Acc. Chem. Res. **1997,** *30,* 452−459

Protein Glycosylation: The Clash of the Titans

BARBARA IMPERIALI

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received May 16, 1997

Enzymes catalyze a variety of transformations with spectacular rate enhancements while demonstrating exquisite selectivities. Information from complementary structural and mechanistic studies has yielded insight into the inner workings of many enzyme-catalyzed reactions. In particular, studies of transformations that involve the interaction of one or more small-molecule substrates with soluble enzymes have met with notable success. There remain, however, a number of reactions, including those involving the covalent modification of macromolecular substrates, that confound us by their complexity.[1,2] One such transformation, asparagine-linked glycosylation, involves the enzyme-catalyzed modification of an asparagine side chain in a nascent polypeptide with a triantennary tetradecasaccharide ($GlcNAc_2Man_9Glc_3$) moiety.[3−5] This first committed step in the biosynthesis of *N*-linked glycoproteins is catalyzed by oligosaccharyl transferase (OT), a heteromeric membrane-associated enzyme complex, found in the lumen of the endoplasmic reticulum (ER) of eukaryotic cells.[6,7] Glycosyl transfer to the nitrogen of the carboxamide side chain of an asparagine occurs through the intermediacy of a dolichol-linked pyrophosphate donor as illustrated in Figure 1. The primary peptide sequence requirements for the glycosylation process include a minimum -Asn-Xaa-Ser/Thr- tripeptide recognition motif where Xaa can be any of the 20 natural amino acids except proline.[8,9] Thus, while OT exhibits rather simple substrate requirements, the enzyme catalyzes an unusual and specific reaction wherein the nucleophilicity of the asparagine side chain must be greatly enhanced to form a covalent linkage with the oligosaccharide. After transfer of the initial tetradecasaccharide, subsequent diversification of the primary glycoprotein conjugates arises from enzyme-catalyzed processing steps that occur in both the ER and the Golgi apparatus.[4] Specifically, the core oligosaccharide structures are subject to the action of an elaborate array of glycosyl hydrolase and glycosyl transferase enzymes. The latter class of enzymes catalyzes the incorporation of a broader diversity of carbohydrate units including fucose, sialic acid, and galactose. The collective action of these enzymes generates the structural diversity associated with mature *N*-linked glycoproteins in eukaryotic cells.

Barbara Imperiali was born in London, England. She received her B.Sc. in medicinal chemistry from University College London in 1979 and her Ph.D. in synthetic organic chemistry from the Massachusetts Institute of Technology in 1983. She is currently Professor of Chemistry at the California Institute of Technology.

The dimensions of the appended carbohydrate moieties are significant when compared with the size of the proteins that they decorate (see Figure 2 illustrating the glycoprotein CD2[10]); therefore, it is not surprising that *N*-linked glycosylation has a major impact on many aspects of protein structure and function. Protein glycosylation has been implicated in such varied processes as the immune response, intracellular targeting, intercellular recognition, and protein folding, stability, and solubility.[11,12] An understanding of the biosynthesis of *N*-linked glycoproteins is of intense interest because of the critical roles played by these key biomolecules.

Role of Protein Substrate Conformation in Asparagine-Linked Glycosylation

Oligosaccharyl transferase acts cotranslationally as membrane-associated ribosomes direct a newly translated polypeptide into the lumen of the rough ER. Recent studies have shown that approximately 14 residues of the newly synthesized peptide must clear the luminal surface of the ER membrane before OT-mediated glycosylation can occur.[13] This implies that the active site of OT resides in a soluble domain of the enzyme, distal to the membrane surface. The fact that the protein remains attached to the ribosome during glycosylation suggests that, while the local secondary structure of the nascent peptide may influence the propensity for the glycosylation reaction to occur, the global folding of the native protein does not play a role in the recognition events that lead to the modification reaction. To probe the importance of secondary structure in *N*-linked glycosylation, the conformational preferences of a family of tripeptides were examined.[14] These analyses were complemented by a parallel determination of the glycosyl acceptor properties of each peptide in an in vitro assay in which the truncated substrate $DolPP(GlcNAc)_2$[15,16] was employed as the glycosyl donor.[17] Since amino- and carboxy-terminal capped tripeptides are the simplest substrates for *N*-linked glycosylation, the hydrogen-bonded motifs for recognition

(1) Han, K.-K.; Martinage, A. *Int. J. Biochem.* **1992**, *24,* 19−28.
(2) Yan, S. C. B.; Grinnell, B. W.; Wold, F. *Trends Biochem. Sci.* **1989**, *14,* 264−268.
(3) Silberstein, S.; Gilmore, R. *FASEB J.* **1996**, *10,* 849−858.
(4) Kornfeld, R.; Kornfeld, S. *Annu. Rev. Biochem.* **1985**, *54,* 631−664.
(5) Imperiali, B.; Hendrickson, T. L. *Bioorg. Med. Chem.* **1995**, *3,* 1565−1578.
(6) Hirschberg, C. B.; Snider, M. D. *Annu. Rev. Biochem.* **1987**, *56,* 63−87.
(7) Abeijon, C.; Hirschberg, C. B. *Trends Biochem. Sci.* **1992**, *17,* 32−36.
(8) Gavel, Y.; von Heijne, G. *Protein Eng.* **1990**, *3,* 433−442.
(9) Marshall, R. D. *Biochem. Soc. Symp.* **1974**, *40,* 17−26.
(10) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Aulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269,* 1273−1278.
(11) Dwek, R. A. *Biochem. Soc. Trans.* **1994**, *23,* 1−25.
(12) Varki, A. *Glycobiology* **1993**, *3,* 97−130.
(13) Nilsson, I.; von Heijne, G. *J. Biol. Chem.* **1993**, *268,* 5798−5801.
(14) Imperiali, B.; Shannon, K. L. *Biochemistry* **1991**, *30,* 4374−4380.
(15) Sharma, C. B.; Lehle, L.; Tanner, W. *Eur. J. Biochem.* **1981**, *116,* 101−108.
(16) Imperiali, B.; Zimmerman, J. W. *Tetrahedron Lett.* **1990**, *31,* 6485−6488.
(17) It should be noted that there is considerable species variation between the OT kinetic parameters K_M and V_{max}. For example, the K_{Mapp} for Bz-Asn-Leu-Thr-NHMe is approximately 10-fold smaller in yeast OT than in porcine liver OT. Therefore the comparison of these values is only applicable when measurements from the same species are being considered.

S0001-4842(95)00226-3 CCC: $14.00 © 1997 American Chemical Society

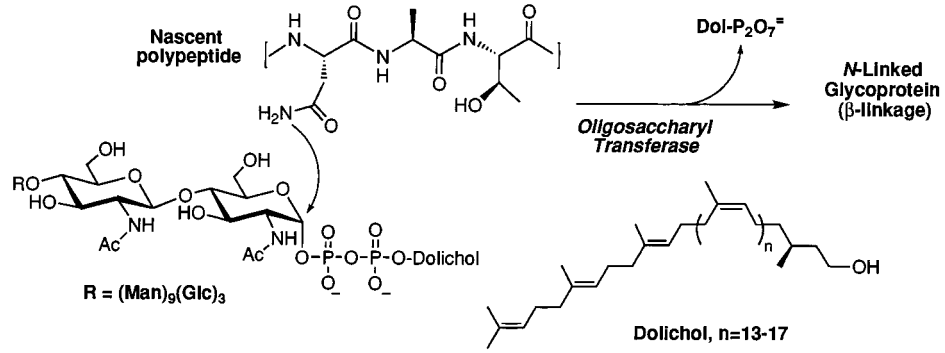


FIGURE 1. Glycosylation reaction catalyzed by oligosaccharyl transferase.

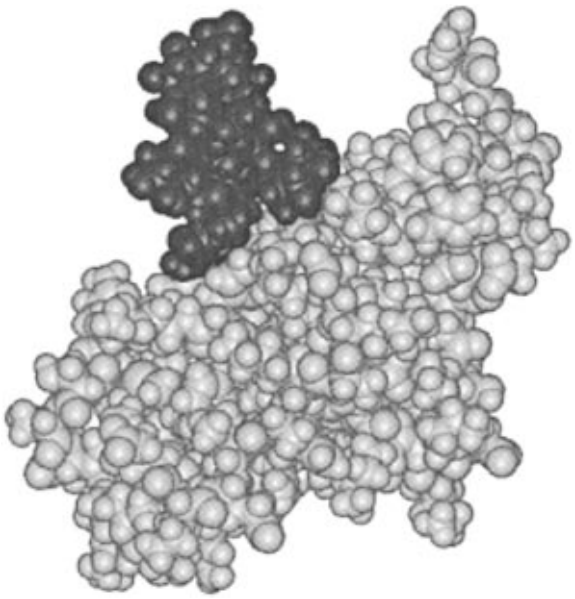


FIGURE 2. Space-filling model of a portion of the glycoprotein CD2[10] (pdb file 1GYA) illustrating the relative dimensions of the carbohydrate (11 saccharide units; black) and protein (105 amino acid residues; gray) constituents of the glycoprotein conjugate.

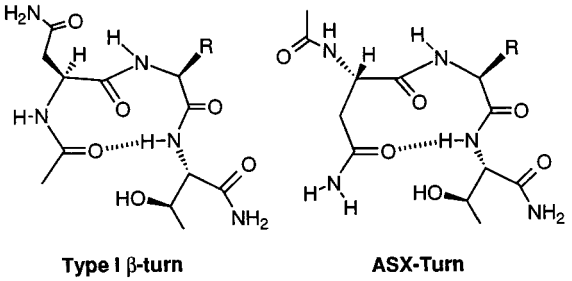


FIGURE 3. Comparison between an Asx-turn and a β-turn conformation for the tripeptide Ac-Asn-Xaa-Thr-NH_2.

are limited to turns; other common structural elements such as α-helices or β-sheets require longer peptides for complete formation.

Two distinct turn conformers, the β-turn[18] and the Asx-turn[19] (illustrated in Figure 3) are compatible with the tripeptide substrates that were studied in this laboratory. The β-turn motif results in a peptide chain reversal and is characterized by a hydrogen bond between the threonine amide proton and the carbonyl group immediately preceding the asparagine. In the Ac-Asn-Xaa-Thr-NH_2 tripeptide illustrated, the hydrogen-bonding carbonyl is provided by the acetyl group. The Asx-turn involves a 10-membered hydrogen-bonding array, similar to that of a β-turn. In this case, however, the carbonyl oxygen of the asparagine side chain acts as the hydrogen bond acceptor, resulting in a more extended peptide backbone conformation. The conformational studies revealed that the tripeptides which served as effective substrates also showed spectroscopic features that were diagnostic of the Asx-turn in solution.[14] The Asx-turn has been frequently observed in protein structures; it has been estimated that 55% of all hydrogen bonds to the carbonyl side chain of asparagine residues are provided by the backbone NH of the $(i + 2)$ residue.[20] In this context, it is noteworthy that the homologous residue, glutamine, is never glycosylated.[21] This apparent lack of reactivity may be explained by the lower preference for an Asx-turn-like conformation in the corresponding glutamine-containing peptides.

Conformationally constrained peptide substrates provided corroborating evidence for the studies performed on flexible, linear peptides. Peptides forced to adopt an Asx-turn conformation showed enhanced glycosylation when compared to the unconstrained analogues. In sharp contrast, peptides constrained to a β-turn motif, with asparagine at either the (i) or $(i + 1)$ position, were devoid of glycosyl acceptor properties.[22] Modeling studies were used to evaluate several compounds with main-chain to side-chain lactam cyclizations between the amino-terminus and the Xaa side chain. These studies predict that the cyclic compound cyclo[Asn-Add]-Thr-NHMe (**1**), which contains the amino acid (*S*)-2-aminodecanedioic acid (Add), would show ideal Asx-turn properties in a minimum energy conformer.[23] The constrained peptide, **1**, was prepared and subject to detailed structural characterization in aqueous media using nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopies. Structural studies confirmed the proposed Asx-turn conformation in aqueous solution. Most importantly, the glycosyl acceptor properties of **1** were compared with those of similar, but unconstrained analogues, thereby enabling a direct assessment of the role of the Asx-turn in *N*-linked glycosylation. The cyclic compound, **1**, has an apparent K_M of 78 μM, while the corresponding value for the linear counterpart, **2**, is 800 μM. Thus, these

(18) Hutchinson, E. G.; Thornton, J. M. *Prot. Sci.* **1994**, *3*, 2207−2216.
(19) Abbadi, A.; Mcharfi, M.; Aubry, A.; Premilat, S.; Boussard, G.; Marraud, M. *J. Am. Chem. Soc.* **1991**, *113*, 2729.
(20) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97.
(21) Bause, E. *Biochem. J.* **1979**, *103*, 296−299.
(22) Imperiali, B.; Shannon, K. L.; Rickert, K. W. *J. Am. Chem. Soc.* **1992**, *114*, 7942−7943.
(23) Imperiali, B.; Spencer, J. R.; Struthers, M. D. *J. Am. Chem. Soc.* **1994**, *116*, 8424−8425.

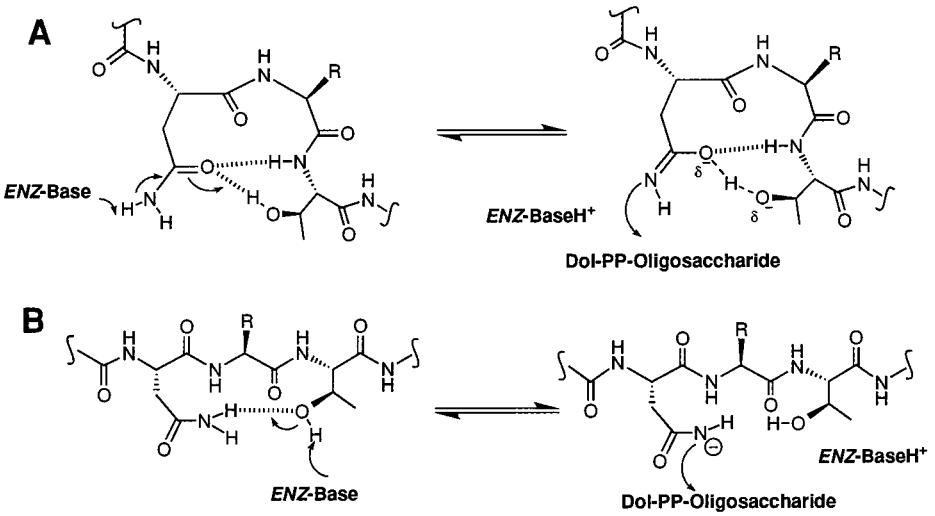
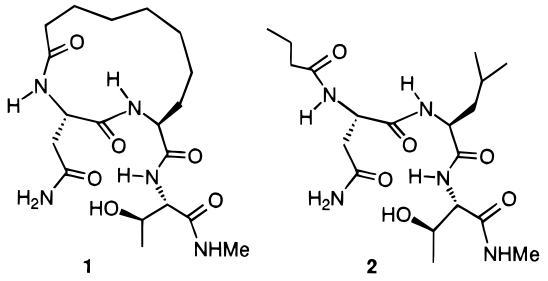


FIGURE 4. Proposed mechanism of asparagine activation for *oligosaccharyl transferase*: (A) mechanism involving amide tautomerization;[25] (B) mechanism involving amide deprotonation.[27, 31]

studies clearly demonstrate that the preorganization of **1** to an Asx-turn conformer significantly improves enzyme affinity.



Previous studies by Bause[24] have examined the effects of disulfide constraints on the glycosylation properties of short polypeptides. In light of the proposal that an Asx-turn is the recognition motif for *N*-linked glycosylation, a reexamination of their results supports the hypothesis that the acceptor ability of an Asn-Xaa-Thr/Ser sequence may be related to the ability of the peptide to access an Asx-turn motif. In all cases where the presence of the disulfide enhanced a β-turn structure, glycosyl acceptor properties diminished. In the one case in which a cysteine was positioned to potentially facilitate the formation of an Asx-turn, glycosyl transfer was enhanced in the oxidized state.

In some cases, Asn-Xaa-Ser/Thr sequences remain unglycosylated following translation and translocation into the lumen of the ER. This apparent failure of the glycosylation machinery most likely results from conformational effects imposed by the neighboring polypeptide sequence that may override the potential for the tripeptide acceptor sequence to adopt an Asx-turn and therefore limit its compatibility with the OT active site.

Toward an Understanding of the Mechanism of Asparagine Amide Activation

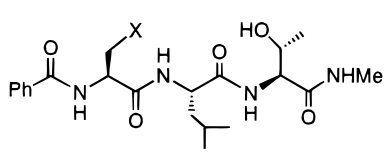
Examples of enhanced amide nitrogen nucleophilicity are rare in mechanistic enzymology. Therefore, a prevailing question with *N*-linked glycosylation concerns the mode of amide activation that renders the nitrogen nucleophilic enough to attack the lipid-linked glycosyl donor. In glycosylation, the electrophile is likely to be represented by a highly reactive cationic species. A reactive electrophile alone, however, is insufficient to guarantee the specific modification of asparagine, particularly when one considers the numerous competing, and potentially more reactive, nucleophiles present in protein substrates. Therefore, to achieve selective *N*-linked glycosylation, cooperation between OT and the local peptide substrate structure may result in the activation of a normally non-nucleophilic asparagine side chain. Thus, the generation of amide reactivity and reaction specificity must be considered in the context of the previous conformational studies identifying the Asx-turn as the likely recognition motif for *N*-linked glycosylation. With an understanding of the necessary secondary structural features required for *N*-linked glycosylation, mechanistic investigations can be designed to probe the roles of the hydroxy amino acid and the amide backbone in enhancing the nucleophilicity of the asparagine side chain.

The conformational studies presented above prompted the development of a mechanistic proposal for OT-mediated catalysis. In this proposal, the unique hydrogen-bonding array provided by the Asx-turn is suggested to facilitate protonation of the carbonyl of the asparagine side chain while simultaneous, enzyme-mediated deprotonation at the nitrogen effects the tautomerization of the carboxamide to an imidol species.[25] Thus, tautomerization would afford a neutral, nucleophilic species, which could then react with the electrophilic lipid-linked oligosaccharide (Figure 4A). This mechanism incorporates both the issues of specificity and reactivity as well as the absolute requirement for a hydroxy amino acid by integrating substrate structural requirements with participation of enzyme active site residues. Thus, the likelihood of an Asn-Xaa-Thr/Ser sequence to undergo glycosylation would be governed by the ability of each potential substrate to adopt an Asx-turn conformation within the active site of the enzyme.

The lack of structural information and limited availability of pure enzyme currently limits mechanistic pro-

(24) Bause, E.; Hettkamp, H.; Legler, G. *Biochem. J.* **1982**, *203*, 761−768.
(25) Imperiali, B.; Shannon, K. L.; Unno, M.; Rickert, K. *J. Am. Chem. Soc.* **1992**, *114*, 7944−7945.

posals to those based primarily upon kinetic studies. Evidence for different mechanistic proposals is therefore obtained by examining the kinetic properties of substrate analogues. By selectively altering the natural substrate, proposals concerning the features of the active site or the reaction intermediates can be investigated. To this end, a number of different substrate analogues have been prepared and examined.[25] Bz-Asp-Leu-Thr-NHMe (**3**) and



3; X = CO_2H
4; X = CO_2Me
5; X = $C(=S)NH_2$
6; X = $C(=O)NH_2$
7; X = $CH_2NH_3^+$
8; X = $C(=O)CH_3$

Bz-Asp(OγMe)-Leu-Thr-NHMe (**4**) exhibited no binding to OT, while the tripeptide Bz-Asn(γS)-Leu-Thr-NHMe (**5**) is a substrate for the enzyme. The heteroatom replacement in **5** increases the acidity of the hydrogen-bond donor site and decreases the basicity at the hydrogen-bond acceptor site relative to the corresponding oxoamide. Although the sulfur substitution introduces some steric perturbations into the system, the binding of **5** to OT is barely affected by the thiocarbonyl replacement (K_{Mapp} = 0.26 vs 0.24 mM for **6**). However, the relative maximal velocity is reduced to 8.4% of that observed for the parent compound, **6**. Incorporation of γ-aminobutyrine (Amb) in place of asparagine yields Bz-Amb-Leu-Thr-NHMe (**7**), a peptide with a significantly lower pK_a at the Amb side chain. Tripeptide **7** is a competitive inhibitor of OT with a K_i of 1 mM, which is comparable to the K_M of **6**. The methyl ketone analogue, **8**, shows weak competitive inhibition of OT. Additionally, tripeptides with N^{δ}-methylasparagine in place of asparagine are not tolerated by the enzyme.[26] For this peptide, binding of the energetically favorable *trans* amide isomer (of methylasparagine) would interfere with the deprotonation step in the proposed mechanism.

Studies with substrate analogues in which the hydroxy amino acid is replaced have also revealed that structural perturbations in this component of the tripeptide sequence significantly affect glycosylation. For example, tripeptides containing valine,[21] threonine(β-OMe),[27,28] β-hydroxynorvaline,[29] and *allo*-threonine[30,31] in the C-terminal position are neither substrates nor inhibitors of OT, even at elevated peptide concentrations, while peptides which incorporate cysteine in this position show low, but measurable, glycosylation. Bause and co-workers have also recently reported that peptides which include (epoxyethyl)glycine in place of the essential hydroxyamino acid caused OT inactivation and labeling of one of the enzyme subunits.[31−33] However, the high concentrations of peptide and low labeling efficiencies reported in these studies should be taken into consideration in the interpretation of these results.

Considered together, substrate analogue studies suggest that for OT substrate recognition, there is a prerequisite for side chain functionality at the asparagine and hydroxy amino acid positions with ionization properties similar to those of the native substrate. Peptides **3** and **4** fail to be recognized as either substrates or inhibitors for OT, while tripeptide **5** is a substrate and **7** is a competitive inhibitor. For peptide **5**, the K_M remains similar to that of the native substrate, even though the reduction in the basicity of the sulfur heteroatom has resulted in less efficient glycosylation. With **7**, the lower pK_a of the nitrogen appears to have allowed recognition by OT, despite the loss of the Asx-turn motif, but glycosyl acceptor properties have been totally lost. These observations and the weak inhibition by the methyl ketone analogue, **8**, are all consistent with the proposed tautomerization mechanism.

An alternative proposal for amide activation has focused on the generation of an anionic species following deprotonation at the asparagine carboxamide (Figure 4B).[27,31] This intermediate is proposed to be accessible in the enzyme-catalyzed reaction if the essential hydroxy amino acid becomes deprotonated by an enzyme-bound base and the incipient alkoxide subsequently deprotonates the asparagine amide. However, the proponents of this mechanism state that this interaction is favored only when the peptide adopts a β-turn or loop conformation.[31] This proposal may therefore be invalidated by the reported kinetic studies on conformationally constrained cyclic peptides which demonstrate a requirement for an Asx-turn and a loss of glycosyl acceptor properties when the substrates are constrained to β-turns.[22,23] Additionally, the kinetic results with peptides **3** and **5** argue against this mechanism.

Finally, a mechanism for asparagine-linked glycosylation has been proposed which incorporates aspects of the favored mechanism for some of the glutamine amidotransferase family of enzymes.[34] This proposal involves the release of ammonia from the asparagine side chain in conjunction with the formation of a cyclic isoasparagine (or succinimide) intermediate. However, kinetic studies with synthetic preparations of the isoasparagine (isoAsn) containing tripeptide Bz-isoAsn-Leu-Thr-NH_2 revealed that this compound was not a competent intermediate in the glycosylation process.[35] Additionally, it should be noted that any mechanism involving backbone amide attack (from the central amino acid amide nitrogen) onto the asparagine side chain is unlikely since this mode of attack would be strongly disfavored in the conformationally constrained substrate, **1**, due to stereoelectronic effects.

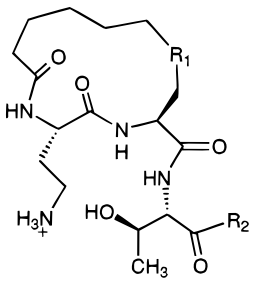
Development of Potent Oligosaccharyl Transferase Inhibitors

The mechanistic studies with the tripeptide Bz-Amb-Leu-Thr-NHMe (**7**) revealed competitive inhibition (K_i = 1 mM) of porcine OT.[25] Although the binding of this

(26) Welply, J. K.; Shenbagamurthi, P.; Lennarz, W. J.; Naider, F. *J. Biol. Chem.* **1983**, *258*, 11856−11863.
(27) Bause, E. *Biochem. Soc. Trans.* **1984**, *12*, 514−517.
(28) Bause, E.; Legler, G. *Biochem. J.* **1981**, *195*, 639−644.
(29) Hortin, G.; Boime, I. *J. Biol. Chem.* **1980**, *255*, 8007−8010.
(30) Shannon, K. L. Ph.D. Thesis, California Institute of Technology, 1992.
(31) Bause, E.; Bruer, W.; Peters, S. *Biochem. J.* **1995**, *312*, 979−985.
(32) Bause, E. *Biochem. J.* **1983**, *209*, 323−330.
(33) Bause, E.; Wesemann, M.; Bartoschek, A.; Breuer, W. *Biochem. J.* **1997**, *322*, 95−102.
(34) Zalkin, H. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1993**, *66*, 203−309.
(35) Clark, R. S.; Banerjee, S.; Coward, J. K. *J. Org. Chem.* **1990**, *55*, 5, 6275−6285.

peptide to the enzyme is weak, it was considered a viable starting point from which to evolve tighter binding analogues.[36] At the onset of these studies, no inhibitors specifically targeted to OT were available. The microbial product tunicamycin[37] is commonly employed to disrupt *N*-linked glycosylation; however, it functions indirectly by inhibiting the first step in the assembly of the oligosaccharide donor (Dol-P-P-$(GlcNAc)_2$-$(Man)_9$-$(Glc)_3$). Thus, use of tunicamycin requires several cell cycles before the supply of the lipid-linked donor is sufficiently depleted to arrest glycosylation.

Building from the simple inhibitory tripeptide (**7**), two priorities were established; to enhance peptide affinity for the target enzyme while developing an efficient synthetic approach. We rationalized that a conformational constraint similar to that employed to enhance substrate acceptor properties might also increase inhibitor affinity. The preparation and analysis of **9** showed that the



9; R_1=CH_2, R_2=NHMe
10; R_1=S, R_2=NHMe
11; R_1=S, R_2=ValThrNphNH$_2$
(Nph=*p*-nitrophenylalanine)

macrolactam constraint in a peptide including the inhibitory amine functionality afforded a peptide with a K_i of 50 μM. To further improve the binding, we turned to statistical studies of *N*-linked glycosylation sites[8] which suggest that the protein modification might be modulated by the identity of the residues beyond the consensus tripeptide sequence, due to interactions between OT and extended binding substrate determinants.

The practical synthesis of the inhibitors involves a facile solid phase method for the assembly of new analogues that utilizes readily available building blocks. In these analogues the macrocycle is formed by a cysteine thiolate alkylation.[36] This procedure was patterned after chemistry developed by Ellman and co-workers for the preparation of constrained β-turn analogues.[38] While the tripeptide **10** exhibited a K_i of 100 μM, similar to the value observed for **7**, extension of the peptide in the *C*-terminal direction dramatically improved binding. Most notably, peptide **11**, which includes the *C*-terminal extension -Val-Thr-Nph-$CONH_2$ (Nph = *p*-nitrophenylalanine), demonstrated slow, tight binding. (The Nph residue was included at the *C*-terminus of the sequence to allow for facile quantification of inhibitor solutions.) Detailed kinetic evaluations using progress curve analyses to afford dissociation and association rate constants reveal a K_i for **11** of 37nM. This constrained peptide is at least 3 orders of magnitude more potent than any previously reported inhibitor of OT. Importantly, **11** selectively inhibits *N*-linked glycosylation; no measurable effect on *O*-linked glycosylation was detected in the in vitro assay of polypeptide *N*-acetylgalactosaminyl transferase.[39]

This class of cyclic peptides provides the first example of a readily available and adaptable family of *N*-linked glycosylation inhibitors.[36] The modular nature of these inhibitors provides immediate opportunities for structural diversification through combinatorial synthesis. For example, while OT has been characterized from several different species and shows significant structural homology throughout eukaryotic evolution, it may be possible to exploit extended-binding interactions to achieve species selectivity. Additionally, these compounds present a readily modifiable platform for the further development of specific glycosylation inhibitors as diagnostic tools to evaluate the role of glycoproteins in biological systems, as potential therapeutic agents, and for the preparation of carbohydrate-depleted glycoproteins for structural studies.

Conformational Consequences of Asparagine-Linked Glycosylation

Co-translational asparagine-linked glycosylation affects the structure and stability of glycoprotein conjugates throughout the protein lifetime.[40] Notably, this co-translational process may facilitate protein folding and assist in the solubilization of partially folded polypeptide intermediates during protein biosynthesis. Without co-translational glycosylation, immature proteins may misfold, aggregate, and be degraded before leaving the ER. After protein synthesis and secretion, the carbohydrate may be implicated in stabilizing specific protein folds to allow for molecular recognition and may also be involved in shielding portions of the protein surface from proteolytic degradation or the immune response.[12] While the significance of glycosylation is not always fully understood, some of the structural effects of glycosylation are now emerging.

A variety of spectroscopic techniques have been employed for the analysis of glycopeptide and glycoprotein structure. For example, NMR, CD, and fluorescence resonance energy transfer (FRET) have been used to provide a structural picture of the polypeptide and to investigate the effect of the oligosaccharide on the structure of the corresponding modified species. In general, X-ray crystallography is not easily applied to glycoproteins due to the heterogeneity of glycan structures and the conformational mobility of the saccharide branches that hamper crystallization.

Studies on the conformational behavior of small peptides and glycopeptides that are patterned after conserved and essential glycosylation sites in native proteins can provide insight into how *N*-linked glycosylation affects protein secondary structure. The advantage of small peptide systems is that a variety of complementary spectroscopic techniques can be enlisted and the direct comparison of glycosylated and nonglycosylated peptides can provide specific information on the conformational consequences of the modification. Additionally, the size

(36) Hendrickson, T. L.; Spencer, J. R.; Kato, M.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 7636−7637.
(37) Elbein, A. D. *Trends Biochem. Sci.* **1981**, *6*, 291−293.
(38) Virgilio, A. A.; Ellman, J. A. *J. Am. Chem. Soc.* **1994**, *116*, 11580−11581.
(39) Wang, Y.; Abernathy, J. L.; Eckhardt, A. E.; Hill, R. L. *J. Biol. Chem.* **1992**, *267*, 12709−12716.
(40) O'Connor, S. E.; Imperiali, B. *Chem. Biol.* **1996**, *3*, 803−812.

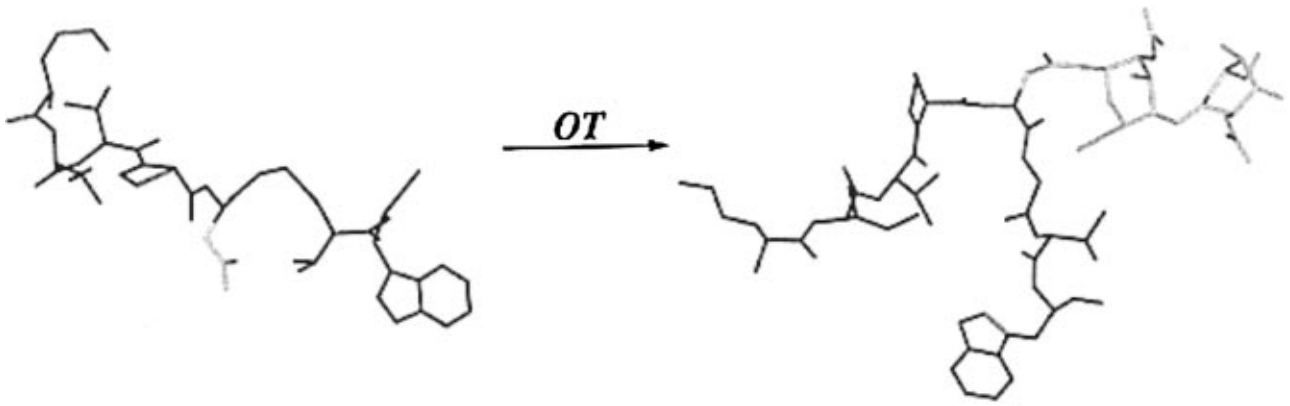


FIGURE 5. Conformational switching of the hemagglutinin A282−A288 peptide by asparagine-linked glycosylation. The asparagine side chain and the disaccharide chitobiose are indicated in gray, and the remaining polypeptide is shown in black.

and flexibility of peptides may mimic the transient and fluxional secondary structure of partially folded, nascent polypeptides in vivo. The comparative analyses of peptides and asparagine-linked glycopeptides in aqueous media has revealed that glycosylation exerts structural effects on the oligopeptide secondary structure that may be critical in the overall folding of many native proteins.[41−48]

In studies from this laboratory, peptides representing the sequences of β-turn surface loops from hemagglutinin A have been examined using time-resolved FRET and NMR techniques.[46,48] In the FRET study, a fluorescently labeled analogue of the A282−A288 sequence (IleThr-ProAsnGlySerIle) of hemagglutinin was examined and the conformational effects of glycosylation analyzed in light of the high-resolution X-ray structure of the intact protein.[49,50] This sequence incorporates a β-turn into a surface loop in the native protein, which is of considerable interest as this motif represents a common feature among the *final* structures of glycosylation sites in many glycoproteins.[51] Additionally, β-turns have been implicated as structural nucleation elements in protein folding. When examined in aqueous solution, the fluorescence studies revealed that glycosylation with a chitobiosyl moiety promoted the adoption of a more compact peptide secondary structure.[48] Subsequent 2D ^{1}H NMR studies[46] supported the FRET analysis and revealed that the peptide adopts an open and extended Asx-turn conformation prior to glycosylation (see Figure 5); in contrast, the glycopeptide exhibits a compact type I β-turn conformation, quite similar to that observed in the final native protein structure. The complementary FRET and NMR studies provide direct evidence that protein glycosylation induces a discrete conformational switch in a surface polypeptide from hemagglutinin. This observation has important implications for the key role played by glycosylation in the correct folding of glycoprotein conjugates. It suggests that in some cases the modification may serve as a critical trigger to bring the peptide to a conformational state that is reflected in the native folded protein, but not in the nascent unmodified sequence.

Site-directed mutagenesis studies also suggest that specific, highly conserved glycosylation sites may play a critical role in subunit folding and oligomerization of multimeric proteins. One example of this phenomenon is found in the nicotinic acetylcholine receptor (nAChR).[52] This pentameric protein complex includes highly conserved 15-residue loop peptides in the soluble, extracellular domains of each subunit. These polypeptides are thought to be involved in intersubunit interactions that are important for the assembly of intact $\alpha_2\beta\gamma\delta$ complexes. The loop sequence in the mature α-1 subunit of the neuromuscular nAChR from *Torpedo Californica* comprises residues 128−143 and includes an asparagine-linked glycosylation site (-AsnCysThr-) and a proline residue (Pro136) at the remote end of the loop formed by the disulfide between Cys128 and Cys142. While site-directed mutagenesis studies have revealed that the glycosylation site is important for the proper assembly and stability of the receptor complex, the mechanism whereby glycosylation influences the structure of the complex was unknown.[53,54] NMR studies of the glycosylated and nonglycosylated loop peptide (AcTyrCysGluIleIleValThrHisPheProPheAspGlnGlnAsnCysThrNH$_2$) of the nAChR α-subunit revealed that derivatization with a truncated chitobiosyl group had a significant influence on the conformational dynamics of the system.[47] After glycosylation, the intramolecular dithiol/disulfide equilibrium was shifted in favor of the oxidized species, suggesting that this modification brings the termini of the loop into closer proximity. Furthermore, in the oxidized peptide, glycosylation significantly altered the *cis*−*trans* proline amide equilibrium (relative to the nonglycosylated species), to favor the *trans* isomer. Since both disulfide bond forma-

(41) Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352−3353.
(42) Davis, J. T.; Hirani, S.; Bartlett, C.; Reid, B. R. *J. Biol. Chem.* **1994**, *269*, 3331−3338.
(43) Live, D. H.; Kumar, R. A.; Beene, X.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12759−12761.
(44) Otvos, L.; Urge, L.; Xiang, Z. Q.; Krivulka, G. R.; Nagy, L.; Szendrei, G. I.; Ertl, H. C. J. *Biochim. Biophys. Acta* **1994**, *1224*, 68−71.
(45) Wormald, M. R.; Wooten, E. W.; Bazzo, R.; Edge, C. J.; Feinstein, A.; Rademacher, T. W.; Dwek, R. A. *Eur. J. Biochem.* **1991**, *198*, 131−139.
(46) O'Connor, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **1997**, *119*, 2295−2296.
(47) Rickert, K. W.; Imperiali, B. *Chem. Biol.* **1995**, *2*, 751−759.
(48) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 97−101.
(49) Wilson, I. A.; Ladner, R. C.; Skehel, J. J.; Wiley: D. C. *Biochem. Soc. Trans.* **1983**, *11*, 145−147.
(50) Wilson, I. A.; Skehel, J. J.; Wiley, D. C. *Nature* **1981**, *289*, 366−373.
(51) Beintema, J. J. *Biosci. Rep.* **1986**, *6*, 709−714.
(52) Stroud, R. M.; McCarthy, M. P.; Shuster, M. *Biochemistry* **1990**, *29*, 9, 11009−11023.
(53) Mishina, M.; Tobimatsu, T.; Imoto, K.; Tanaka, K.; Fujita, Y.; Fukuda, K.; Kurasaki, M.; Takahashi, H.; Morimoto, Y.; Hirose, T.; Inayama, S.; Takahashi, T.; Kuno, M.; Numa, S. *Nature* **1985**, *313*, 364−369.
(54) Gehle, V. M.; Sumikawa, K. *Mol. Brain Res.* **1991**, *11*, 17−25.

tion and *cis*−*trans* proline amide equilibration are among the slower steps in the protein folding process, it becomes clear that *N*-linked glycosylation has the potential to significantly alter the course of protein folding as well as the final folded protein structure. Interestingly, proton NMR analysis of the glycopeptide did not reveal any *specific* hydrogen-bonding interactions between the carbohydrate and the polypeptide, suggesting that the conformational effects of the oligosaccharide arise either from a steric perturbation or from subtle modulations of the local water environment surrounding the polypeptide.

These studies have demonstrated that glycosylation may effect the relative energies of the different conformations that the peptide may sample, thereby altering their representation during the folding process. In the case of the hemagglutinin peptide, the more compact conformation of the glycosylated peptide better approximates the structure of the corresponding region in the final folded protein. Additionally, studies of the nAChR loop peptide have demonstrated that glycosylation can have a significant influence on both disulfide formation and proline isomerization in a local peptide sequence. As the adoption of correct disulfide bonds and proline amide isomers are considered critical steps in protein folding, it is evident that *N*-linked glycosylation plays essential, yet indirect roles in defining the course of the complex folding process. Together, these observations imply that cotranslational glycosylation can instigate structural transitions at a critical time in protein biosynthesis and thus assist in the complex process of protein folding. These studies provide insight as to why the correct folding of many glycoproteins, including the influenza virus coat protein and the nicotinic acetylcholine receptor, may be dependent on cotranslational protein glycosylation.

Biochemical Characterization of Oligosaccharyl Transferase

A review of this fascinating enzyme would be not be complete without a discussion of the current status of the protein biochemistry of the system.[3] Oligosaccharyl transferase is a multimeric, membrane-associated enzyme wherein the functional domains of the enzyme are localized in the lumen of the ER. Future research on the mechanism of action of OT will be greatly advanced by the availability of significant quantities of pure enzyme; however, purification efforts toward this end have been hindered by the lability of OT activity on solubilization from the native membrane-bound environment. Nevertheless, several research groups have recently succeeded in purifying the enzyme to homogeneity from various sources. Evidence now confirms that this enzymatic step has been conserved throughout eukaryotic evolution from yeast to mammals.[3]

The system that has been studied in the most detail is that from the yeast *Saccharomyces cerevisiae*.[3,55] A total of eight distinct subunits have been associated with the activity of the yeast enzyme and each subunit contains at least one trans-membrane hydrophobic domain.[56−63] Additionally, many of the subunits include amino terminus

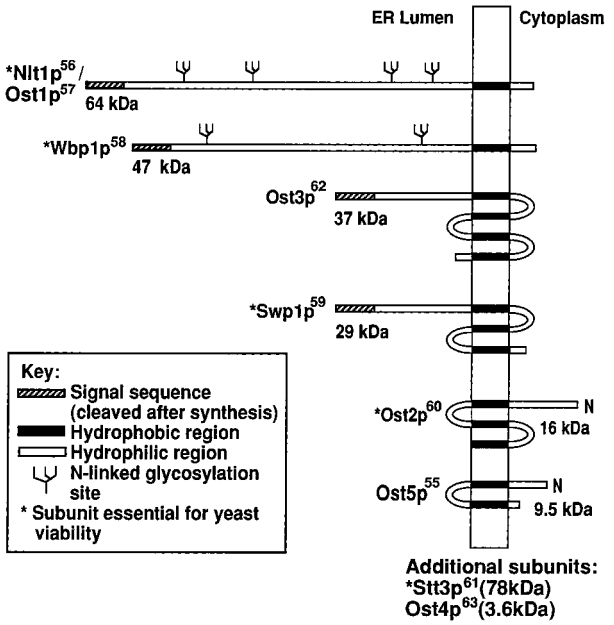


FIGURE 6. Current understanding of the subunit composition of the *S. cerevisiae* oligosaccharyl transferase.

signal sequences which are cleaved prior to complete maturation of the protein. Several of the subunits are glycosylated with *N*-linked high-mannose oligosaccharides, indicating that mature OT is a self-processing enzyme. Genetic knockout experiments have revealed that five of the eight subunits are essential for cell viability and absolutely required for in vivo and in vitro OT activity, while the remaining subunits appear to influence the glycosylation efficiency, but are not essential for catalytic activity. A schematic of the subunit composition of yeast OT is presented in Figure 6. Current efforts in OT biochemistry focus on defining the functional roles of the subunits of the enzyme complex and developing viable expression systems.

The size of the OT enzyme complex inspires several questions. Most importantly, why is the enzyme so complicated? The answer lies at the heart of protein quality control in the secretory pathway. Nascent proteins, bearing an ER-translocation signal sequence must encounter the glycosylation machinery with high fidelity; failures in glycosylation will result in aberrant folding and targeting. Therefore, the complexity of the enzyme is required in part to create the inter-protein contacts that establish "communication" between the ribosome docking and translocation machineries. Additionally, OT must be poised to receive only the fully assembled glycosyl donor. While the truncated disaccharide glycosyl donor DolPP-$(GlcNAc)_2$ may be employed for the in vitro analysis of the enzyme, it is critical that only the unique tetrade-

(55) Kelleher, D. J.; Gilmore, R. *J. Biol. Chem.* **1994**, *269*, 12908−12917.
(56) Pathak, R.; Parker, C. S.; Imperiali, B. *FEBS Lett.* **1995**, *362*, 229−234.
(57) Silberstein, S.; Collins, P. G.; Kelleher, D. J.; Rapiejko, P. J.; Gilmore, R. *J. Cell Biol.* **1995**, *128*, 525−536.
(58) te Heesen, S.; Janetzky, B.; Lehle, L.; Aebi, M. *EMBO J.* **1992**, *11*, 2071−2075.
(59) te Heesen, S.; Knauer, R.; Lehle, L.; Aebi, M. *EMBO J.* **1993**, *12*, 279−284.
(60) Silberstein, S.; Collins, P. G.; Kelleher, D. J.; Gilmore, R. *J. Cell Biol.* **1995**, *131*, 371−383.
(61) Zufferey, R.; Knauer, R.; Burda, P.; Stagljar, I.; te Heesen, S.; Lehle, L.; Aebi, M. *EMBO J.* **1995**, *14*, 4949−4960.
(62) Karaoglu, D.; Kelleher, D. J.; Gilmore, R. *J. Cell Biol.* **1995**, *130*, 567−577.
(63) Chi, J. H.; Roos, J.; Dean, N. *J. Biol. Chem.* **1996**, *271*, 3132−3140.

casaccharide moiety be transferred to the peptide substrate in the OT reaction in vivo as subsequent processing and translocation events depend specifically on the identity of the oligosaccharide that is attached to the protein. Finally, the enzyme makes binding contacts with two substrates of significant proportions: a nascent polypeptide and the dolichol-linked tetradecasaccharide.

Conclusions

The consensus sequence rules for asparagine-linked glycosylation have been known for over two decades. To date, considerable progress has been made in understanding the details of the process, and a clear, but static, definition of the molecular players in the process has evolved. As is evidenced by the identities of the two complex substrates and the elaborate multimeric enzyme that catalyzes the glycosylation process, this reaction, without doubt, represents a molecular version of "The Clash of the Titans"! The next phase of investigations into *N*-linked glycosylation will be focused on generating a molecular description of the catalytic machinery. In this regard, the complementary application of state-of-the-art bioorganic and biophysical methods, together with the tools of contemporary molecular biology, will hopefully lead to our ultimate goal of defining the inner workings of the OT enzyme complex.

The work described in this Account was only possible because of the outstanding intellectual and experimental contribution of my graduate students and postdoctorals. I gratefully acknowledge their dedicated efforts. This program is funded the National Institutes of Health (GM 39334).

AR950226K