From Peptide to Protein: Comparative Analysis of the Substrate Specificity of N-Linked Glycosylation in *C. jejuni*†

Mark M. Chen, Kerney Jebrell Glover, and Barbara Imperiali*

Departments of Chemistry and Biology, Massachusetts Institute of Technology 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Received December 22, 2006; Revised Manuscript Received March 15, 2007

ABSTRACT: The gram-negative bacterium *Campylobacter jejuni* was recently discovered to contain a general N-linked protein glycosylation pathway. Central to this pathway is PglB, a homologue of the Stt3p subunit of the eukaryotic oligosaccharyl transferase (OT), which is involved in the transfer of an oligosaccharide from a polyisoprenyl pyrophosphate carrier to the asparagine side chain of proteins within the conserved glycosylation sites D/E-X_1-**N**-X_2-S/T, where X_1 and X_2 can be any amino acids except proline. Using a library of peptide substrates and a quantitative radioactivity-based in vitro assay, we assessed the amino acids at each position of the consensus glycosylation sequence for their impact on glycosylation efficiency, whereby the sequence DQNAT was found to be the optimal acceptor substrate. In the context of a full-length folded protein, the differences between variations of the glycosylation sequences were found to be consistent with the trends observed from their peptidyl counterparts, though less dramatic because of additional influences. In addition to characterizing the acceptor preferences of PglB, we also assessed the selectivity toward the glycan donor. Interestingly, despite recent reports of relaxed selectivity toward the glycan donor, PglB was not found to be capable of utilizing glycosyl donors such as dolichyl-pyrophosphate-chitobiose, which is the minimum substrate for the eukaryotic OT process.

N-linked glycosylation is an essential protein modification common to all eukaryotic cells. This process involves the transfer of a preassembled oligosaccharide from a polyisoprenyl pyrophosphate carrier to an asparagine side chain within a consensus sequence of the protein acceptor (*1*, *2*). This remarkable reaction, catalyzed by a multimeric protein complex known as the oligosaccharyl transferase (OT[1]), occurs co-translationally when a nascent polypeptide enters the lumen of the endoplasmic reticulum (*3*). The consensus glycosylation sequence in eukaryotic cells has been identified to be **N**-X-S/T, where X can be any amino acid except proline (*4*, *5*). The efficiency of glycosylation can be influenced by the choice of the hydroxyamino acid S/T, the amino acid in the X-position, the amino acids directly adjacent to the glycosylation site, the proximity to the C-terminus as well as the local structural conformation near the glycosylation site (*6*). As a result, not all potential sites are glycosylated (*7*).

Recently, the gram-negative bacterium *Campylobacter jejuni* (*C. jejuni*) was found to contain a general N-linked protein glycosylation system (Figure 1) that shows homology to the eukaryotic pathway (*8*, *9*). Central to this pathway is PglB, an 82 kDa integral membrane protein with significant sequence homology to the Stt3p subunit of the eukaryotic OT (*10*, *11*), which catalyzes the transfer of the assembled oligosaccharide from the undecaprenyl pyrophosphate carrier to the polypeptide acceptor. We previously demonstrated the oligosaccharyl transferase activity of PglB in vitro using a synthetic disaccharide glycosyl donor and a short peptide acceptor (*12*) (Figure 2). Although PglB is capable of glycosylating the peptide Bz-N-L-T-NH_2 in vitro, efficient glycosylation in vivo requires an extended consensus sequence of D/E-X_1-**N**-X_2-S/T, where X_1 and X_2 can be any amino acid except proline (*13*). However, little is known about the acceptor preferences for this prokaryotic OT, which is crucial for understanding the factors that govern glycan site occupancy as well as the engineering of recombinant glycoproteins.

In this study, we synthesized and evaluated a library of peptidyl acceptor substrates for PglB using a quantitative radioactivity-based in vitro assay. Specifically, we compared the preference of PglB for the acidic residues at the −2 position, the hydroxyamino acid residues at the +2 position, and various amino acids in the X_1 and X_2 positions (Figure 2). These comparisons are difficult to address quantitatively using in vivo assays. Furthermore, in order to demonstrate the relevance of the observed peptide-based trends in the context of a full-length folded protein, we have engineered a set of glycosylation consensus sequences into a well-

† This work was supported by National Institutes of Health Grant GM039334 to B.I. and Postdoctoral Fellowship GM65699 to K.J.G.

* To whom correspondence should be addressed. Phone: (617) 253-1838. Fax: (617) 452-2419. E-mail: imper@mit.edu.

[1] Abbreviations: Ac, acetyl; Bac, *N*,*N*′-diacetyl-bacillosamine; Bz, benzoyl; CD, circular dichroism; Dap, 2,3-diaminopropionic acid; DMSO, dimethylsulfoxide; dpm, disintegration per minute; ESI-MS, electrospray ionization mass spectrometry; GalNAc, *N*-acetyl-galactosamine; GlcNAc, *N*-acetyl-glucosamine; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HPLC, high performance liquid chromatography; Hse, homoserine; OT, oligosaccharyl transferase; *p*NF, *para*-nitrophenylalanine; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; UDP, uridine diphosphate; UV, ultraviolet.

10.1021/bi602633n CCC: $37.00
Published on Web 04/17/2007

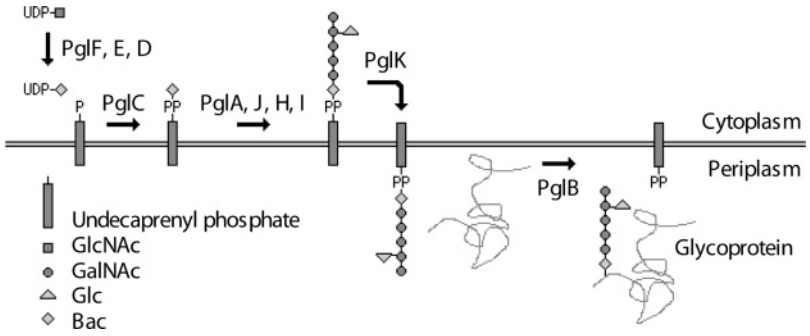
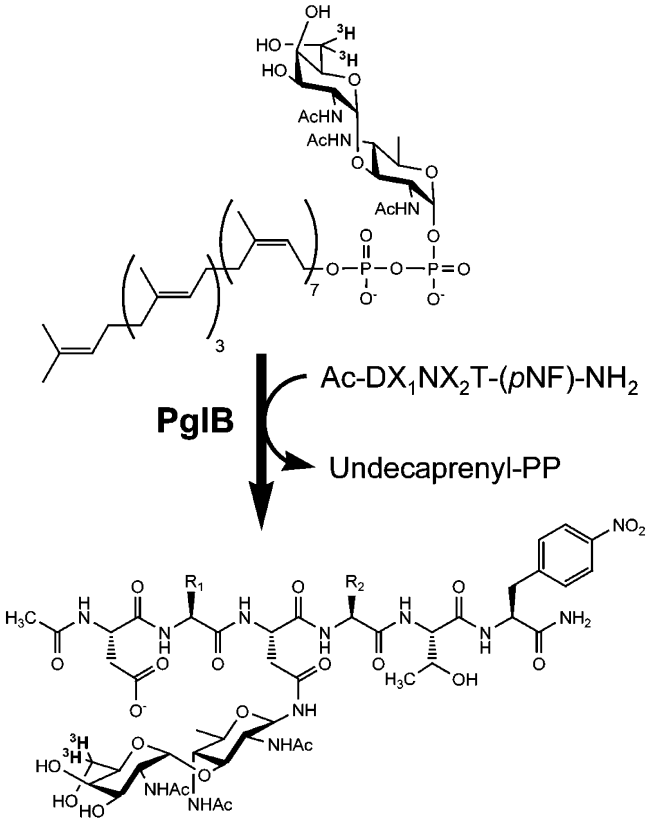


FIGURE 1: *C. jejuni* N-linked glycosylation pathway.



Glycosylation Consensus Sequence

Eukaryotic - **N** - X - S/T -

C. jejuni - D/E - X_1 - **N** - X_2 - S/T -
-2 -1 0 +1 +2

where X ≠ P

FIGURE 2: PglB glycosylation reaction using synthetic radiolabeled undecaprenyl pyrophosphate-linked disaccharide donor and peptide acceptor.

defined loop region within Im7, a non-*C. jejuni* model protein system. Finally, we evaluated the reactivity of PglB toward the eukaryotic glycosyl donor dolichyl-pyrophosphate-chitobiose.

MATERIALS AND METHODS

Peptide Synthesis. All peptides were synthesized by automated peptide synthesis on an ABI 431A peptide synthesizer (Applied Biosystems) using standard Fmoc-based peptide synthesis conditions on PAL-PEG-PS resin. Each peptide was acetylated at the N-terminus and includes a *para*-nitrophenylalanine (*p*NF) at the C-terminus. The peptides were cleaved from the resin using a trifluoroacetic acid cocktail containing 2.5% water and 2.5% triisopropylsilane, purified to ≥95% purity by preparative reverse-phase HPLC using a standard water/acetonitrile gradient, and quantified using the UV absorbance of the *p*NF amino acid at 280 nm (ϵ = 12,500 M^{-1} cm^{-1}).

Cloning and Mutagenesis of Im7. A pTrc (Im7) vector was provided by Professor Sheena Radford from the University of Leeds, encoding the native Im7 protein with an N-terminal hexa-histidine tag described elsewhere (*14*−*16*). Introduction of point mutations in Im7 was performed using the Quikchange mutagenesis kit (Stratagene). All mutants were sequenced to ensure that the gene contained the desired change.

Expression and Purification of Im7. Im7 mutants were expressed from *E. coli* and sequentially purified by Ni-affinity and gel filtration chromatography. Final samples used were confirmed by ESI-MS to be within 1 Da of expected mass and ≥95% pure by Coomassie blue staining of SDS−PAGE. Details of the expression and purification procedure can be found in Supporting Information.

Synthesis of Radioactive Disaccharide Donor. Radiolabeled [^{3}H]GalNAc-(α-1,4)-Bac-(α-1)-PP-undecaprenyl was synthesized following a protocol previously described (*12*). Briefly, chemically synthesized UDP-Bac and undecaprenyl phosphate, and commercially available [^{3}H]UDP-GalNAc were incubated with the *C. jejuni* glycosyltransferases PglC and PglA. The radiolabeled disaccharide product was extracted into a 2:1 mixture of chloroform/methanol, aliquoted, and dried into 1.5 mL tubes, each with approximately 0.3 nmol of disaccharide product (specific activity = 356,000 dpm/nmol, ∼100,000 dpm/tube).

Glycosylation of Peptide Substrates Using PglB. To a tube containing 0.3 nmol of dried radiolabeled [^{3}H]GalNAc-Bac-PP-undecaprenyl was added 10 μL of DMSO. Following vigorous vortexing and sonication (water bath) to resuspend the isoprene-based substrate, 100 μL of 2× assay buffer (280 mM sucrose, 2.4% Triton X-100 (v/v), and 280 mM HEPES at pH 7.5), 2 μL of 1 M $MnCl_2$, and 6 μL of PglB membrane fraction containing approximately 50 ng of enzyme were added and the volume increased to 190 μL with water. Reactions were initiated by the addition of 10 μL of peptide substrate dissolved in DMSO. Aliquots (35 μL) of the reaction mixture were removed at 4 min time intervals up to 20 min and quenched into 1 mL of 3:2 chloroform/

methanol + 200 μL of 4 mM $MgCl_2$. The aqueous layer was extracted, and the organic layer was washed twice with 300 μL of pure solvent upper phase. The aqueous layers were combined, mixed with 5 mL of EcoLite scintillation fluid (MP Biomedicals), and subjected to scintillation counting. Eight reactions were set up in parallel in each assay, and all assays were carried out in duplicate or better.

Glycosylation of Protein Substrates Using PglB. The reaction protocol is similar to the above procedure for peptide glycosylation, with the following modification: 50 μL of PglB membrane fraction was used to glycosylate 13 μL of a 0.75 mM stock of Im7 protein acceptor in buffer P. The final concentrations for the protein and the glycan were 50 and 1.5 μM, respectively. The reaction was left shaking at room temperature for 3 h before being spun at high speed to pellet the membrane fraction. The supernatant was applied to a Ni-NTA spin cartridge (Qiagen), washed, and eluted according to the manufacturer's instructions. A 100 μL aliquot of the elution fraction was injected onto an analytical reverse-phase C18 HPLC column and eluted under a standard water/acetonitrile gradient. Fractions were collected every minute, mixed with EcoLite scintillation fluid (MP Biomedicals), and subjected to scintillation counting. Experiments using eukaryotic glycosyl donors were carried out using the same general procedure. For mass spectral analysis, a 5 nmol aliquot of unlabeled sugar donor was used as the glycosyl donor.

Circular Dichroism Spectra of Acceptor Proteins. Far-UV CD spectra were acquired on an Aviv Model 202 CD Spectrometer (Aviv Biomedical) using a 1 mm path length cell and a protein concentration of 250 μg/mL in buffer C (50 mM sodium phosphate and 400 mM sodium sulfate) at 25 °C.

RESULTS

Glycosylation of Peptide Acceptors by Pg1B. A library of hexapeptides, based on the known glycosylation site ^{88}DFNVS^{92} from the *C. jejuni* glycoprotein PEB3, was synthesized (*17*, *18*). Each peptide differs at a single amino acid in order to systematically evaluate the preference of PglB for each residue within glycosylation sequence. A *para*-nitrophenylalanine (*p*NF) was appended to the C-terminus of each peptide to facilitate accurate peptide concentration determination by UV spectroscopy. When these peptides were assayed with PglB, the kinetic parameters indicated that PglB has a significant preference for threonine as the hydroxyamino acid at position +2 and aspartic acid as the acidic amino acid at position −2 (Table 1). Interestingly, peptides containing cysteine, homoserine, and diaminopropionic acid in place of serine at the +2 position were not glycosylated by PglB and could not inhibit the glycosylation of other acceptor peptides (data not shown). Although PglB is known to glycosylate sequences containing any amino acid in the +1 and −1 positions except proline, clear preferences for specific residues at these positions were observed (Figure 3). Alanine, serine, and positively charged residues appear to be favored at +1, whereas aromatic and amido residues were favored at −1. Like the eukaryotic OT, PglB also demonstrates binding determinants beyond the minimal consensus sequence (*6*). The extended PEB3 peptide Ac-GKDFNVSKI-(*p*NF)-NH_2 was found to have a lower $K_{m(app)}$ than that of the corresponding hexapeptide (Table 1). However, the *p*NF residue being in a different a position relative to the rest of the peptides within the library precludes direct comparisons.

Table 1: Kinetic Parameters of PglB Peptide Substrates[a]

peptide	$V_{max(app)}$ nM/min	$K_{m(app)}$ μM	$V_{max(app)}/K_{m(app)}$ $min^{-1} \times 10^3$
Ac-DFNVA-(*p*NF)-NH_2		no activity	
Ac-DFNVT-(*p*NF)-NH_2	34.6 ± 0.5	1.22 ± 0.07	28.4 ± 1.80
Ac-DFNVS-(*p*NF)-NH_2	34.6 ± 1.2	3.83 ± 0.55	9.03 ± 1.45
Ac-AFNVT-(*p*NF)-NH_2		no activity	
Ac-EFNVT-(*p*NF)-NH_2	6.9 ± 0.3	23.3 ± 2.98	0.30 ± 0.04
Ac-EFNVS-(*p*NF)-NH_2			
Ac-GKDFNVSKI-(*p*NF)-NH_2	11.3 ± 0.3	1.17 ± 0.14	9.67 ± 1.26
Ac-DFNVC-(*p*NF)-NH_2		no activity	
Ac-DFNV-(Hse)-(*p*NF)-NH_2		no activity	
Ac-DFNV-(Dap)-(*p*NF)-NH_2		no activity	

[a] Unnatural amino acids: *p*NF = *para*-nitrophenylalanine, Hse = homoserine, Dap = 2,3-diaminopropionic acid.

Preparation of Im7 Mutants. In these studies, the small bacterial protein Im7 (*14*−*16*) was used as a model system for the incorporation of selected N-linked glycosylation sequences. Specifically, a *C. jejuni* glycosylation consensus sequence of the form D/E-X_1-**N**-X_2-S/T (*13*) was incorporated within the protein by making the point mutations N26D and A28N. This resulted in the formation of a glycosylation consensus sequence ^{26}DVNAT^{30} within the loop region between the first and second helices of Im7 (Figure 4). Likewise, a small library of mutants varying in the glycosylation sequence was created to evaluate the efficiency of glycosylation in the context of a full-length folded protein. The overall structures of the Im7 mutants were compared with that of the wild-type using circular dichroism spectroscopy in order to verify that introduction of the glycosylation sequence did not significantly disrupt the overall fold of the protein (Figure S2, Supporting Information).

Glycosylation of Im7 Mutants by PglB. Each Im7 mutant was incubated with PglB and the radiolabeled glycosyl donor for 3 h. When the sample was subjected to a reverse-phase HPLC analysis, a significant amount of radioactivity was found to co-elute with a peak (228 nm absorption) corresponding to the Im7 protein (Figure 5). No appreciable radioactivity was found to be associated with the wild-type Im7 ^{26}NVAAT^{30} and the non-glycosylatable Im7 mutant ^{26}DVAAT^{30}, missing the asparagine glycosylation site, as the protein acceptor (Figure 6), or in the absence of PglB (data not shown). Mass spectral analysis confirmed that the disaccharide is covalently linked to the Im7 protein (Figure S3, Supporting Information). The Im7 mutant containing the optimized consensus sequence DQNAT was found to be glycosylated to a greater extent than the other mutants. This demonstrates that the glycosylation of fully folded proteins by PglB are influenced by the amino acid primary sequence within the glycosylation site.

Evaluation of PglB with the Eukaryotic Glycan Donor. To determine whether PglB is capable of utilizing the eukaryotic glycosyl donor substrate, we carried out the glycosylation of the Im7 mutants using radiolabeled dolichyl-pyrophosphate chitobiose (Figure 7) as the glycosyl donor. After a 24-h incubation, no significant radioactivity was found to be associated with the Im7 mutant protein (data not shown).

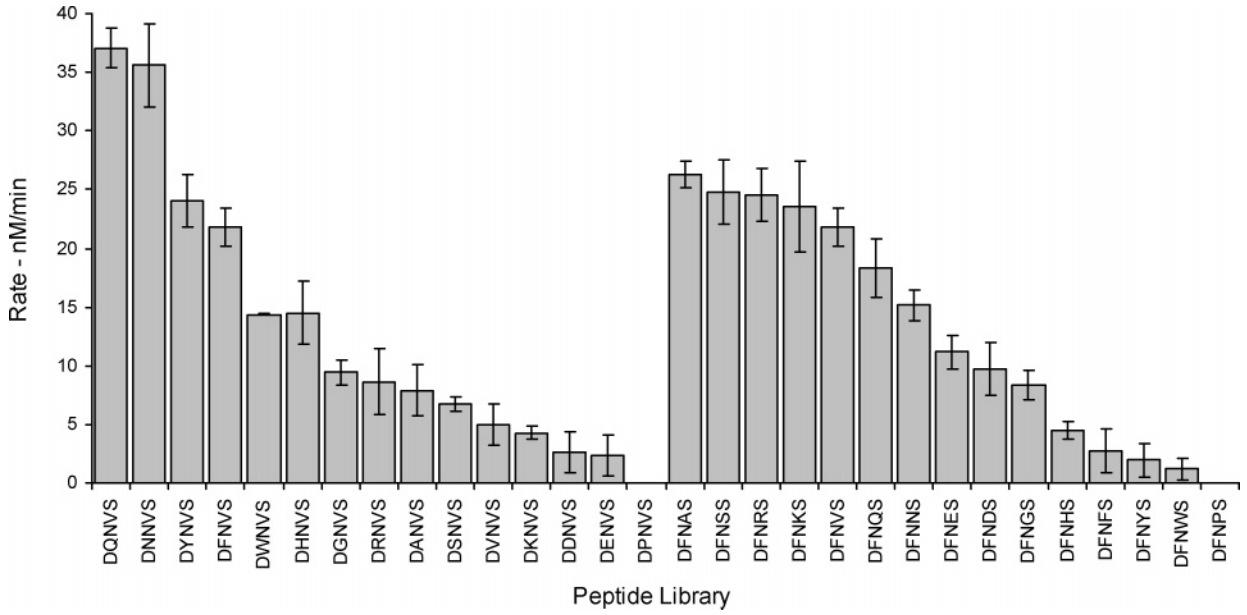


FIGURE 3: Comparison of PglB initial rates with peptides substrates varying at the X_1- and X_2-positions. Peptides were assayed at 10 μM concentration, close to the average peptide $K_{m(app)}$.

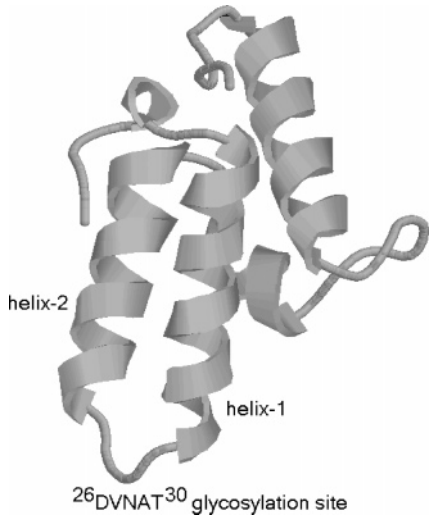


FIGURE 4: Native Im7 structure, indicating the chosen glycosylation site between helix-1 and helix-2 (*30*), pdb code 1CEI.

DISCUSSION

In eukaryotic cells, N-linked glycosylation is known to influence many properties of the resultant glycoproteins, including folding, stability, solubility, antigenicity, trafficking, and enzyme activity (*19*, *20*). Glycosylation occurs at the consensus sequence **N**-X-S/T, where X can be any amino acid except proline (*4*, *5*). However, not all potential sites are glycosylated (*7*), as many factors have been shown to influence site occupancy, such as the choice of serine or threonine at the +2 position, the choice of amino acid at the X-position, local amino acids flanking the glycosylation sequon, and the local structural conformation of the polypeptide (*6*).

In *C. jejuni*, N-linked glycosylation takes place within the consensus sequence D/E-X_1-**N**-X_2-S/T (Figure 2) (*13*), with the additional requirement of an acidic residue at the −2 position for efficient glycosylation, in comparison to the eukaryotic system. In order to understand the factors governing site occupancy in *C. jejuni*, we synthesized a library of peptides, based on the known glycosylation site from periplasmic protein PEB3 (*18*), systematically differing at each position to elucidate the acceptor-binding determinants for PglB. All peptides were capped at both termini because this has been previously shown to improve the binding of peptides to the eukaryotic OT (*4*). We have found that peptides containing threonine as the hydroxyamino acid at position +2 exhibit a 3-fold lower $K_{m(app)}$ in comparison to that of the serine counterpart (Table 1). The eukaryotic OT has a similar preference for threonine as the hydroxyamino acid at position +2, whereby the substitution of serine for threonine can dramatically improve the glycosylation efficiency of a poorly glycosylated site (*21*, *22*). Likewise, although PglB has been known to glycosylate sites with either aspartic or glutamic acid at position −2 in *C. jejuni*, we have found that it exhibited a distinct preference for aspartic acid at that position (Table 1). Between the two model peptides Ac-EVNAT-(*p*NF)-NH_2 and Ac-DVNAT-(*p*NF)-NH_2, having glutamate as the acidic residue at position −2 increased the $K_{m(app)}$ by 6-fold, while decreasing the $V_{max(app)}$ by 5-fold. The effects of the substitutions appear to be additive. When both unfavorable amino acids were present within the same peptide Ac-EFNVS-*p*NF, glycosylation was negligible under the experimental conditions employed. This was initially surprising because the *C. jejuni* glycoprotein HisJ has been found to be glycosylated at the site ^{27}ESNAS^{31}, although less efficiently than other *C. jejuni* glycoproteins (*17*).

The unusual peptides containing cysteine, homoserine, or diaminopropionic acid at the +2 position were not found to be glycosylated by PglB to a noticeable degree under the assay conditions employed, nor did they exert any inhibitory effect on the glycosylation of consensus acceptor peptides, even at a concentration of 200 μM. This suggests that PglB has a more stringent acceptor recognition requirement than the eukaryotic OT, which has been known to glycosylate sites with cysteine in place of the hydroxyl amino acid, albeit at lower efficiencies (*21*, *23*).

In the same manner, we evaluated the preferences of PglB for the amino acids in the two X-positions (Figure 3). For

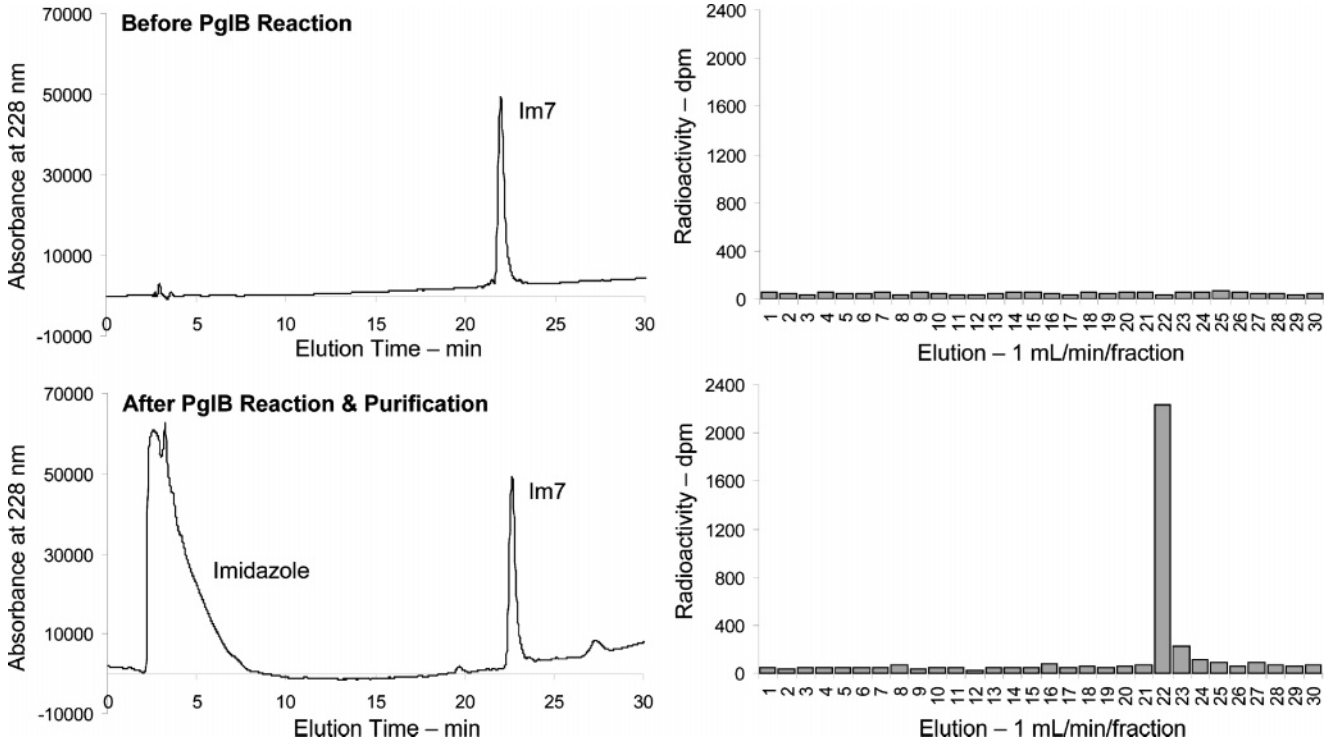


FIGURE 5: Reverse-phase HPLC traces of Im7 mutant ^{26}DVNAT^{30} before and after glycosylation by PglB. Elution of protein was monitored by absorbance at 228 nm and collected in 1 mL fractions for scintillation counting. The degree of glycosylation was quantified by the amount of radiolabeled disaccharide co-eluting with the protein acceptor.

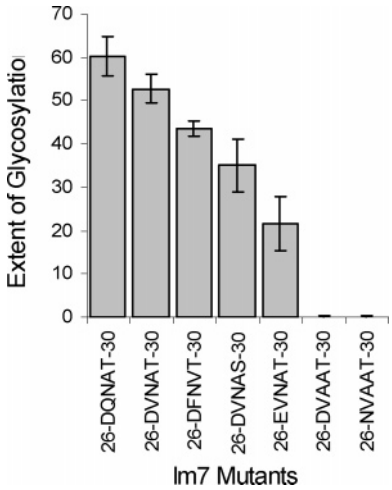


FIGURE 6: Comparison of various Im7 mutants differing in their glycosylation sequences. The Extent of glycosylation corresponds to the percentage of (limiting) radiolabeled glycan donor consumed in the 3 h reaction.

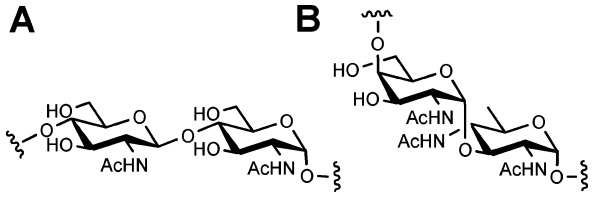


FIGURE 7: Structure of the first two sugars within the (A) eukaryotic glycan GlcNAc-β-(1 → 4)-GlcNAc and (B) *C. jejuni* glycan GalNAc-α-(1 → 3)-Bac.

the X_1-position, PglB revealed a clear preference for the amido functionalities of asparagine and glutamine as well as large hydrophobic functionalities, whereas charged amino acids were clearly disfavored at that position. For the X_2-position, PglB exhibited a trend similar to that of the eukaryotic OT (*24*), favoring the positively charged functionalities of lysine and arginine, along with alanine and serine, while disfavoring large hydrophobic functionalities. The differences within this set of peptides mainly come from the differences in $K_{m(app)}$, whereas the relative $V_{max(app)}$ values are similar and do not follow the same trend. It is interesting that PglB has such contrasting preferences for the two positions directly adjacent to the glycosylation site. Within the peptide library presented, the DQNAT sequence is the optimal peptidyl substrate for PglB.

Recently, it has been demonstrated that unlike the eukaryotic OT, PglB is capable of glycosylating fully folded proteins in vitro (*25*). To demonstrate that the trends observed from the comparison of peptides are relevant to the glycosylation of full-length folded proteins, we engineered variations of the five amino acid consensus glycosylation sequence into a well-defined loop of the non-*C. jejuni*, model protein, Im7, which is structurally well-characterized (*14*−*16*). Circular dichroism spectroscopy studies indicated that the introduction of the various glycosylation sequences did not result in any significant changes in the overall folding of the mutants when compared to the native Im7 (Figure S2, Supporting Information). When assayed with PglB, the relative extents of glycosylation were consistent with the glycosylation efficiencies of the corresponding peptides. However, the differences among the full-length proteins were less dramatic (Table 2). Our interpretation of this finding is that similar to the eukaryotic OT, the acceptor binding of PglB is influenced by residues beyond the five amino acid consensus sequence and that the constraints conferred by the full length protein restricts the conformations that the short flexible peptide naturally adopts, be it favorable or unfavorable. Nevertheless, the overall degree of glycosylation was

Table 2: Kinetic Parameters of PglB Peptide Substrates in Comparison to the Glycosylation Efficiency of Corresponding Im7 Mutants[a]

peptide	$V_{max(app)}$ nM/min	$K_{m(app)}$ μM	$V_{max(app)}/K_{m(app)}$ $min^{-1} \times 10^3$	Im7 mutant	extent of glycosylation
Ac-NVAAT-(*p*NF)-NH_2		no activity		native	0 ± 0.0
Ac-DVAAT-(*p*NF)-NH_2		no activity		^{26}DVNAT^{30}	0 ± 0.0
Ac-EVNAT-(*p*NF)-NH_2	50.6 ± 1.6	22.5 ± 2.56	2.3 ± 0.29	^{26}EVNAT^{30}	22 ± 6.3
Ac-DVNAS-(*p*NF)-NH_2	31.0 ± 1.8	3.33 ± 0.43	9.3 ± 1.47	^{26}DVNAS^{30}	35 ± 5.0
Ac-DVNVT-(*p*NF)-NH_2	38.9 ± 1.9	1.44 ± 0.21	27.0 ± 4.61		
Ac-DFNVT-(*p*NF)-NH_2	34.6 ± 0.5	1.22 ± 0.07	28.4 ± 1.80	^{26}DFNVT^{30}	43 ± 1.6
Ac-DVNAT-(*p*NF)-NH_2	32.1 ± 1.1	1.06 ± 0.05	30.3 ± 1.96	^{26}DVNAT^{30}	53 ± 3.2
Ac-DFNAT-(*p*NF)-NH_2	29.5 ± 1.6	0.87 ± 0.12	33.9 ± 5.59		
Ac-DQNAT-(*p*NF)-NH_2	34.8 ± 1.2	0.80 ± 0.11	43.4 ± 6.73	^{26}DQNAT^{30}	60 ± 4.5

[a] The extent of glycosylation corresponds to the percentage of (limiting) radiolabeled glycosyl donor consumed in the 3-h reaction.

improved by having the optimal glycosylation sequence DQNAT in the Im7 loop.

Our previous in vitro studies have revealed that PglB is capable of transferring undecaprenyl pyrophosphate-linked glycans of varying lengths as well as glycans with GlcNAc at the reducing end (*12*) as opposed to the native *N*,*N*′-diacetyl-bacillosamine sugar from *C. jejuni*. Recent studies further illustrate the nonspecificity of PglB toward the glycosyl donor by showing that it is capable of transferring diverse undecaprenyl pyrophosphate-linked carbohydrates from various O-antigen biosynthetic pathways in vivo (*26*, *27*). In these studies, we have found that the eukaryotic dolichyl-pyrophosphate-GlcNAc-GlcNAc disaccharide is a poor substrate for PglB. This may be due to selectivity toward the isoprene carrier, as dolichol is 30−50 carbons larger than undecaprenol and differs in the oxidation state of the terminal isoprene unit (*28*, *29*). However, on the basis of a structural evaluation of all carbohydrates reported to be accepted by PglB, we believe that the β1−4-linkage between the first two GlcNAc residues might be unfavorable to PglB because the native linkage in *C. jejuni* is β-1,3 (Figure 7). All examples of the O-antigen carbohydrates transferred by PglB involve either α-1,6, α-1,3, or the native β-1,3 linkages between the first two sugars (*26*, *27*). It is possible that although the N-acetyl group at the C4-position of bacillosamine is also not critical, replacement of this functionality with the attachment of a bulky sugar may disrupt the binding to PglB.

In conclusion, we have characterized the acceptor specificity of PglB in vitro, in the context of flexible peptide substrates as well as full-length folded proteins. We have determined that the optimal glycosylation consensus sequence for PglB is DQNAT, although additional binding determinants and local peptide conformations are likely to also affect glycosylation efficiency in full-length proteins. Understanding the factors that influence N-linked glycosylation by PglB is a crucial start to the elucidation of the poorly understood mechanism of oligosaccharyl transferases involved in N-linked protein glycosylation. Furthermore, information regarding the acceptor specificity of PglB can assist in the design and synthesis of novel recombinant glycoproteins for research and clinical applications as well as PglB inhibitors as antibiotics for *C. jejuni*-induced gastrointestinal disorders.

ACKNOWLEDGMENT

We thank Dr. Nelson Olivier, Langdon Martin, and Dr. Jerry Troutman for critical reading of this manuscript and helpful discussions, and Deborah Pheasant for technical assistance with use of the CD spectrometer. We thank Professor Sheena Radford for the plasmid containing wild-type Im7. We gratefully acknowledge the Biophysical Instrumentation Facility for the Study of Complex Macromolecular Systems (NSF-0070319 and NIH GM68762) for the use of the CD spectrometer.

SUPPORTING INFORMATION AVAILABLE

Detailed expression and purification protocols for Im7 mutants, PglB peptide assay data, circular dichroism spectra of Im7 mutants, and ESI-MS of the glycosylated Im7 mutant. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Weerapana, E., and Imperiali, B. (2006) Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems, *Glycobiology 16*, 91−101.
2. Burda, P., and Aebi, M. (1999) The dolichol pathway of N-linked glycosylation, *Biochim. Biophys. Acta 1426*, 239−257.
3. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum, *Annu. Rev. Biochem. 73*, 1019−1049.
4. Hart, G. W., Brew, K., Grant, G. A., Bradshaw, R. A., and Lennarz, W. J. (1979) Primary structural requirements for the enzymatic formation of the N-glycosidic bond in glycoproteins. Studies with natural and synthetic peptides, *J. Biol. Chem. 254*, 9747−9753.
5. Bause, E. (1983) Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes, *Biochem. J. 209*, 331−336.
6. Jones, J., Krag, S. S., and Betenbaugh, M. J. (2005) Controlling N-linked glycan site occupancy, *Biochim. Biophys. Acta 1726*, 121−137.
7. Heijne, G. V., and Gavel, Y. (1990) Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering, *Protein. Eng. 3*, 433−442.
8. Young, N. M., Brisson, J. R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., Jarrell, H. C., Cadotte, N., St Michael, F., Aberg, E., and Szymanski, C. M. (2002) Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*, *J. Biol. Chem. 277*, 42530−42539.
9. Linton, D., Dorrell, N., Hitchen, P. G., Amber, S., Karlyshev, A. V., Morris, H. R., Dell, A., Valvano, M. A., Aebi, M., and Wren, B. W. (2005) Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway, *Mol. Microbiol. 55*, 1695−1703.
10. Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A., Wren, B. W., and Aebi, M. (2002) N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*, *Science 298*, 1790−1793.
11. Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., and Guerry, P. (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni*, *Mol. Microbiol. 32*, 1022−1030.

12. Glover, K. J., Weerapana, E., Numao, S., and Imperiali, B. (2005) Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*, *Chem. Biol. 12*, 1311−1315.
13. Kowarik, M., Young, N. M., Numao, S., Schulz, B. L., Hug, I., Callewaert, N., Mills, D. C., Watson, D. C., Hernandez, M., Kelly, J. F., Wacker, M., and Aebi, M. (2006) Definition of the bacterial N-glycosylation site consensus sequence, *EMBO J. 25*, 1957−1966.
14. Capaldi, A. P., Kleanthous, C., and Radford, S. E. (2002) Im7 folding mechanism: misfolding on a path to the native state, *Nat. Struct. Biol. 9*, 209−216.
15. Friel, C. T., Capaldi, A. P., and Radford, S. E. (2003) Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins, *J. Mol. Biol. 326*, 293−305.
16. Hackenberger, C. P., Friel, C. T., Radford, S. E., and Imperiali, B. (2005) Semisynthesis of a glycosylated Im7 analogue for protein folding studies, *J. Am. Chem. Soc. 127*, 12882−12889.
17. Nita-Lazar, M., Wacker, M., Schegg, B., Amber, S., and Aebi, M. (2005) The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation, *Glycobiology 15*, 361−367.
18. Linton, D., Allan, E., Karlyshev, A. V., Cronshaw, A. D., and Wren, B. W. (2002) Identification of N-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in *Campylobacter jejuni*, *Mol. Microbiol. 43*, 497−508.
19. Helenius, A., and Aebi, M. (2001) Intracellular functions of N-linked glycans, *Science 291*, 2364−2369.
20. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct, *Glycobiology 3*, 97−130.
21. Bause, E., and Legler, G. (1981) The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr(Ser) for the N-glycosylation step during glycoprotein biosynthesis, *Biochem. J. 195*, 639−644.
22. Kasturi, L., Chen, H. G., and ShakinEshleman, S. H. (1997) Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors, *Biochem. J. 323*, 415−419.
23. Sato, C., Kim, J. H., Abe, Y., Saito, K., Yokoyama, S., and Kohda, D. (2000) Characterization of the N-oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor, *J. Biochem. (Tokyo) 127*, 65−72.
24. Shakin-Eshleman, S. H., Spitalnik, S. L., and Kasturi, L. (1996) The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency, *J. Biol. Chem. 271*, 6363−6366.
25. Kowarik, M., Numao, S., Feldman, M. F., Schulz, B. L., Callewaert, N., Kiermaier, E., Catrein, I., and Aebi, M. (2006) N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase, *Science 314*, 1148−1150.
26. Wacker, M., Feldman, M. F., Callewaert, N., Kowarik, M., Clarke, B. R., Pohl, N. L., Hernandez, M., Vines, E. D., Valvano, M. A., Whitfield, C., and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems, *Proc. Natl. Acad. Sci. U.S.A. 103*, 7088−7093.
27. Feldman, M. F., Wacker, M., Hernandez, M., Hitchen, P. G., Marolda, C. L., Kowarik, M., Morris, H. R., Dell, A., Valvano, M. A., and Aebi, M. (2005) Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A. 102*, 3016−3021.
28. Fang, X., Gibbs, B. S., and Coward, J. K. (1995) Synthesis and evaluation of synthetic analogues of dolichyl-PP-chitobiose as oligosaccharyltransferase substrates, *Bioorg. Med. Chem. Lett. 5*, 2701−2706.
29. Ye, X. Y., Lo, M. C., Brunner, L., Walker, D., Kahne, D., and Walker, S. (2001) Better substrates for bacterial transglycosylases, *J. Am. Chem. Soc. 123*, 3155−3156.
30. Chak, K. F., Safo, M. K., Ku, W. Y., Hsieh, S. Y., and Yuan, H. S. (1996) The crystal structure of the immunity protein of colicin E7 suggests a possible colicin-interacting surface, *Proc. Natl. Acad. Sci. U.S.A. 93*, 6437−6442.

BI602633N