Rationally Designed Short Polyisoprenol-Linked PglB Substrates for Engineered Polypeptide and Protein N-Glycosylation

Feng Liu,†,§ Balakumar Vijayakrishnan,†,§ Amirreza Faridmoayer,‡,§ Thomas A. Taylor,†
Thomas B. Parsons,† Gonçalo J.L. Bernardes,† Michael Kowarik,‡ and Benjamin G. Davis*†

†Department of Chemistry, Chemistry Research Laboratory, Oxford University, Oxford OX1 3TA United Kingdom
‡GlycoVaxyn AG, 8952 Schlieren, Switzerland

ABSTRACT: The lipid carrier specificity of the protein N-glycosylation enzyme C. jejuni PglB was tested using a logical, synthetic array of natural and unnatural C10, C20, C30, and C40 polyisoprenol sugar pyrophosphates, including those bearing repeating cis-prenyl units. Unusual, short, synthetically accessible C20 prenols (nerylnerol 1d and geranylnerol 1e) were shown to be effective lipid carriers for PglB sugar substrates. Kinetic analyses for PglB revealed clear Kcat-only modulation with lipid chain length, thereby implicating successful in vitro application at appropriate concentrations. This was confirmed by optimized, efficient in vitro synthesis allowing >90% of Asn-linked β-N-GlcNAc-ylated peptide and proteins. This reveals a simple, flexible biocatalytic method for glycoconjugate synthesis using PglB N-glycosylation machinery and varied chemically synthesized glycosylation donor precursors.

Figure 1. Bacterial N-linked glycosylation and designed unnatural candidate polyisoprenols 1a−g as alternative short lipid carriers for PglB-catalyzed glycosylation. The polyisoprenols were synthesized from building blocks 2a−d and 3a and 3b.

as the fact that C. jejuni PglB can be readily overexpressed in functional form in Escherichia coli,16 highlight PglB’s potential as a synthetic biocatalyst. They suggest it as a potentially ideal model from which to generate a ready synthetic system for in vitro protein glycosylation. However, the donor substrates used normally by PglB in vivo (C55 lipid pyrophosphoryl-linked oligosaccharides containing the rare, bacterial sugar bacillosamine (Bac)) would restrict this system (both in substrate accessibility and product relevance).

In an attempt to optimize the PglB protein N-glycosylation platform for practical, synthetic (and hence in vitro) use, we designed an array of chemically generated polyisoprenol variants to find those simpler and shorter than the natural undecaprenol (Figure 1) and that might serve as alternative lipid carriers that could be recognized by PglB. Insightful prior work has elucidated some aspects of the polyisoprenol specificity of PglB.17,18 However, estimated conversions for these reactions were ≤20%. In addition, many of these prior substrates were prepared enzymatically from polyisoprenols isolated from natural sources, enabling only nmol analysis of lipid pyrophosphates containing the natural Campylobacter (GalNAc-GalNAc-BacdiNAc) glycoconjugate synthesis using PglB N-glycosylation enzyme (Bac) would restrict this system (both in substrate accessibility and product relevance).

Published: December 16, 2013
Received: September 11, 2013

© 2013 American Chemical Society

dx.doi.org/10.1021/ja409409h | J. Am. Chem. Soc. 2014, 136, 566−569
cans. This revealed tantalizing activity for two shorter (mixed cis/trans prenol-9 and prenol-8) variants. However, despite some individual pioneering examples,9–28 synthetic access to other prenols has been rare, and homologated families of lipids have not yet been probed. Therefore, with the intention of probing fuller lipid and sugar substrate plasticity and breadth in PglB and with the intention of creating synthetically accessible substrates, we first systematically varied the lipid carriers (Figure 1).

As a starting point for lipid variation, we speculated that the use of repeating cis-prenyl units at the hydroxyl terminus would confer binding affinity due to a resemblance to the alcohol-terminus that bears the glycan-pyrophosphate found in the natural substrate and likely enters a binding pocket in PglB.29 This notion is supported by a crystal structure which has a Mg2+ bound at the active site.30 Subsequent modeling suggested that the lipid carrier would likely locate first in a narrow pocket that would tightly accommodate two isoprenyl units (and thus impose tighter stereocchemical requirements) and then along a broader hydrophobic groove with potentially more relaxed requirements (Figure S12). Thus, compounds 1a–g, all bearing one or more terminal repeating cis-prenyl units were designed and synthesized as primary candidate lipid carriers (Figure 1). In brief (see SI for further details), compounds 1a–f were synthesized from head (compound 3a and 3b) and tail building blocks (compound 2a–d). The requisite building blocks were prepared and elongated using coupling between appropriate sulfones and allyl chloride. Utilization of a convergent synthetic strategy in the case of long lipids (Scheme 1) allowed useful flexibility in the generation of lipids with different stereochernistry. Removal of sulfonic groups in one step using LiEt3BH/(dppp)PdCl2 valuably reduced both the total number of synthetic steps and the formation of isomeric products (different stereo-chemistry). The resulting, pure synthetic polyprenols were shown to have a d.r. (>95%) according to the characteristic allylic methyl group signals in 1H NMR (see Figure S1); no unwanted isomeric product (<2%, 1H NMR signals at δ 2.7, 5.9 ppm)30 was detected.

These polyprenols were readily phosphorylated using mono-(tetrabutyrammonium) phosphate (TBAP) activated with trichloroacetonitrile (TCA).31 We next attached just a single residue atypical glycan to provide a stringent test of the catalytic activity of PglB. Importantly, we used the sugar that would be found as the first N-linked residue in eukaryotic glycoproteins (but not prokaryotic and so not that normally employed by Campylobacter): GlcNAc. The resulting polyprenyl monophosphates 5a–g were therefore coupled to 2-acetamido-3,4,6-
tri-O-acetyl-2-deoxy-α-D-glucopyranose 1-phosphate (6) using carbonyldiimidazole (CDI)-mediated phosphoesterification.32,33 Deacetylation with catalytic NaOMe in MeOH (Zemplén conditions) gave the lipid pyrophosphate-linked saccharides (LPPS) 8a–g in 9–38% yield over three steps (Scheme 2).

With these first synthetic LPPS's in hand, 8a–h were then tested as substrates of PglB. The glycosylation of peptides bearing the required D/E-X-N-X-S/T consensus motif was determined by examining the increase of molecular weight (16%/6 M urea tricine-SDS-PAGE)34 that corresponds to the addition of GlcNAc (Figure 2). An extract from E. coli LPPS containing C. jejuni heptasaccharide-linked undecaprenyl pyrophosphate was used as a positive control. Excitingly, 8a–e were active substrates for PglB-catalyzed N-glycosylation of the fluorescent peptide Tamra-DANYTK as indicated by the appearance of glycopeptide product bands with higher molecular weight consistent with the addition of GlcNAc. This importantly revealed that the lipid carrier for the substrate of PglB can contain as few as two cis-head repeating units. The activity of 8c and 8e as substrates for PglB also indicated that PglB can tolerate lipid carriers bearing trans geometry close to the hydroxyl terminus. However, activity was lost as the length of lipid became shorter than four prenyl units: nerol (1g)-PP-GlcNAc (8g), cis,cis-farnesol (1f)-PP-GlcNAc conjugate (8f), and citronellol-PP-GlcNAc conjugate (8h).32,33 did not show any detectable activity.
in the *in vitro* assay. These activities ‘mapped’ tight proximal-site activity and relaxed distal-site activity consistent with model shown in Figure S12.

These first kinetic parameters (Table 1) for PglB suggested key features. In particular, the variation of activity with lipid length in the substrate is strikingly only dependent on $K_{cat}$, $k_{cat}$ remains essentially unaltered. This suggests that the lipid may not play a primary role in catalytic turnover but is a key regulator of substrate uptake. This suggested too that in *in vitro* reactions conducted at sufficiently high concentrations >$K_M$ would allow transfer efficiencies equal to those found for full length lipid substrates. This was valuably confirmed in synthetic reactions that allowed the synthesis of GlcNAc-ylated glycopeptide in yields >90% using 0.1 mM glycosyl donor substrate with 20 μM of acceptor peptide. These >90% reactions usefully extend the synthetic utility of PglB.

Having elucidated valuable plasticity toward unnatural lipid-variant substrates, we next examined glycan breadth beyond the atypical monosaccharide GlcNAc already demonstrated. Conjugates (13 and 14) that contain both unnatural sugar and lipid carrier (Scheme 2) were prepared by coupling nerylnerol (1d) and nerylnerol (1d) conjugates were much better substrates than those of longer lipids bearing repeating cis units.24,41,42 Investigation of the crystal structures of MurG and/or PglB with lipid carrier bound, once available, would shed light on these clear differences in lipid carrier specificities.

The discovery of the breadth of PglB and these accessible lipid carriers now effectively enables the synthesis of lipid-pyrophosphate-linked substrates suitable for the *in vitro* generation of tailor-made glycoproteins. Importantly, to our knowledge, this currently represents the only *in vitro* biocatalytic system for the formation of the vital GlcNAc-β-1,2-N-Asn linkage (and importantly can be driven to >95% on proteins, here for AcrA); the recently discovered *Methanococcus* AglB, e.g.,11 does

![Figure 2. Peptide and protein glycosylation. (a) Fluorescent electrophoretic analysis of peptide glycosylation with lipids (8a–h) (− = no lipidd + = lipid extracted and enriched from E. coli cells producing C. jejuni heptasaccharide-linked undecaprenyl pyrophosphate); [glycolipid] = 57 [peptide] 20 μM; [PglB] = 0.44 μM. (b) ES-MS of in *vitro* N-linked protein (AcrA) glycosylation with 8c; >95% diglycosylation.

![Figure 2. Peptide and protein glycosylation. (a) Fluorescent electrophoretic analysis of peptide glycosylation with lipids (8a–h) (− = no lipidd + = lipid extracted and enriched from E. coli cells producing C. jejuni heptasaccharide-linked undecaprenyl pyrophosphate); [glycolipid] = 57 [peptide] 20 μM; [PglB] = 0.44 μM. (b) ES-MS of in *vitro* N-linked protein (AcrA) glycosylation with 8c; >95% diglycosylation.

Table 1. Kinetic Parameters for PglB with GlcNAc Lipids

<table>
<thead>
<tr>
<th>glycolipid substrate</th>
<th>$k_{cat}$ [min$^{-1}$]</th>
<th>$K_M$ [μM]</th>
<th>$k_{cat}/K_M$ [min$^{-1}$·μM$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>8d GlcNAc-PP-oZ$_2$</td>
<td>0.0234 ± 0.0021</td>
<td>0.077 ± 0.016</td>
<td>0.30</td>
</tr>
<tr>
<td>8c GlcNAc-PP-oE$_3$</td>
<td>0.0231 ± 0.0016</td>
<td>0.055 ± 0.01</td>
<td>0.42</td>
</tr>
<tr>
<td>8a GlcNAc-PP-preno$_6$</td>
<td>0.0225 ± 0.0021</td>
<td>0.034 ± 0.01</td>
<td>0.66</td>
</tr>
</tbody>
</table>

*UPLC using TAMRA-fluorescence intensity; substrate concentrations: [glycolipid] = 1, 5, 10, 50, 100, 200 μM, [peptide] = 20 μM; enzyme concentration [PglB] = 0.44 μM; reaction time <2h, 30 °C; all conducted in duplicate.*
not transfer GlcNAc and requires an unusual disaccharide. This discovery complements prior GalNAc transfer and confirms a predicted activity.\textsuperscript{18} Notably, WeCA\textsubscript{33} is the enzyme that would generate PP-linked glycolipid substrates for PglB, is membrane-associated and cannot be readily exploited in vitro. In vitro biocatalytic installation of GlcNAc, shown here, also creates a useful precursor sugar site for carbohydrate-processing enzyme-mediated extension, as shown here (Scheme 2). The >90% in vitro efficiencies shown here therefore make PglB a highly viable synthetic biocatalyst for varied glycopeptides, coupled with substrate accessibility and potential for further enzymatic transformation.

\section*{ASSOCIATED CONTENT}
\textsection Supporting Information
Full procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

\section*{AUTHOR INFORMATION}
\textsection Corresponding Author
Ben.Davis@chem.ox.ac.uk

\textsection Author Contributions
These authors contributed equally.

\section*{Notes}
The authors declare the following competing financial interest(s): M.K. and A.F. are employees of Glycovaxyn. A patent has been filed and will afford inventors royalties, if licensed, in line with university guidelines.

\section*{ACKNOWLEDGMENTS}
We thank the BBSRC (T.A.T.), Glycovaxyn (F.L., B.V., B.G.D.) for funding, Drs P. Carranza, C. Bich for method development and glycopeptide analysis, and Drs L. Gong, J. Wickens for HRMS analysis. B.G.D. is a Royal Society Wolfson Merit Research Award Recipient.

\section*{REFERENCES}
(1) Spiro, R. G. Glycobiology 2002, 12, 43R.
(3) Ohtsubo, K.; Marth, J. D. Cell 2006, 126, 855.
(24) Schagger, H. Nat. Protoc. 2006, 1, 16.