythymidine 5'-diphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
e.g.	from Latin <i>exempli gratia</i> , meaning “for example”
EPS	exopolysaccharide
et al.	from Latin <i>et alii</i> , meaning “and others”
<i>E</i> -value	expectation value
(D/L-)FucNAc	<i>N</i> -acetyl-D/L-fucosamine
g	gram
(D-)GlcNAc	<i>N</i> -acetyl-D-glucosamine
h	hour
HA	hyaluronic acid
His <sub>6</sub>	hexahistidine
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography
Hz	hertz
IC <sub>50</sub>	half maximal inhibitory concentration
i.e.	from Latin <i>id est</i> , meaning “that is to say”
IPTG	isopropyl-β-D-thiogalactopyranoside
K antigen	capsular antigen
kb	kilobase
<i>K<sub>m</sub></i>	Michaelis-Menten constant
KPi	potassium phosphate buffer

I	litre
LB	lysogeny broth
LDS	lithium dodecyl sulfate
lipid II	undecaprenyl-pyrophosphoryl- <i>N</i> -acetylmuramyl-(pentapeptide)- <i>N</i> -acetylglucosamine
lipid I <sub>cap</sub>	undecaprenyl-pyrophosphoryl- <i>N</i> -acetyl-D-fucosamine
lipid II <sub>cap</sub>	undecaprenyl-pyrophosphoryl-D-FucNAc- <i>N</i> -acetyl-L-fucosamine
lipid III <sub>cap</sub>	undecaprenyl-pyrophosphoryl-D-FucNAc-L-FucNAc- <i>N</i> -acetyl-D-mannosaminuronic acid
lipid A	glycolipid component of lipopolysaccharide
LPS	lipopolysaccharide
λ <sub>max</sub>	wavelength of the most intense UV/VIS absorption
M	mol/litre
m	metre
[M-H] <sup>-</sup>	deprotonated molecular ion
MALDI-TOF MS	matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
(D-)ManNAc	<i>N</i> -acetyl-D-mannosamine
(D-)ManNAcA	<i>N</i> -acetyl-D-mannosaminuronic acid
MeOPN	<i>O</i> -methyl phosphoramidate
min	minute
MOPS	3-( <i>N</i> -morpholino)propanesulfonic acid
<i>M<sub>r</sub></i>	relative molecular mass
MRSA	methicillin-resistant <i>Staphylococcus aureus</i>
MS	mass spectrometry
(D-)MurNac	<i>N</i> -acetyl-D-muramic acid
<i>m/z</i>	mass-to-charge ratio
NAD(P) <sup>+</sup>	β-nicotinamide adenine dinucleotide (2'-phosphate), oxidized form
NAD(P)H	β-nicotinamide adenine dinucleotide (2'-phosphate), reduced form
nanoLC-MS/MS	nanoscale liquid chromatography coupled to tandem mass spectrometry
NaPi	sodium phosphate buffer
NCBI	National Center for Biotechnology Information (United States)
NCTC	National Collection of Type Cultures (Public Health England)
Ni-NTA	nickel-nitrilotriacetic acid
NMR	nuclear magnetic resonance
O antigen	highly variable glycan polymer component of lipopolysaccharide
OD <sub>600</sub>	optical density at a wavelength of 600 nm
PAP	3'-phosphoadenosine 5'-phosphate
PCR	polymerase chain reaction
pET	plasmid expression by T7-RNA-polymerase
PG	peptidoglycan
pH	from Latin <i>pondus hydrogenii</i> ; hydrogen ion concentration
PHP	histidinol phosphate phosphoesterase
polybrene	hexadimethrine bromide
ppm	parts per million
psi	pounds per square inch
PSSMID	identifier for a position specific scoring matrix
<i>p</i>	calculated probability

(r)RNA	(ribosomal) ribonucleic acid
rpm	rounds per minute
RT	room temperature
s	second
SDR	short-chain dehydrogenase/reductase
SDS(-PAGE)	sodium dodecyl sulfate (polyacrylamide gel electrophoresis)
Sugp	2-acetamido-2,6-dideoxy-D-xylo-4-hexulose
SWISS-PROT	Swiss protein database
T4/T7	bacteriophage T4/T7
TLR	Toll-like receptor
Tris	tris(hydroxymethyl)-aminomethan
Triton X-100	polyethylene glycol <i>tert</i> -octylphenyl ether
UDP	uridine 5'-diphosphate
UV (light)	ultraviolet light
V	volt
VIS	visible light
$V_{\max}$	maximum velocity
v/v	volume per volume
W	watt
WTA	wall teichoic acid
w/v	weight per volume
2YT broth	double-strength yeast extract-tryptone broth

Decimal multiples and submultiples of units are formed using metric prefixes. Abbreviations for amino acids and nucleotide bases follow the recommendations of the IUPAC-IUB (International Union of Pure and Applied Chemistry & International Union of Biochemistry; [www.chem.qmul.ac.uk/iupac/](http://www.chem.qmul.ac.uk/iupac/)).



## Table of contents

1	Introduction .....	1
1.1	Structural diversity of bacterial capsular polysaccharides.....	1
1.2	Functions of bacterial capsules .....	3
1.3	Bacterial capsular polysaccharides as vaccine agents.....	7
1.4	Biochemistry and genetics of bacterial capsular polysaccharide production.....	9
1.5	Regulation of bacterial capsule expression .....	13
1.6	Capsule biosynthesis in <i>Staphylococcus aureus</i> .....	17
1.7	Short-chain dehydrogenases/reductases .....	22
1.8	Objectives of this work.....	23
2	Materials and methods .....	25
2.1	Materials .....	25
2.1.1	Reagents and chemicals .....	25
2.1.2	Plasmids and oligonucleotides.....	26
2.1.3	Bacterial strains and culture media .....	28
2.1.4	Sterilization procedures .....	29
2.2	Growth, maintenance and preservation of bacterial strains .....	29
2.3	Methods in molecular genetics.....	30
2.3.1	Isolation of genomic DNA from <i>S. aureus</i> .....	30
2.3.2	Preparation of plasmid DNA .....	30
2.3.3	Polymerase Chain Reaction (PCR) .....	30
2.3.4	Agarose gel electrophoresis .....	31
2.3.5	Cleavage of DNA by restriction endonucleases .....	32
2.3.6	Ligation of DNA fragments.....	32
2.3.7	Preparation of CaCl <sub>2</sub> -competent <i>E. coli</i> cells.....	32
2.3.8	Transformation of CaCl <sub>2</sub> -competent <i>E. coli</i> cells .....	32
2.4	Protein and biochemical methods .....	33
2.4.1	Overexpression and purification of hexahistidine fusion proteins .....	33
2.4.2	Sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE).....	35
2.4.3	Determination of protein concentrations according to Bradford.....	35
2.4.4	<i>In vitro</i> synthesis of soluble capsule precursors.....	36
2.4.5	Determination of kinetic parameters for CapD and CapE.....	37
2.4.6	Investigation of CapD and CapE inhibitors .....	37
2.4.7	Capillary electrophoresis analysis of soluble capsule precursors.....	38
2.4.8	CE-based quantification of CapD and CapE reaction products .....	38

2.4.9 MALDI-TOF mass spectrometric analysis of soluble capsule precursors.....	40
2.4.10 <i>In vitro</i> kinase assays .....	41
2.4.11 Identification of phosphorylation sites by nanoLC-MS/MS .....	41
2.4.12 <i>In vitro</i> modulation of CapD and CapE catalytic activity by the CapA1B2 kinase complex.....	42
2.5 <i>In silico</i> methods .....	43
2.5.1 Similarity searches and sequence analysis.....	43
2.5.2 Identification of Suga 4-keto reductases and phylogenetic tree reconstruction .....	43
2.5.3 Genomic context analysis of Suga 4-keto reductases.....	45
3 Results .....	47
3.1 UDP- <i>N</i> -acetylglucosamine 4,6-dehydratases.....	47
3.2 Comparative genomic and phylogenetic analysis of Suga 4-keto reductases.....	49
3.3 Construction of <i>E. coli</i> expression plasmids .....	57
3.4 Expression and purification of recombinant proteins .....	58
3.5 <i>In vitro</i> reconstitution of <i>S. aureus</i> UDP-D-FucNAc biosynthesis .....	59
3.6 Kinetic characterization of CapD- and CapE-catalyzed reactions.....	65
3.7 Screening for inhibitors of CapD and CapE enzymatic activity .....	68
3.8 CapAB-mediated modulation of <i>S. aureus</i> capsule biosynthesis.....	71
4 Discussion.....	75
4.1 Inhibitors of capsule biosynthesis as potential anti-virulence agents .....	75
4.2 Characterization of <i>S. aureus</i> UDP-GlcNAc 4,6-dehydratases.....	76
4.3 Suga 4-keto reductases—a novel protein family .....	81
4.4 <i>In vitro</i> enzymatic synthesis of UDP-D-FucNAc.....	83
4.5 The approved pharmaceutical suramin is a potent inhibitor of CapE .....	85
4.6 A novel mechanism for regulation of capsule biosynthesis.....	88
4.7 Enzymatic checkpoints within capsule biosynthesis are controlled by the tyrosine kinase complex CapAB.....	90
4.8 Future directions .....	92
5 References.....	95
6 Supporting information .....	121
7 Publications.....	137
8 Curriculum vitae .....	139



# 1 Introduction

## 1.1 Structural diversity of bacterial capsular polysaccharides

Polysaccharide capsules are ubiquitous structures found on the cell surface of a broad range of bacterial species. Capsular structures often constitute the outermost layer of the bacterial cell and may as such mediate interactions between the microorganism and its immediate environment; they fulfill various biological functions and were identified as important factors in the virulence of numerous Gram-positive (e.g. *Streptococcus pneumoniae*, *Staphylococcus aureus*) and Gram-negative (e.g. *Neisseria meningitidis*, *Haemophilus influenzae*, *Escherichia coli*) pathogens.<sup>1–5</sup>

Bacterial capsular polysaccharides (CPs) are structurally diverse molecules composed of monosaccharide subunits connected via glycosidic bonds; they can either be homo- or heteropolymers, the latter normally being made up of repeating units comprising two to six sugar residues.<sup>6</sup> The wide variation in structure is caused not only by the diversity of monosaccharide compositions, but also by the diversity of linkages between the sugars subunits, which may be joined in a number of configurations. Further structural complexity arises from the introduction of branches into the polysaccharide chain, variation of the modal chain length, or from modification with non-sugar moieties such as *O*-acetyl groups or amino acid residues (for more details, readers may refer to [www.glycome-db.org](http://www.glycome-db.org)).

Capsule polymers are highly flexible, extend radially from the cell wall, and are rarely cross-linked to one another by strong bonding forces.<sup>7</sup> It is assumed that the stability of capsular structures is largely due to electrostatic, divalent cation-dependent interactions within the often highly acidic (i.e. rich in carboxyl groups) polymer.<sup>7</sup> CPs of Gram-positive organisms can be covalently linked to different cell envelope components, with attachment to the peptide moiety of peptidoglycan (PG) for *Bacillus anthracis*, attachment to *N*-acetylglucosamine (GlcNAc) of PG for *Streptococcus agalactiae*, and covalent attachment to the PG GlcNAc moiety or to the membrane for *S. pneumoniae*.<sup>8–10</sup> Experimental evidence suggests, that the CPs of Gram-negative bacteria may be anchored to the cell surface by covalent linkages between their reducing ends and either lipid A (in the case of *E. coli* K<sub>LPS</sub>, see 1.5) or phospholipid molecules.<sup>11–14</sup> However, other mechanisms (e.g. electrostatic interactions) might also contribute to CP retention in Gram-negative organisms.<sup>15</sup> The tight association with the cellular surface has been used as a criterion to differentiate

CPs from exopolysaccharides (EPSs), which are secreted by bacteria to form a slime layer or a biofilm matrix. This broad distinction is only approximate, since CPs may also be released into the growth medium.<sup>16</sup> Furthermore, certain EPS molecules seem to be tightly associated with the bacterial cell surface in the absence of any detectable cell envelope anchoring.<sup>17</sup>

TABLE 1

Capsular serotypes of selected pathogenic bacteria

Species	Number of capsular serotypes	Serotypes associated with invasive disease <sup>a</sup>
<i>Bacillus anthracis</i>	1 (polyglutamic acid)	
<i>Escherichia coli</i>	80	K1, K5
<i>Haemophilus influenzae</i>	6 (a–f)	b
<i>Klebsiella pneumoniae</i>	77	K1, K2
<i>Neisseria meningitidis</i>	12	A, B, C
<i>Pseudomonas aeruginosa</i>	1 (alginate)	
<i>Salmonella enterica</i>	1 (Vi antigen)	
<i>Staphylococcus aureus</i>	4 (1, 2, 5, 8)	5, 8
<i>Streptococcus pneumoniae</i>	93	children: 4, 6B, 9V, 14, 18C, 19F, 23F adults: 1, 3, 4, 6, 7, 8, 9, 12, 14, 18, 23
<i>Streptococcus pyogenes</i>	1 (hyaluronan)	

Note. Adapted from Cartee, R.T. & Yother, J. (2006) Capsules. In C.A. Nickerson & M.J. Schurr (editors), *Molecular Paradigms of Infectious Disease: A Bacterial Perspective* (pages 138–174).

<sup>a</sup>indicated for species that have more than one serotype

A large number of different capsule types have been identified in human pathogens by serotyping, but only some of the corresponding CP primary structures have been elucidated biochemically. Various capsule types can be found in one bacterial species (TABLE 1). For example, more than 80 antigenically distinct CP types, also referred to as K antigens (K is short for the German word *Kapsel*), have been identified in *E. coli*, so far.<sup>18</sup> In contrast, only one capsular serotype has been described for *Salmonella*.<sup>19</sup> Except in rare circumstances, a given bacterial cell contains the genes for, and expresses only, one type of CP.<sup>10</sup> Serotyping of bacterial clones is of great importance in epidemiology as it often represents the simplest tool to monitor the spread of a pathogen in an affected community.<sup>20</sup> In addition, the capacity to cause certain human diseases may be limited to a few capsular serotypes of a bacterial species. For instance, *E. coli* strains expressing the K1 antigen are the major cause of Gram-negative bacillary meningitis in neonates.<sup>21</sup>

## 1.2 Functions of bacterial capsules

Bacterial polysaccharide capsules are involved in various functions and processes relevant to pathogenicity, including prevention of desiccation, adherence, and resistance to non-specific and specific host immune defense.<sup>22</sup> Bacterial capsules form a highly hydrated gel-like polymer that may protect the cell from the harmful effects of desiccation.<sup>23</sup> Muroid (i.e. encapsulated) strains of *E. coli* exhibit significantly prolonged survival under conditions of dehydration, compared to isogenic non-muroid mutants.<sup>24</sup> The enhanced survival of capsule-bearing bacteria, also observed in other species, may be particularly relevant for host-to-host transmission.<sup>24,25</sup> These findings may furthermore explain the positive correlation between high osmolarity and CP production seen with many species.<sup>26,27</sup>

The influence of CP expression on bacterial adherence and biofilm formation seems to depend on the polysaccharide type and the organism. In *Staphylococcus epidermidis*, the ability to establish a biofilm and to colonize abiotic surfaces (e.g. catheters and other medical devices) was found to be strongly correlated with the ability to cause infections in a clinical setting.<sup>28</sup> The capsular polysaccharide/adhesin of *S. epidermidis* was demonstrated to mediate initial cell–surface interactions crucial for biofilm development by this organism.<sup>29</sup> In contrast, a negative influence on biofilm formation was observed for several *S. pneumoniae* CP types.<sup>30</sup> Furthermore, downregulation of pneumococcal capsule expression was shown to enhance host cell invasion during establishment of the asymptomatic carrier state, although the capsule must be restored for survival after invasion.<sup>31,32</sup> The reduced capsule thickness leads to exposure of adhesins otherwise shielded by the CP molecules, which promote adhesion to host cells and subsequent pathogen uptake.<sup>31,33</sup> The significance of a properly modulated CP biosynthesis has also been demonstrated in case of *N. meningitidis* infections; the “phase-off” state of capsule production enhances host tissue invasion, whereas the “phase-on” state is essential for survival in systemic infections.<sup>34,35</sup>

Most bacteria causing invasive disease produce CPs, which serve as essential immune evasion factors enabling bacterial persistence in the blood-stream of infected hosts, or abscess formation.<sup>4,36–39</sup> Bacterial capsules play an important role in protecting microbes against complement attack.

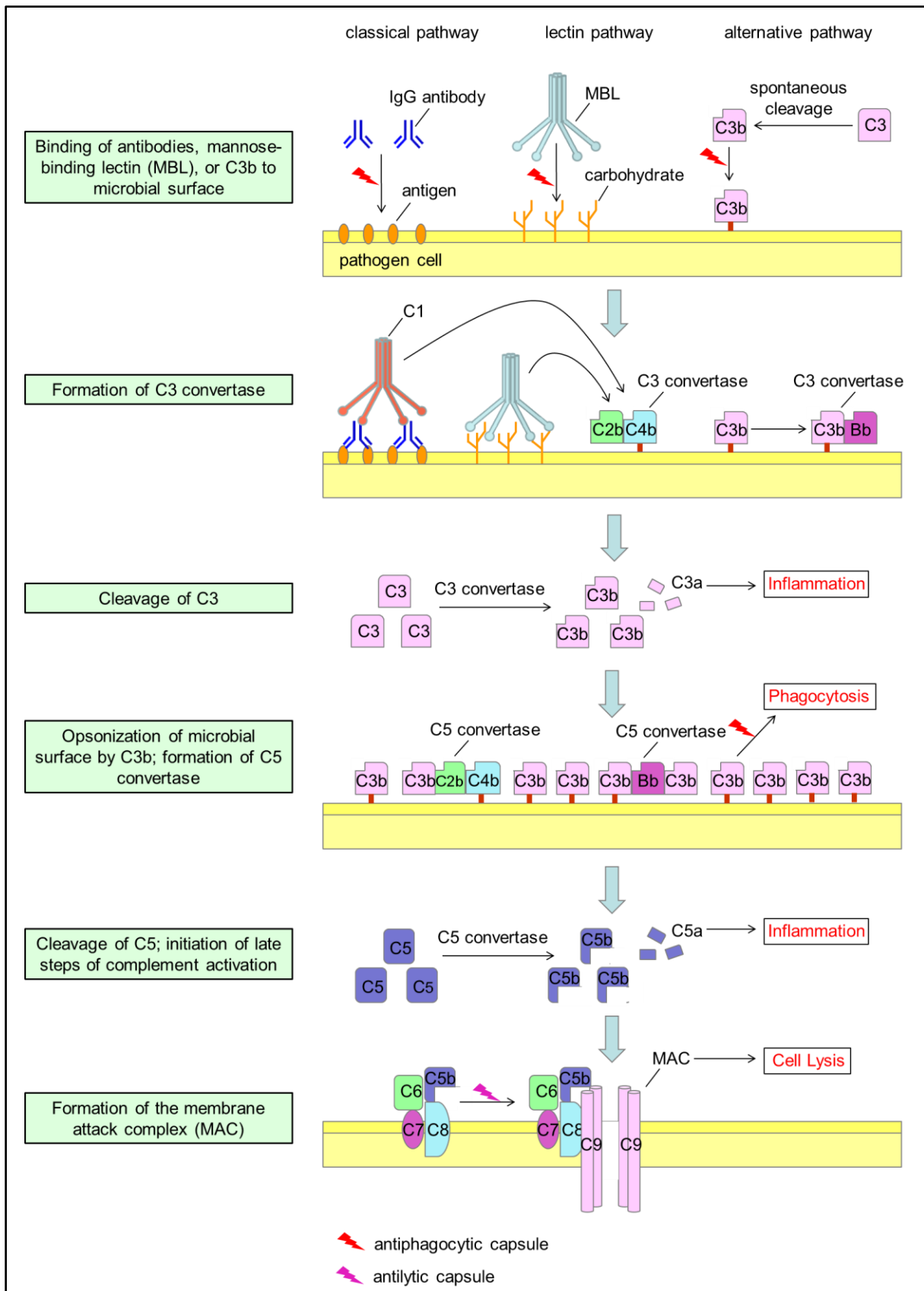


FIGURE 1. Bacterial capsular polysaccharides contribute to complement resistance. Antiphagocytic capsules inhibit activation/deposition of complement, or block access of phagocytic receptors to deposited complement components. Antilytic capsules prevent lysis of Gram-negative bacteria by the complement C5b–9 membrane attack complex.<sup>40</sup>

The complement is a complex system of circulating and cell-surface-bound proteins, which serve as substrates, enzymes, or modulators of a hierarchical series of extracellular proteolytic cascades (FIGURE 1). This system contributes to microbial clearance either by direct killing through formation of a lytic membrane attack complex (only Gram-negative bacteria), or by opsonization of microbes enabling phagocytic uptake (for reviews, see references <sup>41,42</sup>). The complement system represents a functional bridge between innate and adaptive immune responses, and can be activated by antigen-antibody complexes (classical pathway), certain carbohydrates (lectin pathway), or by various surfaces that are not decorated with natural inhibitors (alternative pathway). All three pathways culminate in the cleavage of C3 and in the generation of the principal complement opsonins C3b and iC3b (degradation fragment of C3b). Opsonization of foreign surfaces by C3b enables phagocytosis and leads to the assembly of C5 convertases, which cleave C5 to form C5a and C5b. The C5b fragment associates with components C6, C7, C8, and multiple C9 molecules. On membrane surfaces where the formation of polymeric C9 is not inhibited, such as on the membranes of serum-sensitive Gram-negative bacteria, membrane attack complexes are formed, and the collapse of membrane potential leads to cell death.

Polysaccharide capsules may contribute to complement resistance by a variety of mechanisms (FIGURE 1). Most bacterial CPs are poor activators of the alternative complement pathway.<sup>43</sup> The expression of such polysaccharides may serve to provide a permeability barrier to complement components physically masking underlying surface structures, that would otherwise act as potent activators of the alternative pathway (e.g. cell wall polymers or LPS).<sup>1</sup> Similarly, encapsulation may inhibit the recognition of surface antigens by specific IgG antibodies and the cell surface-binding of mannose-binding lectin, thereby preventing activation of the classical pathway and the lectin pathway, respectively.<sup>44-46</sup> For instance, encapsulated strains of *S. aureus* are not agglutinated by antibodies direct against PG and WTA, which play an important role in the phagocytic clearance of non-encapsulated *S. aureus*.<sup>44</sup> However, even if the complement system is activated and opsonins (C3b, iC3b) are formed, they may become embedded deep in the capsular network and be inaccessible to the respective cell surface receptors on polymorphonuclear leukocytes.<sup>47-49</sup> The accessibility may be restored in the presence of anticapsular antibodies, which direct the deposition of C3 fragments to

the capsule itself.<sup>48,50</sup> For example, acapsular mutants of *S. aureus* are efficiently opsonized for phagocytosis by non-immune serum, whereas phagocytosis of encapsulated strains requires the presence of both, specific anti-CP antibodies and complement.<sup>4,44</sup> In Gram-negative organisms, capsule formation is furthermore thought to confer protection against direct killing, by preventing correct insertion of the cytolytic membrane attack complex (C5b–9) into the outer membrane. Ward & Inzana demonstrated that the *Actinobacillus pleuropneumoniae* CP confers protection against complement-mediated lysis by limiting the deposition of C9 on the bacterial cell surface, although it does not hinder the early steps of complement activation.<sup>51</sup>

Besides being major complement resistance factors, bacterial capsules may also contribute to immune evasion by impairing Toll-like receptor (TLR)-mediated pathogen recognition. The TLRs are germ-line encoded pattern recognition receptors that sense conserved microbial structures, so called pathogen-associated molecular patterns (for a review, see reference<sup>52</sup>). TLR signaling plays an essential role in the host defense to microbial infections, not only by enhancing innate immune responses, but also by providing a link to the adaptive immune system. Activation of TLRs on phagocytes (e.g. macrophages, dendritic cells) stimulates the production of microbicidal substances (e.g. reactive oxygen species, nitric oxid, antimicrobial peptides) and induces the expression of proinflammatory cytokines, chemokines, and co-stimulatory molecules, which subsequently control the activation of antigen-specific adaptive immune responses. Some polysaccharide capsules may downregulate the host immune response by shielding of the bacterial surface, and thus preventing TLR-mediated recognition of pathogen-associated molecular patterns.<sup>46</sup> Certain CPs, however, seem to interfere with TLR-mediated immune responses in a more specific way. It has been suggested that the type C CP of *N. meningitidis* inhibits LPS-induced activation of TLR4 by direct binding to the accessory molecule CD14.<sup>53</sup> The unusual O-methyl phosphoramidate (MeOPN) modification, found on about 70% of *C. jejuni* CPs,<sup>54</sup> has been shown to modulate the host cytokine response by interfering with TLR signaling. The MeOPN modifications on the *Campylobacter jejuni* NCTC 11168 CP were demonstrated to reduce the TLR4-mediated release of the proinflammatory cytokines IL-6 (interleukin-6), IL-10 (interleukin-10) and TNF (tumor necrosis factor) from mouse dendritic cells.<sup>55</sup> MeOPN-mediated downregulation of TLR activation and cytokine production has also

been observed with the structurally distinct *C. jejuni* 81-176 CP, suggesting an active and specific role for the MeOPN moieties in interfering with TLR signaling.<sup>56</sup>

In contrast to the majority of CPs, which are capable of eliciting an antibody response, certain CPs are poorly immunogenic due to similarities to host-associated carbohydrate structures; expression of such CPs does therefore confer some measure of protection against the host's specific immune response.<sup>25</sup> For example, the polysialic acid capsule of *N. meningitidis* serogroup B and *E. coli* K1 strains is antigenically similar to structures expressed on human foetal neuronal cells, and consequently poorly immunogenic.<sup>57,58</sup> Another example of such a molecular mimicry is the *E. coli* K5 antigen strongly resembling human desulpho-heparin.<sup>59</sup>

### 1.3 Bacterial capsular polysaccharides as vaccine agents

Bacterial capsular polysaccharides have the capacity to elicit type-specific antibody responses that confer protection against invasive disease by enabling opsonophagocytic killing of encapsulated bacteria. Thus, they are attractive candidates for vaccine development (for a review, see Weintraub<sup>6</sup>). Moreover, CPs are generally non-toxic and have none of the deleterious adverse effects associated with whole-cell vaccines.<sup>6</sup> Today, several vaccines based on CPs are successfully used in the prevention of infectious diseases (TABLE 2), such as *H. influenzae* serotype B meningitis in children. To overcome the poor immunogenicity of polysaccharide vaccines in infants, conjugate vaccines consisting of polysaccharides linked to highly immunogenic protein carriers (e.g. diphtheria or tetanus toxoids) have been developed.<sup>60,61</sup> Such protein-polysaccharide conjugates act as T cell-dependent antigens.<sup>62</sup> The differences between polysaccharide vaccines and conjugate vaccines are most marked for infants and immunocompromised persons, with conjugate vaccines inducing much stronger immune responses.<sup>63,64</sup> The design of peptides (i.e. T cell-dependent antigens) that mimic polysaccharide antigens may represent a further strategy to overcome T cell independency.<sup>65</sup>

Many vaccine development projects have faced the problem of antigenic variability, which arises from the wide structural heterogeneity among the polysaccharides within and between species (see section 1.1), and the fact that anti-polysaccharide antibodies are usually serotype/serogroup specific. This problem has been approached by multivalent vaccine formulations, which raise immunity individually targeted against the clinically most important capsule types. For example, the

polysaccharide vaccine Pneumovax<sup>®</sup> 23 (see TABLE 2) contains purified CPs from the 23 most prevalent *S. pneumoniae* serotypes, which are accountable for approximately 90% of all pneumococcal infections ([www.who.int](http://www.who.int)).

TABLE 2  
Licensed vaccines based on bacterial capsular polysaccharides

Indication	Trade Name(s) <sup>®</sup>	Vaccine type
<u><i>Haemophilus influenzae</i></u>		
serotype b	ActHIB, Hiberix	1-valent conjugate vaccine <sup>a</sup>
<u><i>Neisseria meningitidis</i></u>		
serogroup C	Meningitec, Menjugate	1-valent conjugate vaccine <sup>b</sup>
	NeisVac-C	1-valent conjugate vaccine <sup>a</sup>
serogroups A, C	Meningokokken- Impfstoff A+C Mérieux	2-valent polysaccharide vaccine
serogroups A, C, W-135, Y	Mencevax ACWY	4-valent polysaccharide vaccine
	Menveo	4-valent conjugate vaccine <sup>b</sup>
	Nimenrix	4-valent conjugate vaccine <sup>a</sup>
<u><i>Salmonella enterica</i></u>		
serovar Typhi	Typhim Vi, Typherix	1-valent polysaccharide vaccine
<u><i>Streptococcus pneumoniae</i></u>		
serotypes 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, 23F	Synflorix	10-valent conjugate vaccine <sup>c</sup>
Synflorix serotypes + 3, 6A, 19A	Prevenar 13	13-valent conjugate vaccine <sup>b</sup>
Prevenar-13 serotypes + 2, 8, 9N, 10A, 11A, 12F, 15B, 17F, 20, 22F, 33F	Pneumovax 23	23-valent polysaccharide vaccine

*Note.* The complete list of vaccines licensed for immunization and distribution in Germany is available at the Paul-Ehrlich-Institute website ([www.pei.de](http://www.pei.de)).

<sup>a</sup>tetanus toxoid as carrier protein

<sup>b</sup>non-toxic mutant of diphtheria toxin (CRM<sub>197</sub>) as carrier protein

<sup>c</sup>*H. influenzae* protein D + diphtheria toxoid + tetanus toxoid as carrier proteins

In some cases, the mimicry of host-associated carbohydrate structures by bacterial polysaccharides represents a key obstacle to vaccine development. Perhaps the best example of this is the meningococcal serogroup B CP, which is structurally identical to the polysialic acid chains found on an important nervous system developmental protein—the neural cell adhesion molecule (N-CAM).<sup>57,58</sup> The search for an effective vaccine against serogroup B meningococcus, which is the most important cause of endemic meningitis in industrialized countries, has been complicated by the inability of its CP to induce a significant antibody response, even when conjugated to a carrier



protein.<sup>66</sup> Furthermore, the use of this polysaccharide in a vaccine has the potential risk of inducing autoantibodies.<sup>58,67</sup> In such cases, other surface-exposed molecules seem to be the more promising candidate antigens for vaccine development.<sup>68</sup> Only recently, Bexsero<sup>®</sup> (Novartis Vaccines and Diagnostics) was licensed for the prevention of serogroup B meningococcal disease, a vaccine that contains three recombinant proteins and outer membrane-vesicles derived from a *N. meningitidis* group B strain as immunogenic components.<sup>69</sup> In the case of other clinically important pathogens, the search for an effective vaccine is still ongoing. For instance, Pfizer's SA4Ag, a multiantigen vaccine for prevention of *S. aureus* disease, which contains CPs from the two most prevalent serotypes, as well as two protein antigens, is currently in Phase II development (<http://clinicaltrials.gov/show/NCT01364571>).<sup>70</sup>

## 1.4 Biochemistry and genetics of bacterial capsular polysaccharide production

Despite the remarkable diversity in bacterial oligo- and polysaccharide structures (see section 1.1), conserved themes are evident in the biosynthesis-export pathways. Central to the understanding of these processes has been the elucidation of LPS O antigen biosynthesis in Gram-negative bacteria, and the identification of three general assembly-export strategies, termed “Wzy-dependent”, “synthase-dependent”, and “ATP-binding cassette (ABC) transporter-dependent”;<sup>71</sup> the bacterial systems for capsule synthesis can be grouped into the same categories (FIGURE 2; reviewed in references<sup>10,14</sup>). The Wzy-dependent pathway of capsule biosynthesis is widespread among both, Gram-negative and Gram-positive bacteria. The respective polymers usually have complex repeating units comprising different sugars and, often, side chains.<sup>72,73</sup> The CPs of essentially all Gram-positive organisms, including those of staphylococci and most streptococci, as well as the group 1 and group 4 capsules of Gram-negative bacteria, are produced via the Wzy-dependent pathway.<sup>72,74–77</sup> Wzy-dependent CP synthesis mechanistically resembles PG biosynthesis and involves the formation of lipid-linked intermediates. Basic repeating units are built on the isoprenoid lipid carrier undecaprenyl-phosphate (C<sub>55</sub>P), by sequential glycosyltransferase reactions occurring at the inner face of the cytoplasmic membrane.<sup>75</sup>

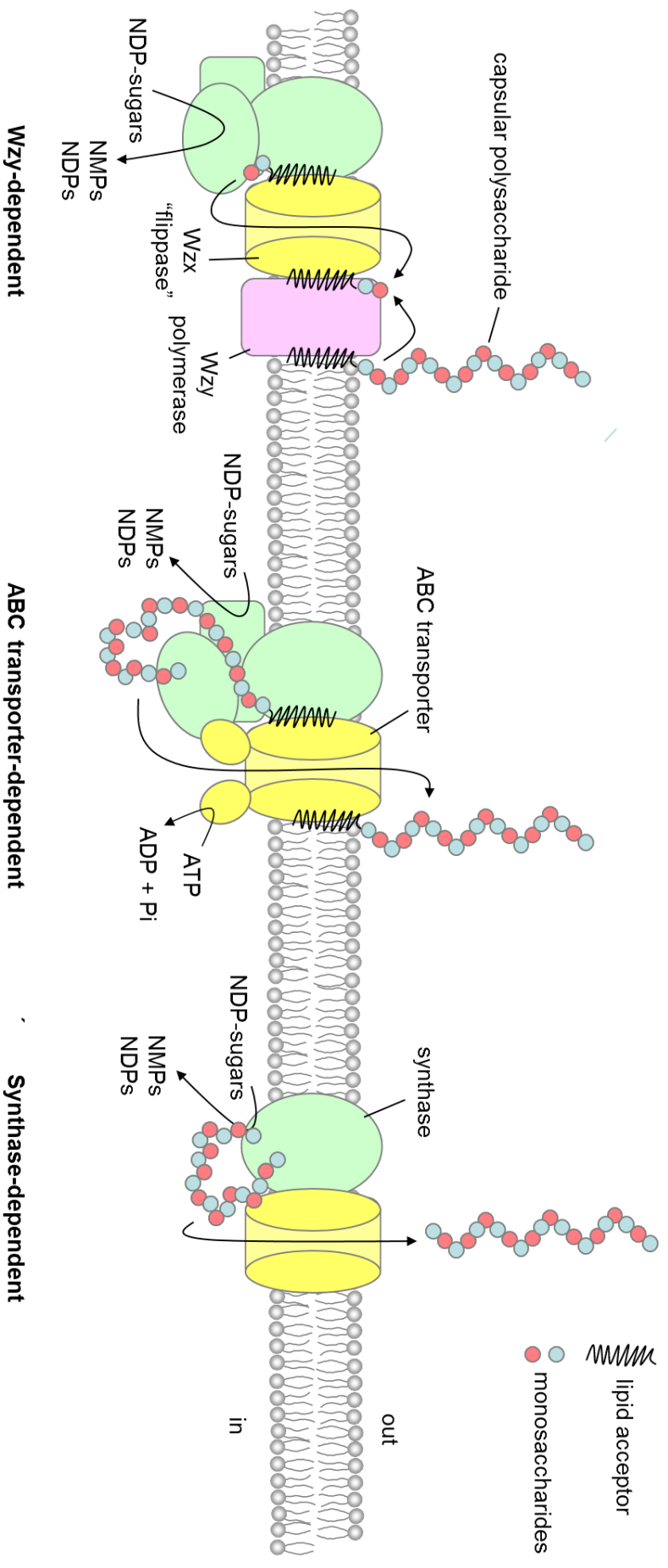


FIGURE 2. Models for the three general assembly/export mechanisms used in the biosynthesis of bacterial capsular polysaccharides (CPs) (adapted from Cutthbertson et al.<sup>78</sup>). In the Wzy-dependent pathway, basic repeat units are synthesized on a lipid carrier by glycosyltransferases (illustrated in green in each model) located at the inner face of the cytoplasmic membrane. The lipid-linked CP building blocks are exported by a process involving a Wzx "flippase", and are polymerized at the outer (periplasmic) surface of the cell membrane by a Wzy polymerase, through growth at the reducing end. In the ABC transporter-dependent pathway, chain elongation occurs by addition of monomers to the non-reducing terminus of a lipid-linked intermediate. The chain is fully polymerized in the cytoplasm and exported by an ABC transporter. In the synthase-dependent pathway, a single protein—the synthase—may catalyze both, chain assembly and export. There is no unifying involvement of a lipid acceptor in the synthase pathway, and the direction of chain growth may differ, depending on the system. In Gram-negative bacteria, additional proteins are required to enable transport of the CP molecules across the outer membrane. NDP, nucleoside diphosphate; NMP, nucleoside monophosphate; Pi, inorganic phosphate.

Once the basic repeating unit is assembled, the polyprenol-linked intermediates are exported across the membrane by a process involving a Wzx “flippase”.<sup>79</sup> The lipid-linked repeat units are polymerized at the outer surface of the cell membrane by a Wzy polymerase, through growth at the reducing end.<sup>80</sup>

In contrast, CP molecules synthesized via ABC transporter-dependent pathways are fully polymerized in the cytoplasm, whereby growth of the polysaccharide chain at the non-reducing end is catalyzed either by a single glycosyltransferase, or by the concerted action of several glycosyltransferases.<sup>81–83</sup> The completed polymer is exported across the membrane via a pathway-defining transporter of the ABC-2 family.<sup>84</sup> So far, ABC transporter-dependent capsule synthesis has only been described in Gram-negative bacteria. The *E. coli* K1 and K5 capsules, as well as CP molecules from a number of other pathogens, including *C. jejuni*, *N. meningitidis* and *H. influenzae*, are synthesized by this mechanism.<sup>85–89</sup> ABC transporter-dependent polymers tend to be unbranched, simple structures.<sup>14</sup> However, this trend does not apply for all species. For example, *C. jejuni* NCTC 11168 produces complex, branched CP molecules that are exported by an ABC transporter.<sup>86,90</sup> Initiation of CP biosynthesis in ABC transporter-dependent systems remains poorly understood. Neither the acceptor used for chain extension, nor the enzymes responsible for initiation, have been identified. CPs assembled via ABC transporter-dependent pathways bear terminal phospholipid moieties at their reducing ends, which likely serve as outer membrane anchors.<sup>11–14</sup> In *E. coli* and *N. meningitidis*, the predominant lipid is a lysophosphatidylglycerol moiety that is attached to the repeat-unit domain of the CP via multiple residues of 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo).<sup>14</sup> The precise point in the biosynthetic pathway at which the phospholipids are added is unknown, though it has been suggested that they might represent the endogenous acceptors for ABC transporter-dependent polymer assembly.<sup>14,91</sup> However, the glycan chain might also be built on a different acceptor molecule (e.g. C<sub>55</sub>P) and transferred to the phospholipid anchor prior to export.<sup>73</sup>

Only a few capsules composed of relatively simple polysaccharides are produced via the synthase-dependent pathway, such as the *S. pneumoniae* type 3 capsule, and the *Streptococcus pyogenes* and *Pasteurella multocida* hyaluronic acid (HA) capsules.<sup>92–94</sup> These CPs comprise either a monosaccharide or a heterodisaccharide repeat unit, and require only a single glycosyltransferase—the synthase—for initiation and polymerization.<sup>10,95</sup> Bifunctional synthases, which catalyze the addition of two

different sugars and/or the formation of different glycosidic bonds, represent an exception to the usual “one enzyme, one sugar linkage” dogma of glycobiology.<sup>96</sup> Such dual-action glycosyltransferases are involved in the synthesis of some important biopolymers, including bacterial cellulose, alginate, hyaluronan, and chondroitin.<sup>97–100</sup> In the case of the *S. pneumoniae* serotype 3 synthase, the ability to transfer two different monosaccharides seems to involve a single carbohydrate-binding site that recognizes two different UDP-sugars.<sup>101</sup> In contrast, the *P. multocida* HA synthase comprises two independent glycosyltransferase sites, as revealed by mutation analysis.<sup>93</sup> Most bacterial polysaccharide synthases are integral membrane proteins that share a similar topology with four transmembrane domains and a large cytoplasmic loop that harbours the glycosyltransferase activity.<sup>102–105</sup> The nature of the acceptor on which the polymer is built is often ambiguous, as well as the process by which the polymer is exported. It has been demonstrated that *S. pneumoniae* utilizes phosphatidylglycerol as lipid acceptor for the synthesis of type 3 CP, whereas HA assembly in *Streptococcus equisimilis* does not seem to involve lipid-linked intermediates.<sup>106,107</sup> Several synthases are proposed to be sufficient for both, polymerization and export.<sup>94,108</sup> To date, the only synthase for which polysaccharide export activity has been definitely proven is the HA synthase from *S. equisimilis*.<sup>107</sup> However, the exact mechanism through which synthases mediate extrusion of nascent polymers remains elusive. In other synthase-dependent systems, additional components may be required for CP transport across the cytoplasmic membrane. For example, an ABC transporter has been implicated in HA export in *P. multocida*.<sup>109</sup> The capsule biosynthetic genes, which encode enzymes involved in the production of activated nucleotide sugar precursors and the respective glycosyltransferases, generally form a cluster on the bacterial chromosome (often named *cps/kps/cap* for “capsule” or *eps/exo* for “exopolysaccharide”).<sup>110</sup> These clusters furthermore comprise genes required for CP polymerization and export, and genes with regulatory functions (as described in the following text and in section 1.5). Within mechanistic types (Wzy-dependent, synthase-dependent, ABC transporter-dependent), the genetic loci for CP production demonstrate similar organization;<sup>10,95,111</sup> genes encoding for pathway-defining proteins (i.e. Wzx “flippases”, Wzy polymerases, transmembrane and nucleotide-binding domains of ABC-2 transporters) can be predicted by their putative membrane topologies, or by the presence of conserved sequence motifs.<sup>84,112,113</sup> In Gram-negative bacteria, the

export apparatus used in the synthase-dependent pathway is clearly distinct from the OPX (outer membrane polysaccharide export) and PCP (polysaccharide co-polymerase) proteins required for the secretion of Wzy- and ABC transporter-dependent polymers.<sup>114,115</sup> It comprises a tetratricopeptide repeat (TPR)-containing scaffold protein, which is thought to protect the polymer from degradation once it reaches the periplasm, and a  $\beta$ -barrel protein, that forms a pore in the outer membrane.<sup>116,117</sup> Due to their predictability by sequence analysis programs, these protein families serve as identifiable hallmarks of synthase-dependent systems in Gram-negative bacteria.<sup>95</sup> Most of the enzymes necessary for CP synthesis are encoded in the polysaccharide-specific loci. However, the lipid acceptors and many of the nucleotide-diphospho-sugar precursors are common to other biosynthetic pathways (e.g. PG, WTA and LPS biosynthesis) and are obtained from cellular pools without the need for unique enzymes.<sup>10,118</sup>

## 1.5 Regulation of bacterial capsule expression

The fact, that the biosyntheses of CP and PG share a common pool of limited precursors, makes a well-orchestrated temporal and spatial regulation of these processes essential for the viability of the cell.<sup>119–124</sup> This hypothesis is supported by the frequent observation that growth and division reduce CP production, while growth conditions which reduce the demand for new cell wall synthesis tend to enhance CP production.<sup>125–127</sup> Furthermore, it has long been appreciated that CP expression must be modulated in order to ensure a balance between the protective functions of the capsule and the surface exposure of molecules required for substrate-binding and/or colonization.<sup>128</sup> For instance, bacteria that colonize the nasopharynx may reduce CP production to expose adhesins, which are important for interactions with the host epithelium, but would increase capsule production when in the bloodstream to ensure protection against opsonophagocytosis (see section 1.2).

Temperature provides a key signal for transcriptional regulation of capsule genes in many organisms, and this may serve to limit capsule production to the environments in which it will confer a selective advantage.<sup>128</sup> In *E. coli*, the expression of group 2 capsule genes is driven by two thermoregulated promoters, with maximal CP production occurring at or near 37°C, while capsule expression is negligible at temperatures below 20°C.<sup>129,130</sup> Osmolarity is another environmental factor influencing CP expression. For example, the expression of the Vi (capsular) antigen

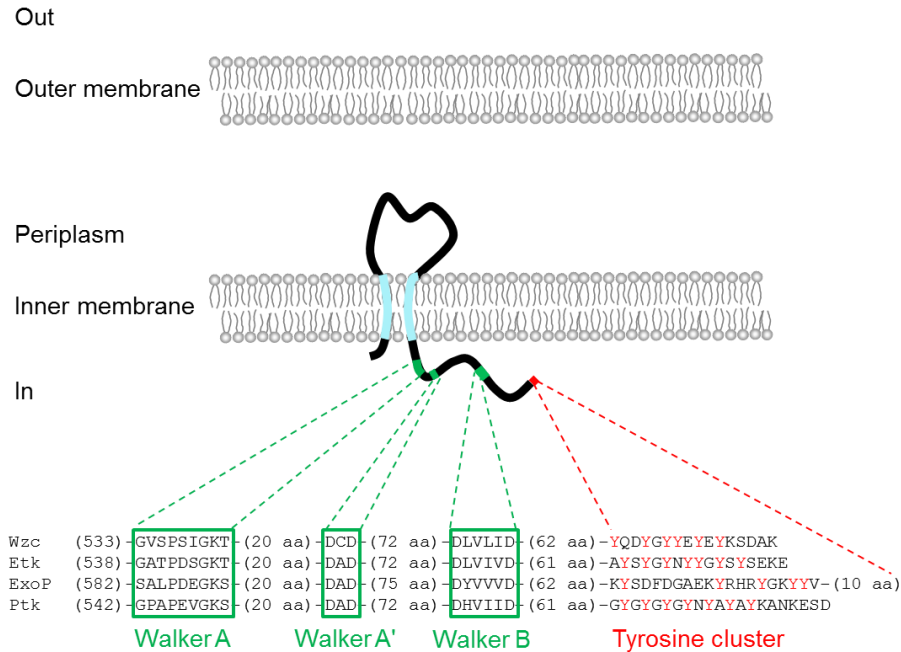
in *Salmonella enterica typhi* is known to be inhibited by high concentrations of NaCl (> 300 mM), such as those found in the human gut.<sup>131,132</sup> This inhibition enables the release of invasion-promoting Sip proteins, which, whilst the bacterium is expressing a capsule, are not secreted, but stockpiled in the cell.<sup>131</sup> A positive correlation between CP expression and osmolarity is seen with many other species, indicating a potential function of the capsule in the prevention of desiccation (see section 1.2).<sup>26,27</sup> Further physicochemical parameters were identified as signals for modulation of bacterial CP expression, such as Mg<sup>2+</sup> and iron levels, pH, and oxygen availability.<sup>125,133–136</sup> The CrgA transcriptional regulator of *N. meningitidis* represents an example how more specific, host-cell derived stimuli may be integrated in the regulation of capsule expression. This regulator, which acts as repressor of capsule gene expression, is induced upon intimate contact with epithelial cells, thus reducing meningococcal capsule production.<sup>137</sup>

Given the substantial energy investment that CP biosynthesis and export represent to bacteria, and that the formation of a capsule is not appropriate for all circumstances (see above), it is not astonishing that capsule expression is tightly regulated and controlled at multiple levels.<sup>138</sup> Besides regulation through phase-variation or transcription factors, bacterial CP expression is controlled on the post-translational level enabling short-term modulation.<sup>34,139–143</sup> The post-translational control is exerted by tyrosine kinase phosphoregulatory systems, which are associated with Wzy- and synthase-dependent polymerization systems of both, Gram-positive and Gram-negative bacteria.<sup>115</sup>

Most protein tyrosine kinases found in bacteria, including those involved in the regulation of CP and EPS production, belong to the family of bacterial tyrosine (BY-)kinases.<sup>144</sup> BY-kinases are structurally and functionally unrelated to their eukaryotic counterparts.<sup>145–147</sup> At first they were believed to be exclusively autophosphorylating enzymes functioning as “co-polymerases” in the biosynthesis of CP and EPS.<sup>115</sup> However, in recent times, members of the BY-kinase family were shown to phosphorylate endogenous protein substrates, such as biosynthetic enzymes involved in polysaccharide production, but also RNA polymerase sigma factors and single-stranded DNA binding proteins.<sup>148–150</sup> Therefore, BY-kinases can be considered as promiscuous enzymes involved in diverse biological processes including stress response, DNA metabolism, and cell-cycle control.<sup>148,149,151,152</sup>

*Proteobacteria*

N-terminus TM1 periplasmic region TM2 cytoplasmic domain C-terminus



*Firmicutes*

N-terminus TM1 TM2 C-terminus N cytoplasmic domain C

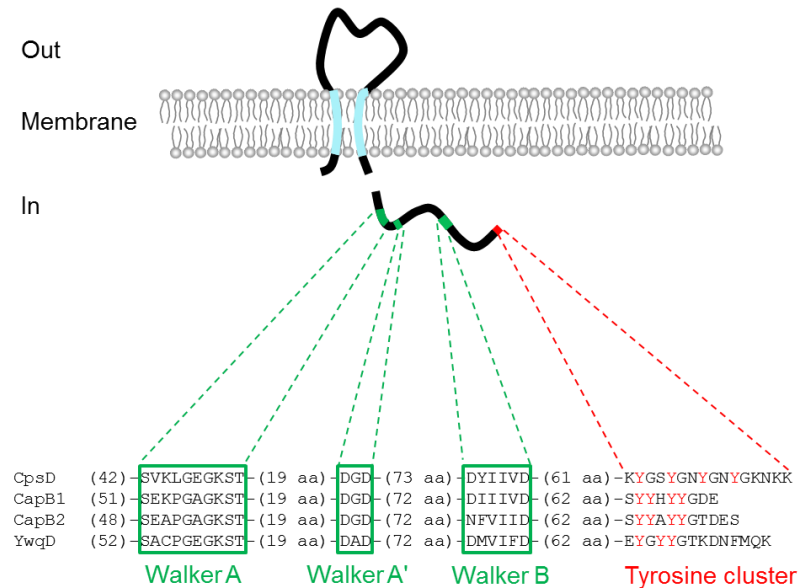


FIGURE 3. Defining structural features of bacterial tyrosine kinases (adapted from Grangeasse et al.<sup>144</sup>). All bacterial tyrosine (BY-)kinases share several common features: an external sensory loop is linked to an intracellular catalytic domain, either directly (*Proteobacteria*) or via specific protein-protein interactions (*Firmicutes*). The catalytic domain contains the Walker-like motifs A, A' and B (green) and a C-terminal tyrosine cluster comprising up to 7 (auto)phosphorylatable tyrosine residues (red). Amino acid sequence alignments are provided for BY-kinases from the following species: *Escherichia coli* Wzc, GenBank accession number CDU38087.1, and Etk, P38134.2; *Acinetobacter johnsonii* Ptk, CAA75431.1; *Sinorhizobium meliloti* ExoP, AAA16042.1; *Streptococcus pneumoniae* CpsD, ADQ39181.1; *Staphylococcus aureus* CapB1, BAF66368.1, and CapB2, BAF68834.1; *Bacillus subtilis* YwqD, CAB15642.1. aa, amino acids; TM, transmembrane domain.

The proteins of the BY-kinase family are characterized by a number of common structural features (FIGURE 3). They possess an external “sensor” domain that is thought to modulate tyrosine kinase activity upon yet unidentified stimuli, and an internal catalytic domain. The catalytic domain encompasses three Walker-like motifs (named A, A' and B) and a C-terminal tyrosine cluster comprising up to seven tyrosine residues, which can be autophosphorylated at the expense of ATP.<sup>145,149,153–157</sup> This autophosphorylation seems to influence the oligomerization state of the BY-kinase.<sup>147</sup> Moreover, the load of tyrosine phosphorylation is supposed to affect the ability of the BY-kinase to interact with other components of the capsule assembly and translocation machinery.<sup>156,158</sup>

The BY-kinases of *Proteobacteria* are encoded in the capsule loci as a single polypeptide; they are composed of an N-terminal membrane-spanning activation domain and a C-terminal cytoplasmic tyrosine kinase domain.<sup>144</sup> In *Firmicutes*, the two domains are located on separate, interacting proteins, which are usually encoded by adjacent genes in the capsule loci.<sup>144</sup> Based on studies in *S. aureus* and *S. pneumoniae*, the cytoplasmic kinase protein alone is not sufficient for phosphotransfer, but has to interact with the C-terminus of the transmembrane adaptor to undergo autophosphorylation.<sup>159–161</sup> In contrast, *Bacillus subtilis* YwqD alone shows tyrosine autokinase activity *in vitro*.<sup>149</sup> Nevertheless, this activity is enhanced in the presence of the cognate transmembrane adaptor YwqC.<sup>149</sup>

The dynamic regulation of cellular processes by protein phosphorylation is achieved through a fine balance of opposing kinase and phosphatase activities. In *Proteobacteria*, the phosphatases involved in the regulation of CP production belong to the family of low-molecular-weight phosphatases (LMW-PTPs), whereas in *Firmicutes*, they belong to the DNA polymerase and histidinol phosphate phosphoesterase (PHP) family.<sup>162–164</sup> The phosphatases are encoded within the capsule loci, by genes located either immediately upstream (both types) or downstream (only PHPs) of those encoding the cognate kinases.<sup>162–164</sup> These enzymes efficiently dephosphorylate BY-kinases and, at least in some cases, the respective BY-kinase protein targets.<sup>162–164</sup>

The autophosphorylation state of BY-kinases, regulated by the corresponding phosphatases, influences the amount, the length and other properties of the polysaccharide produced.<sup>156,160,165,166</sup> In *E. coli* K30, the capsule building blocks can undergo one of two fates. High-level polymerization can occur to generate



high-molecular-weight CP or, alternatively, one to a few repeat units can be ligated onto a lipid A-core and expressed on the cell surface as so-called K<sub>LPS</sub>.<sup>167</sup> Autophosphorylation of the WzC<sub>CP</sub> BY-kinase, as well as its dephosphorylation by the cognate phosphatase Wzb, are both required for assembly of a high-molecular-weight capsule, though not for low-level polymerization or assembly of K<sub>LPS</sub>.<sup>156,166</sup> Similarly, the phosphorylation state of the *Sinorhizobium meliloti* BY-kinase ExoP influences the ratio of low- to high-molecular weight succinoglycan.<sup>165</sup> Current hypotheses suggest that BY-kinases may, depending on their phosphorylation and oligomerization state, interact with the export and polymerization machinery to control the level of CP production and/or CP chain length.<sup>147,168,169</sup> However, their exact mode of function remains enigmatic. Recently, it has been demonstrated that the modulatory effects of BY-kinases are at least partially due to the phosphorylation of biosynthetic enzymes.<sup>149,161,175,176</sup> For instance, the *E.coli* BY-kinase WzC<sub>CA</sub> was shown to phosphorylate the enzyme UDP-glucose dehydrogenase, thereby increasing its catalytic activity *in vitro*, and probably the formation of precursors for polysaccharide production *in vivo*.<sup>175</sup> In addition to modulating enzyme activities, BY-kinase-mediated phosphorylation seems to affect protein targeting.<sup>151</sup>

## 1.6 Capsule biosynthesis in *Staphylococcus aureus*

*Staphylococcus aureus* is an important opportunistic pathogen that may colonize moist skin areas, skin glands, and mucous membranes without causing symptoms of infection.<sup>177</sup> The anterior nares (30% of individuals), perineum (15%) and axillae (2–7%) represent the main habitats for *S. aureus* on healthy humans; other body sites may be colonized as well, such as the vagina, pharynx, rectum, and the toe-web spaces.<sup>178–181</sup> The Gram-positive, coccus-shaped bacterium is capable of causing a broad spectrum of human diseases, such as skin and soft tissue infections, endovascular infections, pneumonia, septic arthritis, endocarditis, osteomyelitis, foreign-body infections, and sepsis.<sup>182</sup> Its pathogenic nature led to the first description of the Gram-positive bacterium over 130 years ago by the Scottish surgeon Sir Alexander Ogston, who identified the organism in human pus and coined the name *Staphylococcus* (derived from the Greek words *staphyle* and *kokkos*, meaning “bunch of grapes” and “berry”, respectively).<sup>183</sup> Since then, it has become clear that *S. aureus* is highly versatile, and adept at acquiring and developing resistance against virtually all antibiotics.<sup>184</sup> The excessive use of antibiotics has

avored the evolution of highly virulent, drug-resistant *S. aureus* strains, thus leading to a rise in the occurrence and severity of infections over the last 20 years.<sup>185,186</sup> Methicillin-resistant *S. aureus* (MRSA) strains, which are resistant to all available penicillins and most other  $\beta$ -lactam antibiotics (with exception of the advanced-generation cephalosporines ceftaroline and ceftobiprole), were first identified in 1961 and soon became endemic in hospitals and health-care facilities worldwide.<sup>186–188</sup> For the following decades, MRSA infections were largely confined to the nosocomial setting, and did predominantly affect patients with health care-associated risk factors, such as hospitalization or residence in a long-term care facility, presence of indwelling medical devices, ulcers or surgical wounds, or chronic dialysis treatment.<sup>186,189–193</sup> Since the mid-1990s, however, there has been a rise in the number of MRSA infections occurring in a non-hospital environment in otherwise healthy individuals.<sup>186</sup> This increase in disease burden has been attributed to so-called “community-associated MRSA” strains, which show enhanced virulence potential, and are capable of causing invasive disease in immunocompetent hosts.<sup>186</sup> Hospital- and community-associated multidrug-resistant MRSA strains, which display resistance to  $\beta$ -lactam antibiotics combined with resistance to other classes of antibiotics, do currently challenge our ability to treat bacterial infections, and evoke fears of an upcoming “post-antibiotic era”.<sup>194–197</sup> Within the health care setting alone, MRSA infections are estimated to affect more than 150,000 patients annually in the European Union, causing more than 5000 attributable deaths, and extra in-hospital costs of € 380 million.<sup>198</sup> In the US, MRSA is the second leading cause of death by infectious disease, with more than 11,000 MRSA-related deaths in 2011 ([www.cdc.gov/drugresistance/threat-report-2013/](http://www.cdc.gov/drugresistance/threat-report-2013/)).

The pathogenic properties of *S. aureus* are due to an armament of virulence factors, ranging from secreted toxins to immune evasion molecules.<sup>199</sup> The expression of a polysaccharide capsule, which confers protection against opsonophagocytosis, contributes substantially to the ability to cause invasive disease (see section 1.2 “Functions of bacterial capsules”); the majority (~80%) of *S. aureus* strains recovered from patients are encapsulated.<sup>200,201</sup> Serotype 5 and 8 capsular polysaccharide types (CP5 and CP8) are dominant among clinical isolates.<sup>200,202</sup> They share similar trisaccharide repeating units, which are identical in monosaccharide composition and sequence, only differing in the glycosidic linkages between the sugars and the sites of O-acetylation.<sup>203–205</sup>

Although polysaccharide structures of CP5 (( $\rightarrow$ 4)- $\beta$ -D-ManNAcA-(1 $\rightarrow$ 4)-3-O-acetyl- $\alpha$ -L-FucNAc-(1 $\rightarrow$ 3)- $\beta$ -D-FucNAc-(1 $\rightarrow$ )<sub>n</sub>) and CP8 (( $\rightarrow$ 3)-4-O-acetyl- $\beta$ -D-ManNAcA-(1 $\rightarrow$ 3)- $\alpha$ -L-FucNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-FucNAc-(1 $\rightarrow$ )<sub>n</sub>) were elucidated decades ago, only little is known about the biochemistry underlying the assembly of these clinically important polymers.

The serotype 5 and 8 capsule gene clusters are allelic and comprise a ~17.5 kb region on the staphylococcal chromosome composed of 16 genes transcribed in one orientation (*cap5(8)A* through *cap5(8)P*).<sup>77,206</sup> Since CP5 and CP8 are built from the same three sugar residues, it is not surprising that 12 of the 16 genes are nearly identical (> 97% sequence identity). The genes responsible for serotype specificity (*cap5HIJK* and *cap8HIJK*) are located in the central regions of the loci, and show little sequence similarity between the two gene clusters. The type-specific regions comprise genes that encode putative Wzx “flippases” (*cap5J*, *cap8I*) and Wzy polymerases (*cap5K*, *cap8K*) (see section 1.4). Furthermore, these regions contain genes encoding putative glycosyltransferases (*cap5I*, *cap8H*) and O-acetyltransferases (*cap5H*, *cap8J*).

Database homology searches with amino acid sequences of *cap5* operon gene products allowed for the prediction of individual enzymatic functions and the proposal of a pathway for capsule biosynthesis in *S. aureus*.<sup>77</sup> Within this Wzy-dependent pathway, synthesis of soluble capsule precursors occurs via three distinct reaction cascades, through which the universal substrate UDP-*N*-acetyl-D-glucosamine (UDP-D-GlcNAc) is converted into the three different nucleotide-coupled sugars UDP-*N*-acetyl-D-fucosamine (UDP-D-FucNAc), UDP-*N*-acetyl-L-fucosamine (UDP-L-FucNAc) and UDP-*N*-acetyl-D-mannosaminuronic acid (UDP-D-ManNAcA) (FIGURE 4). The synthesis of the first soluble precursor UDP-D-FucNAc is allegedly catalyzed in a two-step process by the enzymes CapD and CapN, although experimental evidence is lacking so far. The integral membrane protein CapD is proposed to be a 4,6-dehydratase, that generates the intermediate UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose (UDP-Sugp) by eliminating a water molecule from UDP-D-GlcNAc. This intermediate is further converted to UDP-D-FucNAc by the action of the membrane-associated reductase CapN. Subsequently, CapM is supposed to transfer the *N*-acetyl-D-fucosamine moiety of UDP-D-FucNAc to the membrane-anchored lipid carrier bactoprenol (C<sub>55</sub>P), yielding lipid I<sub>cap</sub>.

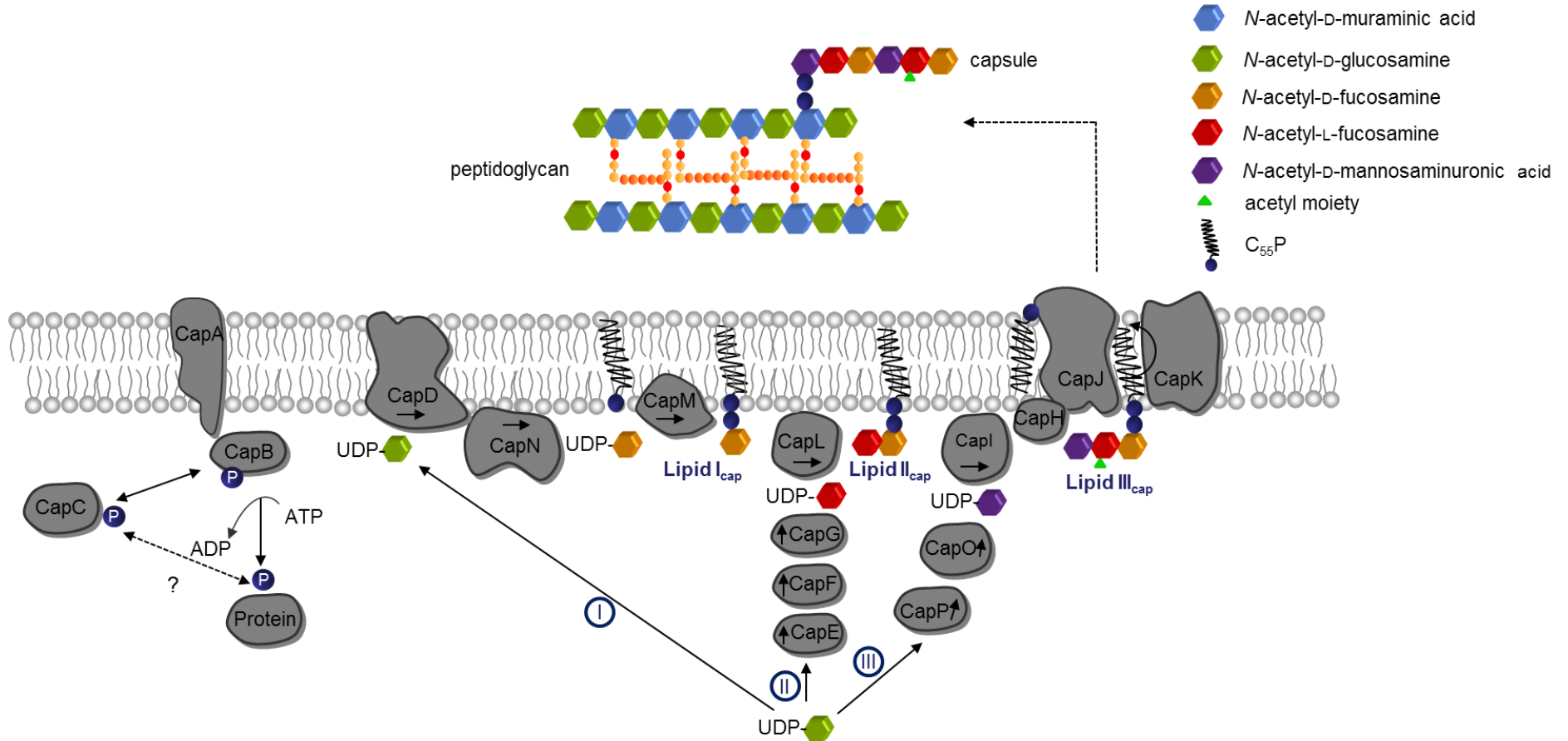


FIGURE 4. Model for the Wzy-dependent capsule biosynthesis pathway in *Staphylococcus aureus* serotype 5. Synthesis of soluble capsule precursors occurs in the cytoplasm via three reaction cascades, through which the universal substrate UDP-*N*-acetyl-D-glucosamine is converted into the three different nucleotide-coupled sugars UDP-*N*-acetyl-D-fucosamine, UDP-*N*-acetyl-L-fucosamine and UDP-*N*-acetyl-D-mannosaminuronic acid. Trisaccharide repeating units are assembled on the lipid carrier undecaprenyl-phosphate (C<sub>55</sub>P), and further modified by *O*-acetylation, before they are exported to the outer surface of the cell membrane where polymerization takes place. The polymerized CP5 is attached to the peptidoglycan; the carrier lipid C<sub>55</sub>P is released and recycled. Post-translational control of capsule biosynthesis is exerted by the bacterial tyrosine kinase complex CapAB. Upon activation by the transmembrane adaptor CapA, the cytoplasmic kinase CapB undergoes autophosphorylation and may phosphorylate target proteins (e.g. biosynthetic enzymes), thus modulating their activity. The phosphorylation state of CapB is regulated by the cognate phosphatase CapC.

The second reaction cascade generating the soluble precursor UDP-L-FucNAc involves the enzymes CapE, CapF and CapG, the enzymatic functions of which have already been elucidated biochemically.<sup>207</sup> The UDP-2-acetamido-2,6-dideoxy-L-talose 2-epimerase CapG finally produces the nucleotide-activated sugar UDP-L-FucNAc, which is assumed to be attached to lipid I<sub>cap</sub> by the transferase CapL, leading to the formation of the second capsule lipid intermediate, lipid II<sub>cap</sub>. The third nucleotide-activated monosaccharide required for CP5 production is generated by the enzymes CapP and CapO, which have already been characterized biochemically.<sup>38,208</sup> CapP is a 2-epimerase that converts UDP-D-GlcNAc into UDP-D-ManNAc. This enzymatic product is further modified into UDP-D-ManNAcA by the enzymatic action of the dehydrogenase CapO. The transmembrane protein CapI has been proposed to transfer the D-ManNAcA moiety to lipid II<sub>cap</sub>, thereby generating the final capsule lipid precursor lipid III<sub>cap</sub>. The C<sub>55</sub>P-coupled trisaccharide is most likely further modified by the putative acetyltransferase CapH,<sup>209</sup> which catalyzes the O-acetylation of L-FucNAc residues at position C3. This modification is characteristic for staphylococcal CPs and has been implicated in virulence.<sup>209</sup> The complete, modified precursor is then translocated to the outer surface of the cell membrane, where polymerization takes place. These processes are proposed to involve the putative Wzx “flippase” CapK and the putative Wzy polymerase CapJ. The polymerized CP5 is attached to the MurNAc moiety of PG by a yet unknown mechanism,<sup>210</sup> assumed to release the carrier lipid C<sub>55</sub>P, which may enter a new synthesis cycle.

In *S. aureus*, CP production is maximal during the post-exponential phase of cell growth, whereas only little CP is produced in the logarithmic growth phase;<sup>125,134,211,212</sup> the underlying regulation is not only achieved by transcriptional control,<sup>143,213–215</sup> but has moreover been linked to tyrosine phosphorylation (FIGURE 4).<sup>175,176,236</sup> Two triplets of adjacent genes encoding for a transmembrane adaptor, a cytoplasmic BY-kinase, and a PHP family phosphatase, were identified in the genome of *S. aureus* serotype 5: the *capA1/capB1/capC1* triplet located at the 5'-end of the *cap* operon, and the highly similar *capA2/capB2/capC2* triplet, which is located elsewhere on the bacterial chromosome, and probably arose from gene duplication.<sup>161,206</sup> *In vitro* kinase activity could be demonstrated for CapB2, but not for CapB1, with CapA1 being the more potent activator protein.<sup>161,217</sup> The nature of the stimulus that triggers cellular CapAB-signaling and the exact mode of signal

transduction are still elusive. Nevertheless, an enzyme participating in CP5 biosynthesis could be identified as endogenous substrate of the CapAB BY-kinase. Soulat et al. demonstrated that the CapA1B2 complex phosphorylates the dehydrogenase CapO, and that this modification is required for catalytic activity *in vitro*.<sup>216,218</sup>

## 1.7 Short-chain dehydrogenases/reductases

Short-chain dehydrogenases/reductases (SDRs) constitute a large, evolutionary old family of NAD(P)(H)-dependent oxidoreductases, sharing sequence motifs and displaying similar mechanisms.<sup>219,220</sup> Members of the SDR superfamily are found in nearly all organisms, from bacteria to man.<sup>221</sup> They are involved in the metabolism of cofactors, carbohydrates, lipids and amino acids, and participate in redox sensing.<sup>220</sup> SDRs are a functionally diverse collection of proteins, and catalyze a wide range of chemical reactions, including oxidation, reduction, epimerization, dehydration and decarboxylation.<sup>220</sup> Sequence identity between different SDR enzymes is generally low (typically 15–30% residue identity), and homology-based function prediction is especially difficult.<sup>219,222</sup> The only unifying criterion for SDRs is the occurrence of a structurally conserved Rossmann fold dinucleotide-binding motif, consisting of a six-stranded  $\beta$ -sheet flanked by  $\alpha$ -helices.<sup>219,223</sup> This structural motif is characterized by the presence of a glycine-rich sequence pattern, which stabilizes the central  $\beta$ -sheet and mediates the interaction with the pyrophosphate moiety of the nucleotide cofactor.<sup>224,225</sup> Most SDRs display a simple one-domain architecture comprising an N-terminal NADP(H)-binding fold, a largely conserved N[ST]YK active site tetrad, and a highly variable C-terminal region, that harbors the substrate binding site.<sup>226–228</sup> Two main types of SDR enzymes can be distinguished: “classical” SDRs have a core structure with a chain length of about 250 amino acid residues, while “extended” SDRs have an additional C-terminal region of about 100 amino acids, that shows only little sequence conservation.<sup>219,220</sup> In addition, transmembrane domains or signal peptides may be attached to these core structures.<sup>226,229</sup>

## 1.8 Objectives of this work

Infections with multiresistant bacteria represent a serious threat to human health. Moreover, such infectious diseases create a substantial financial burden having detrimental effects on global health care systems.<sup>230</sup> Since established antibiotics are becoming increasingly inefficient, there is an urgent need to explore new treatment strategies to counteract drug-resistant pathogens.<sup>231,232</sup> Capsule biosynthesis is not essential for bacterial cell viability per se. However, inhibitors targeting this process may have the potential for use as anti-virulence agents (see Discussion, section 5.1). To explore constituents of the capsule biosynthetic pathway as drug targets, we need a profound understanding of the underlying molecular processes.

The present study is designed to contribute to the elucidation of capsule biosynthesis in the important opportunistic pathogen *S. aureus*. In serotype 5 and 8 strains, which account for most clinical isolates, three nucleotide-activated sugars synthesized in the cytoplasm, which constitute the major building units of the CP polymer, are required for CP production (see section 1.6).<sup>77</sup> The enzymatic reactions involved in the syntheses of the soluble capsule precursors UDP-L-FucNAc and UDP-D-ManNAcA have already been studied *in vitro*.<sup>38,207,208</sup> In contrast, *in vitro* characterization of the biosynthesis of the soluble capsule precursor UDP-D-FucNAc has long been hampered by the fact that CapD, the enzyme catalyzing the initial step of the biosynthetic reaction cascade, is a protein with multiple transmembrane domains.

In this thesis, reconstitution of the biosynthetic reactions underlying UDP-D-FucNAc production shall be achieved using purified recombinant enzymes. For this purpose, the full-length membrane protein CapD and the cytoplasmic enzyme CapN of the serotype 5 strain *S. aureus* Newman will be overexpressed as recombinant hexahistidine fusion proteins, purified via immobilized metal-ion affinity chromatography, and characterized at the molecular level. Capillary electrophoresis (CE) and mass spectrometry (MS) are the methods of choice for analyzing the interconversion of nucleotide-activated dideoxy sugars, and these techniques will be utilized to monitor enzymatic activities. An enzymatic platform for production of the rare amino sugar UDP-D-FucNAc may not only enable future reconstitution of the complete *S. aureus* capsule biosynthesis pathway, but may also help in reconstructing FucNAc-containing glycan structures from other medically important bacteria (e.g. O glycans from *P. aeruginosa*).

Biosynthesis of *S. aureus* capsular polysaccharide consumes essential precursors, i.e. the universal lipid carrier C<sub>55</sub>P and the diphospho-nucleotide sugar UDP-*N*-acetylglucosamine, also required for the syntheses of the major cell wall components peptidoglycan and wall teichoic acid.<sup>120,122,124</sup> Since the cellular undecaprenol pool is limited, precursor fate and prioritization for the different biosynthesis pathways need to be tightly controlled in time and space to ensure bacterial viability.<sup>119,121,123</sup>

The present study aims to provide evidence that enzymatic checkpoints within the CP biosynthetic pathway are controlled by the tyrosine kinase complex CapAB to regulate consumption of shared precursors. Preliminary experiments identified the biosynthetic enzyme CapE as potential regulatory target of the CapAB complex.<sup>233</sup> Thus, the influence of CapAB-mediated phosphorylation on CapE catalytic activity shall be analyzed *in vitro* using a CE-based detection method for quantification of the CapE enzymatic product UDP-2-acetamido-2,6-dideoxy-β-L-*arabino*-hexulose.<sup>207,234</sup> Besides substantial progress in the understanding of capsule biosynthesis and the underlying regulation, the *in vitro* systems established in this study shall provide the molecular basis to screen for inhibitors of capsule biosynthesis as potential anti-virulence agents. Moreover, they may enable *in vitro* synthesis of lipid-linked capsule intermediates as molecular tools for target-based screening (compare Müller & Ulm et al.<sup>235</sup>).



## 2 Materials and methods

### 2.1 Materials

#### 2.1.1 Reagents and chemicals

All chemicals and reagents utilized in this study were of analytical grade or better; the respective suppliers are given in TABLE 3. Compounds used for inhibitor screening can be found in SUPPLEMENTARY TABLE 1.

TABLE 3

Chemicals and reagents used in this study

Manufacturer/Supplier	Chemical/Reagent
B. Braun, Melsungen	<i>aqua ad iniectabilia</i>
Becton Dickinson, Heidelberg	Difco™ 2YT broth
Bio-Rad Laboratories, München	(Bradford) protein assay dye reagent concentrate
Carl Roth, Karlsruhe	ethidium bromide, glycerol, kanamycin, saccharose, SDS, yeast extract
Clontech Laboratories, Heidelberg	Talon® metal affinity resin
Eurofins MWG Operon, Ebersberg	oligonucleotides
Glycon Biochemicals, Luckenwalde	<i>n</i> -dodecyl- $\beta$ -D-maltoside (DDM)
Hartmann Analytic, Braunschweig	$\gamma$ -labeled [ <sup>33</sup> P]-ATP
J.T. Baker; Distributor: VWR International, Langenfeld	chloroform, ethanol, methanol, water (HPLC grade)
Life Technologies, Darmstadt	Invitrogen™ NuPAGE® Novex® 4–12% Bis-Tris gels with reducing agent, LDS sample buffer and MOPS SDS running buffer
Merck, Darmstadt	agar-agar, bromophenol blue, calcium chloride, EDTA, glacial acetic acid, orthophosphoric acid, peptone, potassium phosphate monobasic/dibasic, sodium chloride, sodium phosphate monobasic/dibasic, tryptic soy broth
MP Biomedicals, Eschwege	tryptone
New England BioLabs, Frankfurt am Main	dNTPs, Phusion® High-Fidelity DNA polymerase with HF buffer, Type II restriction endonucleases
Qiagen, Hilden	DNeasy® blood & tissue kit, MinElute® PCR purification kit, Ni-NTA agarose, QIAprep® spin miniprep kit, QIAquick® gel extraction kit
Roche, Mannheim	T4 DNA ligase with ligation buffer

TABLE 3 (continued)

Sigma-Aldrich, Taufkirchen	ATP, ammonium bicarbonate, 6-aza-2-thiothymine, bovine serum albumin, CAPS, diammonium hydrogen citrate, DMSO, DTT, deoxyribonuclease I, formic acid, imidazole, iodoacetamide, lysostaphin, lysozyme, magnesium chloride, NAD(P) <sup>+</sup> , NAD(P)H, PAP, polybrene, ribonuclease A, sodium acetate, triethylammonium acetate buffer (1 M, pH 7), Triton X-100, Tris base/hydrochloride, trypsin (sequencing grade), UDP, UDP-D-GlcNAc
Thermo Fisher Scientific, Schwerte	Fermentas™: GeneRuler™ 1 kb DNA ladder, IPTG, PageBlue™ protein staining solution, PageRuler™ Plus prestained protein ladder, TopVision™ LE GQ agarose; Oxoid blood agar base number 2, Pierce® Zeba™ desalt spin columns (5 ml, 7 kDa molecular weight cut-off)
VWR International, Langenfeld	HiPerSolv Chromanorm® acetonitrile

*Note.* Suppliers are located in Germany. Abbreviated names of chemicals are explained in “Abbreviations and symbols”.

### 2.1.2 Plasmids and oligonucleotides

The Novagen® pET (plasmid expression by T7-RNA-polymerase) system is based on the T7 expression system established by Studier et al. and was developed for cloning and expression of recombinant proteins in *E. coli*.<sup>236</sup> Target genes are cloned in pET plasmids under control of a strong bacteriophage T7 promoter; high-level expression of recombinant proteins is induced by providing a source of T7 RNA polymerase in the host cell.

Recombinant plasmids enabling overexpression and purification of hexahistidine fusion proteins for *in vitro* reconstitution of capsule biosynthesis reactions were constructed on the basis of the pET vectors 24a(+) and 28a(+) (TABLE 4). The oligonucleotide primers utilized for plasmid construction are listed in TABLE 5. The identity of recombinant plasmids was confirmed by sequencing (Sequiserve, Vaterstetten, Germany).

TABLE 4

Plasmids used in this study

Plasmid	Characteristics/Use	Reference/Source
pET24a(+)	<i>Escherichia coli</i> vector, expression of C-terminal His <sub>6</sub> fusion proteins, Km <sup>R</sup> , T7 promoter	Novagen <sup>®</sup> by Merck
pET28a(+)	<i>E. coli</i> vector, expression of N-terminal His <sub>6</sub> fusion proteins, Km <sup>R</sup> , T7 promoter	Novagen <sup>®</sup>
pET24a- <i>capD</i>	<i>Staphylococcus aureus</i> Newman <i>capD</i> in pET24a	Prof J.C. Lee <sup>a</sup>
pET28a- <i>capDtr</i>	truncated <i>capD</i> in pET28a, encodes soluble CapD derivative (amino acids 249–611)	this thesis
pET24a- <i>capE</i> (pKBK50d)	<i>S. aureus</i> Newman <i>capE</i> in pET24a	reference <sup>207</sup>
pET24a- <i>capN</i>	<i>S. aureus</i> Newman <i>capN</i> in pET24a	Prof J.C. Lee <sup>a</sup>
pET28a- <i>capA1</i>	<i>S. aureus</i> Newman <i>capA1</i> in pET28a	reference <sup>233</sup>
pET28a- <i>capB2</i>	<i>S. aureus</i> Newman <i>capB2</i> in pET28a	reference <sup>233</sup>
pET28a- <i>pglF</i>	truncated <i>Campylobacter jejuni pglF</i> in pET28a, encodes soluble PglF derivative (amino acids 130–590)	this thesis, according to reference <sup>237</sup>

Note. His<sub>6</sub>, hexahistidine; Km<sup>R</sup>, kanamycin-resistance marker.

<sup>a</sup>Brigham & Women's Hospital/Harvard Medical School, Boston, USA

TABLE 5

Sequences of oligonucleotide primers used for plasmid construction

Primer name	Sequence (5'-3') <sup>a</sup>	Restriction site
<i>capD</i> for	TATACATATGGCTAGC <u>TTATCTGTGAAATTG</u>	<i>NheI</i>
<i>capD</i> rev	TTTGTCTCGAGATAATTATCTCCCTTTTGC	<i>XhoI</i>
<i>capDtr</i> for	GCGCGGCTAGC <u>ATGTCTGGTGAGTTAGAAGTG</u>	<i>NheI</i>
<i>capDtr</i> rev	GCGCGCTCGAGTCATCGAACATAATTATCTCCC	<i>XhoI</i>
<i>capE</i> for	GGAGGCTAGC <u>ATGTTTCGATGACAAAATT</u>	<i>NheI</i>
<i>capE</i> rev	CTCTCGAGTCTCATTGAAGCTTTATAAT	<i>XhoI</i>
<i>capN</i> for	GTGCAGCTAGC <u>ATGAGAAAAATATT</u>	<i>NheI</i>
<i>capN</i> rev	AATAGATCTCGAGTGCCTTATCTTTG	<i>XhoI</i>
<i>capA1</i> for	GCGCGGCTAGC <u>ATGGAAAGTACATTAGAATTAAC</u>	<i>NheI</i>
<i>capA1</i> rev	GCGCGCTCGAGT <u>TAATTAATTTTTGAATTGAACCC</u>	<i>XhoI</i>
<i>capB2</i> for	GCGCGGCTAGC <u>ATGACGAATACACGAAGAAGTA</u>	<i>NheI</i>
<i>capB2</i> rev	GCGCGCTCGAGT <u>CATTCATCAGTCCCATAATATG</u>	<i>XhoI</i>
<i>pglF</i> for	GCGCGCATATG <u>CTTGTGGATTTTAAACCTTC</u>	<i>NdeI</i>
<i>pglF</i> rev	GCGCGCTCGAGT <u>TATACACCTTCTTTATTGTGT</u>	<i>XhoI</i>

<sup>a</sup>Underlining indicates the restriction enzyme recognition site engineered into the primer.

### 2.1.3 Bacterial strains and culture media

Bacterial strains used in this thesis are shown below in TABLE 6; the different growth media utilized for their maintenance are given in TABLE 7.

TABLE 6

Bacterial strains used in this study

Bacterial strain	Characteristics/Use	Reference/Source
<u><i>Campylobacter jejuni</i></u>		
clinical isolate	DNA template for PCR amplification of <i>pglF</i>	IMMIP
<u><i>Escherichia coli</i></u>		
BL21 (DE3)	general purpose expression host, $\lambda$ prophage DE3 carries T7 RNA polymerase gene under control of an IPTG-inducible promoter ( <i>lacUV5</i> )	Novagen <sup>®</sup> by Merck
C43 (DE3)	derivative of <i>E. coli</i> BL21 (DE3), optimized for expression of membrane proteins	reference <sup>238</sup>
DH5 $\alpha$	non-expression host, general purpose cloning, plasmid propagation and storage	reference <sup>239</sup>
<u><i>Staphylococcus aureus</i></u>		
Newman	methicillin-sensitive, capsular serotype 5, DNA template for PCR amplification of capsule biosynthesis genes	reference <sup>240</sup>

*Note.* IMMIP, Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Germany.

TABLE 7

Culture media used in this study

Culture medium	Composition
Blood agar	40 g/l Oxoid <sup>™</sup> blood agar base number 2, 7% (v/v) sterile defibrinated sheep blood <sup>a</sup>
Lysogeny broth (LB)	10 g/l tryptone, 10 g/l NaCl, 5 g/l yeast extract; + 15 g/l agar-agar for LB agar plates
Tryptic soy broth (Merck)	17 g/l peptone from casein, 3 g/l peptone from soymeal, 5 g/l NaCl, 2.5 g/l K <sub>2</sub> HPO <sub>4</sub> , 2.5 g/l glucose monohydrate
Double-strength yeast extract-tryptone broth (Difco <sup>™</sup> 2YT)	16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl

*Note.* All media were prepared in deionized distilled water.

<sup>a</sup>added after autoclaving and cooling down to 50°C

### 2.1.4 Sterilization procedures

Growth media, solutions, plastic vials and pipette tips were sterilized by autoclaving at 121°C for 20 min. (Varioklav<sup>®</sup> 500E; H+P Labortechnik, Oberschleißheim, Germany). Heat-labile substances, such as antibiotics, were sterilized by filtration using Acrodisc<sup>®</sup> syringe filters (sterile, pore size 0.2 µm; Pall, Dreieich, Germany). The sterilization of glassware was carried out in an oven (model SLM 700; Memmert, Schwabach, Germany) at 200°C for 4 h.

### 2.2 Growth, maintenance and preservation of bacterial strains

Bacterial strains used in this thesis are listed in TABLE 6; the different growth media utilized for their maintenance are denoted in TABLE 7. Glycerol cultures were prepared for long-term storage of bacterial strains. For this purpose, aliquots of overnight cultures were mixed 1:1 (v/v) with sterile glycerol and stored in cryovials at -70°C. For recovery from glycerol stocks, bacterial strains were streaked out onto appropriate agar plates. *S. aureus* Newman was cultivated on blood agar plates, which were incubated at 37°C for 24 to 48 h. *E. coli* strains were streaked onto lysogeny broth (LB) agar plates and grown overnight at 37°C. Transformed *E. coli* DH5α strains used for long-term storage of recombinant plasmids were cultivated on LB agar containing 25 µg/ml kanamycin. The agar plates with the recovered bacterial strains were stored at 4°C for up to three weeks.

Prolonged storage of transformed *E. coli* expression strains, at 4°C as well as at -70°C, may lead to plasmid instability and/or plasmid loss, resulting in drastically reduced recombinant protein productivity. Therefore, the *E. coli* strains BL21 (DE3) and C43 (DE3), utilized as expression hosts in this study, were transformed with the appropriate plasmids immediately before use. The transformed cells were streaked onto LB agar plates containing 25 µg/ml kanamycin and grown overnight at 37°C. These working cultures were stored at 4°C for two weeks maximum.

If not indicated otherwise, liquid cultures were grown aerobically at 37°C in a shaking water bath (model 1083; Gesellschaft für Labortechnik, Burgwedel, Germany), using a 1% (v/v) inoculum from an overnight starter culture. *S. aureus* Newman was cultivated in tryptic soy broth. *E. coli* strains were grown in LB or in double-strength yeast extract-tryptone (2YT) broth (Becton Dickinson); the media were supplemented with 25 µg/ml kanamycin, if appropriate.

## 2.3 Methods in molecular genetics

### 2.3.1 Isolation of genomic DNA from *S. aureus*

For preparation of genomic DNA, 4 ml of a *S. aureus* Newman overnight culture were harvested by centrifugation (13,000 rpm, 2 min, RT; MiniSpin<sup>®</sup> microcentrifuge, rotor F-45-12-11; Eppendorf, Hamburg, Germany), and cells were resuspended in 180 µl ATL buffer (from DNeasy<sup>®</sup> blood & tissue kit; Qiagen). The suspension was supplemented with lysozyme (100 mg/ml) and lysostaphin (5 mg/ml), and incubated at 37°C for 1 h to enable cell wall degradation and complete lysis of the staphylococcal cells. Subsequently, extraction of genomic DNA from the lysate was carried out using the DNeasy<sup>®</sup> blood & tissue kit as described in the manufacturer's instructions. Genomic DNA was eluted from the spin columns with 2 x 100 µl *aqua ad iniectabilia* (*aqua a.i.*) and stored at -20°C until further use. The concentration of the purified genomic DNA was determined photometrically by measuring the absorbance at 260 nm in a UV/VIS spectrophotometer (NanoPhotometer<sup>®</sup>; Implen, München, Germany); purity was assessed by the ratio of absorbance at 260 nm and 280 nm. Quality and quantity of the genomic DNA were also analyzed by agarose gel electrophoresis.

### 2.3.2 Preparation of plasmid DNA

Bacterial overnight cultures were prepared by picking a single *E. coli* colony from an LB agar plate and inoculating 16 ml of LB, to which 25 µg/ml kanamycin were added for selection. The next day, cells were harvested by centrifugation (13,000 rpm, 2 min, RT; MiniSpin<sup>®</sup> microcentrifuge, rotor F-45-12-11; Eppendorf, Hamburg, Germany), and plasmid DNA was isolated from the cells using the QIAprep<sup>®</sup> spin miniprep kit (Qiagen) according to the manufacturer's instructions. Plasmid DNA was eluted from the spin columns with 50 µl *aqua a.i.* and stored at -20°C until further use. Quality and quantity of the purified plasmid DNA were analyzed by agarose gel electrophoresis and by measurement of absorbance (see 2.3.1).

### 2.3.3 Polymerase Chain Reaction (PCR)

PCR amplification of DNA fragments was performed on the iCycler<sup>™</sup> thermal cycler (Bio-Rad Laboratories, München, Germany), using Phusion<sup>®</sup> High-Fidelity DNA polymerase (New England Biolabs) following the manufacturer's instructions, and

according to standard PCR procedures.<sup>241</sup> Oligonucleotide primers (TABLE 5) were obtained from Eurofins MWG Operon; dNTPs were purchased from New England Biolabs.

Amplification of *pglF* was performed using genomic DNA from a clinical isolate of *C. jejuni* as template. For this purpose, several colonies were picked from a blood agar plate and resuspended in *aqua a.i.* The cells were lysed by heating (10 min, 100°C), and cell debris was removed by centrifugation (13,000 rpm, 5 min, RT; MiniSpin<sup>®</sup> microcentrifuge, rotor F-45-12-11; Eppendorf, Hamburg, Germany). The genomic DNA in the supernatant was used as PCR template without further purification. Purified genomic DNA from *S. aureus* Newman (see 2.3.1) was utilized as template for the amplification of *capDtr*.

Size and purity of PCR amplicates were controlled by agarose gel electrophoresis. To enable restriction enzyme digestion, DNA fragments were purified from PCR mixtures using the MinElute<sup>®</sup> PCR purification kit (Qiagen) according to the manufacturer's instructions. PCR products were eluted from the silica columns with 20 µl *aqua a.i.* and stored at -20°C. Quality and quantity of the purified PCR amplicates were determined by absorbance measurements (see 2.3.1).

### 2.3.4 Agarose gel electrophoresis

DNA fragments were analyzed by standard agarose gel electrophoresis,<sup>241</sup> on a horizontal electrophoresis apparatus (Renner Laborbedarf, Dannstadt, Germany). TopVision<sup>™</sup> agarose (Fermentas<sup>™</sup>) and TAE buffer (Tris-acetate EDTA buffer; 40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA, pH 8) were used to prepare 1% (w/v) agarose gels, which were supplemented with 0.5 µg/ml ethidium bromide to enable detection of DNA fragments. DNA-containing samples were mixed with 10% (v/v) loading dye (25% (w/v) saccharose, 0.1% (w/v) SDS, 0.05% (w/v) bromophenol blue, 5 mM sodium acetate) before they were applied onto the agarose gel. The GeneRuler<sup>™</sup> 1 kb DNA ladder (Fermentas<sup>™</sup>) was utilized as length standard. The electrophoresis was performed in TAE buffer, at constant 100 V, for 30 min to 3 h. DNA fragment patterns were visualized under UV light and documented using the Molecular Imager<sup>®</sup> Gel Doc<sup>™</sup> XR+ system (Bio-Rad Laboratories, München, Germany).

### 2.3.5 Cleavage of DNA by restriction endonucleases

Type II restriction endonucleases were purchased from New England Biolabs; restriction of purified plasmid DNA and PCR fragments was performed according to the manufacturer's instructions. PCR fragments were purified from restriction digestion mixtures using the MinElute<sup>®</sup> PCR purification kit; linearized plasmids were isolated by agarose gel electrophoresis, and purified from agarose gel slices using the QIAquick<sup>®</sup> gel extraction kit (both kits were obtained from Qiagen). Restricted DNA fragments were eluted from the silica columns with 20  $\mu$ l *aqua a.i.* and stored at  $-20^{\circ}\text{C}$  until further use. Quality and quantity of restricted DNA fragments were analyzed by absorbance measurements (see 2.3.1).

### 2.3.6 Ligation of DNA fragments

T4 DNA ligase obtained from Roche was utilized to enable molecular cloning. Ligation reactions were performed following the manufacturer's instructions and according to standard cloning procedures.<sup>241</sup> The mixture with the ligation products was either directly utilized for transformation of chemocompetent *E. coli* cells, or stored at  $-20^{\circ}\text{C}$  until further use.

### 2.3.7 Preparation of $\text{CaCl}_2$ -competent *E. coli* cells

For preparation of chemocompetent cells, *E. coli* strains were grown in 500 ml LB at  $37^{\circ}\text{C}$ . At an  $\text{OD}_{600}$  of 0.5, cultures were harvested by centrifugation (7,000 rpm, 12 min,  $4^{\circ}\text{C}$ ; Sorvall<sup>®</sup> Evolution<sup>™</sup> RC centrifuge, rotor SLC-6000; Thermo Fisher Scientific, Schwerte, Germany). Subsequently, the cells were resuspended in 10 ml ice-cold 0.1 M  $\text{CaCl}_2$ , incubated on ice for 20 min, and again harvested by centrifugation (see above). The cell pellet was resuspended in 2 ml ice-cold 0.1 M  $\text{CaCl}_2$  supplemented with 15% (v/v) glycerol, and the suspension was incubated on ice for 1 h. Afterwards, aliquots of 50  $\mu$ l were prepared, which were stored at  $-70^{\circ}\text{C}$  until further use.

### 2.3.8 Transformation of $\text{CaCl}_2$ -competent *E. coli* cells

50- $\mu$ l aliquots of chemocompetent *E. coli* cells were thawed on ice and mixed either with 3  $\mu$ l plasmid DNA, or with up to 20  $\mu$ l ligation mixture. After 30 min on ice, the cells were heat-shocked for 90 s at  $42^{\circ}\text{C}$ , and incubated on ice again for about 3 min.



The cells were pipetted into 1 ml pre-warmed LB, and recovered at 37°C for 1 h in a shaking water bath to enable expression of the resistance marker. Finally, aliquots of the suspension were streaked onto LB agar plates containing 25 µg/ml kanamycin, which were incubated at 37°C until the next day.

## 2.4 Protein and biochemical methods

### 2.4.1 Overexpression and purification of hexahistidine fusion proteins

*E. coli* strain BL21 (DE3) was used as host for recombinant expression of cytoplasmic enzymes (CapB2-His<sub>6</sub>, CapE-His<sub>6</sub>, PglF-His<sub>6</sub>) and of an integral membrane protein (CapA1-His<sub>6</sub>) involved in capsule biosynthesis. Furthermore, *E. coli* BL21 was utilized for the expression of a soluble CapD mutant protein (CapDtr-His<sub>6</sub>). Cells were grown at 37°C in 2 l LB (or 0.5 l LB for overexpression of PglF-His<sub>6</sub>) containing 25 µg/ml kanamycin. At an OD<sub>600</sub> of 0.6, IPTG was added to induce expression of the recombinant protein, either at a concentration of 0.5 mM (for overexpression of CapE-His<sub>6</sub> and PglF-His<sub>6</sub>) or at 1 mM (for CapA-His<sub>6</sub>, CapB-His<sub>6</sub> and CapDtr-His<sub>6</sub>). Expression times and temperatures were optimized for the individual constructs. Cultures for overexpression of PglF-His<sub>6</sub> and CapE-His<sub>6</sub> were induced at 37°C for 4 h; overexpression of CapA1-His<sub>6</sub>, CapB2-His<sub>6</sub> and CapDtr-His<sub>6</sub> was performed at 30°C for 4 h. After induction, cells were harvested (8,000 rpm, 10 min, 4°C; Sorvall<sup>®</sup> Evolution<sup>™</sup> RC centrifuge, rotor SLC-6000; Thermo Fisher Scientific, Schwerte, Germany) and resuspended in lysis buffer (50 mM NaPi, 300 mM NaCl, pH 7.5), which was supplemented with 1% (v/v) Triton X-100 (Sigma-Aldrich) for purification of the integral membrane protein CapA1-His<sub>6</sub>. Lysozyme (250 µg/ml), deoxyribonuclease I (50 µg/ml) and ribonuclease A (10 µg/ml) were added. The suspension was incubated on ice for 30 min and sonicated (50 W, 30 kHz, 10 x 20 s, with cooling periods of 1 min between each burst of sonication; Sonifier UP50H; Hielscher Ultrasonics, Teltow, Germany). Cell debris was removed by centrifugation (14,000 rpm, 30 min, 4°C; Sigma 2K15 centrifuge, rotor 12148; B. Braun, Melsungen, Germany). The supernatant was incubated with 1 ml Ni-NTA-agarose slurry (50% (w/v); Qiagen) for 2 h at 4°C under gentle stirring. The mixture was then loaded onto a column support (5 ml polypropylene cartridge; Qiagen, Hilden, Germany). After washing with lysis buffer, weakly bound material was removed with 10 and 20 mM imidazole. His<sub>6</sub>-tagged recombinant proteins were eluted with lysis buffer containing 300 mM imidazole. Three fractions were collected

each, and stored in 30% (v/v) glycerol at  $-20^{\circ}\text{C}$ . Preparations of CapE-His<sub>6</sub> and PglF-His<sub>6</sub> were dialysed against 10 mM KPi, pH 7.5 before storage (see 2.4.4). Purity of elution fractions was controlled by SDS-PAGE; protein concentration was measured by Bradford's method.<sup>242</sup>

Recombinant proteins CapD-His<sub>6</sub> and CapN-His<sub>6</sub> were expressed in host strain *E. coli* C43 (DE3). Cells transformed with the respective plasmids were grown at  $30^{\circ}\text{C}$  in 2 l 2YT broth (Becton Dickinson) supplemented with 25  $\mu\text{g}/\text{ml}$  kanamycin. At an OD<sub>600</sub> of 0.6, inducer IPTG was added at a concentration of 0.5 mM. Cultures were incubated at  $25^{\circ}\text{C}$  for 18 h to allow expression of the recombinant proteins. The membrane-associated protein CapN-His<sub>6</sub> was purified as described above, using lysis buffer supplemented with 1% (v/v) Triton X-100. Preparations of CapN-His<sub>6</sub> were supplemented with 50% (v/v) glycerol (final concentration) and stored at  $-20^{\circ}\text{C}$  until use. A different protocol was applied to facilitate purification of CapD-His<sub>6</sub>, an enzyme comprising four membrane-spanning domains. Cells containing recombinant CapD-His<sub>6</sub> were harvested (8,000 rpm, 10 min,  $4^{\circ}\text{C}$ ; Sorvall<sup>®</sup> Evolution™ RC centrifuge, rotor SLC-6000) and resuspended in lysis buffer (50 mM NaPi, 300 mM NaCl, 20% (v/v) glycerol, pH 7.5). The suspension was sonicated on ice (see above) and centrifuged (13,000 rpm, 30 min,  $4^{\circ}\text{C}$ ; Sorvall<sup>®</sup> Evolution™ RC centrifuge, rotor SS-34). The pellet was resuspended in lysis buffer supplemented with 49 mM *n*-dodecyl  $\beta$ -D-maltoside (DDM; Glycon Biochemicals). For solubilization of membrane-embedded proteins, the mixture was incubated on ice for 30 min with gentle shaking. Insoluble components were removed by centrifugation (14,000 rpm, 30 min,  $4^{\circ}\text{C}$ ; Sigma 2K15 centrifuge, rotor 12148), and the cleared supernatant was mixed with 1 ml Talon<sup>®</sup> metal affinity resin (50% (w/v) slurry; Clontech Laboratories). The suspension was incubated for 1 h at  $4^{\circ}\text{C}$  under gentle stirring, before it was filled into a support cartridge (see above). The resulting column was washed with lysis buffer containing 1 mM DDM. A second wash was performed with lysis buffer containing 1 mM DDM and 15 mM imidazole. To elute recombinant proteins, lysis buffer supplemented with 1 mM DDM and 200 mM imidazole was applied to the column. Three fractions were collected, and analyzed regarding purity and protein content as described above. CapD-His<sub>6</sub> turned out to be inactivated by freezing, even in the presence of high glycerol concentrations. For this reason, recombinant CapD-His<sub>6</sub> was prepared freshly for all assays.

### **2.4.2 Sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)**

Protein-containing elution fractions were analyzed by discontinuous SDS polyacrylamide gel electrophoresis using pre-cast NuPAGE<sup>®</sup> Novex<sup>®</sup> 4–12% Bis-Tris mini gels (Invitrogen<sup>™</sup>). In contrast to the widely used gel system according to Laemmli, the NuPAGE<sup>®</sup> Bis-Tris system operates at a neutral pH value enabling stability of proteins and gel matrix during electrophoresis, and better band resolution.<sup>243,244</sup> Sample preparation for SDS-PAGE analysis was performed according to the manufacturer's instructions, with the exception that samples containing membrane proteins were heated to 55°C for 3 min after addition of LDS sample buffer. If indicated, NuPAGE<sup>®</sup> reducing agent was omitted from the samples. PageRuler<sup>™</sup> Plus prestained protein ladder (Fermentas<sup>™</sup>) was employed as molecular weight standard; NuPAGE<sup>®</sup> MOPS SDS running buffer was used as running buffer. The electrophoresis was carried out at constant 200 V for 50 min in an XCell SureLock<sup>®</sup> electrophoresis cell (Invitrogen<sup>™</sup> by Life Technologies, Darmstadt, Germany). Subsequently, gels were washed with distilled water, stained to the desired intensity with PageBlue<sup>™</sup> protein staining solution (Fermentas<sup>™</sup>), and again incubated in distilled water, until the background was completely destained. The Molecular Imager<sup>®</sup> Gel Doc<sup>™</sup> XR+ imaging system (Bio-Rad Laboratories, München, Germany) was used for documentation.

### **2.4.3 Determination of protein concentrations according to Bradford**

Sample protein concentrations were determined by Bradford's method using the Bio-Rad protein assay dye reagent following the manufacturer's protocol for microplate assays<sup>242</sup>. In brief, 10- $\mu$ l aliquots of protein samples were pipetted into the wells of a flat-bottom polypropylene microplate (Anicrin; Distributor: VWR International, Langenfeld, Germany) and mixed with 200  $\mu$ l appropriately diluted Bradford reagent. After 10 min of incubation at room temperature, the absorption at a wavelength of 595 nm was measured in a microplate reader (Infinite<sup>®</sup> M200 NanoQuant equipped with Magellan<sup>™</sup> 6 software; Tecan, Crailsheim, Germany). All samples were analyzed in duplicates. A bovine serum albumin standard curve (0–500  $\mu$ g/ml) was prepared analogously.

#### 2.4.4 *In vitro* synthesis of soluble capsule precursors

For functional reconstitution, capsule biosynthetic enzymes were expressed as hexahistidine fusion proteins and purified via Ni-NTA affinity chromatography.<sup>245</sup> Preparations of CapE-His<sub>6</sub> and PglF-His<sub>6</sub> were dialysed against 10 mM KPi, pH 7.5, using Zeba™ desalt spin columns (5 ml, 7 kDa molecular weight cut-off; Pierce™) according to the manufacturer's instructions. Dialysates were supplemented with 30% (v/v) glycerol and stored at -20°C until further use. Elution fractions containing CapD-His<sub>6</sub> and CapN-His<sub>6</sub> were used for enzyme assays without prior buffer exchange. CapN-His<sub>6</sub> was stored in 50% (v/v) glycerol at -20°C; CapD-His<sub>6</sub> was prepared freshly for all assays.

Syntheses of soluble capsule precursors UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose (UDP-Sugp) and UDP-2-acetamido-2,6-dideoxy-L-arabino-4-hexulose, catalyzed by CapD and CapE, respectively, were carried out in a total volume of 50 µl, containing UDP-D-GlcNAc (concentrations of UDP-D-GlcNAc varied according to assay type) in 10 mM KPi, 10 mM MgCl<sub>2</sub>, pH 7.5. CapD assays additionally contained 0.8% (v/v) Triton X-100 and, if not stated otherwise, 2 mM NADP<sup>+</sup> as cofactor. Reactions were initiated by the addition of 3 µg CapD-His<sub>6</sub>, or of 8.5 µg CapE-His<sub>6</sub>, and incubated at 30°C for 30 min to 12 h.

The synthesis of the precursor UDP-Sugp was also performed with a recombinant truncated version of the enzyme PglF from *C. jejuni*, as described by Schoenhofen et al.<sup>237</sup> In brief, 5 µg PglF-His<sub>6</sub> and 3 mM UDP-GlcNAc were incubated in 50 µl 10 mM KPi, pH 7.5 at 30°C for 16 h.

The PglF-His<sub>6</sub> enzymatic product was used as substrate for reconstitution of CapN catalytic activity. For this purpose, PglF reactions were carried out overnight and quenched by heating (5 min, 90°C). CapN-catalyzed synthesis of UDP-D-FucNAc was carried out in a total volume of 40 µl. CapN-His<sub>6</sub> (12 µg) was incubated in the presence of ~3 mM UDP-Sugp, 0.8% (v/v) Triton X-100 and 10 mM KPi, pH 7.5, for 2 h at 30°C. Alternatively, synthesis of UDP-D-FucNAc was performed in a one-pot-assay containing 3.5 µg PglF, 12 µg CapN, 3 mM UDP-D-GlcNAc, 0.8% (v/v) Triton X-100 and 10 mM KPi, in a total volume of 60 µl. Cofactors NADH or NADPH were added at a concentration of 1.875 mM, if indicated.

All enzymatic reactions were quenched by heating (5 min, 90°C) to enable reliable quantifications. For capillary electrophoresis (CE) analysis, samples were prepared as described above. Synthesis of larger quantities of soluble capsule precursors for

mass spectrometric analysis was achieved by a 10-fold upscale of the synthesis reaction.

#### 2.4.5 Determination of kinetic parameters for CapD and CapE

To determine the linear range of the enzymatic reactions, formation of CapD and CapE enzymatic products was monitored over time (six time points, 10 min to 6 h) at 30°C with a substrate (UDP-D-GlcNAc) concentration of 1 mM. For determination of the kinetic parameters  $K_m$  (Michaelis-Menten constant) and  $V_{max}$  (maximum velocity),<sup>246</sup> eight different substrate concentrations were chosen. For CapD assays, the cofactor NADP<sup>+</sup> was added at a concentration of 2 mM. Each analysis was repeated three times in independent experiments.

#### 2.4.6 Investigation of CapD and CapE inhibitors

A library of selected compounds (73 compounds for CapD; 46 compounds for CapE; SUPPLEMENTARY TABLE 1) was initially screened at high concentrations of 100 μM (for highly water-soluble compounds), or at 10 μM. Subsequently, full concentration–inhibition curves were determined for the most potent inhibitors, and IC<sub>50</sub> values (concentration required for 50% inhibition *in vitro*) were calculated. The substrate (UDP-D-GlcNAc) concentration for CapD and CapE inhibition testing was 1 mM, and substrate conversion was strictly controlled to be below 10%. For CapD assays, the cofactor NADP<sup>+</sup> was added at a concentration of 600 μM. The IC<sub>50</sub> values of inhibitors were obtained by testing a suitable range of inhibitor concentrations (eight data points spanning three orders of magnitude). The incubation time was chosen to be within the linear range (steady-state phase; 30 min to 1 h). Negative control reactions were performed in the presence of heat-inactivated (10 min, 100°C) enzymes. For further details, see below, in the CE quantification section. Each analysis was repeated three times in independent experiments. For the determination of the inhibition mechanism, eight different substrate concentrations and three different inhibitor concentrations were used, and each analysis was carried out with duplicate measurements.

### 2.4.7 Capillary electrophoresis analysis of soluble capsule precursors

Capillary electrophoresis (CE) analyses described in this thesis were performed on a P/ACE™ MDQ capillary electrophoresis apparatus (Beckman Coulter, Krefeld, Germany) equipped with a diode array detector. The electrophoretic separations were carried out using uncoated fused-silica capillaries of 60 cm total length (50 cm effective length) × 75.5 μm (inner diameter) × 363.7 μm (outer diameter) obtained from Polymicro Technologies (Kehl, Germany). Data collection and corrected peak area analysis were performed with 32 Karat™ software (Beckman Coulter). Further data analysis was carried out using the software GraphPad Prism 4 (GraphPad Software Inc., La Jolla, United States). The capillary temperature was kept constant at 15°C. The following conditions were applied:  $\lambda_{\max}$  260 nm, separation voltage -15 kV, running buffer 68 mM CAPS, 0.002% polybrene, pH 12.4 (adjusted with NaOH), electrokinetic injection (-1 kV, 20 s). The capillary was washed with 0.2 M NaOH for 2 min (20 psi) and with running buffer for 3 min (20 psi) before each injection. Analytical samples were diluted 1:40 in autoclaved ultrapure water to give a volume of 50 μl, and 50 μl of 3'-phosphoadenosine 5'-phosphate (PAP) dissolved in sterile ultrapure water were added as an internal standard (final concentration 6.45 μM). An aliquot of the test mixture (90 μl) was pipetted into the sample vial and subsequently measured by CE.

CE analysis of CapN enzymatic reactions was performed as described by Kneidinger et al.<sup>207</sup>

### 2.4.8 CE-based quantification of CapD and CapE reaction products

The CapD and CapE reaction products are UDP-Sugp and UDP-2-acetamido-2,6-dideoxy-L-arabino-4-hexulose, respectively. Since both compounds are not commercially available, CE calibration was performed with enzymatic products.

Recombinant CapE and PglF are both capable of catalyzing 100% conversion of the substrate UDP-D-GlcNAc. However, the final product concentrations in the reaction mixtures are lower than the initial substrate concentrations, due to simultaneous hydrolysis of substrate and products yielding UDP. To determine the actual product concentration in the samples, we quantified the degradation product UDP, subtracted the respective values from the initial substrate amount and considered the difference to represent the amount of UDP-Sugp and UDP-2-acetamido-2,6-dideoxy-L-arabino-4-hexulose, respectively.

The amount of UDP was determined as follows: UDP was diluted with water to different concentrations; then 50- $\mu$ l aliquots of the different UDP dilutions were mixed with 50  $\mu$ l of the internal standard PAP (final concentration 3.125  $\mu$ M), which was prepared in biological matrix (PgIF reaction mixture diluted 1:40 (v/v) with water). Since UDP is also formed during *in vitro* PgIF assays, UDP calibration was performed using biological matrix as a blank value.

TABLE 8

Quantification of CapD and CapE reaction products: method validation

Parameters	CapD product UDP-2-acetamido-2,6-dideoxy- $\alpha$ -D-xylo-hexulose	CapE product UDP-2-acetamido-2,6-dideoxy- $\beta$ -L-arabino-hexulose
Regression equation <sup>a</sup>	$y = 7.533x - 0.35$	$y = 7.873x - 0.767$
R <sup>2</sup>	0.9927	0.9801
LOD ( $\mu$ M)	0.71	0.81
LOQ ( $\mu$ M)	2.15	2.45
Calibration range ( $\mu$ M)	0.22–7.09	0.33–5.34

	Concentration ( $\mu$ M)	Accuracy (% recovery) <sup>a</sup>	Accuracy acceptable range <sup>b</sup>	Precision (RSD %) <sup>a</sup>	Precision acceptable range <sup>b</sup>
CapD product	0.44	100.9	80–120%	14.4	20%
	0.89	100.4	80–120%	2.3	20%
	3.55	101.0	80–120%	5.4	20%
	7.09	99.9	80–120%	7.0	20%
CapE product	0.67	100.9	80–120%	4.4	20%
	1.34	93.6	80–120%	7.0	20%
	2.67	101.5	80–120%	11	20%
	5.35	99.9	80–120%	11	20%

Note. Performed by Wenjin Li (Pharmaceutical Chemistry I, University of Bonn), according to ICH and FDA guidelines.<sup>247,248</sup> LOD, limit of detection; LOQ, limit of quantification; R<sup>2</sup>, coefficient of determination; RSD, relative standard deviation.

<sup>a</sup>triplicate samples with duplicate measurements

<sup>b</sup>according to ICH guidelines

The CE-based detection method for monitoring of UDP-D-GlcNAc 4,6-dehydratase activity was validated according to the guidelines published by the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, and by the US Food and Drug Administration

(FDA), demonstrating its suitability for the analysis of enzyme kinetics and for inhibitor testing.<sup>247,248</sup> A summary of the parameters of method validation is provided in TABLE 8.

The linearity of the quantitative determination and the limits of detection and quantification of the enzymatic products UDP-Sugp and UDP-2-acetamido-2,6-dideoxy-L-*arabino*-4-hexulose were determined by processing six-point calibration curves (triplicate samples with duplicate measurements) in the presence of 3.125  $\mu\text{M}$  PAP as internal standard. The precision and accuracy values were determined at the following concentrations: 0.44, 0.89, 3.55 and 7.09  $\mu\text{M}$  for UDP-Sugp; 0.67, 1.34, 2.67 and 5.35  $\mu\text{M}$  for UDP-2-acetamido-2,6-dideoxy-L-*arabino*-4-hexulose (triplicate samples with duplicate measurements; 3.125  $\mu\text{M}$  PAP as internal standard). Accuracy was determined by calculating the ratios of the predicted concentrations and the spiked values, and relative standard deviations were calculated.

#### **2.4.9 MALDI-TOF mass spectrometric analysis of soluble capsule precursors**

Synthesis of larger quantities of soluble capsule precursors for mass spectrometric analysis was achieved by a 10-fold upscale of the synthesis reaction. Reaction mixtures were quenched by acidification to pH  $\sim$ 2.5 (addition of 3–5  $\mu\text{l}$  20% (v/v)  $\text{H}_3\text{PO}_4$ ). Precipitated material was removed by centrifugation (13,000 rpm, 10 min, RT; MiniSpin<sup>®</sup> microcentrifuge, rotor F-45-12-11; Eppendorf, Hamburg, Germany). UDP-coupled monosaccharides were purified from salts, detergents and cofactors by ion-pair reversed-phase high pressure liquid chromatography using a Gilson HPLC system equipped with Trilution LC software (Gilson, Limburg an der Lahn, Germany). Sample separation was performed under isocratic conditions (20 mM triethylammonium acetate, pH 7, 1 ml/min) on a Hypersil<sup>™</sup> octadecyl silane (C18) column (particle size 5  $\mu\text{m}$ , 250 mm x 4.6 mm; Schambeck SFD, Bad Honnef, Germany), with detection set at 260 nm. Fractions containing UDP-D-GlcNAc and its derivatives were lyophilized to dryness (Beta freeze dryer; Martin Christ Gefriertrocknungsanlagen, Osterode am Harz, Germany), resuspended in methanol–chloroform (1:1, v/v) and subjected to matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS).

MALDI-TOF MS analyses were performed as follows: Samples were spotted onto a ground steel MALDI-TOF target plate and allowed to dry at room temperature. Subsequently, each sample was overlaid with 1  $\mu\text{l}$  of matrix (saturated solution of



6-aza-2-thiothymine in 50% (v/v) ethanol, 20 mM diammonium hydrogen citrate) and air dried at again. Spectra were recorded in the reflector negative mode at a laser frequency of 9 Hz within a mass range from 300 to 3000 Da on a Biflex™ III mass spectrometer (Bruker Daltonik, Bremen, Germany). Data analysis was performed using flexAnalysis™ software (Bruker Daltonik).

#### **2.4.10 *In vitro* kinase assays**

*In vitro* phosphorylation of different purified proteins (compare reference <sup>161</sup>) was carried out at 30°C in a total volume of 15 µl containing 2 µg of the recombinant target protein, 0,5 µg of CapA1-His<sub>6</sub> and CapB2-His<sub>6</sub>, 1 mM DTT, 1 mM EDTA, and 4 µCi [ $\gamma$ -<sup>33</sup>P]-ATP (~300 nM; Hartmann Analytic) in 25 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, pH 7.0. After 30 min of incubation, reactions were stopped by addition of NuPAGE® LDS sample buffer and analyzed by SDS-PAGE (see 2.4.2). SDS-PAGE separation of phosphorylated proteins was performed using electrophoresis buffer without NuPAGE® reducing agent. Radioactive protein bands were visualized using a storage phosphor screen in a Storm imaging system (GE Healthcare, München, Germany).

#### **2.4.11 Identification of phosphorylation sites by nanoLC-MS/MS**

Nanoscale liquid chromatography coupled to tandem mass spectrometry (nanoLC-MS/MS) was utilized to identify phosphorylation sites on capsule biosynthetic enzymes. For this purpose, recombinant His<sub>6</sub>-tagged proteins were phosphorylated *in vitro* and separated by SDS-PAGE as described above, with omission of radiolabeled ATP from the reaction mixture.

NanoLC-MS/MS and data analysis were performed as follows: For peptide preparation, protein bands were excised from the Coomassie-stained polyacrylamide gels and subjected to tryptic in gel digestion.<sup>249,250</sup> In brief, proteins were reduced with 20 mM DTT, slices were washed with 50 mM ammonium bicarbonate, and proteins were alkylated with 40 mM iodoacetamide. The slices were washed again and dehydrated with acetonitrile. Slices were dried in a vacuum concentrator and incubated with 400 ng sequencing grade trypsin (Sigma-Aldrich) at 37°C overnight. The peptide extract was dried in a vacuum concentrator and stored at -20°C. Dried peptides were dissolved in 10 µl 0.1% (v/v) formic acid (solvent A) and aliquots (1 µl) were injected onto a C18 trap column (20 mm x 100 µm; NanoSeparations, Nieuwkoop, Netherlands). Bound peptides were eluted onto a C18 analytical column

(200 mm x 75  $\mu$ m; NanoSeparations). Peptides were separated in a linear gradient from 0% to 55% solvent B (80% (v/v) acetonitrile, 0.1% (v/v) formic acid) within 40 min at a flow rate of 400 nl/min. The nanoHPLC was coupled online to an LTQ Orbitrap Velos™ mass spectrometer (Thermo Fisher Scientific, Schwerte, Germany). Peptide ions between 395 and 1800  $m/z$  were scanned in the orbitrap detector with a resolution of 30,000 (maximum fill time 400 ms, automatic gain control target at  $10^6$ ). The 25 most intense precursor ions (threshold intensity 5000) were subjected to collision-induced dissociation, and fragments were analyzed in the linear ion trap. Fragmented peptide ions were excluded from repeat analysis for 15 s.

Raw data processing and analysis of database searches were performed with Proteome Discoverer software version 1.40.288 (Thermo Fisher Scientific, Schwerte, Germany). Peptide identification was done with an in-house Mascot server version 2.3 (Matrix Science, London, United Kingdom). MS2 data was searched against *Staphylococcus aureus* strain N315 NCBI nr (release 20130323) and *E. coli* sequences from SWISS-PROT (release 2013\_03). Precursor ion  $m/z$  tolerance was 8 ppm, fragment ion tolerance was 0.6 Da, b- and y-ion series were included. Semitryptic peptides with up to one missed cleavage were searched. The following dynamic modifications were set: Alkylation of Cys by iodoacetamide or acrylamide, phosphorylation (Tyr), and oxidation (Met). The PhosphoRS3.0 node was used for scoring of the phosphosite assignment.<sup>251</sup> Mascot results from searches against SWISS-PROT were sent to the percolator algorithm version 2.04 as implemented in Proteome Discoverer 1.4.<sup>252</sup>

#### **2.4.12 *In vitro* modulation of CapD and CapE catalytic activity by the CapA1B2 kinase complex**

The influence of CapAB-mediated tyrosine phosphorylation on CapD and CapE catalytic activity was examined *in vitro* using purified recombinant proteins. Assays were performed in a total volume of 100  $\mu$ l containing the respective target protein (17  $\mu$ g CapE-His<sub>6</sub> or 3  $\mu$ g CapD-His<sub>6</sub>) and 8.5  $\mu$ g of purified CapA1-His<sub>6</sub> and CapB2-His<sub>6</sub>. CapAB-mediated modulation of CapE catalytic activity was tested in the presence of 3 mM UDP-D-GlcNAc and 1 mM ATP in 10 mM KPi, 10 mM MgCl<sub>2</sub>, pH 7.5 supplemented with 1 mM DTT. Assays for investigation of *in vitro* modulation of CapD enzymatic activity additionally contained 0.8% (v/v) Triton X-100 and 2 mM NADP<sup>+</sup>. Samples were incubated for 10 min to 16 h at 30°C. Reactions were

quenched by heating (5 min, 90°C), then subjected to CE analysis. Control reactions were performed with heat-inactivated (10 min, 100°C) proteins CapA-His<sub>6</sub> and CapB-His<sub>6</sub>.

## 2.5 *In silico* methods

### 2.5.1 Similarity searches and sequence analysis

Protein and gene sequences were retrieved from databases of the National Center for Biotechnology Information (NCBI) ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). Homology searches against SWISS-PROT and the NCBI non-redundant protein database were performed on the NCBI BLAST server<sup>253</sup> ([www.ncbi.nlm.nih.gov/blast](http://www.ncbi.nlm.nih.gov/blast)) using BLASTp with a BLOSUM62 weight matrix<sup>254</sup> and default parameters. Similarity searches with nucleotide sequences were performed using BLASTx (with default settings) against the NCBI non-redundant nucleotide database. The BLASTp algorithm as implemented in the BLAST-2-sequences option<sup>255</sup> was employed to determine percent identity and percent similarity of protein sequences. Global pairwise sequence alignments were performed with the Needleman-Wunsch alignment tool<sup>256</sup> available at the NCBI BLAST server. If not stated otherwise, multiple sequence alignments were carried out using NCBI's COBALT aligner<sup>257</sup> ([www.st-va.ncbi.nlm.nih.gov/tools/cobalt/re\\_cobalt.cgi](http://www.st-va.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi)) with default settings.

Amino acid sequences were analyzed with bioinformatics tools to predict transmembrane helices and potential phosphorylation sites: TMHMM ([www.cbs.dtu.dk/services/TMHMM-2.0/](http://www.cbs.dtu.dk/services/TMHMM-2.0/)), a hidden Markov model-based program,<sup>258</sup> was utilized to predict transmembrane helices; potential phosphorylation sites were predicted using the neural-network-based program NetPhos2.0<sup>259</sup> (available at [www.cbs.dtu.dk/services/NetPhos/](http://www.cbs.dtu.dk/services/NetPhos/)).

### 2.5.2 Identification of *Sugp* 4-keto reductases and phylogenetic tree reconstruction

Potential orthologs for *Staphylococcus aureus* CapN were retrieved from the OrtholugeDB database<sup>260</sup> ([www.pathogenomics.sfu.ca/ortholugedb/](http://www.pathogenomics.sfu.ca/ortholugedb/)). Bacterial sequences with a length of 260–340 aa and a putative nucleotide-binding motif at the N-terminus were considered as potential orthologs. Furthermore, BLASTp searches (see 2.5.1) were performed against the NCBI non-redundant protein database using the amino acid sequences of the (proposed) *Sugp* 4-keto reductases CapN of

*Staphylococcus aureus* (GenBank accession number BAF66380.1), WbpK (AAC45865.1) and WbpV (AAF23991.1) of *Pseudomonas aeruginosa*, WreQ of *Rhizobium etli* (AAQ93037.1), and WbcT of *Francisella tularensis* (CAG46095.1) as query sequences. The query sequences, as well as selected BLASTp hits with *E*-values < 1e-25, identity > 35%, and query length coverage > 75%, were included in the analysis. After removal of duplicates, an initial data set of 402 sequences was obtained. Multiple sequence alignments were performed with ClustalW2<sup>261</sup> using the Gonnet substitution matrix<sup>262</sup> and default parameters (gap opening penalty, 10; gap extension penalty, 0.2; gap distance penalty, 5) at the European Bioinformatics Institute website ([www.ebi.ac.uk](http://www.ebi.ac.uk)). When identical or nearly identical sequences were detected, only a single representative sequence was selected and the other sequences were discarded (with a few exceptions). The sequences were re-aligned, and a distance-based tree was generated from the alignment using the neighbor-joining option implemented in ClustalW2 (default settings; without exclusion of gaps). Based on the clustering results, a set of 184 sequences was chosen for further analysis (SUPPLEMENTARY TABLE 2). As revealed by homology searches, the data set did not only contain potential Sugp 4-keto reductases, but also proteins from dTDP-4-dehydrorhamnose reductase<sup>263,264</sup> and C<sub>55</sub>PP-GlcNAc epimerase<sup>265</sup> protein families.

For further differentiation, a phylogenetic analysis was performed on the Phylogeny.fr platform (<http://phylogeny.lirmm.fr>) as follows: Sequences were aligned with MUSCLE<sup>266</sup> (version 3.7) configured for highest accuracy (default settings). The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program<sup>267</sup> (version 3.0). The default substitution model was selected assuming an estimated proportion of invariant sites and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data. Reliability for internal branches was assessed by an approximate likelihood-ratio test (Shimodaira–Hasegawa-like).<sup>267</sup> TreeDyn<sup>268</sup> (version 198.3) was utilized for tree visualization. In this phylogenetic reconstruction, the homologs from dTDP-4-dehydrorhamnose reductase and C<sub>55</sub>PP-GlcNAc epimerase protein families clustered according to their (proposed) functions with high statistical support (> 0.99). The sequences contained in the respective clades were excluded from the analysis, as well as the sequences of *Nitrosococcus oceani*, *Methylomonas methanica*, *Aromateolum aromaticum* and

*Desulfomicrobium baculatum* forming a separate clade apart from the alleged proteobacterial UDP-Sugp reductase sequences (see SUPPLEMENTARY TABLE 2). Since the proposed UDP-Sugp reductases CapN, WbjF, WbpK, WbpV and WreQ (see above) had been found to align well with the multi-domain models cd05232<sup>269</sup> (PSSMID 187543; “UDP-glucose 4 epimerase, subgroup 4”) and COG0451<sup>270</sup> (PSSMID 223528; “WcaG”), a batch CD-Search<sup>269</sup> (online tool available at [www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi](http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi)) against the NCBI conserved domain database (version 3.12) was performed with the remaining 132 sequences. For most of these protein sequences, cd05232 and COG0451 were the top-ranking domain models. However, three sequences of *Rhizobiales* (NCBI accession numbers YP\_001627391.1, YP\_002732468.1, YP\_001371115.1) were assigned to domain model pfam01370<sup>271</sup> (“epimerase”) instead. They were thus removed from the data set, together with six further sequences found in the same clade in the maximum likelihood tree, which also produced relatively low alignment scores with cd05232 and COG0451 (compare SUPPLEMENTARY TABLE 2).

Phylogenetic analysis was again performed with the reduced data set. Two archaeal sequences (potential CapN orthologs retrieved from OrtholugeDB; GenBank accession numbers ADG90620.1, AGT34472.1) were included as outgroup for tree rooting. Alignments were obtained with MUSCLE, as well as with ClustalW, and phylogenetic trees were constructed from both of the sequence alignments, using either PhyML (see above) or the neighbor-joining method implemented in BioNJ<sup>272</sup> (all analyses were performed at the Phylogeny.fr website). On the basis of these analyses, several “rogue sequences”<sup>273</sup> were identified and removed, thus improving the overall support. A set of 119 proteins was chosen for reconstruction of the final maximum likelihood tree of the Sugp 4-keto reductase family. The phylogenetic reconstruction was performed using MUSCLE and PhyML as described above. The statistical significance of the phylogenetic tree obtained was assessed by bootstrap analysis<sup>274</sup> with 100 replicates.

### 2.5.3 Genomic context analysis of Sugp 4-keto reductases

The genomic neighborhood of the 119 potential UDP-Sugp keto reductases identified in the phylogenetic analysis (see 2.5.2) was surveyed for genes encoding large membrane-bound UDP-GlcNAc 4,6-dehydratases and WbpL- or CapM-type glycosyltransferases. To facilitate the analysis, genomic context networks were

obtained by searching the STRING database<sup>275</sup> (<http://string-db.org>) with the amino acid sequences of the proposed reductases CapN, WbpK, WbpV, WreQ, and WbcT (for accession numbers, see 2.5.2). Furthermore, gene annotation tables were retrieved from the NCBI gene and nucleotide databases, along with the respective gene product sequences. If no ortholog of *capD* and/or *capM/wbpL* was found in the neighborhood of a reductase gene, tBLASTn searches of the respective genome or nucleotide sequence were performed. The BLASTp algorithm as implemented in the BLAST-2-sequences option<sup>255</sup> was employed to assess the similarity between gene product sequences and reference proteins (*S. aureus* Newman CapD and CapM, *P. aeruginosa* WbpL; GenBank accession numbers BAF66370.1, BAF66379.1, and AAF23990.1, respectively). Proteins fulfilling the following criteria were accepted as orthologs: CapD, 540–718 aa length, *E*-value  $\leq 2e-104$ ,  $\geq 61\%$  query coverage, domain model COG1086;<sup>270</sup> CapM, 164–240 aa (the CapM homolog from *Methylococcus capsulatus* is 464 aa long; AAU90617.1), *E*-value  $\leq 6e-32$ ,  $\geq 59\%$  query coverage; WbpL, 319–380 aa, *E*-value  $\leq 1e-19$ ,  $\geq 75\%$  query coverage, domain model cd06854.<sup>269</sup>

### 3 Results

#### 3.1 UDP-*N*-acetylglucosamine 4,6-dehydratases

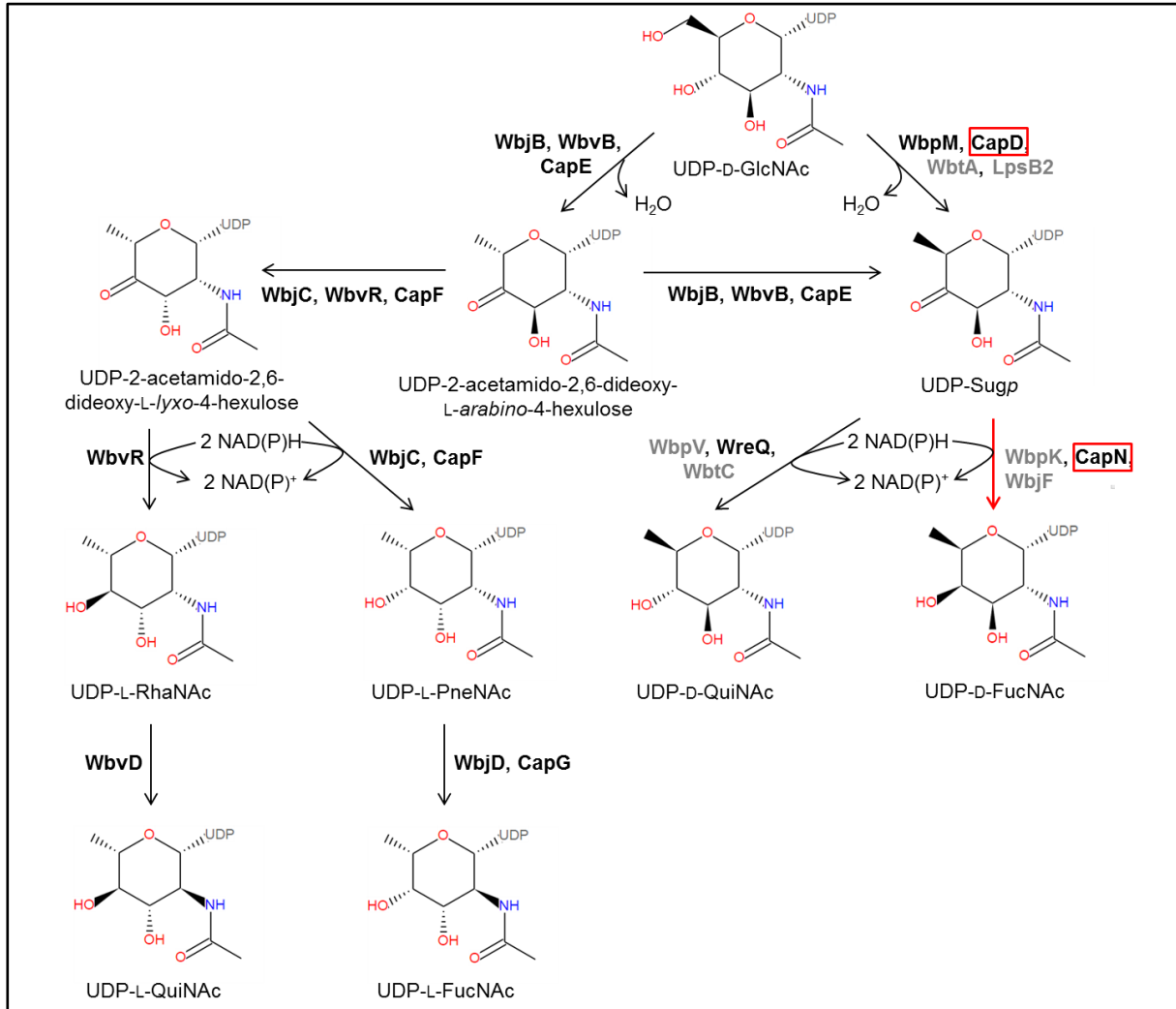


FIGURE 5. Proposed divergent pathways for the biosynthesis of 2-acetamido-2,6-dideoxy-hexoses from the common precursor substrate UDP-D-GlcNAc (chemical structures were retrieved from [www.metacyc.org](http://www.metacyc.org)). Biochemical evidence for the last reaction step in UDP-D-FucNAc synthesis (red arrow) has been lacking so far. In the present study, functional reconstitution of the complete biosynthetic pathway was achieved for the first time using recombinant purified enzymes from *Staphylococcus aureus* serotype 5 (red frames). Names of enzymes that have not been characterized at the biochemical level are shaded grey. Biosynthetic enzymes are from the following species: *Pseudomonas aeruginosa*, WbjB/C/D/F & WbpK/M/V; *S. aureus*, CapD/E/F/G/N; *Francisella tularensis*, WbtA/C; *Rhizobium etli*, LpsB2 & WreQ; *Vibrio cholerae*, WbvB/D/R. D(L)-FucNAc, *N*-acetyl-D(L)-fucosamine (2-acetamido-2,6-dideoxy-D(L)-galactose); D-GlcNAc, *N*-acetyl-D-glucosamine (2-acetamido-2-deoxy-D-glucose); L-PneNAc, *N*-acetyl-L-pneumosamine (2-acetamido-2,6-dideoxy-L-talose); D(L)-QuiNAc, *N*-acetyl-D(L)-quinovosamine (2-acetamido-2,6-dideoxy-D(L)-glucose); L-RhaNAc, *N*-acetyl-L-rhamnosamine (2-acetamido-2,6-dideoxy-L-mannose); Sugp, 2-acetamido-2,6-dideoxy-D-xylo-4-hexulose.

Various rare deoxysugars recognized as important functional components of bacterial cell surface glycans are synthesized from the activated sugar nucleotide precursor UDP-D-GlcNAc. Many of the respective biosynthetic pathways are initiated by the activity of UDP-GlcNAc 4,6-dehydratases.<sup>276,277</sup> These enzymes are members of the extended SDR superfamily (see Introduction, section 1.7) and catalyze 4,6-dehydration of UDP-D-GlcNAc to form a 4-keto-6-deoxy intermediate (FIGURE 5).<sup>278,279</sup>

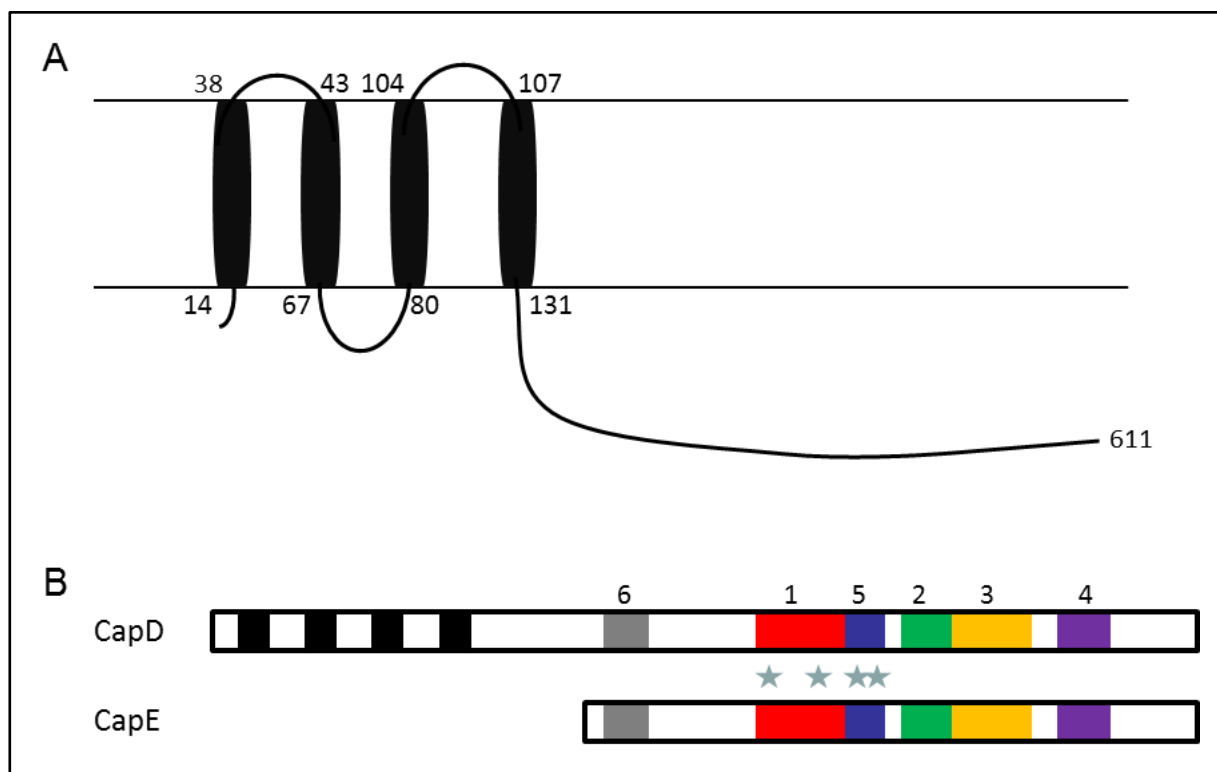


FIGURE 6. Topological model of CapD (A) showing transmembrane domains (black) as predicted with TMHMM 2.0. (B) Schematic representation of CapD and CapE showing six conserved motifs identified using MEME motif discovery software as described by Creuzenet et al.<sup>280</sup> Motifs 1–6 of *S. aureus* Newman CapD (GenBank accession number BAF66370.1) encompass amino acids 289–308 (motif 6), 367–416 (motif 1), 449–477 (motif 2), 478–519 (motif 3), 535–563 (motif 4), and 417–439 (motif 5). The NTMK active site residues are marked by stars.

The *S. aureus* serotype 5 capsule biosynthetic enzymes CapD (GenBank accession number BAF66370.1) and CapE (BAF66371.1) belong to the extended SDR family and possess the typical TGXXGXXG nucleotide-binding signature.<sup>220</sup> CapD and CapE are homologous to a wide range of enzymes involved in the production of surface-exposed polysaccharides in bacterial pathogens. Several of these enzymes



were demonstrated to function as UDP-GlcNAc 4,6-dehydratases, including FlaA1 from *Helicobacter pylori*, PglF from *C. jejuni*, and WbpM from *P. aeruginosa*.<sup>237,280–282</sup> Many of the homologs of CapD and CapE are large membrane proteins (~600 aa) with four N-terminal transmembrane helices, and a cytoplasmic C-terminal half containing five to six conserved sequence motifs arranged in the same pattern. As revealed by prediction of transmembrane topology and multiple sequence alignments, this is also true for CapD (FIGURE 6).<sup>280</sup> However, several much shorter homologs (~350 aa), including CapE, do only correspond to the soluble C-terminal half of CapD.<sup>280</sup> Besides size and cellular localization, the two protein subfamilies seem to differ in function. All soluble enzymes that have been characterized biochemically are UDP-GlcNAc 5-inverting 4,6-dehydratases, whereas all transmembrane proteins investigated so far function as 5-retaining 4,6-dehydratases (i.e. the configuration at C5 remains unaltered during the enzymatic reaction).<sup>207,234,237,280–287</sup>

### 3.2 Comparative genomic and phylogenetic analysis of Sugp 4-keto reductases

The biosyntheses of the nucleotide-activated sugars UDP-*N*-acetyl-D-quinovosamine (UDP-D-QuiNAc) and UDP-*N*-acetyl-D-fucosamine (UDP-D-FucNAc) have been proposed to start from UDP-D-GlcNAc involving two steps (see FIGURE 5).<sup>276</sup> The first step is the generation of the 4-keto-6-deoxy intermediate UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose (UDP-Sugp) by 4,6-dehydratation of UDP-D-GlcNAc. In a second step, the 4-keto group of this intermediate can be reduced either to an axial or to an equatorial hydroxyl moiety by different stereospecific 4-keto reductases, generating UDP-D-QuiNAc or its C4-epimer UDP-D-FucNAc, respectively. Only recently, the synthesis of UDP-D-QuiNAc has been reconstituted *in vitro* using purified recombinant enzymes from *P. aeruginosa* and *Rhizobium etli*.<sup>288</sup> In contrast, biochemical evidence confirming the proposed route for UDP-D-FucNAc biosynthesis is still lacking.

The stereospecific 4-keto reductases which catalyze production of UDP-D-QuiNAc and UDP-D-FucNAc belong to the same protein family (compare reference <sup>289</sup>). Members of this protein family are common among bacteria, but routinely misannotated as UDP-glucose 4-epimerases in the NCBI and SWISS-PROT databases. Starting from an *in silico* analysis of *S. aureus* CapN, an extensive survey

of the protein family was undertaken, to facilitate allocation of gene function in polysaccharide gene clusters. Correct annotation of such reductase genes may not only help to prevent confusion and misinterpretation, but also allow for prediction of structural features of bacterial polysaccharides.

TABLE 9

Sequence similarities of selected bacterial proteins to *Staphylococcus aureus* Newman CapN

Bacterial strain	Protein	GenBank accession number	Sequence similarity to CapN		Reference
			Identity	Overlap <sup>a</sup>	
<u><i>Escherichia coli</i></u>					
O157:H7	Gnu	KFF50896.1	26%	267/331	290
<u><i>Francisella tularensis</i></u>					
SCHU S4	WbtC	CAG46095.1	37%	261/263	291
<u><i>Pseudomonas aeruginosa</i></u>					
PAO1 (O5)	WbpK	AAC45865.1	26%	280/320	287
IATS O6	WbpV	AAF23991.1	28%	284/320	287
PA103 (O11)	WbjF	AAB39483.1	33%	209/314	292
<u><i>Rhizobium etli</i></u>					
CFN42	WreQ	AAQ93037.1	27%	280/309	288
<u><i>Staphylococcus aureus</i></u>					
Newman (serotype 5)	CapF	BAF66372.1	25%	146/369	264
MW2 (serotype 8)	Cap8N	BAB94002.1	98%	295/295	206
M (serotype 1)	Cap1N	AAL26668.1	42%	262/287	293
<u><i>Streptococcus iniae</i></u>					
SF1 (serotype I)	CpsG	AGM98615.1	39%	259/286	294
<u><i>Vibrio cholerae</i></u>					
O139	WbfT	AAC46247.1	29%	285/328	295

Note. Sequence similarities to CapN (GenBank accession number BAF66380.1) were determined using the NCBI BLAST-2-sequences option.

<sup>a</sup>length of overlap with CapN/total number of amino acids

The *S. aureus* serotype 5 capsule biosynthetic enzyme CapN (GenBank accession number BAF66380.1) displays sequence patterns characteristic of the extended SDR family, such as a canonical NTYK active site tetrad, and a TGXXGXXG nucleotide-binding signature found near the N-terminus of the protein.<sup>220</sup> CapN was postulated to function as 4-keto reductase catalyzing the formation of UDP-D-FucNAc by stereospecific reduction of the proposed CapD reaction product UDP-Sugp.<sup>77</sup> Corroborating this assignment, genes encoding proteins with high sequence identity

to serotype 5 CapN were identified in the capsule loci of *S. aureus* serotype 1 and 8 strains, which produce D-FucNAc-containing CPs (TABLE 9).<sup>204,206,293,296,297</sup>

The BLASTp algorithm was used to search the NCBI non-redundant protein database for genes encoding CapN homologs. Astonishingly, most homologs identified in *Firmicutes* species (100 best hits), such as CpsG from *Streptococcus iniae* (see TABLE 9), have a G to S substitution in their nucleotide-binding motif (consensus sequence TGXXSYIG; see SUPPLEMENTARY FIGURE 1). None of these proteins have been characterized so far, but serological differences between *S. iniae* strains have been attributed to mutations in the *cpsG* gene, indicating that the gene product possesses enzymatic activity.<sup>294</sup> The TGXXSYIG motif is also found in some homologous proteins from Gram-negative species, for example in WbtC from *Francisella tularensis*. As deduced from knockout mutant studies, WbtC likely catalyzes the reduction of the 4-keto group of UDP-Sugp to yield UDP-D-QuiNAc, suggesting that this protein is capable to bind NAD(P)H.<sup>291</sup> Notably, a second sequence consistent with a cofactor-binding motif was detected in WbtC, as well as in many other CapN homologs (SUPPLEMENTARY FIGURE 1).

Genetic evidence suggests, that the enzymes WbpK and WbjF, involved in O antigen biosynthesis in *Pseudomonas aeruginosa* O5 and O11 strains, respectively, mediate the same reaction as CapN.<sup>287,292,298</sup> However, CapN shares only moderate identity (~30%) at the amino acid level with the two putative reductases (TABLE 9). Similar results were obtained upon comparison with the protein WreQ from *R. etli*, and with its homolog WbpV from *P. aeruginosa*, enzymes which convert UDP-Sugp to UDP-D-QuiNAc (TABLE 9).<sup>287,288,292</sup> Moreover, CapN shares only 25% sequence identity (over a length of ~150 aa) with the 4-keto reductase domain of *S. aureus* Newman CapF, which is not surprising, since the two enzymes are assumed to reduce stereochemically distinct substrates and exhibit different overall folds.<sup>77,207,264</sup> A BLASTp search against the well-annotated SWISS-PROT database identified the putative UDP-glucose 4-epimerase WbfT from *Vibrio cholerae* O139 and the C<sub>55</sub>PP-GlcNAc 4-epimerase Gnu from *E. coli* O157:H7 as closest homologs of CapN, but sequence identities were rather low (< 30%; see TABLE 9). Notably, Fallarino et al. suggested that WbfT had been falsely annotated.<sup>299</sup> Significant similarities to CapN were also found with proteins from the dTDP-4-dehydrorhamnose reductase family.<sup>263</sup>

To estimate the evolutionary relationships among CapN homologs from different protein families, selected amino acid sequences were aligned with MUSCLE 3.7, and maximum likelihood trees were calculated from the alignments using PhyML 3.0 (see section 2.5). In this phylogenetic reconstruction, the homologs from dTDP-4-dehydrorhamnose reductase<sup>263,264</sup> and C<sub>55</sub>PP-GlcNAc epimerase<sup>265</sup> protein families clustered according to their (proposed) functions (data not shown). The sequences likely corresponding to Sugp 4-keto reductases formed a group together with proteins annotated as UDP-glucose 4-epimerases. However, there is hardly any experimental evidence supporting the assumption that proteins from this group might function as epimerases, and sequence similarities to biochemically characterized UDP-glucose epimerase proteins are fairly low (data not shown).

The clustering results were confirmed using the NCBI conserved domain search tool as described in the “Materials and methods” section. Based on this analysis, a set of 119 potential Sugp 4-keto reductase proteins was chosen for phylogenetic tree assembly. The majority of these proteins were from *Proteobacteria* (76 sequences) and *Firmicutes* (24 sequences). Furthermore, the data set included sequences from the phyla *Bacteroidetes* (8), *Chlorobi* (3), *Fusobacteria* (3), *Actinobacteria* (2), *Cyanobacteria* (1), *Fibrobacteres* (1) and *Spirochaetes* (1). The N[ST]YK active site tetrad was found to be strictly conserved among this set of proteins (a multiple sequence alignment is shown in SUPPLEMENTARY FIGURE 1). A phylogenetic tree calculated by the maximum likelihood method<sup>300</sup> showed five major clades supported by bootstrap values  $\geq 0.75$  (i.e. these clades occurred in at least 75 of 100 random bootstrap replicates), which were labeled I to V (FIGURE 7A).

As seen in FIGURE 7, Clade I is composed exclusively of sequences of *Bacteroidetes*. The proteins found in Clade I have a size of 296–339 aa. Notably, Clade I comprises two putative 4-keto reductases from *Bacteroides fragilis* NCTC 9343. This organism has the capacity to synthesize eight distinct capsular polysaccharides, designated polysaccharide A through polysaccharide H.<sup>301</sup> The proteins identified in the phylogenetic analysis are encoded in the polysaccharide B and C biosynthetic loci, by genes named *wcgW* and *wcfK*, respectively.<sup>302,303</sup> Polysaccharide B is known to contain D-QuiNAc and, based on sequence homologies, polysaccharide C was proposed to contain D-QuiNAc as well.<sup>303,304</sup> However, experimental evidence for the functions of the *wcgW* and *wcfK* gene products is still lacking. This is also true for all other proteins included in Clade I.

FIGURE 7. Maximum likelihood tree of 119 potential *Sugp* 4-keto reductase proteins (A). Full-length amino acid sequences were aligned with MUSCLE 3.7, and phylogenies were calculated from the alignment using PhyML 3.0. The tree was rooted using two archaeal homologs of unknown function (GenBank ADG90620.1, AGT34472.1) as outgroup (not shown). Bootstrap values (100 replicates) are shown for nodes with  $\geq 50\%$  support. Accession numbers from GenBank or NCBI are given for all sequences. Leaves are colored according to their taxonomic affiliation. Horizontal branch lengths are drawn to scale, with the bar indicating 0.3 amino acid substitutions per site. Organisms known to produce D-QuiNAc or D-FucNAc are marked by black and red diamonds, respectively. (B) Schematic representation of the genomic context of bacterial *Sugp* 4-keto reductases. *Sugp* 4-keto reductase genes (symbolized by an "R") co-occur in potential operon structures with genes encoding membrane-bound UDP-D-GlcNAc 4,6-dehydratases ("D") and putative D-FucNAc-/D-QuiNAc-1-phosphate transferases. The glycosyltransferase genes form two distinct groups, typified by *Staphylococcus aureus capM* (black "T") and *Pseudomonas aeruginosa wbpL* (red "T"). Other genes are denoted by an "x" and, in the case of multiple genes, by numbers indicating the number of genes. Direction of transcription is from left to right, unless indicated otherwise by arrows above gene symbols. Genes that are not closely linked ( $> 12$  intervening genes) are separated by commas. D-FucNAc, *N*-acetyl-D-fucosamine; D-QuiNAc, *N*-acetyl-D-quinovosamine; UDP-D-GlcNAc, UDP-*N*-acetyl-D-glucosamine; UDP-*Sugp*, UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose.

Clade II consists of rather short sequences (260–295 aa), which are mainly from *Firmicutes*. Most obviously, Clade II is divided into two moderately well-supported subclades (FIGURE 7). The smaller one (bootstrap support 0.88) contains several putative reductases from staphylococci with canonical nucleotide-binding site,<sup>220</sup> including CapN from *S. aureus* Newman. Moreover, this subclade contains a protein sequence from *Listeria grayei* with altered TGXXSYIG nucleotide-binding motif. All other proteins in the data set comprising the TGXXSYIG motif are found in the second subclade of Clade II (bootstrap support 0.5); this includes various proteins from *Firmicutes*, as well as three proteins from *Gammaproteobacteria*, and a single protein from each of the phyla *Bacteroidetes*, *Spirochaetes* and *Fibrobacteres*. This grouping is not consistent with the established bacterial phylogeny and could be the result of horizontal gene transfer or, in the case of the sequences of *Spirochaetes* and *Fibrobacteres*, of the limited data set used (these sequences were the only representatives of the respective phyla).<sup>305,306</sup> The vast majority of polysaccharides produced by the organisms in Clade II are of unknown structure. As already mentioned, D-FucNAc residues are found in several CP types of *S. aureus*.<sup>204,296</sup> Moreover, D-FucNAc is a component of the *Bacillus halodurans* C-125 teichuronic acid.<sup>307</sup> The only protein in Clade II, the function of which has been elucidated so far, is the reductase WbtC, which catalyzes the final step in UDP-D-QuiNAc synthesis in *Francisella tularensis*.<sup>291</sup>

The protein sequences from *Epsilonproteobacteria* (285–289 aa in length), as well as those from *Fusobacteria* (290–297 aa), clustered to form monophyletic groups, which were named Clade III and Clade IV. Inconsistent with established 16S rRNA phylogenies, the sequences of *Epsilonproteobacteria* seem to be more closely related to the sequences of *Fusobacteria*, than to those of other *Proteobacteria*.<sup>306</sup> This discrepancy might be the result of horizontal gene transfer. None of the proteins in Clade III and Clade IV have been characterized so far. However, one protein sequence in Clade III is derived from a gene located in the LPS biosynthetic locus of *Fusobacterium nucleatum* ATCC 25586, which is known to incorporate D-QuiNAc into its LPS.<sup>308</sup>

Clade V is the largest clade within the UDP-Sugp reductase tree and comprises protein sequences of *Actinobacteria*, *Chlorobi*, *Cyanobacteria*, and *Proteobacteria*, amongst those the sequences of the (proposed) 4-keto reductases WbjF, WbpK, WbpV, and WreQ (TABLE 3). The proteins in Clade V have a length of 300 to

336 aa. The sequence from *Cyanobacterium aponium* found in Clade V is the only representative of the phylum *Cyanobacteria* included in the analysis. Thus, this sequence grouping loosely together with the proteobacterial reductase proteins is likely the result of the limited data set used. The same might be true for the two sequences of *Actinobacteria* forming a deep branch within Clade V. Alternatively, this grouping might result from lateral gene transfer. Several gene transfer events seem to have occurred between members of the phylum *Chlorobi* and *Pseudomonas* species. However, the deeper branches in Clade V are not well-supported, preventing the tracking of the reductase evolutionary history to full extent. Nevertheless, an overlay of polysaccharide structural information onto the phylogenetic tree revealed a correlation between expression of either D-QuiNAc- or D-FucNAc-containing polysaccharides and sequence-based clustering of gammaproteobacterial protein sequences (FIGURE 7A).<sup>309–320</sup>

STRING<sup>275</sup> ([www.string-db.org](http://www.string-db.org)) reported three conserved gene sets widely found with SugaP 4-keto reductase genes. As seen in panel B of FIGURE 7, most putative reductase genes are associated with genes predicted to encode membrane-bound UDP-GlcNAc 4,6-dehydratases. For instance, *wbfT* from *Vibrio cholerae* O139 is located only 3 kb upstream of the gene *wbfY*, which encodes a protein (AAC46251.1) sharing 40% sequence identity with CapD.<sup>295</sup> Furthermore, the majority of reductase genes co-occur with either *wbpL/wbcO* (Clade I, III and V) or *capM* (Clade II, IV and V) homologs (FIGURE 7B), which encode different types of glycosyltransferases (FIGURE 8). Similar to MraY and WecA proteins, *P. aeruginosa* WbpL and its homologs, such as WbcO from *Yersinia enterocolitica*, are large hydrophobic proteins belonging to the polyprenol phosphate:*N*-acetylhexosamine-1-phosphate transferase family.<sup>321,322</sup> Members of this protein family catalyze glycosyl transfer reactions involving a membrane-associated long-chain prenyl phosphate acceptor (typically C<sub>55</sub>P) and a soluble UDP-*N*-acetyl-D-hexosamine sugar nucleotide donor substrate.<sup>321,322</sup> Though biochemical evidence is lacking so far, WbpL has been proposed to catalyze the transfer of D-QuiNAc/D-FucNAc residues onto C<sub>55</sub>P.<sup>292,323</sup> In contrast, *S. aureus* CapM and homologous proteins (e.g. LpsB1 from *R. etli*) are anchored to the cytoplasmic membrane by a single transmembrane domain that is linked to a large C-terminal catalytic domain located in the cytoplasm (FIGURE 8).

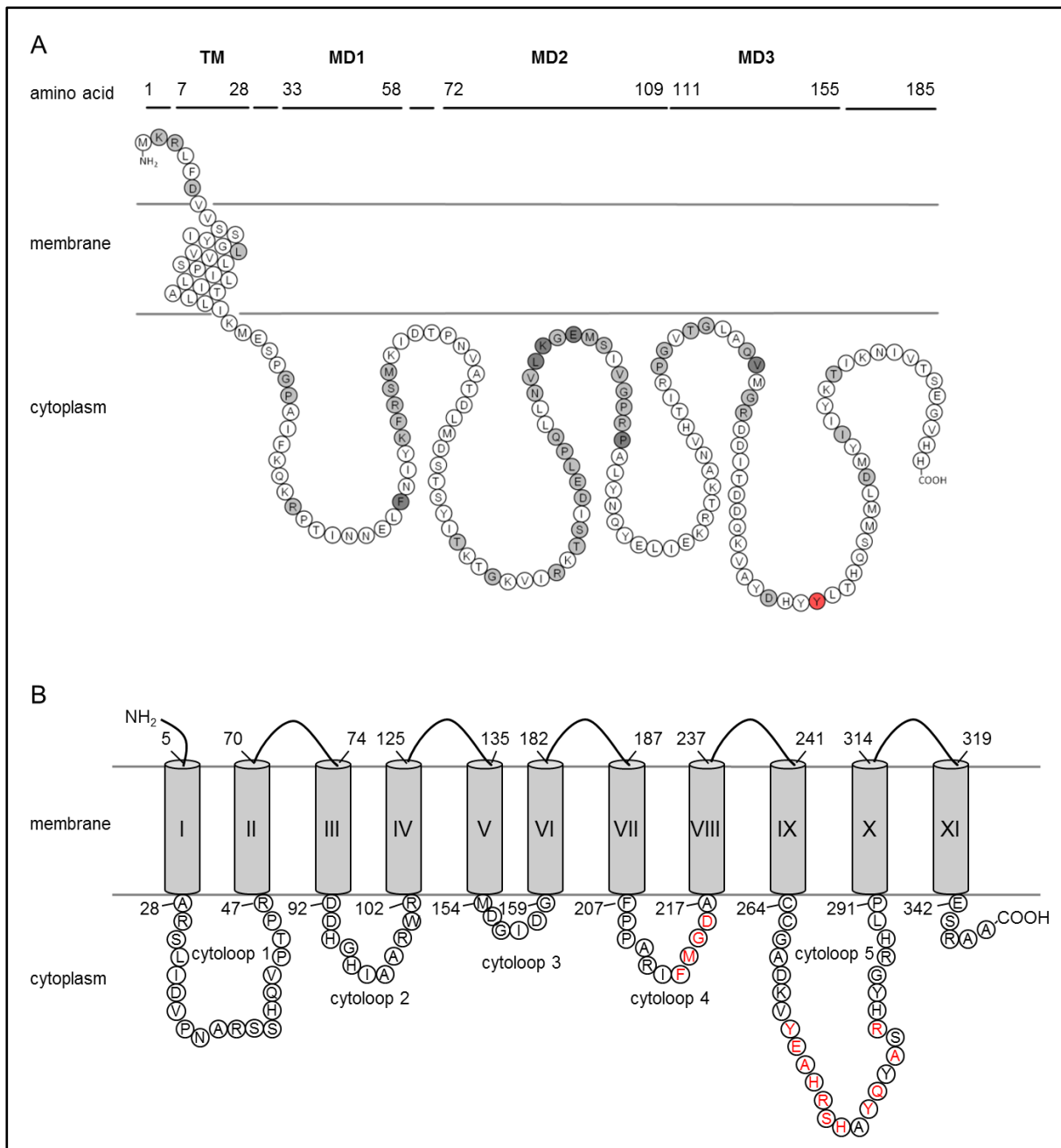


FIGURE 8. Structural models of the putative initiating *N*-acetyl-D-fucosamine-1-phosphate transferases CapM and WbpL based on TMHMM server 2.0 predictions. (A) The *Staphylococcus aureus* glycosyltransferase CapM (GenBank accession number BAF66379.1) is anchored to the membrane by one transmembrane domain (TM). The catalytic domain of CapM is located in the cytoplasm; microdomains 1 and 2 (MD1 & MD2) are predicted to interact with the membrane-standing substrate C<sub>55</sub>P, microdomain 3 (MD3) is assumed to interact with the sugar substrate. The depth of grey shading indicates the degree of residue conservation across species; a potential tyrosine-phosphorylation site is highlighted in red. (B) *Pseudomonas aeruginosa* WbpL (AAF23990.1) is a large hydrophobic protein comprising 11 membrane-spanning domains (designated I through XI), which define five cytoplasmic loops. The conserved FMGD active site motif (cytloop 4) and the residues of the putative carbohydrate-recognition domain (cytloop 5) are highlighted in red.



In most cases, the gene topology is consistent with reductase, dehydratase and glycosyltransferase forming an operon (or being co-transcribed from a bi-directional promoter, as in *V. cholerae* O37). Therefore, it seems likely that the encoded proteins have associated functions.<sup>324–326</sup> Moreover, the gene context analysis strongly supports the assumption that the proteins identified in the phylogenetic analysis are UDP-Sugp reductases involved either in the synthesis of UDP-D-FucNac or of UDP-D-QuiNac. However, only one protein of the whole group has been characterized at the biochemical level, so far.<sup>288</sup> Thus, additional experimental data is highly required to validate the *in silico* predictions.

### 3.3 Construction of *E. coli* expression plasmids

*In vitro* reconstitution of PglF catalytic activity was performed using a soluble derivative of the native *C. jejuni* membrane protein, lacking transmembrane domains, as originally described by Schoenhofen et al.<sup>237</sup> For this purpose, a 5'-truncated version of the *pglF* gene from a clinical isolate of *C. jejuni* was amplified by PCR. The respective oligonucleotide primers (see TABLE 5) were designed based on the nucleotide sequence of the *pglF* gene from *C. jejuni* NCTC 11168 (NCBI GeneID 905411). The amplicon was cloned into pET28a in *E. coli* DH5 $\alpha$ , using the restriction sites *Nde*I and *Xho*I. The genetic construct enabling the expression of an N-terminal hexahistidine fusion protein was analyzed by DNA sequencing (Sequiserve, Vaterstetten, Germany). A GenBank BLASTx search with the deduced amino acid sequence revealed 100% sequence identity to residues 130–590 of the PglF enzyme from *C. jejuni* subsp. *jejuni* 1997-1 (GenBank accession EIB51738.1).

Creuzenet & Lam described an active truncated variant of the membrane-anchored UDP-GlcNAc 4,6-dehydratase WbpM from *P. aeruginosa* PAO1 (GenBank accession number AAC45867.1).<sup>281</sup> This enzyme shares substantial sequence similarity with CapD from *S. aureus* Newman (41% identity and 64% similarity over 446 aa). Thus, a recombinant plasmid for expression of a soluble CapD variant was generated analogously to the genetic construct reported by Creuzenet & Lam.<sup>281</sup> Oligonucleotide primers were designed based on a global alignment of the WbpM and CapD amino acid sequences. Purified genomic DNA from *S. aureus* Newman was used as template for PCR amplification of 5'-truncated *capD*. The purified PCR product was digested with *Nhe*I and *Xho*I, and ligated into appropriately restricted pET28a. After cloning in *E. coli* DH5 $\alpha$ , the recombinant plasmid pET28a-*capDtr* was

confirmed by sequencing. It enables expression of an N-terminally His<sub>6</sub>-tagged CapD mutant (aa 249–611) lacking the four transmembrane domains and the first conserved sequence motif (compare FIGURE 6).

Recombinant expression of CapD-His<sub>6</sub> and CapN-His<sub>6</sub> was performed using plasmids pET24a-*capD* and pET24a-*capN*, which were constructed in the laboratory of Professor Jean C. Lee, at Brigham & Women's Hospital/Harvard Medical School, Boston, USA. Plasmids pET28a-*capA1*, pET28a-*capB2*, and pET24a-*capE* (pKBK50d) used for overexpression of the hexahistidine fusion proteins CapA1-His<sub>6</sub>, CapB2-His<sub>6</sub> and CapE-His<sub>6</sub>, respectively, have been described previously.<sup>207,233</sup>

### 3.4 Expression and purification of recombinant proteins

Full-length CapD from *S. aureus* Newman serotype 5 was overexpressed as N-terminal His<sub>6</sub>-tag fusion protein using *E. coli* C43.<sup>238</sup> After sonication and ultracentrifugation, most of the recombinant protein resided in the membrane fraction. CapD-His<sub>6</sub> could be purified to about 80–90% purity by nickel chelation chromatography after solubilization in DDM. Typically, 5 mg of recombinant protein were obtained per litre of culture. *E. coli* BL21 was employed for overexpression of the soluble truncated CapD variant (CapDtr-His<sub>6</sub>; see previous section). As expected, the mutant protein could be isolated at high yield from the cell lysate without use of detergent.

*S. aureus* CapN was predicted to reside in the cytoplasm, since no putative transmembrane helices were identified using TMHMM server version 2.0 (see 2.5.1).<sup>327</sup> However, attempts to isolate the fusion protein CapN-His<sub>6</sub> from the cytoplasmic content of IPTG-induced *E. coli* C43 cells, did not succeed. However, subjecting the Triton X-100-solubilized membrane fraction to affinity chromatography enabled purification of CapN-His<sub>6</sub>. On the average, 2.5 mg of recombinant protein (> 90% pure) were isolated from 1 l of culture. The finding that CapN is a membrane-associated protein, together with the fact that it is supposed to act on UDP-Sugp, suggests possible interactions between CapN and other membrane-bound capsule biosynthesis proteins, such as CapD and CapM.

*S. aureus* CapA1, CapB2 and CapE, and truncated *C. jejuni* PglF were expressed as hexahistidine fusion proteins and purified by affinity chromatography as previously described.<sup>207,233,237</sup> All recombinant proteins were obtained at high yield and in sufficient purity for use in *in vitro* enzyme assays. The apparent molecular masses of

the proteins as determined by SDS-PAGE were in good agreement with the calculated molecular masses (i.e. CapA1-His<sub>6</sub>: 27.0 kDa, CapB2-His<sub>6</sub>: 27.4 kDa, CapD-His<sub>6</sub>: 70.2 kDa, CapDtr-His<sub>6</sub>: 42.9 kDa, CapE-His<sub>6</sub>: 39.9 kDa, CapN-His<sub>6</sub>: 34.7 kDa, PglF-His<sub>6</sub>: 53.1 kDa; FIGURE 9).

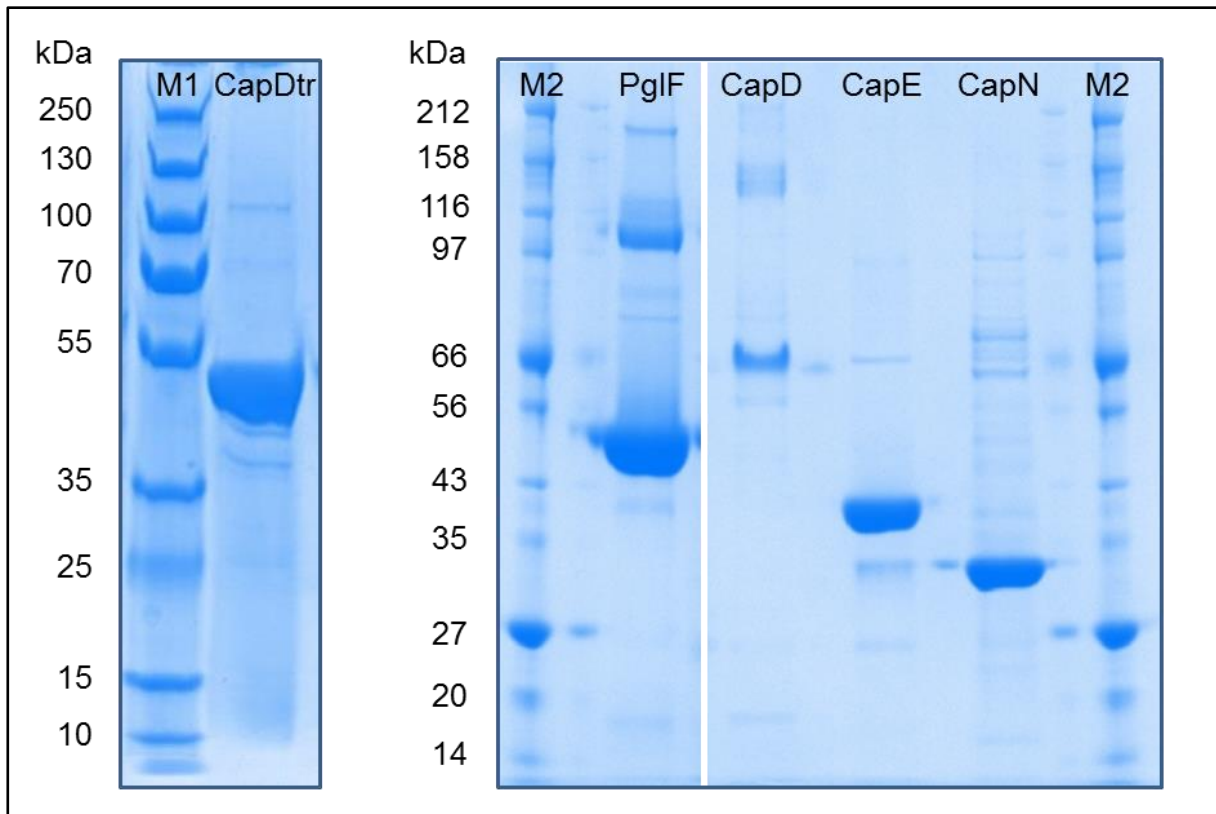


FIGURE 9. SDS-PAGE analysis of purified recombinant enzymes. The molecular masses (in kDa) of the marker proteins are indicated on the left of each panel (M1, PageRuler Plus prestained protein ladder, Thermo Scientific; M2, protein marker broad range, New England BioLabs). Calculated molecular masses of N-terminal hexahistidine (His<sub>6</sub>) fusion proteins are as follows: CapDtr(uncated), 42.9 kDa; PglF, 53.1 kDa; CapD, 70.2 kDa; CapE, 39.9 kDa; CapN, 34.7 kDa.

### 3.5 *In vitro* reconstitution of *S. aureus* UDP-D-FucNAc biosynthesis

Homology searches with CapD and CapN suggested that both proteins are involved in the synthesis of the first soluble capsule precursor UDP-D-FucNAc.<sup>77</sup> Corroborating this assignment, Burrows et al. reported that *cap8D* provided *in trans* can complement the knockout of *wbpM* in *P. aeruginosa*, thus restoring O antigen biosynthesis.<sup>329</sup> However, the pathway proposed for UDP-D-FucNAc biosynthesis in *S. aureus* has yet to be verified biochemically.

*In vitro* characterization of capsule biosynthetic enzymes was performed in close collaboration with the group of Professor Christa E. Müller of the Department of

Pharmacy (Pharmaceutical Chemistry I) at the University of Bonn. To assess the catalytic activity of CapD, a capillary electrophoresis (CE)-based method applying micellar electrokinetic chromatography<sup>330</sup> coupled with UV detection was developed for separation and quantification of UDP-D-GlcNAc and dehydrated derivatives (Wenjin Li, Pharmaceutical Chemistry I). CE is a simple, economic, and powerful method for monitoring enzyme reactions, especially such that generate charged products.<sup>331–333</sup> Based on predicted net charge differences (MarvinSketch software,<sup>334</sup> available at [www.chemaxon.com/free-software/](http://www.chemaxon.com/free-software/)) between UDP-D-GlcNAc and the expected product UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose (UDP-Sugp), CAPS buffer at pH 12.4 was chosen as the running buffer. Since both analytes are negatively charged in alkaline buffer, CE was performed in reversed polarity mode, using fused-silica capillaries, and the cationic surfactant polybrene as running buffer additive for dynamic coating.<sup>333,335</sup> During the running process, polybrene micelles are adsorbed onto the inner capillary wall, thus creating a positively charged surface and causing a reversal of the electro-osmotic flow.<sup>336</sup> With this setup, electro-osmotic flow and migration of anionic compounds are both directed towards the positive electrode.<sup>332</sup>

To confirm its suitability for the investigation of CapD catalytic activity, PglF *in vitro* reactions were analyzed using the new CE method. Truncated, soluble PglF from *C. jejuni* has been demonstrated to catalyze 4,6-dehydratation of UDP-D-GlcNAc, leading to the formation of UDP-Sugp, the same product expected to be generated in the CapD-catalyzed reaction.<sup>237</sup> The reaction involves oxidation at C4 and conversion of the hydroxymethyl group at C6 to a methyl group (see FIGURE 5).<sup>278,279</sup> The resulting 4-keto-6-deoxy intermediate has a reduced charge-to-mass ratio compared to UDP-D-GlcNAc and was therefore expected to display enhanced electrophoretic mobility in the CE system.

As revealed by CE analysis of the PglF reaction, the peak corresponding to UDP-D-GlcNAc disappeared in favor of a new peak (FIGURE 10A); no such conversion was observed with a heat-inactivated negative control (data not shown). As predicted, the PglF enzymatic product migrated faster in the capillary compared to UDP-GlcNAc, as result of a reduced net negative charge. Furthermore, the degradation product UDP was detected in samples containing active PglF. Since 4-keto-6-deoxy sugar nucleotides are relatively unstable compounds, it seems likely that UDP is formed by non-enzymatic hydrolysis of the reaction product.<sup>237,279</sup>

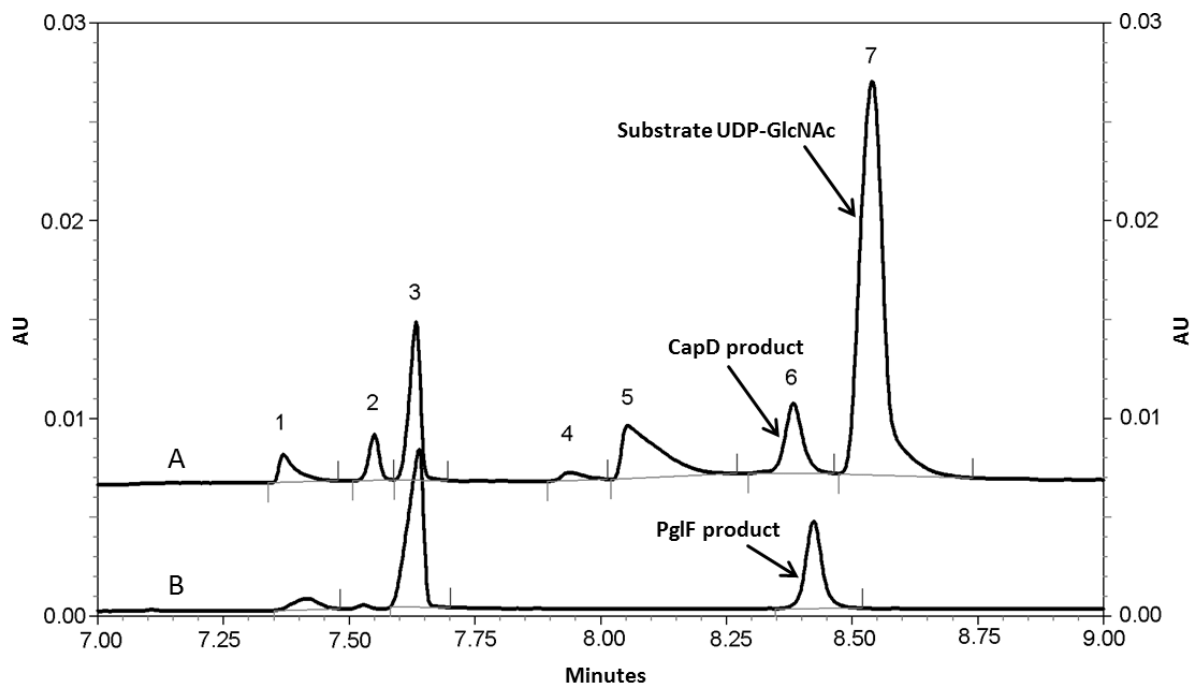


FIGURE 10. Capillary electrophoresis (CE) analysis of *in vitro* UDP-GlcNAc 4,6-dehydratase activity of PglF and CapD. Recombinant purified enzymes were incubated in the presence of substrate UDP-D-GlcNAc (3 mM). Reaction mixtures were analyzed by CE, with UV detection set at 260 nm (arbitrary units, AU). (A) UDP-GlcNAc converted with CapD (3  $\mu$ g) in the presence of 3 mM NADP<sup>+</sup> and 0.8% (v/v) Triton X-100. (B) UDP-GlcNAc converted with PglF (3.5  $\mu$ g). Peak identities were confirmed by spiking samples with reference compounds. 1, UDP; 2, NADP<sup>+</sup> impurity; 3, internal standard PAP (3'-phosphoadenosine 5'-phosphate); 4 and 5, NADP<sup>+</sup> impurities; 6, product UDP-Sugp (UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose); 7, substrate UDP-GlcNAc and cofactor NADP<sup>+</sup>.

To investigate the enzymatic activity of the predicted 4,6-dehydratase,<sup>77</sup> purified recombinant CapD was incubated with UDP-D-GlcNAc in the presence of detergent and cofactors for 2 h. CE analysis of the quenched reaction mixture revealed a peak with identical UV spectrum and retention time as observed for the PglF enzymatic product (FIGURE 10B). The identity of each peak was confirmed by spiking with reference compounds. In contrast to PglF, which displayed full *in vitro* activity without externally added cofactors, CapD activity was dependent on the presence of detergent and NADP<sup>+</sup>; addition of NAD<sup>+</sup> or the absence of any cofactor resulted in less than 5% substrate conversion (see section 3.6, FIGURE 16). Since NADP<sup>+</sup> co-migrated with UDP-GlcNAc in CE analysis, the percentage of substrate conversion could only be assessed indirectly by quantifying the amount of UDP-Sugp and UDP in the samples. According to these calculations, CapD is, like PglF, capable to catalyze almost complete conversion of the substrate UDP-GlcNAc.

MALDI-TOF mass spectrometry was used to further characterize the product of the CapD *in vitro* reaction. To this end, sugar nucleotides were purified from the CapD reaction mixture by ion-pair reversed-phase HPLC as described in the “Materials and Methods” section (2.4.9). Fractions containing sugar nucleotides were lyophilized and subjected to MS analysis; enzymatic product generated with PgIF-His<sub>6</sub> was processed in identical manner and served as standard for MALDI-TOF MS. MS spectra were kindly recorded by Michael Josten (Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn).

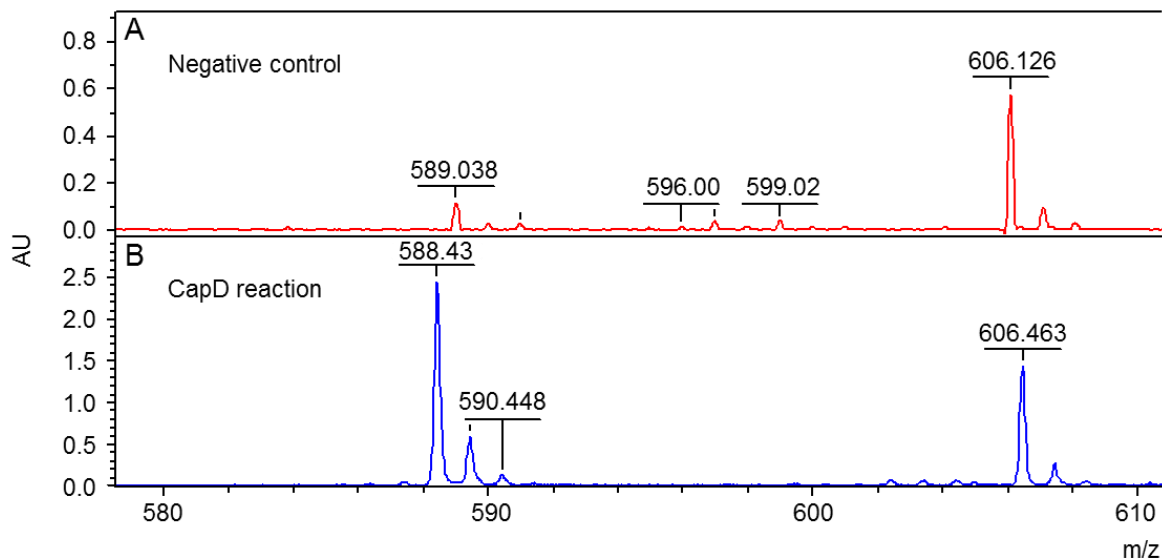


FIGURE 11. Mass spectrometric analysis of *in vitro* UDP-GlcNAc 4,6-dehydratase activity of CapD. Purified recombinant CapD (30  $\mu$ g) was incubated with substrate UDP-D-GlcNAc (0.9 mg) in the presence of 3 mM NADP<sup>+</sup> and 0.8% (v/v) Triton X-100. HPLC-purified sugar nucleotides were analyzed by negative-mode MALDI-TOF mass spectrometry. (A) Mass spectrum obtained from a control reaction without recombinant enzyme. (B) UDP-GlcNAc converted with CapD. The peak at  $m/z$  588.4 corresponds to the deprotonated molecular ion of UDP-Sugp (CapD product). Peaks at  $m/z$  606.5 arise either from UDP-GlcNAc or from the hydrated (diol) form of UDP-Sugp. AU, arbitrary units; UDP-Sugp, UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose.

A molecular ion at  $m/z$  588.4 (negative mode;  $[M-H]^-$ ), consistent with the expected mass of 589.3 for UDP-Sugp, was visible in both MS spectra (FIGURE 11). Although complete substrate conversion was expected under the conditions applied, a second peak at  $m/z$  606.5 was observed with both samples, which could arise from residual substrate UDP-D-GlcNAc (neutral mass 607.4). In aqueous environment, the 4-keto group of UDP-Sugp undergoes hydration, resulting in the formation of a 4-dihydroxy (diol) derivative, that has the same molecular mass as UDP-D-GlcNAc.<sup>279</sup> Thus, the

peak at  $m/z$  606.5 might also correspond to the hydrated form of the PglF/CapD reaction product.

Taken together, these results demonstrate that CapD functions as dehydratase initiating the synthesis of the soluble capsule precursor UDP-D-FucNAc by conversion of the substrate UDP-D-GlcNAc. As demonstrated by *in vitro* assays performed with a soluble CapD variant (see section 3.3), the C-terminal domain is sufficient for catalytic activity (Marvin Rausch, personal communication). It should be noted that accurate mass determination alone does not allow for the differentiation between isomers, and is thus not sufficient to unequivocally confirm the identity of a compound. Taking into account the genetic evidence, and the knowledge gained from biochemical studies on homologous 4,6-dehydratases, it seems highly likely that the CapD reaction product is indeed UDP-Sugp.<sup>237,281,283,284,287,329</sup> However, further experiments should be performed for structure elucidation.

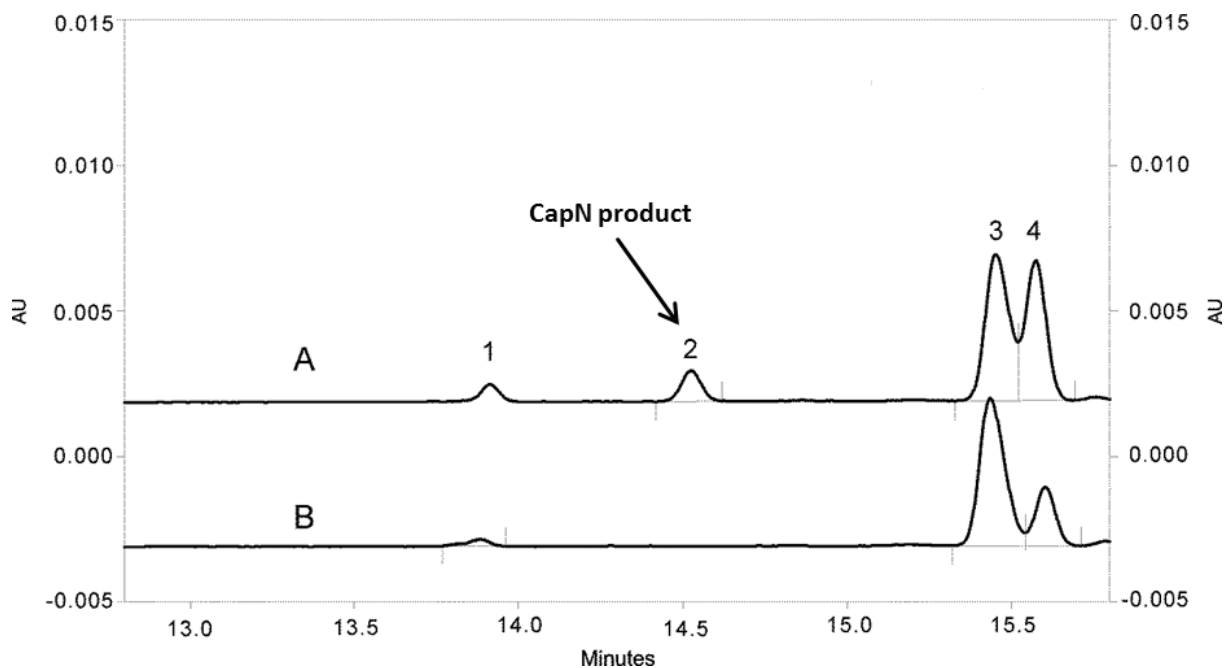


FIGURE 12. Capillary electrophoresis (CE) analysis of *in vitro* UDP-Sugp 4-keto reductase activity of CapN. In a combined assay, purified recombinant CapN (12  $\mu$ g) and PglF (3.5  $\mu$ g) were incubated simultaneously with substrate UDP-D-GlcNAc (3 mM), in the presence of 1.9 mM NADPH and 0.8% (v/v) Triton X-100. Reaction mixtures were analyzed by CE, with UV detection set at 260 nm (arbitrary units, AU). (A) UDP-D-GlcNAc converted by sequential action of PglF and CapN. (B) Electropherogram of a control reaction containing heat-inactivated CapN. Peak identities were confirmed by spiking samples with reference compounds. 1, UDP-D-GlcNAc; 2, UDP-D-FucNAc (CapN product); 3, UDP-Sugp (UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose; PglF product); 4, NADP<sup>+</sup>.

CapN is proposed to further convert the CapD reaction product to UDP-D-FucNAc by stereospecific reduction of the C4 keto group.<sup>77</sup> To assess the catalytic activity of the putative reductase, purified CapN-His<sub>6</sub> was incubated in the presence of the intermediate UDP-Sugp. PglF-His<sub>6</sub> was utilized to produce the substrate for CapN functional reconstitution, as it enabled 100% conversion of UDP-D-GlcNAc and was not dependent on externally added cofactors or detergent for activity. PglF assays were quenched by heating and directly used for CapN *in vitro* reactions, or, alternatively, synthesis of UDP-D-FucNAc was performed in a combined assay.

For analysis of the CapN reaction, a CE method described by Kneidinger et al. was applied.<sup>207</sup> Purified CapN-His<sub>6</sub> was able to convert UDP-Sugp into a new carbohydrate species displaying altered CE migration behavior (FIGURE 12A); no such peak was observed with the heat-inactivated negative controls (FIGURE 12B).

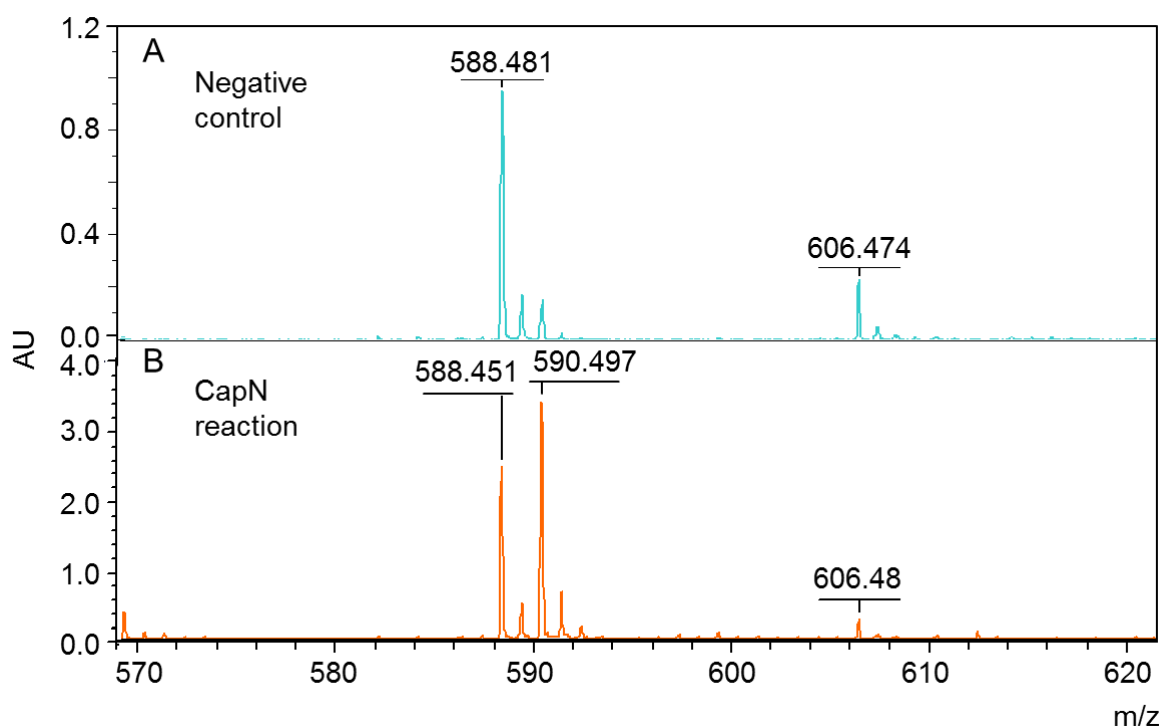


FIGURE 13. Mass spectrometric analysis of CapN *in vitro* activity. In a combined assay, purified recombinant CapN (120  $\mu$ g) and PglF (35  $\mu$ g) were incubated simultaneously with substrate UDP-D-GlcNAc (1.1 mg), in the presence of 1.9 mM NADPH and 0.8% (v/v) Triton X-100. HPLC-purified sugar nucleotides were analyzed by negative-mode MALDI-TOF mass spectrometry. (A) Mass spectrum of a control reaction containing heat-inactivated CapN. (B) UDP-GlcNAc converted with PglF and CapN. The peak at  $m/z$  590.5 corresponds to the deprotonated molecular ion of UDP-D-FucNAc (CapN product). Peaks at  $m/z$  588.5 correspond to the singly charged UDP-Sugp molecule (PglF product); peaks at  $m/z$  606.5 arise either from UDP-GlcNAc or from the hydrated (diol) form of UDP-Sugp. AU, arbitrary units; UDP-Sugp, UDP-2-acetamido-2,6-dideoxy-D-xyl/o-4-hexulose.



In good agreement with the studies of Kneidinger et al. on UDP-L-FucNAc biosynthesis, the CapN enzymatic product migrated faster in the capillary than the PglF enzymatic product, but slower than UDP-D-GlcNAc (FIGURE 12).<sup>207</sup> Analogous results were obtained using the CapD enzymatic product as a substrate for CapN reactions (data not shown).

MALDI-TOF MS analysis of HPLC-purified sugar nucleotides from the CapN reaction mixture revealed an ion ( $[M-H]^-$ ) at  $m/z$  590.4 (FIGURE 13), consistent with the formation of UDP-D-FucNAc (neutral mass 591.4). These results confirm that CapN functions as a reductase and, using NADPH as a cofactor, is able to convert the substrate UDP-Sug $p$  to form the first soluble capsule precursor UDP-D-FucNAc.

### 3.6 Kinetic characterization of CapD- and CapE-catalyzed reactions

The newly established CE method was applied for *in vitro* characterization of the UDP-GlcNAc 5-inverting 4,6-dehydratase CapE (see FIGURE 5). In the CapE enzyme–substrate reaction, the peak corresponding to UDP-D-GlcNAc disappeared, and a faster-migrating peak appeared instead in the CE chromatogram (FIGURE 14).

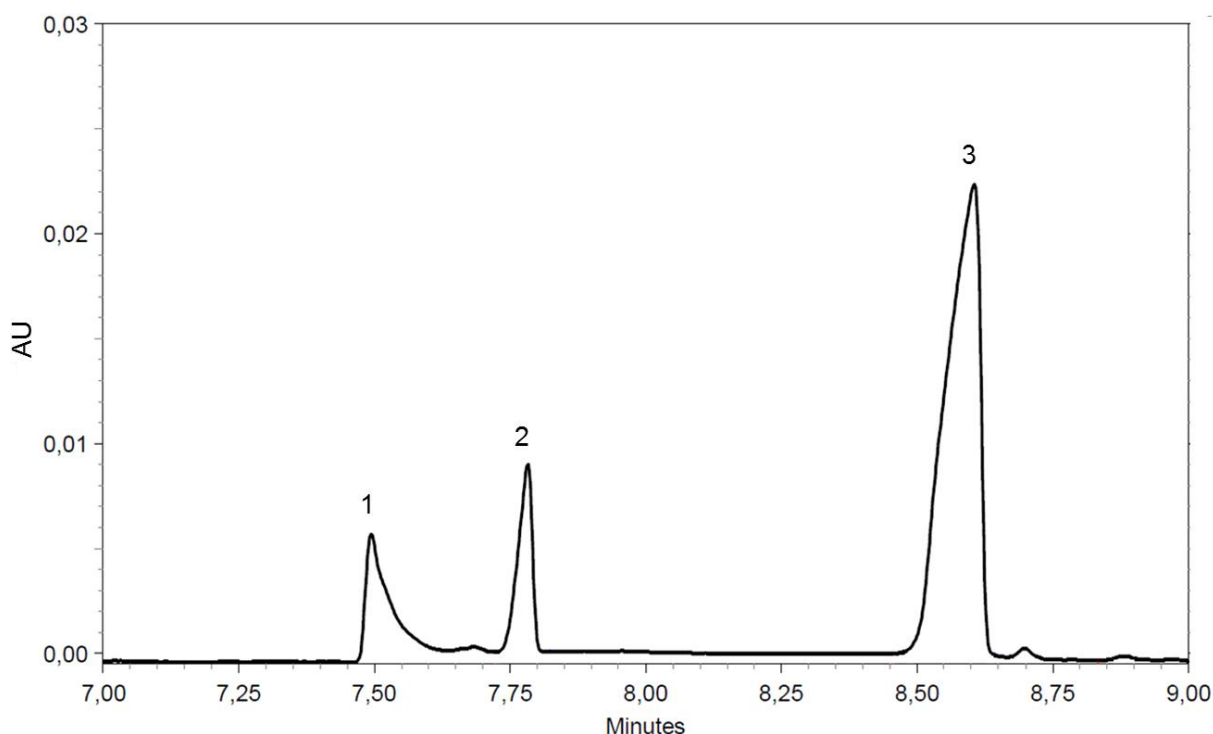


FIGURE 14. Capillary electropherogram of CapE *in vitro* reaction run to apparent completion. Peaks: 1, UDP; 2, Internal Standard PAP; 3, CapE product UDP-2-acetamido-2,6-dideoxy-L-arabino-4-hexulose (likely together with epimeric by-product UDP-Sug $p$ ). AU, arbitrary units; PAP, 3'-phosphoadenosine 5'-phosphate; UDP-Sug $p$ , UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose.

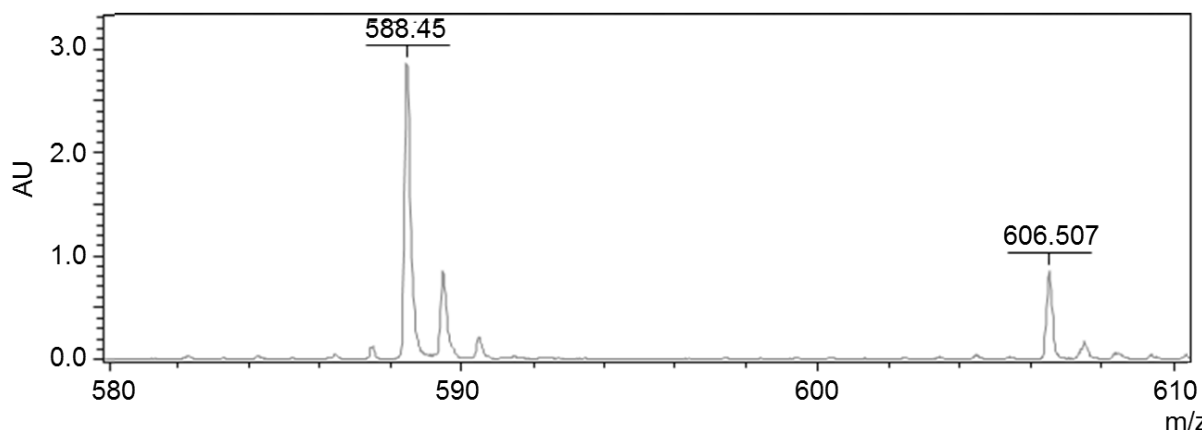


FIGURE 15. Mass spectrum obtained from a CapE reaction. MALDI-TOF MS analysis (negative mode) of HPLC-purified sugar nucleotides gave the expected molecular ion peak for UDP-Sugp ( $[M-H]^-$ ,  $m/z$  588.5). The peak at  $m/z$  606.5 may correspond to the hydrated (diol) forms of these compounds, or to residual substrate UDP-D-GlcNAc. AU, arbitrary units; UDP-Sugp, UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose.

The 4,6-dehydratase/5-epimerase activity of CapE, which converts UDP-GlcNAc to UDP-2-acetamido-2,6-dideoxy-L-*arabino*-4-hexulose, has already been demonstrated *in vitro*.<sup>207,234</sup> CapE was also found to catalyze slow “back-epimerization” of the initial reaction product with L-*arabino*-configuration, leading to formation of the by-product UDP-Sugp.<sup>234</sup> Thus, the CapE product peak on CE analysis probably represents a mixture of the two isomeric UDP-amino sugars. As expected, a molecular mass of 588.5 ( $[M-H]^-$ ) was determined for the HPLC-purified CapE reaction product(s) (FIGURE 15).

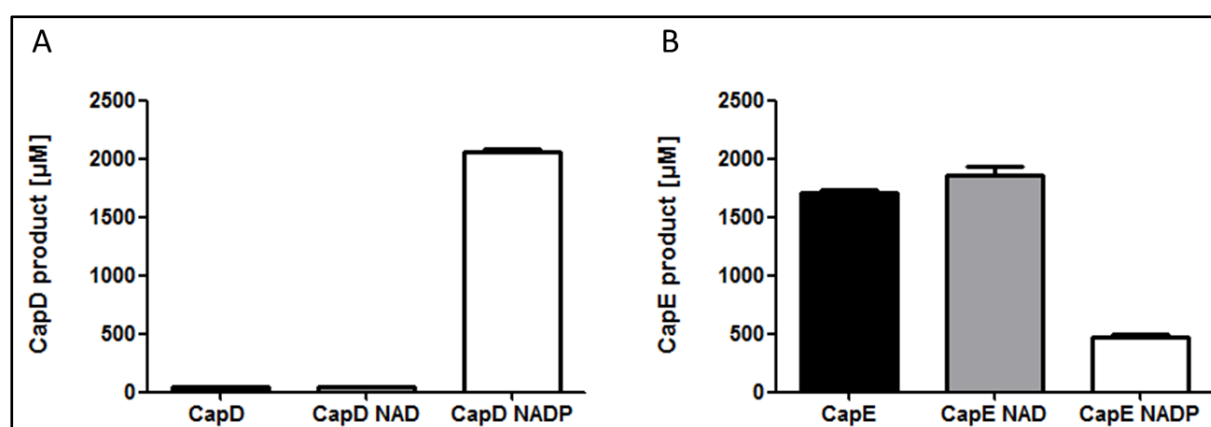


FIGURE 16. Cofactor dependency of CapD and CapE *in vitro* enzymatic activity. Recombinant purified enzymes were incubated with substrate UDP-GlcNAc (3 mM), in the presence or absence of cofactors (3 mM NAD<sup>+</sup>/NADP<sup>+</sup>) and substrate conversion was quantified by capillary electrophoresis (triplicate samples with duplicate measurements). (A) CapD catalytic activity is strictly NADP<sup>+</sup>-dependent. (B) CapE catalytic activity is maximal without exogenous cofactors. Addition of NAD<sup>+</sup> has no effect on CapE-mediated substrate conversion, whereas addition of NADP<sup>+</sup> inhibits CapE *in vitro* activity

CapE-His<sub>6</sub> was able to catalyze complete conversion of the substrate UDP-D-GlcNAc, and was not dependent on externally added cofactors for full enzymatic activity. Astonishingly, CapE enzymatic activity was inhibited when the assay was supplemented with high concentrations of NADP<sup>+</sup> (FIGURE 16).

The CE-based detection method for monitoring UDP-GlcNAc 4,6-dehydratase activity was validated according to the guidelines published by the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, and by the US Food and Drug Administration (FDA), demonstrating its suitability for the analysis of enzyme kinetics and for inhibitor testing.<sup>247,248</sup> A summary of the parameters of method validation is provided in TABLE 8.

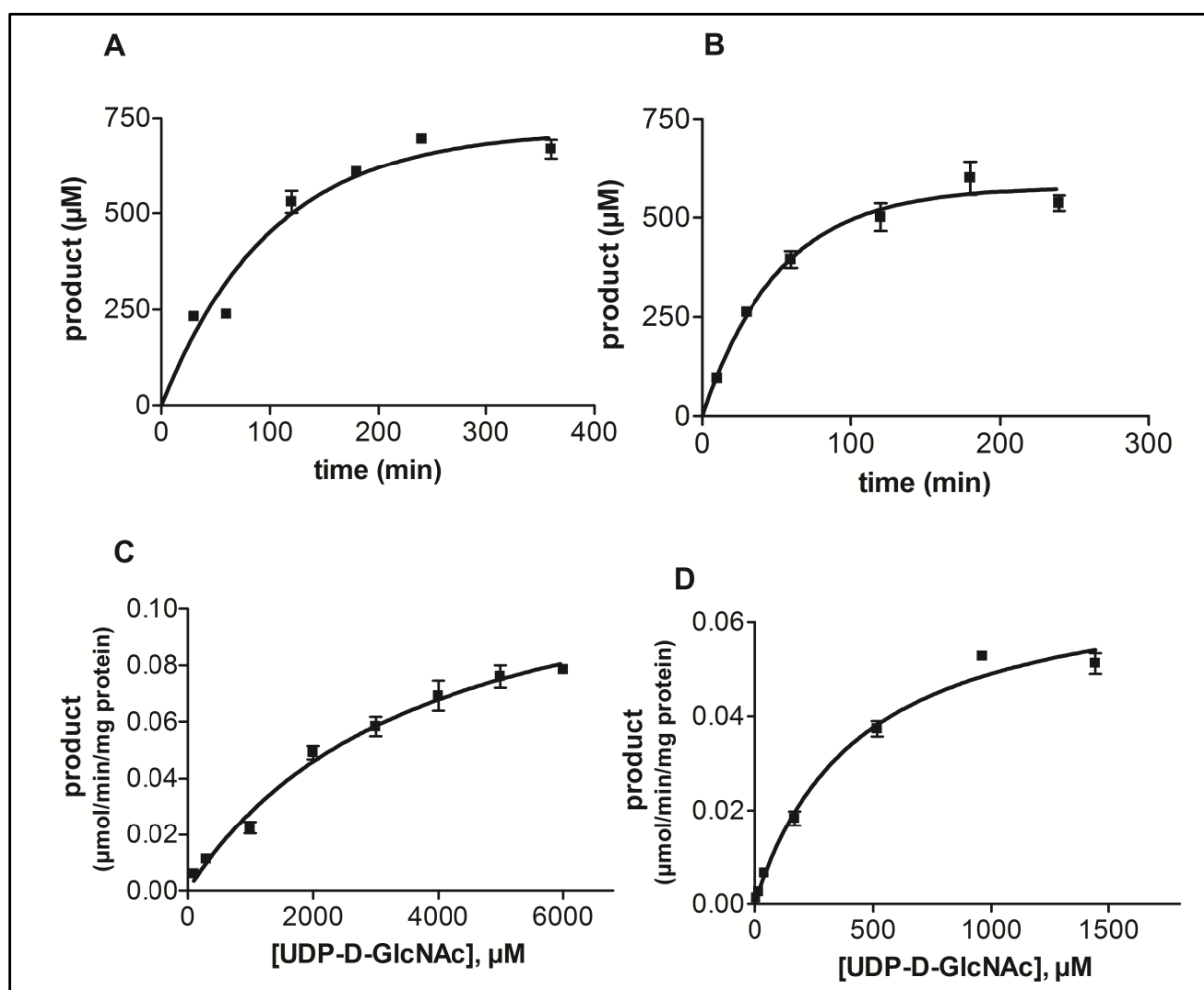


FIGURE 17. Enzyme kinetics of CapD and CapE. Time-dependent enzymatic reaction of CapD (A) and CapE (B). (C) Michaelis–Menten plot for CapD,  $K_m$   $3650 \pm 270$   $\mu\text{M}$ ,  $V_{max}$   $0.127 \pm 0.011$   $\mu\text{mol}/\text{min}/\text{mg}$ ; (D) Michaelis–Menten plot for CapE,  $K_m$   $457 \pm 76$   $\mu\text{M}$ ,  $V_{max}$   $0.071 \pm 0.004$   $\mu\text{mol}/\text{min}/\text{mg}$ .

Enzyme kinetics (Michaelis-Menten constant,  $K_m$ , and maximum velocity,  $V_{max}$ ) were determined for both CapD and CapE (FIGURE 17). CapD displayed a  $K_m$  value of  $3650 \pm 270 \mu\text{M}$  and a  $V_{max}$  value of  $0.127 \pm 0.011 \mu\text{mol}/\text{min}/\text{mg}$ . For CapE, a  $K_m$  value of  $457 \pm 76 \mu\text{M}$  and a  $V_{max}$  value of  $0.071 \pm 0.004 \mu\text{mol}/\text{min}/\text{mg}$  were determined. The CapD *in vitro* reaction was in the linear range (steady-state phase) for up to 120 min, whereas CapE measurements were in the linear range for 60 min under the conditions applied.

### 3.7 Screening for inhibitors of CapD and CapE enzymatic activity

The validated CE method (see TABLE 8) was used for compound screening, in order to identify CapD and/or CapE inhibitors. A series of selected small molecules was tested, including uracil nucleotide derivatives, compounds known to bind nucleotide-binding sites, and clinically applied antibiotics.<sup>331,337,338</sup>

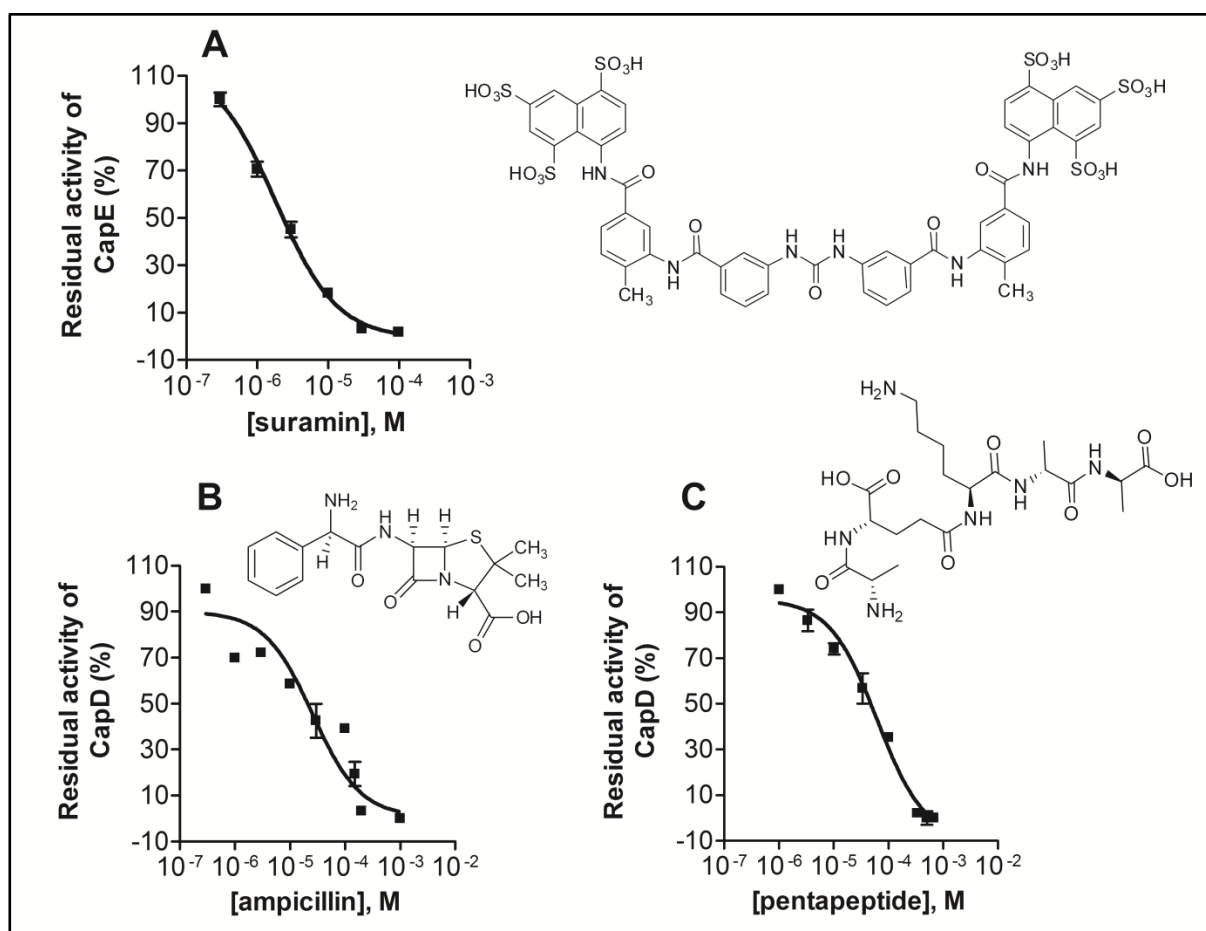


FIGURE 18. Concentration-dependent inhibition of CapD and CapE. (A) CapE inhibition by suramin,  $IC_{50} 1.82 \pm 0.40 \mu\text{M}$ ; (B) CapD inhibition by ampicillin,  $IC_{50} 40.1 \pm 14.9 \mu\text{M}$ ; (C) CapD inhibition by the pentapeptide L-Ala-D-Glu-L-Lys-D-Ala-D-Ala,  $IC_{50} 73.9 \pm 12.8 \mu\text{M}$ . Substrate (UDP-D-GlcNAc) concentration was 1 mM.



competitive inhibitor of CapD, while the peptidoglycan pentapeptide showed a competitive inhibition mechanism at CapD.

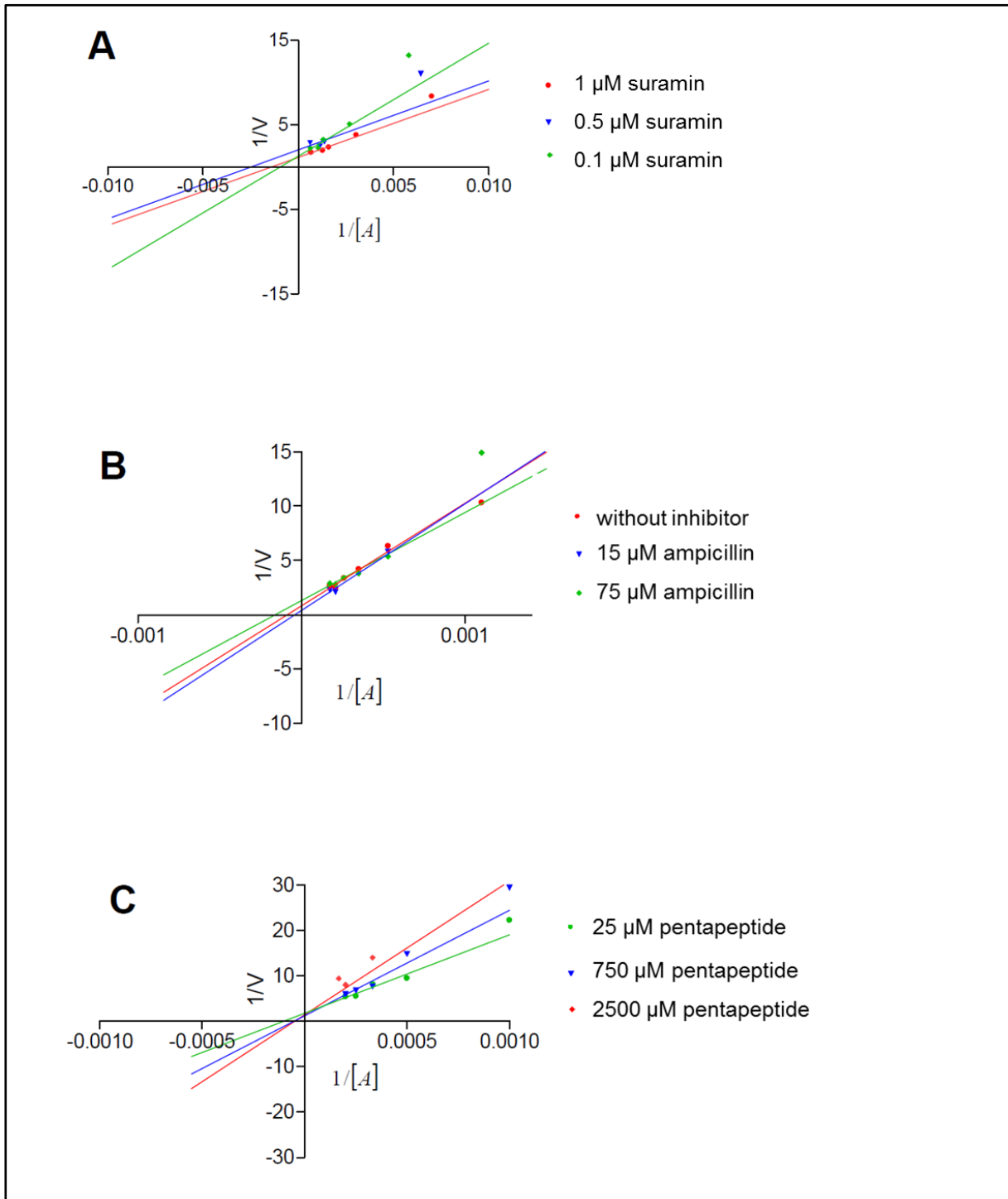


FIGURE 19. Determination of enzyme inhibition mechanism. (A) CapE inhibition by suramin is non-competitive; (B) CapD inhibition by ampicillin is non-competitive; (C) CapD inhibition by the pentapeptide L-Ala-D-Glu-L-Lys-D-Ala-D-Ala is competitive.

### 3.8 CapAB-mediated modulation of *S. aureus* capsule biosynthesis

The reversible phosphorylation of proteins is a fundamental mechanism for regulation of various cellular processes.<sup>341</sup> In *S. aureus*, the bacterial tyrosine (BY-)kinase complex CapAB has been proposed to regulate capsule biosynthesis on the protein level.<sup>160,161,216</sup> So far, one endogenous protein substrate of the CapAB complex has been identified: the dehydrogenase CapO is activated by CapAB phosphorylation, as judged by measuring its capacity to catalyze conversion of NAD<sup>+</sup> to NADH.<sup>216,218</sup>

To identify endogenous protein substrates of the CapAB kinase complex, selected proteins involved in the syntheses of soluble capsule precursors were examined for phosphotransfer (Marvin Rausch, unpublished results and reference<sup>233</sup>). The full length integral membrane protein CapA1 and the cytoplasmic kinase CapB2 were overexpressed as His<sub>6</sub>-tagged fusion proteins, purified by Ni-NTA chromatography, and utilized to reconstitute CapAB tyrosine kinase activity *in vitro* in the presence of  $\gamma$ -labeled [<sup>33</sup>P]-ATP (FIGURE 20A).

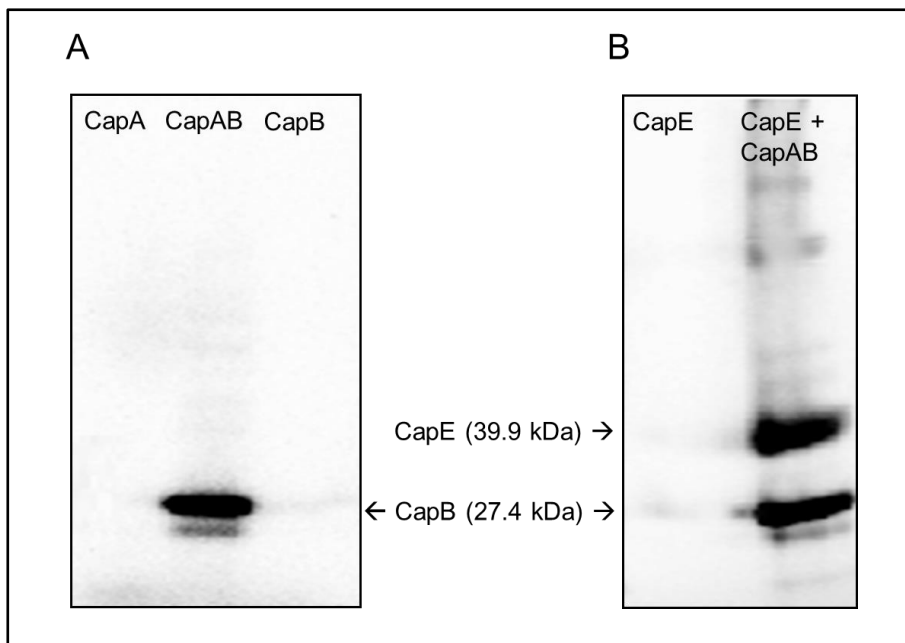


FIGURE 20. CapE is phosphorylated by the CapAB tyrosine kinase complex. (A) Autophosphorylation of CapB2 in the presence of CapA1. (B) CapE is intensely phosphorylated in the presence of CapA1B2. Recombinant purified proteins CapE (3  $\mu$ g), CapA1, and CapB2 were incubated in the presence of  $\gamma$ -labeled [<sup>33</sup>P]-ATP, and phosphotransfer was analyzed by SDS-PAGE and phosphoimaging. The apparent molecular masses of the radiolabeled protein bands were in good agreement with the calculated masses of the respective hexahistidine (His<sub>6</sub>) fusion proteins, i.e. CapB-His<sub>6</sub>, 27.4 kDa; CapE-His<sub>6</sub>, 39.9 kDa.

Investigation of CapA1B2-mediated phosphotransfer onto putative target proteins revealed tyrosine phosphorylation of the dehydratase CapE (FIGURE 20B). In contrast, only marginal phosphotransfer to CapD, and no phosphorylation of recombinant CapN, CapF, or CapG could be detected (data not shown).

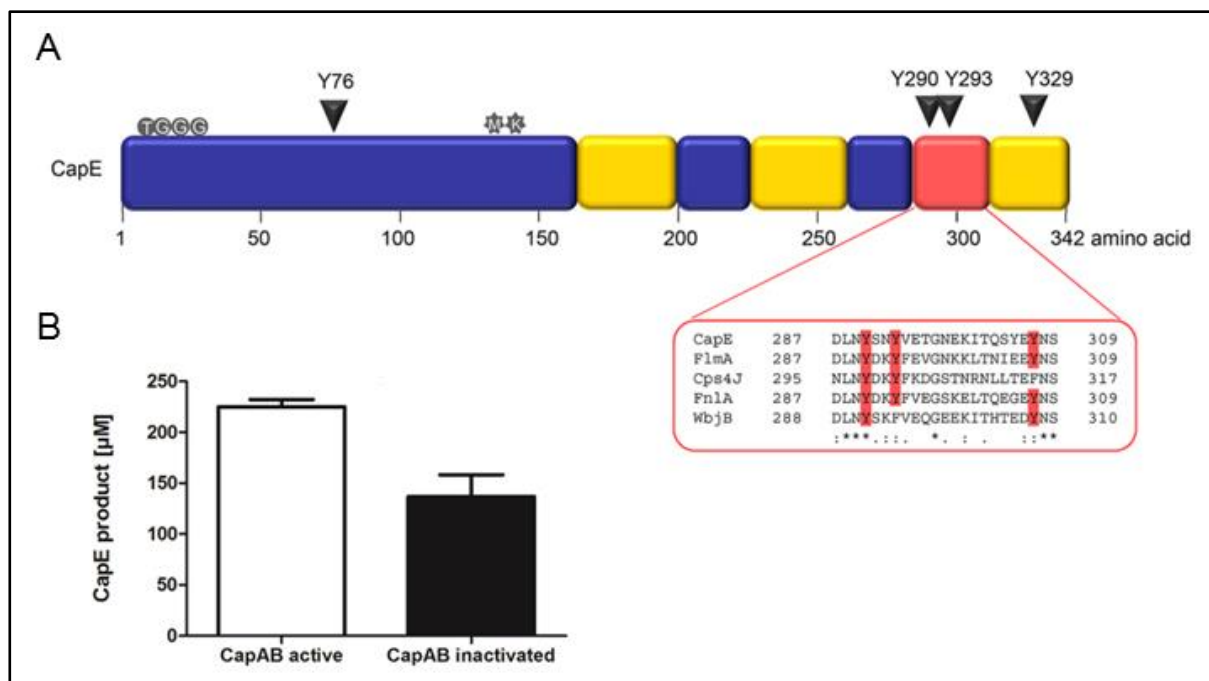


FIGURE 21. CapE enzymatic activity is modulated by tyrosine phosphorylation. (A) Mapping of CapE phosphorylation sites by NanoLC-MS/MS. Confirmed tyrosine phosphorylation sites are marked by arrows. Cofactor-binding domain, substrate-binding domain and the mobile latch are marked blue, yellow and red, respectively. The active site motif MXXXK (stars) and the cofactor-binding motif TGXXGXXG (circles) are highlighted. (B) Impact of CapAB kinase activity on CapE-mediated substrate conversion. CapE was incubated in the presence of either active CapAB (white) or heat-inactivated CapAB (black) with UDP-D-GlcNAc for 30 min. Substrate conversion was quantified by capillary electrophoresis (triplicate samples with duplicate measurements). Statistical significance was analyzed by an unpaired *t*-test ( $p = 0.0029$ ).

Using the NetPhos 2.0 prediction software ([www.cbs.dtu.dk/services/NetPhos/](http://www.cbs.dtu.dk/services/NetPhos/)),<sup>259</sup> eight potential tyrosine phosphorylation sites were identified within the CapE amino acid sequence, with Tyr76 having the highest probability score (FIGURE 21). Tryptic fragments of *in vitro* phosphorylated CapE-His<sub>6</sub> were analyzed by nanoscale liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) to experimentally validate the *in silico* results. LC-MS/MS analysis was performed at the MS service lab of Bonn University, by Dr Marc Sylvester. In good agreement with the computational predictions, Tyr76 of the CapE protein was found to be phosphorylated. As revealed by sequence analysis, this residue is located in proximity to the conserved



TGXXGXXG cofactor-binding motif and the MXXXK catalytic site (FIGURE 21A).<sup>220</sup> Three other phosphotyrosine residues (Tyr290, Tyr293 and Tyr329), which had also been predicted by NetPhos 2.0, were mapped to a C-terminal loop of CapE (FIGURE 21A; red).

More recently, crystal structure analysis of CapE revealed that the protein assembles into a hexamer (trimer of dimers), and that the C-terminal mobile loop connects two contiguous subunits of CapE, involving contacts to Tyr290 and Tyr293 (see Discussion, FIGURE 27).<sup>228</sup> Moreover, the mobile loop appears to occupy a strategic position in close proximity to the entrance of the substrate-binding pocket, and its phosphorylation might enhance the access of UDP-D-GlcNAc to the active site.<sup>228</sup> No phosphotyrosine residues could be identified when CapE-His<sub>6</sub> was subjected to MS analysis without prior *in vitro* phosphorylation, demonstrating that the recombinant enzyme is not modified by protein kinases of the *E. coli* expression host.

To study the impact of CapAB-mediated phosphorylation on CapE enzymatic activity, CapE assays were performed in the presence of either active or of heat-inactivated CapAB kinase complex, and product yields were quantified by CE. As shown in panel B of FIGURE 21, the product yield was increased by 60% in the presence of CapAB, compared to the control reaction, indicating that the CapE-catalyzed reaction is positively modulated through phosphorylation.

Taken together these results support the hypothesis, that the CapE reaction represents a regulatory checkpoint within the CP biosynthesis pathway, with CapE Tyr290 (NetPhos probability score: 0.547) and Tyr293 (NetPhos probability score: 0.837) being the most promising candidates for regulatory tyrosine phosphorylation sites. Moreover, the experimental data presented here identify UDP-GlcNAc 4,6-dehydratases belonging to the extended SDR protein family as a further enzyme class that is regulated by BY-kinases.<sup>219,220</sup>



## 4 Discussion

### 4.1 Inhibitors of capsule biosynthesis as potential anti-virulence agents

The number of new antibiotics developed and approved has steadily decreased in the past three decades, leaving fewer options to treat infections caused by resistant bacteria.<sup>342</sup> As a consequence thereof, doctors now routinely encounter patients with infections that do not respond to available treatment.<sup>343</sup> Since the output of the drug pipelines is not well-positioned to control the growing number of resistant pathogens, there is an urgent need to develop new treatment strategies. Capsule biosynthesis is not essential for bacterial cell viability; however, inhibitors targeting this process may have the potential for use as anti-virulence drugs. Development of anti-virulence drugs represents an emerging strategy to combat multi drug-resistant bacteria.<sup>344</sup> This strategy is especially attractive since anti-infective drugs, that act by inhibiting virulence rather than bacterial growth, might apply a milder evolutionary pressure for the development of resistance.<sup>344</sup>

Inhibition of capsule formation may unmask patterns on the bacterial cell surface, which are important for pathogen recognition and clearance by the innate immune system (see Introduction, section 1.2). Encapsulated *S. aureus* strains are more virulent than isogenic acapsular mutants in rodent models of staphylococcal bacteremia,<sup>4,345</sup> surgical wound infection,<sup>346</sup> septic arthritis,<sup>37</sup> and subcutaneous and renal abscess formation.<sup>38,346</sup> *In vitro* assays revealed that capsule-deficient mutants are more efficiently phagocytosed and killed by human neutrophils<sup>4,345</sup> and mouse macrophages,<sup>37</sup> than the parental strains. Recently, the marine sponge metabolite fascioquinol E, impairing *S. pneumoniae* capsule formation, has been put forward as a novel anti-virulence agent. It could be demonstrated that bacterial cells produced less CP and were more readily captured by murine macrophages upon fascioquinol treatment.<sup>347</sup> Inhibition of capsule formation leading to enhanced complement opsonization and phagocytosis<sup>211,348</sup> could as well have beneficial effects on already established infections.

Inhibitors of capsule formation might also be used in combination with established antibiotics, to extend the life-span of these drugs. The combination of CP inhibitors and cell wall antibiotics appears especially promising for drug development since these compounds may have synergistic effects. CP precursors are assembled onto

the lipid carrier bactoprenol (C<sub>55</sub>P),<sup>122</sup> which is also crucial for biosynthesis of the cell wall polymers peptidoglycan (PG) and wall teichoic acid (WTA). Unlike PG, WTA is not required for the viability of *S. aureus* cells.<sup>121,349</sup> Nevertheless, inhibition of the late-stage WTA biosynthesis reactions leads to cell death, probably due to accumulation of toxic intermediates, or due to depletion of the universal lipid carrier C<sub>55</sub>P, resulting in impaired PG synthesis.<sup>121</sup> WTA inhibitors have therefore been proposed as potential anti-MRSA  $\beta$ -lactam combination agents.<sup>350,351</sup> In analogy to the WTA biosynthesis in *S. aureus*, loss-of-function mutations in enzymes involved in *S. pneumoniae* CP polymerization and transport are lethal unless the function of the initiating glycosyltransferase Cps2E is disrupted by suppressor mutations abolishing flux of C<sub>55</sub>P into capsule biosynthesis.<sup>123</sup>

To explore constituents of the capsule biosynthetic pathway as drug targets, we need a profound understanding of the underlying molecular processes. The present study contributes to the elucidation of capsule biosynthesis in the important opportunistic pathogen *S. aureus*. The enzymatic reactions involved in the syntheses of the soluble capsule precursors UDP-L-FucNAc and UDP-D-ManNAcA have already been studied *in vitro*.<sup>38,207,208</sup> However, *in vitro* characterization of the biosynthesis of the first soluble capsule precursor UDP-D-FucNAc has long been hampered by the fact that CapD, the enzyme catalyzing the initiating step, is a protein with multiple transmembrane domains (FIGURE 6). This thesis describes the successful reconstitution and thorough characterization of CapD enzymatic activity *in vitro* using recombinant, full-length CapD. The established *in vitro* system provided the basis for subsequent reconstitution of the complete biosynthetic pathway, and for inhibitor screening. Furthermore, the present study reports new insights into the post-translational regulation of CP biosynthesis in *S. aureus*.

## 4.2 Characterization of *S. aureus* UDP-GlcNAc 4,6-dehydratases

In the course of evolution, proteins deriving from a common ancestral protein can accumulate sequence divergence, giving rise to families of homologous proteins. Amino acid residues which are important for protein function, stability, or folding are in general less susceptible to mutational alteration, and this is directly reflected in characteristic patterns of conserved residues (for reviews, see references<sup>352–354</sup>). There are many aspects of protein function that contribute to the evolution of protein (sub)families. These may include the global conservation of catalytic mechanisms (in

the case of enzymes), specific binding to substrates and cofactors, as well as the interaction with other proteins in processes such as cell signaling, the regulation of reactions, and the formation of macromolecular complexes.<sup>354</sup> It is commonly accepted that fully conserved amino acid residues are related to functional features common to all family members, whereas positions that are differentially conserved within subfamilies reflect on functional specificity.<sup>354</sup>

The family of UDP-GlcNAc 4,6-dehydratases can be further divided into two subfamilies, one consisting of short soluble enzymes, and the other of large membrane proteins.<sup>280</sup> Besides size and cellular localization, the two protein subfamilies seem to differ in function (see section 3.1). Based on *in silico* predictions and genetic experiments, *S. aureus* CapD was grouped into the subfamily of large, membrane-bound proteins and predicted to function as 5-retaining 4,6-dehydratase.<sup>77,329,355</sup> The results obtained in this thesis provide biochemical evidence for this prediction. Reconstituted CapD converted UDP-GlcNAc to a carbohydrate species with a molecular mass of 589.4, consistent with the formation of UDP-Sugp (FIGURE 10 + FIGURE 11). Taking into account the genetic evidence, and the knowledge gained from biochemical studies on homologous 4,6-dehydratases, it seems highly likely that the CapD reaction product is indeed UDP-Sugp.<sup>237,281,283,284,287,329</sup> However, further experiments should be performed for structure elucidation.

Antisense-mediated depletion of CapD was shown to abolish capsule production in *S. aureus*.<sup>355</sup> Similarly, a naturally occurring mutation in the *capE* gene of *S. aureus* HG001, leading to an M to R substitution in the active site, renders the strain CP-negative.<sup>207,356</sup> *capE* provided *in trans* restored CP expression in HG001.<sup>355</sup> Since *S. aureus* strains defective in CapD or CapE expression are viable, these enzymes are not expected to represent suitable targets for monotherapeutic antibiotic intervention. However, inhibition of CapD or CapE resulting in full repression of capsule biosynthesis might support immune clearance. Thus, compounds targeting these enzymatic steps may have the potential for use as anti-virulence drugs (e.g. for surgical prophylaxis<sup>357</sup>).

The enzyme CapD consists of 4 transmembrane domains (aa 1–130), a linker region (aa 131–281), and a C-terminal catalytic domain (aa 282–611) (FIGURE 6). The biological relevance of the membrane domains is still elusive. It has been demonstrated with soluble variants (i.e. variants lacking transmembrane domains) of

*C. jejuni* PglF and *P. aeruginosa* WbpM, that the cytoplasmic C-terminal domain is sufficient for catalytic activity *in vitro*.<sup>237,281</sup> Expression of the WbpM catalytic domain from a recombinant plasmid restored LPS production in a *wbpM* knock-out mutant. However, the complemented strain failed to produce high-molecular-weight B-band LPS, suggesting that the WbpM membrane domains are crucial for high-level O antigen polymerization *in vivo*.<sup>281</sup> The membrane localization of CapD is especially intriguing since substrate and product of this enzyme are water-soluble and reside in the cytoplasm. In contrast to CapD and homologs from other pathogens, the vast majority of enzymes synthesizing nucleotide-activated sugar precursors (for diverse biosynthetic pathways) are located in the cytoplasm. For instance, all other enzymes involved in the synthesis of soluble CP precursors in *S. aureus* are devoid of membrane-spanning domains (as predicted with TMHMM 2.0, see 2.5.1). Assembly of capsule building blocks involves the transfer of activated sugar monomers to the membrane lipid C<sub>55</sub>P.<sup>122</sup> As revealed by *in vitro* assays using a truncated mutant protein, the transmembrane domains of CapD are not required for catalytic activity per se (see section 3.5). Thus, it seems possible that CapD acts as a docking station, recruiting proteins of the CP biosynthetic machinery to the cytoplasmic membrane, bringing them in close vicinity to the membrane-bound acceptor C<sub>55</sub>P. Moreover, the membrane-spanning domains may be required for the modulation of CapD enzymatic activity (see also 4.6). Changes in the membrane lipid composition during cell growth, such as an increase of cardiolipin,<sup>358</sup> might influence CapD activity. For several membrane proteins, it has been shown that protein–lipid interactions are crucial for *in vitro* or *in vivo* enzymatic activity, as they may affect correct insertion, folding, and topology.<sup>359–361</sup>

The large membrane-bound dehydratases predominantly harbor an altered catalytic tetrad, in which the conserved tyrosine is replaced by a methionine, while the short soluble proteins contain NTMK and NTYK tetrads in roughly equal portions.<sup>362</sup> Furthermore, a few enzymes seem to contain a NTLK tetrad (FIGURE 22). The reason for the emergence of the unusual methionine-based active site, found in CapD as well as in CapE, is still unclear. It has been suggested that its catalytic inertness at physiological pH (in comparison to tyrosine-based active sites) may represent a way to regulate the conversion of the essential cell wall precursor UDP-GlcNAc.<sup>362,363</sup>

A		#	#	#	#	
CapE	95	EAVKTNII	IGTENVLQSAIHQNVKVICLS	TDKAAYPINAMG	ISKAMMEKV	144
FlaA1	102	ECIKTNI	IMGASNVINACLKNAISQVIALS	TDKAAANPINLY	GATKLCSDKL	151
WbjB	95	EAVKTNV	IGTENVLES	AIQNGVKKVVCL	STDKAVYPINAMG	ISKAMMEKV 144
WbvB	95	EAVKTNV	LGTENVLEAAIFHGVERVVCL	STDKAVYPINAMG	ISKAMMEKV 144	
PseB	96	ECIKTNI	IHGAQNVIDACFENGVKKCIALS	TDKACNPVNLY	GATKCLASDKL 145	
B						
CapD	383	EAVRNNI	LGTKN	TAEAAKNAEVK	KFVMISTDKAVNPPNVM	GASKRIAEMI 432
WbpM	399	EGVLNNV	IGTLHAVQAAVQVGVQNFVL	ISTDKAVRPTNVM	GSTRKLAEMV 448	
PglF	366	SAVINNI	LGTKILCDSAKENKVA	KFVMISTDKAVRPTNIM	GCTKRVCELY 415	
WbtA	354	RAIRNNI	LGTKN	AIDLAEAGVESFIL	ISTDKAVRPTNVM	GATKRVCELY 403
LpsB2	378	EGIKNNV	MGTLITARAANKYGVSN	FLISTDKAVRPTNVM	GASKRLAEMV 427	
CapD <sub>xy1</sub>	381	EAVRNNV	LATENVARHCVAAKIQT	FVLISTDKAVEPVNVL	GATKRYAEMV 430	

FIGURE 22. Three variations of the extended SDR family catalytic tetrad are found within UDP-GlcNAc 4,6-dehydratases. The soluble enzymes (A) contain NTYK (canonical) and NTMK tetrads in roughly equal portions, while most membrane-bound dehydratases (B) contain a NTMK active site. A NTLK tetrad is found in a few proteins. Protein sequences were aligned using COBALT; active site residues are highlighted in yellow. GenBank accession numbers are as follows: *Staphylococcus aureus* CapD, BAF66370.1, CapE, BAF66371.1; *Helicobacter pylori* FlaA1, O25511.1; *Pseudomonas aeruginosa* WbjB, AAD45265.1, WbpM, AAD45269.1; *Vibrio cholerae* WbvB, AAM22595.1; *Campylobacter jejuni* PglF, AAD51388.1, PseB, ADT73107.1; *Francisella tularensis* WbtA, AAS60264.1; *Rhizobium etli* LpsB2, ABC93046.1; *Xylella fastidiosa* CapD<sub>xy1</sub>, EWG15203.1. SDR, short-chain dehydrogenase/reductase.

The kinetic data presented here clearly support this hypothesis (FIGURE 17). CapE was found to exhibit similar kinetic characteristics as a FlaA1 mutant protein containing an altered methionine-based active site ( $K_m$  FlaA1 Y141M:  $565 \pm 6 \mu\text{M}$ ;  $K_m$  CapE:  $457 \pm 76 \mu\text{M}$ ).<sup>363</sup> The enzyme kinetics of full-length CapD ( $K_m$  at pH 7:  $3.65 \pm 0.27 \text{ mM}$ ) resemble those reported for a truncated WbpM mutant protein ( $K_m$  at pH 10:  $2.77 \pm 0.007 \text{ mM}$ ).<sup>281</sup> As revealed by crystallographic studies, FlaA1 and CapE protomers self-assemble to form “doughnut-shaped” hexamers in solution.<sup>228,282</sup> This oligomerization is thought to be crucial for “fine-tuning” of the active site conformation and for cofactor-binding.<sup>234,363</sup> Consistent with this hypothesis, a FlaA1 mutant defective in oligomer formation was found to be devoid of catalytic activity.<sup>363</sup> To date, there is no crystal structure of a membrane-bound UDP-GlcNAc 4,6-dehydratase. However, it is well conceivable that oligomerization of CapD and other large homologs may proceed in a different manner, and via different residues.<sup>362</sup> Furthermore, it could be speculated that different modes of oligomer

formation may account for the differences in enzyme kinetics observed between soluble and membrane-bound enzymes.

The observations made in this thesis regarding cofactor requirements for CapD and CapE catalytic activity (FIGURE 16), as well as the published results of others, suggest that different modes of cofactor-binding exist among UDP-GlcNAc 4,6-dehydratases. Some recombinantly expressed enzymes (e.g. CapE) are purified with tightly bound nicotinamide cofactor, and thus do not depend on the addition of exogenous cofactor for activity,<sup>228,237,280</sup> while other proteins (e.g. CapD) require exogenous NAD(P)<sup>+</sup> for catalytic activity.<sup>207,281</sup> These different modes of cofactor binding have been attributed to differentially conserved residues within the cofactor-binding site.<sup>282,362</sup> Within a given constellation of side chains, a single residue can have a significant impact on the selectivity and avidity of an enzyme for a given cofactor.<sup>364–366</sup> For instance, wild-type glucose–fructose oxidoreductase from *Zymomonas mobilis* contains tightly bound NADP<sup>+</sup> as non-dissociable redox cofactor reacting in a “ping-pong type” mechanism; a single amino acid exchange is sufficient to generate an active mutant protein utilizing dissociable NAD(P)<sup>+</sup> as cosubstrate and displaying a sequential reaction type.<sup>364</sup> Ishiyama et al. reported that the soluble UDP-GlcNAc dehydratase FlaA1 contains several positively charged amino acid residues that associate with the 2'-phosphate group of NADP(H), while these interactions seem to be absent in the case of the large homolog WbpM.<sup>362</sup> As revealed by multiple sequence alignments, the respective residues are present in CapE, but absent in CapD (FIGURE 23).

				##	#	
FlaA1	14	LI	TGGTGSF	GKCFVRKVLDTTNAKKIIVYSR	DELKQSEM	52
CapE	8	LI	TGGTGSFG	NAVMKQFLDS-NIKEIRIFSR	DEKKQDDI	45
WbjB	8	LI	TGGTGSFG	NAVLKRFLDT-DIGEIRVFSR	DEKKQDDM	45
WbpM	301	MV	TGAGGSI	GSELCRQIMSC-SPSVLILFEHSEYNLYSI		338
CapD	288	LV	TGAGGSI	GSEICRQVCNF-YPERIILLGHGENSIYLI		325

FIGURE 23. Alignment of UDP-GlcNAc 4,6-dehydratase protein sequences showing the region around the TGXXGXXG nucleotide-binding site (red). The cationic residues R44, D45 and K48 of *Helicobacter pylori* FlaA1 (yellow highlights) have been implied in tight binding of the nucleotide cofactor. Sequences were aligned using COBALT; their GenBank accession numbers are as follows: *Staphylococcus aureus* CapD, BAF66370.1, CapE, BAF66371.1; *H. pylori* FlaA1, O25511.1; *Pseudomonas aeruginosa* WbjB, AAD45265.1, WbpM, AAD45269.1.



However, the activity of recombinant WbjB has been found to depend on the addition of NADP<sup>+</sup>,<sup>207</sup> though this protein contains the residues identified by Ishiyama et al. (FIGURE 23). Thus, in the WbjB protein background, other amino acid residues might play a decisive role in determining cofactor specificity. In contrast to WbjB, CapE was inhibited in the presence of 1 mM NADP<sup>+</sup> (FIGURE 16). One possible explanation for this finding is that NADP<sup>+</sup> acts as competitor of NADPH-binding, and that the concluding reduction step within CapE catalysis is inhibited when the oxidized form of the nicotinamide cofactor is present in excess (compare references <sup>282,285,367</sup>). Similarly, CapE enzymatic activity *in vivo* might be modulated by the cellular NADP<sup>+</sup>/NADPH ratio.<sup>367</sup> Since the coenzyme moiety is not observed in the structure of CapE in complex with substrate analog,<sup>228</sup> it is also conceivable that the inhibition in the presence of NADP<sup>+</sup> is due to interference with substrate binding.

### 4.3 Suga 4-keto reductases—a novel protein family

Phylogenomic inference of protein molecular function is a multistep process involving selection of homologs, multiple sequence alignment, and phylogenetic tree construction, overlaying annotations on the tree topology, discriminating between orthologs and paralogs, and, finally, inferring the function of a protein based on the orthologs identified by this process and the annotations retrieved (for reviews, see references <sup>368,369</sup>). Phylogenomic analysis is thought to be more accurate than function prediction based on mere sequence homology.<sup>369,370</sup> This is especially true for SDR enzymes, for which pairwise-sequence identities are generally low, rendering homology-based function predictions difficult.<sup>219,222</sup> Here, phylogenetic trees were used to assess orthology among homologs of *S. aureus* CapN, leading to the identification of the novel protein family of “Suga 4-keto reductases” (FIGURE 7). The members of the newly defined protein family likely catalyze stereospecific reduction of UDP-Suga, to form either UDP-D-QuiNAc or UDP-D-FucNAc (alternatively, they might catalyze the reduction of lipid-bound Suga; see FIGURE 24). Notably, a similar cluster of proteins with “unknown function” was identified in a phylogenetic analysis of gammaproteobacterial UDP-hexose 4-epimerase protein families conducted by Cunneen et al.<sup>265</sup> The results of Cunneen et al., along with the data presented in section 3.2, suggest a close relationship between Suga 4-keto reductases and UDP-hexose 4-epimerases. UDP-hexose 4-epimerases, such as

*E. coli* GalE, utilize  $\text{NAD}^+$  to alter the configuration of the chiral C4 atom of the saccharide moiety by transient oxidation–reduction, which involves the formation of a transient 4-keto intermediate equivalent to UDP-Sugp.<sup>371</sup> Sugp 4-keto reductases may have evolved from UDP-hexose 4-epimerase proteins by sequence changes stabilizing the conformational shape adopted during the transition state.

The assumption that the proteins identified in the phylogenetic analysis function as Sugp reductases is strengthened by the genomic context analysis, which revealed co-occurrence (in all cases except one) and genomic association of membrane-bound UDP-GlcNAc 4,6-dehydratases, alleged Sugp 4-keto reductases, and two different types of glycosyltransferases (CapM- or WbpL-like; FIGURE 8), suggesting functional, as well as physical interactions between these proteins.<sup>324–326</sup> The arrangement of these genes (i.e. relative order and orientation) in potential operon structures was relatively well preserved within bacterial phyla. Gene order conservation could be due to any of the three following reasons: first, the species have diverged only recently and gene order has not yet been destroyed; second, there has been lateral gene transfer of a block of genes; and third, the integrity of the cluster is important to the fitness of the cell, and thus selectable.<sup>372</sup> Gene pairing was most obvious in the case of reductases and glycosyltransferases, which were encoded by directly adjacent genes in ~90% of analyzed genomes. The exclusive coding patterns of the *capM* and *wbpL* homologs provide an example of non-orthologous gene displacement,<sup>373</sup> the encoded glycosyltransferases, though they are non-orthologous, are supposed to perform analogous functions as priming *N*-acetyl-D-hexosamine-1-phosphate transferases.

Having said that, the question arises: why have bacteria evolved two types of glycosyltransferases that are, at first glance, functionally redundant? A full description of a protein's function ideally requires knowledge of all partner proteins with which it specifically associates, in a direct way, by physical interactions, as well as in an indirect way, e.g. through participation in the same metabolic pathway.<sup>374</sup> Intriguingly, the BY-kinase proteins CapB1 (high confidence, score: 0.84) and CapB2 (medium confidence, score: 0.572) were identified as functional partner of CapM based on co-occurrence and gene neighborhood in different genomes ([www.string-db.org](http://www.string-db.org)).<sup>275</sup> No such association was reported using the *P. aeruginosa* WbpL protein sequence as input for STRING, suggesting that the presence of different glycosyltransferase types reflects differences in the regulation of polysaccharide production.

#### 4.4 *In vitro* enzymatic synthesis of UDP-D-FucNAc

The epimeric dideoxy sugars D-QuiNAc and D-FucNAc (for structures, see FIGURE 5) are present in various polysaccharides important for bacterial virulence (see Results, section 3.2), but are absent from human glycans. Thus, their biosynthetic pathways might serve as targets for antibacterial drug action. Furthermore, oligosaccharide antigens containing these amino sugars may be interesting candidates for vaccine development.<sup>375</sup> Only recently, the synthesis of UDP-D-QuiNAc has been reconstituted *in vitro* using a purified recombinant version of the enzyme WreQ of *R. etli* for reduction of the key intermediate UDP-Sugp.<sup>288</sup> In the present study, *in vitro* enzymatic synthesis of UDP-D-FucNAc was achieved for the first time, by functional reconstitution of the Sugp 4-keto reductase CapN from *S. aureus* serotype 5.

Recombinantly expressed CapN resided in the membrane fraction and could be purified by nickel chelation chromatography after extraction with Triton X-100 (FIGURE 9). Subsequent removal of the detergent by dialysis resulted in precipitation of the protein. Similar observations have been reported for recombinant WreQ.<sup>288</sup> One possible explanation is that CapN and homologous proteins are membrane-associated, and that the interactions with the lipid bilayer are crucial for the conformational stability of these proteins.<sup>288</sup> Moreover, heterocomplex formation with partner proteins (e.g. glycosyltransferases) may be required for protein stability.<sup>376</sup> Corroborating the *in silico* predictions, purified serotype 5 CapN protein catalyzed NADPH-dependent reduction of UDP-Sugp, producing a nucleotide sugar species with a molecular mass of 591.5 (FIGURE 12 + FIGURE 13), consistent with the formation of UDP-D-FucNAc. Since the composition of *S. aureus* CP5 is known,<sup>204,205</sup> and because previous biochemical studies have elucidated the biosynthetic routes leading to the synthesis of the other soluble capsule precursors, UDP-L-FucNAc and UDP-D-ManNAc,<sup>38,207,208,286</sup> it can be concluded that the reaction product is UDP-D-FucNAc.

The pathways for the biosynthesis of D-FucNAc and D-QuiNAc have been subject of a debate, lately. *P. aeruginosa* O5 strains have D-FucNAc at the reducing terminus of their O antigen,<sup>310</sup> while serotype O6 strains have D-QuiNAc at the same position.<sup>309</sup> Nevertheless, *wbpL*<sub>O5</sub> is able to complement the loss of *wbpL* in serotype O6.<sup>292</sup> This has been attributed to donor substrate promiscuity of WbpL.<sup>292</sup> Similarly, a lack of selectivity in CapM-like glycosyltransferases might represent an explanation for the

incorporation of Suga instead of D-QuiNAc seen in the *R. etli* mutant strain CE166.<sup>377</sup> However, it is also possible that WbpL and CapM are specific for the donor substrate UDP-Suga, and that the reduction step in D-QuiNAc/D-FucNAc biosynthesis occurs after transfer of Suga to C<sub>55</sub>P (FIGURE 24B).<sup>287</sup> This would allow the glycosyltransferases to be selective, yet still provide a mechanism by which different amino sugars may be incorporated at the first position of the repeating unit. Furthermore, this would represent a very economical way for enlarging the diversity of O antigen structures and for phase variation. Notably, a close resemblance between Suga 4-keto reductases and C<sub>55</sub>PP-GlcNAc epimerases was noted in this study, as well as by others.<sup>265,290</sup> C<sub>55</sub>PP-Suga being the natural substrate for Suga 4-keto reductases might also explain the low catalytic efficiency for reduction of UDP-Suga observed with WreQ,<sup>288</sup> as well as with CapN.

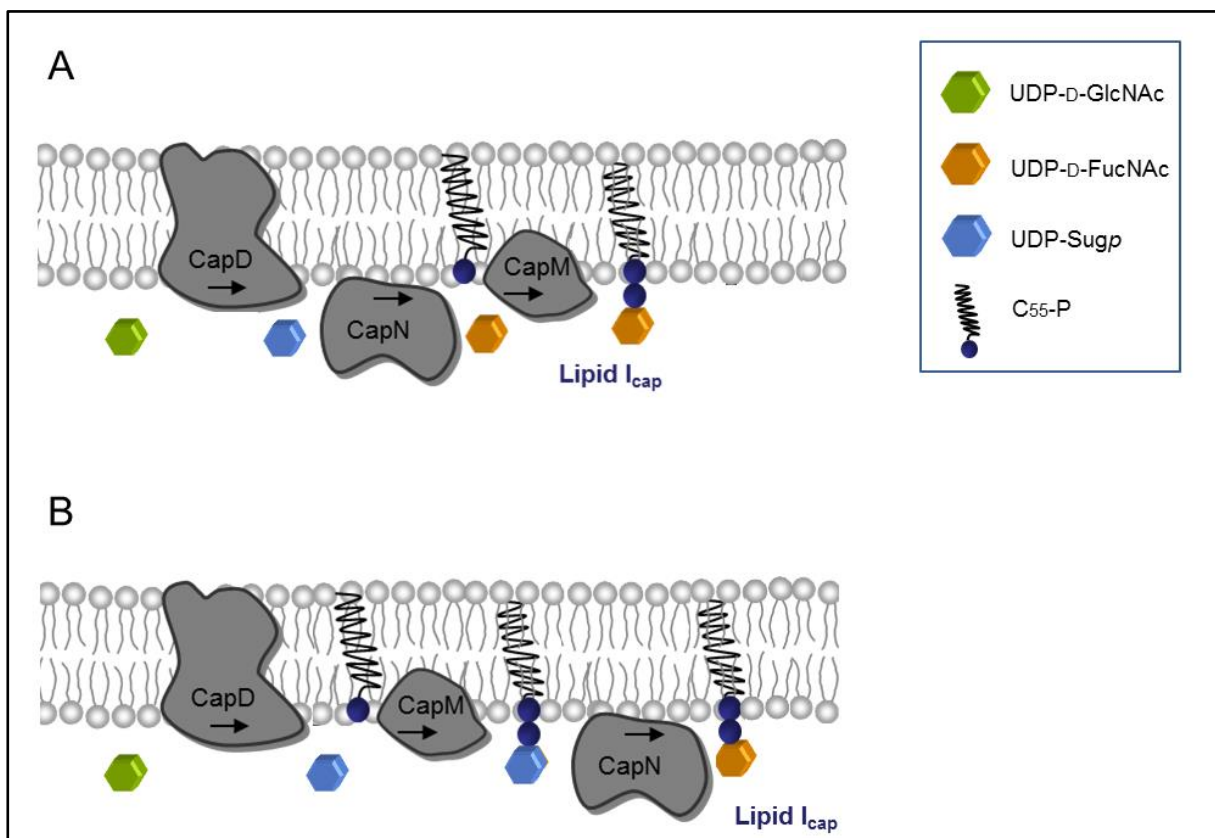


FIGURE 24. Alternative pathways for the biosynthesis of UDP-D-FucNAc in *S. aureus*. The integral membrane protein CapD is a 4,6-dehydratase, that generates the intermediate UDP-Suga (UDP-2-acetamido-2,6-dideoxy-D-xylo-4-hexulose) by eliminating a water molecule from UDP-D-GlcNAc. In the first scenario (A), this intermediate is further converted to UDP-D-FucNAc by the action of the 4-keto reductase CapN. Subsequently, the glycosyltransferase CapM transfers the *N*-acetyl-D-fucosamine moiety of UDP-D-FucNAc to the membrane-anchored lipid carrier bactoprenol (C<sub>55</sub>P), yielding lipid I<sub>cap</sub>. In the second scenario (B), the reduction step occurs after transfer of Suga to C<sub>55</sub>P.

## 4.5 The approved pharmaceutical suramin is a potent inhibitor of CapE

Establishment of a robust, validated test system allowed for inhibitor screening of UDP-GlcNAc 4,6-dehydratases. The developed CE method (see “Materials and methods”) enabled quantitation of the enzymatic reaction products with a low detection limit such that substrate concentrations around the  $K_m$  value of the enzymes (see FIGURE 16) could be employed for inhibitor testing.<sup>378</sup> Suramin, a pharmaceutical used for the treatment of onchocerciasis and trypanosome infections (for reviews, see references<sup>339,379,380</sup>), showed selective inhibition of CapE with an  $IC_{50}$  value of 1.82  $\mu$ M tested at a substrate concentration of 1 mM (FIGURE 18A). It acted as a non-competitive inhibitor with respect to the substrate UDP-D-GlcNAc (FIGURE 19A).

Suramin (Germanin<sup>™</sup>; manufactured by Bayer) was introduced in the early 1920s, and to this day remains the drug of choice for treatment of the early phase of *Trypanosoma brucei rhodesiense* infections.<sup>381,382</sup> Notably, no significant clinical resistance to suramin has emerged despite almost 100 years of use as trypanocidal agent.<sup>380</sup> Suramin is a symmetrical naphthylamine derivative of urea carrying six sulfonate groups, which are negatively charged at physiological pH (FIGURE 18). By virtue of its polyanionic nature, suramin can form tight complexes with various protein targets, resulting in a broad range of biological activities (e.g. immunomodulatory, antiproliferative, antiangiogenic, antineoplastic, contraceptive, anthelmintic, antiprotozoal, antibacterial, and antiviral activities<sup>339,379,380,383–385</sup>). It is thus considered a “dirty drug”.<sup>386</sup> Several enzymes from microbial pathogens are inhibited by suramin *in vitro*, for instance the *Mycobacterium tuberculosis* DNA repair protein RecA,<sup>387</sup> HIV-1 reverse transcriptase,<sup>388</sup> and various glycolytic enzymes from *T. brucei*.<sup>389</sup>

Although suramin binds to a variety of proteins, it can also discriminate between closely related members of the same protein family, suggesting the existence of rather specific docking sites.<sup>390</sup> For instance, suramin is a potent inhibitor of human elastase, neutrophil protease 3, and cathepsin G, but has only little or no inhibitory activity towards pancreatic trypsin and chymotrypsin.<sup>391</sup> Differential inhibition of proteins from different UDP-GlcNAc 4,6-dehydratase subfamilies was observed in this thesis. Due to its relatively high potency and selectivity for CapE compared to CapD, suramin may be a useful tool for studying capsule biosynthetic enzymes.

Suramin is a non-aggregator (i.e. it does not inhibit its targets in an unspecific manner by formation of self-aggregates), and its inhibitory activity was, at least in one case, shown to be insensitive to the presence of Triton X-100 and bovine serum albumin.<sup>392,393</sup> Thus, it seems unlikely that differences in inhibitor binding are caused by the detergents used for purification and reconstitution of full-length CapD. However, further UDP-GlcNAc dehydratases should be assayed for inhibition to evaluate the discriminatory capacity of suramin. The evaluation of suramin derivatives may yield even more potent inhibitors.

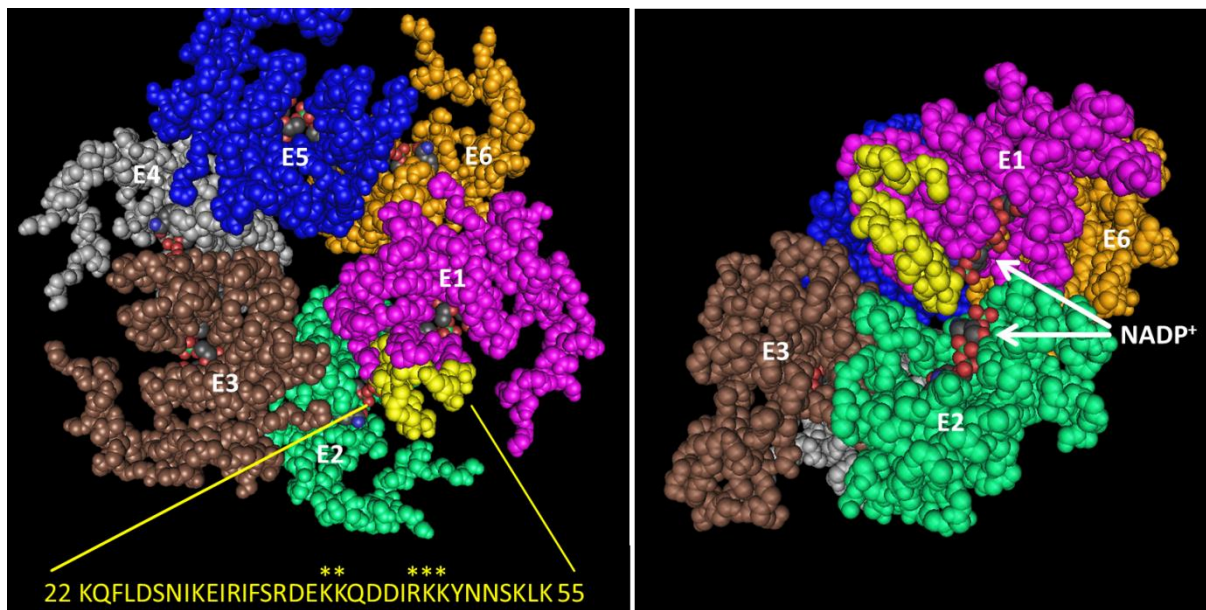


FIGURE 25: Space-filling representation of the three-dimensional structure of the hexameric CapE holoenzyme in top view (left) and side view (right). The structural model was created using NCBI's interactive 3D structure viewer Cn3D, and is based on the crystal structure published by Miyafusa et al.<sup>228</sup> The holoenzyme is shown with bound cofactor NADP<sup>+</sup>; the cofactor molecules are depicted as Corey-Pauling-Koltun (CPK) models. The six CapE protomers are colored differently, and labeled E1 through E6. Within the cofactor-binding domain of E1 (pink protomer), the region comprising amino acids 22–55 is highlighted in yellow. This sequence stretch is rich in basic residues (K, R) and contains a heparin-binding motif, which might represent a docking site for the CapE inhibitor suramin. The respective amino acid sequence is given at the bottom of the left panel; the heparin-binding motif is denoted by stars.

Suramin is structurally similar to heparin and acts as a competitor of glycosaminoglycan binding.<sup>394</sup> The ability of suramin to bind to heparin-binding growth factors was shown to depend on the correct spatial alignment between sulfonate groups on the terminal naphthalene rings and basic amino acid residues (especially K and R) on the surface of the target protein, in a manner similar to that

found for heparin–antithrombin III interactions.<sup>395–397</sup> The cofactor-binding domain of CapE harbors a region rich in basic amino acids (residues 22–55; FIGURE 25), which contains a motif consistent with a heparin-binding site<sup>398</sup> (K<sup>40</sup>KXXXXRKK). This motif might represent a docking site for the heparin analog suramin. According to the crystal structure published by Miyafusa et al.,<sup>228</sup> the respective residues are located on the outer surface of the CapE hexamer, in a cleft formed between two contiguous CapE monomers (FIGURE 25). Since K41 is part of the NADP<sup>+</sup>-binding pocket,<sup>228</sup> binding of suramin to this residue might result in displacement of the cofactor. In good agreement with this hypothesis, a different mode of cofactor binding was observed for CapD and CapE (FIGURE 16). As revealed by aligning the amino acid sequences of CapE and CapD, the latter protein lacks an equivalent motif (data not shown).

Suramin has been shown to penetrate into eukaryotic cells by an active transport mechanism and can even cross the blood-brain barrier.<sup>399,400</sup> It is thought to adsorb to cell membranes by the sulfonate groups interacting with positively charged membrane moieties (e.g. phosphatidylcholine), and by the internaphthalene bridge hydrophobically interacting with the acyl chains.<sup>399</sup> Suramin is likely internalized by a combination of pinocytosis and adsorptive endocytosis, with the endosomal/lysosomal compartment as endpoint.<sup>399</sup> However, its large molecular size (*Mr* 1429) and high negative net charge impair gastrointestinal absorption, and thus peroral bioavailability.<sup>401</sup> Two close analogs of suramin, trypan blue (*Mr* 961) and trypan red (*Mr* 1003), are routinely used as so-called vital stains, because they cannot cross the intact cytoplasmic membrane of viable cells.<sup>402,403</sup> Similarly, suramin is likely not able to passively diffuse across the highly polarized (interior negative), highly anionic (i.e. high content of anionic phospholipids in the outer leaflet) cytoplasmic membrane of *S. aureus*.<sup>404,405</sup> Nevertheless, whole-cell effects have been reported for *Firmicutes* species: inhibition of  $\beta$ -lactam-mediated cell lysis by suramin was shown in *S. aureus*,<sup>406</sup> and impairment of cell division, combined with altered “rough” cell surface properties, was observed in *E. faecalis* upon exposure to the compound.<sup>407</sup> The mode of action that underlies these effects remains elusive. It has been suggested, that suramin interferes with the function of bacterial autolysins, though there are no indications of a direct inhibitory effect on such cell wall-degrading enzymes.<sup>408</sup> The observed phenotypical changes may be related to binding of suramin to transmembrane protein kinases, such as PknB,<sup>409</sup> or Walk.<sup>410</sup>

Although suramin has been extensively used to treat parasitic infectious diseases, it is not an ideal drug. Therapy requires intravenous administration in a clinical setting ([www.who.int](http://www.who.int)). Serum level monitoring is mandatory, since the drug has a narrow therapeutic window and can cause serious toxicities, including coagulopathy, adrenal necrosis, and motor neuropathy (Guillain-Barré-like syndrome).<sup>411,412</sup> Clinical pharmacokinetic studies have demonstrated that suramin is highly (> 99.7%) bound to plasma proteins and stays in the body for a considerable period of time (terminal half-life of ~50 days).<sup>413</sup> Due to its high toxicity and unfavourable pharmacokinetics, suramin has only limited potential for use in antibacterial chemotherapy. Moreover, the immunosuppressant activities of suramin may impair bacterial clearance by the human immune system.<sup>385,414,415</sup>

#### 4.6 A novel mechanism for regulation of capsule biosynthesis

In CapD inhibition studies, we identified the  $\beta$ -lactam antibiotic ampicillin as an inhibitor of CapD (IC<sub>50</sub> 40.1  $\mu$ M) showing a non-competitive mechanism of enzyme inhibition (FIGURE 18B + FIGURE 19B). Dicloxacillin was similarly potent as ampicillin in the screening assay (see SUPPLEMENTARY TABLE 1), while all other investigated  $\beta$ -lactams did not show any significant inhibition at a test concentration of 10  $\mu$ M. Ampicillin is a perorally bioavailable drug, able to penetrate eukaryotic lipid bilayers. Ampicillin also appears to cross the staphylococcal cell membrane (see below). The inhibitory concentration of ampicillin at CapD is not much higher than typically observed therapeutic drug levels.<sup>416</sup>

The fact that the C<sub>55</sub>P carrier is only present in limited amounts within the bacterial cell,<sup>119</sup> and is required for the biosynthesis of diverse cell envelope components, such as CP, WTA and PG,<sup>120,122,124</sup> makes a well-orchestrated spatial and temporal regulation of these processes crucial for viability. In *S. aureus*, CP biosynthesis occurs predominantly in the late-exponential and post-exponential growth phase.<sup>211</sup> Post-translational control of *S. aureus* capsule expression is exerted by the bacterial tyrosine (BY-)kinase complex CapAB.<sup>161</sup> It has been reported that the synthesis of UDP-D-ManNAcA is regulated by CapAB-mediated tyrosine phosphorylation of the dehydrogenase CapO.<sup>216</sup> Moreover, the present thesis identifies the dehydratase CapE, involved in the synthesis of UDP-L-FucNAc, as regulatory target of the CapAB complex. In contrast, no CapAB protein target could be identified among the



enzymes of the UDP-D-FucNAc biosynthesis pathway; neither CapD, nor CapN was phosphorylated by CapAB *in vitro* (see section 3.8).

As revealed by *in vitro* inhibition studies, CapD catalytic activity is inhibited in the presence of the PG pentapeptide side chain (FIGURE 18C) and of the PG precursor lipid II (SUPPLEMENTARY TABLE 1). The findings that ampicillin (FIGURE 18B) and the structurally analogous D-Ala-D-Ala dipeptide (SUPPLEMENTARY TABLE 1) inhibit CapD activity suggest that the C-terminal end of the pentapeptide side chain may represent the minimal structure required for inhibition.

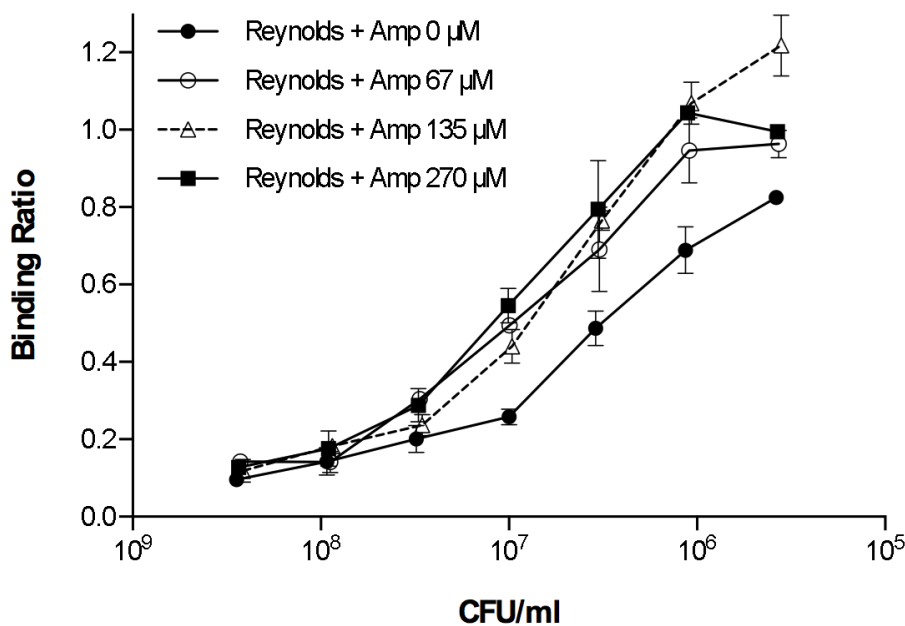


FIGURE 26. ELISA inhibition experiment showing that capsular polysaccharide production by the serotype 5 strain *Staphylococcus aureus* Reynolds is inhibited when the culture is incubated in the presence of ampicillin (Amp). The binding ratio was calculated by comparing the OD<sub>405 nm</sub> of the serum samples incubated with different bacterial concentrations to that of the serum sample containing no bacteria. For quantitation, the 50% inhibitory bacterial concentration was compared between cultures grown in the presence or absence of ampicillin. CFU, colony-forming units; OD<sub>405 nm</sub>, optical density at a wavelength of 405 nm.

Corroborating the *in vitro* findings, staphylococcal cells grown in the presence of subinhibitory concentrations of ampicillin displayed drastically reduced CP production, suggesting an integrated mechanism of regulation for cell envelope biosynthesis pathways in *S. aureus* that controls the coordinated assembly of vital

cell wall polymers in time and space (results from the laboratory of Prof Jean C. Lee, Brigham & Women's Hospital/Harvard Medical School, Boston, USA; FIGURE 26).

Effects of  $\beta$ -lactam antibiotics on capsule expression have been described in other studies. For instance, De Hennezel et al. reported that non-capsular antibodies were protective in a mouse-model of pneumococcal pneumonia when combined with subcurative doses of ampicillin.<sup>418</sup> Since non-capsular antibodies are not promoting phagocytic clearance in the presence of a capsule,<sup>44</sup> this points to ampicillin-induced reduction of CP production. It should be noted that the alleged capacity of ampicillin to penetrate into staphylococcal cells has not been experimentally proven, and that other mechanisms than inhibition of CapD are conceivable for the observed phenomenon, for instance direct or indirect modulatory effects of ampicillin on the activity of bacterial “receptor kinases”, such as CapAB and PknB.<sup>161,409</sup>

Taken together, the results of the present study strongly indicate that synchronization of CP and PG biosynthesis in *S. aureus* is, besides its regulation by transcription factors<sup>143,213,214</sup> and tyrosine kinases,<sup>161,216</sup> facilitated through kinetic modulation of CapD activity by peptidoglycan precursors. Moreover, the inhibitory effect of ampicillin on capsule production observed in whole cells might be of clinical relevance, and raises the possibility of additional target sites for  $\beta$ -lactam antibiotics.

#### **4.7 Enzymatic checkpoints within capsule biosynthesis are controlled by the tyrosine kinase complex CapAB**

Protein phosphorylation is one of the most common modes of regulation of protein function in both prokaryotes and eukaryotes.<sup>341</sup> Several biosynthetic enzymes involved in polysaccharide production have been identified as endogenous substrates of bacterial tyrosine kinases (see section 1.6), for example the UDP-glucose 6-dehydrogenases TuaA and YwqF of *Bacillus subtilis*,<sup>149</sup> the homologous *E. coli* dehydrogenase Ugd,<sup>175</sup> and the UDP-ManNAc dehydrogenase CapO involved in *S. aureus* capsule biosynthesis.<sup>216</sup>

The fact that CapO (FIGURE 4), is a regulatory target of the CapAB kinase complex<sup>216,218</sup> led us to assume that earlier checkpoints within the biosynthetic pathway are also controlled by CapAB. As deduced from kinase assays, the dehydratase CapE is an endogenous protein substrate of CapAB (FIGURE 20), and phosphorylation of CapE results in enhanced conversion of the precursor UDP-D-GlcNAc *in vitro* (FIGURE 21B).

The identification of phosphorylation sites revealed phosphotransfer to four tyrosine residues, with Tyr76, Tyr290, and Tyr293 being widely conserved among homologous proteins from different pathogens (FIGURE 21A). Interestingly, these residues are located in strategic functional regions of CapE, i.e. next to the cofactor binding site, in close proximity to the active site, and within a recently described mobile loop (FIGURE 27A). Crystal structure analysis identified this mobile “latch” to connect two CapE protomers within the hexameric complex, and showed that the latch of one dimerization partner is associated with the substrate-binding domain of the contiguous CapE monomer (FIGURE 27B), suggesting that this mobile loop is involved in regulating the access of the substrate UDP-GlcNAc to the active site.<sup>228</sup> Since contacts to Tyr290 and Tyr293 appear to be involved in the interaction, phosphorylation might induce conformational changes that facilitate access of UDP-GlcNAc, increasing CapE enzymatic activity.

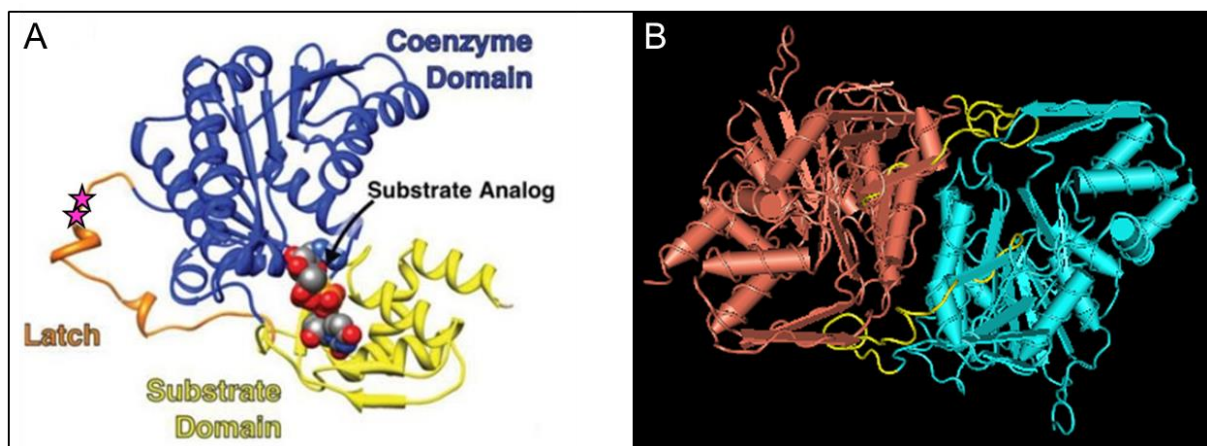


FIGURE 27. Structure of CapE monomer bound to substrate analog UDP-6N3-GlcNAc (A; adapted from Miyafusa et al.<sup>228</sup>). The coenzyme-binding domain, substrate-binding domain, and a mobile loop (“latch”) are colored in blue, yellow, and orange, respectively. The potential regulatory phosphorylation sites Tyr290 and Tyr293 are indicated by stars. (B) Worm representation of two contiguous subunits within the hexameric CapE complex, created using NCBI's 3D structure viewer Cn3D. The two protomers are connected via their mobile latches (highlighted in yellow in each subunit) making contacts with the substrate-binding domain of the dimerization partner.

Two couples of adjacent genes encoding for a transmembrane adaptor and a cytoplasmic BY-kinase were identified in the genome of *S. aureus* serotype 5: the *capA1/capB1* couple located at the 5'-end of the *cap* operon, and the highly similar *capA2/capB2* triplet, which is located elsewhere on the bacterial chromosome, and probably arose from gene duplication.<sup>161,206</sup> The *in vitro* assays presented here were

performed with the tyrosine kinase CapB2, since attempts to reconstitute CapB1 autokinase activity have failed so far.<sup>161,233</sup> However, a recent biochemical study indicates that BY-kinases have evolved rather relaxed substrate specificities; despite originating from phylogenetically distant bacteria, six out of seven tested BY-kinases (*S. aureus* CapB2, *S. pneumoniae* CpsD, *B. subtilis* PtkA, as well as BY-kinases of *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Bifidobacterium animalis*) were able to catalyze *in vitro* phosphorylation of protein substrates from *B. subtilis* (YwqF, YvyG, and YjoA).<sup>419</sup> The authors suggested that BY-kinases maintain the capability to phosphorylate the same set of substrates across very distant bacterial phyla, since they recognize the overall structure of a target protein, and not a particular short sequence or phosphorylation site.<sup>419</sup> With regard to these findings, it seems likely that CapB1 and CapB2 recognize largely overlapping spectra of substrates. Moreover, since potential phosphorylation sites are determined by the sequence properties of the target protein,<sup>420,421</sup> CapB1 and CapB2 are expected to phosphorylate the same tyrosine residues in a given protein. Nevertheless, efforts to reconstitute CapB1 kinase activity are ongoing in our group.

A recent study identified an exopolysaccharide of *B. subtilis* as quorum-sensing molecule modulating BY-kinase signaling by binding to the extracellular sensor domain of EpsA.<sup>422</sup> In addition, it was demonstrated that the *S. aureus* exopolysaccharide PIA (polysaccharide intercellular adhesin) binds to CapA *in vitro*.<sup>422</sup> The signal detected by CapA *in vivo* has not yet been identified, but it is tempting to speculate that the activator protein senses the presence or absence of membrane-bound cell envelope precursors, such as lipid III<sub>cap</sub>, resulting in regulation of enzymatic checkpoints that ensures the sufficient supply of precursors for CP polymerization. Moreover, CapAB might be required for correct localization and assembly of the capsule biosynthetic machinery.

## 4.8 Future directions

In the present study, *in vitro* enzymatic synthesis of the nucleotide-activated dideoxy sugar precursor UDP-*N*-acetyl-D-fucosamine was achieved for the first time. As revealed by CE and MS analysis, the purified recombinant enzymes CapD and CapN from *Staphylococcus aureus* serotype 5 converted the substrate UDP-D-GlcNAc into a sugar nucleotide species having a molecular mass of 591.5, consistent with the formation of UDP-D-FucNAc. Taking into account the genetic evidence and the

knowledge gained in studies on homologous enzymes,<sup>237,281,283,284,287,288,329</sup> it seems highly likely that the enzymatic product generated by sequential action of the 4,6-dehydratase CapD and the 4-keto reductase CapN is indeed UDP-D-FucNAc—the first soluble precursor required for CP repeat unit assembly in *S. aureus*. However, NMR spectroscopy should be performed for structure elucidation, to unequivocally confirm the identity of the reaction product.

Moreover, substrate specificity and preferences of the reductase CapN should be determined. *In vitro*, CapN is obviously capable of catalyzing reduction of the substrate UDP-Sugp, though with low catalytic efficiency. However, it is conceivable that *in vivo* the reduction step could occur after transfer of Sugp to C<sub>55</sub>P (see FIGURE 24). Preliminary results from our laboratory indicate that CapM may indeed accept UDP-Sugp as donor substrate, which provides the opportunity to test C<sub>55</sub>PP-Sugp as alternative substrate for CapN, in order to elucidate the exact biosynthetic route leading to the formation of D-FucNAc.

If both D-FucNAc and Sugp may occupy the first position in the repeating unit, inactivation of Sugp 4-keto reductases such as CapN might represent a simple and economic mechanism for serotype switching or phase variation. Sugp is an extremely rare sugar, and has so far only been detected in a few bacterial polysaccharides.<sup>287,377,423–426</sup> Sugp is present in the CP produced by *S. pneumoniae* type 5 strains,<sup>426</sup> and the expression of a type 5 capsule has been linked with high virulence and invasiveness.<sup>427,428</sup> Sugp is also found in the LPS outer core of the enteric pathogen *Yersinia enterocolitica* serotype O:3.<sup>287</sup> The outer core hexasaccharide is required for full virulence, and plays a role in outer membrane integrity relevant in the resistance to polycations and hydrophobic agents.<sup>429</sup> *Y. enterocolitica* O:3 mutants lacking the LPS outer core, but expressing the O antigen, were more susceptible than the wild-type bacteria to polymyxin B, melittin, poly-L-lysine, and poly-L-ornithine.<sup>429</sup> Interestingly, substitution of the outer core Sugp residue by D-QuiNAc or D-FucNAc (by providing *wbpV* or *wbpK in trans*) was reported to enhance resistance against the phage tail-like bacteriocin enterocolitacin.<sup>287</sup> In this context, it would be interesting to analyze capsule production in a *capN* knockout mutant of *S. aureus* and, in case formation of CP occurs, to analyze the virulence properties of the mutant strain, as well as its susceptibility to antimicrobial peptides.

The finding that selected  $\beta$ -lactam antibiotics as well as structurally similar peptidoglycan precursors inhibit CapD enzymatic activity *in vitro* suggests an integrated mechanism of regulation for cell envelope biosynthesis pathways. Construction of non-binding mutants of CapD by site-directed mutagenesis and expression of these mutant proteins in a *capD* negative background may allow for conclusions about the *in vivo* relevance of the observed inhibitory effect of PG precursors on CapD catalytic activity. First of all, this requires mapping of the inhibitor-binding sites on CapD, ideally by crystallographic studies.

The present study provides evidence that the dehydratase CapE, which initiates the synthesis of the second soluble capsule precursor UDP-L-FucNAc, is a regulatory target of the CapAB tyrosine kinase complex. As assessed by CE quantification of the CapE enzymatic product, CapE catalytic activity *in vitro* was markedly enhanced in the presence of the reconstituted CapAB complex. In depth biochemical characterization revealed phosphotransfer onto four tyrosine residues, suggesting that CapE is positively modulated by CapAB-mediated phosphorylation. Site-directed mutagenesis studies are currently performed in our laboratory, to confirm the role of the four tyrosine residues as regulatory phosphorylation sites of CapE.

Two couples of adjacent genes encoding for a transmembrane adaptor and a cytoplasmic BY-kinase were identified in the genome of *S. aureus* serotype 5: the *capA1/capB1* gene couple located at the 5'-end of the *cap* operon, and the highly similar *capA2/capB2* couple, which is located elsewhere on the bacterial chromosome, and probably arose from gene duplication.<sup>161,206</sup> Since attempts to reconstitute CapB1 autokinase activity have failed so far, biochemical assays were performed with the homologous tyrosine kinase CapB2. There is, as yet, no *in vivo* experimental evidence supporting a regulatory role for the CapA2B2 complex in *S. aureus* capsule biosynthesis. No differences in the amount of cell wall-attached CP were observed with *capA2* and *capB2* knock-out mutant strains (Prof Jean C. Lee, unpublished results). However, at least *in vitro*, CapB1 and CapB2 are expected to recognize largely overlapping spectra of substrates (compare Shi et al.<sup>419</sup>). Moreover, since potential phosphorylation sites are determined by the sequence properties of the target protein,<sup>420,421</sup> CapB1 and CapB2 are expected to phosphorylate the same tyrosine residues in a given protein. Nevertheless, efforts to reconstitute CapB1 kinase activity are ongoing in our group.

## 5 References

1. Howard, C. J. & Glynn, A. A. The virulence for mice of strains of *Escherichia coli* related to the effects of K antigens on their resistance to phagocytosis and killing by complement. *Immunology* **20**, 767 (1971).
2. Mackinnon, F. G. *et al.* Demonstration of lipooligosaccharide immunotype and capsule as virulence factors for *Neisseria meningitidis* using an infant mouse intranasal infection model. *Microb. Pathog.* **15**, 359–366 (1993).
3. Moxon, E. R. & Vaughn, K. A. The type b capsular polysaccharide as a virulence determinant of *Haemophilus influenzae*: studies using clinical isolates and laboratory transformants. *J. Infect. Dis.* **143**, 517–524 (1981).
4. Thakker, M., Park, J.-S., Carey, V. & Lee, J. C. *Staphylococcus aureus* serotype 5 capsular polysaccharide is antiphagocytic and enhances bacterial virulence in a murine bacteremia model. *Infect. Immun.* **66**, 5183–5189 (1998).
5. Wood, W. B. & Smith, M. R. The inhibition of surface phagocytosis by the capsular 'slime layer' of *Pneumococcus* type III. *J. Exp. Med.* **90**, 85–96 (1949).
6. Weintraub, A. Immunology of bacterial polysaccharide antigens. *Carbohydr. Res.* **338**, 2539–2547 (2003).
7. Beveridge, T. J. & Graham, L. L. Surface layers of bacteria. *Microbiol. Rev.* **55**, 684 (1991).
8. Candela, T. & Fouet, A. *Bacillus anthracis* CapD, belonging to the  $\gamma$ -glutamyltranspeptidase family, is required for the covalent anchoring of capsule to peptidoglycan. *Mol. Microbiol.* **57**, 717–726 (2005).
9. Deng, L., Kasper, D. L., Krick, T. P. & Wessels, M. R. Characterization of the linkage between the type III capsular polysaccharide and the bacterial cell wall of group B *Streptococcus*. *J. Biol. Chem.* **275**, 7497–7504 (2000).
10. Yother, J. Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation. *Annu. Rev. Microbiol.* **65**, 563–581 (2011).
11. Gotschlich, E. C., Fraser, B. A., Nishimura, O., Robbins, J. B. & Liu, T. Y. Lipid on capsular polysaccharides of gram-negative bacteria. *J. Biol. Chem.* **256**, 8915–8921 (1981).
12. Schmidt, M. A. & Jann, K. Phospholipid substitution of capsular (K) polysaccharide antigens from *Escherichia coli* causing extraintestinal infections. *FEMS Microbiol. Lett.* **14**, 69–74 (1982).
13. Tzeng, Y.-L. *et al.* Translocation and surface expression of lipidated serogroup B capsular polysaccharide in *Neisseria meningitidis*. *Infect. Immun.* **73**, 1491–1505 (2005).
14. Willis, L. M. *et al.* Conserved glycolipid termini in capsular polysaccharides synthesized by ATP-binding cassette transporter-dependent pathways in Gram-negative pathogens. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7868–7873 (2013).
15. Jiménez, N. *et al.* Effects of lipopolysaccharide biosynthesis mutations on K1 polysaccharide association with the *Escherichia coli* cell surface. *J. Bacteriol.* **194**, 3356–3367 (2012).
16. Valle, J. *et al.* Broad-spectrum biofilm inhibition by a secreted bacterial polysaccharide. *Proc. Natl. Acad. Sci.* **103**, 12558–12563 (2006).

17. Quintero, E. J. & Weiner, R. M. Evidence for the adhesive function of the exopolysaccharide of *Hyphomonas* Strain MHS-3 in its attachment to surfaces. *Appl. Environ. Microbiol.* **61**, 1897–1903 (1995).
18. Orskov, I., Orskov, F., Jann, B. & Jann, K. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. *Bacteriol. Rev.* **41**, 667 (1977).
19. Robbins, J. D. & Robbins, J. B. Reexamination of the protective role of the capsular polysaccharide (Vi antigen) of *Salmonella typhi*. *J. Infect. Dis.* **150**, 436–449 (1984).
20. Díaz-Quiñonez, A. *et al.* Outbreak of *Vibrio cholerae* serogroup O1, serotype Ogawa, biotype El Tor strain—La Huasteca Region, Mexico, 2013. *MMWR Morb. Mortal. Wkly. Rep.* **63**, 552–553 (2014).
21. Scott, C. L., Iyasu, S., Rowley, D. & Atrash, H. K. Postneonatal mortality surveillance—United States, 1980–1994. *MMWR CDC Surveill. Summ. Morb. Mortal. Wkly. Rep. CDC Surveill. Summ. Cent. Dis. Control* **47**, 15–30 (1998).
22. Taylor, C. M. & Roberts, I. S. in *Contributions to Microbiology* (eds. Russell, W. & Herwald, H.) **12**, 55–66 (KARGER, 2004).
23. Roberson, E. B. & Firestone, M. K. Relationship between desiccation and exopolysaccharide production in a soil *Pseudomonas* sp. *Appl. Environ. Microbiol.* **58**, 1284–1291 (1992).
24. Ophir, T. & Gutnick, D. L. A role for exopolysaccharides in the protection of microorganisms from desiccation. *Appl. Environ. Microbiol.* **60**, 740–745 (1994).
25. Roberts, I. S. The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu. Rev. Microbiol.* **50**, 285–315 (1996).
26. Berry, A., DeVault, J. D. & Chakrabarty, A. M. High osmolarity is a signal for enhanced algD transcription in mucoid and nonmucoid *Pseudomonas aeruginosa* strains. *J. Bacteriol.* **171**, 2312–2317 (1989).
27. Sledjeski, D. D. & Gottesman, S. Osmotic shock induction of capsule synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **178**, 1204–1206 (1996).
28. Deighton, M. A. & Balkau, B. Adherence measured by microtiter assay as a virulence marker for *Staphylococcus epidermidis* infections. *J. Clin. Microbiol.* **28**, 2442–2447 (1990).
29. McKenney, D. *et al.* The *ica* locus of *Staphylococcus epidermidis* encodes production of the capsular polysaccharide/adhesin. *Infect. Immun.* **66**, 4711–4720 (1998).
30. Moscoso, M., García, E. & López, R. Biofilm formation by *Streptococcus pneumoniae*: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. *J. Bacteriol.* **188**, 7785–7795 (2006).
31. Hammerschmidt, S. *et al.* Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infect. Immun.* **73**, 4653–4667 (2005).
32. Magee, A. D. & Yother, J. Requirement for capsule in colonization by *Streptococcus pneumoniae*. *Infect. Immun.* **69**, 3755–3761 (2001).
33. Schembri, M. A., Dalsgaard, D. & Klemm, P. Capsule shields the function of short bacterial adhesins. *J. Bacteriol.* **186**, 1249–1257 (2004).
34. Hammerschmidt, S. *et al.* Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. *EMBO J.* **15**, 192–198 (1996).



35. Virji, M. *et al.* Opc- and pilus-dependent interactions of meningococci with human endothelial cells: molecular mechanisms and modulation by surface polysaccharides. *Mol. Microbiol.* **18**, 741–754 (1995).
36. Horwitz, M. A. & Silverstein, S. C. Influence of the *Escherichia coli* capsule on complement fixation and on phagocytosis and killing by human phagocytes. *J. Clin. Invest.* **65**, 82–94 (1980).
37. Nilsson, I. M., Lee, J. C., Bremell, T., Rydén, C. & Tarkowski, A. The role of staphylococcal polysaccharide microcapsule expression in septicemia and septic arthritis. *Infect. Immun.* **65**, 4216–4221 (1997).
38. Portolés, M., Kiser, K. B., Bhasin, N., Chan, K. H. N. & Lee, J. C. *Staphylococcus aureus* Cap5O has UDP-ManNAc dehydrogenase activity and is essential for capsule expression. *Infect. Immun.* **69**, 917–923 (2001).
39. Tzianabos, A. O., Wang, J. Y. & Lee, J. C. Structural rationale for the modulation of abscess formation by *Staphylococcus aureus* capsular polysaccharides. *Proc. Natl. Acad. Sci.* **98**, 9365–9370 (2001).
40. Cartee, R. T. & Yother, J. *Molecular Paradigms of Infectious Disease: A Bacterial Perspective.* (Springer Science & Business Media, 2006).
41. Rautemaa, R. & Meri, S. Complement-resistance mechanisms of bacteria. *Microbes Infect.* **1**, 785–794 (1999).
42. Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J. D. Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* **11**, 785–797 (2010).
43. Joiner, K. A. Complement evasion by bacteria and parasites. *Annu. Rev. Microbiol.* **42**, 201–230 (1988).
44. Karakawa, W. W., Sutton, A., Schneerson, R., Karpas, A. & Vann, W. F. Capsular antibodies induce type-specific phagocytosis of capsulated *Staphylococcus aureus* by human polymorphonuclear leukocytes. *Infect. Immun.* **56**, 1090–1095 (1988).
45. Krarup, A., Sørensen, U. B. S., Matsushita, M., Jensenius, J. C. & Thiel, S. Effect of capsulation of opportunistic pathogenic bacteria on binding of the pattern recognition molecules mannan-binding lectin, L-ficolin, and H-ficolin. *Infect. Immun.* **73**, 1052–1060 (2005).
46. Thurlow, L. R., Thomas, V. C., Fleming, S. D. & Hancock, L. E. *Enterococcus faecalis* capsular polysaccharide and mechanisms of host innate immune evasion. *Infect. Immun.* **77**, 5551–5557 (2013).
47. Taylor, P. W. & Robinson, M. K. Determinants that increase the serum resistance of *Escherichia coli*. *Infect. Immun.* **29**, 278–280 (1980).
48. Wilkinson, B. J., Sisson, S. P., Kim, Y. & Peterson, P. K. Localization of the third component of complement on the cell wall of encapsulated *Staphylococcus aureus* M: implications for the mechanism of resistance to phagocytosis. *Infect. Immun.* **26**, 1159–1163 (1979).
49. Winkelstein, J. A., Abramovitz, A. S. & Tomasz, A. Activation of C3 via the alternative complement pathway results in fixation of C3b to the pneumococcal cell wall. *J. Immunol.* **124**, 2502–2506 (1980).
50. Brown, E. J., Joiner, K. A., Cole, R. M. & Berger, M. Localization of complement component 3 on *Streptococcus pneumoniae*: anti-capsular antibody causes complement deposition on the pneumococcal capsule. *Infect. Immun.* **39**, 403–409 (1983).

51. Ward, C. K. & Inzana, T. J. Identification and characterization of a DNA region involved in the export of capsular polysaccharide by *Actinobacillus pleuropneumoniae* serotype 5a. *Infect. Immun.* **65**, 2491–2496 (1997).
52. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat. Immunol.* **11**, 373–384 (2010).
53. Kocabas, C. *et al.* *Neisseria meningitidis* type C capsular polysaccharide inhibits lipooligosaccharide-induced cell activation by binding to CD14. *Cell. Microbiol.* **9**, 1297–1310 (2007).
54. McNally, D. J. *et al.* Commonality and biosynthesis of the O-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *J. Biol. Chem.* **282**, 28566–28576 (2007).
55. Rose, A., Kay, E., Wren, B. W. & Dallman, M. J. The *Campylobacter jejuni* NCTC11168 capsule prevents excessive cytokine production by dendritic cells. *Med. Microbiol. Immunol. (Berl.)* **201**, 137–144 (2012).
56. Maue, A. C. *et al.* The polysaccharide capsule of *Campylobacter jejuni* modulates the host immune response. *Infect. Immun.* **81**, 665–672 (2013).
57. Jennings, H. J. *et al.* Structure, conformation and immunology of sialic acid-containing polysaccharides of human pathogenic bacteria. *Pure Appl. Chem.* **56**, 893–905 (1984).
58. Finne, J., Bitter-Suermann, D., Goridis, C. & Finne, U. An IgG monoclonal antibody to group B meningococci cross-reacts with developmentally regulated polysialic acid units of glycoproteins in neural and extraneural tissues. *J. Immunol.* **138**, 4402–4407 (1987).
59. Vann, W. F., Schmidt, M. A., Jann, B. & Jann, K. The structure of the capsular polysaccharide (K5 antigen) of urinary-tract-infective *Escherichia coli* 010:K5:H4. *Eur. J. Biochem.* **116**, 359–364 (1981).
60. Beuvery, E. C., Miedema, F., van Delft, R. & Haverkamp, J. Preparation and immunochemical characterization of meningococcal group C polysaccharide-tetanus toxoid conjugates as a new generation of vaccines. *Infect. Immun.* **40**, 39–45 (1983).
61. Ward, J. *et al.* Enhanced immunogenicity in young infants of a new *Haemophilus influenzae* type B(Hib) capsular polysaccharide(prp)-diphtheria toxoid(D) conjugate vaccine. *Pediatr. Res.* **18**, 287A–287A (1984).
62. Schneerson, R., Barrera, O., Sutton, A. & Robbins, J. B. Preparation, characterization, and immunogenicity of *Haemophilus influenzae* type b polysaccharide-protein conjugates. *J. Exp. Med.* **152**, 361–376 (1980).
63. O'Brien, K. L. *et al.* Immunologic priming of young children by pneumococcal glycoprotein conjugate, but not polysaccharide, vaccines. *Pediatr. Infect. Dis. J.* **15**, 425–430 (1996).
64. De Roux, A. *et al.* Comparison of pneumococcal conjugate polysaccharide and free polysaccharide vaccines in elderly adults: conjugate vaccine elicits improved antibacterial immune responses and immunological memory. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **46**, 1015–1023 (2008).
65. Moe, G. R., Tan, S. & Granoff, D. M. Molecular mimetics of polysaccharide epitopes as vaccine candidates for prevention of *Neisseria meningitidis* serogroup B disease. *FEMS Immunol. Med. Microbiol.* **26**, 209–226 (1999).

66. Girard, M. P., Preziosi, M.-P., Aguado, M.-T. & Kieny, M. P. A review of vaccine research and development: meningococcal disease. *Vaccine* **24**, 4692–4700 (2006).
67. Häyrynen, J. *et al.* Antibodies to polysialic acid and its N-propyl derivative: binding properties and interaction with human embryonal brain glycopeptides. *J. Infect. Dis.* **171**, 1481–1490 (1995).
68. Caesar, N. M., Myers, K. A. & Fan, X. Neisseria meningitidis serogroup B vaccine development. *Microb. Pathog.* **57**, 33–40 (2013).
69. Gorringe, A. R. & Pajón, R. Bexsero: a multicomponent vaccine for prevention of meningococcal disease. *Hum. Vaccines Immunother.* **8**, 174–183 (2012).
70. Frenc, R. W. Rapid rises in antibody titers observed following single dose administration of a novel 4-antigen Staphylococcus aureus vaccine (SA4Ag) to healthy adults. in *IDWeek 2014* (Idsa, 2014). at <<https://idsa.confex.com/idsa/2014/webprogram/Paper44838.html>>
71. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700 (2002).
72. Bentley, S. D. *et al.* Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet.* **2**, e31 (2006).
73. Whitfield, C. Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annu Rev Biochem* **75**, 39–68 (2006).
74. Amor, P. A. & Whitfield, C. Molecular and functional analysis of genes required for expression of group IB K antigens in Escherichia coli: characterization of the his-region containing gene clusters for multiple cell-surface polysaccharides. *Mol. Microbiol.* **26**, 145–161 (1997).
75. Drummel-Smith, J. & Whitfield, C. Gene products required for surface expression of the capsular form of the group 1 K antigen in Escherichia coli (O9a:K30). *Mol. Microbiol.* **31**, 1321–1332 (1999).
76. Nakhamchik, A., Wilde, C. & Rowe-Magnus, D. A. Identification of a Wzy polymerase required for group IV capsular polysaccharide and lipopolysaccharide biosynthesis in Vibrio vulnificus. *Infect. Immun.* **75**, 5550–5558 (2007).
77. O’Riordan, K. & Lee, J. C. Staphylococcus aureus capsular polysaccharides. *Clin. Microbiol. Rev.* **17**, 218–234 (2004).
78. Cuthbertson, L., Kos, V. & Whitfield, C. ABC transporters involved in export of cell surface glycoconjugates. *Microbiol. Mol. Biol. Rev.* **74**, 341–362 (2010).
79. Rick, P. D. *et al.* Evidence that the wzxE gene of Escherichia coli K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid intermediate in the assembly of enterobacterial common antigen. *J. Biol. Chem.* **278**, 16534–16542 (2003).
80. Woodward, R. *et al.* In vitro bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat. Chem. Biol.* **6**, 418–423 (2010).
81. Hodson, N. *et al.* Identification that KfiA, a protein essential for the biosynthesis of the Escherichia coli K5 capsular polysaccharide, is an  $\alpha$ -UDP-GlcNAc glycosyltransferase. The formation of a membrane-associated K5 biosynthetic complex requires KfiA, KfiB, and KfiC. *J. Biol. Chem.* **275**, 27311–27315 (2000).
82. Ninomiya, T. *et al.* Molecular cloning and characterization of chondroitin polymerase from Escherichia coli strain K4. *J. Biol. Chem.* **277**, 21567–21575 (2002).

83. Steenbergen, S. M., Wrona, T. J. & Vimr, E. R. Functional analysis of the sialyltransferase complexes in *Escherichia coli* K1 and K92. *J. Bacteriol.* **174**, 1099–1108 (1992).
84. Reizer, J., Reizer, A. & Saier, M. A New subfamily of bacterial ABC-type transport-systems catalyzing export of drugs and carbohydrates. *Protein Sci.* **1**, 1326–1332 (1992).
85. Frosch, M., Edwards, U., Bousset, K., Krausse, B. & Weisgerber, C. Evidence for a common molecular-origin of the capsule gene loci in Gram-negative bacteria expressing group-II capsular polysaccharides. *Mol. Microbiol.* **5**, 1251–1263 (1991).
86. Karlyshev, A. V., Linton, D., Gregson, N. A., Lastovica, A. J. & Wren, B. W. Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Mol. Microbiol.* **35**, 529–541 (2000).
87. Kroll, J., Loynds, B., Brophy, L. & Moxon, E. The *bex* locus in encapsulated *Haemophilus influenzae* - a chromosomal region involved in capsule polysaccharide export. *Mol. Microbiol.* **4**, 1853–1862 (1990).
88. Pavelka, M. S., Wright, L. F. & Silver, R. P. Identification of two genes, *kpsM* and *kpsT*, in region 3 of the polysialic acid gene cluster of *Escherichia coli* K1. *J. Bacteriol.* **173**, 4603–4610 (1991).
89. Smith, A. N., Boulnois, G. J. & Roberts, I. S. Molecular analysis of the *Escherichia coli* K5 *kps* locus: identification and characterization of an inner-membrane capsular polysaccharide transport system. *Mol. Microbiol.* **4**, 1863–1869 (1990).
90. Michael, F. S. *et al.* The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur. J. Biochem.* **269**, 5119–5136 (2002).
91. Willis, L. M. & Whitfield, C. *KpsC* and *KpsS* are retaining 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) transferases involved in synthesis of bacterial capsules. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20753–20758 (2013).
92. Cartee, R. T., Forsee, W. T., Jensen, J. W. & Yother, J. Expression of the *Streptococcus pneumoniae* type 3 synthase in *Escherichia coli*. Assembly of type 3 polysaccharide on a lipid primer. *J. Biol. Chem.* **276**, 48831–48839 (2001).
93. Jing, W. & DeAngelis, P. L. Dissection of the two transferase activities of the *Pasteurella multocida* hyaluronan synthase: two active sites exist in one polypeptide. *Glycobiology* **10**, 883–889 (2000).
94. Tlapak-Simmons, V. L., Baggenstoss, B. A., Kumari, K., Heldermon, C. & Weigel, P. H. Kinetic characterization of the recombinant hyaluronan synthases from *Streptococcus pyogenes* and *Streptococcus equisimilis*. *J. Biol. Chem.* **274**, 4246–4253 (1999).
95. Whitney, J. C. & Howell, P. L. Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends Microbiol.* **21**, 63–72 (2013).
96. Hagopian, A. & Eylar, E. H. Glycoprotein biosynthesis: Studies on the receptor specificity of the polypeptidyl:N-acetylgalactosaminyl transferase from bovine submaxillary glands. *Arch. Biochem. Biophys.* **128**, 422–433 (1968).
97. DeAngelis, P. L. Evolution of glycosaminoglycans and their glycosyltransferases: implications for the extracellular matrices of animals and the capsules of pathogenic bacteria. *Anat. Rec.* **268**, 317–326 (2002).

98. Remminghorst, U. & Rehm, B. H. Bacterial alginates: from biosynthesis to applications. *Biotechnol. Lett.* **28**, 1701–1712 (2006).
99. Römling, U. Molecular biology of cellulose production in bacteria. *Res. Microbiol.* **153**, 205–212 (2002).
100. Weigel, P. H. & DeAngelis, P. L. Hyaluronan synthases: a decade-plus of novel glycosyltransferases. *J. Biol. Chem.* **282**, 36777–36781 (2007).
101. Forsee, W. T., Cartee, R. T. & Yother, J. Biosynthesis of type 3 capsular polysaccharide in *Streptococcus pneumoniae*. Enzymatic chain release by an abortive translocation process. *J. Biol. Chem.* **275**, 25972–25978 (2000).
102. Arrecubieta, C., García, E. & López, R. Sequence and transcriptional analysis of a DNA region involved in the production of capsular polysaccharide in *Streptococcus pneumoniae* type 3. *Gene* **167**, 1–7 (1995).
103. DeAngelis, P. L., Papaconstantinou, J. & Weigel, P. H. Molecular cloning, identification, and sequence of the hyaluronan synthase gene from group A *Streptococcus pyogenes*. *J. Biol. Chem.* **268**, 19181–19184 (1993).
104. Keenleyside, W. J. & Whitfield, C. A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar Borreze. *J. Biol. Chem.* **271**, 28581–28592 (1996).
105. Saxena, I. M., Brown, R. M., Fevre, M., Geremia, R. A. & Henrissat, B. Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. *J. Bacteriol.* **177**, 1419–1424 (1995).
106. Cartee, R. T., Forsee, W. T. & Yother, J. Initiation and synthesis of the *Streptococcus pneumoniae* type 3 capsule on a phosphatidylglycerol membrane anchor. *J. Bacteriol.* **187**, 4470–4479 (2005).
107. Hubbard, C., McNamara, J. T., Azumaya, C., Patel, M. S. & Zimmer, J. The hyaluronan synthase catalyzes the synthesis and membrane translocation of hyaluronan. *J. Mol. Biol.* **418**, 21–31 (2012).
108. Ventura, C. L., Cartee, R. T., Forsee, W. T. & Yother, J. Control of capsular polysaccharide chain length by UDP-sugar substrate concentrations in *Streptococcus pneumoniae*. *Mol. Microbiol.* **61**, 723–733 (2006).
109. Chung, J. Y., Zhang, Y. & Adler, B. The capsule biosynthetic locus of *Pasteurella multocida* A:1. *FEMS Microbiol. Lett.* **166**, 289–296 (1998).
110. Reeves, P. R. *et al.* Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol.* **4**, 495–503 (1996).
111. Willis, L. M. & Whitfield, C. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydr. Res.* **378**, 35–44 (2013).
112. Cunneen, M. M. & Reeves, P. R. Membrane topology of the *Salmonella enterica* serovar Typhimurium group B O-antigen translocase Wzx. *FEMS Microbiol. Lett.* **287**, 76–84 (2008).
113. Islam, S. T., Taylor, V. L., Qi, M. & Lam, J. S. Membrane topology mapping of the O-Antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. *Mbio* **1**, e00189–10 (2010).

114. Cuthbertson, L., Mainprize, I. L., Naismith, J. H. & Whitfield, C. Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol. Mol. Biol. Rev.* **73**, 155–177 (2009).
115. Morona, R., Bosch, L. V. D. & Daniels, C. Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. *Microbiology* **146**, 1–4 (2000).
116. Keiski, C.-L. *et al.* AlgK is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin. *Structure* **18**, 265–273 (2010).
117. Whitney, J. C. *et al.* Structural basis for alginate secretion across the bacterial outer membrane. *Proc. Natl. Acad. Sci.* **108**, 13083–13088 (2011).
118. Hanson, B. R. & Neely, M. N. Coordinate regulation of Gram-positive cell surface components. *Curr. Opin. Microbiol.* **15**, 204–210 (2012).
119. Barreteau, H. *et al.* Quantitative high-performance liquid chromatography analysis of the pool levels of undecaprenyl phosphate and its derivatives in bacterial membranes. *J. Chromatogr. B* **877**, 213–220 (2009).
120. Bouhss, A., Trunkfield, A. E., Bugg, T. D. H. & Mengin-Lecreux, D. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol. Rev.* **32**, 208–233 (2008).
121. D'Elia, M. A. *et al.* Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. *J. Bacteriol.* **188**, 4183–4189 (2006).
122. Guo, H., Yi, W., Song, J. K. & Wang, P. G. Current understanding on biosynthesis of microbial polysaccharides. *Curr. Top. Med. Chem.* **8**, 141–151 (2008).
123. Xayarath, B. & Yother, J. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J. Bacteriol.* **189**, 3369–3381 (2007).
124. Xia, G., Kohler, T. & Peschel, A. The wall teichoic acid and lipoteichoic acid polymers of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* **300**, 148–154 (2010).
125. Dassy, B., Stringfellow, W. T., Lieb, M. & Fournier, J. M. Production of type 5 capsular polysaccharide by *Staphylococcus aureus* grown in a semi-synthetic medium. *J. Gen. Microbiol.* **137**, 1155–1162 (1991).
126. Duguid, J. P. & Wilkinson, J. F. The influence of cultural conditions on polysaccharide production by *Aerobacter aerogenes*. *J. Gen. Microbiol.* **9**, 174–189 (1953).
127. Dunican, L. K. & Seeley, H. W. Extracellular polysaccharide synthesis by members of the genus *Lactobacillus*: conditions for formation and accumulation. *J. Gen. Microbiol.* **40**, 297–308 (1965).
128. Reid, A. N. & Cuthbertson, L. *Bacterial Glycomics: Current Research, Technology and Applications*. (Horizon Scientific Press, 2012).
129. Cieslewicz, M. & Vimr, E. Thermoregulation of *kpsF*, the first region 1 gene in the *kps* locus for polysialic acid biosynthesis in *Escherichia coli* K1. *J. Bacteriol.* **178**, 3212–3220 (1996).
130. Rowe, S., Hodson, N., Griffiths, G. & Roberts, I. S. Regulation of the *Escherichia coli* K5 capsule gene cluster: evidence for the roles of H-NS, BipA, and integration host factor in regulation of group 2 capsule gene clusters in pathogenic *E. coli*. *J. Bacteriol.* **182**, 2741–2745 (2000).

131. Arricau, N. *et al.* The RcsB–RcsC regulatory system of *Salmonella typhi* differentially modulates the expression of invasion proteins, flagellin and Vi antigen in response to osmolarity. *Mol. Microbiol.* **29**, 835–850 (1998).
132. Pickard, D. *et al.* Characterization of defined ompR mutants of *Salmonella typhi*: ompR is involved in the regulation of Vi polysaccharide expression. *Infect. Immun.* **62**, 3984–3993 (1994).
133. Bonet, R., Simon-Pujol, M. D. & Congregado, F. Effects of nutrients on exopolysaccharide production and surface properties of *Aeromonas salmonicida*. *Appl. Environ. Microbiol.* **59**, 2437–2441 (1993).
134. Dassy, B. & Fournier, J. Respiratory activity is essential for post-exponential-phase production of type 5 capsular polysaccharide by *Staphylococcus aureus*. *Infect Immun* **64**, 2408–2414 (1996).
135. Lee, J. C., Takeda, S., Livolsi, P. J. & Paoletti, L. C. Effects of in vitro and in vivo growth conditions on expression of type 8 capsular polysaccharide by *Staphylococcus aureus*. *Infect. Immun.* **61**, 1853–1858 (1993).
136. Weiser, J. N. *et al.* Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. *Infect. Immun.* **69**, 5430–5439 (2001).
137. Deghmane, A.-E., Giorgini, D., Larribe, M., Alonso, J.-M. & Taha, M.-K. Down-regulation of pili and capsule of *Neisseria meningitidis* upon contact with epithelial cells is mediated by CrgA regulatory protein. *Mol. Microbiol.* **43**, 1555–1564 (2002).
138. Jarman, T. R. & Pace, G. W. Energy requirements for microbial exopolysaccharide synthesis. *Arch. Microbiol.* **137**, 231–235 (1984).
139. Bacon, D. J. *et al.* A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. *Mol. Microbiol.* **40**, 769–777 (2001).
140. Lai, Y.-C., Peng, H.-L. & Chang, H.-Y. RmpA2, an activator of capsule biosynthesis in *Klebsiella pneumoniae* CG43, regulates K2 cps gene expression at the transcriptional level. *J. Bacteriol.* **185**, 788–800 (2003).
141. Trisler, P. & Gottesman, S. Ion transcriptional regulation of genes necessary for capsular polysaccharide synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **160**, 184–191 (1984).
142. Waite, R. D., Penfold, D. W., Struthers, J. K. & Dowson, C. G. Spontaneous sequence duplications within capsule genes cap8E and tts control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37. *Microbiology* **149**, 497–504 (2003).
143. Wamel, W. van *et al.* Regulation of *Staphylococcus aureus* type 5 capsular polysaccharides by agr and sarA in vitro and in an experimental endocarditis model. *Microb. Pathog.* **33**, 73–79 (2002).
144. Grangeasse, C., Cozzone, A. J., Deutscher, J. & Mijakovic, I. Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem. Sci.* **32**, 86–94 (2007).
145. Doublet, P., Vincent, C., Grangeasse, C., Cozzone, A. J. & Duclos, B. On the binding of ATP to the autophosphorylating protein, Ptk, of the bacterium *Acinetobacter johnsonii*. *FEBS Lett.* **445**, 137–143 (1999).
146. Lee, D. C., Zheng, J., She, Y.-M. & Jia, Z. Structure of *Escherichia coli* tyrosine kinase Etk reveals a novel activation mechanism. *EMBO J.* **27**, 1758–1766 (2008).

147. Olivares-Illana, V. *et al.* Structural basis for the regulation mechanism of the tyrosine kinase CapB from *Staphylococcus aureus*. *PLoS Biol.* **6**, e143 (2008).
148. Klein, G., Dartigalongue, C. & Raina, S. Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*. *Mol. Microbiol.* **48**, 269–285 (2003).
149. Mijakovic, I. *et al.* Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J.* **22**, 4709–4718 (2003).
150. Mijakovic, I. *et al.* Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.* **34**, 1588–1596 (2006).
151. Jers, C. *et al.* *Bacillus subtilis* BY-kinase PtkA controls enzyme activity and localization of its protein substrates. *Mol. Microbiol.* **77**, 287–299 (2010).
152. Petranovic, D. *et al.* *Bacillus subtilis* strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication. *Mol. Microbiol.* **63**, 1797–1805 (2007).
153. Doublet, P., Grangeasse, C., Obadia, B., Vaganay, E. & Cozzone, A. J. Structural organization of the protein-tyrosine autokinase Wzc within *Escherichia coli* cells. *J. Biol. Chem.* **277**, 37339–37348 (2002).
154. Grangeasse, C., Doublet, P. & Cozzone, A. J. Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K12 occurs through a two-step process. *J. Biol. Chem.* **277**, 7127–7135 (2002).
155. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945 (1982).
156. Paiment, A., Hocking, J. & Whitfield, C. Impact of phosphorylation of specific residues in the tyrosine autokinase, Wzc, on its activity in assembly of group 1 capsules in *Escherichia coli*. *J. Bacteriol.* **184**, 6437–6447 (2002).
157. Duclos, B., Grangeasse, C., Vaganay, E., Riberty, M. & Cozzone, A. J. Autophosphorylation of a bacterial protein at Tyrosine. *J. Mol. Biol.* **259**, 891–895 (1996).
158. Collins, R. F. *et al.* Periplasmic protein-protein contacts in the inner membrane protein Wzc form a tetrameric complex required for the assembly of *Escherichia coli* group 1 capsules. *J. Biol. Chem.* **281**, 2144–2150 (2006).
159. Bender, M. H., Cartee, R. T. & Yother, J. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. *J. Bacteriol.* **185**, 6057–6066 (2003).
160. Morona, J. K., Paton, J. C., Miller, D. C. & Morona, R. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. *Mol. Microbiol.* **35**, 1431–1442 (2000).
161. Soulat, D. *et al.* *Staphylococcus aureus* operates protein-tyrosine phosphorylation through a specific mechanism. *J. Biol. Chem.* **281**, 14048–14056 (2006).
162. Grangeasse, C. *et al.* Functional characterization of the low-molecular-mass phosphotyrosine-protein phosphatase of *Acinetobacter johnsonii*. *J. Mol. Biol.* **278**, 339–347 (1998).
163. Mijakovic, I. *et al.* In vitro characterization of the *Bacillus subtilis* protein tyrosine phosphatase YwqE. *J. Bacteriol.* **187**, 3384–3390 (2005).



164. Morona, J. K., Morona, R., Miller, D. C. & Paton, J. C. Streptococcus pneumoniae capsule biosynthesis protein CpsB is a novel manganese-dependent phosphotyrosine-protein phosphatase. *J. Bacteriol.* **184**, 577–583 (2002).
165. Niemeyer, D. & Becker, A. The molecular weight distribution of succinoglycan produced by *Sinorhizobium meliloti* is influenced by specific tyrosine phosphorylation and ATPase activity of the cytoplasmic domain of the ExoP protein. *J. Bacteriol.* **183**, 5163–5170 (2001).
166. Wugeditsch, T. *et al.* Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *J. Biol. Chem.* **276**, 2361–2371 (2001).
167. MacLachlan, P. R., Keenleyside, W. J., Dodgson, C. & Whitfield, C. Formation of the K30 (group I) capsule in *Escherichia coli* O9:K30 does not require attachment to lipopolysaccharide lipid A-core. *J. Bacteriol.* **175**, 7515–7522 (1993).
168. Collins, R. F. *et al.* The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2390–2395 (2007).
169. Whitfield, C. & Larue, K. Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. *Nat. Struct. Mol. Biol.* **15**, 121–123 (2008).
170. Bastin, D. A., Stevenson, G., Brown, P. K., Haase, A. & Reeves, P. R. Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. *Mol. Microbiol.* **7**, 725–734 (1993).
171. Marolda, C. L., Tatar, L. D., Alaimo, C., Aebi, M. & Valvano, M. A. Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide O antigen. *J. Bacteriol.* **188**, 5124–5135 (2006).
172. McNulty, C. *et al.* The cell surface expression of group 2 capsular polysaccharides in *Escherichia coli*: the role of KpsD, RhsA and a multi-protein complex at the pole of the cell. *Mol. Microbiol.* **59**, 907–922 (2006).
173. Morona, R., van den Bosch, L. & Manning, P. A. Molecular, genetic, and topological characterization of O-antigen chain length regulation in *Shigella flexneri*. *J. Bacteriol.* **177**, 1059–1068 (1995).
174. Tocilj, A. *et al.* Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat. Struct. Mol. Biol.* **15**, 130–138 (2008).
175. Grangeasse, C. *et al.* Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J. Biol. Chem.* **278**, 39323–39329 (2003).
176. Minic, Z. *et al.* Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in *Streptococcus thermophilus*, by EpsD tyrosine kinase. *J. Bacteriol.* **189**, 1351–1357 (2007).
177. Vos, P. *et al.* *Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes*. (Springer Science & Business Media, 2011).
178. Armstrong-Esther, C. A. Carriage patterns of *Staphylococcus aureus* in a healthy non-hospital population of adults and children. *Ann. Hum. Biol.* **3**, 221–227 (1976).

179. Guinan, M. E. *et al.* Vaginal colonization with *Staphylococcus aureus* in healthy women: a review of four studies. *Ann. Intern. Med.* **96**, 944–947 (1982).
180. Rimland, D. & Roberson, B. Gastrointestinal carriage of methicillin-resistant *Staphylococcus aureus*. *J. Clin. Microbiol.* **24**, 137–138 (1986).
181. Gillaspay, A. F. & Landolo, J. J. *Desk Encyclopedia of Microbiology*. (Academic Press, 2003).
182. Lowy, F. D. *Staphylococcus aureus* infections. *N. Engl. J. Med.* **339**, 520–532 (1998).
183. Ogston, A. Micrococcus poisoning. *J. Anat. Physiol.* **16**, 526–567 (1882).
184. Chambers, H. F. & DeLeo, F. R. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* **7**, 629–641 (2009).
185. Boucher, H. W. & Corey, G. R. Epidemiology of methicillin-resistant *Staphylococcus aureus*. *Clin. Infect. Dis.* **46**, S344–S349 (2008).
186. David, M. Z. & Daum, R. S. Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. *Clin. Microbiol. Rev.* **23**, 616–687 (2010).
187. Jevons, M. P. ‘Celbenin’-resistant staphylococci. *Br. Med. J.* **1**, 124–125 (1961).
188. Barber, M. Methicillin-resistant staphylococci. *J. Clin. Pathol.* **14**, 385 (1961).
189. Boyce, J. M. Methicillin-resistant *Staphylococcus aureus*. Detection, epidemiology, and control measures. *Infect. Dis. Clin. North Am.* **3**, 901–913 (1989).
190. Centers for Disease Control and Prevention. Invasive methicillin-resistant *Staphylococcus aureus* infections among dialysis patients—United States, 2005. *MMWR Morb. Mortal. Wkly. Rep.* **56**, 197 (2007).
191. Coello, R., Glynn, J. R., Gaspar, C., Picazo, J. J. & Fereres, J. Risk factors for developing clinical infection with methicillin-resistant *Staphylococcus aureus* (MRSA) amongst hospital patients initially only colonized with MRSA. *J. Hosp. Infect.* **37**, 39–46 (1997).
192. Mulhausen, P. L., Harrell, L. J., Weinberger, M., Kochersberger, G. G. & Feussner, J. R. Contrasting methicillin-resistant *Staphylococcus aureus* colonization in veterans affairs and community nursing homes. *Am. J. Med.* **100**, 24–31 (1996).
193. Voss, A., Milatovic, D., Wallrauch-Schwarz, C., Rosdahl, V. T. & Braveny, I. Methicillin-resistant *Staphylococcus aureus* in Europe. *Eur. J. Clin. Microbiol. Infect. Dis.* **13**, 50–55 (1994).
194. Dowzicky, M. *et al.* Characterization of isolates associated with emerging resistance to quinupristin/dalfopristin (Synercid) during a worldwide clinical program. *Diagn. Microbiol. Infect. Dis.* **37**, 57–62 (2000).
195. Limbago, B. M. *et al.* Report of the 13th vancomycin-resistant *Staphylococcus aureus* isolate from the United States. *J. Clin. Microbiol.* **52**, 998–1002 (2014).
196. Marty, F. M. *et al.* Emergence of a clinical daptomycin-resistant *Staphylococcus aureus* isolate during treatment of methicillin-resistant *Staphylococcus aureus* bacteremia and osteomyelitis. *J. Clin. Microbiol.* **44**, 595–597 (2006).
197. Meka, V. G. *et al.* Linezolid resistance in sequential *Staphylococcus aureus* isolates associated with a T2500A mutation in the 23S rRNA gene and loss of a single copy of rRNA. *J. Infect. Dis.* **190**, 311–317 (2004).

198. European Centre for Disease Prevention and Control/European Medicines Agency (ECDC/EMA). *Joint technical report. The bacterial challenge: time to react.* (ECDC/EMA, 2009).
199. Archer, G. L. Staphylococcus aureus: a well-armed pathogen. *Clin. Infect. Dis.* **26**, 1179–1181 (1998).
200. Cocchiari, J. L. *et al.* Molecular characterization of the capsule locus from non-typeable Staphylococcus aureus. *Mol. Microbiol.* **59**, 948–960 (2006).
201. Verdier, I. *et al.* Identification of the capsular polysaccharides in Staphylococcus aureus clinical isolates by PCR and agglutination tests. *J. Clin. Microbiol.* **45**, 725–729 (2007).
202. Arbeit, R. D., Karakawa, W. W., Vann, W. F. & Robbins, J. B. Predominance of two newly described capsular polysaccharide types among clinical isolates of Staphylococcus aureus. *Diagn. Microbiol. Infect. Dis.* **2**, 85–91 (1984).
203. Fournier, J. M., Vann, W. F. & Karakawa, W. W. Purification and characterization of Staphylococcus aureus type 8 capsular polysaccharide. *Infect. Immun.* **45**, 87–93 (1984).
204. Jones, C. Revised structures for the capsular polysaccharides from Staphylococcus aureus types 5 and 8, components of novel glycoconjugate vaccines. *Carbohydr. Res.* **340**, 1097–1106 (2005).
205. Moreau, M. *et al.* Structure of the type 5 capsular polysaccharide of Staphylococcus aureus. *Carbohydr. Res.* **201**, 285–297 (1990).
206. Sau, S. *et al.* The Staphylococcus aureus allelic genetic loci for serotype 5 and 8 capsule expression contain the type-specific genes flanked by common genes. *Microbiology* **143**, 2395–2405 (1997).
207. Kneidinger, B. *et al.* Three highly conserved proteins catalyze the conversion of UDP-N-acetyl-D-glucosamine to precursors for the biosynthesis of O antigen in Pseudomonas aeruginosa O11 and capsule in Staphylococcus aureus type 5. Implications for the UDP-N-acetyl-L-fucosamine biosynthetic pathway. *J. Biol. Chem.* **278**, 3615–27 (2003).
208. Kiser, K. B., Bhasin, N., Deng, L. & Lee, J. C. Staphylococcus aureus cap5P encodes a UDP-N-acetylglucosamine 2-epimerase with functional redundancy. *J. Bacteriol.* **181**, 4818–4824 (1999).
209. Bhasin, N. *et al.* Identification of a gene essential for O-acetylation of the Staphylococcus aureus type 5 capsular polysaccharide. *Mol. Microbiol.* **27**, 9–21 (1998).
210. Chan, Y. G.-Y., Kim, H. K., Schneewind, O. & Missiakas, D. The capsular polysaccharide of Staphylococcus aureus is attached to peptidoglycan by the LytR-CpsA-Psr (LCP) family of enzymes. *J. Biol. Chem.* **289**, 15680–15690 (2014).
211. Cunnion, K. M., Lee, J. C. & Frank, M. M. Capsule production and growth phase influence binding of complement to Staphylococcus aureus. *Infect. Immun.* **69**, 6796–6803 (2001).
212. Poutrel, B., Gilbert, F. & Lebrun, M. Effects of culture conditions on production of type 5 capsular polysaccharide by human and bovine Staphylococcus aureus strains. *Clin Diagn Lab Immunol* **2**, 166–171 (1995).
213. Dassy, B., Hogan, T., Foster, T. J. & Fournier, J.-M. Involvement of the accessory gene regulator (agr) in expression of type 5 capsular polysaccharide by Staphylococcus aureus. *J. Gen. Microbiol.* **139**, 1301–1306 (1993).

214. Luong, T. T. & Lee, C. Y. The *arl* locus positively regulates *Staphylococcus aureus* type 5 capsule via an *mgrA*-dependent pathway. *Microbiology* **152**, 3123–3131 (2006).
215. Pohlmann-Dietze, P. *et al.* Adherence of *Staphylococcus aureus* to endothelial cells: influence of capsular polysaccharide, global regulator *agr*, and bacterial growth phase. *Infect Immun* **68**, 4865–4871 (2000).
216. Soulat, D., Grangeasse, C., Vaganay, E., Cozzone, A. J. & Duclos, B. UDP-acetyl-mannosamine dehydrogenase is an endogenous protein substrate of *Staphylococcus aureus* protein-tyrosine kinase activity. *J. Mol. Microbiol. Biotechnol.* **13**, 45–54 (2007).
217. Gruszczuk, J. *et al.* Comparative analysis of the Tyr-kinases CapB1 and CapB2 fused to their cognate modulators CapA1 and CapA2 from *Staphylococcus aureus*. *PLOS ONE* **8**, e75958 (2013).
218. Gruszczuk, J. *et al.* Structure analysis of the *Staphylococcus aureus* UDP-N-acetyl-mannosamine dehydrogenase Cap5O involved in capsular polysaccharide biosynthesis. *J. Biol. Chem.* **286**, 17112–17121 (2011).
219. Jörnvall, H. *et al.* Short-chain dehydrogenases/reductases (SDR). *Biochemistry (Mosc.)* **34**, 6003–6013 (1995).
220. Kavanagh, K. L., Jörnvall, H., Persson, B. & Oppermann, U. Medium- and short-chain dehydrogenase/reductase gene and protein families. *Cell. Mol. Life Sci.* **65**, 3895–3906 (2008).
221. Jörnvall, H., Höög, J.-O. & Persson, B. SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett.* **445**, 261–264 (1999).
222. Kallberg, Y., Oppermann, U. & Persson, B. Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. *FEBS J.* **277**, 2375–2386 (2010).
223. Rossmann, M. G., Moras, D. & Olsen, K. W. Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**, 194–199 (1974).
224. Kleiger, G. & Eisenberg, D. GXXXG and GXXXA motifs stabilize FAD and NAD(P)-binding Rossmann folds through C $\alpha$ -H $\cdots$ O hydrogen bonds and van der Waals interactions. *J. Mol. Biol.* **323**, 69–76 (2002).
225. Wierenga, R. K., De Maeyer, M. C. H. & Hol, W. G. J. Interaction of pyrophosphate moieties with  $\alpha$ -helices in dinucleotide-binding proteins. *Biochemistry (Mosc.)* **24**, 1346–1357 (1985).
226. Filling, C. *et al.* Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J. Biol. Chem.* **277**, 25677–25684 (2002).
227. Ghosh, D. *et al.* Three-dimensional structure of holo 3- $\alpha$ , 20- $\beta$ -hydroxysteroid dehydrogenase: a member of a short-chain dehydrogenase family. *Proc. Natl. Acad. Sci.* **88**, 10064–10068 (1991).
228. Miyafusa, T., Caaveiro, J. M. M., Tanaka, Y., Tanner, M. E. & Tsumoto, K. Crystal structure of the capsular polysaccharide synthesizing protein CapE of *Staphylococcus aureus*. *Biosci. Rep.* **33**, 463–474 (2013).
229. Kallberg, Y. & Persson, B. Prediction of coenzyme specificity in dehydrogenases/reductases. *FEBS J.* **273**, 1177–1184 (2006).

230. Giske, C. G., Monnet, D. L., Cars, O., Carmeli, Y. & others. Clinical and economic impact of common multidrug-resistant gram-negative bacilli. *Antimicrob. Agents Chemother.* **52**, 813–821 (2008).
231. Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **48**, 1–12 (2009).
232. Talbot, G. H. *et al.* Bad bugs need drugs: an update on the development pipeline from the Antimicrobial Availability Task Force of the Infectious Diseases Society of America. *Clin. Infect. Dis.* **42**, 657–668 (2006).
233. Rausch, M. Charakterisierung von Kapselbiosynthese-Proteinen aus *Staphylococcus aureus*. *Diploma Thesis Univ. Bonn* (2012).
234. Miyafusa, T., Caaveiro, J. M., Tanaka, Y. & Tsumoto, K. Dynamic elements govern the catalytic activity of CapE, a capsular polysaccharide-synthesizing enzyme from *Staphylococcus aureus*. *FEBS Lett.* **587**, 3824–3830 (2013).
235. Müller, A., Ulm, H., Reder-Christ, K., Sahl, H.-G. & Schneider, T. Interaction of type A lantibiotics with undecaprenol-bound cell envelope precursors. *Microb. Drug Resist.* **18**, 261–270 (2012).
236. William Studier, F., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. in *Methods in Enzymology* (ed. David V. Goeddel) **Volume 185**, 60–89 (Academic Press, 1990).
237. Schoenhofen, I. C. *et al.* Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*. Enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J. Biol. Chem.* **281**, 723–732 (2006).
238. Miroux, B. & Walker, J. E. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298 (1996).
239. Woodcock, D. M. *et al.* Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.* **17**, 3469–3478 (1989).
240. Baba, T., Bae, T., Schneewind, O., Takeuchi, F. & Hiramatsu, K. Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. *J. Bacteriol.* **190**, 300–310 (2008).
241. Sambrook, J., Fritsch, E. F., Maniatis, T. & others. *Molecular cloning*. **2**, (Cold spring harbor laboratory press New York, 1989).
242. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
243. Laemmli, U. K. Most commonly used discontinuous buffer system for SDS electrophoresis. *Nature* **227**, 680–685 (1970).
244. Moos, M., Nguyen, N. Y. & Liu, T. Y. Reproducible high yield sequencing of proteins electrophoretically separated and transferred to an inert support. *J. Biol. Chem.* **263**, 6005–6008 (1988).
245. Porath, J., Carlsson, J. A. N., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598–599 (1975).
246. Menten, L. & Michaelis, M. I. Die Kinetik der Invertinwirkung. *Biochem Z* **49**, 333–369 (1913).

247. FDA, C. Guidance for industry: bioanalytical method validation. US Department of Health and Human Services. (2001).
248. ICH. Validation of analytical procedures: text and methodology, harmonized tripartite guideline. (2005).
249. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **203**, 173–179 (1992).
250. Jenö, P., Mini, T., Moes, S., Hintermann, E. & Horst, M. Internal sequences from proteins digested in polyacrylamide gels. *Anal. Biochem.* **224**, 75–82 (1995).
251. Taus, T. *et al.* Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* **10**, 5354–5362 (2011).
252. Käll, L., Storey, J. D., MacCoss, M. J. & Noble, W. S. Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. *J. Proteome Res.* **7**, 29–34 (2007).
253. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
254. Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10915–10919 (1992).
255. Tatusova, T. A. & Madden, T. L. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* **174**, 247–250 (1999).
256. Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453 (1970).
257. Papadopoulos, J. S. & Agarwala, R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinforma. Oxf. Engl.* **23**, 1073–1079 (2007).
258. Sonnhammer, E. L., Von Heijne, G., Krogh, A. *et al.* A hidden Markov model for predicting transmembrane helices in protein sequences. in *Ismb* **6**, 175–182 (1998).
259. Blom, N., Gammeltoft, S. & Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **294**, 1351–1362 (1999).
260. Whiteside, M. D., Winsor, G. L., Laird, M. R. & Brinkman, F. S. L. OrthoLugeDB: a bacterial and archaeal orthology resource for improved comparative genomic analysis. *Nucleic Acids Res.* **41**, D366–376 (2013).
261. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinforma. Oxf. Engl.* **23**, 2947–2948 (2007).
262. Gonnet, G. H., Cohen, M. A. & Benner, S. A. Exhaustive matching of the entire protein sequence database. *Science* **256**, 1443–1445 (1992).
263. Vallon, O. New sequence motifs in flavoproteins: Evidence for common ancestry and tools to predict structure. *Proteins Struct. Funct. Bioinforma.* **38**, 95–114 (2000).
264. Takamitsu, M., Jose, M. M., Yoshikazu, T. & Kouhei, T. Crystal structure of the enzyme CapF of *Staphylococcus aureus* reveals a unique architecture composed of two functional domains. *Biochem. J.* **443**, 671–680 (2012).

265. Cunneen, M. M., Liu, B., Wang, L. & Reeves, P. R. Biosynthesis of UDP-GlcNAc, UndPP-GlcNAc and UDP-GlcNAcA involves three easily distinguished 4-epimerase enzymes, Gne, Gnu and GnaB. *PLoS ONE* **8**, e67646 (2013).
266. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
267. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
268. Chevenet, F., Brun, C., Bañuls, A.-L., Jacq, B. & Christen, R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* **7**, 439 (2006).
269. Marchler-Bauer, A. *et al.* CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**, D225–229 (2011).
270. Tatusov, R. L. *et al.* The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**, 41 (2003).
271. Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–301 (2012).
272. Gascuel, O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.* **14**, 685–695 (1997).
273. Sanderson, M. J. & Shaffer, H. B. Troubleshooting molecular phylogenetic analyses. *Annu. Rev. Ecol. Syst.* **33**, 49–72 (2002).
274. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783 (1985).
275. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–815 (2013).
276. King, J. D., Kocíncová, D., Westman, E. L. & Lam, J. S. Review: Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Innate Immun.* **15**, 261–312 (2009).
277. Morrison, M. J. & Imperiali, B. The renaissance of bacillosamine and its derivatives: pathway characterization and implications in pathogenicity. *Biochemistry (Mosc.)* **53**, 624–638 (2014).
278. Allard, S. T. M. *et al.* Toward a structural understanding of the dehydratase mechanism. *Struct. Lond. Engl.* **1993** **10**, 81–92 (2002).
279. McNally, D. J. *et al.* Identification of Labile UDP-Ketosugars in *Helicobacter pylori*, *Campylobacter jejuni* and *Pseudomonas aeruginosa*: key metabolites used to make glycan virulence factors. *ChemBioChem* **7**, 1865–1868 (2006).
280. Creuzenet, C., Schur, M. J., Li, J., Wakarchuk, W. W. & Lam, J. S. FlaA1, a new bifunctional UDP-GlcNAc C6 dehydratase/C4 reductase from *Helicobacter pylori*. *J. Biol. Chem.* **275**, 34873–80 (2000).
281. Creuzenet, C. & Lam, J. S. Topological and functional characterization of WbpM, an inner membrane UDP-GlcNAc C6 dehydratase essential for lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **41**, 1295–310 (2001).
282. Ishiyama, N. *et al.* Structural studies of FlaA1 from *Helicobacter pylori* reveal the mechanism for inverting 4,6-dehydratase activity. *J. Biol. Chem.* **281**, 24489–24495 (2006).

283. Hartley, M. D. *et al.* Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N'-diacetylglucosamine. *Biochemistry (Mosc.)* **50**, 4936–4948 (2011).
284. Morrison, M. J. & Imperiali, B. Biosynthesis of UDP-N,N'-diacetylglucosamine in *Acinetobacter baumannii*: biochemical characterization and correlation to existing pathways. *Arch. Biochem. Biophys.* **536**, 72–80 (2013).
285. Morrison, J. P., Schoenhofen, I. C. & Tanner, M. E. Mechanistic studies on PseB of pseudaminic acid biosynthesis: A UDP-N-acetylglucosamine 5-inverting 4, 6-dehydratase. *Bioorganic Chem.* **36**, 312–320 (2008).
286. Mulrooney, E. F., Poon, K. K. H., McNally, D. J., Brisson, J.-R. & Lam, J. S. Biosynthesis of UDP-N-acetyl-L-fucosamine, a precursor to the biosynthesis of lipopolysaccharide in *Pseudomonas aeruginosa* serotype O11. *J. Biol. Chem.* **280**, 19535–19542 (2005).
287. Pinta, E. *et al.* Identification and role of a 6-deoxy-4-keto-hexosamine in the lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O: 3. *Chem.-Eur. J.* **15**, 9747–9754 (2009).
288. Li, T., Simonds, L., Kovrigin, E. L. & Noel, K. D. In vitro biosynthesis and chemical identification of UDP-N-acetyl-D-quinovosamine (UDP-D-QuiNAc). *J. Biol. Chem.* **289**, 18110–18120. (2014).
289. Thibodeaux, C. J., Melançon, C. E. & Liu, H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed Engl.* **47**, 9814–9859 (2008).
290. Rush, J. S., Alaimo, C., Robbiani, R., Wacker, M. & Waechter, C. J. A novel epimerase that converts GlcNAc-PP-undecaprenol to GalNAc-PP-undecaprenol in *Escherichia coli* O157. *J. Biol. Chem.* **285**, 1671–1680 (2010).
291. Twine, S. M., Vinogradov, E., Lindgren, H., Sjostedt, A. & Conlan, J. W. Roles for *wbtC*, *wbtI*, and *kdtA* genes in lipopolysaccharide biosynthesis, protein glycosylation, virulence, and immunogenicity in *Francisella tularensis* strain SCHU S4. *Pathog. Basel Switz.* **1**, 12–29 (2012).
292. Bélanger, M., Burrows, L. L. & Lam, J. S. Functional analysis of genes responsible for the synthesis of the B-band O antigen of *Pseudomonas aeruginosa* serotype O6 lipopolysaccharide. *Microbiology* **145**, 3505–3521 (1999).
293. Lin, W. S., Cunneen, T. & Lee, C. Y. Sequence analysis and molecular characterization of genes required for the biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*. *J. Bacteriol.* **176**, 7005–7016 (1994).
294. Millard, C. M. *et al.* Evolution of the capsular operon of *Streptococcus iniae* in response to vaccination. *Appl. Environ. Microbiol.* **78**, 8219–8226 (2012).
295. Stroehrer, U. H., Jedani, K. E. & Manning, P. A. Genetic organization of the regions associated with surface polysaccharide synthesis in *Vibrio cholerae* O1, O139 and *Vibrio anguillarum* O1 and O2: a review. *Gene* **223**, 269–282 (1998).
296. Liau, D. F. & Hash, J. H. Structural analysis of the surface polysaccharide of *Staphylococcus aureus* M. *J. Bacteriol.* **131**, 194–200 (1977).
297. Murthy, S. V. K. N., Ann Melly, M., Harris, T. M., Hellerqvist, C. G. & Hash, J. H. The repeating sequence of the capsular polysaccharide of *Staphylococcus aureus* M. *Carbohydr. Res.* **117**, 113–123 (1983).



298. Dean, C. R. *et al.* Characterization of the serogroup O11 O-antigen locus of *Pseudomonas aeruginosa* PA103. *J. Bacteriol.* **181**, 4275–4284 (1999).
299. Fallarino, A., Mavrangelos, C., Stroehrer, U. H. & Manning, P. A. Identification of additional genes required for O-antigen biosynthesis in *Vibrio cholerae* O1. *J. Bacteriol.* **179**, 2147–2153 (1997).
300. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
301. Krinos, C. M. *et al.* Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature* **414**, 555–558 (2001).
302. Comstock, L. E. *et al.* Analysis of a capsular polysaccharide biosynthesis locus of *Bacteroides fragilis*. *Infect. Immun.* **67**, 3525–3532 (1999).
303. Coyne, M. J., Kalka-Moll, W., Tzianabos, A. O., Kasper, D. L. & Comstock, L. E. *Bacteroides fragilis* NCTC9343 produces at least three distinct capsular polysaccharides: cloning, characterization, and reassignment of polysaccharide B and C biosynthesis loci. *Infect. Immun.* **68**, 6176–6181 (2000).
304. Baumann, H., Tzianabos, A. O., Brisson, J. R., Kasper, D. L. & Jennings, H. J. Structural elucidation of two capsular polysaccharides from one strain of *Bacteroides fragilis* using high-resolution NMR spectroscopy. *Biochemistry (Mosc.)* **31**, 4081–4089 (1992).
305. Koonin, E. V. & Wolf, Y. I. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Res.* **36**, 6688–6719 (2008).
306. Yarza, P. *et al.* The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.* **31**, 241–250 (2008).
307. Aono, R. & Uramoto, M. Presence of fucosamine in teichuronic acid of the alkalophilic *Bacillus* strain C-125. *Biochem. J.* **233**, 291–294 (1986).
308. Onoue, S., Niwa, M., Isshiki, Y. & Kawahara, K. Extraction and characterization of the smooth-type lipopolysaccharide from *Fusobacterium nucleatum* JCM 8532 and its biological activities. *Microbiol. Immunol.* **40**, 323–331 (1996).
309. Bystrova, O. V. *et al.* Structural studies on the core and the O-polysaccharide repeating unit of *Pseudomonas aeruginosa* immunotype 1 lipopolysaccharide. *Eur. J. Biochem. FEBS* **269**, 2194–2203 (2002).
310. Sadovskaya, I. *et al.* Structural characterization of the outer core and the O-chain linkage region of lipopolysaccharide from *Pseudomonas aeruginosa* serotype O5. *Eur. J. Biochem. FEBS* **267**, 1640–1650 (2000).
311. Reddy, G. P. *et al.* Purification and determination of the structure of capsular polysaccharide of *Vibrio vulnificus* M06-24. *J. Bacteriol.* **174**, 2620–2630 (1992).
312. Kondakova, A. N. *et al.* Low structural diversity of the O-polysaccharides of *Photobacterium* *asymbiotica* subsp. *asymbiotica* and *australis* and their similarity to the O-polysaccharides of taxonomically remote bacteria including *Francisella tularensis*. *Carbohydr. Res.* **346**, 1951–1955 (2011).
313. Niedziela, T. *et al.* Structures of two novel, serologically nonrelated core oligosaccharides of *Yokenella regensburgei* lipopolysaccharides differing only by a single hexose substitution. *Glycobiology* **20**, 207–214 (2010).

314. Jann, B. *et al.* NMR investigation of the 6-deoxy-L-talose-containing O45, O45-related (O45rel), and O66 polysaccharides of *Escherichia coli*. *Carbohydr. Res.* **278**, 155–165 (1995).
315. Forsberg, L. S., Bhat, U. R. & Carlson, R. W. Structural characterization of the O-antigenic polysaccharide of the lipopolysaccharide from *Rhizobium etli* strain CE3. A unique O-acetylated glycan of discrete size, containing 3-O-methyl-6-deoxy-L-talose and 2,3,4-tri-O-,methyl-L-fucose. *J. Biol. Chem.* **275**, 18851–18863 (2000).
316. Knirel, Y. A. *et al.* Structure of the capsular polysaccharide of *Vibrio cholerae* O139 synonym Bengal containing D-galactose 4,6-cyclophosphate. *Eur. J. Biochem. FEBS* **232**, 391–396 (1995).
317. Bystrova, O. V. *et al.* Full structure of the lipopolysaccharide of *Pseudomonas aeruginosa* immunotype 5. *Biochem. Biokhimiia* **69**, 170–175 (2004).
318. Dean, C. R., Datta, A., Carlson, R. W. & Goldberg, J. B. WbjA adds glucose to complete the O-antigen trisaccharide repeating unit of the lipopolysaccharide of *Pseudomonas aeruginosa* serogroup O11. *J. Bacteriol.* **184**, 323–326 (2002).
319. Bystrova, O. V. *et al.* Structure of the lipopolysaccharide of *Pseudomonas aeruginosa* O-12 with a randomly O-acetylated core region. *Carbohydr. Res.* **338**, 1895–1905 (2003).
320. Ganguly, J., Choudhury, Nair, G. B. & Sen, A. K. Structural characterization of the O-antigenic polysaccharide from the lipopolysaccharide of *Vibrio cholerae* O37. *Indian J. Chem.* **48B**, 729–734 (2009).
321. Anderson, M. S., Eveland, S. S. & Price, N. P. Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. *FEMS Microbiol. Lett.* **191**, 169–175 (2000).
322. Price, N. P. & Momany, F. A. Modeling bacterial UDP-HexNAc:polyprenol-P HexNAc-1-P transferases. *Glycobiology* **15**, 29R–42R (2005).
323. Rocchetta, H. L., Burrows, L. L., Pacan, J. C. & Lam, J. S. Three rhamnosyltransferases responsible for assembly of the A-band D-rhamnan polysaccharide in *Pseudomonas aeruginosa*: a fourth transferase, WbpL, is required for the initiation of both A-band and B-band lipopolysaccharide synthesis. *Mol. Microbiol.* **28**, 1103–1119 (1998).
324. Korbil, J. O., Jensen, L. J., von Mering, C. & Bork, P. Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nat. Biotechnol.* **22**, 911–917 (2004).
325. Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. The use of gene clusters to infer functional coupling. *Proc. Natl. Acad. Sci.* **96**, 2896–2901 (1999).
326. Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D. & Yeates, T. O. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4285–4288 (1999).
327. Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
328. Schoenhofen, I. C. *et al.* Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J. Biol. Chem.* **281**, 723–32 (2006).

329. Burrows, L. L., Urbanic, R. V. & Lam, J. S. Functional conservation of the polysaccharide biosynthetic protein WbpM and its homologues in *Pseudomonas aeruginosa* and other medically significant bacteria. *Infect. Immun.* **68**, 931–936 (2000).
330. Deeb, S. E., Iriban, M. A. & Gust, R. MEKC as a powerful growing analytical technique. *Electrophoresis* **32**, 166–183 (2011).
331. El-Tayeb, A. *et al.* Nucleoside-5'-monophosphates as prodrugs of adenosine A2A receptor agonists activated by ecto-5'-nucleotidase. *J. Med. Chem.* **52**, 7669–7677 (2009).
332. Iqbal, J., Lévesque, S. A., Sévigny, J. & Müller, C. E. A highly sensitive CE-UV method with dynamic coating of silica-fused capillaries for monitoring of nucleotide pyrophosphatase/phosphodiesterase reactions. *Electrophoresis* **29**, 3685–3693 (2008).
333. Lee, S.-Y., Lévesque, S. A., Sévigny, J. & Müller, C. E. A highly sensitive capillary electrophoresis method using p-nitrophenyl 5'-thymidine monophosphate as a substrate for the monitoring of nucleotide pyrophosphatase/phosphodiesterase activities. *J. Chromatogr. B* **911**, 162–169 (2012).
334. Csizmadia, P. MarvinSketch and MarvinView: molecule applets for the World Wide Web. in *Proceedings of ECSOC-3, The Third International Electronic Conference on Synthetic Organic Chemistry, September 1–30* 367–369 (1999).
335. Iqbal, J. An enzyme immobilized microassay in capillary electrophoresis for characterization and inhibition studies of alkaline phosphatases. *Anal. Biochem.* **414**, 226–231 (2011).
336. Melanson, J. E., Baryla, N. E. & Lucy, C. A. Dynamic capillary coatings for electroosmotic flow control in capillary electrophoresis. *TrAC Trends Anal. Chem.* **20**, 365–374 (2001).
337. El-Tayeb, A., Qi, A. & Müller, C. E. Synthesis and structure-activity relationships of uracil nucleotide derivatives and analogues as agonists at human P2Y<sub>2</sub>, P2Y<sub>4</sub>, and P2Y<sub>6</sub> receptors. *J. Med. Chem.* **49**, 7076–7087 (2006).
338. Baqi, Y. & Müller, C. E. Rapid and efficient microwave-assisted copper (0)-catalyzed Ullmann coupling reaction: general access to anilinoanthraquinone derivatives. *Org. Lett.* **9**, 1271–1274 (2007).
339. Voogd, T. E., Vansterkenburg, E. L., Wilting, J. & Janssen, L. H. Recent research on the biological activity of suramin. *Pharmacol. Rev.* **45**, 177–203 (1993).
340. Schneider, T. & Sahl, H. G. An oldie but a goodie-cell wall biosynthesis as antibiotic target pathway. *Int. J. Med. Microbiol.* **300**, 161–169 (2010).
341. Cozzone, A. J. Diversity and specificity of protein-phosphorylating systems in bacteria. *Folia Microbiol. (Praha)* **42**, 165–170 (1997).
342. Guidos, R., J. Combating antimicrobial resistance: policy recommendations to save lives. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **52**, S397–S428 (2011).
343. Souli, M., Galani, I. & Giamarellou, H. Emergence of extensively drug-resistant and pandrug-resistant Gram-negative bacilli in Europe. *Euro Surveill. Bull. Eur. Sur Mal. Transm. Eur. Commun. Dis. Bull.* **13**, (2008).
344. Rasko, D. A. & Sperandio, V. Anti-virulence strategies to combat bacteria-mediated disease. *Nat. Rev. Drug Discov.* **9**, 117–128 (2010).
345. Watts, A. *et al.* *Staphylococcus aureus* strains that express serotype 5 or serotype 8 capsular polysaccharides differ in virulence. *Infect. Immun.* **73**, 3502–11 (2005).

346. McLoughlin, R. M. *et al.* CD4+ T cells and CXC chemokines modulate the pathogenesis of *Staphylococcus aureus* wound infections. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10408–13 (2006).
347. Standish, A. J., Salim, A. A., Zhang, H., Capon, R. J. & Morona, R. Chemical inhibition of bacterial protein tyrosine phosphatase suppresses capsule production. *PLoS One* **7**, e36312 (2012).
348. Cunnion, K. M., Zhang, H.-M. & Frank, M. M. Availability of complement bound to *Staphylococcus aureus* to interact with membrane complement receptors influences efficiency of phagocytosis. *Infect. Immun.* **71**, 656–62 (2003).
349. Weidenmaier, C. *et al.* Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. *Nat. Med.* **10**, 243–5 (2004).
350. Wang, H. *et al.* Discovery of wall teichoic acid inhibitors as potential anti-MRSA  $\beta$ -lactam combination agents. *Chem. Biol.* **20**, 272–84 (2013).
351. Roemer, T., Schneider, T. & Pinho, M. G. Auxiliary factors: a chink in the armor of MRSA resistance to  $\beta$ -lactam antibiotics. *Curr. Opin. Microbiol.* **16**, 538–48 (2013).
352. Orengo, C. A. & Thornton, J. M. Protein families and their evolution—a structural perspective. *Annu Rev Biochem* **74**, 867–900 (2005).
353. Pazos, F., Rausell, A. & Valencia, A. Phylogeny-independent detection of functional residues. *Bioinformatics* **22**, 1440–1448 (2006).
354. Rausell, A., Juan, D., Pazos, F. & Valencia, A. Protein interactions and ligand binding: from protein subfamilies to functional specificity. *Proc. Natl. Acad. Sci.* **107**, 1995–2000 (2010).
355. Jansen, A. *et al.* Production of capsular polysaccharide does not influence *Staphylococcus aureus* vancomycin susceptibility. *BMC Microbiol.* **13**, 65 (2013).
356. Wann, E. R., Dassy, B., Fournier, J.-M. & Foster, T. J. Genetic analysis of the cap5 locus of *Staphylococcus aureus*. *FEMS Microbiol. Lett.* **170**, 97–103 (1999).
357. Haas, B. & Nathens, A. B. Future diagnostic and therapeutic approaches in surgical infections. *Surg. Clin. North Am.* **89**, 539–554 (2009).
358. Short, S. A. & White, D. C. Metabolism of phosphatidylglycerol, lysylphosphatidylglycerol, and cardiolipin of *Staphylococcus aureus*. *J. Bacteriol.* **108**, 219–226 (1971).
359. Bogdanov, M., Heacock, P. N. & Dowhan, W. A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. *EMBO J.* **21**, 2107–2116 (2002).
360. Zhang, W., Bogdanov, M., Pi, J., Pittard, A. J. & Dowhan, W. Reversible topological organization within a polytopic membrane protein is governed by a change in membrane phospholipid composition. *J. Biol. Chem.* **278**, 50128–50135 (2003).
361. Palsdottir, H. & Hunte, C. Lipids in membrane protein structures. *Biochim. Biophys. Acta BBA-Biomembr.* **1666**, 2–18 (2004).
362. Ishiyama, N. Structural studies of the O antigen biosynthetic enzymes from *Pseudomonas aeruginosa*. (2005). at <[http://digitool.library.mcgill.ca/R/?func=dbin-jump-full&object\\_id=100628&local\\_base=GEN01-MCG02](http://digitool.library.mcgill.ca/R/?func=dbin-jump-full&object_id=100628&local_base=GEN01-MCG02)>
363. Creuzenet, C., Urbanic, R. V. & Lam, J. S. Structure-function studies of two novel UDP-GlcNAc C6 dehydratases/C4 reductases. Variation from the SYK dogma. *J. Biol. Chem.* **277**, 26769–26778 (2002).

364. Wiegert, T., Sahm, H. & Sprenger, G. A. The substitution of a single amino acid residue (Ser-116 --> Asp) alters NADP-containing glucose-fructose oxidoreductase of *Zymomonas mobilis* into a glucose dehydrogenase with dual coenzyme specificity. *J. Biol. Chem.* **272**, 13126–13133 (1997).
365. Woodyer, R., van der Donk, W. A. & Zhao, H. Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. *Biochemistry (Mosc.)* **42**, 11604–11614 (2003).
366. Sherbet, D. P. *et al.* Cofactors, redox state, and directional preferences of hydroxysteroid dehydrogenases. *Mol. Cell. Endocrinol.* **265–266**, 83–88 (2007).
367. Ashton, A. R. & Hatch, M. D. Regulation of C4 photosynthesis: regulation of activation and inactivation of NADP-malate dehydrogenase by NADP and NADPH. *Arch. Biochem. Biophys.* **227**, 416–424 (1983).
368. Sjölander, K. Phylogenomic inference of protein molecular function: advances and challenges. *Bioinformatics* **20**, 170–179 (2004).
369. Brown, D. & Sjölander, K. Functional classification using phylogenomic inference. *PLoS Comput Biol* **2**, e77 (2006).
370. Eisen, J. A. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res.* **8**, 163–167 (1998).
371. Allard, S. T. M., Giraud, M.-F. & Naismith, J. H. Epimerases: structure, function and mechanism. *Cell. Mol. Life Sci. CMLS* **58**, 1650–1665 (2001).
372. Tamames, J. Evolution of gene order conservation in prokaryotes. *Genome Biol.* **2**, research0020 (2001).
373. Koonin, E. V., Mushegian, A. R. & Bork, P. Non-orthologous gene displacement. *Trends Genet. TIG* **12**, 334 (1996).
374. Mering, C. von *et al.* STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* **31**, 258–261 (2003).
375. Leonori, D. & Seeberger, P. H. De novo synthesis of the bacterial 2-amino-2,6-dideoxy sugar building blocks D-fucosamine, D-bacillosamine, and D-xylo-6-deoxy-4-ketohexosamine. *Org. Lett.* **14**, 4954–4957 (2012).
376. Dixit, P. D. & Maslov, S. Evolutionary capacitance and control of protein stability in protein-protein interaction networks. *PLoS Comput. Biol.* **9**, e1003023 (2013).
377. Forsberg, L. S., Noel, K. D., Box, J. & Carlson, R. W. Genetic locus and structural characterization of the biochemical defect in the O-antigenic polysaccharide of the symbiotically deficient *Rhizobium etli* mutant, CE166. *J. Biol. Chem.* **278**, 51347–51359 (2003).
378. Brooks, H. B. *et al.* in *Assay Guidance Manual* (eds. Sittampalam, G. S. *et al.*) (Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2004). at <<http://www.ncbi.nlm.nih.gov/books/NBK92007/>>
379. Stein, C. A. Suramin: a novel antineoplastic agent with multiple potential mechanisms of action. *Cancer Res.* **53**, 2239–2248 (1993).
380. Fairlamb, A. H. Chemotherapy of human African trypanosomiasis: current and future prospects. *Trends Parasitol.* **19**, 488–494 (2003).

381. Kleine, F. & Fischer, W. Bericht über die Prüfung von Bayer 205. *Afr. Dt Med Wschr* **48**, 1693 (1922).
382. Brun, R., Blum, J., Chappuis, F. & Burri, C. Human African trypanosomiasis. *The Lancet* **375**, 148–159 (2010).
383. LaPushin, R., Totpal, K., Higuchi, M. & Aggarwal, B. B. Suramin inhibits tumor cell cytotoxicity mediated through natural killer cells, lymphokine-activated killer cells, monocytes, and tumor necrosis factor. *J. Clin. Immunol.* **14**, 39–49 (1994).
384. Jones, R., Parry, R., Lo Leggio, L. & Nickel, P. Inhibition of sperm-zona binding by suramin, a potential 'lead' compound for design of new anti-fertility agents. *Mol. Hum. Reprod.* **2**, 597–605 (1996).
385. Margolles-Clark, E. *et al.* Suramin inhibits the CD40–CD154 costimulatory interaction: a possible mechanism for immunosuppressive effects. *Biochem. Pharmacol.* **77**, 1236–1245 (2009).
386. Pinedo, H. M. & van Rijswijk, R. E. Suramin awakes? *J. Clin. Oncol.* **10**, 875–877 (1992).
387. Nautiyal, A., Patil, K. N. & Muniyappa, K. Suramin is a potent and selective inhibitor of Mycobacterium tuberculosis RecA protein and the SOS response: RecA as a potential target for antibacterial drug discovery. *J. Antimicrob. Chemother.* dku080 (2014). doi:10.1093/jac/dku080
388. Jentsch, K. D., Hunsmann, G., Hartmann, H. & Nickel, P. Inhibition of human immunodeficiency virus type I reverse transcriptase by suramin-related compounds. *J. Gen. Virol.* **68**, 2183–2192 (1987).
389. Willson, M., Callens, M., Kuntz, D. A., Perié, J. & Opperdoes, F. R. Synthesis and activity of inhibitors highly specific for the glycolytic enzymes from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **59**, 201–210 (1993).
390. Hermans, J. M., Haines, D. S., James, P. S. & Jones, R. Kinetics of inhibition of sperm  $\beta$ -acrosin activity by suramin. *FEBS Lett.* **544**, 119–122 (2003).
391. Cadène, M. *et al.* Inhibition of neutrophil serine proteinases by suramin. *J. Biol. Chem.* **272**, 9950–9955 (1997).
392. McGovern, S. L. & Shoichet, B. K. Kinase inhibitors: not just for kinases anymore. *J. Med. Chem.* **46**, 1478–1483 (2003).
393. Torrente, M. P., Castellano, L. M. & Shorter, J. Suramin inhibits Hsp104 ATPase and disaggregase activity. *PLoS ONE* **9**, e110115 (2014).
394. Ganesh, V. K., Muthuvel, S. K., Smith, S. A., Kotwal, G. J. & Murthy, K. H. M. Structural basis for antagonism by suramin of heparin binding to vaccinia complement protein. *Biochemistry (Mosc.)* **44**, 10757–10765 (2005).
395. Schreuder, H. A. *et al.* The intact and cleaved human antithrombin III complex as a model for serpin–proteinase interactions. *Nat. Struct. Mol. Biol.* **1**, 48–54 (1994).
396. Fini, C. *et al.* Boar sperm proacrosin infrared investigation: secondary structure analysis after autoactivation and suramin binding. *Biochem. Mol. Med.* **58**, 37–45 (1996).
397. Gaboriau, D., Howes, E. A., Clark, J. & Jones, R. Binding of sperm proacrosin/ $\beta$ -acrosin to zona pellucida glycoproteins is sulfate and stereodependent. Synthesis of a novel fertilization inhibitor. *Dev. Biol.* **306**, 646–657 (2007).

398. Fromm, J. R., Hileman, R. E., Caldwell, E. E., Weiler, J. M. & Linhardt, R. J. Pattern and spacing of basic amino acids in heparin binding sites. *Arch. Biochem. Biophys.* **343**, 92–100 (1997).
399. Stein, C. A., Khan, T. M., Khaled, Z. & Tonkinson, J. L. Cell surface binding and cellular internalization properties of suramin, a novel antineoplastic agent. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **1**, 509–517 (1995).
400. Sanderson, L., Khan, A. & Thomas, S. Distribution of suramin, an antitrypanosomal drug, across the blood-brain and blood-cerebrospinal fluid interfaces in wild-type and P-glycoprotein transporter-deficient mice. *Antimicrob. Agents Chemother.* **51**, 3136–46 (2007).
401. Yan, L. & Müller, C. E. Preparation, properties, reactions, and adenosine receptor affinities of sulfophenylxanthine nitrophenyl esters: toward the development of sulfonic acid prodrugs with peroral bioavailability. *J. Med. Chem.* **47**, 1031–1043 (2004).
402. Ehrlich, P. u. K. Shiga: Farbtherapeutische Versuche bei Trypanosomenerkrankung. Berlin. *Klin Wochenschr* **41**, 362 (1904).
403. Jørgensen, M. G. Some structural features of vital dyes. *Bull. Exp. Biol. Med.* **83**, 591–595 (1977).
404. Mates, S. M. *et al.* Membrane potential and gentamicin uptake in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci.* **79**, 6693–6697 (1982).
405. Epand, R. M. & Epand, R. F. Lipid domains in bacterial membranes and the action of antimicrobial agents. *Biochim. Biophys. Acta BBA-Biomembr.* **1788**, 289–294 (2009).
406. Wecke, J., Franz, M. & Giesbrecht, P. Inhibition of the bacteriolytic effect of  $\beta$ -lactam-antibiotics on *Staphylococcus aureus* by the polyanionic drugs suramin and Evans Blue. *APMIS* **98**, 71–81 (1990).
407. Lominski, I., Cameron, J. & Wyllie, G. Chaining and unchaining *Streptococcus faecalis*—a hypothesis of the mechanism of bacterial cell separation. *Nature* **181**, 1477–1477 (1958).
408. Giesbrecht, P., Kersten, T., Maidhof, H. & Wecke, J. Staphylococcal cell wall: morphogenesis and fatal variations in the presence of penicillin. *Microbiol. Mol. Biol. Rev.* **62**, 1371–1414 (1998).
409. Yeats, C., Finn, R. D. & Bateman, A. The PASTA domain: a  $\beta$ -lactam-binding domain. *Trends Biochem. Sci.* **27**, 438–440 (2002).
410. Dubrac, S., Boneca, I. G., Poupel, O. & Msadek, T. New insights into the Walk/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in *Staphylococcus aureus*. *J. Bacteriol.* **189**, 8257–8269 (2007).
411. Stein, C. A., LaRocca, R. V., Thomas, R., McAtee, N. & Myers, C. E. Suramin: an anticancer drug with a unique mechanism of action. *J. Clin. Oncol.* **7**, 499–508 (1989).
412. De Bruijn, E. A. *et al.* Therapeutic drug monitoring of suramin and protein binding. *J. Liq. Chromatogr.* **14**, 3719–3733 (1991).
413. Collins, J. M. C. *et al.* Clinical pharmacokinetics of suramin in patients With HTLV-III/LAV Infection. *J. Clin. Pharmacol.* **26**, 22–26 (1986).
414. Hart, P. & Young, M. R. Interference with normal phagosome-lysosome fusion in macrophages, using ingested yeast cells and suramin. *Nature* **256**, 47–49 (1975).

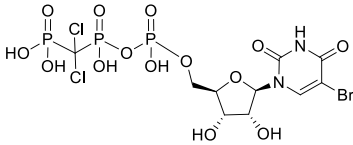
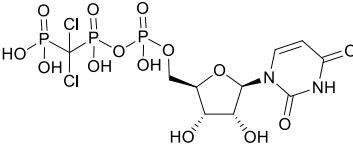
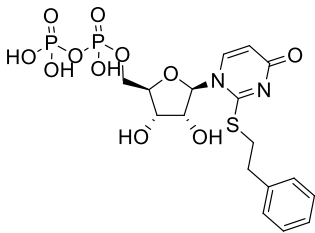
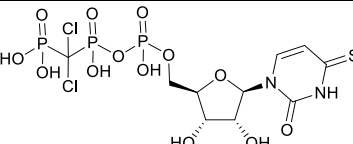
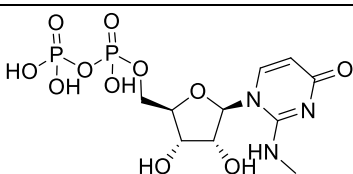
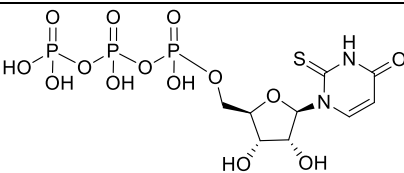
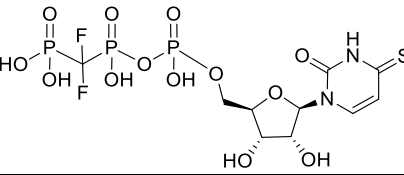
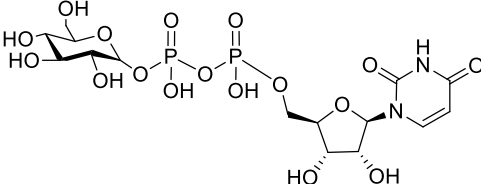
415. Roilides, E., Paschalides, P., Freifeld, A. & Pizzo, P. A. Suppression of polymorphonuclear leukocyte bactericidal activity by suramin. *Antimicrob. Agents Chemother.* **37**, 495–500 (1993).
416. Tan, J. S. & Salstrom, S. J. Bacampicillin, ampicillin, cephalothin, and cephapirin levels in human blood and interstitial fluid. *Antimicrob. Agents Chemother.* **15**, 510–2 (1979).
417. Li, W. & Ulm, H. *et al.* Analysis of the *Staphylococcus aureus* capsule biosynthesis pathway in vitro: characterization of the UDP-GlcNAc C6 dehydratases CapD and CapE and identification of enzyme inhibitors. *Int. J. Med. Microbiol. IJMM* **304**, 958–969 (2014).
418. De Hennezel, L., Ramisse, F., Binder, P., Marchal, G. & Alonso, J.-M. Effective combination therapy for invasive pneumococcal pneumonia with ampicillin and intravenous immunoglobulins in a mouse model. *Antimicrob. Agents Chemother.* **45**, 316–318 (2001).
419. Shi, L. *et al.* Evolution of bacterial protein-tyrosine kinases and their relaxed specificity toward substrates. *Genome Biol. Evol.* **6**, 800–817 (2014).
420. Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).
421. Macek, B. *et al.* The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol. Cell. Proteomics* **6**, 697–707 (2007).
422. Elsholz, A. K., Wacker, S. A. & Losick, R. Self-regulation of exopolysaccharide production in *Bacillus subtilis* by a tyrosine kinase. *Genes Dev.* **28**, 1710–1720 (2014).
423. Sadovskaya, I., Brisson, J.-R., Khieu, N. H., Mutharia, L. M. & Altman, E. Structural characterization of the lipopolysaccharide O-antigen and capsular polysaccharide of *Vibrio ordalii* serotype O: 2. *Eur. J. Biochem.* **253**, 319–327 (1998).
424. MacLean, L. L., Perry, M. B., Crump, E. M. & Kay, W. W. Structural characterization of the lipopolysaccharide O-polysaccharide antigen produced by *Flavobacterium columnare* ATCC 43622. *Eur. J. Biochem.* **270**, 3440–3446 (2003).
425. Kilcoyne, M. *et al.* The structure of the O-polysaccharide of the *Pseudoalteromonas rubra* ATCC 29570 T lipopolysaccharide containing a keto sugar. *Carbohydr. Res.* **340**, 2369–2375 (2005).
426. Jansson, P.-E., Lindberg, B. & Lindquist, U. Structural studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 5. *Carbohydr. Res.* **140**, 101–110 (1985).
427. Briles, D. E., Crain, M. J., Gray, B. M., Forman, C. & Yother, J. Strong association between capsular type and virulence for mice among human isolates of *Streptococcus pneumoniae*. *Infect. Immun.* **60**, 111–116 (1992).
428. Kelly, T., Dillard, J. P. & Yother, J. Effect of genetic switching of capsular type on virulence of *Streptococcus pneumoniae*. *Infect. Immun.* **62**, 1813–1819 (1994).
429. Skurnik, M., Venho, R., Bengoechea, J. A. & Moriyón, I. The lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3 is required for virulence and plays a role in outer membrane integrity. *Mol. Microbiol.* **31**, 1443–1462 (1999).



## 6 Supporting information

SUPPLEMENTARY TABLE 1

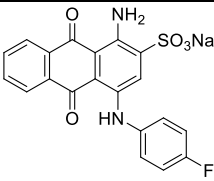
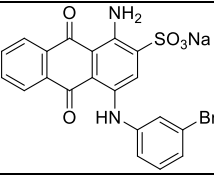
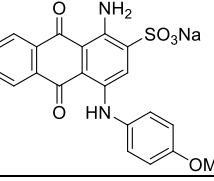
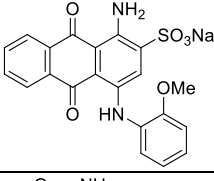
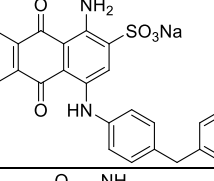
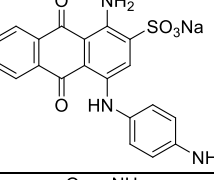
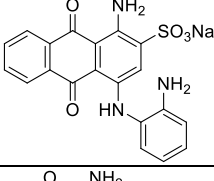
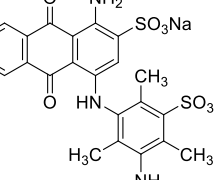
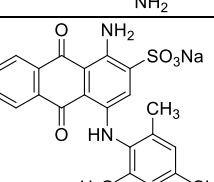
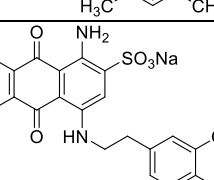
Screening results

Compound code/name (test concentration)	Structure	% Inhibition	
		CapD	CapE
<b>UDP/UTP derivatives and analogs</b>			
Ali 179 <sup>a</sup> (100 μM)		1%	33%
Ali 412C <sup>a</sup> (100 μM)		6%	27%
Ali 404B <sup>a</sup> (100 μM)		7%	33%
Ali 411A <sup>a</sup> (100 μM)		15%	35%
Ali 409A <sup>a</sup> (100 μM)		4%	22%
Ali 149 <sup>a</sup> (100 μM)		-18%	18%
Ali 452 <sup>a</sup> (100 μM)		-4%	25%
UDP-glucose (100 μM) <sup>a</sup>		13%	17%

SUPPLEMENTARY TABLE 1 (continued)

UDP <sup>a</sup> (100 μM)		15%	22%
UTP <sup>a</sup> (100 μM)		-4%	20%
<b>anthraquinone derivatives</b>			
YB001 <sup>a</sup> (10 μM)		10%	17%
YB003 <sup>a</sup> (10 μM)		2%	18%
YB005 <sup>a</sup> (10 μM)		0%	21%
YB009 <sup>a</sup> (10 μM)		25%	15%
YB010 <sup>b</sup> (10 μM)		17%	21%
YB011 <sup>a</sup> (10 μM)		25%	7%
YB012 <sup>a</sup> (10 μM)		26%	12%

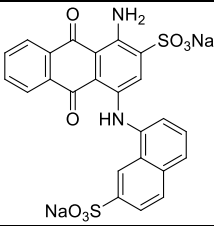
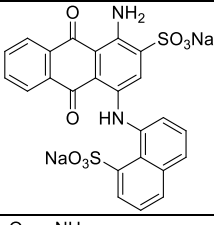
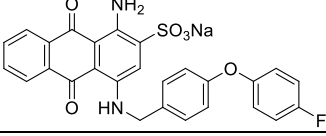
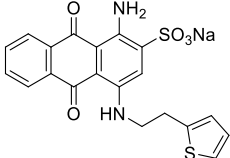
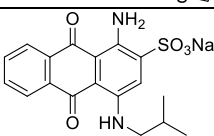
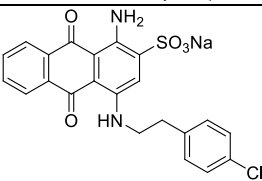
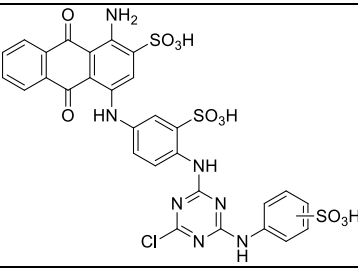
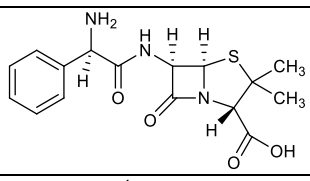
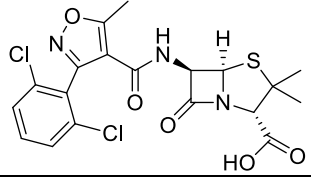
SUPPLEMENTARY TABLE 1 (continued)

YB013 <sup>a</sup> (10 μM)		24%	9%
YB014 <sup>a</sup> (10 μM)		21%	5%
YB015 <sup>a</sup> (10 μM)		10%	7%
YB016 <sup>b</sup> (10 μM)		-1%	7%
YB017 <sup>b</sup> (10 μM)		14%	-1%
YB019 <sup>a</sup> (10 μM)		11%	9%
YB020 <sup>b</sup> (10 μM)		34%	15%
YB024 <sup>a</sup> (10 μM)		13%	9%
YB028 <sup>b</sup> (10 μM)		29%	19%
YB029 <sup>a</sup> (10 μM)		13%	13%

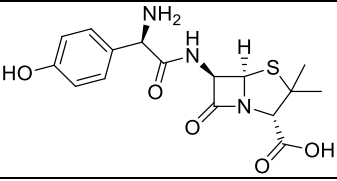
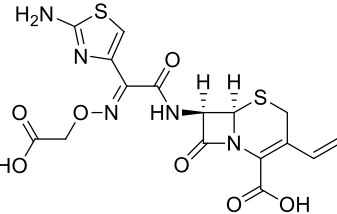
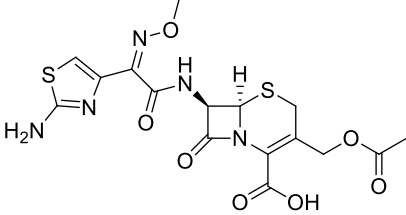
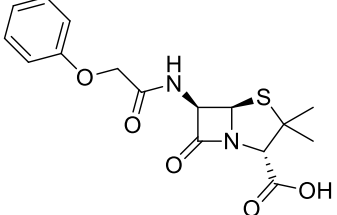
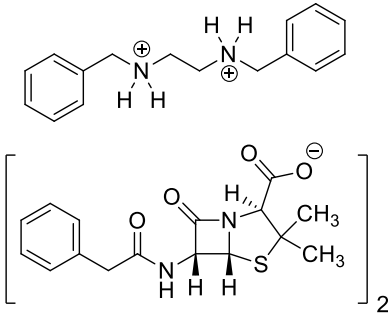
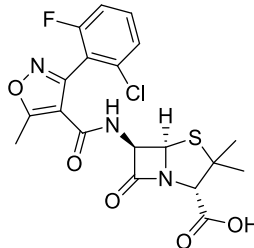
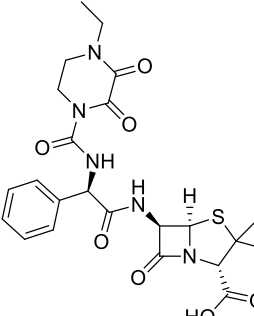
SUPPLEMENTARY TABLE 1 (continued)

YB031 <sup>a</sup> (10 μM)		-4%	5%
YB032 <sup>b</sup> (10 μM)		25%	18%
YB035 <sup>b</sup> (10 μM)		32%	18%
YB038 <sup>a</sup> (10 μM)		23%	15%
YB042 <sup>b</sup> (10 μM)		0%	12%
YB047 <sup>a</sup> (10 μM)		20%	8%
YB053 <sup>b</sup> (10 μM)		33%	7%
YB055 <sup>a</sup> (10 μM)		-10%	-1%
YB056 <sup>b</sup> (10 μM)		27%	4%

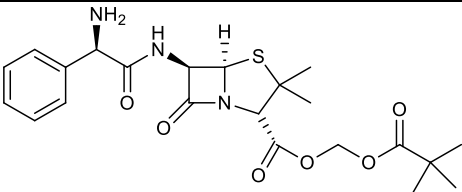
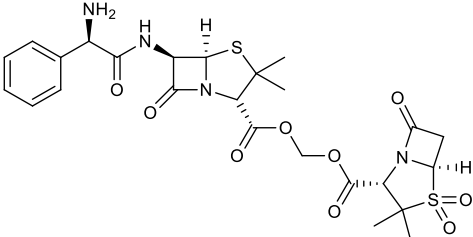
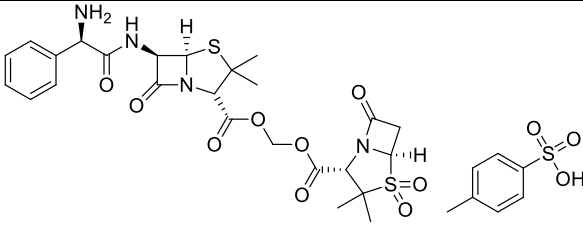
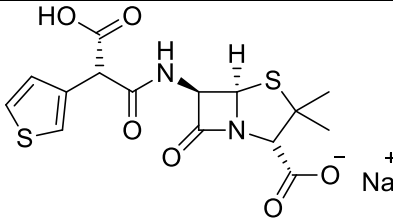
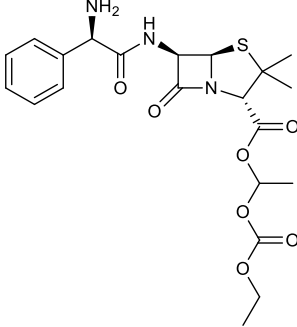
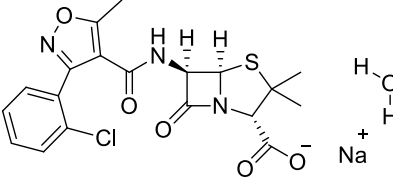
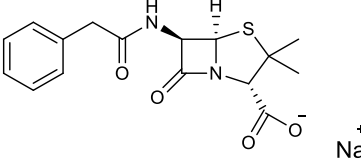
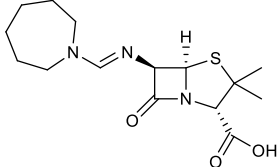
SUPPLEMENTARY TABLE 1 (continued)

YB057 <sup>b</sup> (10 μM)		32%	13%
YB058 <sup>b</sup> (10 μM)		25%	0%
YB068 <sup>b</sup> (10 μM)		17%	12%
YB083 <sup>b</sup> (10 μM)		35%	6%
YB086 <sup>b</sup> (10 μM)		25%	9%
YB087 <sup>b</sup> (10 μM)		32%	17%
RB2 <sup>a</sup> (10 μM)		17%	15%
<b>β-lactam drugs</b>			
ampicillin <sup>b</sup> (10 μM)		41%	5%
dicloxacillin <sup>b</sup> (10 μM)		44%	6%

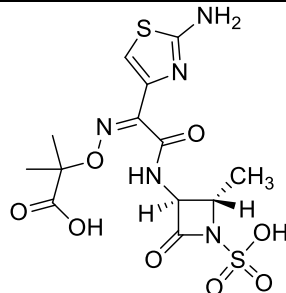
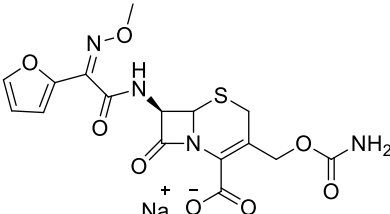
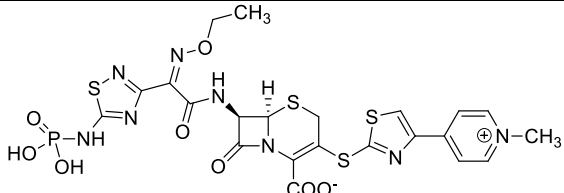
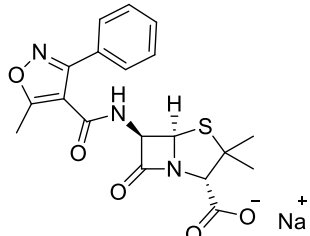
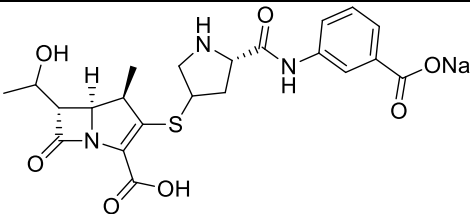
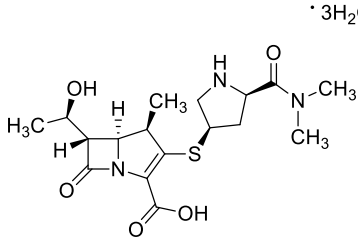
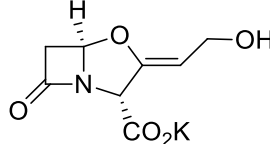
SUPPLEMENTARY TABLE 1 (continued)

amoxicillin <sup>b</sup> (10 μM)		-3%	n.d.
cefixime <sup>b</sup> (10 μM)		3%	n.d.
cefotaxime sodium <sup>b</sup> (10 μM)		5%	n.d.
phenoxymethylpenicillinic acid <sup>b</sup> (10 μM)		-8%	n.d.
benzathine benzylpenicillin <sup>b</sup> (10 μM)		-3%	n.d.
flucloxacillin sodium <sup>b</sup> (10 μM)		-5%	n.d.
piperacillin <sup>b</sup> (10 μM)		-4%	n.d.

SUPPLEMENTARY TABLE 1 (continued)

pivampicillin <sup>b</sup> prodrug (10 μM)		-1%	n.d.
sultamicillin <sup>b</sup> prodrug (10 μM)		-6%	n.d.
sultamicillin tosylate <sup>b</sup> prodrug (10 μM)		1%	n.d.
ticarcillin monosodium <sup>b</sup> (10 μM)		-8%	n.d.
bacampicillin hydrochloride <sup>b</sup> prodrug (10 μM)		4%	n.d.
cloxacillin sodium <sup>b</sup> (10 μM)		0%	n.d.
penicillin G sodium salt <sup>b</sup> (10 μM)		-33%	n.d.
mecillinam <sup>b</sup> (10 μM)		-6%	n.d.

SUPPLEMENTARY TABLE 1 (continued)

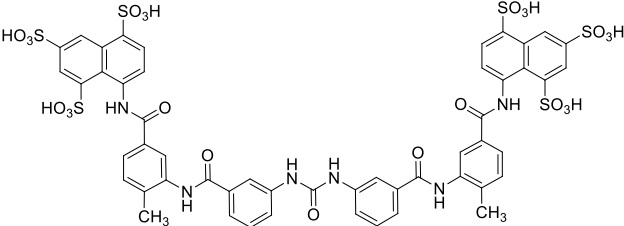
aztreonam <sup>b</sup> (10 μM)		-11%	n.d.
cefuroxime sodium salt <sup>b</sup> (10 μM)		-12%	n.d.
ceftaroline fosamil <sup>b</sup> (10 μM)		9%	n.d.
oxacillin sodium <sup>b</sup> (10 μM)		16%	n.d.
ertapenem sodium salt <sup>b</sup> (10 μM)		1%	n.d.
meropenem trihydrate <sup>b</sup> (10 μM)		-26%	n.d.
potassium clavulanate <sup>b</sup> (100 μM)		-23%	n.d.



SUPPLEMENTARY TABLE 1 (continued)

bacterial metabolites			
peptidoglycan pentapeptide <sup>a</sup> L-Ala-γ-D-Glu-L- Lys-D-Ala-D-Ala (100 μM)		51%	n.d.
lipid II <sup>a</sup> (100 μM)		67%	n.d.
UDP-D-MurNAc- pentapeptide <sup>a</sup> (100 μM)		14%	n.d.
D-Ala-D-Ala <sup>a</sup> (100 μM)		26%	n.d.
N <sup>ε</sup> ,N <sup>ε</sup> -diacetyl-L- Lys-D-Ala-D-Ala <sup>a</sup> (100 μM)		28%	n.d.

SUPPLEMENTARY TABLE 1 (continued)

		other compounds	
suramin <sup>b</sup> (10 µM)		-14%	75%

*Note.* Compounds were screened either at 10 µM or at 100 µM, depending on their water-solubility. Substrate (UDP-D-GlcNAc) concentration was 1 mM in all screening tests. Percent inhibition of CapD/CapE enzymatic activity at the screening concentration is provided; negative values indicate activation of enzymatic activity; n.d. = not determined. For potent inhibitors, full concentration–inhibition curves were recorded, and IC<sub>50</sub> values were determined (see main manuscript).

<sup>a</sup>compound screened in buffer without DMSO

<sup>b</sup>compound screened in buffer with 1% DMSO

SUPPLEMENTARY TABLE 2

Sequence set for phylogenetic analysis of the Sugp 4-keto reductase family

Accession number	Organism	Source
YP_001516388.1	<i>Acaryochloris marina</i> MBIC11017	OrtholugeDB
AFA48881.1	<i>Acetobacterium woodii</i> DSM 1030	BLASTp search CapN
YP_001235365.1 <sup>b</sup>	<i>Acidiphilium cryptum</i> JF-5	OrtholugeDB
EXS24708.1	<i>Acinetobacter baumannii</i> 573719	BLASTp search WbpK
EGJ67248.1	<i>A. baumannii</i> 6014059	BLASTp search WbpK
AGK44886.1	<i>A. baumannii</i> AB0057	OrtholugeDB
AEA00564.1	<i>Aerococcus urinae</i> ACS-120-V-Col10a	OrtholugeDB
AEB50812.1	<i>Aeromonas veronii</i> B565	OrtholugeDB
CAQ80715.1	<i>Aliivibrio salmonicida</i> LFI1238	OrtholugeDB
YP_003239525.1 <sup>a</sup>	<i>Ammonifex degensii</i> KC4	OrtholugeDB
ACL67752.1	<i>Anaeromyxobacter dehalogenans</i> 2CP-1	OrtholugeDB
ADG94190.1	<i>Arcobacter nitrofigilis</i> DSM 7299	OrtholugeDB
YP_157917.1	<i>Aromatoleum aromaticum</i> EbN1	OrtholugeDB
YP_003180026.1 <sup>a</sup>	<i>Atopobium parvulum</i> DSM 20469	OrtholugeDB
CAL96199.1	<i>Azoarcus</i> species BH72	OrtholugeDB
CBS87488.1	<i>Azospirillum</i> species B510	OrtholugeDB
YP_004096659.1 <sup>a</sup>	<i>Bacillus cellulosilyticus</i> DSM 2522	OrtholugeDB
NP_977659.1 <sup>a</sup>	<i>Bacillus cereus</i> ATCC 10987	OrtholugeDB
EEK97603.1	<i>Bacillus cereus</i> AH187	OrtholugeDB
YP_177182.1 <sup>a</sup>	<i>Bacillus clausii</i> KSM-K16	OrtholugeDB
AEH52402.1	<i>Bacillus coagulans</i> 2-6	OrtholugeDB

SUPPLEMENTARY TABLE 2 (continued)

BAB07434.1	<i>Bacillus halodurans</i> C-125	OrtholugeDB
ADI00528.1	<i>Bacillus selenitireducens</i> MLS10	OrtholugeDB
YP_003868079.1 <sup>a</sup>	<i>Bacillus subtilis</i> subspecies <i>spizizenii</i> strain W23	OrtholugeDB
NP_391661.3 <sup>a</sup>	<i>Bacillus subtilis</i> subspecies <i>subtilis</i> strain 168	OrtholugeDB
YP_003667391.1 <sup>a</sup>	<i>Bacillus thuringiensis</i> BMB171	OrtholugeDB
ABY46214.1	<i>Bacillus weihenstephanensis</i> KBAB4	OrtholugeDB
EDM19840.1	<i>Bacteroides caccae</i> ATCC 43185	BLASTp search CapN
CAH06763.1	<i>Bacteroides fragilis</i> NCTC 9343	OrtholugeDB
CAH07611.1	<i>B. fragilis</i> NCTC 9343	OrtholugeDB
YP_212215.1 <sup>a</sup>	<i>B. fragilis</i> NCTC 9343	OrtholugeDB
YP_211209.1 <sup>a</sup>	<i>B. fragilis</i> NCTC 9343	OrtholugeDB
YP_213273.1 <sup>a</sup>	<i>B. fragilis</i> NCTC 9343	OrtholugeDB
YP_100924.1 <sup>a</sup>	<i>B. fragilis</i> YCH46	OrtholugeDB
YP_098834.1 <sup>a</sup>	<i>B. fragilis</i> YCH46	OrtholugeDB
YP_099850.1 <sup>a</sup>	<i>B. fragilis</i> YCH46	OrtholugeDB
ADV44138.1	<i>Bacteroides helcogenes</i> P 36-108	OrtholugeDB
YP_004161543.1 <sup>a</sup>	<i>B. helcogenes</i> P 36-108	OrtholugeDB
YP_004161803.1 <sup>a</sup>	<i>B. helcogenes</i> P 36-108	OrtholugeDB
YP_004258251.1 <sup>a</sup>	<i>Bacteroides salanitronis</i> DSM 18170	OrtholugeDB
NP_811799.1 <sup>a</sup>	<i>Bacteroides thetaiotaomicron</i> VPI-5482	OrtholugeDB
YP_001300216.1 <sup>a</sup>	<i>Bacteroides vulgatus</i> ATCC 8482	OrtholugeDB
YP_001301242.1 <sup>a</sup>	<i>B. vulgatus</i> ATCC 8482	OrtholugeDB
YP_001633455.1	<i>Bordetella petrii</i> DSM 12804	OrtholugeDB
NP_772563.1	<i>Bradyrhizobium diazoefficiens</i> USDA 110	OrtholugeDB
YP_002732468.1 <sup>b</sup>	<i>Brucella melitensis</i> ATCC 23457	OrtholugeDB
YP_001627391.1 <sup>b</sup>	<i>Brucella suis</i> ATCC 23445	OrtholugeDB
AAU49962.1	<i>Burkholderia mallei</i> ATCC 23344	OrtholugeDB
ACQ98521.1	<i>Burkholderia pseudomallei</i> MSHR346	OrtholugeDB
YP_001681771.1 <sup>a</sup>	<i>Caulobacter</i> species K31	OrtholugeDB
YP_001983909.1	<i>Cellvibrio japonicus</i> Ueda107	OrtholugeDB
ABG64558.1 <sup>b</sup>	<i>Chelativorans</i> species BNC1	OrtholugeDB
YP_003124384.1 <sup>a</sup>	<i>Chitinophaga pinensis</i> DSM 2588	OrtholugeDB
YP_001943893.1 <sup>a</sup>	<i>Chlorobium limicola</i> DSM 245	OrtholugeDB
ABB24695.1	<i>Chlorobium luteolum</i> DSM 273	OrtholugeDB
ABL66603.1	<i>Chlorobium phaeobacteroides</i> DSM 266	OrtholugeDB
ABX43855.1	<i>Clostridium phytofermentans</i> ISDg	OrtholugeDB
EMS70345.1	<i>Clostridium termitidis</i> CT1112	BLASTp search CapN
AAZ28046.1	<i>Colwellia psychrerythraea</i> 34H	OrtholugeDB
YP_003549169.1 <sup>a</sup>	<i>Coralimargarita akajimensis</i> DSM 45221	OrtholugeDB
BAC17172.1	<i>Corynebacterium efficiens</i> YS-314	OrtholugeDB

SUPPLEMENTARY TABLE 2 (continued)

AFZ53196.1	<i>Cyanobacterium aponinum</i> PCC 10605	BLASTp search WbpK
AFT68008.1	<i>Cycloclasticus</i> species P1	BLASTp search WbpK
YP_003808718.1	<i>Desulfarculus baarsii</i> DSM 2075	OrtholugeDB
YP_002431668.1 <sup>a</sup>	<i>Desulfatibacillum alkenivorans</i> AK-01	OrtholugeDB
YP_002460906.1 <sup>a</sup>	<i>Desulfitobacterium hafniense</i> DCB-2	OrtholugeDB
ADW19313.1	<i>Desulfobulbus propionicus</i> DSM 2032	OrtholugeDB
ABW67105.1	<i>Desulfococcus oleovorans</i> Hxd3	OrtholugeDB
YP_003159756.1	<i>Desulfomicrobium baculatum</i> DSM 4028	OrtholugeDB
YP_004498346.1 <sup>a</sup>	<i>Desulfotomaculum carboxydivorans</i> CO-1-SRB	OrtholugeDB
ADH86493.1	<i>Desulfurivibrio alkaliphilus</i> AHT2	OrtholugeDB
ACT08204.1	<i>Dickeya zeae</i> Ech1591	OrtholugeDB
YP_002250211.1 <sup>a</sup>	<i>Dictyoglomus thermophilum</i> H-6-12	OrtholugeDB
YP_002932741.1 <sup>a</sup>	<i>Edwardsiella ictaluri</i> 93-146	OrtholugeDB
EZR14264.1	<i>Enterobacter</i> species BWH 27	BLASTp search WbpK
NP_815853.1 <sup>a</sup>	<i>Enterococcus faecalis</i> V583	OrtholugeDB
AGE29663.1	<i>Enterococcus faecium</i> NRRL B-2354	BLASTp search CapN
AFC91450.1	<i>Escherichia coli</i> O45:H10 strain H61	BLASTp search WbpK
YP_002271284.1 <sup>a</sup>	<i>E. coli</i> O157:H7 strain EC4115	OrtholugeDB
ADO36951.1	<i>Eubacterium limosum</i> KIST612	OrtholugeDB
ACB62080.1	<i>Exiguobacterium sibiricum</i> 255-15	OrtholugeDB
ADL24802.1	<i>Fibrobacter succinogenes</i> S85	OrtholugeDB
ACU08375.1	<i>Flavobacteriaceae bacterium</i> 3519-10	OrtholugeDB
ABQ03373.1	<i>Flavobacterium johnsoniae</i> UW101	OrtholugeDB
ABZ87470.1	<i>Francisella philomiragia</i> subspecies <i>philomiragia</i> ATCC 25017	OrtholugeDB
CAG46095.1	<i>Francisella tularensis</i> subspecies <i>tularensis</i> SCHU S4	reference sequence WbcT
EYD70022.1	<i>Fusobacterium necrophorum</i> subspecies <i>funduliforme</i> B35	BLASTp search WbcT
AAL93809.1	<i>Fusobacterium nucleatum</i> ATCC 25586	OrtholugeDB
EHO82815.1	<i>Fusobacterium ulcerans</i> 12-1B	BLASTp search WbcT
ADL56823.1	<i>Gallionella capsiferiformans</i> ES-2	OrtholugeDB
ACH38643.2	<i>Geobacter bemidjiensis</i> Bem	OrtholugeDB
ACT18628.1	<i>Geobacter</i> species M21	OrtholugeDB
AGM38884.1	<i>Haemophilus parasuis</i> 84-15995	BLASTp search CapN
ADJ62001.1	<i>Herbaspirillum seropedicae</i> SmR1	OrtholugeDB
CAL61328.1	<i>Herminiimonas arsenicoxydans</i>	OrtholugeDB
YP_003756842.1 <sup>b</sup>	<i>Hyphomicrobium denitrificans</i> ATCC 51888	OrtholugeDB
AAV81403.1	<i>Idiomarina loihiensis</i> L2TR	OrtholugeDB
CDG81600.1	<i>Janthinobacterium</i> species Marseille	OrtholugeDB

SUPPLEMENTARY TABLE 2 (continued)

EJN56509.1	<i>Lactobacillus coryniformis</i> CECT 5711	BLASTp search CapN
YP_001271650.1 <sup>a</sup>	<i>Lactobacillus reuteri</i> DSM 20016	OrtholugeDB
YP_003352668.1 <sup>a</sup>	<i>Lactococcus lactis</i> KF147	OrtholugeDB
YP_003455544.1	<i>Legionella longbeachae</i> NSW150	OrtholugeDB
YP_094797.1	<i>Legionella pneumophila</i> strain Philadelphia 1	OrtholugeDB
EUJ27787.1	<i>Listeria grayi</i> FSL F6-1183	BLASTp search CapN
BAE48842.1	<i>Magnetospirillum magneticum</i> AMB-1	OrtholugeDB
ABM19688.1	<i>Marinobacter aquaeolei</i> VT8	OrtholugeDB
ABR69745.1	<i>Marinomonas</i> species MWYL1	OrtholugeDB
BAK21004.1	<i>Melissococcus plutonius</i> ATCC 35311	OrtholugeDB
ABE49544.1	<i>Methylobacillus flagellatus</i> KT	OrtholugeDB
YP_001639501.1 <sup>b</sup>	<i>Methylobacterium extorquens</i> PA1	OrtholugeDB
AAU93112.1	<i>Methylococcus capsulatus</i> strain Bath	OrtholugeDB
AEF99624.1	<i>Methylomonas methanica</i> MC09	OrtholugeDB
EEF78408.1	<i>Methylophaga thiooxydans</i> DMS010	BLASTp search WbpK
ACT47946.1	<i>Methylotenera mobilis</i> JLW8	OrtholugeDB
ADI29637.1	<i>Methylotenera versatilis</i> 301	OrtholugeDB
YP_318993.1	<i>Nitrobacter winogradskyi</i> Nb-255	OrtholugeDB
YP_344157.1	<i>Nitrosococcus oceani</i> ATCC 19707	OrtholugeDB
CAD84411.1	<i>Nitrosomonas europaea</i> ATCC 19718	OrtholugeDB
EAR60515.1	<i>Oceanospirillum</i> species MED92	BLASTp search WbpV
YP_001371115.1 <sup>b</sup>	<i>Ochrobactrum anthropi</i> ATCC 49188	OrtholugeDB
YP_004254159.1 <sup>a</sup>	<i>Odoribacter splanchnicus</i> DSM 220712	OrtholugeDB
YP_810971.1 <sup>a</sup>	<i>Oenococcus oeni</i> PSU-1	OrtholugeDB
YP_003949006.1 <sup>a</sup>	<i>Paenibacillus polymyxa</i> SC2	OrtholugeDB
YP_004275600.1 <sup>a</sup>	<i>Pedobacter saltans</i> DSM 12145	OrtholugeDB
ACF42880.1	<i>Pelodictyon phaeoclathratiforme</i> BU-1	OrtholugeDB
CAG21077.1	<i>Photobacterium profundum</i> SS9	OrtholugeDB
CAQ86427.1	<i>Photorhabdus asymbiotica</i> ATCC43949	OrtholugeDB
ABM38779.1	<i>Polaromonas naphthalenivorans</i> CJ2	OrtholugeDB
AEA21732.1	<i>Prevotella denticola</i> F0289	OrtholugeDB
YP_003575047.1 <sup>a</sup>	<i>Prevotella ruminicola</i> 23	OrtholugeDB
YP_002016412.1 <sup>a</sup>	<i>Prosthecochloris aestuarii</i> DSM 271	OrtholugeDB
AAC45865.1	<i>Pseudomonas aeruginosa</i> O5 PAO1	reference sequence WbpK
AAB39483.1	<i>P. aeruginosa</i> O11 PA103	reference sequence WbjF
AAF23991.1	<i>P. aeruginosa</i> IATS O6	reference sequence WbpV
EOT14879.1	<i>P. aeruginosa</i> O10 PA14	OrtholugeDB
ABR86892.1	<i>P. aeruginosa</i> O12 PA7	OrtholugeDB

SUPPLEMENTARY TABLE 2 (continued)

ABA75791.1	<i>Pseudomonas fluorescens</i> Pf0-1	OrtholugeDB
AAY96279.1	<i>Pseudomonas protegens</i> Pf-5	OrtholugeDB
ABQ80062.1	<i>Pseudomonas putida</i> F1	OrtholugeDB
EIK67266.1	<i>Pseudomonas synxantha</i> BG33R	BLASTp search WbpK
AAY38670.1	<i>Pseudomonas syringae</i> pathovar <i>syringae</i> B728a	OrtholugeDB
ABY23173.1	<i>Renibacterium salmoninarum</i> ATCC 33209	OrtholugeDB
AAQ93037.1	<i>Rhizobium etli</i> CFN 42	reference sequence WreQ
ACS57111.1	<i>Rhizobium leguminosarum</i> biovar <i>trifolii</i> WSM1325	OrtholugeDB
YP_003578087.1 <sup>a</sup>	<i>Rhodobacter capsulatus</i> SB 1003	OrtholugeDB
EXM38188.1	<i>Ruminococcus albus</i> SY3	BLASTp search CapN
AHG21859.1	<i>Serratia fonticola</i> RB-25	BLASTp search WbpK
YP_004501405.1 <sup>a</sup>	<i>Serratia</i> species AS12	OrtholugeDB
ACK45991.1	<i>Shewanella baltica</i> OS223	OrtholugeDB
AAN56173.1	<i>Shewanella oneidensis</i> MR-1	OrtholugeDB
ABP76250.1	<i>Shewanella putrefaciens</i> CN-32	OrtholugeDB
YP_001879836.1 <sup>a</sup>	<i>Shigella boydii</i> CDC 3083-94	OrtholugeDB
ADE13107.1	<i>Sideroxydans lithotrophicus</i> ES-1	OrtholugeDB
ADY12094.1	<i>Sphaerochaeta globosa</i> strain Buddy	OrtholugeDB
ADZ78392.1	<i>Sphingobacterium</i> species 21	OrtholugeDB
YP_616633.1 <sup>b</sup>	<i>Sphingopyxis alaskensis</i> RB2256	OrtholugeDB
BAB94002.1	<i>Staphylococcus aureus</i> serotype 8 MW2	OrtholugeDB
BAF66380.1	<i>S. aureus</i> serotype 5 Newman	reference sequence CapN
AAL26668.1	<i>S. aureus</i> strain M	BLASTp search CapN
ADC86484.1	<i>Staphylococcus lugudensis</i> HKU09-01	OrtholugeDB
YP_330067.1 <sup>a</sup>	<i>Streptococcus agalactiae</i> A909	OrtholugeDB
AGM98615.1	<i>Streptococcus iniae</i> SF1	BLASTp search CapN
YP_003485108.1 <sup>a</sup>	<i>Streptococcus mutans</i> NN2025	OrtholugeDB
ABN45584.1	<i>Streptococcus sanguinis</i> SK36	OrtholugeDB
YP_004401334.1 <sup>a</sup>	<i>Streptococcus suis</i> ST3	OrtholugeDB
AGS58366.1	<i>Streptococcus suis</i> 14A	BLASTp search CapN
ADN09535.1	<i>Sulfurimonas autotrophica</i> DSM 16294	OrtholugeDB
ACZ12787.1	<i>Sulfurospirillum deleyianum</i> DSM 6946	OrtholugeDB
NP_896554.1 <sup>b</sup>	<i>Synechococcus</i> species WH 8102	OrtholugeDB
BAK93722.1	<i>Tetragenococcus halophilus</i> NBRC 12172	BLASTp search CapN
ADC72044.1	<i>Thioalkalivibrio</i> species K90mix	OrtholugeDB
AAZ97826.1	<i>Thiobacillus denitrificans</i> ATCC 25259	OrtholugeDB
EKF08323.1	<i>Thalassospira profundimaris</i> WP0211	BLASTp search WbpV
AGT34472.1	<i>Thermofilum</i> species 1910b	outgroup
ADG90620.1	<i>Thermosphaera aggregans</i> DSM 11486	outgroup

SUPPLEMENTARY TABLE 2 (continued)

NP_972045.1 <sup>a</sup>	<i>Treponema denticola</i> ATCC 35405	OrtholugeDB
ADU38842.1	<i>Variovorax paradoxus</i> EPS	OrtholugeDB
AEH32018.1	<i>Vibrio anguillarum</i> 775	OrtholugeDB
AGA82312.1	<i>Vibrio cholerae</i> O1 biovar El Tor strain N16961	OrtholugeDB
AAM22609.1	<i>V. cholerae</i> O37 1322-69	BLASTp search WbpV
AAC46247.1	<i>V. cholerae</i> O139 AI1837	BLASTp search WbpV
ACH64821.1	<i>Vibrio fischeri</i> MJ11	OrtholugeDB
CCO59607.1	<i>Vibrio nigripulchritudo</i> SFn1	BLASTp search WbpK
EGF42689.1	<i>Vibrio parahaemolyticus</i> 10329	BLASTp search WbpK
WP_017074481.1	<i>Vibrio splendidus</i> ZS-139	BLASTp search WbpK
AAO09281.1	<i>Vibrio vulnificus</i> CMCP6	OrtholugeDB
ABD38630.1	<i>V. vulnificus</i> MO6-24/O	OrtholugeDB
YP_001418446.1 <sup>a</sup>	<i>Xanthobacter autotrophicus</i> Py2	OrtholugeDB
YP_001721901.1 <sup>a</sup>	<i>Yersinia pseudotuberculosis</i> YPIII	OrtholugeDB
EHM45344.1	<i>Yokenella regensburgei</i> ATCC 43003	BLASTp search WbpK

*Note.* Accession numbers are from GenBank, or from the NCBI non-redundant protein database. A grey background indicates protein sequences used for assembly of the final phylogenetic tree of the Supg 4-keto reductase family (see main text).

<sup>a</sup>putative dTDP-4-dehydrorhamnose reductase or C<sub>55</sub>PP-GlcNAc epimerase

<sup>b</sup>sequence producing low alignment score with domain models cd05232 and COG0451

WbpK	1	-----MKAVMVTGASGFVGSALC-CELARTGYAVIAVRRVVERIPS--VTYIEADL	167	PLIYANDAPGNFGRLLKLVASGLPLPLDGV-RNARSLVSRNIVGFLSLCAEHPDAAGEL
WreQ	1	-----MRCLVTGAAGFVGSPLVKRLHAEKIYDLVATTRSQTPAFPPEVAHFP-IEI	167	PLVYGPARGNFALLVNLVRRKLLPLPFASL-KNHRTLVAVQNLVDLIIACATHPAAPGEI
WbpV	1	-----MTRHNVLVTGATGFIGAALVNSLSSGQYKVVAGCRRRGGAWPRGVTPLLLGEL	172	VLVYGPVGKANVQTMWRWLKRGVPLPLGAI-HNRRSLVSLDNLVDLIIITCIEHPAAVQV
WbfT	1	MCTGDRKMPKSILLTGSTGFVGTNLVKSLTLKSDYIVKSAVRHAVNK--DDGLLFEVVDI	177	TIVYGPVGKANFASLMRLVSKGIPLPFGSITQNKRSLSVINNLVDLIVTCIDHPKAANQV
WbtC	1	-----MKKRILVTGLSSYIGNSFAAKYNSD-----FS---IDKISL	151	PMVYGEKSGKNYPKLVKLA--KYTFIFPNI-NNQRSVISIDNLSKEIAEIIQLTK--HGV
CpsG	1	-----MKKVLITGANSYIGTSLEKWLQQS-----EEQYHVDTLDM	148	PMVYGPQATGNYSRLSKLS--KFTPIFPKV-ANKRSMIYLDNLLFVRLSIETEL--SGI
CapN	1	-----MRKNILITGVHGYIGNALKDKLIEQ-----GHQVDQINV	152	PMIYGAKCPGNFQRLMQLS--KRLPIIPNI-NNQRSALYIKHLTAFIDQLISLEV--TGV
Cap1N	1	-----MKNILITGKNGFVGNQFQLLLNND-----NYKVDRVSL	151	PMIYGKKGPNFAKLIKVA--NLSRIFPKY-HNERSVIHIDNLYKHILHLLKNEK--SDI
		::** .:* :		::* . * : . : * * : : . : :
WbpK	50	TDPATFAGEFPTVDCIIHLAGRAHILTDKVADPLAAFREVNRDATVRLATRALEAGVKRF	226	FLVADGEDVSIAQMIEA-LSRGMGRRPALFTFP-AVLLKLVMLLGGKSMHEQLCGSLQV
WreQ	51	TGGTDWTAAL EGVQVIVHLAARVHIMNDRAADPLAEFRRTNTAAALNLAEQAASAGVKRF	226	FLAGDGEDLSTPALIRG-IAAGLVKPMPLVFPF-PALLQMAAKALGKEAVYQRLCGSLQV
WbpV	55	GSSVWDAE-SAIDTVVHCAARVHVMSSETASDPLVEFRKANVQGTLDLAREAVSRGVRRF	231	FLVSDGEDLSTTELLRR-MGRALGAPARLLPVP-ASWIGAAAKVLNRQAFARRLCGSLQV
WbfT	59	NASTDFELPKNTTVVHCAARAHVMDKAEPLTLYREVNTAGTVNLAKQAIDSGVKRF	237	FLVSDGHDVSTAEMVRE-LAIALDKPTWQLPVP-IWCYKLFGKLFGKSDIVDRLTGTLQV
WbtC	34	RDVSWANIDLSGYDAVLHVAGIAH--TSKDPKLEKEYKINTQLTYDLAKQAKDQGVRF	206	FLLQDNEYFCTSQFIKRYRQDVLGKRTYLTIF-NPII---RLLAKKVDFINKVFGNLT
CpsG	36	IDPNWKTDFSPYDSIFHVAIVH--KNEKQMNSDLYEVNTKLPTELATIAKHSGLRQF	203	HFPQNKDYVTTSQLVNV-IRQVNGKSTLLTSLF-NPII---KSL-KGFSQINKLFGNLVY
CapN	35	RNQLWKSTSFKDYDVLIIHTAALVH--NNSPQARLSDYMQVNMLLTKQLAKAKAEDVKQF	207	YHPQDSFYFDTSVMEY-IRRQSHRKTVLINM--PSML---NKYFNKLSVFRKLFGNLIY
Cap1N	34	KNNDWKFCSEFENYDVIHHLAALVH--NNQPNAKIIDYMNNTNYHLTRELAKKAKDDGVSQF	206	THPQNMMEYMNNTALSL-IRNHLGKSSELIETIPVNFII---NKIIEKSNIANKIYGNLT
		::* * . * . : . * ** * : : *		: . : . . : * * : : * *
WbpK	110	VFVSSIGVNGNSTRQ--QAFNEDSPAGPHAPYAIISKYEAQEGLTLRGGKGMELVVVRP	284	DASKARRLLGWVPVETIGAGLQAAGREYILRQRERRK
WreQ	111	VFVSTIKVNGEEN--D--RPFRRHDDRPKPIDPYGISKLECEIGLREIAARTGMEVVIIRP	284	DITRARDVLGWSPVTPREGLKLAVE-----
WbpV	114	FISSIKVNGEGTEPG--RPYTADSPNPVDPYGVSKREAEQALLDLAEETGLEVVIIRP	289	DIMKTRQVLGWTPPVGVDQALEKTARSLDRQ----
WbfT	119	FISSIKVNGEGTLVG--CPFKTEDNHAPEDDYGLSKSEAEKQLVALAKDSSMEVVIIRP	295	DISHTKETLGWKPPQTLQEGFKQTAQAFQANNR---
WbtC	92	VFLSSIIIVYGDSAPIGQQKVIITKYTEPKPDDFYGDSKLQTEIKLNSLASD-DFNIAIIRP	262	EK-----
CpsG	94	IFLSSMSVYGNDE----EITRETRENPSYGYKSKLAAEIGLKDQLSD-SFKVILIRP	257	SKEMSQEAFDYN--VT--GFEESIRISERNNEKI--
CapN	93	IFMSTMAVYGKEGHVGSQDQVDTQTPMNPNTNYGISKKFAEQALQELISD-SFKVAIVRP	261	SNTLYENNALE--IIPGKMSLVIADIMDETTTKDKA
Cap1N	92	IFFSTMGIFGMNGLINKKCEITQLTPYKPKSAYDYSKLLAERDIQKLIDD-KFVVNIVRP	262	SKNIDDR-----EINNMHFDNFNQTISKTLK
		:* . * : : * * * * * : : : : * *		

SUPPLEMENTARY FIGURE 1. Multiple alignment of *Sugp* 4-keto reductase sequences of *Proteobacteria* and *Firmicutes* (computed by Clustal Omega version 1.2.1). The residues of the active site tetrad (N[ST]YK) are marked by yellow stars. Potential cofactor-binding sites are shaded grey. An \* (asterisk) indicates positions which have a single, fully conserved residue. A : (colon) indicates conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet PAM 250 matrix). A . (period) indicates conservation between groups of weakly similar properties (scoring ≤ 0.5).



## 7 Publications

### Several parts of this thesis were published in the following article:

Li W.,\* **Ulm H.**\* (\*contributed equally), Rausch M., Li X., O'Riordan K., Lee J.C., Schneider T., and Müller C.E. Analysis of the *Staphylococcus aureus* capsule biosynthesis pathway *in vitro*: characterization of the UDP-GlcNAc C6 dehydratases CapD and CapE and identification of enzyme inhibitors. *Int. J. Med. Microbiol. IJMM* **304**, 958–969 (2014).

### Further publications:

**Ulm H.**,\* Wilmes M.\* (\*contributed equally), Shai Y., and Sahl H.G. Antimicrobial host defensins – specific antibiotic activities and innate defense modulation. *Front. Immunol.* **3**, 249 (2012)

Müller A.,\* **Ulm H.**\* (\*contributed equally), Reder-Christ K., Sahl H.G., and Schneider, T. Interaction of type A lantibiotics with undecaprenol-bound cell envelope precursors. *Microb. Drug Resist.* **18**, 261–270 (2012).

Arnusch C.J., **Ulm H.**, Josten M., Shadkchan Y., Oshero N., Sahl H.G., and Shai Y. Ultrashort peptide bioconjugates are exclusively antifungal agents and synergize with cyclodextrin and amphotericin B. *Antimicrob Agents Chemother.* **56**, 1–9 (2012)

Hardt P., Engels I., Gajdiss M., **Ulm H.**, Sass P., Ohlsen K., Sahl H.G., Bierbaum G., Schneider T., and Grein F. The Ser/Thr kinase PknB crosstalks with the essential WalKR two-component system in *S. aureus* and interacts with the cell wall precursor lipid II. *PLoS pathog.* (submitted)

Rausch M.\*, **Ulm H.**\* (\*contributed equally), Li W., Hardt P., Sylvester M., Josten M., Sahl H.G., Müller C.E., Lee J.C., and Schneider T. Analysis of the *Staphylococcus aureus* capsule biosynthesis pathway *in vitro* – role of the CapAB kinase complex in pathway regulation. (in preparation)

