bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Bacterial glycoengineering as a biosynthetic route to customized glycomolecules**

2

3 Laura E. Yates, Dominic C. Mills and Matthew P. DeLisa*

4

5 School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, New York, USA

6

7 *Address correspondence to: Matthew P. DeLisa, School of Chemical and Biomolecular
8 Engineering, Cornell University, Ithaca, NY 14853. Tel: 607-254-8560; Fax: 607-255-9166;
9 Email: md255@cornell.edu

10

11 **Abstract**

12 Bacteria have garnered increased interest in recent years as a platform for the biosynthesis of a
13 variety of glycomolecules such as soluble oligosaccharides, surface-exposed carbohydrates
14 and glycoproteins. The ability to flexibly engineer commonly used laboratory species such as
15 *Escherichia coli* to efficiently synthesize non-native sugar structures by recombinant expression
16 of enzymes from various carbohydrate biosynthesis pathways has allowed for the facile
17 generation of important products such as conjugate vaccines, glycosylated outer membrane
18 vesicles, and a variety of other research reagents for studying and understanding the role of
19 glycans in living systems. This chapter highlights some of the key discoveries and technologies
20 for equipping bacteria with the requisite biosynthetic machinery to generate such products. As
21 the bacterial glyco-toolbox continues to grow, these technologies are expected to expand the
22 range of glycomolecules produced recombinantly in bacterial systems, thereby opening up this
23 platform to an even larger number of applications.

24

25

26

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

27 **Keywords**

28 Bacterial oligosaccharyltransferase, bacterial protein glycosylation, bacterial polysaccharides,
29 carbohydrate biosynthesis pathways, conjugate vaccines, glycoengineering, glycosyltransferase

30

31 **Abbreviations**

32 CPS: Capsular polysaccharide. LPS: lipopolysaccharide. LOS: lipo-oligosaccharide. ECA:
33 enterobacterial common antigen. S-layer: surface layer. ABC-transporter: ATP-binding cassette
34 transporter. Und-PP: Undecaprenyl pyrophosphate. HA: Hyaluronic acid. Gb_3:
35 Globotriaosylceramide. STX: Shiga toxin Gal: Galactose. Glc: Glucose STEC: Shigatoxin
36 producing *Escherichia coli.* Gb_4 Globotetraosylceramide. GM: monosialotetrahexosylganglioside
37 NeuNAc: *N*-acetylneuraminic acid. GlcNAc: *N*-acetylglucosamine. GalNAc: *N*-
38 acetylgalactosamine Le^X: Lewis X antigen. Le^y: Lewis Y antigen. LacNAc: *N*-acetyllactosamine
39 PolySia: Polysialic acid. NCAM: Neural cell adhesion molecule. T-antigen: Thomsen-
40 Friedenreich antigen Hib: *Haemophilus influenza* type b. hGH: human growth hormone PEG:
41 Polyethylene glycol. diNAcBac: Bacillosamine. EPA: Exotoxin A from *Pseudomonas*
42 *aeruginosa.* MBP: Maltose binding protein. IgG: Immunoglobulin G. Man: Mannose EPO:
43 erythropoietin. OMV: Outer membrane vesicle.

44

45

46

47

48

49

50

51

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

52 **Bacteria as a platform for polysaccharide and glycoconjugate production.**

53 In recent years, there has been growing interest in developing bacterial species as hosts for
54 glycoengineering applications involving the biosynthesis of structurally diverse polysaccharides,
55 which can be produced as free glycans or as conjugates to lipids or proteins. The most obvious
56 advantage of this approach is the vastly simpler and cheaper culturing conditions required for
57 maintenance of bacterial cells when compared to a eukaryotic cell culture. However, bacteria
58 are in fact highly proficient producers of carbohydrates, with more than 140 unique
59 monosaccharide base types identified in bacterial species, in contrast to the 14 base types
60 produced by mammalian species (Herget et al. 2008). Many of these bacterial monosaccharides
61 are then assembled into an even more diverse array of polysaccharides, often as part of surface
62 structures such as capsular polysaccharide (CPS) and the O-antigen component of
63 lipopolysaccharide (LPS), which are often important virulence factors in pathogenic species. In
64 *Escherichia coli* alone, 187 unique O-antigen structures and 80 CPS structures have been
65 identified to date (Stenutz et al. 2006; Senchenkova et al. 2016; Whitfield 2006). Other bacterial
66 polysaccharides have important structural functions (e.g., peptidoglycan), or play a role in
67 adaptation to environmental conditions by mechanisms such as osmoregulation (e.g.,
68 enterobacterial common antigen, ECA) (Kuhn et al. 1988).

69 The pathways responsible for production of mono- and polysaccharides are frequently
70 well-defined in bacteria, especially in commonly used host species such as *E. coli* (Aoki-
71 Kinoshita and Kanehisa 2015). Furthermore, with the exception of the ubiquitous structural
72 polysaccharide peptidoglycan, bacterial polysaccharides are typically non-essential for viability,
73 meaning biosynthesis pathways are amenable to genetic manipulation and deletion. For
74 example, metabolic engineering studies have identified routes to enhance the availability of
75 relevant nucleotide-activated sugars, leading to improved polysaccharide yields (Ruffing and

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

76 Chen 2006). As a result of these and other related efforts, bacteria represent a tractable, well-
77 defined platform for engineering the biosynthesis of polysaccharides.

78 While the ability of bacteria to produce polysaccharides and glycolipids is established, it
79 was long believed that bacteria were incapable of modifying proteins with carbohydrate
80 moieties, a process known as glycosylation. However, this paradigm was overturned in the
81 1970s with the identification of glycosylated surface layer (S-layer) proteins in *Halobacterium*
82 *salinarum*, *Clostridium thermosaccharolyticum* and *Clostridium thermohydrosulfuricum* (Sleytr
83 1975; Sleytr and Thorne 1976). Although examples of bacterial protein glycosylation remain
84 relatively uncommon, in the past 15 years a diverse array of systems have been discovered and
85 characterized, including examples of sequential and *en bloc* transfer of both *N*-linked and *O*-
86 linked glycans (Szymanski et al. 1999; Castric 1995; Grass et al. 2003; Thibault et al. 2001).

87 From an engineering perspective, perhaps the most significant advance came in 2002
88 with the functional transfer of a complete protein *N*-glycosylation system from the
89 gastrointestinal pathogen *Campylobacter jejuni* into a laboratory strain of *E. coli*, which is
90 naturally incapable of protein glycosylation (Wacker et al. 2002). The versatility of this system
91 was further enhanced by a series of experiments demonstrating the modularity of the bacterial
92 glycosylation machinery, which was found to tolerate a number of different glycan structures
93 and protein substrates (Feldman et al. 2005; Kowarik et al. 2006; Wacker et al. 2006).
94 Importantly, the newfound ability to generate glycoproteins in a genetically tractable host
95 organism like *E. coli* provided a unique opportunity to both understand and exploit the
96 glycosylation process in ways that were not previously possible with eukaryotic systems. This is
97 because even though the pathways involved in the production of protein-linked polysaccharides
98 in eukaryotic cells are well understood, the essential nature of many of these mechanisms limits
99 the potential for manipulation.

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

100 **Polysaccharide production in bacteria.** Enzymatic synthesis of polysaccharides utilizes
101 nucleotide-activated sugars as glycosyl donors, to supply the necessary energy for the reaction.
102 In bacteria, these nucleotide sugars are typically only present in the cytoplasm where they are
103 synthesized. Consequently, all initial polysaccharide biosynthesis in bacteria also takes place
104 within the cytoplasm. The majority of polysaccharides are synthesized by one of three
105 pathways: the Wzy-dependent pathway, the ATP-binding cassette (ABC) transporter-dependent
106 pathway, and the synthase-dependent pathway (**Fig. 1**), although shorter oligosaccharides may
107 be formed by the direct action of glycosyltransferases on a substrate such as lipid A in the case
108 of the LPS core or lipooligosaccharides (LOS) (Kalynych et al. 2014). The Wzy-dependent
109 pathway involves the sequential action of glycosyltransferases on a lipid anchor, undecaprenyl
110 diphosphate (Und-PP), on the inner leaflet of the cytoplasmic membrane, followed by
111 translocation of a completed subunit across the membrane by the flippase Wzx. The subunits
112 then undergo polymerization by the polymerase Wzy. The number of repeat units is modulated
113 somewhat by Wzz, the chain-length regulator, although the resulting polymers are not strictly
114 uniform in length. Completed polysaccharides are then removed from Und-PP and transferred
115 to a target location, which differs depending on the species in question and the type of
116 polysaccharide produced (Raetz and Whitfield 2002). Common examples of polysaccharides
117 produced by this mechanism include the majority of O-antigen polysaccharides, and a
118 significant proportion of capsules, as well as specific examples such as ECA, a surface
119 polysaccharide common to most *Enterobacteriacae*, but limited to this family (Kuhn et al. 1988).
120 In contrast, the ABC transporter-dependent pathway involves the assembly of the entire
121 polysaccharide on a lipid anchor at the inner face of the cytoplasmic membrane, before the
122 chain is capped to indicate completion, and the entire structure is transported across the
123 membrane by the ABC-transporter complex (Cuthbertson et al. 2010). As with Wzy-dependent
124 systems, however, the polysaccharide is then removed from the lipid anchor and transferred to

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

125 a permanent point of attachment. Polysaccharides assembled by this method typically form O-
126 antigen polysaccharides or capsules. Synthase-dependent polysaccharide assembly is unique
127 in that it can occur in the presence or absence of a lipid anchor. A transmembrane
128 glycosyltransferase simultaneously catalyzes formation of the polymer and translocation across
129 the membrane (Whitney and Howell 2013). Polysaccharides produced by this mechanism may
130 be attached to the exterior of the cell, but are more frequently released into the extracellular
131 environment to form non-covalently associated exopolysaccharides such as hyaluronic acid
132 (HA), alginate or cellulose.

133 **Bioengineering of secreted oligosaccharides in bacteria.** Small, soluble oligosaccharides
134 play many important roles in biological systems, and as such have a multitude of potential uses
135 in research, medicine and industry. However, owing to the extremely high heterogeneity of such
136 structures, together with low yield and complex purification when isolating from natural sources,
137 engineered production has been the focus of much research. Chemical synthesis is complex
138 and costly, and the resulting oligosaccharides are subject to the same issues regarding
139 heterogeneity, limiting their usefulness without significant downstream purification. Chemo- and
140 in vitro-enzymatic methods have also been widely explored, and have shown great
141 improvements with respect to yield and structural homogeneity, but isolation of the required
142 enzymes is a demanding process, and the necessary nucleotide-activated sugars are extremely
143 expensive to supply for such large-scale synthesis; consequently production beyond the
144 milligram scale, especially for larger tri- and tetrasaccharides remains unfeasible by this
145 method.

146 The development of a metabolically-engineered *E. coli* strain that could produce human
147 milk oligosaccharides in a fermentation process represented a significant advance within the
148 field (Priem et al. 2002). The engineered strain utilizes glycerol as an affordable carbon source,
149 relying on native metabolic pathways within the bacterium to produce a continuous supply of the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

150 required nucleotide sugars. The approach also relies on the presence of a soluble acceptor
151 sugar in the cytoplasm as an assembly platform. In this case lactose, which can be imported
152 from the growth medium, was used. However, methods for the in situ synthesis of acceptor
153 sugars have also been developed (Samain et al. 1997). Such engineered strains have been
154 shown to produce quantities of up to 34 g/L of secreted oligosaccharide, and the scalable nature
155 of production means the manufacture of kilogram quantities of sugar are entirely feasible
156 (Drouillard et al. 2010). This approach has since been used for the production of more than 25
157 different oligosaccharides ranging from disaccharides to pentasaccharides, including structures
158 that are known to have immunomodulatory effects or to be associated with cancer in humans
159 (Ruffing and Chen 2006).
160 **Bioengineering of exopolysaccharides in bacteria.** Many exopolysaccharides produced by
161 bacteria have significant commercial value (Schmid et al. 2015), the most widely studied of
162 which are listed in **Table 1**. Some of these polymers are naturally-occurring in bacteria, while
163 others have been engineered via heterologous gene expression, particularly in cases where the
164 original source or isolation method was undesirable. One example is HA, an extremely
165 hydrophilic polymer of alternating β-D-glucuronic acid and β-D-*N*-acetyl-glucosamine residues
166 that is a desirable material in medicine and cosmetics owing to its high water retention capacity
167 and lack of toxicity. Initially, this polysaccharide was purified from rooster combs, although the
168 majority of production is now achieved via microbial fermentation (Liu et al. 2011). Native
169 bacterial production of HA was first achieved from *Streptococcus zooepidemicus* (Thonard et al.
170 1964), but due to co-production of the streptolysin exotoxin, recombinant production remained a
171 priority. Indeed, recombinant HA was eventually achieved using the host organism *Bacillus*
172 *subtilis* (Widner et al. 2005), and subsequently *E. coli* (Yu and Stephanopoulos 2008). Such
173 approaches achieve yields of ~10 g/L, which is thought to be near the production limit owing to
174 the effect of the exopolysaccharide on the viscosity of the growth medium (Yu and

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

175 Stephanopoulos 2008). Key advances have come instead in the area of polymer length
176 regulation, allowing for better control of physiochemical properties, and achieved largely through
177 metabolic engineering and tighter control of the availability of the precursor nucleotide sugars
178 (Jia et al. 2013).

179 In other cases, such as the commercially valuable xanthan, metabolic engineering has
180 enabled yields of up to 50 g/L, also thought likely to be the highest level feasible for bioreactor
181 processing (Seviour et al. 2011). Further increases will rely on additional engineering strategies
182 to alter the molecular structure of the polysaccharide and reduce the resulting viscosity via
183 modifications such as limiting polymer length or altering the degree of acylation or pyruvylation
184 of a compound (Schmid et al. 2015). Bacterial production also offers unprecedented levels of
185 purity when compared to extraction methods from other sources – for example, cellulose free
186 from the common plant contaminants ligin and hemicellulose (Klemm et al. 2005). Furthermore,
187 with the growing understanding of the pathways behind bacterial synthesis of such
188 exopolysaccharides and recent advances in bioinformatics and systems biology, it may soon be
189 possible to engineer bacteria to produce entirely novel polysaccharides with useful chemical
190 properties. Indeed, a metabolic engineering approach was recently used to synthesize a variant
191 form of cellulose containing a proportion of GlcNAc monomers in addition to the usual glucose.
192 This modification resulted in the production of a biopolymer that is far more readily
193 biodegradable than the standard form (Yadav et al. 2010).

194

195 **Table 1.** Extensively studied bacterial exopolysaccharides: composition, sources and uses.

196

EPS	Components	Organism	Main applications*
Cellulose	Glucose	*Gluconacetobacter xylinus*	Foods (indigestible fiber)

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

			Wound healing Engineered blood vessels Audio speaker diaphragms
Xanthan	Glucose Mannose Glucuronic acid Acetate Pyruvate	*Xanthomonas campestris*	Foods Petroleum industry Pharmaceuticals Cosmetics and personal care products Agriculture
Alginate	Guluronic acid Mannuronic acid Acetate	*Pseudomonas aeruginosa, Azotobacter vinelandii*	Surgical dressings Wound management Controlled drug release
Gellan	Glucose Rhamnose Glucuronic acid Acetate Glycerate	*Sphingomonas paucimobilis*	Foods Pet food Pharmaceuticals Agar substitute
Dextran	Glucose	*Leuconostoc mesenteroides*	Foods Blood volume expander Chromatographic media
Curdlan	Glucose	*Agrobacterium tumefaciens,*	Foods Pharmaceuticals

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

		Alcaligenes faecalis	Heavy metal removal Concrete additive
Hyaluronic acid	Glucuronic acid *N*-acetylglucosamine	*Streptococcus zooepidemicus, Bacillus subtilis*	Medicine Solid culture media
Succinoglycan	Glucose Galactose Acetate Pyruvate Succinate	*Sinorhizobium meliloti*	Food Oil recovery
Levan	Fructose	*Bacillus subtilis, Zymomonas mobilis*	Food (prebiotic) Medicines Cosmetics

197 *summarized from (Schmid et al. 2015)

198

199 **Bioengineering of intracellular and cell-associated polysaccharides in bacteria.** The most
200 widely manipulated cellular polysaccharide biosynthesis system is probably the LPS pathway
201 (**Fig. 2**), in part due to the significance of this polysaccharide in pathogenesis, but also owing to
202 the conserved mechanistic nature of the pathway combined with the highly variable glycan
203 structures produced.

204 The tendency for genes responsible for production of a bacterial polysaccharide to be organized
205 as a single, continuous operon, especially in the case of O-antigens and CPS has greatly
206 facilitated the transfer of polysaccharide coding loci from their native species into a heterologous
207 host, typically *E. coli*. Early methods generally centered around the generation of a cosmid
208 library from fragmented genomic DNA, followed by screening of individual cosmids at the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

209 genomic or phenotypic level to locate clones conferring production of the polysaccharide of
210 interest. This approach has been used to produce a variety of O-antigens from Gram-negative
211 organisms including *Pseudomonas aeruginosa*, *Salmonella typhimurium* and *Yersinia*
212 *enterocolitica* in an *E. coli* strain background (Goldberg et al. 1992). A similar approach has also
213 been employed for the production of CPS from the Gram-positive organism *Streptococcus*
214 *pneumoniae* in the Gram-positive host *Lactococcus lactis* (Gilbert et al. 2000). Cloning of
215 sequenced, annotated polysaccharide biosynthetic loci has enabled production in *E. coli* of
216 polysaccharides from diverse Gram-negative species such as *Bukholderia pseudomallei*
217 (Garcia-Quintanilla et al. 2014) and *Francisella tularensis* (Cuccui et al. 2013). A further
218 advance was the recent demonstration that various CPS structures from the Gram-positive
219 bacterium *S. pneumoniae* could be produced in a Gram-negative host, namely *E. coli*, using the
220 *en bloc* transfer of the entire CPS coding locus (Kay et al. 2016; Price et al. 2016). The
221 recombinant CPS structures are produced essentially as an O-antigen in *E. coli*, and some
222 features of processing appear to be borrowed from the host, including the action of the O-
223 antigen ligase WaaL in attaching the polymerized polysaccharide to the outer core on lipid A,
224 and subsequent transport to the outer surface of the cell. These findings demonstrated an
225 unexpected cross-compatibility between systems from two disparate sources, and highlighted
226 the mechanistic similarity of CPS biosynthesis in Gram-positive bacteria and O-antigen
227 biosynthesis in Gram-negative bacteria.

228 Recently, the wide availability of whole-genome sequences and a thorough
229 understanding of the mechanisms behind bacterial polysaccharide biosynthesis has led to a
230 more informed approach to the production of heterologous polysaccharides. A recent study
231 produced two different *Staphylococcus aureus* CPS structures by expressing combinations of *P.*
232 *aeruginosa* and *S. aureus* glycosyltransferases in *E. coli*, with sugar precursors provided by a
233 combination of *P. aeruginosa* enzymes along with native enzymes in the *E. coli* host. The

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

234 resulting glycans were confirmed by MALDI-TOF/TOF tandem mass spectrometry analysis as
235 having the same structure as the native CPS, and were recognized by capsular serotype-
236 specific typing antiserum (Wacker et al. 2014). Hence, bacterial glycosyltransferase enzymes
237 may be regarded as modular entities defined only by function, opening up a new approach to
238 polysaccharide bioengineering in host species such as *E. coli*. This insight also facilitates the
239 engineering of bacterial glycans in cases where information regarding the biosynthesis of a
240 target polysaccharide (and/or its intermediates) is incomplete or incompatible with further
241 processing as a result of assembly on a lipid other than Und-PP. For example, the Vi
242 polysaccharide of *Salmonella enterica* serovar Typhi is currently licensed as a purified
243 polysaccharide vaccine for typhoid fever, but represents an interesting candidate for further
244 development as a glycoconjugate. Unfortunately, recombinant production of this polysaccharide
245 is challenging because the lipid on which it is assembled in the native host is not currently
246 known. To circumvent this issue, Wetter et al. modified the *E. coli* O121 O-antigen, a structure
247 that is well known to build on Und-PP, to resemble the Vi polysaccharide. Following transfer of
248 the resulting Vi-like polysaccharide to a carrier protein, a glycoconjugate was produced that
249 elicited antibodies that were immunoreactive with *E. coli* O121 LPS (Wetter et al. 2013).

250 **Bioengineering of eukaryotic polysaccharides on the LPS core in bacteria.** The ability to
251 expand the bacterial polysaccharide production system to engineer structures beyond
252 prokaryotic polysaccharides is crucial if this approach is to become broadly applicable and
253 useful. Several human-like glycans have been assembled on a truncated LPS outer core
254 structure. Typical mutations involve the disruption of the second glycosyltransferase enzyme of
255 the outer core, resulting in an intact lipid A molecule, coupled to a complete inner core structure,
256 but with only a single glucose residue from the outer core added to the second heptose residue
257 of the inner core (see **Fig 2.**). This exposed glucose then becomes the attachment site for

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

258 recombinant glycans, while the Lpt export system translocates the resulting LOS structure to the
259 surface of the cell, ensuring the recombinant glycan is exposed (Merritt et al. 2013).

260 The human glycosphingolipid globotriaosylceramide (Gb_3) is the receptor for Shiga-toxin
261 (Stx), a potent AB_5 toxin produced by pathogenic species such as *Shigella dysenteriae* and *E.*
262 *coli* O157. This receptor is composed of a trisaccharide, Gal(α1–4)Gal(β1–4)Glc, and is present
263 on many eukaryotic cell types, but is found at the highest concentrations in renal tissue and in
264 microvascular endothelial cells (Paton et al. 2000). An analogous structure to the Gb_3 receptor
265 is produced by *Neisseria spp*. as a component of LOS and is representative of a common
266 strategy employed by mucosal pathogens whereby surface display of host glycan epitopes aids
267 immune evasion (Cress et al. 2014). Expression of the glycosyltransferases LtxC from *Neisseria*
268 *meningitidis*, and LtxE from *Neisseria gonhorroeae* in *E. coli* resulted in the production of a
269 novel LPS-associated Gb_3 polysaccharide structure. When administered to mice, the
270 engineered *E. coli* were found to protect against challenge with a Shiga-toxin producing *E. coli*
271 (STEC) strain, suggesting an effective molecular mimic of the toxin binding site had been
272 recreated that sequestered the secreted toxin (Paton et al. 2000). An analogous approach has
273 been used to engineer *E. coli* cells that express molecular mimics for other receptors implicated
274 in bacterial toxin binding: globotetraosylceramide (Gb_4), and the gangliosides GM_1 and GM_2
275 (Focareta et al. 2006; Hostetter et al. 2014). These engineered bacterial strains have also
276 proven efficacious in animal models for the treatment of toxin-associated bacterial infections
277 such as cholera and STEC.

278 A similar approach was used to produce the ganglioside GM_3 epitope,
279 NeuNAcα(2,3)Galβ(1,4), as an attachment to the exposed glucose residue of truncated lipid A
280 (Ilg et al. 2010). This feat was accomplished by expressing the *Neisseria* enzymes SiaB, a
281 CMP-sialic acid synthetase, together with the galactosyltransferase LgtE and the sialic acid
282 transferase Lst, which together generated a GM_3-like structure that was displayed on the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

283 surface of the cell. This strain may be useful for investigating the effects of sialic acid-containing
284 bacterial LOS structures and their role in development of post-infection autoimmune diseases
285 such as Guillain-Barre syndrome. Other human-like glycans with a role in bacterial attachment
286 have also been expressed in *E. coli*, including fucosylated oligosaccharides: the blood group H,
287 Lewis X (Le^X) and Lewis Y (Le^Y) antigens (Yavuz et al. 2011) and poly-*N*-acetyllactosamine
288 (Mally et al. 2013). Fucose is a common component of human glycans, and is thought to play a
289 role in the binding of various pathogenic bacteria including *P. aeruginosa* and *C. jejuni*, and it is
290 envisioned that these strains may prove useful for studying specific bacterial interactions with
291 human receptors, as well as informing the design of competitive inhibitors for novel probiotic-
292 based therapies.

293 A further example of a eukaryotic glycan that may also be produced as a bacterial mimic
294 is polysialic acid (PolySia), a linear homopolymer of α-2,8-linked sialic acid residues. In humans,
295 this glycan is most notably found as an elaboration of the *N*-linked glycan on neural cell
296 adhesion molecule (NCAM), but is also expressed by *E. coli* K1 and *N. meningitidis* group B as
297 the K1 capsule and CPS A, respectively (Moe et al. 2009). Owing to its occurrence on these
298 pathogens as well as its enhanced expression on some malignant tumors (Livingston et al.
299 1988; Komminoth et al. 1991), PolySia represents an intriguing target for vaccine or therapeutic
300 antibody development. By expressing a combination of glycosyltransferases from *N.*
301 *gonorrhoeae*, *C. jejuni* and *E. coli*, Valentine and co-workers were able to produce PolySia
302 directly on the LPS core of an *E. coli* strain not normally capable of synthesizing this structure.
303 Interestingly, where the aforementioned GM_3 production study supplied NeuAc via the growth
304 medium and relied on a single synthetase enzyme to convert the sugar into the nucleotide
305 activated form CMP-NeuAc (Ilg et al. 2010), here the authors reconstituted the entire
306 biosynthesis pathway capable of converting the readily available housekeeping sugar UDP-

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

307 GlcNAc into CMP-NeuAc (Valentine et al. 2016), highlighting the flexibility and versatility of
308 bacteria as hosts for glycoengineering.
309 **Bioengineering of eukaryotic polysaccharides on the lipid anchor Und-PP in bacteria.**
310 Because direct conjugation to the LPS core is not always possible or desirable, alternative sites
311 for polysaccharide assembly have also been explored such as the common lipid anchor Und-
312 PP. In *E. coli* K-12, the ECA and O-antigen biosynthesis pathways involve installation of a
313 GlcNAc residue on Und-PP by an initiating glycosyltransferase called WecA. By introducing
314 glycosyltransferases from the *Haemophilus influenzae* LOS biosynthesis pathway that were
315 capable of modifiying this Und-PP-linked GlcNAc in the recombinant system, a tetrasaccharide
316 resembling the Le^X antigen (minus the fucose residue) was assembled on Und-PP (Hug et al.
317 2011). The use of this lipid as a carrier enabled subsequent conjugation of the glycan to a
318 protein using an oligosaccharyltransferase-mediated mechanism that is described in greater
319 detail below. To complete the Le^X structure, the purified glycoconjugate was subjected to in vitro
320 enzymatic elaboration to add the fucose residue (Hug et al. 2011). The use of engineered
321 bacteria to produce Le^X containing glycoproteins is significant because these proteins are
322 known to function as immunomodulatory molecules (Atochina et al. 2001; van Die et al. 2003;
323 Srivastava et al. 2014), and have been shown to ameliorate symptoms associated with
324 autoimmune disorders in animal models (Atochina and Harn 2006).
325 Another human-like glycan produced in a similar manner is the Thomsen-Friedenreich
326 antigen (T antigen), a Galβ1-3GalNAc disaccharide. Valentine et al. used UndPP-linked GlcNAc
327 as a primer for producing the T antigen disaccharide (Valentine et al. 2016). This was
328 accomplished by addition of two heterologous glycosyltransferases and a nucleotide sugar
329 epimerase to ensure availability of the required substrate UDP-GalNAc. Because T antigen is
330 overexpressed on a number of malignancies including breast, colon, prostate and stomach

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

331 cancer (Heimburg-Molinaro et al. 2011), recombinant biosynthesis could yield highly
332 immunogenic glycoconjugates that elicit antibodies against this important glycan epitope.
333 A final example of engineering human-like glycans in a bacterial host involved the
334 bottom-up creation of a eukaryotic *N*-glycan biosynthesis pathway. Specifically, the conserved
335 core of all human *N*-glycans, the oligosaccharide $Man_3GlcNAc_2$, was successfully produced on
336 Und-PP by co-expression of four eukaryotic glycosyltransferases, including the yeast uridine
337 diphosphate-*N*-acetylglucosamine transferases Alg13 and Alg14 and the mannosyltransferases
338 Alg1 and Alg2 (Valderrama-Rincon et al. 2012). By including a bacterial
339 oligosaccharyltransferase PglB from *C. jejuni*, glycans were successfully transferred to
340 eukaryotic target proteins as discussed below. The $Man_3GlcNAc_2$ structure has been shown to
341 be the minimal structure required for efficacy of a glycoprotein therapeutic (Van Patten et al.
342 2007), and is the predominant glycoform conjugated to proteins expressed in a baculovirus host
343 system. Furthermore, as the conserved core of human *N*-glycans, this structure has enormous
344 potential as a precursor for further modification, either in vivo or in vitro.

345 **Glycoprotein expression in bacterial hosts: current applications and future opportunities.**

346 The above findings demonstrate the remarkable versatility of bacterial systems for the
347 biosynthesis of a vast array of carbohydrate structures. However, to exploit the full potential of
348 carbohydrates, it is often necessary to conjugate these structures to additional biomolecules
349 such as proteins. Two different mechanisms are responsible for making the majority of proteins
350 that become covalently modified with sugar molecules (*i.e.*, glycoproteins). These mechanisms
351 are defined based on the amino acid residue onto which the glycan is installed. In *N*-linked
352 glycosylation, the glycan is attached to the nitrogen atom of an asparagine residue, while in O-
353 linked glycosylation the sugar moiety is attached to the oxygen atom of either a serine or a
354 threonine side chain. While both types of glycosylation were long believed to occur exclusively
355 in eukaryotes, multiple bacterial machineries for the generation of both types of modifications

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

356 have been discovered over the last 15 years. These bacterial glycosylation systems, or hybrids
357 thereof, have opened the door to using bacteria for the production of two important classes of
358 glycoproteins: (1) glycoconjugate vaccines, whereby immunogenic carbohydrates from
359 pathogens including bacteria and viruses are linked to proteins; and (2) therapeutic proteins that
360 are glycosylated in their natural form and require the modification for full function, for example,
361 monoclonal antibodies.

362 Glycoconjugates are amongst the most successful vaccines generated to date, eliciting
363 a robust T-cell dependent immune response and conferring protection across all age groups
364 (Vella and Pace 2015). For three important bacterial pathogens in particular, *H. influenza* type B
365 (Hib), *S. pneumoniae* and *N. meningitidis*, glycoconjugates have proven to be highly effective in
366 countries where they have been introduced (Ladhani 2012; Grijalva et al. 2007). The standard
367 production method for these conjugates involves the separate generation and purification of the
368 protein and the carbohydrate moiety, chemical activation thereof, and conjugation as well as
369 subsequent purification of the resulting glycoprotein (Lees et al. 2008). Even though it is an
370 established and accepted method, there are several drawbacks to this approach. Firstly, it
371 requires culturing large volumes of a pathogenic species of interest for the generation of the
372 native carbohydrate, followed by harvesting and purification of the carbohydrate. Depending on
373 the biosafety level of the species of interest, as well as the ease of culturing, this step can
374 present a major hurdle regarding the expansion of the technique to novel pathogenic species.
375 Secondly, the activation and chemical conjugation steps required to couple the glycan to the
376 carrier protein can be technically challenging and inefficient, resulting in low yields, as well as a
377 heterogeneous population of glycoproteins with different numbers of target glycans attached at
378 different locations throughout the protein. Therefore, alternative methods for generating
379 glycoconjugates that overcome some of these limitations are desired.

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

380 In addition to glycoconjugate vaccines, many proteins of therapeutic interest are also
381 glycoproteins. In fact, 70% of therapeutic proteins approved by regulatory agencies or currently
382 in clinical and preclinical trials are decorated with glycans in their native form (Sethuraman and
383 Stadheim 2006). Historically, this has limited the use of *E. coli* to proteins and peptides that are
384 not natively glycosylated such as insulin and homologues thereof or to those that are natively
385 glycosylated but are functional without the addition of the glycan moiety, such as human growth
386 hormone (hGH) and interferon α (Ferrer-Miralles et al. 2009). It should be pointed out that these
387 proteins often require additional post-translational modifications such as the addition of
388 polyethylene glycol (PEG) to increase serum half-life (Bailon and Won 2009). While some
389 notable breakthroughs have been made (Valderrama-Rincon et al. 2012), the routine use of *E.*
390 *coli* as a production platform for therapeutic glycoproteins and glycopeptides requires further
391 engineering of glycosylation pathways in this host.

392 ***N*-linked glycoprotein expression in bacteria.** The discovery of an *N*-glycosylation machinery
393 in the human intestinal bacterial pathogen *C. jejuni* (Szymanski et al. 1999) and the subsequent
394 functional transfer of the complete machinery into the more tractable species *E. coli* (Wacker et
395 al. 2002) demonstrated for the first time that bacteria could be an alternative source of
396 recombinant *N*-glycoproteins. Subsequent studies showed that a single enzyme, an
397 oligosaccharyltransferase named *Cj*PglB (PglB from *C. jejuni*), was responsible for transferring
398 the glycan to the acceptor protein. Interestingly, this enzyme was shown to share sequence
399 homology with the STT3 catalytic subunit of the eukaryotic oligosaccharyltransferase enzyme
400 complex (Wacker et al. 2002). A functional study of the genes within the glycosylation locus
401 demonstrated that the substrate glycan was assembled on the lipid carrier Und-PP (Linton et al.
402 2005), in a fashion similar to the O-antigen biosynthesis pathway present in many Gram-
403 negative species of bacteria (Hug and Feldman 2011). It was further demonstrated that the
404 *Cj*PglB enzyme possesses remarkably relaxed glycan substrate specificity. That is, in addition

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

405 to its native substrate oligosaccharide -- a heptasaccharide glycan with the structure
406 diNAcBacGalNAc$_5$Glc (Young et al. 2002) -- the enzyme was also able to recognize much larger
407 polysaccharides such as structurally different bacterial O-antigens and transfer these to proteins
408 (Feldman et al. 2005). Around the same time, a five amino acid glycosylation sequon for *Cj*PglB
409 was discovered (Kowarik et al. 2006), which could be engineered either into flexible secondary
410 structures within a protein (Kowarik et al. 2006) or at either the N- or the C-terminus (Fisher et
411 al. 2011). Altogether, these studies provided the requisite ingredients for making customized
412 recombinant bacterial glycoproteins, where potentially any protein of interest could be modified
413 with any glycan moiety at a desired position by co-expression of *Cj*PglB, the glycan of interest
414 assembled on Und-PP, and the desired acceptor protein modified to contain one or more
415 glycosylation sequon(s) (**Fig. 3**).

416 **Customized *N*-glycoproteins produced recombinantly in *E. coli.*** To date, the predominant
417 class of glycoproteins produced using the above components are conjugates in which bacterial
418 surface glycan structures are site-specifically linked to immunogenic carrier proteins. In the
419 majority of published cases, the glycans are O-antigen polysaccharides built on Und-PP (see
420 above for in-depth discussion of different methods used for the recombinant production of these
421 structures) and installed on the carrier protein by *Cj*PglB. **Table 2** summarizes the
422 glycoconjugate vaccine candidates generated and tested to date. While multiple studies have
423 demonstrated the generation of specific, and potentially protective antibody responses against
424 *E. coli*-derived glycoconjugate vaccine candidates*,* it is particularly noteworthy that two have
425 been successfully tested in Phase I trials. The first is a conjugate vaccine candidate against *S.*
426 *dysenteriae* type 1 composed of the O-antigen glycan coupled to the exotoxin A of *P.*
427 *aeruginosa* (EPA). Testing of this vaccine candidate in healthy adults at two different doses with
428 or without co-adminstration of adjuvant revealed it to be well tolerated and capable of eliciting
429 statistically significant antigen-specific humoral immune responses (Hatz et al. 2015). A second

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

430 conjugate vaccine candidate comprised of the *S. flexneri* 2a O-antigen conjugated to EPA was
431 also tested in healthy adults, with similar results regarding tolerance and immunogenicity
432 (Riddle et al. 2016). Hence, recombinant production of glycogonjugates in *E. coli* appears to be
433 a promising alternative to the traditional methods used for biomanufacturing conjugate vaccines.
434
435 **Table 2:** List of bacterial glycoconjugate vaccine candidates that have been produced using a
436 bacterial glycoengineering approach.

Target species	Carrier protein	Carbohydrate	Animal model	Safety in humans	Immunogenicity in humans	References
Shigella dysenteriae type 1	*C. jejuni* model glycoprotein AcrA, Exotoxin A from *P. aeruginosa* (EPA)	O-antigen		X	X	(Ihssen et al. 2010; Hatz et al. 2015; Ravenscroft et al. 2016)
Shigella flexneri 2a	EPA	O-antigen		X	X	(Kampf et al. 2015; Riddle et al. 2016)
Francisella	EPA	O-antigen	mice			(Cuccui et

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tularensis						al. 2013)
Burkholderia pseudomallei	AcrA	O-antigen	mice			(Garcia-Quintanilla et al. 2014)
Brucella abortus	AcrA	O-antigen of *Yersinia enterocolitica* O9*	mice			(Iwashkiw et al. 2012a)
Escherichia coli serotypes O1, O2, O6 and O25	EPA	O-antigens of the four strains	mice, rabbits, rats			(van den Dobbelsteen et al. 2016)
Escherichia coli O157:H7	Maltose binding protein (MBP)	O-antigen**	mice			(Ma et al. 2014)
Salmonella typhi	EPA	Vi capsule (O-antigen of *E. coli* O121)***	mice			(Wetter et al. 2013)
Staphylococcus aureus serotpyes 5 and 8	EPA and *S. aureus* α toxin	Capsular polysaccharide	mice			(Wacker et al. 2014)

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

437 *glycoconjugate generated by recombinant expression of the glycosylation machinery in *Y.*
438 *enterocolitica* serotype O9, taking advantage of the structural identity of the O-antigen of this
439 species with the target species.
440 **glycoconjugate generated by recombinant expression of the glycosylation machinery directly
441 in the target *E. coli* serotype.
442 ***glycoconjugate generated by recombinant expression of a related O-antigen structure
443 modified to resemble the target glycan.
444
445 Glycoconjugate proteins produced recombinantely in *E. coli* have found uses in other
446 applications as well. For instance, bacterial glycoconjugates have been successfully used as
447 diagnostic tools for human and bovine brucellosis (Ciocchini et al. 2013; Ciocchini et al. 2014)
448 as well as for the Shiga-toxin-producing *E. coli* serotypes O157, O145, and O121 (Melli et al.
449 2015). Additionally, Shang and co-workers generated a glycoconjugate comprised of the
450 maltose binding protein (MBP) and the *E. coli* O86:B7 O-antigen, which bears structural
451 similarity to the blood group B antigen epitope. This glycoconjugate functioned as a 'molecular
452 sponge' to lower the levels of blood group B antibodies in plasma without negatively affecting
453 the clotting function of the plasma (Shang et al. 2016).
454 While there is a lot of promise for glycoconjugates where the sugar moiety is derived
455 from an immunogenic bacterial glycan, these types of glycans will not be useful in applications
456 where the goal is to install native, eukaryotic glycans onto therapeutic proteins. Several efforts
457 have been made to leverage the bacterial protein glycosylation machinery for the generation of
458 glycoproteins carrying mammalian glycans. Perhaps the most notable example is Valderrama-
459 Rincon et al. who demonstrated the complete recombinant assembly and transfer to protein of
460 the eukaryotic *N*-linked core glycan $GlcNAc_2Man_3$ (Valderrama-Rincon et al. 2012) (see above
461 for description of the approach used for biosynthesis of the glycan). Transfer of the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

462 $GlcNAc_2Man_3$ glycan to asparagine residues in several different target proteins including the Fc
463 domain of human immunoglobulin G (IgG) was achieved with *Cj*PglB, which as mentioned
464 above has fairly relaxed specificity towards the glycan substrate. One can envision an extension
465 of this glycan, either *in vivo* or potentially *in vitro*, to generate additional structures found in
466 mammalian N-glycans. In an alternative approach, post-processing of a purified pre-form of the
467 glycoconjugate outside of the bacterial cell can be performed to generate the final product. For
468 instance, the same $GlcNAc_2Man_3$ glycan structure was installed on a protein by a combination of
469 recombinant *in vivo* glycosylation of the protein with the *C. lari* heptasaccharide glycan,
470 $GalNAc_5GlcNAc$, followed by *in vitro* enzymatic trimming of the glycan down to a single GlcNAc
471 residue, and finally transglycosylation of the trimmed glycan with a preassembled $Man_3GlcNAc$
472 sugar to obtain the final structure (Schwarz et al. 2010). A similar combined method of in vitro
473 and in vivo glycosylation and modification was used to install the blood group antigen Le^x on a
474 protein (Hug et al. 2011). The recombinantely expressed tetrasaccharide $GalNAc_2Gal_2$ was
475 produced on Und-PP in *E. coli* and this glycan was subsequently transferred *in vivo* to an
476 acceptor protein using *Cj*PglB. Following purification, *in vitro* fucosylation was performed to yield
477 the final Le^x glycan on the protein. Although these combined *in vivo* and *in vitro* methods of
478 glycoprotein biosynthesis are potentially less applicable to large scale production of
479 glycoproteins, they nevertheless expand the range of glycan modifications on proteins, which
480 may be beneficial for the generation of glycoproteins carrying sugars that are potentially too
481 challenging for the expression and transfer *in vivo* alone.

482 **Expanding glycosylation through identification of alternative oligosaccharyltransferases.**

483 While *Cj*PglB remains one of the best-characterized bacterial oligosaccharyltransferases, there
484 are two main limitations that restrict its use for the coupling of designer glycans to acceptor
485 proteins. First, compared to the canonical eukaryotic glycosylation sequon, N-X-S/T (where X
486 can be any amino acid except proline), used by eukaryotic oligosaccharyltransferases, *Cj*PglB

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

487 requires an extended sequon (D/E-X_1-N-X_2-S/T) for the attachment of glycans to proteins
488 (Kowarik et al. 2006). One consequence of this requirement is that, at a minimum, these five
489 amino acids need to be engineered into the protein of interest either by addition of the residues
490 as a terminal or internal tag or by changing of a native stretch of amino acids to render it a
491 substrate for glycosylation. If these modifications are added to either of the termini, it can be
492 speculated that this will not have a major impact on the overall structure and function of the
493 protein. However, it may be desirable to engineer the site of glycan attachment into the protein,
494 in which case these modifications may interfere with protein folding and/or function. Another
495 consequence is that native *N*-glycoproteins of mammalian origin will need to have their shorter
496 sequons extended to include a D or E in the -2 position to be glycosylated by *Cj*PglB.

497 To address this limitation, several groups have used bioinformatics to identify
498 orthologues of *Cj*PglB, which were then functionally characterized in glyco-competent *E. coli*
499 cells (Jervis et al. 2010; Ielmini and Feldman 2011; Schwarz et al. 2011b; Ollis et al. 2015; Mills
500 et al. 2016). From these studies, oligosaccharyltransferases were identified from two species of
501 *Desulfovibrio*, in particular, that did not require the negatively charged amino acid at position -2
502 and were therefore able to glycosylate the shorter eukaryotic N-X-S/T sequon (Ielmini and
503 Feldman 2011; Ollis et al. 2015). Of these, only the PglB orthologue of *D. gigas* was able to
504 modify the native QYNST sequon in the Fc domain of human IgG (Ollis et al. 2015), suggesting
505 that additional factors govern acceptor-site specificity and must be satisfied to allow for the
506 installation of glycans onto shorter eukaryotic sequons. Additionally, the orthologue from *D.*
507 *desulfuricans* showed markedly lower efficiency in transferring the *E. coli* O7 O-antigen
508 polysaccharide (Ielmini and Feldman 2011), suggesting this enzyme may not be as flexible as
509 *Cj*PglB regarding the glycan structure. As no other polysaccharides were tested as substrates
510 for the *D. desulfuricans* PglB, it is unclear whether the low efficiency of transfer the O7 O-
511 antigen is specific to this substrate or an inherent property of the enzyme. The ability of the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

512 orthologues from *D. vulgaris* and *D. gigas* to transfer mono- and polysaccharides was not
513 tested, so it remains unclear whether these enzymes may be useful in the generation of custom
514 glycoconjugates.

515 In parallel to the functional characterization of *Cj*PglB orthologues, a directed evolution
516 approach has applied to *Cj*PglB with the goal of relaxing the acceptor-sequon specificity. Using
517 the crystal structure of the closely related PglB enzyme of *C. lari* (Lizak et al. 2011) as guidance,
518 combined with a high-throughput genetic screen using a secreted acceptor protein, a library of
519 *Cj*PglB mutants was screened for the ability of the enzyme to glycosylate non-canonical
520 acceptor protein sites (Ollis et al. 2014). This screen identified three *Cj*PglB variants that no
521 longer required the negatively charged residue at the -2 position. The three mutants
522 glycosylated a eukaryotic protein at its native N-X-S/T sequon, suggesting that these enyzmes
523 may be useful for authentically glycosylating eukaryotic proteins and peptides. While the glycan
524 specificity was not specifically tested, the fact that the mutants were derived from *Cj*PglB
525 suggests that the relaxed glycan specificity of the parent enzyme will remain.

526 A second limitation of *Cj*PglB is the requirement of the native enzyme for an acetamido
527 group at the monosaccharide that constitutes the reducing end of the oligo- or polysaccharide
528 (Wacker et al. 2006). Many glycans of interest do not terminate in a glycan that conforms to this
529 requirement, such as most capsular glycans of *S. pneumoniae* serotypes that terminate in either
530 galactose or glucose residues (Bentley et al. 2006). While a natural variant among the
531 orthologues of *Cj*PglB enzymes from other species may lack this requirement, evidence for this
532 has yet to be reported. In fact, two studies analyzing the protein *N*-glycan diversity within the
533 *Campylobacter* genus and in one species of *Helicobacter* identified exclusively sugars
534 containing an acetamido group at the reducing end (Jervis et al. 2012; Nothaft et al. 2012),
535 suggesting that this is a shared feature among many of the bacterial species that possess
536 protein *N*-glycosylation machineries. The same appears to be true for the sugar attached to an

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

537 identified glycoprotein in *D. gigas*, which was *N*-glycosylated with a di-saccharide of GlcNAc and
538 *N*-acetylallosamine (Santos-Silva et al. 2007). To address this issue, one study used structure-
539 guided mutagenesis to engineer a *Cj*PglB variant that was able to transfer two O-antigens from
540 *S. typhimurium* which both contain non-acetylated sugars (galactose residues) at the reducing
541 end (Ihssen et al. 2015). This work demonstrates that the glycan specificity of *Cj*PglB can be
542 engineered to a certain extent, and suggests that in the future it will be possible to transfer
543 virtually any glycan to any protein using modified versions of *Cj*PglB.
544 **Alternative routes for bacterial protein *N*-linked glycosylation.** A novel family of bacterial
545 enzymes has recently emerged that may be of potential use in bacterial glycoengineering. In
546 contrast to the enzymes described in the previous section, these enzymes: (1) are active in the
547 bacterial cytoplasm, not the periplasm; (2) use nucleotide-activated glycans instead of lipid-
548 linked glycans as a substrate; and (3) recognize the shorter, bacterial N-X-S/T glycosylation
549 sequon (McCann and St Geme 2014). The first member of the family was discovered in *H.*
550 *influenzae*, and shown to be involved in the glycosylation of the high molecular weight adhesin
551 protein HMW1 (Grass et al. 2003). The glycans attached to the adhesin protein were identified
552 predominantely as hexose sugars, and glycosylation of the adhesin protein was demonstrated
553 to be important for correct secretion of the adhesin as well as adhesion of the bacteria to airway
554 epithelial cells (Grass et al. 2010). Further members of the family have been identified in several
555 other species of bacteria (McCann and St Geme 2014), and in vitro experiments confirmed
556 activity of the orthologues from *Y. enterocolitica* and *Actinobacillus pleuropneumoniae* (Schwarz
557 et al. 2011a). The preferred substrate for the *A. pleuropneumoniae* enzyme (termed ApNGT)
558 was demonstrated to be UDP-Glc (Schwarz et al. 2011a), and a downstream gene was shown
559 to encode a glycosyltranferase enzyme that was able to extend the Glc moiety installed by
560 ApNGT with further Glc residues. Additionally, when expressed in *E. coli*, ApNGT was shown to
561 glycosylate recombinantly co-expressed auto-transporter proteins from the same species (the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

562 enzyme's native substrate), as well as co-expressed human erythropoietin (EPO) and several
563 native *E. coli* proteins (Naegeli et al. 2014). A polypeptide modified with a glucose moiety by
564 ApNGT was also successfully elaborated through *in vitro* transglycosylation mediated by
565 endoglycosidase enzymes (Lomino et al. 2013). This suggests that ApNGT and other enzymes
566 from this family may be useful tools for installation of a priming glucose residue on proteins of
567 interest followed by either *in vitro* or *in vivo* elaboration of the glycan. It can also be envisioned
568 that directed evolution of the enzyme from this family may allow for the modulation of the
569 carbohydrate specificity in a similar way to *Cj*PglB.

570 **Customized *O*-glycoproteins produced recombinantly in *E. coli.*** In addition to the bacterial
571 *N*-glycosylation mechanisms discussed above, pathways that lead to the modification of serine
572 or threonine residues (*O*-linked glycosylation) have also been identified in several bacterial
573 species. These mechanisms are more commonly found in bacteria than their *N*-glycosylation
574 counterparts (Iwashkiw et al. 2013), and are currently being pursued for recombinant protein
575 glycosylation. The following section will highlight similarities and differences between the *N*- and
576 *O*-linked pathways.

577 Over the last decade, *O*-glycosylation machineries that share mechanistic similarities
578 with the *N*-glycosylation pathways described above have been identified and characterized in
579 several bacterial species (Iwashkiw et al. 2013). It was initially observed that the type IV pilus
580 subunit protein PilA in *P. aeruginosa* strain 1244 was modified with a glycan in a manner
581 dependent on the product of the gene adjacent to *pilA* named PilO/TfpO (Castric 1995). A
582 similar machinery was identified in *N. mengingitidis*, whereby deletion of a gene termed *pglL* led
583 to the loss of a carbohydrate moiety from the pilus subunit protein PilE (Power et al. 2006).
584 Interestingly, both the *P. aeruginosa* PilO/TfpO and the *N. meningitidis* PglL proteins showed
585 homology to O-antigen ligase proteins that are involved in transfer of the O-antigen subunit from
586 the lipid carrier Und-PP onto the lipid A moiety during LPS biogenesis (Whitfield et al. 1997).

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

587 This suggested that these enzymes may use Und-PP-linked glycans as substrate. Analysis of
588 the glycan structure present on *P. aeruginosa* PilA showed the presence of a single O-antigen
589 repeat unit, further strengthening the hypothesis that Und-PP-linked glycans may be the
590 substrate for this enzyme family (Castric et al. 2001). When PilO/TfpO and PilA from *P.*
591 *aeruginosa* (or PglL and PilE from *N. meningitidis*) were recombinantly co-expressed in *E. coli*
592 along with a Und-PP-linked oligo- or polysaccharide, transfer of the glycan to the pilin protein
593 was observed (Faridmoayer et al. 2007). These results not only demonstrated recombinant
594 activity of this new family of bacterial O-oligosaccharyltransferase enzymes, but also confirmed
595 the substrate identity as Und-PP-linked glycans. Further analysis of the glycan specificity of
596 PglL demonstrated a remarkable promiscuity with regards to the glycan. Diverse glycan
597 structures were shown to be transferred to PilE by PglL in vivo including structures containing a
598 Gal residue at the reducing end such as the *S. typhimurium* LT2 O-antigen and the di-
599 saccharide-pentapeptide peptidoglycan building block, none of which are substrates for the *C.*
600 *jejuni* oligosaccharyltransferase *Cj*PglB (Faridmoayer et al. 2008). Additionally, in vitro
601 glycosylation experiments revealed that the enzyme displayed flexibility toward the lipid carrier
602 (Faridmoayer et al. 2008; Musumeci et al. 2013). Altogether, these characteristics suggest that
603 this enzyme is a very promising tool for the generation of designer glycoproteins with *O*-linked
604 sugars.

605 To date, however, the biotechnological use of this enzyme family has been hampered by
606 one major bottleneck. Unlike in the case of *Cj*PglB, there is a lack of a consensus sequon for
607 glycosylation that would allow for the 'tagging' of any protein as a substrate for *O*-glycosylation.
608 Analysis of the *O*-glycome of several organisms that possess PglL-like *O*-glycosylation systems
609 identified multiple glycosylated proteins, and while these helped to determine that the amino
610 acid residues around the glycan attachment site were rich in serine, proline and alanine, they
611 did not reveal the presence of any consensus sequence (Vik et al. 2009; Iwashkiw et al. 2012b;

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

612 Lithgow et al. 2014; Elhenawy et al. 2016). Towards a more universal glycosylation strategy,
613 Qutyan and coworkers showed that a C-terminal fusion of *E. coli* alkaline phosphatase with the
614 final 15 amino acids from the C-terminus of PilA was glycosylated by PilO/TfpO when expressed
615 in *P. aeruginosa*; however, the observed glycosylation was not very efficient (Qutyan et al.
616 2010). Additionally, while it has been shown that PilO/TfpO has relatively relaxed specificity and
617 was able to transfer multiple different serotype *O*-glycans of *P. aeruginosa* (DiGiandomenico et
618 al. 2002), the enzyme was only able to transfer a single O-antigen subunit both in the native
619 organism as well as recombinantly in *E. coli* (DiGiandomenico et al. 2002; Faridmoayer et al.
620 2007). Hence, alternative PilO/TfpO *O*-oligosaccharyltransferases will need to be identified or
621 engineered for transferring longer polysaccharides, which are often desirable for
622 glycoengineering purposes. This issue appears to have been solved recently by Pan and co-
623 workers who reported the development and optimization of an *O*-linked 'glycosylation tag'
624 consisting of an 8 amino acid motif flanked by two approximately 10 amino acid sequences
625 containing mainly hydrophilic residues (Pan et al. 2016). This tag was successfully fused to both
626 the N- and C-termini of three potential vaccine carrier proteins -- the cholera toxin B subunit,
627 exotoxin A from *P. aeruginosa*, and the detoxified variant of diphtheria toxin CRM197 -- and
628 glycosylated with two different sugars including the *S. typhimurium* LT2 O-antigen, which as
629 discussed above is not a substrate for *Cj*PglB. Recombinant *O*-glycoproteins produced with this
630 method were tested in a series of animal experiments and elicited a glycan-specific antibody
631 response (Pan et al. 2016). The ability to tag proteins for PglL-dependent *O*-glycosylation opens
632 up this enzyme family for biotechnological applications, in particular in cases where the glycan
633 of interest may not be an optimal substrate for *N*-glycosylation by *Cj*PglB.

634 **Alternative routes for bacterial protein *O*-linked glycosylation.**

635 Many bacterial species possess *O*-glycosylated flagellar proteins, with the glycosylation patterns
636 ranging from a single glycan at a single site to multiple glycans attached to different sites on the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

637 protein (Nothaft and Szymanski 2010). These glycans are installed in a processive manner, with
638 individual glycosyltransferases adding the glycans sequentially to the protein. This mechanism
639 is similar to the installation of *O*-linked glycans in eukaryotic mucin-like glycosylation (Kudelka et
640 al. 2015). It could therefore be hypothesized that enzymes from these machineries could
641 potentially be used/engineered to install mucin-like glycans on human proteins. The successful
642 recombinant installation of the first monosaccharide of the core of human mucin-like glycan, a
643 GalNAc residue, has been demonstrated in the cytoplasm of *E. coli* using a recombinantly
644 expressed human GalNAc transferase enzyme (Henderson et al. 2011). However, no further
645 elaboration of this priming glycan with other sugars has been demonstrated.

646 **Alternative therapeutic bacterial conjugates.** Although some unconjugated polysaccharides
647 are currently licensed as vaccines, they often elicit a T-cell independent immune response
648 stimulated by the extensive cross-linking of receptors on the surface of B cells. As such, they
649 are poorly immunogenic in children under two years of age and elderly patients, greatly limiting
650 their usefulness (De Gregorio and Rappuoli 2014). While protein conjugation is the most widely
651 studied approach to counter this problem, the field of bacterial glycobiology is opening up
652 alternative approaches to boost the immunogenicity of carbohydrate epitopes.

653 One such approach is based on bacterial outer membrane vesicles (OMVs), which are
654 small (20-200 nm in size) liposomes released from the outer membrane of nearly all Gram-
655 negative bacterial species. These vesicles are non-replicating versions of their bacterial 'parent',
656 and contain many of the same components as the bacterial outer membrane including
657 membrane proteins, CPS, and LOS and LPS, as well as some of the luminal components of the
658 bacterial periplasm (Kulp and Kuehn 2010). OMVs have garnered interest as vaccine
659 candidates because vesicles from several bacterial pathogens have been shown to possess
660 potent immunogenic capacities (Schild et al. 2008; Ellis et al. 2010; Alaniz et al. 2007).
661 Intriguingly, OMVs appear to also possess intrinsic adjuvant properties, potentially removing the

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

662 need to include adjuvants in the formulation (Sanders and Feavers 2011; Chen et al. 2010)
663 OMVs derived directly from pathogenic *N. meningitidis* have been successfully incorporated into
664 a commercial vaccine formulation, the recently licensed Bexsero (Holst et al. 2009; Gorringe
665 and Pajon 2012). Native OMVs have been further engineered to carry additional immunogenic
666 proteins, which are recombinantly displayed on the surface of the OMV through genetic fusion
667 to outer membrane proteins or in the OMV lumen through periplasmic expression (Chen et al.
668 2010; Muralinath et al. 2011). Importantly, robust immune responses against these recombinant
669 immunogens have been demonstrated (Chen et al. 2010; Muralinath et al. 2011). Three recent
670 reports highlight a novel bacterial glycoengineering approach to OMV-based vaccines whereby
671 immunogenic glycans are recombinantly displayed on the exterior of OMVs. The approach
672 takes advantage of the following: (1) the fact that standard laboratory strains of *E. coli* have lost
673 the ability to produce a native O-antigen glycan due to the insertion of an IS element in the
674 second glycosyltransferase gene *wbbL* (Liu and Reeves 1994) while the rest of the mechanism
675 including the flippase and ligase genes remain intact; (2) the ability to recombinantly express
676 non-native polysaccharides in *E. coli*; (3) the fact that the O-antigen ligase WaaL has relative
677 relaxed glycan specificity and will efficiently transfer engineered glycans from Und-PP to the
678 lipid A-core in cells that lack the native O-antigen (Han et al. 2012); and importantly, (4) the
679 recombinant O-antigen is efficiently transported to the cell surface and packaged into released
680 OMVs. Using this approach, *E. coli*-derived glycosylated OMVs (glycOMVs) have been
681 decorated with the O-antigens of eight Gram-negative bacterial species including *F. tularensis*
682 (Chen et al. 2016), PolySia (Valentine et al. 2016), the CPS of *S. pneumoniae* serotype 14, and
683 the *N*-linked heptasaccharide of *C. jejuni* (Price et al. 2016). Following immunization, the
684 glycOMVs carrying the *F. tularensis* O-antigen were shown to elicit significant serum titers of
685 class-switched, glycan-specific IgG antibodies in mice, and prolonged survival upon challenge
686 with the highly virulent *F. tularensis* subsp. *tularensis* (type A) strain Shu S4 (Chen et al. 2016).

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

687 Likewise, glycOMVs decorated with PolySia also elicited glycan-specific IgG antibodies in
688 mouse immunization studies, and the serum antibodies had potent bactericidal activity, killing *N.*
689 *meningitidis* serogroup B bacteria that possess a PolySia capsular glycan (Valentine et al.
690 2016). GlycOMVs carrying the *S. pneumoniae* serotype 14 CPS also elicited glycan-specific
691 antibodies in mice, and the serum antibodies were shown to possess potent bactericidal
692 properties when tested in an opsonophagocytic assay. In fact, the bacterial killing of the serum
693 from mice vaccinated with the glycOMVs carrying the capsular glycan was as efficient as the
694 serum from mice that had been vaccinated with the commercial glycoconjugate vaccine
695 Prevnar13® (Price et al. 2016). And finally, glycOMVs displaying the *C. jejuni* *N*-linked glycan
696 were shown to lower levels *of C. jejuni* colonization in chickens by almost 4 logs (Price et al.
697 2016). The expansion of the technology to cover further species or serotypes is envisioned to
698 be relatively straightforward, simply requiring the recombinant expression of a pathogen-specific
699 glycan structure on the surface of *E. coli* cells.

700 A related approach to glycOMV vaccines is the development of whole-cell vaccines
701 displaying recombinant glycan epitopes. This strategy also leverages the fact that recombinant
702 polysaccharides assembled on Und-PP are often efficiently transferred to lipid A and displayed
703 as recombinant chimeric LPS on the surface of Gram-negative bacteria. This approach has
704 been evaluated using several different species of Gram-negative bacteria as hosts (*S. enterica*
705 serovar Typhi, *S. enterica* serovar Typhimurium and *E. coli*) carrying biosynthesis gene clusters
706 for immunogenic carbohydrates of *S. dysenteriae* serotype O1 (Xu et al. 2007), shiga-toxin
707 producing *E. coli* serotype O111 (Wang et al. 1999), and *C. jejuni* (Nothaft et al. 2016). In
708 contrast to glycOMV vaccine candidates, these whole-cell vaccine candidates replicate. While it
709 is desirable to control their ability to replicate, a balance needs to be found between controlling
710 the replication of the bacteria and ensuring they persist long enough in the vaccinated
711 organisms to generate a desired immune response. Genetic inactivation of the *aroA* gene

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

712 encoding a 5-enolpyruvylshikimate-3-phosphate synthetase, involved in the shikimate pathway
713 that directly connects glycolysis to the synthesis of aromatic amino acids (Bentley 1990), is a
714 commonly used strategy to attenuate live bacterial vaccine candidates. This is particularly useful
715 in species of *Salmonella* as these mutants are able to grow in rich media *in vitro* but become
716 self-limiting *in vivo*, where aromatic amino acids are not freely available (Ruby et al. 2012).
717 However, recent data suggests that deletion of *aroA*, at least in *S. enterica* serovar
718 Typhimurium, can lead to additional effects in cellular physiology that may have an influence on
719 the behavior of the recombinant bacteria within the host (Felgner et al. 2016). Nonetheless,
720 attenuated, glycan epitope-expressing bacteria offer an additional opportunity for
721 glycoengineering of vaccine candidates, in particular in areas where minimal cost of production
722 may be a priority, such as in poultry and other livestock vaccines.

723 **Concluding remarks and outlook.** In summary, bacterial expression systems have been
724 successfully used for the production of a variety of carbohydrate structures ranging from small
725 secreted oligosaccharides to repeating polymers of high molecular weight, and spanning
726 structures found in all kingdoms of life. Furthermore, the characterization of both *N*- and *O*-
727 linked protein glycosylation systems in a variety of bacterial species has greatly enhanced the
728 potential of bacterial systems for the generation of therapeutically relevant glycoconjugates.
729 These bacterial conjugation systems have been employed to generate well-defined therapeutic
730 compounds including the first conjugate vaccines produced entirely in bacteria as well as novel
731 immunogenic entities such as glycosylated outer membrane vesicles. Two of these bacterially-
732 derived glycoconjugates have recently undergone successful Phase I clinical trials, and new
733 candidates are also emerging.

734 Owing to their versatility and ease of manipulation, bacteria are an ideal host for the
735 production of a diverse array of structurally defined polysaccharides and glycoconjugates that
736 will be of interest as medical and industrial products. Furthermore, the low costs associated with

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

737 the culturing of bacterial strains, especially *E. coli*, opens up this technology to a far wider range
738 of laboratories than existing chemical/chemoenzymatic synthesis methods or mammalian cell
739 culture approaches. The findings from a recent report commissioned by the National Academy
740 of Sciences states that ‘glycans play roles in almost every biological process and are involved in
741 every major disease’ and further asserts that ‘the development of transformative methods for
742 the facile synthesis of carbohydrates and glycoconjugates should be a high priority’ (National-
743 Research-Council 2012). Bacterial glycoengineering represents an emerging field with the
744 potential to play a major role in meeting these goals.

745

746 **References**

747 Alaniz RC, Deatherage BL, Lara JC, Cookson BT (2007) Membrane vesicles are immunogenic
748 facsimiles of Salmonella typhimurium that potently activate dendritic cells, prime B and T
749 cell responses, and stimulate protective immunity in vivo. J Immunol 179 (11):7692-7701

750 Aoki-Kinoshita KF, Kanehisa M (2015) Glycomic analysis using KEGG GLYCAN. Methods Mol
751 Biol 1273:97-107

752 Atochina O, Daly-Engel T, Piskorska D, McGuire E, Harn DA (2001) A schistosome-expressed
753 immunomodulatory glycoconjugate expands peritoneal Gr1(+) macrophages that
754 suppress naive CD4(+) T cell proliferation via an IFN-gamma and nitric oxide-dependent
755 mechanism. J Immunol 167 (8):4293-4302

756 Atochina O, Harn D (2006) Prevention of psoriasis-like lesions development in fsn/fsn mice by
757 helminth glycans. Exp Dermatol 15 (6):461-468

758 Bailon P, Won CY (2009) PEG-modified biopharmaceuticals. Expert Opin Drug Deliv 6 (1):1-16

759 Bentley R (1990) The shikimate pathway--a metabolic tree with many branches. Crit Rev
760 Biochem Mol Biol 25 (5):307-384

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

761 Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabbinowitsch E, Collins M, Donohoe K,
762 Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kaltoft MS, Barrell B, Reeves
763 PR, Parkhill J, Spratt BG (2006) Genetic analysis of the capsular biosynthetic locus from
764 all 90 pneumococcal serotypes. PLoS Genet 2 (3):e31
765 Castric P (1995) pilO, a gene required for glycosylation of Pseudomonas aeruginosa 1244 pilin.
766 Microbiology 141 (Pt 5):1247-1254
767 Castric P, Cassels FJ, Carlson RW (2001) Structural characterization of the Pseudomonas
768 aeruginosa 1244 pilin glycan. J Biol Chem 276 (28):26479-26485
769 Chen DJ, Osterrieder N, Metzger SM, Buckles E, Doody AM, DeLisa MP, Putnam D (2010)
770 Delivery of foreign antigens by engineered outer membrane vesicle vaccines. Proc Natl
771 Acad Sci U S A 107 (7):3099-3104
772 Chen L, Valentine JL, Huang CJ, Endicott CE, Moeller TD, Rasmussen JA, Fletcher JR, Boll
773 JM, Rosenthal JA, Dobruchowska J, Wang Z, Heiss C, Azadi P, Putnam D, Trent MS,
774 Jones BD, DeLisa MP (2016) Outer membrane vesicles displaying engineered
775 glycotopes elicit protective antibodies. Proc Natl Acad Sci U S A 113 (26):E3609-3618
776 Ciocchini AE, Rey Serantes DA, Melli LJ, Iwashkiw JA, Deodato B, Wallach J, Feldman MF,
777 Ugalde JE, Comerci DJ (2013) Development and validation of a novel diagnostic test for
778 human brucellosis using a glyco-engineered antigen coupled to magnetic beads. PLoS
779 Negl Trop Dis 7 (2):e2048
780 Ciocchini AE, Serantes DA, Melli LJ, Guidolin LS, Iwashkiw JA, Elena S, Franco C, Nicola AM,
781 Feldman MF, Comerci DJ, Ugalde JE (2014) A bacterial engineered glycoprotein as a
782 novel antigen for diagnosis of bovine brucellosis. Vet Microbiol 172 (3-4):455-465
783 Cress BF, Englaender JA, He W, Kasper D, Linhardt RJ, Koffas MA (2014) Masquerading
784 microbial pathogens: capsular polysaccharides mimic host-tissue molecules. FEMS
785 Microbiol Rev 38 (4):660-697

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

786 Cuccui J, Thomas RM, Moule MG, D'Elia RV, Laws TR, Mills DC, Williamson D, Atkins TP, Prior
787 JL, Wren BW (2013) Exploitation of bacterial N-linked glycosylation to develop a novel
788 recombinant glycoconjugate vaccine against Francisella tularensis. Open Biol 3
789 (5):130002
790 Cuthbertson L, Kos V, Whitfield C (2010) ABC transporters involved in export of cell surface
791 glycoconjugates. Microbiol Mol Biol Rev 74 (3):341-362
792 De Gregorio E, Rappuoli R (2014) From empiricism to rational design: a personal perspective of
793 the evolution of vaccine development. Nat Rev Immunol 14 (7):505-514
794 DiGiandomenico A, Matewish MJ, Bisaillon A, Stehle JR, Lam JS, Castric P (2002)
795 Glycosylation of Pseudomonas aeruginosa 1244 pilin: glycan substrate specificity. Mol
796 Microbiol 46 (2):519-530
797 Drouillard S, Mine T, Kajiwara H, Yamamoto T, Samain E (2010) Efficient synthesis of 6'-
798 sialyllactose, 6,6'-disialyllactose, and 6'-KDO-lactose by metabolically engineered E. coli
799 expressing a multifunctional sialyltransferase from the Photobacterium sp. JT-ISH-224.
800 Carbohydr Res 345 (10):1394-1399
801 Elhenawy W, Scott NE, Tondo ML, Orellano EG, Foster LJ, Feldman MF (2016) Protein O-
802 linked glycosylation in the plant pathogen Ralstonia solanacearum. Glycobiology 26
803 (3):301-311
804 Ellis TN, Leiman SA, Kuehn MJ (2010) Naturally produced outer membrane vesicles from
805 Pseudomonas aeruginosa elicit a potent innate immune response via combined sensing
806 of both lipopolysaccharide and protein components. Infect Immun 78 (9):3822-3831
807 Faridmoayer A, Fentabil MA, Haurat MF, Yi W, Woodward R, Wang PG, Feldman MF (2008)
808 Extreme substrate promiscuity of the Neisseria oligosaccharyl transferase involved in
809 protein O-glycosylation. J Biol Chem 283 (50):34596-34604

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

810 Faridmoayer A, Fentabil MA, Mills DC, Klassen JS, Feldman MF (2007) Functional
811 characterization of bacterial oligosaccharyltransferases involved in O-linked protein
812 glycosylation. J Bacteriol 189 (22):8088-8098
813 Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A,
814 Valvano MA, Aebi M (2005) Engineering N-linked protein glycosylation with diverse O
815 antigen lipopolysaccharide structures in Escherichia coli. Proc Natl Acad Sci U S A 102
816 (8):3016-3021
817 Felgner S, Frahm M, Kocijancic D, Rohde M, Eckweiler D, Bielecka A, Bueno E, Cava F,
818 Abraham WR, Curtiss R, 3rd, Haussler S, Erhardt M, Weiss S (2016) aroA-Deficient
819 Salmonella enterica Serovar Typhimurium Is More Than a Metabolically Attenuated
820 Mutant. MBio 7 (5)
821 Ferrer-Miralles N, Domingo-Espin J, Corchero JL, Vazquez E, Villaverde A (2009) Microbial
822 factories for recombinant pharmaceuticals. Microb Cell Fact 8:17
823 Fisher AC, Haitjema CH, Guarino C, Celik E, Endicott CE, Reading CA, Merritt JH, Ptak AC,
824 Zhang S, DeLisa MP (2011) Production of secretory and extracellular N-linked
825 glycoproteins in Escherichia coli. Appl Environ Microbiol 77 (3):871-881
826 Focareta A, Paton JC, Morona R, Cook J, Paton AW (2006) A recombinant probiotic for
827 treatment and prevention of cholera. Gastroenterology 130 (6):1688-1695
828 Garcia-Quintanilla F, Iwashkiw JA, Price NL, Stratilo C, Feldman MF (2014) Production of a
829 recombinant vaccine candidate against Burkholderia pseudomallei exploiting the
830 bacterial N-glycosylation machinery. Front Microbiol 5:381
831 Gilbert C, Robinson K, Le Page RW, Wells JM (2000) Heterologous expression of an
832 immunogenic pneumococcal type 3 capsular polysaccharide in Lactococcus lactis. Infect
833 Immun 68 (6):3251-3260

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

834 Goldberg JB, Hatano K, Meluleni GS, Pier GB (1992) Cloning and surface expression of
835 Pseudomonas aeruginosa O antigen in Escherichia coli. Proc Natl Acad Sci U S A 89
836 (22):10716-10720
837 Gorringe AR, Pajon R (2012) Bexsero: a multicomponent vaccine for prevention of
838 meningococcal disease. Hum Vaccin Immunother 8 (2):174-183
839 Grass S, Buscher AZ, Swords WE, Apicella MA, Barenkamp SJ, Ozchlewski N, St Geme JW,
840 3rd (2003) The Haemophilus influenzae HMW1 adhesin is glycosylated in a process that
841 requires HMW1C and phosphoglucomutase, an enzyme involved in lipooligosaccharide
842 biosynthesis. Mol Microbiol 48 (3):737-751
843 Grass S, Lichti CF, Townsend RR, Gross J, St Geme JW, 3rd (2010) The Haemophilus
844 influenzae HMW1C protein is a glycosyltransferase that transfers hexose residues to
845 asparagine sites in the HMW1 adhesin. PLoS Pathog 6 (5):e1000919
846 Grijalva CG, Nuorti JP, Arbogast PG, Martin SW, Edwards KM, Griffin MR (2007) Decline in
847 pneumonia admissions after routine childhood immunisation with pneumococcal
848 conjugate vaccine in the USA: a time-series analysis. Lancet 369 (9568):1179-1186
849 Han W, Wu B, Li L, Zhao G, Woodward R, Pettit N, Cai L, Thon V, Wang PG (2012) Defining
850 function of lipopolysaccharide O-antigen ligase WaaL using chemoenzymatically
851 synthesized substrates. J Biol Chem 287 (8):5357-5365
852 Hatz CF, Bally B, Rohrer S, Steffen R, Kramme S, Siegrist CA, Wacker M, Alaimo C, Fonck VG
853 (2015) Safety and immunogenicity of a candidate bioconjugate vaccine against Shigella
854 dysenteriae type 1 administered to healthy adults: A single blind, partially randomized
855 Phase I study. Vaccine 33 (36):4594-4601
856 Heimburg-Molinaro J, Lum M, Vijay G, Jain M, Almogren A, Rittenhouse-Olson K (2011) Cancer
857 vaccines and carbohydrate epitopes. Vaccine 29 (48):8802-8826

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

858 Henderson GE, Isett KD, Gerngross TU (2011) Site-specific modification of recombinant
859 proteins: a novel platform for modifying glycoproteins expressed in E. coli. Bioconjug
860 Chem 22 (5):903-912
861 Herget S, Toukach PV, Ranzinger R, Hull WE, Knirel YA, von der Lieth CW (2008) Statistical
862 analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): characteristics
863 and diversity of bacterial carbohydrates in comparison with mammalian glycans. BMC
864 Struct Biol 8:35
865 Holst J, Martin D, Arnold R, Huergo CC, Oster P, O'Hallahan J, Rosenqvist E (2009) Properties
866 and clinical performance of vaccines containing outer membrane vesicles from Neisseria
867 meningitidis. Vaccine 27 Suppl 2:B3-12
868 Hostetter SJ, Helgerson AF, Paton JC, Paton AW, Cornick NA (2014) Therapeutic use of a
869 receptor mimic probiotic reduces intestinal Shiga toxin levels in a piglet model of
870 hemolytic uremic syndrome. BMC Res Notes 7:331
871 Hug I, Feldman MF (2011) Analogies and homologies in lipopolysaccharide and glycoprotein
872 biosynthesis in bacteria. Glycobiology 21 (2):138-151
873 Hug I, Zheng B, Reiz B, Whittal RM, Fentabil MA, Klassen JS, Feldman MF (2011) Exploiting
874 bacterial glycosylation machineries for the synthesis of a Lewis antigen-containing
875 glycoprotein. J Biol Chem 286 (43):37887-37894
876 Ielmini MV, Feldman MF (2011) Desulfovibrio desulfuricans PglB homolog possesses
877 oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein
878 acceptor sequence requirements. Glycobiology 21 (6):734-742
879 Ihssen J, Haas J, Kowarik M, Wiesli L, Wacker M, Schwede T, Thony-Meyer L (2015) Increased
880 efficiency of Campylobacter jejuni N-oligosaccharyltransferase PglB by structure-guided
881 engineering. Open Biol 5 (4):140227

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

882 Ihssen J, Kowarik M, Dilettoso S, Tanner C, Wacker M, Thony-Meyer L (2010) Production of
883 glycoprotein vaccines in Escherichia coli. Microb Cell Fact 9:61
884 Ilg K, Yavuz E, Maffioli C, Priem B, Aebi M (2010) Glycomimicry: display of the GM3 sugar
885 epitope on Escherichia coli and Salmonella enterica sv Typhimurium. Glycobiology 20
886 (10):1289-1297
887 Iwashkiw JA, Fentabil MA, Faridmoayer A, Mills DC, Peppler M, Czibener C, Ciocchini AE,
888 Comerci DJ, Ugalde JE, Feldman MF (2012a) Exploiting the Campylobacter jejuni
889 protein glycosylation system for glycoengineering vaccines and diagnostic tools directed
890 against brucellosis. Microb Cell Fact 11:13
891 Iwashkiw JA, Seper A, Weber BS, Scott NE, Vinogradov E, Stratilo C, Reiz B, Cordwell SJ,
892 Whittal R, Schild S, Feldman MF (2012b) Identification of a general O-linked protein
893 glycosylation system in Acinetobacter baumannii and its role in virulence and biofilm
894 formation. PLoS Pathog 8 (6):e1002758
895 Iwashkiw JA, Vozza NF, Kinsella RL, Feldman MF (2013) Pour some sugar on it: the expanding
896 world of bacterial protein O-linked glycosylation. Mol Microbiol 89 (1):14-28
897 Jervis AJ, Butler JA, Lawson AJ, Langdon R, Wren BW, Linton D (2012) Characterization of the
898 structurally diverse N-linked glycans of Campylobacter species. J Bacteriol 194 (9):2355-
899 2362
900 Jervis AJ, Langdon R, Hitchen P, Lawson AJ, Wood A, Fothergill JL, Morris HR, Dell A, Wren B,
901 Linton D (2010) Characterization of N-linked protein glycosylation in Helicobacter
902 pullorum. J Bacteriol 192 (19):5228-5236
903 Jia Y, Zhu J, Chen X, Tang D, Su D, Yao W, Gao X (2013) Metabolic engineering of Bacillus
904 subtilis for the efficient biosynthesis of uniform hyaluronic acid with controlled molecular
905 weights. Bioresour Technol 132:427-431

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

906 Kalynych S, Morona R, Cygler M (2014) Progress in understanding the assembly process of
907 bacterial O-antigen. FEMS Microbiol Rev 38 (5):1048-1065
908 Kampf MM, Braun M, Sirena D, Ihssen J, Thony-Meyer L, Ren Q (2015) In vivo production of a
909 novel glycoconjugate vaccine against Shigella flexneri 2a in recombinant Escherichia
910 coli: identification of stimulating factors for in vivo glycosylation. Microb Cell Fact 14:12
911 Kay EJ, Yates LE, Terra VS, Cuccui J, Wren BW (2016) Recombinant expression of
912 Streptococcus pneumoniae capsular polysaccharides in Escherichia coli. Open Biol 6
913 (4):150243
914 Klemm D, Heublein B, Fink HP, Bohn A (2005) Cellulose: fascinating biopolymer and
915 sustainable raw material. Angew Chem Int Ed Engl 44 (22):3358-3393
916 Komminoth P, Roth J, Lackie PM, Bitter-Suermann D, Heitz PU (1991) Polysialic acid of the
917 neural cell adhesion molecule distinguishes small cell lung carcinoma from carcinoids.
918 Am J Pathol 139 (2):297-304
919 Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC,
920 Hernandez M, Kelly JF, Wacker M, Aebi M (2006) Definition of the bacterial N-
921 glycosylation site consensus sequence. EMBO J 25 (9):1957-1966
922 Kudelka MR, Ju T, Heimburg-Molinaro J, Cummings RD (2015) Simple sugars to complex
923 disease--mucin-type O-glycans in cancer. Adv Cancer Res 126:53-135
924 Kuhn HM, Meier-Dieter U, Mayer H (1988) ECA, the enterobacterial common antigen. FEMS
925 Microbiol Rev 4 (3):195-222
926 Kulp A, Kuehn MJ (2010) Biological functions and biogenesis of secreted bacterial outer
927 membrane vesicles. Annu Rev Microbiol 64:163-184
928 Ladhani SN (2012) Two decades of experience with the Haemophilus influenzae serotype b
929 conjugate vaccine in the United Kingdom. Clin Ther 34 (2):385-399

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

930 Lees A, Puvanesarajah V, Frasch CE (2008) Conjugation Chemistry. In: Siber GR, Klugman
931 KP, Makela PH (eds) Pneumococcal Vaccines: The Impact of Conjugate Vaccines. ASM
932 Press, Washington DC,
933 Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi
934 M, Wren BW (2005) Functional analysis of the Campylobacter jejuni N-linked protein
935 glycosylation pathway. Mol Microbiol 55 (6):1695-1703
936 Lithgow KV, Scott NE, Iwashkiw JA, Thomson EL, Foster LJ, Feldman MF, Dennis JJ (2014) A
937 general protein O-glycosylation system within the Burkholderia cepacia complex is
938 involved in motility and virulence. Mol Microbiol 92 (1):116-137
939 Liu D, Reeves PR (1994) Escherichia coli K12 regains its O antigen. Microbiology 140 (Pt
940 1):49-57
941 Liu L, Liu Y, Li J, Du G, Chen J (2011) Microbial production of hyaluronic acid: current state,
942 challenges, and perspectives. Microb Cell Fact 10:99
943 Livingston BD, Jacobs JL, Glick MC, Troy FA (1988) Extended polysialic acid chains (n greater
944 than 55) in glycoproteins from human neuroblastoma cells. J Biol Chem 263 (19):9443-
945 9448
946 Lizak C, Gerber S, Numao S, Aebi M, Locher KP (2011) X-ray structure of a bacterial
947 oligosaccharyltransferase. Nature 474 (7351):350-355
948 Lomino JV, Naegeli A, Orwenyo J, Amin MN, Aebi M, Wang LX (2013) A two-step enzymatic
949 glycosylation of polypeptides with complex N-glycans. Bioorg Med Chem 21 (8):2262-
950 2270
951 Ma Z, Zhang H, Shang W, Zhu F, Han W, Zhao X, Han D, Wang PG, Chen M (2014)
952 Glycoconjugate vaccine containing Escherichia coli O157:H7 O-antigen linked with
953 maltose-binding protein elicits humoral and cellular responses. PLoS One 9 (8):e105215

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

954 Mally M, Fontana C, Leibundgut-Landmann S, Laacisse L, Fan YY, Widmalm G, Aebi M (2013)
955 Glycoengineering of host mimicking type-2 LacNAc polymers and Lewis X antigens on
956 bacterial cell surfaces. Mol Microbiol 87 (1):112-131

957 McCann JR, St Geme JW, 3rd (2014) The HMW1C-like glycosyltransferases--an enzyme family
958 with a sweet tooth for simple sugars. PLoS Pathog 10 (4):e1003977

959 Melli LJ, Ciocchini AE, Caillava AJ, Vozza N, Chinen I, Rivas M, Feldman MF, Ugalde JE,
960 Comerci DJ (2015) Serogroup-specific bacterial engineered glycoproteins as novel
961 antigenic targets for diagnosis of shiga toxin-producing-escherichia coli-associated
962 hemolytic-uremic syndrome. J Clin Microbiol 53 (2):528-538

963 Merritt JH, Ollis AA, Fisher AC, DeLisa MP (2013) Glycans-by-design: engineering bacteria for
964 the biosynthesis of complex glycans and glycoconjugates. Biotechnol Bioeng 110
965 (6):1550-1564

966 Mills DC, Jervis AJ, Abouelhadid S, Yates LE, Cuccui J, Linton D, Wren BW (2016) Functional
967 analysis of N-linking oligosaccharyl transferase enzymes encoded by deep-sea vent
968 proteobacteria. Glycobiology 26 (4):398-409

969 Moe GR, Bhandari TS, Flitter BA (2009) Vaccines containing de-N-acetyl sialic acid elicit
970 antibodies protective against neisseria meningitidis groups B and C. J Immunol 182
971 (10):6610-6617

972 Muralinath M, Kuehn MJ, Roland KL, Curtiss R, 3rd (2011) Immunization with Salmonella
973 enterica serovar Typhimurium-derived outer membrane vesicles delivering the
974 pneumococcal protein PspA confers protection against challenge with Streptococcus
975 pneumoniae. Infect Immun 79 (2):887-894

976 Musumeci MA, Hug I, Scott NE, Ielmini MV, Foster LJ, Wang PG, Feldman MF (2013) In vitro
977 activity of Neisseria meningitidis PglL O-oligosaccharyltransferase with diverse synthetic
978 lipid donors and a UDP-activated sugar. J Biol Chem 288 (15):10578-10587

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

979 Naegeli A, Neupert C, Fan YY, Lin CW, Poljak K, Papini AM, Schwarz F, Aebi M (2014)
980 Molecular analysis of an alternative N-glycosylation machinery by functional transfer
981 from Actinobacillus pleuropneumoniae to Escherichia coli. J Biol Chem 289 (4):2170-
982 2179
983 National-Research-Council (2012) Transforming Glycoscience: A Roadmap for the Future. The
984 National Academies Press, Washington, DC
985 Nothaft H, Davis B, Lock YY, Perez-Munoz ME, Vinogradov E, Walter J, Coros C, Szymanski
986 CM (2016) Engineering the Campylobacter jejuni N-glycan to create an effective chicken
987 vaccine. Sci Rep 6:26511
988 Nothaft H, Scott NE, Vinogradov E, Liu X, Hu R, Beadle B, Fodor C, Miller WG, Li J, Cordwell
989 SJ, Szymanski CM (2012) Diversity in the protein N-glycosylation pathways within the
990 Campylobacter genus. Mol Cell Proteomics 11 (11):1203-1219
991 Nothaft H, Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. Nat Rev
992 Microbiol 8 (11):765-778
993 Ollis AA, Chai Y, Natarajan A, Perregaux E, Jaroentomeechai T, Guarino C, Smith J, Zhang S,
994 DeLisa MP (2015) Substitute sweeteners: diverse bacterial oligosaccharyltransferases
995 with unique N-glycosylation site preferences. Sci Rep 5:15237
996 Ollis AA, Zhang S, Fisher AC, DeLisa MP (2014) Engineered oligosaccharyltransferases with
997 greatly relaxed acceptor-site specificity. Nat Chem Biol 10 (10):816-822
998 Pan C, Sun P, Liu B, Liang H, Peng Z, Dong Y, Wang D, Liu X, Wang B, Zeng M, Wu J, Zhu L,
999 Wang H (2016) Biosynthesis of Conjugate Vaccines Using an O-Linked Glycosylation
1000 System. MBio 7 (2):e00443-00416
1001 Paton AW, Morona R, Paton JC (2000) A new biological agent for treatment of Shiga toxigenic
1002 Escherichia coli infections and dysentery in humans. Nat Med 6 (3):265-270

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1003 Power PM, Seib KL, Jennings MP (2006) Pilin glycosylation in Neisseria meningitidis occurs by
1004 a similar pathway to wzy-dependent O-antigen biosynthesis in Escherichia coli. Biochem
1005 Biophys Res Commun 347 (4):904-908
1006 Price NL, Goyette-Desjardins G, Nothaft H, Valguarnera E, Szymanski CM, Segura M, Feldman
1007 MF (2016) Glycoengineered Outer Membrane Vesicles: A Novel Platform for Bacterial
1008 Vaccines. Sci Rep 6:24931
1009 Priem B, Gilbert M, Wakarchuk WW, Heyraud A, Samain E (2002) A new fermentation process
1010 allows large-scale production of human milk oligosaccharides by metabolically
1011 engineered bacteria. Glycobiology 12 (4):235-240
1012 Qutyan M, Henkel M, Horzempa J, Quinn M, Castric P (2010) Glycosylation of pilin and nonpilin
1013 protein constructs by Pseudomonas aeruginosa 1244. J Bacteriol 192 (22):5972-5981
1014 Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. Annu Rev Biochem 71:635-700
1015 Ravenscroft N, Haeuptle MA, Kowarik M, Fernandez FS, Carranza P, Brunner A, Steffen M,
1016 Wetter M, Keller S, Ruch C, Wacker M (2016) Purification and characterization of a
1017 Shigella conjugate vaccine, produced by glycoengineering Escherichia coli. Glycobiology
1018 26 (1):51-62
1019 Riddle MS, Kaminski RW, Di Paolo C, Porter CK, Gutierrez RL, Clarkson KA, Weerts HE,
1020 Duplessis C, Castellano A, Alaimo C, Paolino K, Gormley R, Gambillara Fonck V (2016)
1021 Safety and immunogenicity of a candidate bioconjugate vaccine against Shigella flexneri
1022 2a administered to healthy adults: a single blind, randomized phase I study. Clin Vaccine
1023 Immunol
1024 Ruby T, McLaughlin L, Gopinath S, Monack D (2012) Salmonella's long-term relationship with
1025 its host. FEMS Microbiol Rev 36 (3):600-615
1026 Ruffing A, Chen RR (2006) Metabolic engineering of microbes for oligosaccharide and
1027 polysaccharide synthesis. Microb Cell Fact 5:25

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1028 Samain E, Drouillard S, Heyraud A, Driguez H, Geremia RA (1997) Gram-scale synthesis of
1029 recombinant chitooligosaccharides in Escherichia coli. Carbohydr Res 302 (1-2):35-42
1030 Sanders H, Feavers IM (2011) Adjuvant properties of meningococcal outer membrane vesicles
1031 and the use of adjuvants in Neisseria meningitidis protein vaccines. Expert Rev
1032 Vaccines 10 (3):323-334
1033 Santos-Silva T, Dias JM, Dolla A, Durand MC, Goncalves LL, Lampreia J, Moura I, Romao MJ
1034 (2007) Crystal structure of the 16 heme cytochrome from Desulfovibrio gigas: a
1035 glycosylated protein in a sulphate-reducing bacterium. J Mol Biol 370 (4):659-673
1036 Schild S, Nelson EJ, Camilli A (2008) Immunization with Vibrio cholerae outer membrane
1037 vesicles induces protective immunity in mice. Infect Immun 76 (10):4554-4563
1038 Schmid J, Sieber V, Rehm B (2015) Bacterial exopolysaccharides: biosynthesis pathways and
1039 engineering strategies. Front Microbiol 6:496
1040 Schwarz F, Fan YY, Schubert M, Aebi M (2011a) Cytoplasmic N-glycosyltransferase of
1041 Actinobacillus pleuropneumoniae is an inverting enzyme and recognizes the NX(S/T)
1042 consensus sequence. J Biol Chem 286 (40):35267-35274
1043 Schwarz F, Huang W, Li C, Schulz BL, Lizak C, Palumbo A, Numao S, Neri D, Aebi M, Wang
1044 LX (2010) A combined method for producing homogeneous glycoproteins with
1045 eukaryotic N-glycosylation. Nat Chem Biol 6 (4):264-266
1046 Schwarz F, Lizak C, Fan YY, Fleurkens S, Kowarik M, Aebi M (2011b) Relaxed acceptor site
1047 specificity of bacterial oligosaccharyltransferase in vivo. Glycobiology 21 (1):45-54
1048 Senchenkova SN, Guo X, Naumenko OI, Shashkov AS, Perepelov AV, Liu B, Knirel YA (2016)
1049 Structure and genetics of the O-antigens of Escherichia coli O182-O187. Carbohydr Res
1050 435:58-67
1051 Sethuraman N, Stadheim TA (2006) Challenges in therapeutic glycoprotein production. Curr
1052 Opin Biotechnol 17 (4):341-346

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1053 Seviour RJ, McNeil B, Fazenda ML, Harvey LM (2011) Operating bioreactors for microbial
1054 exopolysaccharide production. Crit Rev Biotechnol 31 (2):170-185
1055 Shang W, Zhai Y, Ma Z, Yang G, Ding Y, Han D, Li J, Zhang H, Liu J, Wang PG, Liu XW, Chen
1056 M (2016) Production of human blood group B antigen epitope conjugated protein in
1057 Escherichia coli and utilization of the adsorption blood group B antibody. Microb Cell
1058 Fact 15 (1):138
1059 Sleytr UB (1975) Heterologous reattachment of regular arrays of glycoproteins on bacterial
1060 surfaces. Nature 257 (5525):400-402
1061 Sleytr UB, Thorne KJ (1976) Chemical characterization of the regularly arranged surface layers
1062 of Clostridium thermosaccharolyticum and Clostridium thermohydrosulfuricum. J
1063 Bacteriol 126 (1):377-383
1064 Srivastava L, Tundup S, Choi BS, Norberg T, Harn D (2014) Immunomodulatory glycan lacto-N-
1065 fucopentaose III requires clathrin-mediated endocytosis to induce alternative activation
1066 of antigen-presenting cells. Infect Immun 82 (5):1891-1903
1067 Stenutz R, Weintraub A, Widmalm G (2006) The structures of Escherichia coli O-polysaccharide
1068 antigens. FEMS Microbiol Rev 30 (3):382-403
1069 Szymanski CM, Yao R, Ewing CP, Trust TJ, Guerry P (1999) Evidence for a system of general
1070 protein glycosylation in Campylobacter jejuni. Mol Microbiol 32 (5):1022-1030
1071 Thibault P, Logan SM, Kelly JF, Brisson JR, Ewing CP, Trust TJ, Guerry P (2001) Identification
1072 of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. J
1073 Biol Chem 276 (37):34862-34870
1074 Thonard JC, Migliore SA, Blustein R (1964) Isolation of Hyaluronic Acid from Broth Cultures of
1075 Streptococci. J Biol Chem 239:726-728

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1076 Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, Heiss C, Azadi
1077 P, Aebi M, DeLisa MP (2012) An engineered eukaryotic protein glycosylation pathway in
1078 Escherichia coli. Nat Chem Biol 8 (5):434-436
1079 Valentine JL, Chen L, Perregaux EC, Weyant KB, Rosenthal JA, Heiss C, Azadi P, Fisher AC,
1080 Putnam D, Moe GR, Merritt JH, DeLisa MP (2016) Immunization with Outer Membrane
1081 Vesicles Displaying Designer Glycotopes Yields Class-Switched, Glycan-Specific
1082 Antibodies. Cell Chem Biol 23 (6):655-665
1083 van den Dobbelsteen GP, Fae KC, Serroyen J, van den Nieuwenhof IM, Braun M, Haeuptle MA,
1084 Sirena D, Schneider J, Alaimo C, Lipowsky G, Gambillara-Fonck V, Wacker M, Poolman
1085 JT (2016) Immunogenicity and safety of a tetravalent E. coli O-antigen bioconjugate
1086 vaccine in animal models. Vaccine 34 (35):4152-4160
1087 van Die I, van Vliet SJ, Nyame AK, Cummings RD, Bank CM, Appelmelk B, Geijtenbeek TB,
1088 van Kooyk Y (2003) The dendritic cell-specific C-type lectin DC-SIGN is a receptor for
1089 Schistosoma mansoni egg antigens and recognizes the glycan antigen Lewis x.
1090 Glycobiology 13 (6):471-478
1091 Van Patten SM, Hughes H, Huff MR, Piepenhagen PA, Waire J, Qiu H, Ganesa C, Reczek D,
1092 Ward PV, Kutzko JP, Edmunds T (2007) Effect of mannose chain length on targeting of
1093 glucocerebrosidase for enzyme replacement therapy of Gaucher disease. Glycobiology
1094 17 (5):467-478
1095 Vella M, Pace D (2015) Glycoconjugate vaccines: an update. Expert Opin Biol Ther 15 (4):529-
1096 546
1097 Vik A, Aas FE, Anonsen JH, Bilsborough S, Schneider A, Egge-Jacobsen W, Koomey M (2009)
1098 Broad spectrum O-linked protein glycosylation in the human pathogen Neisseria
1099 gonorrhoeae. Proc Natl Acad Sci U S A 106 (11):4447-4452

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1100 Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines
1101 ED, Valvano MA, Whitfield C, Aebi M (2006) Substrate specificity of bacterial
1102 oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and
1103 eukaryotic systems. Proc Natl Acad Sci U S A 103 (18):7088-7093
1104 Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR,
1105 Dell A, Wren BW, Aebi M (2002) N-linked glycosylation in Campylobacter jejuni and its
1106 functional transfer into E. coli. Science 298 (5599):1790-1793
1107 Wacker M, Wang L, Kowarik M, Dowd M, Lipowsky G, Faridmoayer A, Shields K, Park S,
1108 Alaimo C, Kelley KA, Braun M, Quebatte J, Gambillara V, Carranza P, Steffen M, Lee JC
1109 (2014) Prevention of Staphylococcus aureus infections by glycoprotein vaccines
1110 synthesized in Escherichia coli. J Infect Dis 209 (10):1551-1561
1111 Wang L, Curd H, Reeves PR (1999) Immunization of mice with live oral vaccine based on a
1112 Salmonella enterica (sv Typhimurium) aroA strain expressing the Escherichia coli O111
1113 O antigen. Microb Pathog 27 (1):55-59
1114 Wetter M, Kowarik M, Steffen M, Carranza P, Corradin G, Wacker M (2013) Engineering,
1115 conjugation, and immunogenicity assessment of Escherichia coli O121 O antigen for its
1116 potential use as a typhoid vaccine component. Glycoconj J 30 (5):511-522
1117 Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in Escherichia coli.
1118 Annu Rev Biochem 75:39-68
1119 Whitfield C, Amor PA, Koplin R (1997) Modulation of the surface architecture of gram-negative
1120 bacteria by the action of surface polymer:lipid A-core ligase and by determinants of
1121 polymer chain length. Mol Microbiol 23 (4):629-638
1122 Whitney JC, Howell PL (2013) Synthase-dependent exopolysaccharide secretion in Gram-
1123 negative bacteria. Trends Microbiol 21 (2):63-72

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1124 Widner B, Behr R, Von Dollen S, Tang M, Heu T, Sloma A, Sternberg D, Deangelis PL, Weigel
1125 PH, Brown S (2005) Hyaluronic acid production in Bacillus subtilis. Appl Environ
1126 Microbiol 71 (7):3747-3752
1127 Xu DQ, Cisar JO, Osorio M, Wai TT, Kopecko DJ (2007) Core-linked LPS expression of
1128 Shigella dysenteriae serotype 1 O-antigen in live Salmonella Typhi vaccine vector
1129 Ty21a: preclinical evidence of immunogenicity and protection. Vaccine 25 (33):6167-
1130 6175
1131 Yadav V, Paniliatis BJ, Shi H, Lee K, Cebe P, Kaplan DL (2010) Novel in vivo-degradable
1132 cellulose-chitin copolymer from metabolically engineered Gluconacetobacter xylinus.
1133 Appl Environ Microbiol 76 (18):6257-6265
1134 Yavuz E, Maffioli C, Ilg K, Aebi M, Priem B (2011) Glycomimicry: display of fucosylation on the
1135 lipo-oligosaccharide of recombinant Escherichia coli K12. Glycoconj J 28 (1):39-47
1136 Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St
1137 Michael F, Aberg E, Szymanski CM (2002) Structure of the N-linked glycan present on
1138 multiple glycoproteins in the Gram-negative bacterium, Campylobacter jejuni. J Biol
1139 Chem 277 (45):42530-42539
1140 Yu H, Stephanopoulos G (2008) Metabolic engineering of Escherichia coli for biosynthesis of
1141 hyaluronic acid. Metab Eng 10 (1):24-32

1142

1143 **Figure 1. Biosynthesis of bacterial polysaccharides**. The majority of bacterial
1144 polysaccharides are assembled by one of three mechanisms, the Wzy-dependent, the ABC
1145 transporter-dependent or the synthase-dependent pathway. The key protein components for
1146 each mechanism are indicated on the diagram, and are located in the inner membrane of Gram-
1147 negative organisms or the membrane of Gram-positive organisms. Polysaccharides are
1148 synthesized from nucleotide diphosphate (NDP) sugars. For the Wzy-dependent pathway,

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1149 multiple glycosyltransferases (GTases) in the cytoplasm synthesize oligosaccharides on Und-
1150 PP. Oligosaccharides typically contain diverse monosaccharides and may be branched;
1151 consequently this assembly mechanism is responsible for the production of most high-
1152 complexity sugars. The completed oligosaccharide repeat unit is transported across the relevant
1153 membrane by the translocase or flippase enzyme Wzx. Multiple repeat units are then linked
1154 together by the polymerase enzyme Wzy to form a repeating heteropolymer. The final length of
1155 the polymer may be controlled by the chain length regulator Wzz. In the ABC transporter-
1156 dependent pathway, a homopolymer or simple heteropolymer is assembled on Und-PP on the
1157 cytoplasmic face of the membrane, often by just a single GTase. The completed polysaccharide
1158 is capped with a moiety such as a phosphate group, and transported through the membrane by
1159 the ATP-binding cassette (ABC) transporter. For synthase-dependent biosynthesis, the
1160 polysaccharide is simultaneously polymerized and transported across the membrane. In the
1161 absence of a membrane anchor, a receptor protein for a signaling molecule such as bis-(3′-5′)-
1162 cyclic dimeric guanosine monophosphate (c-di-GMP) may play a role in initiation of
1163 polysaccharide assembly. In Gram-negative organisms, polysaccharides are frequently
1164 transported across the outer membrane by an additional export system to enable surface
1165 display.

1166 **Figure 2. LPS biosynthesis**. Multiple glycosyltransferases (GTases) transfer NDP-sugars to
1167 the nascent oligosaccharide to form an O-antigen repeat unit on Und-PP in the cytoplasm of the
1168 cell. The completed oligosaccharide is transported across the inner membrane by the flippase,
1169 Wzx, and multiple repeat units are joined together by the polymerase enzyme Wzy to form the
1170 completed O-antigen portion of the LPS. The structure shown is representative and does not
1171 indicate a specific serotype. Simultaneously, the LPS core, comprising lipid A, and the inner and
1172 outer core sugars is assembled in the cytoplasm. The inner core consists of two or three Kdo
1173 monosaccharides (diamond shapes), which are added during synthesis of the lipid A molecule,

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1174 and three heptose sugars (heptagon shapes) which are added by the sequential action of three
1175 GTases. The outer core shown is an R1 structure, consisting of three glucose and two
1176 galactose residues (hexagons), and is assembled by the sequential action of a further five
1177 GTases. The completed LPS core is transported across the inner membrane by the ABC
1178 transporter MsbA. The O-antigen repeat unit is removed from the Und-PP membrane anchor
1179 and attached to the first galactose on the R1 outer core by the ligase enzyme WaaL. The entire
1180 LPS structure is then extracted from the inner membrane and transported across the periplasm
1181 and through the outer membrane to the extracellular face by the Lpt protein complex, where
1182 lipid A becomes a component of the outer face of the outer membrane with the polysaccharide
1183 displayed on the surface of the cell.

1184 **Figure 3. Recombinant protein glycosylation in *E. coli* using the bacterial**
1185 **oligosaccharyltransferase.** Co-expression of three components is required for recombinant
1186 glycosylation in *E. coli*: (1) The glycan biosynthetic locus for the production of the carbohydrate
1187 of interest on the lipid carrier undecaprenol pyrophosphate; (2) the oligosaccharyltransferase
1188 (*e.g.*, *Cj*PglB); and (3) the acceptor protein of interest that has been engineered with a signal
1189 peptide for export into the periplasm and an acceptor sequon (*e.g.*, D/E-X_1-N-X_2-S/T, where X
1190 can be any amino acid except proline) for glycosylation by the oligosaccharyltransferase.
1191 Sequons can be engineered into an exposed, flexible loop or at either the N- or the C-terminus
1192 of the protein. The glycoprotein can then be purified from the bacterial cells using standard
1193 methods.

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

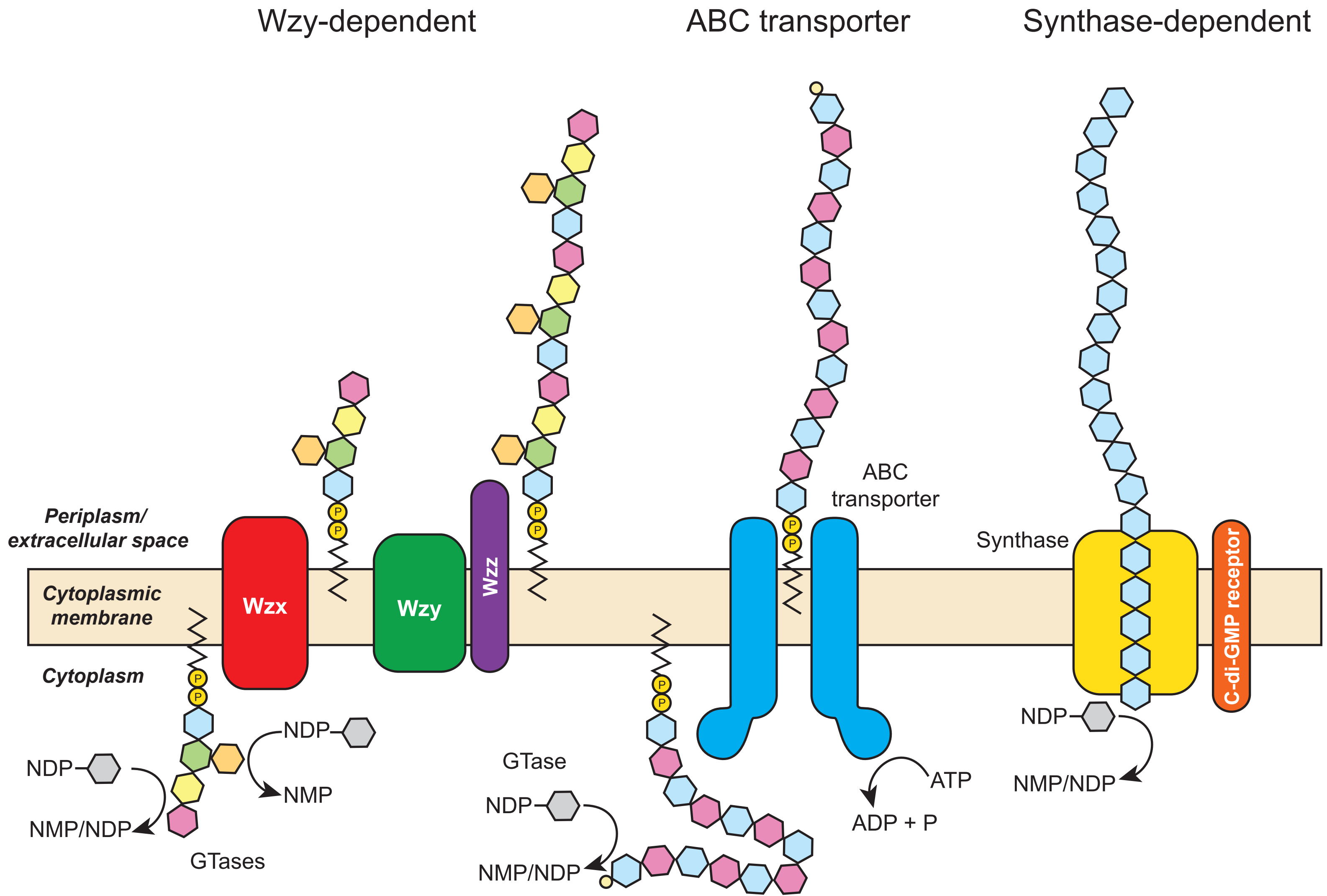

Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

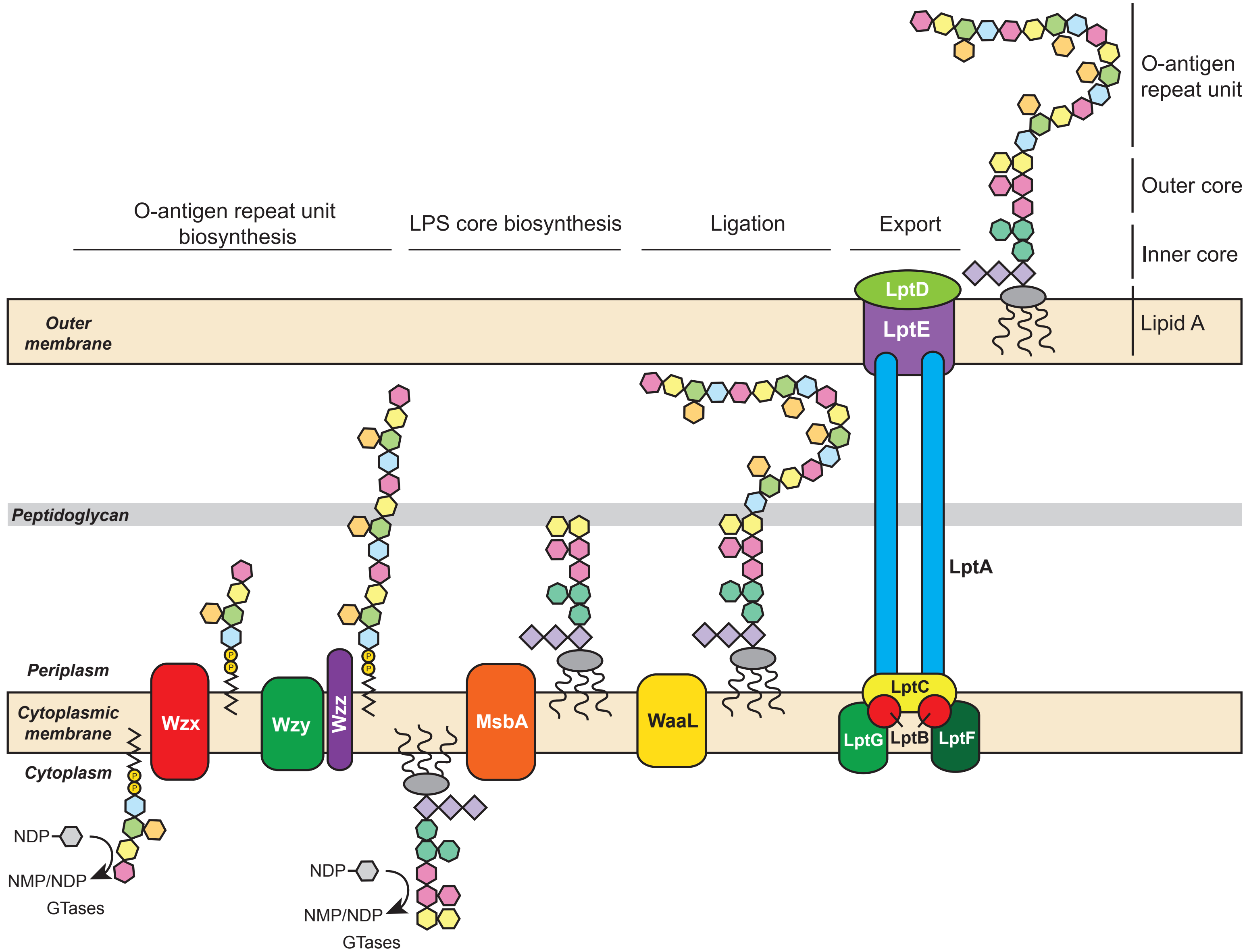

Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/118224; this version posted March 18, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

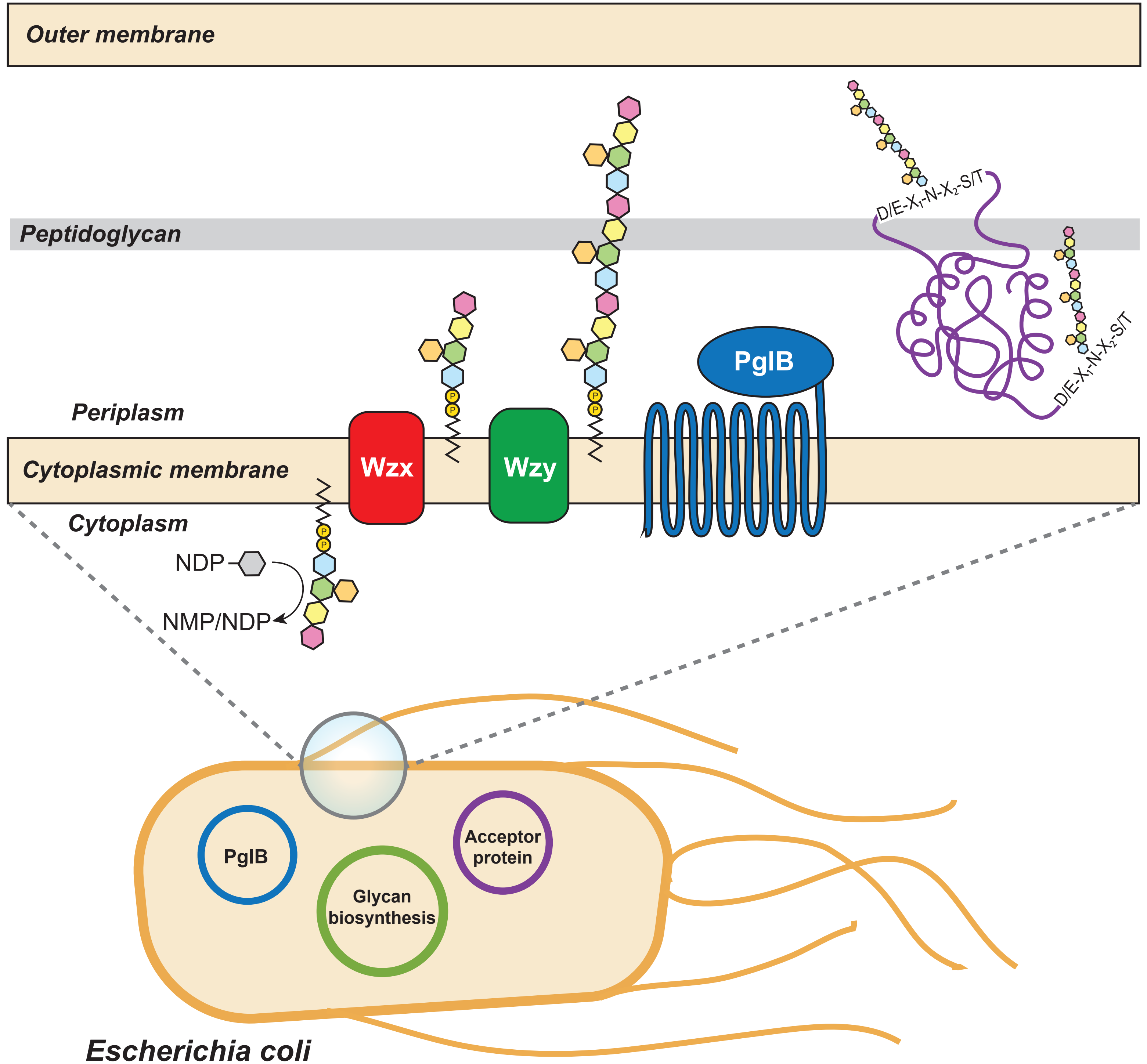

Figure 3