Investigating the Pathway of Asparagine-Linked Glycoprotein Biosynthesis

by

Mary K. O'Reilly
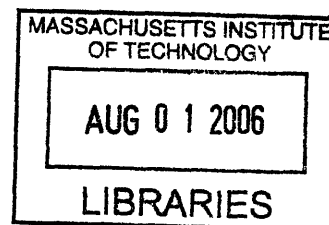
B.S. Chemistry
Purdue University, 1999

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 2006

Signature of Author: _____
Department of Chemistry
March 2, 2006

Certified by: _____
Barbara Imperiali
Class of 1922 Professor of Chemistry and Professor of Biology
Thesis Supervisor

Accepted by: _____
Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor JoAnne Stubbe _____
Chair

Professor Barbara Imperiali _____
Thesis Supervisor

Professor Sarah E. O'Connor _____

Professor Catherine L. Drennan _____

Investigating the Pathway of Asparagine-Linked Glycoprotein Biosynthesis

by
Mary O'Reilly

Submitted to the Department of Chemistry
on March 2, 2006 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ABSTRACT

The biosynthesis of asparagine-linked glycoproteins, highly conserved throughout all eukaryotes, requires a dolichylpyrophosphate-linked tetradecasaccharide precursor (Dol-PP-GlcNAc$_2$Man$_9$Glc$_3$), from which the tetradecasaccharide is transferred co-translationally to nascent polypeptides in the lumen of the ER by the multimeric membrane-bound enzyme, oligosaccharyl transferase (OT). The saccharide donor is assembled by a series of membrane-bound enzymes, which together comprise the dolichol pathway.

Despite over two decades of genetic and bioinformatics approaches that have identified the vast majority of dolichol pathway genes in yeast, the roles of two mannosyltransferases in the pathway, Alg2 and Alg11, remained ambiguous. This thesis describes the biochemical studies that were carried out to clarify these roles. The substrate specificity of Alg1, the first mannosyltransferase in the pathway, was studied, and this enzyme was also used as a tool to prepare Manβ1,4-GlcNAc$_2$-PP-Dol from synthetic GlcNAc$_2$-PP-Dol. Access to this trisaccharide intermediate facilitated the characterization of Alg2 function, proposed to be involved in addition of the second and/or third mannose. A cell membrane fraction isolated from *E. coli* overexpressing thioredoxin-tagged Alg2 was used to demonstrate that this enzyme carries out an α1,3-mannosylation, followed by an α1,6-mannosylation, to form the branched pentasaccharide intermediate of the dolichol pathway. Having the means to access this intermediate chemoenzymatically, it was thus possible to define the function of Alg11, which had similarly been proposed to catalyze addition of the fourth and/or fifth mannose. Using the same procedure, TRX-Alg11 was shown to catalyze two sequential α1,2-mannosylations onto the α1,3-branch of the pentasaccharide substrate to afford the heptasaccharide intermediate. The elucidation of the dual function of each of these enzymes thus completes the identification of the entire ensemble of glycosyltransferases that comprise the dolichol pathway.

Finally, peptidyl mimics based on the consensus site for glycosylation by OT on nascent polypeptides, Asn-Xaa-Thr/Ser, were designed and evaluated. Both substrate-based peptide isosteres and product-based neoglycoconjugates were used to investigate the conformational and stereoelectronic preferences of OT binding. Neoglycoconjugates also showed promise as inhibitors of the deglycosylating enzyme Peptide: *N*-glycanase (PNGase), which aids in the degradation of misfolded proteins in the secretory pathway.

Thesis Supervisor: Barbara Imperiali
Title: Class of 1922 Professor of Chemistry and Professor of Biology

Acknowledgements

My advisor, Professor Barbara Imperiali, and I began our journeys at MIT together, and it has been a great pleasure to share mine with her. Her enthusiasm and good humor set the tone for our open and exceedingly supportive working environment, which she very skillfully maintains. I am also grateful for her patience and her understanding. I would also like to thank my thesis committee, Professors JoAnne Stubbe, Cathy Drennan, and Sarah O'Connor, for the generosity with which they share their time, advice, and support.

For the fact that I am here at MIT to begin with, I am indebted to Professor Jean Chmielewski. Not only did she give me the opportunity to work in her lab as an undergraduate, her excitement and creativity inspired me to pursue graduate school. In the end it was her encouragement and her support that brought me here.

I have been fortunate to collaborate with a number of former and current members of the lab on a variety of endeavors. Vincent Tai taught me the secrets of working with dolichyl compounds, Stephane Peluso taught me peptide synthesis with the beloved hula shaker, Maria de L. Ufret set the bar high with the skill and speed that she demonstrated with the OT assay, and Sarah O'Connor and Carlos Bosques were so helpful to me that it is hard to believe now that we didn't actually work together. I have recently been extremely lucky to have Guofeng Zhang join forces with me, and thanks to him there is a story to tell about Alg11 in this thesis.

For showing great prowess at the MALDI, I have Langdon Martin and Matthieu Sainlos to thank. Very big thanks also go to Jebrell Glover, not only for sharing his expertise with me, especially in the cell membrane prep, but for selflessly assuming the role of Imperiali lab expert in all things biological. For the survival of my computer, I have Melissa Shults, and her patience, to thank.

I would like to thank everyone in the lab for cheering me on through the writing of this thesis, especially Eranthie Weerapana, Bianca Sculimbrene, and Beth Vogel, for reading portions of it. Not only has Eranthie read nearly every chapter of this thesis, she has also been an incredible resource, as I was clever enough to wait for her to go through the process first.

So many of my labmates, former and current, have been much more than that to me, and I thank them for their friendship. Rob Dempski set a great example for critical thinking and at the same time was fun outside of the lab. I am also grateful to Mark Nitz, for challenging me and inspiring me, and for good advice both in and out of the lab, and to Kathy Franz, whose eagerness to solve problems outside of just her own created a stimulating environment. When Eranthie Weerapana joined the lab a year after me, I was asked to take her under my wing. The irony of that situation continues to astound me, considering how much I have come to look up to her. Not only has she been willing to let me "hurt her brain" in thinking through innumerable experimental details with me, she has also sustained me through some extraordinarily difficult decisions with wisdom and compassion, and for all these things I thank her. Many thanks also go to Seungjib Choi, for being a fountain of synthetic advice, and for his realism. To Elvedin Lukovic, for his enormous heart, his humor, and for selflessly defending against awkward silences in any situation. To Galen Loving, for making us all laugh at life, and especially

4

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

Standard 3-letter and 1-letter codes are used for the 20 natural amino acids.

2-AB	2-aminobenzamide
Ac	acetyl
AEBSF	4-(2-Aminoethyl)benzenesulphonyl fluoride
ALG	Asparagine-linked glycosylation
Alloc	allyl oxycarbonyl
ATP	adenosine 5'-triphosphate
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
Bz	benzoyl
calcd	calculated
CAZy	Carbohydrate-Active EnZYmes (database)
cAMP	cyclic adenosine 5'-monophosphate
CDG	Congenital Disorders of Glycosylation
CDI	1,1'-carbonyldiimidazole
Ci	Curies
Dab	diaminobutyric acid
DEAD	diethylazodicarboxylate
DIEA	*N,N*-diisopropylethylamine
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulfoxide
Dol	dolichyl
Dol-P	dolichyl phosphate
Dol-PP	dolichylpyrophosphate
dpm	disintegrations per minute
DTT	dithiothreitol
ε	extinction coefficient or molar absorptivity
ER	endoplasmic reticulum
ESI	electrospray ionization mass spectrometry
Fmoc	fluoren-9-ylmethyloxycarbonyl
E. coli	*Escherichia coli*
ERAD	Endoplasmic reticulum-assisted degradation
GDP-Man	guanosine 5'-diphospho-α-D-mannose
GalNAc	*N*-acetyl-D-galactosamine
GlcNAc	*N*-acetyl-D-glucosamine
Gal	D-galactose
Glc	D-glucose
GT-A/B	glycosyltransferase fold families A/B
HPLC	high-performance liquid chromatography
Hz	hertz
K_i	inhibition constant
K_m	Michaelis constant
λ_{em}	emission wavelength

λ_{ex}	excitation wavelength
LB	Luria-Bertani broth
LRMS	low resolution mass spectrometry
MALDI-MS	Matrix assisted laser desorption ionization mass spectrometry
Man	D-mannose
NMR	nuclear magnetic resonance
Nph	*p*-nitrophenylalanine
Nva	norvaline
O.D.	optical density (absorbance at 600 nm)
OT	oligosaccharyl transferase
P	phosphate
PP	pyrophosphate
PAL	5-(4'-aminomethyl-3',5'-dimethoxyphenoxy)valeric acid
PCR	polymerase chain reaction
PEG-PS	polyethyleneglycol-grafted polystyrene
Pfam	Protein families (database)
PMA	phosphomolybdic acid
PMFS	phenylmethylfluorosulfonate
ppm	parts per million
PSUP	pure solvent upper phase
PyAOP	7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium-hexafluorophosphate
PyBOP	benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium-hexafluorophosphate
RP-HPLC	reverse-phase HPLC
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	sodium dodecyl sulfate
tBu	*tert*-butyl
TBDMS	*tert*-butyldimethylsilyl
TFA	trifluoroacetic acid
TLC	thin layer chromatography
THF	tetrahydrofuran
TIS	triisopropylsilane
TMHMM	TransMembrane prediction using Hidden Markov Models
t_R	retention time
Tris	tris(hydroxymethyl)aminomethane
TRX	thioredoxin
TUP	theoretical upper phase
UDP-GlcNAc	uridine 5'diphospho-*N*-acetyl-_-D-glucosamine
UDP-Glc	uridine 5'diphospho-_-D-glucose
UMP	uridine 5'-monophosphate
Und	undecaprenyl
UPR	Unfolded Protein Response
Xaa	used to denote any amino acid

Chapter I: Introduction

Asparagine-linked protein glycosylation is one of the most prevalent, and also most complicated, protein modifications known. Due to the wide variety of monosaccharide units, glycosyl linkages, branching of oligosaccharides, and diversity of processing by glycosyltransferases and hydrolases, glycans on mature glycoproteins represent a highly heterogenous population.[1] Unlike DNA and protein biopolymers, such non-templated diversity allows for a fine-tuning of the functions of glycans.[2] The biological roles of N-linked glycans are numerous, and include facilitation of protein folding, oligomerization, and structural stability, control of intracellular trafficking, secretion, and catalytic activity of enzymes, protection from proteases, as well as regulation of the expression of receptors on the cell surface and cell-cell adhesion.[3] In viral glycoproteins, such as GP120 in Human Immunodeficiency Virus (HIV) and hemagglutinin in the influenza virus, N-linked glycans serve to modulate antigenicity, evade the immune system, aid in fusion and infectivity, and to help support virion production.

In contrast to the vast heterogeneity of mature glycoproteins, the initial protein glycosylation occurs through the co-translational transfer of a homogenous tetradecasaccharide donor containing N-acetylglucosamine (GlcNAc), mannose (Man), and glucose (Glc) units from the dolichylpyrophosphate-linked donor, $Glc_3Man_9GlcNAc_2$-PP-Dol, to an asparagine residue within the consensus sequon, Asn-Xaa-Thr/Ser.[4] This transfer is mediated by the multimeric, membrane-bound enzyme, oligosaccharyltransferase (OT).[5-7] It is the glycosidase-mediated trimming, first in the ER and then in the Golgi, and the Golgi glycosyltransferase-mediated elaboration that creates the eventual diversity.[8] The rate of glycosylation must meet the demands of protein

16

synthesis, and therefore the machinery for the biosynthesis of the tetradecasaccharide donor ($Glc_3Man_9GlcNAc_2$-PP-Dol), as well as the glycosylation and protein synthesis machinery must be co-localized and intimately coordinated. This thesis will focus primarily on biochemical studies toward understanding the glycosylation steps carried out by the glycosyltransferases comprising the dolichol pathway (Chapters II-V). This thesis will also address the protein glycosylation and deglycosylation steps involved in glycoprotein biosynthesis and quality control, respectively.

Overview of *N*-linked glycoprotein biosynthesis

The secretory pathway of eukaryotes begins on the cytosolic face of the ribosome-docked rough ER. The events occurring from the initiation of protein synthesis to glycosylation involve the precise orchestration of the ER machinery depicted in Figure I-1.[6, 9, 10] Ribosomes are targeted to the ER membrane upon translation of a hydrophobic N-terminal signal sequence, which is a feature of proteins expressed in the secretory pathway, including secreted, membrane-bound, and lysosomal proteins. The signal sequence is recognized by the membrane-bound signal recognition particle, thus recruiting the ribosome to the cytosolic face of the rough ER membrane. Here, the growing polypeptide chain is guided through a pore in the membrane created by the translocon. As the nascent polypeptide emerges from the membrane into the lumen of the ER, a signal peptidase (Sec11p in yeast) is present to cleave the signal sequence (though not all signal sequences are removed). OT, also present at the site of the translocon, glycosylates Asn-Xaa-Thr/Ser sequons once they have cleared the membrane surface by approximately fourteen residues, using the tetradecasaccharide donor provided

by the dolichol pathway.[11] Not all of these sequons are glycosylated, and it is generally believed that the neighboring sequences and local conformational preferences may play a role in transfer to a particular sequence. Details of sequon usage are still being examined.
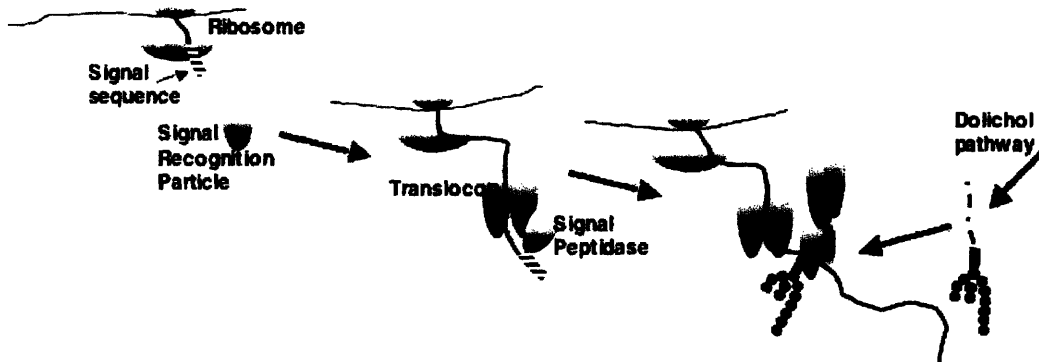
Figure I-1. Protein synthesis and glycosylation machinery in the secretory pathway.

As the biosynthesis of the nascent glycoprotein is proceeding, protein folding ensues. The extent of folding that occurs prior to release from the ribosome is still unclear. To ensure that only correctly folded proteins are produced, eukaryotes have evolved an intricate quality control system, as depicted in Figure I-2.[2, 12, 13] The first steps (in the center of the figure below the membrane) are sequential deglucosylation of the two terminal glucose residues by Glucosidase I and Glucosidase II. The monoglucosylated form of the oligosaacharide is then specifically recognized by the lectin (or carbohydrate-binding) domain in calnexin (CNX) or calreticulin (CRT). These chaperones, with seemingly overlapping functions, are highly homologous to each other, with the major difference being that CNX (not shown in the figure) is membrane-bound, while CRT is a soluble protein in the ER lumen. This lectin/oligosaccharide interaction is

believed to sequester the glycoproteins, affording them an opportunity to fold before

aggregation can occur, and these chaperones also aid in the folding process by recruiting

the protein disulfide isomerase-like protein, ERp57. By the action of Glucosidase II, the

final glucose is removed, thus lowering the affinity of the lectin site for the saccharide

and favoring glycoprotein release.
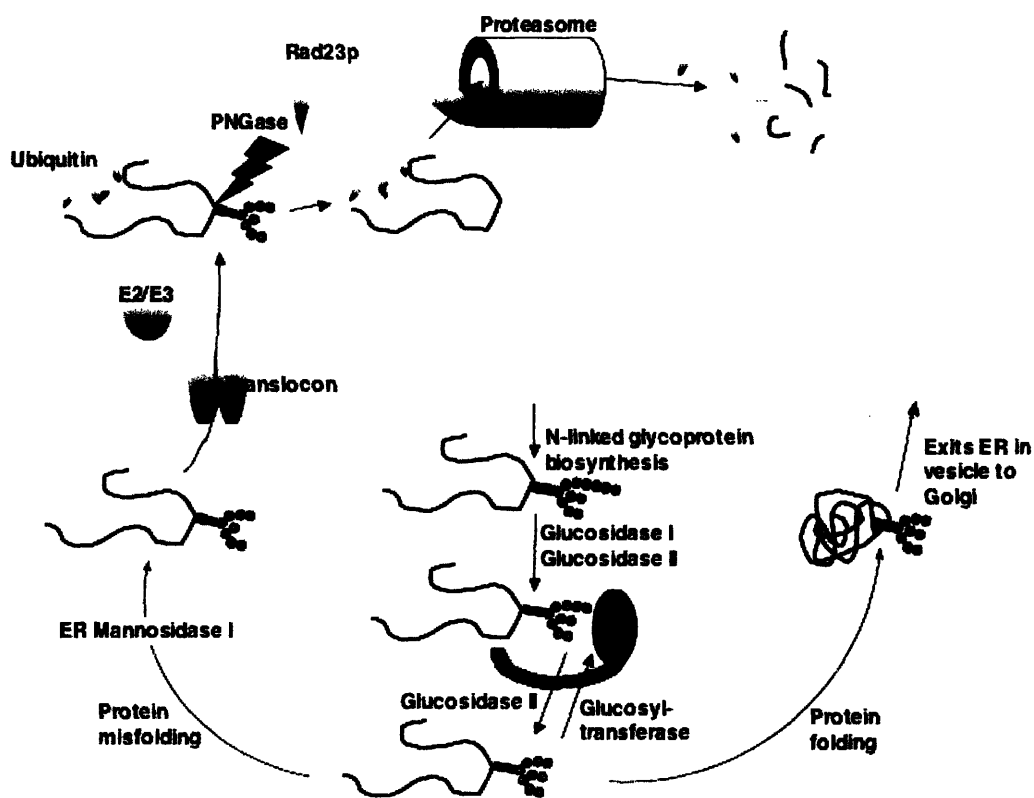
Figure I-2. Quality control pathway of nascent glycoprotein folding.

Upon release from CRT or CNX, proteins may complete the folding process and

then be transported out of the ER in vesicles to undergo saccharide trimming and

elaboration in the Golgi. However, misfolded proteins are recognized by the UDP-Glc:

19

glycoprotein glucosyltransferase (UGGT), and reglucosylated on the $Man_9GlcNAc_2$ core. Regeneration of the monoglucosylated species allows rebinding to CRT or CNX, thus prolonging the opportunity for glycoproteins to fold properly. Glycoproteins may go through several rounds of deglucosylation and glucosylation in what is known as the calnexin/calreticulin (CNX/CRT) cycle, during which time the UGGT effectively acts as a sensor for misfolded protein. This sensing behavior occurs with delicate specificity that favors neither fully folded nor fully misfolded proteins, presumably by recognizing hydrophobic patches on proteins.[14] However, misfolded glycoproteins not associated with CRT/CNX can also be recognized by an α1,2-mannosidase, ER Mannosidase I, which removes a single mannose from the middle branch of the triantennary oligosaccharide species. This species is recognized by the ER degradation-enhanced alpha mannosidase-like protein (EDEM), and is then destined for ER-assisted degradation (ERAD). Thus, nascent proteins that are unable to fold properly do not remain CNX/CRT cycle *ad infinitum*. Upon retrotranslocation to the cytosol, misfolded glycoproteins are ubiquitinated by the action of ubiquitin-conjugating enzyme (E2) and ubiquitin ligase (E3), deglycosylated by the soluble enzyme Peptide: *N*-glycanase (PNGase), and finally degraded by the proteasome. PNGase acts by cleaving the asparagine amide to release the glycan in a manner similar to cysteine proteases. Through the chemical synthesis of the relevant oligosaccharide species, molecular-level details of the specificity and binding affinities of lectin-like domains in UGGT, CRT, and E3 have been elucidated.[15, 16]

Overview of the dolichol pathway

To accomplish the biosynthesis of the dolichylpyrophosphate-linked tetradecasaccharide (Glc$_3$Man$_9$GlcNAc$_2$-PP-Dol) donor for N-linked protein glycosylation (Figure I-3), a series of membrane-bound glycosyltransferases that comprise the dolichol pathway catalyze the sequential transfer of monosaccharide units onto a membrane-bound dolichylpyrophosphate carrier.[4, 17-19]
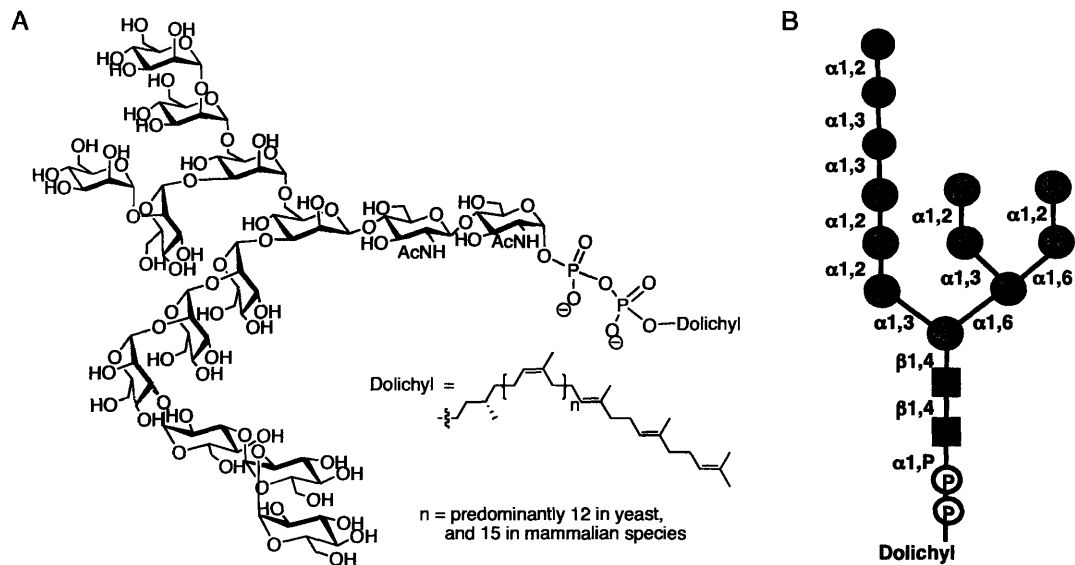
Figure I-3. A) Structure of the tetradecasaccharide product of the dolichol pathway. The gray background highlights the saccharides that are transferred on the cytosolic face of the ER membrane. B) Cartoon diagram of the tetradecasaccharide product indicates the configuration of the glycosidic linkages. (black squares, GlcNAc; gray circles, Man; black circles, Glc; P, phosphate).

As shown in Figure I-4, The first seven transferases, including one glycosyl-phosphate transferase and six glycosyltransferases, build up the intermediate $Man_5GlcNAc_2$-PP-Dol on the cytosplasmic face of the ER membrane prior to translocation of the intermediate to the lumenal face of the membrane. The topology of these events was established by testing the sensitivity or resistance of the glycosyltransferase activities in intact vesicles to proteases, and the recognition of dolichylpyrophosphate-linked saccharide intermediates to the mannose-specific lectin, Concanavalin A.[20, 21]

cytosol

SEC59   ALG7   ALG13/14   ALG1   ALG2?   ALG2?   ALG11?   ALG11?

RFT1

ER lumen

ALG3

ALG9

OT   ALG6   ALG9   ALG12
      ALG8
      ALG10
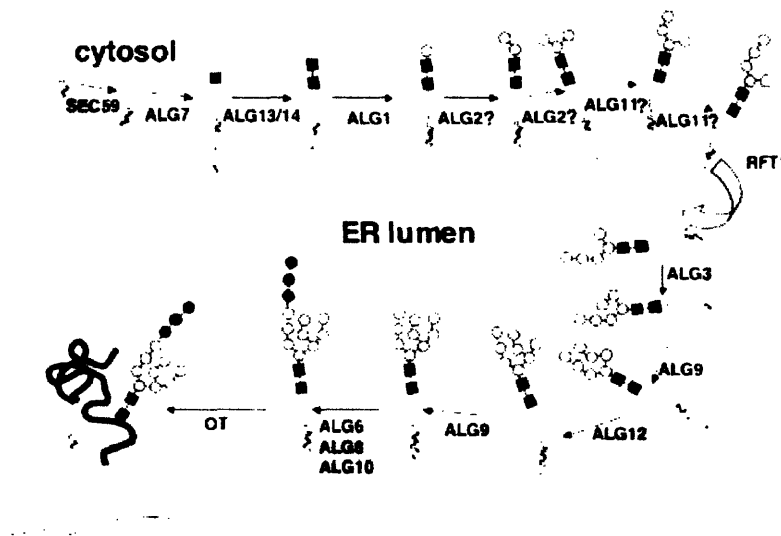
Figure I-4. The sequence and topology of the glycosyltransferase steps, with the proposed enzyme associated with each step indicated. (black squares, GlcNAc; gray circles, Man; black circles, Glc; P, phosphate; wavy lines, dolichyl group).

The translocation, or "flipping", of the heptasaccharide intermediate has long been suspected to be protein-dependent, but only recently was a candidate gene, *RFT1*, identified.[22] The mechanism of this process is still unknown, and it does not appear to be

ATP-dependent. Some evidence exists for the dolichyl-induced modification of phospholipid packing of the membrane to induce a hexagonal phase.[23] This dispruption of the bilayer may aid in the flipping of the heptasaccharide. Interestingly, NMR studies of dolichol reveal a coiled structure in membrane-mimicking environments that may contribute to membrane dispruption.[24-26] Structures solved with the addition of native peptides taken from transmembrane-spanning regions of dolichyl-binding proteins revealed altered conformations that may also be expected to affect the membrane packing in conjuncation with dolichol.

Once translocation has occurred, the remaining glycosyltransferases act on the lumenal side of the membrane to complete the assembly of the tetradecasaccharide. To ensure the proper assembly of N-linked glycoproteins, this assembly process follows a strict order of addition, and both Rft1 and OT are believed to demonstrate a strong preference for the heptasaccharide and tetradecasaccharide structures, respectively.[22, 27] These features are important, as the proper assembly of the full oligosaccharide is crucial for the recognition of nascent glycoproteins by the lectins of the quality control pathway for the prevention of misfolded protein production.

Glycosyl donor substrates for oligosaccharide biosynthesis

The glycosyl donors that provide the monosaccharide units for the growing oligosaccharide differ in the activating group, depending on whether these substrates are being used in the in the cytosol or in the ER lumen (Figure I-5).[18] In the cytosol, the nucleotide-activated sugars UDP-GlcNAc and GDP-Man, which are not

transported into the ER lumen, are used as substrates. Dolichylphosphate-linked monosaccharide donors, Man-P-Dol and Glc-P-Dol, are used as substrates in the lumen for the final seven glycosylation steps. These dolichylphosphate-linked substrates are biosynthesized on the cytoplasmic face of the ER membrane from Dol-P and the nucleotide-activated sugars, GDP-Man and UDP-Glc, by Dpm1 and Alg5 (*S. cerevisiae*), respectively, and then undergo a translocation to face the lumen. There are not yet any candidate proteins for these translocation processes, but studies in liposomes argue against passive transport.[28] While the absence of GDP-Man in the lumen makes Man-P-Dol transport necessary, it is not clear why Glc-P-Dol must be synthesized in the cytosol, since UDP-Glc does get transported into the lumen.
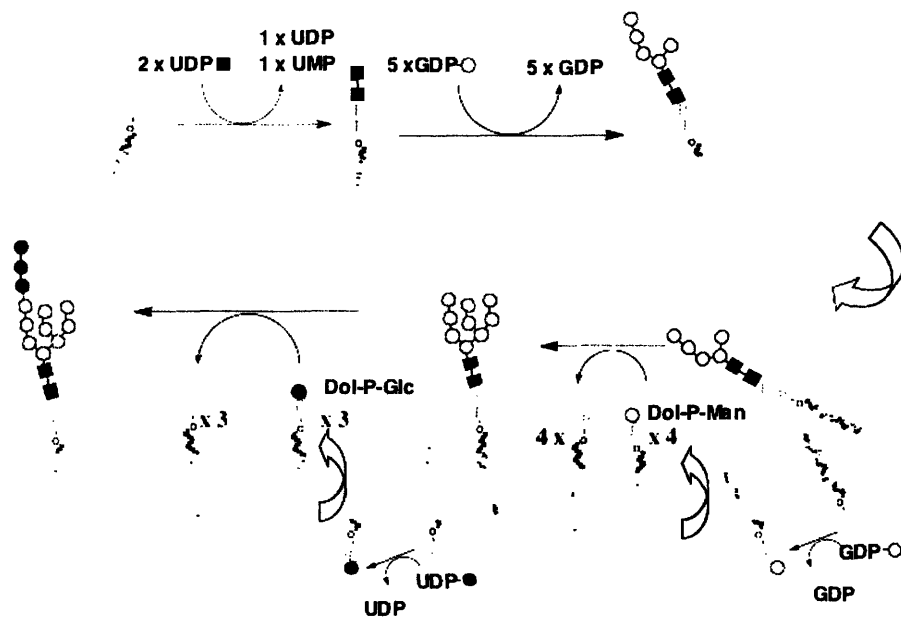
Figure I-5. Summary of the activated monosaccharide donor substrates for the glycosyltransferases in the dolichol pathway. (black squares, GlcNAc; gray circles, Man; black circles, Glc; P, phosphate; wavy lines, Dol; open arrows indicate translocation).

Biosynthesis and recycling of the carrier group, dolichyl phosphate (Dol-P)

Dolichyl phosphate (Dol-P) acts as the carrier for the growing oligosaccharide chain, anchoring it to the membrane, where the glycosyltransferases that transfer the saccharide units reside. In yeast, the biosynthesis of dolichol begins along the same path as sterol and ubiquinone biosynthesis up to the farnesylpyrophosphate intermediate, and then synthesis is completed by a cis-prenyltransferase-catalyzed condensation of isopentenylpyrophosphate groups.[29, 30] Pre-steady state kinetics and X-ray crystallographic data for the cis-prenyltransferase family are beginning to reveal how chain length is controlled in polyisoprene biosynthesis, and site-directed mutations have been demonstrated to produce altered chain lengths.[31, 32] To convert the polyisoprenoid to dolichol, a polyprenol reductase reduces the double bond in the α-isoprene unit, though the details of this step remain unclear.[33] To initiate entry into the dolichol pathway, dolichol is then phosphorylated by a CTP-dependent kinase, an activity that has been assigned to SEC59 in S. cerevisiae and hDK1 in humans.[34, 35] The subsequent pool of Dol-P is available for Man-P-Dol, Glc-P-Dol, and GlcNAc-PP-Dol synthesis, and it has been shown that all three synthases compete for the same pool of Dol-P.[36]

Recycling of Dol-P can occur via three sources (Figure I-6). Following the transfer of the tetradecasaccharide donor to nascent proteins, Dol-PP is generated. A dolichyl pyrophosphate phosphatase has now been identified in both S. cerevisiae (CWH8) and in mammalian cells. This enzyme removes the terminal phosphate and, at a slower rate, the remaining phosphate to generate dolichol.[37, 38] Similarly, the lumenal glycosyltransferases generate Dol-P upon transfer of Man or Glc from Man-P-Dol or Glc-

P-Dol, respectively. Dol-P can then be dephosphorylated, and dolichol is believed to freely diffuse through the ER membrane to return to the cytosolic face, where the dolichol kinase regenerates Dol-P.
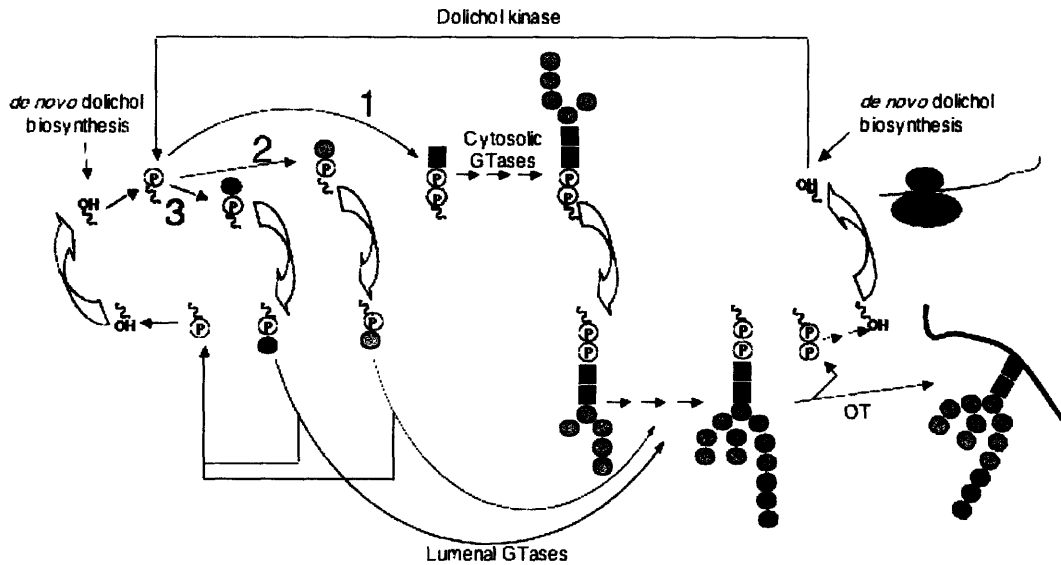
Figure I-6. The three possible paths of dolichyl phosphate are conversion to GlcNAc-PP-Dol (**1**), Man-P-Dol (**2**), or Glc-P-Dol (**3**), and the byproducts of these reactions are used to recycle dolichol through the action of a pyrophosphatase. (black squares, GlcNAc; gray circles, Man; black circles, Glc; P, phosphate; wavy lines, dolichyl group; open arrows indicate translocation; straight arrows indicate recycling pathway).

The glycosyl-1-phosphate transferase and glycosyltransferases of the dolichol pathway

The first GlcNAc is added to the dolichyl phosphate carrier by the transfer of GlcNAc-phosphate from UDP-GlcNAc by the enzyme GlcNAc phosphotransferase (GPT), also known as Enzyme I.[39] This step is inhibited by the natural product tunicamycin. Exploiting the specific inhibition by this natural product, *ALG7* was identified as the gene responsible for this transformation by screening segments of the *S. cerevisiae* genome to find the sequence that, upon overexpression, conferred resistance to tunicamycin.[40] Alg7 is a 448-residue, 50 kDa, enzyme with nine predicted transmembrane domains (Table I-1).[41] Oligomerization of this enzyme has been proposed based on chemical crosslinking studies and dominant negative effects of co-expressing wild-type with mutant GPT enzymes from hamster microsomal membranes.[42] While partial purification from solubilized microsomes has enabled some biochemical characterization, the large number of transmembrane domains makes it challenging to overexpress and purify using current technologies.

The enzyme responsible for the second GlcNAc transfer, referred to as Enzyme II due to its position in the pathway,[39] had remained elusive until very recently, despite numerous genetic screens for defects in this pathway. Bioinformatics has become a powerful tool for identifying genes based on sequence homology to enzymes with a similar function. In the case of Enzyme II, an analogous step is carried out in bacterial cell wall peptidoglycan biosynthesis by the UDP-GlcNAc undecaprenyl-PP-MurNAc pentapeptide: *N*-acetylglucosaminyltransferase known as MurG.[43-45] Due to the similarity

in function with Enzyme II, the sequence of MurG was used to search for homologous sequences that may correspond to that of Enzyme II. A small yeast protein, YGL047w, was identified by searching within the protein family (PF03033) to which the C-terminal domain of MurG belongs.[46] However, due to the small size and the lack of a signal sequence and transmembrane domains, YGL047w initially seemed an unlikely candidate for Enzyme II activity. A search for mammalian sequences in the PF03033 family that are similar to the N-terminal domain of MurG was carried out with no success. Cps14g, a glycosyltransferase involved in bacterial capsular polysaccharide biosynthesis, also bears similarities to the C-terminal domain of MurG, and has an accessory protein, Cps14f, that is critical for full activity.[47] Thus, by homology to Cps14f, the yeast protein YBR070c was identified, which contains a signal sequence and 1-2 predicted transmembrane domains.[46] Down-regulation of either YBR070c or YGL047w in yeast leads to defects in growth and glycosylation.[46, 48] Evidence that an accumulation of GlcNAc-PP-Dol results from down-regulation of either of these genes implicated both gene products in the second step of the dolichol pathway.[46, 48] Thus, the Enzyme II activity was assigned to YGL047w and YBR070c, which were designated *ALG13* and *ALG14* (Table I-1). Physical interaction between Alg13 and Alg14 was demonstrated by co-fractionation of these proteins upon affinity chromatography of Alg14, and localization studies showed that Alg14 is capable of recruiting the soluble Alg13 to the cytoplasmic face of the ER membrane.[48, 49] Current work in the Imperiali lab is focused on validating the function of these proteins *in vitro*, and research has shown that both subunits are necessary and sufficient for the transfer of GlcNAc from UDP-GlcNAc to Dol-PP-GlcNAc to form the disaccharide intermediate in the dolichol pathway.[50]

Step in pathway	Gene	# Amino acids/ MW	Predicted TMD (TMHMM server, v. 2.0)	Accession # Primary reference
1) DolP + UDPGlcNAc	*ALG7*	448 AA 58,368 Da	7	P07286/[40]
2) DolPPGlcNAc + UDPGlcNAc	*ALG13*	202 AA 22,661 Da	0	P53178/[46]
	ALG14	237 AA 27,035 Da	1	P38242/[46]
3) DolPPGlcNAc$_2$+ GDPMan	*ALG1*	449 AA 51,929 Da	1	P16661/[51]
4) DolPPGlcNAc$_2$Man+ GDPMan	*ALG2?*	503 AA 58,047 Da	3	P43636/[53][56][57]
5) DolPPGlcNAc$_2$Man$_2$+ GDPMan	*ALG2?*			
6) DolPPGlcNAc$_2$Man$_3$+ GDPMan	*ALG11?*	548 AA 63,143 Da	1	P53954/[58][61][62]
7) DolPPGlcNAc$_2$Man$_4$+ GDPMan	*ALG11?*			
8) DolPPGlcNAc$_2$Man$_5$+ DolPMan	*ALG3*	458 AA 52,860 Da	9	P38179/[53]
9) DolPPGlcNAc$_2$Man$_6$+ DolPMan	*ALG9*	555 AA 63,776 Da	7	P53869/[65]
10) DolPPGlcNAc$_2$Man$_7$+ DolPMan	*ALG12*	551 AA 62,672 Da	9	P53730[27]
11) DolPPGlcNAc$_2$Man$_8$+ DolPMan	*ALG9*	555 AA 63,776 Da	7	P53869/[66]
12) DolPPGlcNAc$_2$Man$_9$+ DolPMan	*ALG6*	544 AA 62,783 Da	10	Q12001/[71]
13) DolPPGlcNAc$_2$Man$_9$Glc+ DolPMan	*ALG8*	577 AA 67,385 Da	12	P40351/[72]
14) DolPPGlcNAc$_2$Man$_9$Glc$_2$+ DolPMan	*ALG10*	525 AA 61,808 Da	11	P50076/[73]

Table I-1. Summary of dolichol pathway enzymes

ALG1, the yeast gene responsible for the first mannosylation step that forms the Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol intermediate was identified by a yeast genetic screen known as mannose suicide selection that isolates mutants deficient in *N*-linked glycoprotein biosynthesis.[51] Mutants with temperature-sensitive phenotypes were identified by the ability to grow at the permissive temperature after metabolic labeling with tritiated mannose followed by long-term storage. Cells with normal levels of glycosylation were less likely to survive the prolonged radiation damage due to the high-

mannose characteristic of yeast glycoproteins. Alg1 is a 449-residue, 52-kDa, enzyme with one predicted transmembrane domain at the N-terminus (Table I-1).[41] Based on immunoprecipitation and gel filtration experiments of detergent-solubilized Alg1, it appears that this enzyme oligomerizes, though the extent to which this occurs and the physiological relevance are not yet known.[52] Alg1 is discussed further with respect to substrate specificity in Chapter III.

For the final four cytosolic mannosylation steps, only two candidate genes have been identified. It has been suggested that Alg2 is responsible for either the first, second, or both the first and second mannosylation steps,[53-57] and that Alg11 is responsible for either the third, fourth, or both the third and fourth mannosylation steps (Figure I-4).[7, 52, 58] The fact that both *ALG2* and *ALG11* are essential genes precludes the use of complete deletion strains to clarify the roles of these enzymes, although plasmids containing these genes were able to complement the defects in *alg2* and *alg11* yeast mutants, respectively.[54, 58] The use of bioinformatics has greatly facilitated the identification of missing enzymes such as Alg12 and Alg13/14, but no candidates have been generated for any remaining mannosyltransferase activities.

Although Alg2 mutants have been identified and characterized in *S. cerevisiae*,[53] *R. pusillus*,[56] and human,[57] the results have not been entirely consistent. The yeast mutants, *alg2-1* and *alg2-2*, which have a temperature-sensitive phenotype, each contain two point mutations, with one in common, G377R, that is sufficient to confer temperature-sensitivity.[55] The human mutation results in a truncated gene product due to a frame shift that introduces a premature stop codon. None of these mutations, in yeast or human, affects the part of a conserved signature sequence

(XHXGXXXXEXXXXGXXP) that is shared by a number of other glycosyltransferases involved in a wide range of functions in carbohydrate metabolism, including Alg11 and MurG.[59] In the Alg2 sequence, as well as Alg11 and many other glycosyltransferases from the same Pfam family,[60] the signature sequence also contains a conserved EX_7E motif (Figure I-7). The first Glu residue of this motif corresponds to the residue immediately N-terminal to the His residue in the signature sequence, and the second Glu corresponds to the conserved Glu in the signature sequence. Both the human and yeast mutants, with intact signature sequences, accumulate dolichylpyrophosphate-linked trisaccharide and tetrasaccharide intermediates. In contrast, the mutation found in the filamentous fungi, *R. pusillus*, leads to a truncation of Alg2 that is more severe than that found in the human mutant, resulting in a loss of the signature sequence and an accumulation of only the trisaccharide intermediate.[56] Interestingly, *ALG2* is not an essential gene in this organism, and due to the severity of the mutation described in this study, the mutant may in effect represent an *ALG2* deletion strain. Alg2 from *S. cerevisiae* is a 503-residue, 58-kDa mannosyltransferase containing four predicted transmembrane domains, with one pair of domains at the extreme C-terminus, and another pair of domains located near the N-terminus, as predicted by the TMHMM server, v. 2.0 (Table I-1).[41] Chapter IV describes the biochemical validation that was carried out to show that Alg2 is responsible for both the α1,3- and the α1,6-mannosylation steps to form the core pentasaccharide, Manα1,3-(Manα1,6)-Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol.

```
TPAYEHFGIVPLEAMK      Alg2 S. cerevisiae
TPSNEHFGIVPLEAMY      Alg2 human
TPSNEHFGITPVEGMY      Alg2 R. pusillus
AMWNEHFGIAVVEYMA      Alg11 S. cerevisiae

                      G377R

                          alg2 ts yeast mutant

                          CDGIi human mutant

                          R. pusillus mutant
```
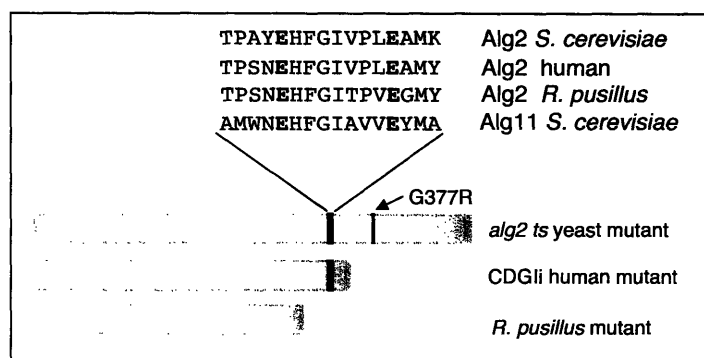
Figure I-7. Summary of Alg2 mutants characterized to date, highlighting the conserved signature sequence also shared by Alg11.

A yeast mutant, *alg11*, was isolated based on resistance to sodium vanadate and shown to affect glycosylation at an early step in the pathway.[61, 62] Efforts to define the precise step for which Alg11 is responsible included extensive characterization of accumulated dolichylpyrophosphate-linked saccharide intermediates and glycoproteins extracted from *alg11* mutant cells.[58] While non-specific translocation into the lumen and elongation by the lumenal glycosyltransferases complicated the interpretation of these studies, Alg11 was proposed, based on these results, to catalyze the fourth or fifth mannosylation. In later studies on the translocation of intermediates from the cytosol to the lumen by Rft1, *alg11* mutant yeast cells were found to accumulate the dolichylpyrophosphate-linked $GlcNAc_2Man_3$ intermediate prior to the flipping step, suggesting that Alg11 may be responsible for the fourth mannosylation.[22] Alg11 from *S. cerevisiae* is a 548-residue, 63-kDa, protein with one predicted transmembrane domain at the N-terminus, as predicted by the TMHMM server, v. 2.0 (Table I-1).[41] Alg2 and Alg11 have been found to interact in separate subcomplexes with Alg1, though not with one another.[52] Chapter V describes the biochemical characterization that was carried out

to show that Alg11 is responsible for the two α1,2-mannosylation steps that convert the pentasaccharide intermediate to the heptasaccharide intermediate, Manα1,2-Manα1,2-Manα1,3-(Manα1,6)-Manβ1,4-GlcNAc-β1,4GlcNAc-PP-Dol.

The lumenal glycosyltransferases are all multi-transmembrane spanning enzymes that use dolichyl-linked substrates as both glycosyl donor and glycosyl acceptor. Since the later steps in the dolichol pathway are not essential for viability, and the glycosylation defects are much more mild, it is more difficult to identify phenotypes associated with defects in these later steps. This limitation can be overcome by generating synthetic lethal mutants. In these screens, mutants are made in combination with mutations in OT subunits that decrease the efficiency of transfer of the truncated saccharides to protein, leading to underglycosylation and a pronounced effect of the desired phenotype.

The first Man-P-Dol dependent mannosyltransferase was discovered by a number of yeast genetic screens. The *alg3* mutant was characterized by accumulation of Man$_5$GlcNAc$_2$-PP-Dol.[53] The *ALG3* gene was later isolated, and found to be a 53.5-kDa enzyme with multiple predicted membrane-spanning domains (Table I-1).[63] Isolation of Alg3 by immunoprecipitation from a yeast solubilized extract enabled the validated of Alg3 as the mannosyltransferase responsible for the first lumenal mannosyltransfer to form Man$_6$GlcNAc$_2$-PP-Dol.[64] Alg3 was further found to act, surprisingly, in a metal-independent fashion, making it the only metal-independent transferase in the pathway.

An *alg9* mutant was isolated that had a synthetic lethal phenotype at 30 °C with a mutant of the OT subunit, *wbp1*.[65] This mutant was found to accumulate the Man$_6$GlcNAc$_2$-PP-Dol intermediate, and *ALG9* was therefore assigned the function of the α1,2-mannosyltransferase directly following the Alg3 step. Recently, microsomes from

an *alg9* mutant strain were shown to elongate a $Man_7GlcNAc_2$-PP-Dol intermediate to $Man_8GlcNAc_2$-PP-Dol, with no evidence of further elongation, suggesting that Alg9 may also carry out the second $\alpha1,2$-mannosylation step in the lumen to form the $Man_9GlcNAc_2$-PP-Dol intermediate.[66] While historically it has been assumed that one glycosyltransferase can only be responsible for the formation of one glycosidic linkage, examples are now emerging that challenge this view. For example, in *S. cyanogenus*, Landomycin A biosynthesis involves the assembly of a hexasaccharide moiety by two monofunctional and two bifunctional glycosyltransferases.[67] In *Campylobacter jejuni*, the $\alpha1,4$-*N*-acetylgalactosaminyltransferase, PglH, acts iteratively to transfer three GalNac units to the growing oligosaccharide chain in the recently discovered Pgl pathway of *N*-linked glycoprotein biosynthesis, which has many similarities to the dolichol pathway.[68]

The remaining mannosyltransferase in the pathway is encoded by *ALG12* and is assigned to the $\alpha1,6$-mannosyltransferase activity in between the two Alg9-mediated steps based on the accumulation of the $Man_7GlcNAc_2$-PP-Dol intermediate. This gene was first identified in a yeast screen for mutants that affected cell surface biosynthesis.[69] The role of Alg12 in the dolichol pathway was later uncovered by using the Alg9 sequence in a database searching algorithm used to identify candidate lumenal mannosyltransferases.[27] The phenotype of the *alg12* mutant is very mild, and full glucosylation is often observed.[70]

Finally, *ALG6*, *ALG8*, and *ALG10* are responsible for the three glucosylations that complete the biosynthesis of the core oligosaccharide.[71-73] While Alg6 and Alg8 are highly homologous and catalyze the same linkage formation, they do not complement one another *in vivo*. Similarly, Alg10 is not able to glucosylate the monoglucosylated

saccharide intermediate. Further highlighting the specificity of the glucosyltransferases, an *alg3* mutant strain transfers the accumulated heptasaccharide intermediate to protein, and only trace amounts of the glycan were found to be glucosylated.[74]

Oligosaccharyltransferase

The dolichol pathway culminates in the OT-catalyzed transfer of the complete tetradecasaccharide structure to Asn in nascent polypeptides in the reaction shown in Figure I-8.

R = GlcNAcMan$_9$Glc$_3$
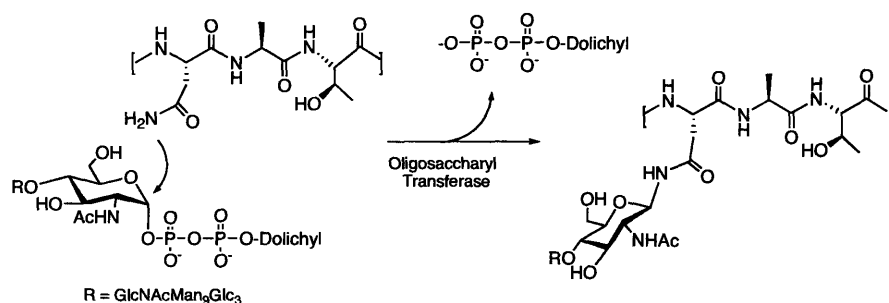
Figure I-8. The glycosylation reaction catalyzed by OT.

While OT has a strong preference for the complete tetradecasaccharide, transfer of truncated saccharides is observed in the analysis of glycoproteins in mutants that are deficient in some step of the dolichol pathway.[75] In yeast, there is evidence for transfer of intermediates as limited as the trisaccharide, ManGlcNAc$_2$, which also highlights the relaxed specificity of the flippase, Rft1, in mutant yeast strains.[75] However, the efficiency of glycosylation is severely reduced in these cases, and *in vitro* studies have revealed a significant degree of specificity for the full donor oligosaccharide. For

example, the absence of the three Glc units causes a reduction in the rate of transfer by 5 to 20-fold.[76] More detailed kinetic studies using purified OT from yeast and canine microsome sources have suggested that cooperativity in oligosaccharide donor binding accounts for the observed preference for the fully assembled form.[7, 76, 77] It has been proposed that there is both a regulatory and a catalytic binding site for the dolichylpyrophosphate-linked donors, and that occupation of the regulatory site tunes the selectivity of the catalytic site by changing the preference for the full-length donor over assembly intermediates. The regulatory site has a similar affinity for the tetradeca- and undecasaccharide donors, and thus, in a mixed population of intermediates, will bind a representative sample of the population. In the presence of incomplete intermediates, binding in the regulatory site changes the preference of the catalytic site for the full-length over truncated intermediates from 1.5-fold to 7 to 10-fold higher. This modulation would ensure that in the presence of truncated sugars, the binding of the full-length donor is favored.

GlcNAc$_2$-PP-Dol is a well-established substrate for OT *in vitro* and it is commonly used for activity assays.[78] The specificity of OT for the GlcNAc units on the glycosyl donor has been investigated *in vitro* using synthetic analogs.[79] These studies revealed that GlcNAc-PP-Dol, while being a poor substrate (11% compared to GlcNAc$_2$-PP-Dol), is the minimal donor. Synthetic disaccharide derivatives in which the *N*-acetyl group of the GlcNAc directly linked to pyrophosphate was replaced with either a fluoro- or trifluoroacetamido- group were not substrates. These results suggested that the *N*-acetyl group is a critical determinant for catalysis. The synthetic disaccharide derivatives do, however, display inhibitory behavior, with K_i values for the fluoro- and

trifluoroacetimido-derivatives of 252 μM and 154 μM, respectively. Many details of the recognition of the acceptor substrate, which is the nascent polypeptide, have also been elucidated through the use of peptide-based OT substrate and inhibitor analogs by comparison to substrate peptides with the native consensus sequence for glycosylation.[80-82] Our recent contributions to these studies are discussed in more detail in Chapter VI.

The oligosaccharyltransferase complex is composed of membrane-bound subunits, which have been assigned to subcomplexes by a combination of biochemical and genetic studies.[7, 83] The specific number of subunits varies according to species, but in yeast, there are nine subunits that each belong to one of three subclusters, as depicted in Figure I-9. Five subunits are essential for viability (Stt3, Swp1, Wbp1, Ost1, and Ost2), and biochemical approaches such as photo-crosslinking have been employed to identify which is the catalytic subunit.
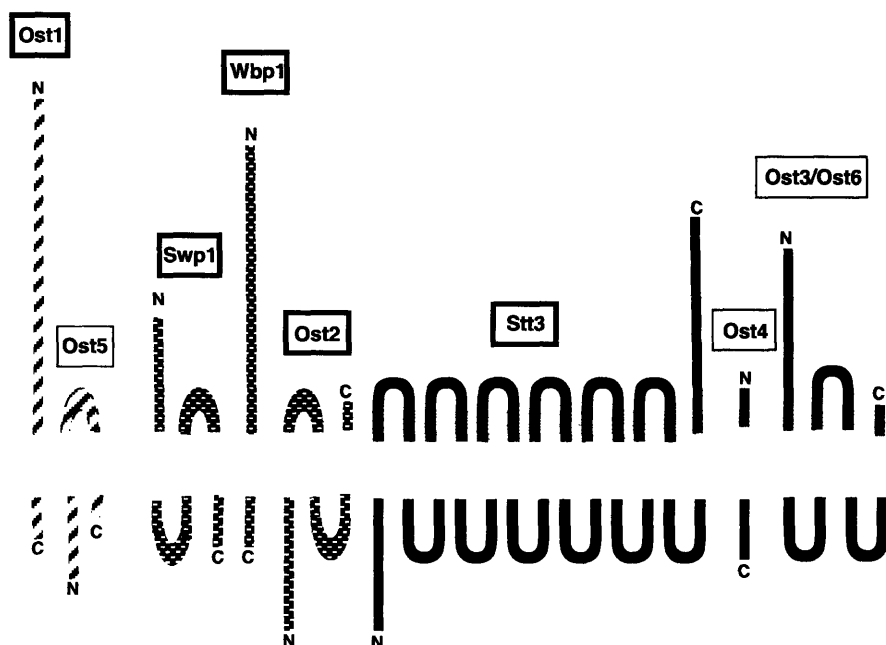
Figure I-9. OT subunits identified in yeast. Subcomplexes are distinguished by pattern. Dark boxes around names indicate essential subunits.

Homology searches have found that Stt3 is the most highly conserved subunit, although the large number of transmembrane-spanning regions casted doubt on using this criterion as evidence that Stt3 is the catalytic domain. However, the recent discovery of the Pgl pathway of *N*-linked glycosylation in the gram-negative bacteria, *Campylobacter jejuni*, with many similarities to the eukaryotic dolichol pathway, revealed a single OT-like protein, PglB, which is a homologue of Stt3.[84] The Pgl gene cluster includes, in addition to PglB, all of the genes necessary for the biosynthesis of the oligosaccharide donor. Expression of the Pgl gene cluster in *E. coli* led to native glycosylation of the *C. jejuni* protein AcrA.[85] Since *N*-linked glycosylation is not present in *E. coli*, these results suggested that PglB is sufficient to carry out the glycan transfer. Due to the high degree of homology between PglB and Stt3, this result makes Stt3 the likely candidate for the catalytic subunit of OT. Interestingly, PglB appears to recognize an expanded version of the known consensus sequon for eukaryotic glycosylation that includes Asp-Xaa-Asn-Xaa-Ser/Thr. Residues within PglB that are critical for catalysis have been localized to a WWDYG motif, which is highly conserved. Mutations in this region have lead to a loss of transferase activity.[86] Efforts are now underway to biochemically confirm that the Stt3 subunit alone is sufficient to carry out protein glycosylation.

Regulation of the dolichol pathway

Aspects of the regulation of tetradecasaccharide donor biosynthesis were briefly introduced previously, including the competition of the Dol-P dependent GlcNAc-1-P transferase, GPT-1 (Alg7 in *S. cerevisiae*), Dol-P-Man synthase, and Dol-P-Glc synthase

for the available pool of Dol-P. Interestingly, further support for this model was provided by the more recent observation that in gently permeabilized Chinese hamster ovary cells, treatment with compounds that block protein translation causes an accumulation of $Glc_3Man_9GlcNAc_2$-PP-Dol with a concomitant loss of activity of the dolichol pathway enzymes.[87] In contrast, full activity of the enzymes was confirmed in microsomes prepared from these treated cells, suggesting that this inhibition of *in vivo* activity is due to the sequestration of Dol-P in the form of the full-length tetradecasaccharide.

Exogenous addition of Man-P-Dol to the retina of embryonic chicks or mouse lymphoma cell line stimulates the synthesis of GlcNAc-PP-Dol by allosteric regulation of GPT enzyme activity,[88-91] and the reciprocal relationship was also found in the chick retina.[92] Glc-P-Dol did not show the same stimulatory effect, and rather inhibited the stimulation by Man-P-Dol.[90] GPT-1 activity has also been shown, in retinas from embryonic chicks, to be inhibited both by product inhibition by GlcNAc-PP-Dol as well as feedback inhibition by $GlcNAc_2$-PP-Dol.[92]

A number of studies have been carried out to show that the early genes in the dolichol pathway, *ALG7*, *ALG1*, and *ALG2* are early growth-response genes.[93] *ALG7* expression is intimately linked to the cell cycle in *S. cerevisiae*, with its activity being regulated with respect to stages of cell proliferation. Due to different check points for *ALG7* in the G1 phase of the cell cycle, the dolichol pathway genes are either differentially or co-ordinately regulated.[93-98] Retinoic acid, a potent regulator of growth and differentiation, was found to enhance the activity of the Alg7 homologue in a mouse teratocarcinoma cell line.[99] A 3-fold activity increase was accompanied by similar increases in transcription and translation of the gene, as well as an elevation of

$Glc_3Man_9GlcNAc_2$-PP-Dol and protein glycosylation levels. Similarly, in a study examining the effect of 8-bromo-cAMP on glycosylation of a short peptide acceptor, a significant increase in glycosylation, accompanied by an increase in the pool of tetradecasaccharide donor, was observed upon pretreatment with the cyclic AMP derivative.[100]

The ER unfolded protein response (UPR) is a system that eukaryotic cells invoke in response to certain stress signals, particularly involving the accumulation of unfolded proteins in the ER, which leads to an increase in expression of ER proteins such as folding chaperones.[101] Using small molecules to induce UPR in human fibroblasts, it was found that the dolichol pathway is also regulated in response to this stress, stimulating the biosynthesis of $Glc_3Man_9GlcNAc_2$-PP-Dol.[102]

Congenital Disorders of Glycosylation (CDG)

Defects in various steps of *N*-linked glycoprotein biosynthesis in humans can cause severe medical conditions. There are two main types of Congenital Disorders of Glycosylation, CDGI and CDGII, which comprise a series of rare but serious diseases resulting from aberrant glycosylation. Members of the CDGI family are specifically characterized by defects in *N*-linked glycoprotein biosynthesis.[103-105] The molecular basis for disease stems from mutations in biosynthetic genes, which leads to loss or truncation of dolichylpyrophosphate-linked saccharides, and these incomplete intermediates are transferred to protein by OT at a significantly reduced rate.[106] The consequence in the patient is expression of proteins in which entire oligosaccharide moieties are missing

from natively glycosylated sequons. The clinical presentations are extremely broad, and can include a range of severities of mental retardation, liver dysfunction, and intestinal disorders.[103, 107] These are summarized in Table I-2, along with the eleven human genes that have been identified to date in which mutations have been linked to different forms of the disease. Some success in treating the symptoms of these disorders has been achieved through the administration of mannose.[108, 109]

Gene	Enzyme	Typical symptoms
PMM2	Phosphomannomutase II	Mental retardation (MR), hypotonia, esotropia, lipodystrophy, cerebellar hypoplasia, seizures
MPI	Phosphomannose isomerase	Hepatic fibrosis, protein-losing enteropathy (PLE), coagulopathy, hypoglycemia
ALG6	Dol-P-Glc: $Man_9GlcNAc_2$-PP-Dol glucosyltransferase	MR, hypotonia, epilepsy
ALG3	Dol-P-Man: $Man_5GlcNAc_2$-PP-Dol mannosyltransferase	Severe MR, optic nerve atrophy
DPM1	Dol-P-Man Synthase I GDP-Man: Dol-P-Mannosyltransferase	Severe MR, epilepsy, hypotonia, mildly dysmorphic, coagulopathy
MPDU1	Mannose-P-Dolichol utilization defect 1/Lec35	Short stature, icthyosis, MR, retinopathy
ALG12	Dol-P-Man: $Man_7GlcNAc_2$-PP-Dol mannosyltransferase	Hypotonia, MR, facial dysmorphism, microcephaly, frequent infections
ALG8	Dol-P-Glc: $Glc_1Man_9GlcNAc_2$-PP-Dol glucosyltransferase	Hepatomegaly, coagulopathy, PLE, renal failure
ALG2	Putative GDP-Man: $ManGlcNAc_2$-PP-Dol mannosyltransferase	Normal at birth, hepatomegaly, coagulopathy, MR, hypomyelination, intractable seizures
DPAGT1	UDP-GlcNAc: -dolichol phosphate N-acetylglucosamine-1 phosphate transferase	Severe MR, hypotonia, seizures, microcephaly
ALG1	GDP-Man: $GlcNAc_2$-PP-Dol mannosyltransferase	Severe MR, hypotonia, acquired microcephaly, intractable seizures, fever, coagulopathy, nephrotic syndrome
ALG9	Dol-P-Man: $Man_{6 and 8}GlcNAc_2$-PP-Dol mannosyltransferase	Severe microcephaly, hepatomegaly, hypotonia, seizures

Table I-2. Summary of CDGs characterized to date. In some cases, the nomenclature for human and yeast genes is identical. Table adapted from Freeze and Aebi.[106]

Glycosyltransferase families and structural motifs

The CAZy database includes classification of over 12,000 glycosytransferases (known or putative), divided into 81 families based on sequence similarities. The large number of families highlights the low level of sequence homology among glycosyltransferases. Alg2 and Alg11 both belong to the CAZy glycosyltransferase family 4, which, in addition to GDP-Man α-mannosyltransferase activity, also includes sucrose and sucrose phosphate synthases, α-glucosyltransferase, lipopolysaccharide N-acetylglucosaminyl transferase, diacylglycerol 3-glucosyltransferase, diglucosyl diacylglycerol synthase, digalactosyldiacylglycerol synthase, trehalose phosphorylase, phosphatidylinositol α-mannosyltransferase, and UDP-galactose α-galactosyltransferase activity.[110]

X-ray crystal structures of glycosyltransferases were not available until the first structure was reported in 1994 for bacteriophage T4-glucosyltransferase.[111] There are now structures for at least 23 different glycosyltransferases, spanning 17 distinct glycosyltransferase families.[112] In contrast to the low sequence homology, a high degree of structural homology exists among the known structures of glycosyltransferases, with only two major folds, GT-A and GT-B,[113] and one recently identified novel fold for a sialyltransferase from *C. jejuni* (Figure I-10).[114] Glycosyltransferase folds are largely comprised of α/β/α sandwiches, with GT-A folds consisting of a Rossman fold, in which the active site is located in a space between the central β-sheet and a smaller β-sheet, and GT-B folds maintain the catalytic site between two distinct Rossman folds separated by a linker region.[112] The Rossman fold is a common motif found in nucleotide-binding

proteins.[115] A common element of the GT-A fold family is a conserved Asp-Xaa-Asp motif, where Xaa can be any amino acid, which is involved in binding of the glycosylacceptor phosphate moiety *via* a divalent metal,[112] and is distinct from the EX_7E motif shared by Alg2, Alg11, and a number of other enzymes in the same Pfam protein family database.[60] As discussed earlier, Alg2 and Alg11, which share 22% homology to each other, also contain an extended signature sequence that is found in over 60 other glycosyltransferases. X-ray crystal structures have been solved for three of the glycosyltransferases from this group, all of them revealing a GT-B fold, and an involvement of these conserved residues in binding of the nucleotide sugar donor.[111, 116, 117] In Alg2 and Alg11, the signature sequence of each lies within the soluble cytosolic region. Threading analysis with sequences of glycosyltransferases for which structures have not been solved reveal that most fall within either the GT-A or GT-B fold family. For Alg2, Alg11, and other members of the same family, threading analysis predicts the GT-B fold. Another emerging commonality among several glycosyltransferases is ordered substrate binding, in which donor substrate binding induces a loop motion, which creates the binding site for the glycosyl acceptor.[118]
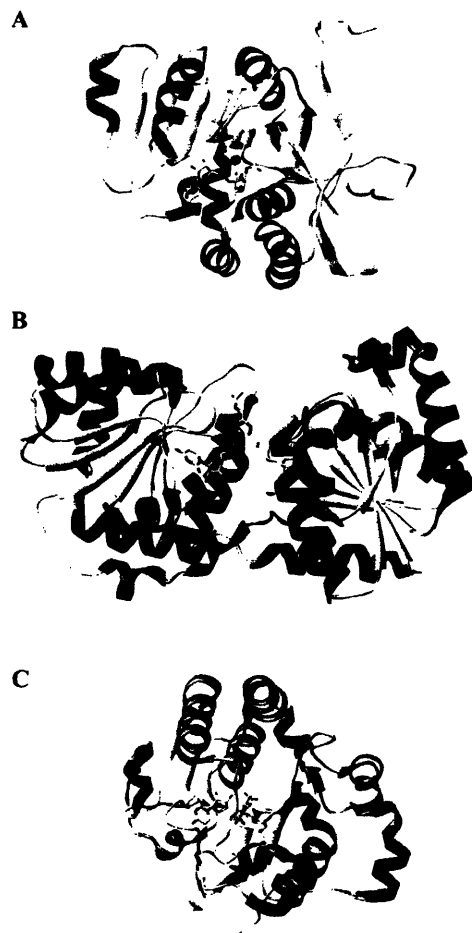
Figure I-10. Representative structures of the three different folds known for glycosyltransferases.[112] A. GT-A fold. B. GT-B fold. C. Sialyltransferase CstII

Mechanisms of Glycosyltransferases

Due to the difficulty in expressing and stabilizing glycosyltransferases, many of which are membrane-bound and have a significant amount of flexibility in their two-domain structures, X-ray crystallographic information is still quite scarce. For the same

reasons, the mechanistic details of this class of enzymes are unclear. The formation of glycosidic bonds can be described as retaining or inverting, depending on whether there is a change in configuration with respect to the anomeric center of the glycosyl donor. To understand these mechanisms, clues were taken from the better understood glycosylhydrolases.[119] The mechanism of inverting glycosyltransferases, based on structural and mechanistic studies, is believed to be similar to the inverting glycosylhydrolases, in which an active site base deprotonates the nucleophilic hydroxyl group of the glycosyl acceptor, facilitating an S_N2 displacement of the activating group of the glycosyl donor (Figure I-11A). In the case of retaining glycosyltransferases, it is not yet clear whether the mechanism involves a double displacement or rather an S_N1-type reaction. The double displacement mechanism requires a nucleophilic active site residue and a covalent intermediate (Figure I-11B), while the S_N1-type mechanism would involve stabilization by the enzyme of the oxocarbenium ion-like transition state of the glycosyl donor after release of the activating group and prior to bond formation with the glycosyl acceptor (Figure I-11C). Mechanistic studies employing secondary kinetic isotope effects, fluorinated derivatives of glycosyl donors, and inhibition by glycosyl donor analogs with positively charged nitrogen species to mimic the oxocarbenium ion transition state all provide evidence for the oxocarbenium ion-like transition state.[120] To date, X-ray crystal structures of retaining glycosyltransferases only represent six families, the majority of which have not been solved in the presence of an enzyme-bound glycosyl donor. Therefore, no candidate nucleophiles have been identified, and despite great efforts, evidence for a covalent intermediate to support the double displacement mechanism has not been found. However, a stable glycosyl-enzyme intermediate was

45

recently observed in a Gln189Glu mutant of the *Neisseria meningitidis* α-galactosyltransferase, LgtC.[121] Residue Gln189 in the wild-type enzyme is positioned properly to be a nucleophilic residue, but no modification was found for the wild-type enzyme. Thus, mutation of this residue to the more nucleophilic Glu and subsequent evidence for a covalent modification of this mutant suggested modification of Glu189 in the mutant. Surprisingly, however, Asp190, and not Glu189, was the modified residue in this mutant. Mutating Asp190 to an Ala residue in the wild type enzyme resulted in a 3,000-fold decrease in activity. Thus, it is still unclear whether the wild-type enzyme acts through a double-displacement mechanism.
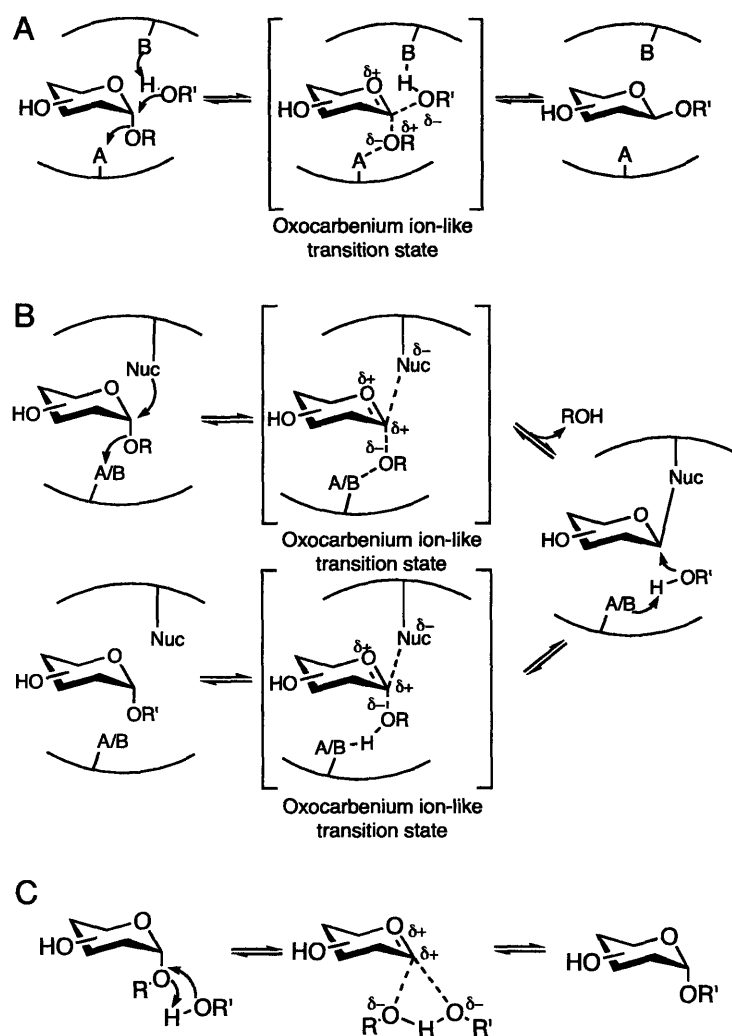
Figure I-11. Proposed mechanisms of glycosyltransferases. The inverting glycosyltransferases use general acid (A)/ general base (B) catalysis in a direct displacement S_N2 reaction to invert the anomeric configuration (A). The retaining glycosyltransferases are proposed to act through either a double displacement S_N2-like reaction, involving the formation of a covalent intermediate (B), or an S_N1-like reaction, without a covalent intermediate (C). These mechanisms were taken from Lairson, et al.[119]

Conclusion

Since the discovery of the Alg enzyme genes began over 20 years ago, many advances have been made in the functional assignment of the glycosyltransferase genes through genetics and bioinformatics, as well as in understanding the regulation and the role of the dolichol pathway in cellular processes. However, with the exception of Alg1, very few biochemical studies have been attempted with these enzymes. These advances have been primarily limited by the same challenges that are responsible for the small number of X-ray crystal structures and the incomplete mechanistic picture of this class of enzymes, which include membrane-association in many cases, and structural flexibility. Both of these features help to make glycosyltransferases notoriously unstable. In order to gain a deeper understanding of the biosynthesis as a whole, it is necessary to unambiguously define the function of each glycosyltransferase in the pathway. Only then can the concerted actions of these enzymes be studied in a controlled manner. Overcoming the inherent obstacles of this system will allow many insights into the structural and mechanistic challenges of bifunctional glycosyltransferases, and continue the improvement of the chemoenzymatic syntheses of valuable oligosaccharide structures.

References

1.   Evolution of glycan diversity. In *Essentials of Glycobiology*, Varki, A.; Cummings, R.; Esko, J.; Freeze, H. H.; Hart, G.; Marth, J., Eds. Cold Spring Harbor Laboratory Press: 1999; pp 31-39.

2.   Helenius, A.; Aebi, M., Intracellular functions of *N*-linked glycans. *Science* **2001**, 291, (5512), 2364-9.

3.   Varki, A., Biological roles of oligosaccharides: all the theories are correct. *Glycobiology* **1993**, 3, 97-130.

4.   Kornfeld, R.; Kornfeld, S., Assembly of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* **1985**, 54, 631-664.

5.   Imperiali, B., Protein glycosylation: the clash of the titans. *Acc. Chem. Res.* **1997**, 30, 452-459.

6.   Dempski, R. E.; Imperiali, B., Oligosaccharyl transferase: gatekeeper to the secretory pathway. *Curr. Opin. Chem. Biol.* **2002**, 6, 844-850.

7.   Kelleher, D. J.; Gilmore, R., An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **2005**, Advanced Access.

8.   Dean, N., Asparagine-linked glycosylation in the yeast Golgi. *Biochim. Biophys. Acta* **1999**, 1426, (2), 309-22.

9.   Alder, N. N.; Johnson, A. E., Cotranslational membrane protein biogenesis at the endoplasmic reticulum. *J. Biol. Chem.* **2004**, 279, (22), 22787-22790.

10.   Johnson, A. E.; M.A., v. W., The translocon: a dynamic gateway at the ER membrane. *Ann. Rev. Dev. Biol.* **1999**, 15, 799-842.

11.   Nilsson, I. M.; von Heijne, G., Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J Biol Chem* **1993**, 268, (8), 5798-801.

12.   Hebert, D. N.; Garman, S. C.; Molinari, M., The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends Cell Biol.* **2005**, 15, (7), 364-70.

13.   Helenius, A.; Aebi, M., Roles of *N*-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **2004**, 73, 1019-1049.

14.   Ito, Y.; Hagihara, S.; Matsuo, I.; Totani, K., Structural approaches to the study of oligosaccharides in glycoprotein quality control. *Curr. Opin. Struct. Biol.* **2005**, 15, (5), 481-9.

15.   Arai, M. A.; Matsuo, I.; Hagihara, S.; Totani, K.; Maruyama, J.; Kitamoto, K.; Ito, Y., Design and synthesis of oligosaccharides that interfere with glycoprotein quality-control systems. *Chembiochem* **2005**, 6, (12), 2281-9.

16.   Totani, K.; Ihara, Y.; Matsuo, I.; Koshino, H.; Ito, Y., Synthetic substrates for an endoplasmic reticulum protein-folding sensor, UDP-glucose: glycoprotein glucosyltransferase. *Angew. Chem., Int. Ed. Engl.* **2005**, 44, (48), 7950-4.

17.   Herscovics, A.; Orlean, P., Glycoprotein biosynthesis in yeast. *FASEB J.* **1993**, 7, 540-550.

18.   Burda, P.; Aebi, M., The dolichol pathway of *N*-linked glycosylation. *Biochim. Biophys. Acta* **1999**, 1426, 239-257.

19.   Hubbard, S. C.; Ivatt, R. J., Synthesis and processing of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* **1981**, 50, 555-583.

20.   Hirschberg, C. B.; Snider, M. D., Topography of glycosylation in the rough endoplasmic reticulum and golgi apparatus. *Ann. Rev. Biochem.* **1987,** 56, 63-87.

21.   Abeijon, C.; Hirschberg, C. B., Topography of Glycosylation Reactions in the Endoplasmic-Reticulum. *Trends Biochem. Sci.* **1992,** 17, (1), 32-36.

22.   Helenius, J.; Ng, D. T.; Marolda, C. L.; Walter, P.; Valvano, M. A.; Aebi, M., Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* **2002,** 415, (6870), 447-50.

23.   Schutzbach, J. S., The role of the lipid matrix in the biosynthesis of dolichyl-linked oligosaccharides. *Glycoconj. J.* **1997,** 14, (2), 175-82.

24.   Zhou, G. P.; Troy, F. A., 2nd, NMR study of the preferred membrane orientation of polyisoprenols (dolichol) and the impact of their complex with polyisoprenyl recognition sequence peptides on membrane structure. *Glycobiology* **2005,** 15, (4), 347-59.

25.   Zhou, G. P.; Troy, F. A., 2nd, NMR studies on how the binding complex of polyisoprenol recognition sequence peptides and polyisoprenols can modulate membrane structure. *Curr. Protein Pept. Sci.* **2005,** 6, (5), 399-411.

26.   Zhou, G. P.; Troy, F. A., 2nd, Characterization by NMR and molecular modeling of the binding of polyisoprenols and polyisoprenyl recognition sequence peptides: 3D structure of the complexes reveals sites of specific interactions. *Glycobiology* **2003,** 13, (2), 51-71.

27.   Burda, P.; Jakob, C. A.; Beinhauer, J.; Hegemann, J. H.; Aebi, M., Ordered assembly of the asymmetrically branched lipid-linked oligosaccharide in the endoplasmic reticulum is ensured by the substrate specificity of the individual glycosyltransferases. *Glycobiology* **1999,** 9, (6), 617-625.

28.   McCloskey, M. A.; Troy, F. A., Paramagnetic isoprenoid carrier lipids. 2. Dispersion and dynamics in lipid membranes. *Biochemistry* **1980,** 19, (10), 2061-6.

29.   Swiezewskaa, E.; Danikiewiczb, W., Polyisoprenoids: structure, biosynthesis and function. *Prog. Lipid Res.* **2005,** 44, (4), 235-58.

30.   Kharel, Y.; Koyama, T., Molecular analysis of cis-prenyl chain elongating enzymes. *Nat. Prod. Rep.* **2003,** 20, (1), 111-8.

31.   Pan, J. J.; Chiou, S. T.; Liang, P. H., Product distribution and pre-steady-state kinetic analysis of *Escherichia coli* undecaprenyl pyrophosphate synthase reaction. *Biochemistry* **2000,** 39, (35), 10936-42.

32.   Kharel, Y.; Takahashi, S.; Yamashita, S.; Koyama, T., Manipulation of prenyl chain length determination mechanism of cis-prenyltransferases. *FEBS J.* **2006,** 273, (3), 647-57.

33.   Szkopinska, A.; Swiezewska, E.; Rytka, J., Interplay between the cis-prenyltransferases and polyprenol reductase in the yeast *Saccharomyces cerevisiae*. *Biochimie* **2005**.

34.   Heller, L.; Orlean, P.; Adair, W. L., Jr., *Saccharomyces cerevisiae sec59* cells are deficient in dolichol kinase activity. *Proc. Natl. Acad. Sci. USA* **1992,** 89, (15), 7013-6.

35.   Fernandez, F.; Shridas, P.; Jiang, S.; Aebi, M.; Waechter, C. J., Expression and characterization of a human cDNA that complements the temperature-sensitive defect in dolichol kinase activity in the yeast *sec59-1* mutant: the enzymatic phosphorylation of dolichol and diacylglycerol are catalyzed by separate CTP-mediated kinase activities in *Saccharomyces cerevisiae*. *Glycobiology* **2002,** 12, (9), 555-62.

36.   Rosenwald, A. G.; Stoll, J.; Krag, S. S., Regulation of glycosylation: three enzymes compete for a common pool of dolichyl phosphate *in vivo*. *J. Biol. Chem.* **1990,** 265, (24), 14544-14553.

37.   Rush, J. S.; Cho, S. K.; Jiang, S.; Hofmann, S. L.; Waechter, C. J., Identification and characterization of a cDNA encoding a dolichyl pyrophosphate phosphatase located in the endoplasmic reticulum of mammalian cells. *J. Biol. Chem.* **2002,** 277, (47), 45226-34.

38.   Fernandez, F.; Rush, J. S.; Toke, D. A.; Han, G. S.; Quinn, J. E.; Carman, G. M.; Choi, J. Y.; Voelker, D. R.; Aebi, M.; Waechter, C. J., The CWH8 gene encodes a dolichyl pyrophosphate phosphatase with a luminally oriented active site in the endoplasmic reticulum of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **2001,** 276, (44), 41455-64.

39.   Sharma, C. B.; Lehle, L.; Tanner, W., Solubilization and characterization of the initial enzymes of the dolichol pathway from yeast. *Eur. J. Biochem.* **1982,** 126, 319-325.

40.   Rine, J.; Hansen, W.; Hardemann, E.; Davis, R. W., Targeted selection of recombinant clones through gene dosage effects. *Proc. Natl. Acad. Sci. USA* **1983,** 80, 6750-6754.

41.   Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.

42.   Dan, N.; Lehrman, M. A., Oligomerization of hamster UDP-GlcNAc:dolichol-P GlcNAc-1-P transferase, an enzyme with multiple transmembrane spans. *J. Biol. Chem.* **1997,** 272, (22), 14214-14219.

43.   Crouvoisier, M.; Mengin-Lecreulx, D.; van Heijenoort, J., UDP-N-acetylglucosamine:N-acetylmuramoyl-(pentapeptide) pyrophosphoryl undecaprenol *N*-acetylglucosamine transferase from *Escherichia coli*: overproduction, solubilization, and purification. *FEBS Lett.* **1999,** 449, (2-3), 289-92.

44.   Mengin-Lecreulx, D.; Texier, L.; Rousseau, M.; van Heijenoort, J., The MurG gene of Escherichia coli codes for the UDP-*N*-acetylglucosamine: *N*-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol *N*-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J. Bacteriol.* **1991,** 173, (15), 4625-36.

45.   Mengin-Lecreulx, D.; Texier, L.; van Heijenoort, J., Nucleotide sequence of the cell-envelope MurG gene of *Escherichia coli. Nucleic Acids Res.* **1990,** 18, (9), 2810.

46.   Chantret, I.; Dancourt, J.; Barbat, A.; Moore, S. E., Two proteins homologous to the N- and C-terminal domains of the bacterial glycosyltransferase MurG are required for the second step of dolichyl-linked oligosaccharide synthesis in *Saccharomyces cerevisiae. J. Biol. Chem.* **2005,** 280, (10), 9236-42.

47.   Kolkman, M. A.; van der Zeijst, B. A.; Nuijten, P. J., Functional analysis of glycosyltransferases encoded by the capsular polysaccharide biosynthesis locus of *Streptococcus pneumoniae* serotype 14. *J. Biol. Chem.* **1997,** 272, (31), 19502-8.

48.   Bickel, T.; Lehle, L.; Schwarz, M.; Aebi, M.; Jakob, C. A., Biosynthesis of lipid-linked oligosaccharides in Saccharomyces cerevisiae: Alg13p and Alg14p form a complex required for the formation of GlcNAc(2)-PP-dolichol. *J. Biol. Chem.* **2005,** 280, (41), 34500-6.

49.   Gao, X. D.; Tachikawa, H.; Sato, T.; Jigami, Y.; Dean, N., Alg14 Recruits Alg13 to the Cytoplasmic Face of the Endoplasmic Reticulum to Form a Novel Bipartite UDP-

N-acetylglucosamine Transferase Required for the Second Step of N-Linked Glycosylation. *J. Biol. Chem.* **2005**, 280, (43), 36254-62.

50.   Zhang, G.; O'Reilly, M. K.; Imperiali, B., *In Vitro* Reconstitution of Enzyme II Activity from the Yeast Dolichol Pathway-YGL047w and YBR070c Proteins Are Essential and Sufficient for Second Sugar Transfer. *Manuscript in Preparation* **2005**.

51.   Huffaker, T. C.; Robbins, P. W., Temperature-sensitive yeast mutants deficient in asparagine-linked glycosylation. *J. Biol. Chem.* **1982**, 257, (6), 3203-3210.

52.   Gao, X.-D.; Nishikawa, A.; Dean, N., Physical interactions between the Alg1, Alg2, and Alg11 mannosyltransferases of the endoplasmic reticulum. *Glycobiology* **2004**, 14, (6), 559-570.

53.   Huffaker, T. C.; Robbins, P. W., Yeast mutants deficient in protein glycosylation. *Proc. Natl. Acad. Sci. USA* **1983**, 80, 7466-7470.

54.   Jackson, B. J.; Kukuruzinska, M. A.; Robbins, P. W., Biosynthesis of asparagine-linked oligosaccharides in *Saccharomyces cerevisiae*: the *alg2* mutation. *Glycobiology* **1993**, 3, (4), 357-364.

55.   Yamakazi, H.; Shiraishi, N.; Takauchi, K.; Ohnishi, Y.; Horinouchi, S., Characterization of *ALG2* encoding a mannosyltransferase in the zygomycete fungus *Rhizomucor pusillus*. *Gene* **1998**, 221, 179-184.

56.   Takeuchi, K.; Yamazaki, H.; Shiraishi, N.; Ohnishi, Y.; Nishikawa, Y.; Horinouchi, S., Characterization of an *alg2* mutant of the zygomycete fungus *Rhizomucor pusillus*. *Glycobiology* **1999**, 9, (12), 1287-1293.

57.   Thiel, C.; Schwarz, M.; Peng, J.; Grzmil, M.; Hasilik, M.; Braulke, T.; Kohlschutter, A.; von Figura, K.; Lehle, L.; Korner, C., A new type of congenital disorder of glycosylation (CDG-Ii) provides new insights into the early steps of dolichol-linked oligosaccharide biosynthesis. *J. Biol. Chem.* **2003**, 278, (25), 22498-22505.

58.   Cipollo, J. F.; Trimble, R. B.; Chi, J. H.; Yan, Q.; Dean, N., The yeast *ALG11* gene specifies addition of the terminal α1,2-Man to the $Man_5GlcNAc_2$-PP-dolichol N-glycosylation intermediate formed on the cytosolic side of the endoplasmic reticulum. *J. Biol. Chem.* **2001**, 276, (24), 21828-40.

59.   Hu, Y.; Walker, S., Remarkable structural similarities between diverse glycosyltransferases. *Chem. Biol.* **2002**, 9, (12), 1287-96.

60.   Bateman, A.; Coin, L.; Durbin, R.; Finn, R. D.; Hollich, V.; Griffiths-Jones, S.; Khanna, A.; Marshall, M.; Moxon, S.; Sonnhammer, E. L. L.; Studholme, D. J.; Yeats, C.; Eddy, S. R., The Pfam protein families database. *Nucleic Acids Res.* **2004**, 32, D138-D141.

61.   Ballou, L.; Hitzeman, R. A.; Lewis, M. S.; Ballou, C. E., Vanadate-resisitant yeast mutants are defective in protein glycosylation. *Proc. Natl. Acad. Sci. USA* **1991**, 88, 3209-3212.

62.   Dean, N., Yeast glycosylation mutants are sensitive to aminoglycosides. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 1287-1291.

63.   Aebi, M.; Gassenhuber, J.; Domdey, H.; te Hesen, S., Cloning and characterization of the *ALG3* gene of *Saccharomyces cerevisiae*. *Glycobiology* **1996**, 9, (3), 7160-7165.

64.   Sharma, C. B.; Knauer, R.; Lehle, L., Biosynthesis of lipid-linked oligosaccharides in yeast: the *ALG3* gene encodes the Dol-P-Man:$Man_5GlcNAc_2$-PP-Dol mannosyltransferase. *Biol. Chem.* **2001**, 382, 321-328.

65.   Burda, P.; Te Heesen, S.; Brachat, A.; Wach, A.; Dusterhoft, A.; Aebi, M., Stepwise assembly of the lipid-linked oligosaccharide in the endoplasmic reticulum of *Saccharomyces cerevisiae*: identification of the *ALG9* gene encoding a putative mannosyl transferase. *Proc. Natl. Acad. Sci. USA* **1996,** 93, (14), 7160-5.

66.   Frank, C. G.; Aebi, M., *ALG9* mannosyltransferase is involved in two different steps of lipid-linked oligosaccharide biosynthesis. *Glycobiology* **2005,** 15, (11), 1156-63.

67.   Luzhetskyy, A.; Fedoryshyn, M.; Durr, C.; Taguchi, T.; Novikov, V.; Bechthold, A., Iteratively acting glycosyltransferases involved in the hexasaccharide biosynthesis of landomycin A. *Chem. Biol.* **2005,** 12, (7), 725-9.

68.   Glover, K. J.; Weerapana, E.; Imperiali, B., *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic *N*-linked glycosylation. *Proc. Natl. Acad. Sci. USA* **2005,** 102, (40), 14255-9.

69.   Lussier, M.; White, A. M.; Sheraton, J.; di Paolo, T.; Treadwell, J.; Southard, S. B.; Horenstein, C. I.; Chen-Weiner, J.; Ram, A. F.; Kapteyn, J. C.; Roemer, T. W.; Vo, D. H.; Bondoc, D. C.; Hall, J.; Zhong, W. W.; Sdicu, A. M.; Davies, J.; Klis, F. M.; Robbins, P. W.; Bussey, H., Large scale identification of genes involved in cell surface biosynthesis and architecture in *Saccharomyces cerevisiae*. *Genetics* **1997,** 147, (2), 435-50.

70.   Cipollo, J. F.; Trimble, R. B., The *Saccharomyces cerevisiae* alg12Δ mutant reveals a role for the middle-arm α1,2Man- and upper-arm α1,2Manα1,6Man- residues of Glc₃Man₉GlcNAc₂-PP-Dol in regulating glycoprotein glycan processing in the endoplasmic reticulum and Golgi apparatus. *Glycobiology* **2002,** 12, (11), 749-62.

71.   Reiss, G.; te Heesen, S.; Zimmerman, J.; Robbins, P. W.; Aebi, M., Isolation of the *ALG6* locus of *Saccharomyces cerevisiae* required for glucosylation in the *N*-linked glycosylation pathway. *Glycobiology* **1996,** 6, (5), 493-8.

72.   Stagljar, I.; te Heesen, S.; Aebi, M., New phenotype of mutations deficient in glucosylation of the lipid-linked oligosaccharide: cloning of the *ALG8* locus. *Proc. Natl. Acad. Sci. USA* **1994,** 91, (13), 5977-81.

73.   Burda, P.; Aebi, M., The *ALG10* locus of *Saccharomyces cerevisiae* encodes the alpha-1,2 glucosyltransferase of the endoplasmic reticulum: the terminal glucose of the lipid-linked oligosaccharide is required for efficient *N*-linked glycosylation. *Glycobiology* **1998,** 8, (5), 455-62.

74.   Verostek, M. F.; Atkinson, P. H.; Trimble, R. B., Glycoprotein biosynthesis in the *alg3 Saccharomyces cerevisiae* mutant. *J. Biol. Chem.* **1993,** 268, (16), 12095-12103.

75.   Jackson, B. J.; Warren, C. D.; Bugge, B.; Robbins, P. W., Synthesis of lipid-linked oligosaccharides in *Saccharomyces cerevisiae*: Man₂GlcNAc₂ and Man₁GlcNAc₂ are transferred from dolichol to protein in vivo. *Arch. Biochem. Biophys.* **1989,** 272, (1), 203-209.

76.   Karaoglu, D.; Kelleher, D. J.; Gilmore, R., Allosteric regulation provides a molecular mechanism for preferential utilization of the fully assembled dolichol-linked oligosaccharide by the yeast oligosaccharyltransferase. *Biochemistry* **2001,** 40, (40), 12193-12206.

77.   Kelleher, D. J.; Karaoglu, D.; Mandon, E. C.; Gilmore, R., Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. *Mol Cell* **2003,** 12, (1), 101-11.

78.   Sharma, C. B.; Lehle, L.; Tanner, W., N-Glycosylation of yeast proteins. Characterization of the solubilized oligosaccharyl transferase. *Eur. J. Biochem.* **1981,** 116, (1), 101-8.

79.   Tai, V. W.; Imperiali, B., Substrate specificity of the glycosyl donor for oligosaccharyl transferase. *J. Org. Chem.* **2001,** 66, (19), 6217-28.

80.   Imperiali, B., Protein glycosylation: The clash of the titans. *Acc. Chem. Res.* **1997,** 30, (11), 452-459.

81.   Ufret. M.; Imperiali, B., Probing the extended binding determinants of oligosaccharyl transferase with synthetic inhibitors of asparagine-linked glycosylation. *Bioorg. Med. Chem. Lett.* **2000,** 10, (3), 281-4.

82.   Weerapana, E.; Imperiali, B., Peptides to peptidomimetics: towards the design and synthesis of bioavailable inhibitors of oligosaccharyl transferase. *Org. Biomol. Chem.* **2003,** 1, (1), 93-9.

83.   Silberstein, S.; Gilmore, R., Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J.* **1996,** 10, (8), 849-858.

84.   Szymanski, C. M.; Logan, S. M.; Linton, D.; Wren, B. W., Campylobacter--a tale of two protein glycosylation systems. *Trends Microbiol.* **2003,** 11, (5), 233-8.

85.   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.; Aebi, M., N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli. Science* **2002,** 298, (5599), 1790-3.

86.   Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B., Chemoenzymatic synthesis of glycopeptides with PgIB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni. Chem. Biol.* **2005,** 12, (12), 1311-5.

87.   Gao, N.; Lehrman, M. A., Coupling of the dolichol-PP-oligosaccharide pathway to translation by perturbation-sensitive regulation of the initiating enzyme, GlcNAc-1-P transferase. *J. Biol. Chem.* **2002,** 277, (42), 39425-39435.

88.   Kean, E. L., Stimulation by GDP-mannose of the biosynthesis of N-acetylglucosaminylpyrophosphoryl polyprenols by the retina. *J. Biol. Chem.* **1980,** 255, (5), 1921-7.

89.   Kean, E. L., Site of stimulation by mannosyl-P-dolichol of GlcNAc-lipid formation by microsomes of embryonic chick retina. *Glycoconj. J.* **1996,** 13, (4), 675-80.

90.   Kean, E. L., Stimulation by dolichol phosphate-mannose and phospholipids of the biosynthesis of N-acetylglucosaminylpyrophosphoryl dolichol. *J. Biol. Chem.* **1985,** 260, (23), 12561-71.

91.   Kean, E. L., Studies on the activation by dolichol-P-mannose of the biosynthesis of GlcNAc-P-P-dolichol and the topography of the GlcNAc-transferases concerned with the synthesis of GlcNAc-P-P-dolichol and $(GlcNAc)_2$-P-P-dolichol: a review. *Biochem. Cell Biol.* **1992,** 70, (6), 413-21.

92.   Kean, E. L.; Wei, Z.; Anderson, V. E.; Zhang, N.; Sayre, L. M., Regulation of the biosynthesis of N-acetylglucosaminylpyrophosphoryldolichol, feedback and product inhibition. *J. Biol. Chem.* **1999,** 274, (48), 34072-82.

93.   Kukuruzinska, M. A.; Lennon, K., Growth-related coordinate regulation of the early N-glycosylation genes in yeast. *Glycobiology* **1994,** 4, (4), 437-43.

94.   Kukuruzinska, M. A.; Lennon, K., Diminished activity of the first *N*-glycosylation enzyme, dolichol-P-dependent *N*-acetylglucosamine-1-P transferase (GPT), gives rise to mutant phenotypes in yeast. *Biochim. Biophys. Acta* **1995**, 1247, (1), 51-9.

95.   Lennon, K.; Bird, A.; Chen, Y. F.; Pretel, R.; Kukuruzinska, M. A., The dual role of mRNA half-lives in the expression of the yeast *ALG7* gene. *Mol. Cell. Biochem.* **1997**, 169, (1-2), 95-106.

96.   Lennon, K.; Bird, A.; Kukuruzinska, M. A., Deregulation of the first *N*-glycosylation gene, *ALG7*, perturbs the expression of G1 cyclins and cell cycle arrest in *Saccharomyces cerevisiae. Biochem. Biophys. Res. Commun.* **1997**, 237, (3), 562-5.

97.   Lennon, K.; Pretel, R.; Kesselheim, J.; te Heesen, S.; Kukuruzinska, M. A., Proliferation-dependent differential regulation of the dolichol pathway genes in *Saccharomyces cerevisiae. Glycobiology* **1995**, 5, (6), 633-42.

98.   Pretel, R.; Lennon, K.; Bird, A.; Kukuruzinska, M. A., Expression of the first *N*-glycosylation gene in the dolichol pathway, *ALG7*, is regulated at two major control points in the G1 phase of the Saccharomyces cerevisiae cell cycle. *Exp. Cell. Res.* **1995**, 219, (2), 477-86.

99.   Meissner, J. D.; Naumann, A.; Mueller, W. H.; Scheibe, R. J., Regulation of UDP-*N*-acetylglucosamine:dolichyl phosphate *N*-acetylglucosamine-1-phosphate transferase by retinoic acid in P19 cells. *Biochem. J.* **1999**, 338, 561-568.

100.   Konrad, M.; Merz, W. E., Regulation of N-glycosylation. Long term effect of cyclic AMP mediates enhanced synthesis of the dolichol pyrophosphate core oligosaccharide. *J. Biol. Chem.* **1994**, 269, (12), 8659-66.

101.   Chapman, R.; Sidrauski, C.; Walter, P., Intracellular signaling from the endoplasmic reticulum to the nucleus. *Annu. Rev. Cell Dev. Biol.* **1998**, 14, 459-85.

102.   Doerrler, W. T.; Lehrman, M. A., Regulation of the dolichol pathway in human fibroblasts by the endoplasmic reticulum unfolded protein response. *Proc. Natl. Acad. Sci. USA* **1999**, 96, (23), 13050-5.

103.   Freeze, H. H., Update and perspectives on congenital disorders of glycosylation. *Glycobiology* **2001**, 11, (12), 129R-143R.

104.   Aebi, M.; Hennet, T., Congenital disorders of glycosylation: genetic model systems lead the way. *Trends Cell Biol.* **2001**, 11, (3), 136-41.

105.   Jaeken, J.; Carchon, H., Congenital disorders of glycosylation: a booming chapter of pediatrics. *Curr. Opin. Pediatr.* **2004**, 16, (4), 434-9.

106.   Freeze, H. H.; Aebi, M., Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr. Opin. Struct. Biol.* **2005**, 15, (5), 490-8.

107.   Freeze, H. H.; Westphal, V., Balancing *N*-linked glycosylation to avoid disease. *Biochimie* **2001**, 83, 791-799.

108.   Rush, J. S.; Panneerselvam, K.; Waechter, C. J.; Freeze, H. H., Mannose supplementation corrects GDP-mannose deficiency in cultured fibroblasts from some patients with Congenital Disorders of Glycosylation (CDG). *Glycobiology* **2000**, 10, (8), 829-35.

109.   Panneerselvam, K.; Freeze, H. H., Mannose corrects altered N-glycosylation in carbohydrate-deficient glycoprotein syndrome fibroblasts. *J. Clin. Invest.* **1996**, 97, (6), 1478-87.

110.   Coutinho, P. M.; Henrissat, B., Carbohydrate-Active Enzymes server. In 1999.

111.	Vrielink, A.; Ruger, W.; Driessen, H. P. C.; Freemont, P. S., Crystal structure of the DNA modifying enzyme betaglucosyltransferase in the presence and absence of substrate uridine diphosphoglucose. *EMBO J.* **1994**, 13, 3413-3422.

112.	Breton, C.; Snajdrova, L.; Jeanneau, C.; Koca, J.; Imberty, A., Structures and mechanisms of glycosyltransferases. *Glycobiology* **2006**, 16, (2), 29R-37R.

113.	Unligil, U. M.; Rini, J. M., Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **2000**, 10, (5), 510-7.

114.	Chiu, C. P.; Watts, A. G.; Lairson, L. L.; Gilbert, M.; Lim, D.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C., Structural analysis of the sialyltransferase CstII from *Campylobacter jejuni* in complex with a substrate analog. *Nat. Struct. Mol. Biol.* **2004**, 11, (2), 163-70.

115.	Lesk, A. M., NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* **1995**, 5, (6), 775-83.

116.	Ha, S.; Walker, D.; Shi, Y.; Walker, S., The 1.9Å crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. *Protein Sci.* **2000**, 9, (1045-1052).

117.	Mulichak, A. M.; Losey, H. C.; Walsh, C. T.; Garavito, R. M., Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure* **2001**, 9, 547-557.

118.	Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E., Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem. Sci.* **2005**, 30, (1), 53-62.

119.	Lairson, L. L.; Withers, S. G., Mechanistic analogies amongst carbohydrate modifying enzymes. *Chem. Commun. (Camb)* **2004**, (20), 2243-8.

120.	Unligil, U. M.; Rini, J. M., Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **2000**, 10, 510-517.

121.	Lairson, L. L.; Chiu, C. P.; Ly, H. D.; He, S.; Wakarchuk, W. W.; Strynadka, N. C.; Withers, S. G., Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. *J. Biol. Chem.* **2004**, 279, (27), 28339-44.

Chapter II: Substrate specificity of the UDP-GlcNAc: GlcNAc-PP-Dolichol, *N*-

acetylglucosamine transferase (Enzyme II)

Introduction

In the second step of the dolichol pathway, an N-acetylglucosamine unit (GlcNAc) is transferred from a UDP-GlcNAc glycosyl donor to a dolichylpyrophosphate-linked GlcNAc acceptor (GlcNAc-PP-Dol) by the UDP-GlcNAc: GlcNAc-PP-dolichol, N-acetylglucosamine transferase that has been traditionally known as Enzyme II (Scheme II-1).[1] Recently this activity has been assigned to the heterooligomeric complex, Alg13/Alg14.[2] Analogous to the other glycosyltransferases involved in the assembly of $Man_5GlcNAc_2$-PP-Dol, this membrane-bound enzyme is localized in the endoplasmic reticulum (ER) membrane with the active site facing the cytosol, where the reaction occurs.[3]
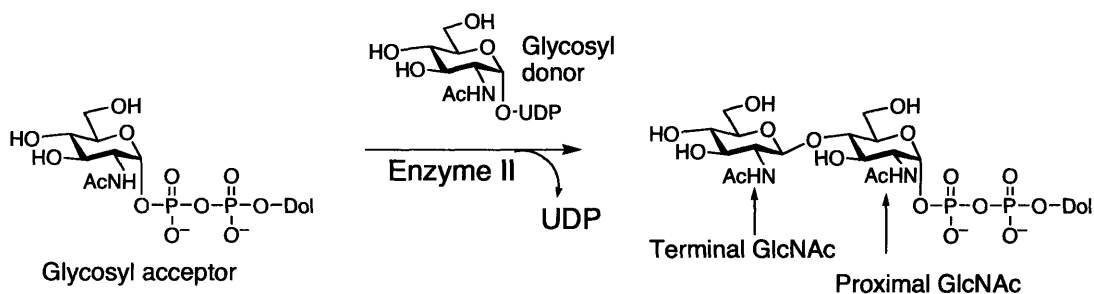
Scheme II-1. Glycosyltransfer reaction catalyzed by Enzyme II.

The product of the Enzyme II reaction, $GlcNAc_2$-PP-Dol, is a competent substrate for oligosaccharyl transferase (OT),[4] and is frequently employed for *in vitro* assays due to the difficulty in isolation of the complete tetradecasaccharide donor from native sources. For this purpose, radiolabeled (^{3}H)GlcNAc-GlcNAc-PP-Dol, where one of the C-6 protons of the terminal GlcNAc is replaced by a tritium, was previously prepared by an

efficient chemoenzymatic synthesis using Enzyme II from a partially purified preparation

of porcine liver microsomes. In the interest of evaluating OT with

dolichylpyrophosphate-linked disaccharides bearing non-natural substitutions at the C-2

position of the proximal GlcNAc,[5] Enzyme II was chosen as a tool to prepare these

analogues chemoenzymatically. This chapter describes how the attempts to prepare non-

natural OT substrate analogues using Enzyme II have revealed the tolerance of Enzyme II

with respect to substrate modification, and thus the limits of the utility of this enzyme for

such chemoenzymatic purposes. Thus, modifications to the C-2 position of the proximal

monosaccharide can provide information about both Enzyme II and, as described

elsewhere,[5] OT (Figure II-1).
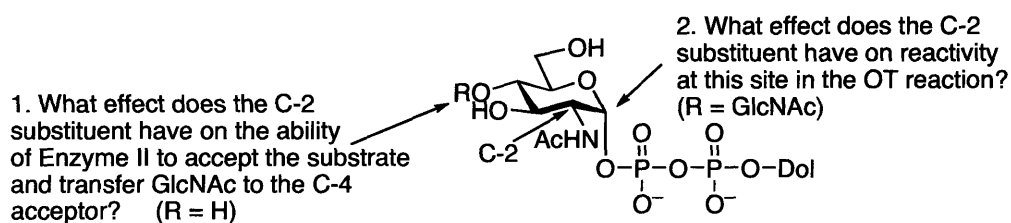
Figure II-1. Modification of the C-2 of GlcNAc-PP-Dol can be used to probe two

distinct enzymes.

The observation that UDP-Glc inhibits the Enzyme II reaction suggests that the C-

2 acetamido group on the glycosyl donor, UDP-GlcNAc, is a key determinant.[6] In

constrast, the importance of the acetamido group on the glycosyl acceptor, GlcNAc-PP-

Dol, had not yet been evaluated. There is a great deal of interest, and there have been

many advances, in using chemoenzymatic methods to prepare synthetically challenging

carbohydrates using glycosyltransferases that exhibit relaxed substrate specificity.[7] Thus,

in addition to studying substrate specificity as a means to unravel mechanistic details, understanding the tolerance of glycosyltransferases for various substrates is crucial to expanding the repertoire of tools for chemoenzymatic carbohydrate synthesis.

There have been a number of reports on the solubilization and partial purification of Enzyme II from natural sources such as mung bean seedling,[6] yeast,[8-11] and rat[1] and pig[12] liver, however, until the very recent discovery of the Alg13/Alg14 complex,[2] the identity of the gene(s) encoding Enzyme II had remained elusive. Nevertheless, characterization of this enzyme has been possible through the use of microsomal preparations, which exhibit no detectable background reactivity from unrelated native enzymes with dolichylpyrophosphate-linked pathway intermediates.[13-15] Partially purified Enzyme II from solubilized yeast microsomes was used to determine an apparent K_m of 40 μM for the GlcNAc$_2$-PP-Dol acceptor at a fixed concentration of 4.9 μM UDP-GlcNAc.

Results and Discussion

Non-natural dolichylpyrophosphate-linked monosaccharides used in this study

A variety of substrate analogues with substituents replacing the C-2 acetamido group of dolichylpyrophosphate-linked GlcNAc (**1a**), including the fluoro, O-ethyl, trifluoroacetamido, and amino functionalities (**1b-e**, respectively) (Scheme II-2) were investigated. Compounds **1a-d** were synthesized and compounds **1a-c** evaluated as substrates by Dr. Vincent W.-F. Tai. Because the synthesis of **1e** directly followed from

the synthesis of **1d**, both the synthesis and biological evaluation of **1d** and **1e** will be described in detail herein.
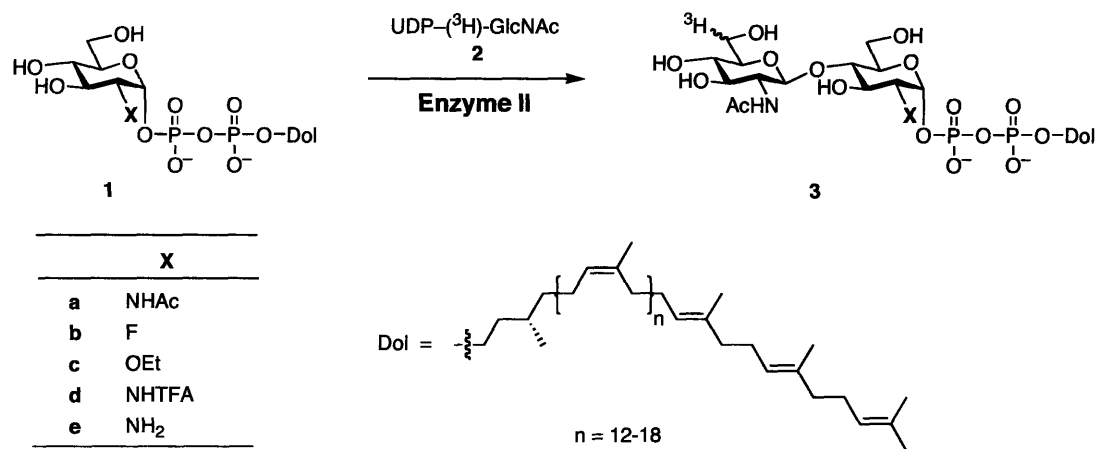
Scheme II-2. Non-natural monosaccharide acceptors used in the Enzyme II reaction.

Synthesis of GlcNHTFA-PP-Dol (1d) and GlcNH₂-PP-Dol (1e)

The general scheme for the synthesis of GlcNHTFA-PP-Dol (**1d**) is outlined in Scheme II-3. GlcNH₂-PP-Dol (**1e**) was synthesized using **1d** as a precursor. 3,4,6-Tri-*O*-acetyl-2-deoxy-2-trifluoroacetamido-α-D-glucopyranosyl phosphate (**4d**) has recently been prepared by two groups.[16] In the method used by the Tanner group, the α-phosphate, **4d**, was prepared from the corresponding reducing sugar using the phosphorochloridate method in 52% yield.[16] Alternatively, Busca and Martin showed that both the α- and β-phosphates could be obtained in high yield from ring opening of oxazoline, **6**, with dibenzyl phosphate under different conditions.[17]

61
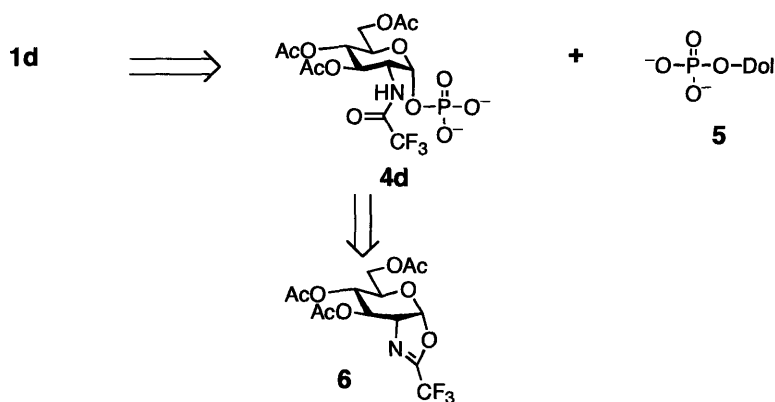
Scheme II-3. Retrosynthetic analysis of GlcNHTFA-PP-Dol (**1d**).

Following the Martin protocol, ring opening of the oxazoline **6** [17, 18] with dibenzyl phosphate in refluxing 1,2-dichloroethane afforded the thermodynamic product **8** in 65% yield (Scheme II-4).[17] The α-anomeric stereochemistry of **8** was confirmed based on the coupling constant between C-1 and C-2 protons in the ^{1}H NMR spectrum ($J_{1,2}$ = 3.3 Hz). Catalytic hydrogenolysis of **8** with 10% palladium on charcoal gave the α-phosphate, **4d**, in quantitative yield. The resulting phosphate was then condensed with Dol–P (**5**) to afford the protected GlcNHTFA-PP-Dol (**9**) in 63% yield. Chemoselective deprotection of the *O*-acetates in **9** with guanidine/guanidinium nitrate solution[19] yielded GlcNHTFA-PP-Dol (**1d**) in 47% yield after silica gel chromatography. Alternatively, treatment of **9** with lithium hydroxide (10 equivalents) provided the amine derivative GlcNH$_2$-PP-Dol (**1e**) in 63% yield.
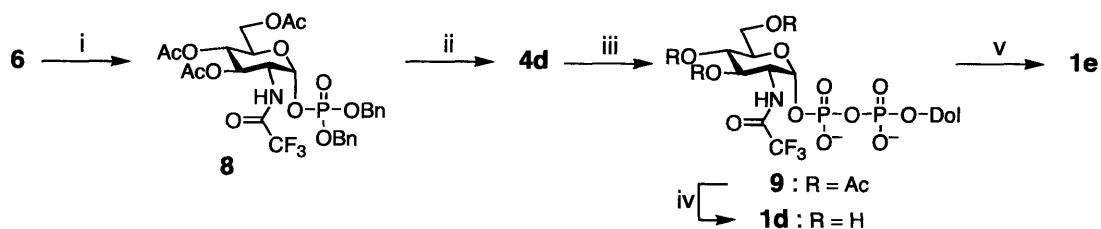
Scheme II-4. Synthesis of GlcNHTFA-PP-Dol (**1d**) and GlcNH$_2$-PP-Dol (**1e**)[a].

[a]Reagents and conditions: (i) dibenzyl phosphate, 1,2-dichloroethane, Δ, 65% (ii) H$_2$, 10% Pd/C, MeOH, 100% (iii) (1) CDI, DMF; (2) Dol–P (**5**), 63% (iv) guanidine/ guanidinium nitrate, CH$_2$Cl$_2$/MeOH, 47%; (v) LiOH, CH$_2$Cl$_2$/MeOH, 63%.

Biological Evaluation

Non-natural dolichylpyrophosphate-linked monosaccharides (**1b-e**) were evaluated as glycosyl acceptors for Enzyme II using a partially-purified pig liver microsome preparation and radiolabeled UDP-(^{3}H)GlcNAc as the glycosyl donor. The radioactive assay used to evaluate this glycosyltransferase, as well as other glycosyltransferases on the pathway, is shown in Figure II-2.
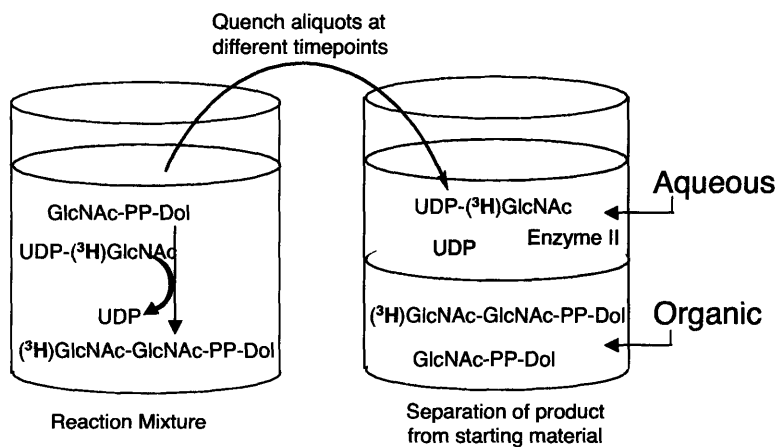
Figure II-2. Radioactive assay used for the dolichol pathway enzymes.

Product formation was quantified by measuring the amount of tritium label transferred to an organic extract [unreacted UDP-(^{3}H)GlcNAc partitions into the aqueous layer] after successive aqueous-organic extractions.[20] Percent conversion was calculated by dividing the disintegrations per minute (dpm) of tritium in the organic layer by the total number of dpms in the vial, and therefore represents the percent of radiolabeled monosaccharide transferred from nucleotide-activated sugar (0.2 µM in this assay) to the dolichylpyrophosphate-linked substrate. Assuming similar kinetic parameters for the non-natural disaccharide derivatives, all assays were performed with the acceptor substrate concentration above K_m. However, due to the fact that the dolichylpyrophosphate-linked substrate is in an unknown physical state in the detergent-containing assay, considering that the critical micelle concentration (CMC) is unknown, this assay is not used to derive kinetic constants, but rather to detect enzyme activity and to make comparisons of activity under varying sets of reaction conditions. In this case, the assay provided information to enable a qualitative comparison among the series of natural and non-natural substrates in order to gain insight into specificity.

Non-natural analogues 2DFGlc-PP-Dol (**1b**) and 2OEtGlc-PP-Dol (**1c**) did not display any significant transferase activity.[21] In contrast, the GlcNHTFA-PP-Dol (**1d**) and GlcNH$_2$-PP-Dol (**1e**) derivatives were shown to be relatively poor and moderate substrates for Enzyme II (Figure II-3), respectively. Even at higher concentrations of **1d** (1.5 mM) and **1e** (600 μM), only 13% and 41% conversion were observed after 14 hours, respectively. These yields are low compared to the 70% conversion achieved after only 30 minutes using 150 µM of the natural substrate and keeping the concentration of the

glycosyl donor constant. They do, however, indicate the feasibility of using these substrates for the chemoenzymatic preparation of non-natural disaccharide analogues.
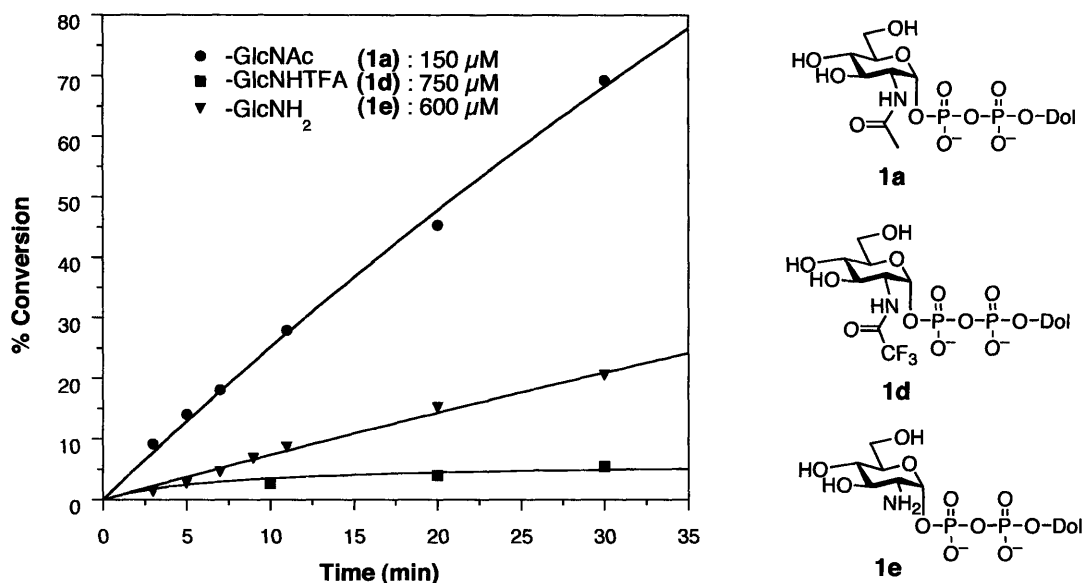
Figure II-3. Activity assays of relative rates of dolichylpyrophosphate-linked GlcNAc (**1a**), GlcNHTFA (**1d**), and GlcNH$_2$ (**1e**) analogues.

Derivatives **1b-d** were further shown to be inhibitors when competed with the natural substrate **1a**. As shown in Table II-1, the percent conversion of GlcNAc-PP-Dol (**1a**) was reduced by 27-32% when equivalent amounts of the non-natural derivatives **1b-1d** were added.

compounds[a]	Relative % conversion[b]
1a	1
1a + **1b**[c]	0.68
1a + **1c**	0.73
1a + **1d**	0.68

[a] 150 μM of each compound and 5 μL of pig liver microsomes were used.

[b] % conversion after 10 minutes of reaction.

[c] 10 μL of pig liver microsomes was used in this case.

Table II-1. Inhibitory activity of non-natural monosaccharide acceptors. Relative % conversion of **1a** when competed with **1b-d**.

Conclusion

In summary, Enzyme II demonstrates high specificity for the glycosyl acceptor. It is noted that even a relatively minor modification of the GlcNAc C-2 position in the GlcNAc-PP-Dol substrate diminishes the activity of this enzyme. Indeed, the electronic effects of the substituents at the C-2 position should induce minimal effects on the nucleophilicity of the C-4 hydroxyl group. Therefore, the acetamido group is postulated to act as a critical binding determinant for this transferase reaction through highly specific hydrogen bonding interactions between the N–H and C=O groups and the enzyme. Neither the 2-deoxy-2-fluoro analogue (**1b**) nor the 2-O-ethyl analogue (**1c**) function as substrates to the limits of detection of this assay (background is approximately 1% conversion),[21] perhaps due to the lack of the proposed essential

acetamido group. The strong electronic effects of the trifluoromethyl group of the trifluoroacetamide analogue (**1d**) may be a cause of detrimental effects on binding. The fact that the amine analogue $GlcNH_2$-PP-Dol (**1e**) is a substrate suggests that the minimal recognition element for this transferase reaction is a C-2 amine substituent.

The substrate specificity observed for Enzyme II is different from that of a 'core-2' GlcNAc transferase reported by Tanner.[16] In this case, UDP–GlcNHTFA was found to be a substrate for 'core-2' GlcNAc transferase but neither the 2-amino (UDP–$GlcNH_2$) nor the 2-hydroxyl (UDP–Glc) derivatives exhibited substrate behavior.

Acknowledgments

Experimental Procedures

General Methods for Synthesis.

Melting points were uncorrected. Optical rotations were measured with a JASCO DIP-1000 automatic digital polarimeter operating at 589 nm at 25 °C and are reported in degrees. Concentration (c) is indicated as units of 10 mg/mL. IR spectra were recorded on a Perkin-Elmer 1600 FT-IR spectrometer. NMR spectra were measured on a Varian

500, Bruker AM-500 (500.15 MHz for ^{1}H, 125 MHz for ^{13}C), or General Electric QE300

(75 MHz for ^{13}C). All chemical shifts were recorded in ppm downfield from

tetramethylsilane on the d scale. ^{31}P NMR chemical shifts are reported in ppm relative to

85% phosphoric acid external standard. Multiplicities are reported in the following

abbreviations: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), dd (doublet of

doublets), *etc*. Mass spectra were obtained either at the Mass Spectrometry Laboratory,

operated by the College of Chemistry, California Institute of Technology, Pasadena,

CA91125, at the UCR Mass Spectrometry Facility, Department of Chemistry, University

of California, Riverside, CA92521, or with a Mariner electrospray mass spectrometer

from Applied Biosystems using the negative ion mode. All reactions were monitored by

analytical thin-layer chromatography (TLC) on glass precoated with silica gel $60F_{254}$ (E.

Merck) and compounds were visualized with 20% w/v dodecamolybdophosphoric acid in

ethanol and subsequent heating. Phosphorous compounds were visualized with

Molybdenum blue spray reagent.[22] All columns were packed wet with E. Merck silica

gel 60 (230-400 mesh) as the stationary phase and eluted by flash chromatography. All

solvents were reagent grade. Methylene chloride and triethylamine were distilled from

calcium hydride. Tetrahydrofuran and toluene were distilled from sodium/benzophenone

ketyl. Other reagents were purchased from commercial suppliers and used without

further purification.

Mono(triethylammonium) salt of 3,4,6-tri-*o*-acetyl-2-deoxy-2-trifluoroacetamido-α-D-glucopyranosyl phosphate (4d).[17]
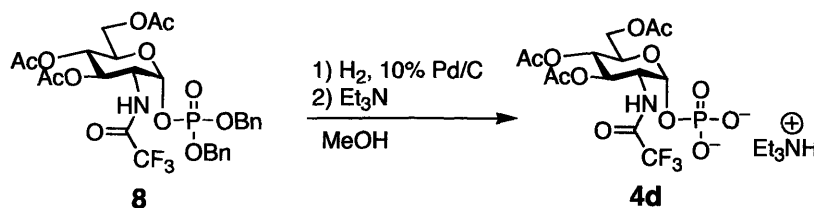
A solution of the protected phosphate **8**[17] (145 mg, 0.220 mmol) and 10% Pd/C

(20 mg) was stirred under H_2 in MeOH (8 mL) until no starting material remained as

shown by TLC. The catalyst was filtered through a pad of Celite and washed with EtOH

(15 mL). Et_3N (1 mL) was added followed by the evaporation of solvent *in vacuo* to give

the mono-(triethylammonium) salt of the phosphate (136 mg, quantitative yield) as a

foam; IR (neat) cm^{-1} 3411, 2995, 1748, 1560, 1458, 1369, 1224, 1189, 1157, 1040, 962,

921, 840, 718; ^{1}H NMR (CDCl$_3$, 500 MHz) δ 1.08 [t, 9H, *J* = 7.3 Hz, N(CH$_2$C*H*$_3$)], 1.75

(s, 3H, Ac), 1.79 (s, 3H, Ac), 1.85 (s, 3H, Ac), 2.87 [q, 6H, *J* = 7.3 Hz, N(C*H*$_2$CH$_3$)], 3.90

(dd, 1H, *J* = 2.0, 12.4 Hz, H-6a), 4.03 (dd, 1H, *J* = 3.4, 12.4 Hz, H-6b), 4.07 (dt, 1H, *J* =

2.9, 10.2 Hz, H-5), 4.14 (dt, 1H, *J* = 2.6, 10.5 Hz, H-2), 4.92 (t, 1H, *J* = 9.6 Hz, H-3/H-4),

5.19 (t, 1H, *J* = 9.6 Hz, H-4/H-3), 5.32 (dd, 1H, *J* = 3.3, 6.8 Hz, H-1); ^{13}C NMR (CDCl$_3$,

125 MHz) δ 7.9, 19.8, 20.0, 20.1, 45.7, 52.3 (d, *J* = 7.5 Hz), 61.3, 67.9, 68.0 (2x), 70.6,

92.8, 157.9 (q, *J* = 37 Hz), 169.7, 170.4, 171.0; ^{31}P NMR (CDCl$_3$, 162 MHz) d –0.9; [a]$_D$

+33.9 ° (*c* = 0.9, CHCl$_3$); MS (–ve FAB): 480 (M⁻); HRMS: Calcd for C$_{14}$H$_{18}$NO$_{12}$F$_3$P:

m/z = 480.051874; Found: 480.051500.

P¹-Dolichyl P²-[3,4,6-tri-*O*-acetyl-2-deoxy-2-trifluoroacetamido-α-D-glucopyranosyl]

diphosphate (9).
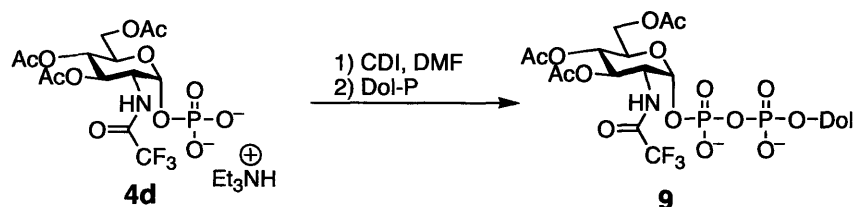
To a solution of the **4d** (0.04 mmol) in DMF (1 mL) at room temperature was

added 1,1'-carbonyldiimidazole (CDI, 0.2 mmol) in one portion. After 2 hours, MeOH

(0.36 mmol) was added (to quench excess CDI) to the mixture and stirred for 30 min.

Dolichylmonophosphate (**5**)[23] (tri-*n*-butyl ammonium form, 0.028 mmol) in CH_2Cl_2 (1

mL) was then added and the mixture was allowed to stir for 3 days at room temperature.

The crude mixture was concentrated followed by chromatography with DE-52 [acetate

form, eluted with increasing concentration of NH_4OAc in $CHCl_3$, MeOH (2:1 v/v)]. The

products were concentrated and further purified by silica gel chromatography.

¹⁹F NMR ($CDCl_3$, 470 MHz) d –76.4; ³¹P NMR ($CDCl_3$, 162 MHz) d –10.1, –12.4; MS

(–ve ESI): 825.7, 859.8, 893.8, 927.9, 961.9, 995.9, 1030.0 (*m/2* for n = 12 – 18).

P¹-Dolichyl P²-[2-deoxy-2-trifluoroacetamido-α-D-glucopyranosyl] diphosphate (1d).

70
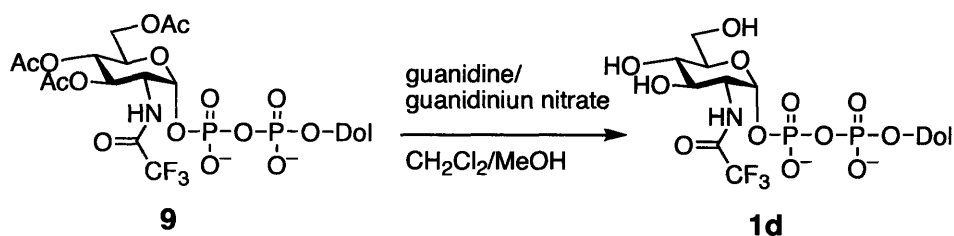
To a stirred solution of **9** (3 μmol) in CH_2Cl_2 (1.5 mL) at room temperature was added a solution of guanidine/guanidinium nitrate (1 mL). [Stock solution of the guanidine/guanidinium nitrate reagent was prepared by dissolving guanidinium nitrate (122.1 mg, 1 mmol) and sodium methoxide (10.8 mg, 0.2 mmol) in MeOH/CH_2Cl_2.] After 20 min, cation exchange resin (Dowex 50X8, pyridinium form, 200 mg) was added and the mixture was stirred for 10 min. The resin was then filtered and washed with $CHCl_3$/MeOH (2:1 v/v). Concentration of the solvent followed by column chromatography afforded the corresponding dolichylpyrophosphate-linked saccharide **1e** in 47% yield.

MS (–ve ESI): 1593.1, 1661.2, 1729.6, 1797.7, 1865.2 (M^- for n = 13 – 17).

P^1-Dolichyl P^2-[2-amino-α-D-glucopyranosyl] diphosphate (1e).
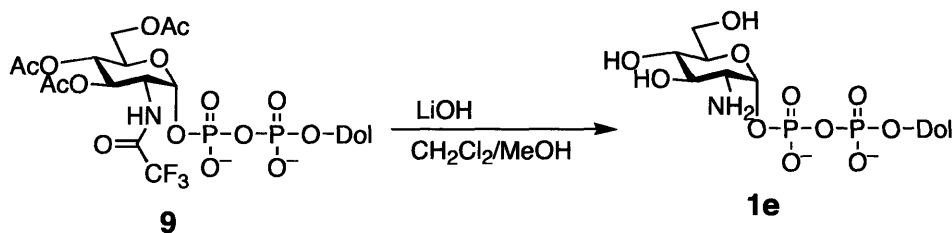
To a stirred solution of peracetylated dolichylpyrophosphate-linked saccharide **9** (4.4 μmol) in CH_2Cl_2/MeOH (1.2 mL, 5:1 v/v) at room temperature was added aqueous lithium hydroxide solution (3 M, 11 μmol). After 20 min, the mixture was worked up as in the preparation of **1d** to yield **1e** in 63% yield.

MS (–ve ESI): 1498.5, 1566.6, 1634.6, 1702.7, 1736.7 (M^- for n = 13 – 17).

Biological Methods

Crude porcine liver microsomes were prepared from fresh pig liver based on a protocol developed for rat liver.[24, 25] Briefly, a 100 g sample of fresh pig liver typically affords 2.5-5 mL of a microsomal preparation (~ 100 mg/mL protein in crude microsomal membranes) that can be stored at –80 °C in the presence of 30% glycerol. This material can be used for several months without significant loss of transferase activity. The natural substrate GlcNAc-PP-Dol **1a** was prepared according to literature procedure.[20] Radiolabeled uridine diphosphate N-acetyl-D-glucosamine-6-^{3}H [UDP-(^{3}H)GlcNAc, **2**, specific activity of 60 Ci/mmol, concentration = 1 mCi/mL] was obtained from American Radiolabeled Chemicals, Inc., St. Louis, MO.

Enzyme II Assay

The assay buffer consisted of 50 mM Tris acetate, pH 7.0, 3 mM dithiothreitol (DTT), 5 mM $MgCl_2$, 0.25 M sucrose and 1% Nonidet P-40. Dolichylpyrophosphate-linked monosaccharides (**1a-e**) were aliquoted from a chloroform/methanol stock solution of weighed material into an Eppendorf tube, and the solvent was evaporated under a stream of nitrogen. To this Eppendorf tube was added UDP-(^{3}H)GlcNAc (60 Ci/mmol for the study of 2DFGlc-PP-Dol (**1b**), and 0.6 Ci/mmol for the other analogues, final concentration = 0.167 μM) in ethanol/water, and the solvent was again evaporated under nitrogen. Assay buffer was added (final volume = 100 μL) to the Eppendorf tube and vortexed vigorously. The reaction was initiated by the addition of freshly thawed pig liver microsomes (5 or 10 μL) and shaken at 180 rpm. Aliquots (10 μL) of the reaction mixture were quenched with 0.6 mL of chloroform/methanol/2.5 mM $MgCl_2$ (3 : 2 : 1

v/v) at different times. The upper aqueous phase was removed and the lower phase washed twice with 100-μL aliquots of chloroform/methanol/water (2.75 : 44 : 53.25 v/v/v). The combined aqueous phases were then mixed with Ecolite (ICN Biomedicals, Inc.) (5.5 mL) and counted on a Beckman LS-5000TD scintillation counter to give the amount of unreacted UDP-(^{3}H)GlcNAc. The organic layer was dried and mixed with Betamax (ICN Biomedicals, Inc.) (5.5 mL) and counted as well. The percent conversion of dolichylpyrophosphate-linked disaccharide was based on [^{3}H$_{org}$/ (^{3}H$_{aq}$ + ^{3}H$_{org}$) × 100].

References

1.   Leloir, R. F.; Staneloni, R. J.; Carminatti, H.; Behrens, N. H., Biosynthesis of a *N,N'*-diacetylchitobiose containing lipid by liver microsomes - probable dolichol pyrophosphate derivative. *Biochem. Biophys. Res. Comm.* **1973,** 52, (4), 1285-1292.
2.   Chantret, I.; Dancourt, J.; Barbat, A.; Moore, S. E., Two proteins homologous to the N- and C-terminal domains of the bacterial glycosyltransferase MurG are required for the second step of dolichyl-linked oligosaccharide synthesis in *Saccharomyces cerevisiae. J. Biol. Chem.* **2005,** 280, (10), 9236-42.
3.   Kean, E. L., Topographical orientation in microsomal vesicles of the *N*-acetylglucosaminyltransferases which catalyze the biosynthesis of *N*-acetylglucosaminylpyrophosphoryldolichol and *N*-acetylglucosaminyl-*N*-acetylglucosaminylpyrophosphoryldolichol. *J. Biol. Chem.* **1991,** 266, (2), 942-946.
4.   Sharma, C. B.; Lehle, L.; Tanner, W., *N*-Glycosylation of yeast proteins. Characterization of the solubilized oligosaccharyl transferase. *Eur. J. Biochem.* **1981,** 116, (1), 101-8.
5.   Tai, V. W.; Imperiali, B., Substrate specificity of the glycosyl donor for oligosaccharyl transferase. *J. Org. Chem.* **2001,** 66, (19), 6217-28.
6.   Kaushal, G. P.; Elbein, A. D., Purification and properties of UDP-GlcNAc: dolichylpyrophosphoryl-GlcNAc GlcNAc transferase from mung bean seedling. *Plant Physiol.* **1986,** 81, 1086-1091.
7.   Hanson, S.; Best, M.; Bryan, M. C.; Wong, C. H., Chemoenzymatic synthesis of oligosaccharides and glycoproteins. *Trends Biochem. Sci.* **2004,** 29, (12), 656-63.
8.   Sharma, C. B.; Lehle, L.; Tanner, W., Solubilization and characterization of the initial enzymes of the dolichol pathway from yeast. *Eur. J. Biochem.* **1982,** 126, 319-325.
9.   Lehle, L.; Tanner, W., Formation of lipid-bound oligosaccharides in yeast. *Biochem. Biophys. Acta* **1975,** 399, (2), 364-374.

10. Lehle, L.; Tanner, W., Glycosyl transfer from dolichyl phosphate sugars to endogenous and exogenous glycoprotein acceptors in yeast. *Eur. J. Biochem.* **1978,** 83, (2), 563-570.

11. Lehle, L.; Tanner, W., The Specific Site of Tunicamycin Inhibition in the Formation of Dolichol-Bound N-Acetylglucosamine Derivatives. *FEBS Lett.* **1976,** 71, (1), 167-170.

12. Kyosseva, S. V., Biosynthesis of lipid-tri-to-heptasaccharides in microsomes from pig embryonic liver. *Dokl. Bolg. Akad. Nauk.* **1990,** 43, (3), 71-74.

13. Herscovics, A.; Warren, C. D.; Bugge, B.; Jeanloz, R. W., Biosynthesis of P1-di-N-acetyl-alpha-chitobiosyl P2-dolichyl pyrophosphate in calf pancreas microsomes. *J. Biol. Chem.* **1978,** 253, (1), 160-5.

14. Kyosseva, S. V.; Zhivkov, V. I., Biosynthesis of Lipid-Linked Oligosaccharides in Embryonic Liver - Formation of Dolichyl Pyrophosphate N-Acetylglucosaminyl Intermediates. *Int. J. Biochem.* **1983,** 15, (8), 1051-1057.

15. Kean, E. L.; Niu, N. Q., Kinetics of formation of GlcNAc-GlcNAc-P-P-dolichol by microsomes from the retina of the embryonic chick. *Glycoconjugate Journal* **1998,** 15, (1), 11-17.

16. Sala, R. F.; MacKinnon, S. L.; Palcic, M. M.; Tanner, M. E., UDP-N-trifluoroacetylglucosamine as an alternative substrate in N-acetylglucosaminyltransferase reactions. *Carbohydr. Res.* **1998,** 306, (1-2), 127-36.

17. Busca, P.; Martin, O. R., A convenient synthesis of alpha- and beta-D-glucosamine-1-phosphate and derivatives. *Tetrahedron Lett.* **1998,** 39, 8101-8104.

18. Wolfrom, M. L.; Bhat, H. B., Trichloroacetyl and trifluoroacetyl as N-blocking groups in nucleoside synthesis with 2-amino sugars. *J. Org. Chem.* **1967,** 32, (6), 1821-3.

19. Ellervic, U.; Magnusson, G., Guanidine/guanidinium nitrate; a mild and selective O-deacetylation reagent that leaves the N-Troc group intact. *Tetrahedron Lett.* **1997,** 38, 1627-1628.

20. Imperiali, B.; Zimmerman, J. W., Synthesis of dolichylpyrophosphate-linked oligosaccharides. *Tetrahedron Lett.* **1990,** 31, 6485-6488.

21. Tai, V. W.-F.; O'Reilly, M. K.; Imperiali, B., Substrate specificity of *N*-acetylglucosaminyl (diphosphodolichol) *N*-acetylglucosaminyl transferase, a key enzyme in the dolichol pathway. *Bioorg. Med. Chem. Lett.* **2001,** 9, 1133-1140.

22. Kundu, S. K.; Chakravarty, S.; Bhaduri, N.; Saha, H. K., A novel spray reagent for phospholipid detection. *J Lipid Res* **1977,** 18, (1), 128-30.

23. Danilov, L. L.; Maltsev, S. D.; Shibaev, V. N., Phosphorylation of polyprenols with tetra-n-butylammonium phosphate and trichloroacetonitrile. *Bioorg. Khim.* **1988,** 14, 1287-1289.

24. Imperiali, B.; Shannon, K. L., Differences between Asn-Xaa-Thr-containing peptides: a comparison of solution conformation and substrate behavior with oligosaccharyltransferase. *Biochemistry* **1991,** 30, (18), 4374-80.

25. Behrens, N. H.; Tabora, E., Dolichol intermediates in the glycosylation of proteins. *Methods Enzymol.* **1978,** 50, 402-35.

Chapter III: Alg1: Specificity studies and chemoenzymatic synthesis of

ManGlcNAc$_2$-PP-Dol

Introduction

ALG1 was first discovered in *Saccharomyces cerevisiae* through the use of a genetic screen for mutants defective in *N*-linked glycosylation.[1] One isolated temperature-sensitive mutant, *alg1-1*, was found to accumulate dolichylpyrophosphate-linked $GlcNAc_2$ ($GlcNAc_2$-PP-Dol) at the non-permissive temperature. The mutation in *alg1-1* was therefore postulated to affect the β1,4-mannosyltransferase that catalyzes the third step in the dolichol pathway, transferring a mannose residue (Man) from GDP-Man to yield Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol (Scheme III-1). The change in configuration at the anomeric center of the mannosyl residue from the α-anomer in the GDP-Man donor to the β-anomer in the $ManGlcNAc_2$-PP-Dol product makes Alg1 an inverting glycosyltransferase.
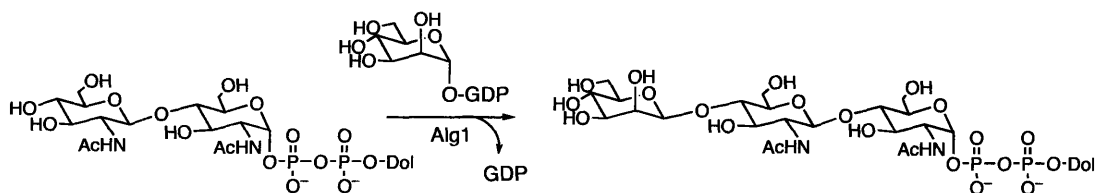
Scheme III-1. β1,4-Mannosylation reaction catalyzed by Alg1.

By complementation of the temperature-sensitive mutation *alg1-1* with a total genomic yeast DNA library, the region of sequence containing *ALG1* was cloned, and *E. coli* lysates containing the expressed protein demonstrated the ability to transfer a mannose residue from GDP-Man to $GlcNAc_2$-PP-Dol.[2] This work led to the subsequent sequencing of the *ALG1* gene, which revealed an open reading frame of 449 amino acids,

and a predicted molecular weight of 51.9 kDa.[3] Further examination of the sequence and

topology has revealed that Alg1 has at least one predicted transmembrane domain at the

N-terminus, as predicted by the TMHMM server, v. 2.0 (Figure III-1).[4] This domain

anchors Alg1 into the ER membrane, and the catalytic domain faces the cytosolic side of

the ER membrane,[5] making it a type II membrane protein belonging to the PFAM

glycosyltransferase 1 family.[6]
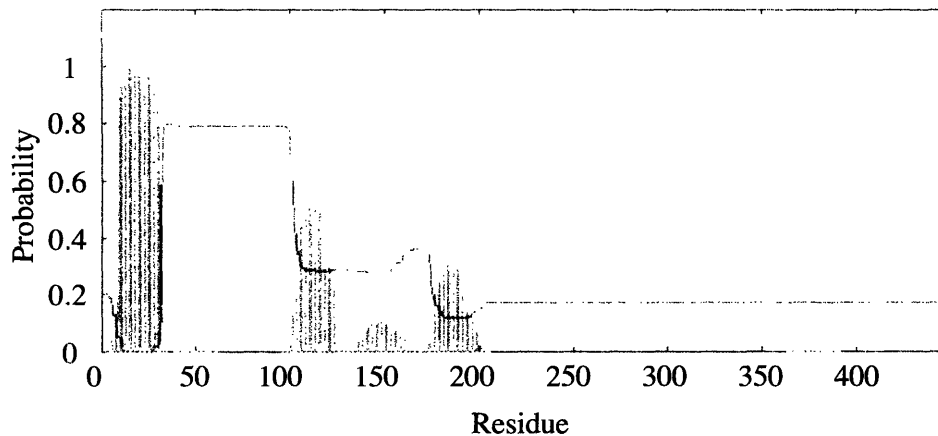
Figure III-1. Predicted topology of Alg1. Plot of the probability that Alg1 residues are

in a membrane-spanning region.

The human homologue of Alg1, Hmt-1, was found by using the amino acid

sequence of yeast Alg1 to search the expressed sequence tag (EST) database.[7] This study

further showed that *HMT1* complemented the temperature-sensitive defect of *alg1-1*.

Mutations in *HMT1* have been linked to a recently discovered congenital disorder of

glycosylation, CDGIk, which can lead to death in early infancy, and is specifically

characterized by reduced Hmt-1 activity and accumulation of $GlcNAc_2$-PP-Dol in

patients.[8-10] Interestingly, it was recently demonstrated that overexpression of the yeast gene *MPG1*, which encodes a GDP-Man pyrophosphorylase, a key enzyme in the biosynthesis of GDP-Man, largely overcame the temperature-sensitive glycosylation defect of the *alg1-1* mutant by increasing the available pool of GDP-Man.[11]

Alg1 has also received considerable attention due to the challenge of preparing β-mannosides by chemical synthesis, and therefore it has been one of the most characterized of the enzymes on the pathway to date. Early studies with solubilized and partially purified enzyme from yeast membranes revealed a distinct pH optimum of 7.5, a specific activity of 21.4 nmol × h^{-1} × mg^{-1} with a fixed, non-saturating $GlcNAc_2$-PP-Dol concentration, and a requirement for a divalent cation, specifically Mg^{2+} or Mn^{2+}, with a 1.7-fold preference for the former.[12] The K_m values of both substrates were also found in this study to be 17 μM (apparent) and 7 μM for $GlcNAc_2$-PP-Dol and GDP-Man, respectively. Interestingly, it was found that substitutions could be made with respect to the isoprenoid, as the same acceptor K_m was also found for $GlcNAc_2$-PP-phytanyl (Figure III-2) using crude yeast microsomes as an enzyme source.[13] In addition to the lack of a requirement for the dolichyl moiety, deletion of the predicted transmembrane domain (residues 1-35) of Alg1 did not result in a significant loss in transferase activity from *Escherichia coli* expression, and both the phytanyl-linked variant and the native substrate were accepted to a similar degree.[14, 15] Immobilization of a His-tagged version of the truncated Alg1 on Ni-NTA beads provided a tool for preparative-scale synthesis of phytanylpyrophosphate-linked $ManGlcNAc_2$ that remained stable at 4 °C for several weeks.[16, 17] To our knowledge there has not been any previous work on the specificity of Alg1 for the saccharide portion of the dolichyl-linked substrate.
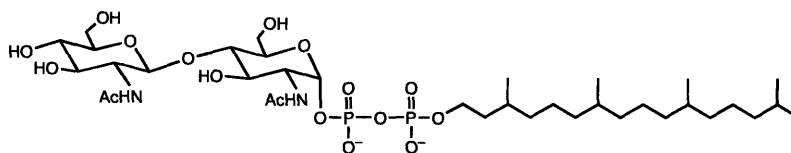
Figure III-2. Structure of GlcNAc$_2$-PP-phytanyl.

Using a variety of mutant yeast strains deficient in various steps in the assembly

of the tetradecasaccharide donor, Glc$_3$Man$_9$GlcNAc$_2$-PP-Dol, it has been found that the

dolichol pathway enzymes follow a defined order of addition of saccharides, and this

behavior was attributed to strict substrate specificity of the glycosyltransferases.[18]

However, this study does not rule out the possibility that complexation of these enzymes

may play a crucial role in the observed fidelity. Recent gel filtration and

immunoprecipitation studies on Alg1, Alg2, and Alg11 have shown that Alg1 homo-

oligomerizes, and also immunoprecipitates Alg2 and, separately, Alg11.[19] Interestingly,

Alg2 and Alg11 were not found to co-immunoprecipitate with one another. In the same

study, the genetic basis for the phenotype of the *alg1-1* yeast temperature-sensitive

mutant was found to be a mutation that leads to a premature stop codon and the

subsequent truncation of the protein, deleting the final sixteen amino acids at the C-

terminus. Using an identically truncated Alg1 construct in the same immunoprecipitation

experiments, it was revealed that while this minor truncation has a moderately

detrimental effect on Alg2 and Alg11 complexation, it completely abolishes Alg1 homo-

oligomerization. Further, mutations at residues E278, G310, and H356 of Alg1 were

found to disrupt catalytic activity without affecting the oligomerization properties. Thus,

as expected, complex formation and catalytic function are determined by different parts of the protein sequence.

One of our goals in the study of Alg1 was to further probe substrate specificity from two different perspectives. One approach began with the question of whether a synthetic glycopeptide with an asparagine-linked $GlcNAc_2$ moiety would be accepted as a substrate of Alg1, which normally recognizes a dolichylpyrophosphate-linked $GlcNAc_2$ substrate (Figure III-3). The ability of Alg1 to mannosylate the glycopeptide would be of particular value due to the need for new methods of building homogenous glycoproteins. This capability would facilitate biochemical studies of glycoproteins as well as the production of glycoproteins for therapeutics, which cannot be administered with the glycan heterogeneity that is found in glycoproteins isolated from native sources. We also wanted to directly address the question of whether specificity drives the fidelity of the pathway. Dolichylpyrophosphate-linked GlcNAc (GlcNAc-PP-Dol) (Figure III-3) was tested as a substrate of Alg1 using purified enzyme to determine whether this earlier intermediate could be accepted and a new product detected.
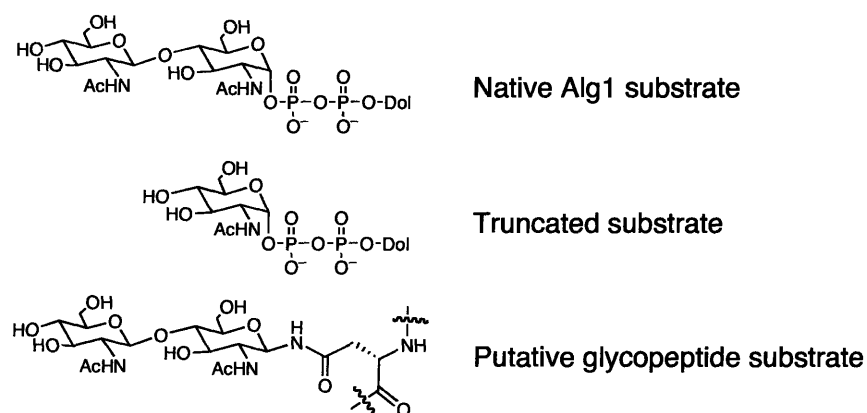
Figure III-3. Comparison of the Alg1 substrate with alternate substrates that were tested.

80

The ultimate goal was to establish an expression system for an active, pure, and stable full-length Alg1 preparation. Such a preparation enables both the chemoenzymatic synthesis of the trisaccharide intermediate, $ManGlcNAc_2$-PP-Dol, as well as future studies toward understanding the coordinated functions of a consecutive series of enzymes in the pathway.

Results and Discussion

Cloning and expression of Alg1 constructs

Expression systems for three different Alg1 constructs were established, including a soluble, truncated Alg1 lacking the N-terminal transmembrane domain (Alg1Δ1-35) expressed in *E. coli*, a thioredoxin fusion protein of full-length Alg1 in *E. coli*, and the full-length Alg1 in *S. cerevisiae*. *ALG1Δ1-35* encodes the Alg1 sequence 36-449, which comprises the full-length protein with the first 35 residues omitted. This gene was cloned into a pET-15b expression vector (Novagen), and therefore bears an N-terminal His tag. This feature enabled Ni-NTA purification of the enzyme from clarified lysate in the presence of the detergent Triton X-100 (1%). Expression of Alg1 in the absence of the transmembrane domain enabled isolation and purification, as described in "Experimental Procedures", of 0.45 mg of Alg1Δ1-35 from a 100-mL expression culture (Figure III-4a).

For expression of full-length Alg1, a cloning strategy was chosen that provided the flexibility to test expression in a variety of different systems from a single clone by homologous recombination with vectors specified for expression in different host organisms. Thus, the Gateway Cloning Technology (Invitrogen) was used to provide

ALG1 constructs in both of the vectors, pBAD-DEST49 and pYES-DEST52, for expression in *E. coli* and *S. cerevisiae*, respectively. Expression of yeast genes in *E. coli* is complicated by the difference in codon usage in these two genomes. In the yeast *ALG1* sequence, 36 of the 450 codons (8.0%) are considered "rare", due to the significant under representation of these codons in the coding sequence of the *E. coli* genome. Similarly, the yeast *ALG2* and *ALG11* genes contain 35 out of 504 (6.9%) and 30 out of 549 (5.5%) rare codons, respectively. Appending thioredoxin to the N-terminus of sequences has been shown to improve expression and solubility of many protein targets.[20] Thus, the pBAD-DEST49 vector, which encodes a fusion protein whereby thioredoxin (TRX) is fused to the N-terminus, was chosen for *ALG1* expression in *E. coli*. A C-terminal His tag enables purification, and the V5 epitope, a 14-amino acid sequence from the P and V proteins of the paramyxovirus,[21] is also included at the C-terminus for detection by Western blot analysis. Another advantage of this expression system is that the arabinose-induction enables very tightly regulated expression. This feature makes the pBAD vector ideal for expression of highly insoluble and/or toxic proteins. The pYES-DEST52 vector does not include the TRX domain, but bears the same C-terminal tags, and is designed to enable regulated, galactose-inducible expression.

TRX-Alg1-V5-His was expressed in *E. coli* and purified by Ni-NTA using procedures analogous to those used for Alg1Δ1-35, as described in "Experimental Procedures". Due to modest expression levels and relatively weak binding of the expressed protein to Ni-NTA (significant amounts washed away with 20 mM imidazole), less than 1 mg per liter of expression culture was isolated (Figure III-4b-c). Several attempts were made to remove the TRX domain from the expressed protein by taking

advantage of a proteolytic cleavage site for enterokinase that had been engineered between TRX and the target protein. Unfortunately, the proteolysis reaction was very low yielding, and generally led to the rapid precipitation of protein.

Similarly low yields were obtained from the expression of Alg1-V5-His in *S. cerevisiae*, despite the fact that this is the homologous host. Expression was induced by exchanging the glucose in the growth media with galactose as the main carbon source in the yeast culture. The Alg1 construct was isolated by fractionating crude yeast microsomes, and then purifying detergent-solubilized microsomes by Ni-NTA affinity chromatography, as described in "Experimental Procedures". Less than 1 mg of purified Alg1 was consistently isolated from 4-L expression cultures (Figure III-4d-e). However, the amount isolated proved to be ample for the experiments described here, based on the activity of this preparation.

His-Alg1Δ1-35, *E. coli* expression (A)

TRX-Alg1-V5-His, *E. coli* expression (B,C)

Alg1-V5-His, *S. cerevisiae* expression (D,E))

| 6x His tag | V5 epitope | TRX domain |
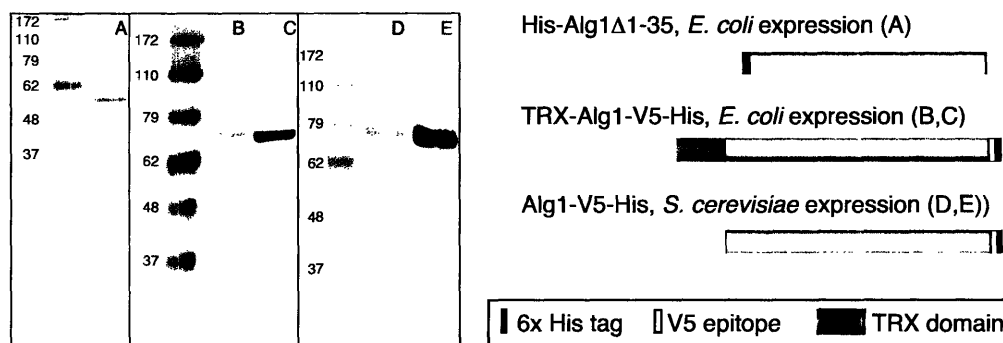
Figure III-4. SDS-PAGE and Western blot analysis of Alg1 constructs from *E. coli* and *S. cerevisiae* expression. A) Coomassie stain of purified Alg1Δ1-35, B) Coomassie stain of purified TRX-Alg1 from *E. coli* expression, C) Western blot analysis of TRX-Alg1, D) Coomassie stain of purified Alg1 from *S. cerevisiae* expression, E) Western blot analysis of Alg1. Molecular weights of the markers are shown in kDa. Western blots were probed with the α-V5 antibody (Invitrogen).

Characterization of Alg1Δ1-35 from *E. coli* expression

Mannosyltransferase activity of Alg1Δ1-35 was detected by the radioactive assay described in Chapter II (Figure II-2), whereby aqueous/organic extractions were used to separate tritiated GDP-(^{3}H)Man from tritiated product, (^{3}H)ManGlcNAc$_2$-PP-Dol, at a number of time points. In order to demonstrate the enzyme dependence of this reaction, initial assays were carried out with different amounts of Alg1Δ1-35 (Figure III-5).
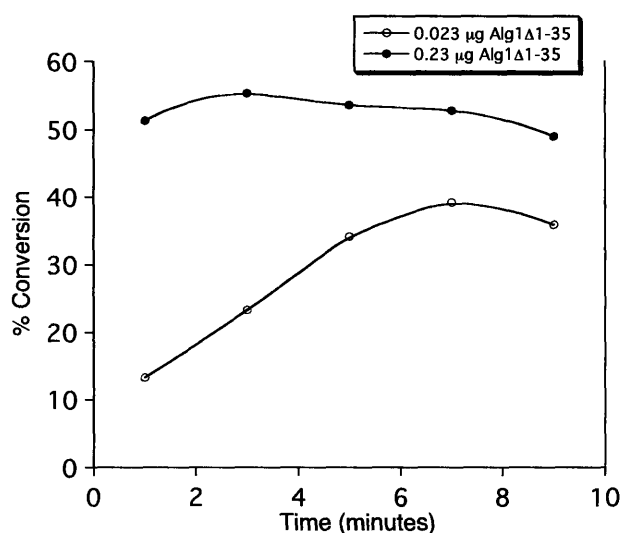
Figure III-5. Alg1Δ1-35 activity assays, varying enzyme concentration (0.48 or 4.8 pM) to show enzyme dependence. Both assays contained 0.15 mM GlcNAc$_2$-PP-Dol and 1.67 μM GDP-(^{3}H)Man (10,000 dpm/assay). Approximated specific activity, based on the first three time points of the assay with 0.023 μg enzyme, is 23 μmol × hr^{-1} × mg^{-1}.

To further demonstrate that the activity seen in this assay was specifically the result of mannosyl transfer to GlcNAc$_2$-PP-Dol, and not an artifact of the extraction

procedure, UDP-(^{3}H)GlcNAc and UDP-(^{3}H)Glc were used in place of GDP-(^{3}H)Man and shown not to be transferred to the organic-soluble layer by Alg1Δ1-35 (Figure III-6).
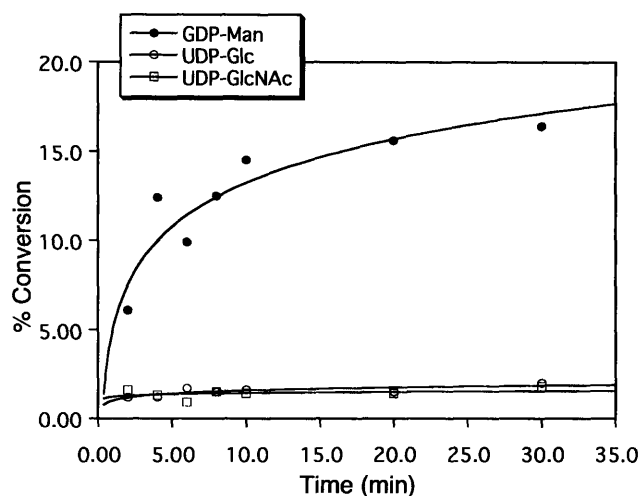
Figure III-6. Activity assays of Alg1Δ1-35 in the presence of UDP-(^{3}H)GlcNAc, (^{3}H)UDP-Glc, or (^{3}H)GDP-Man.

In order to confirm the product of the reaction, organic-soluble products of an Alg1Δ1-35 reaction were analyzed by gel filtration chromatography on BioGel P-2 (BioRad). Products were isolated from a reaction containing 0.23 μg Alg1Δ1-35, 0.45 μM (^{3}H)GDP-Man (22 Ci/mmol), and approximately 5 μM GlcNAc$_2$-PP-Dol, in which 31% conversion with respect to the GDP-(^{3}H)Man substrate was achieved, corresponding to approximately 85,000 dpm of product. Following mild acid hydrolysis of the glycan, as described in "Experimental Procedures", half of this glycan sample was loaded onto the column in a volume of 250 μL. As shown in Figure III-7, a peak was detected that corresponded to a product larger than a disaccharide. Since Alg1 is known to catalyze

the conversion of $GlcNAc_2$-PP-Dol to $ManGlcNAc_2$-PP-Dol, this product was thus characterized as $(^3H)ManGlcNAc_2$.
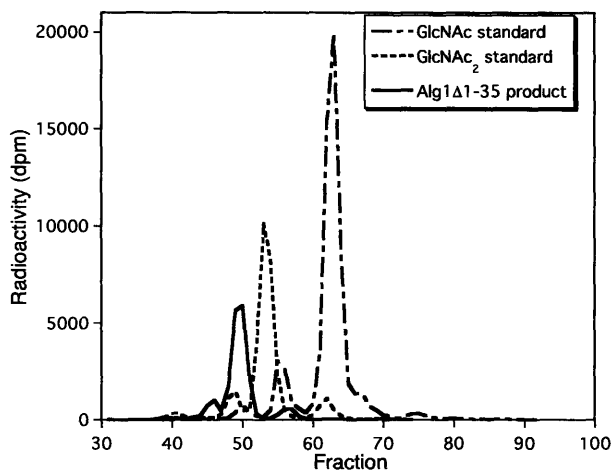
Figure III-7. Gel filtration analysis of the product of the Alg1Δ1-35 reaction. GlcNAc and $GlcNAc_2$ standards are shown superimposed for comparison.

To test the ability of Alg1Δ1-35 to transfer a mannose residue to a glycopeptide substrate, an N_α-Fmoc-protected $GlcNAc\beta1,4$-$GlcNAc\beta1,N_\delta$-Asn amino acid building block for solid-phase peptide synthesis was synthesized in nine steps as described previously.[22] Briefly, the chitobiose moiety was isolated from enzymatically-degraded chitin, aminated at the anomeric position, and coupled to the activated carboxylate side chain of an Fmoc-Asp-OAlloc residue. Finally, the hydroxyl groups of the chitobiose were protected with t-butyldimethylsilyl (TBDMS) groups in preparation for solid phase peptide synthesis. The building block was incorporated into a peptide sequence derived from interferon-β in the position of a natively-glycosylated Asn (Figure III-8). The free N-terminus of the peptide was treated with the reactive fluorescein isothiocyanate derivative to incorporate fluorescein for sensitive detection and quantitation *via* HPLC.

86
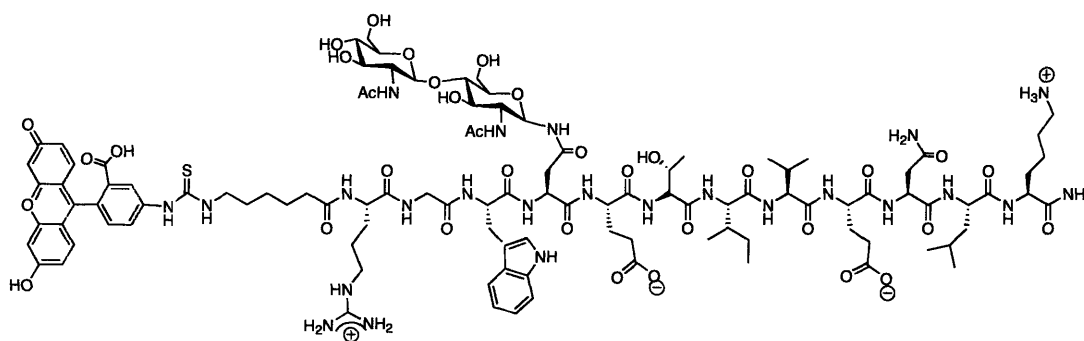
Figure III-8. Structure of the glycopeptide tested as a putative Alg1 substrate.

This glycopeptide was subsequently tested as a substrate of Alg1. Prior to, and

following, incubation with Alg1Δ1-35 in the presence of GDP-Man, the glycopeptide

was analyzed on an analytical C_{18} HPLC column, monitoring UV absorbance, in order to

detect mannosylation by a change in retention time. A range of Alg1Δ1-35

concentrations were used, up to 110 μM, which is 400 times the amount necessary to

achieve nearly 50% conversion with the native substrate. A shift in retention time was

not observed in the HPLC trace of the glycopeptide from the reaction mixture, and mass

spectrometry of the collected peak revealed only starting material (Figure III-9).

Therefore, Alg1 did not accept this alternate substrate to a degree that reached the

detection limit of this assay, suggesting that even in the absence of the transmembrane

domain, which is most likely the region that interacts with the dolichyl unit, replacement

of dolichylpyrophosphate with a peptide is not tolerated by the enzyme.
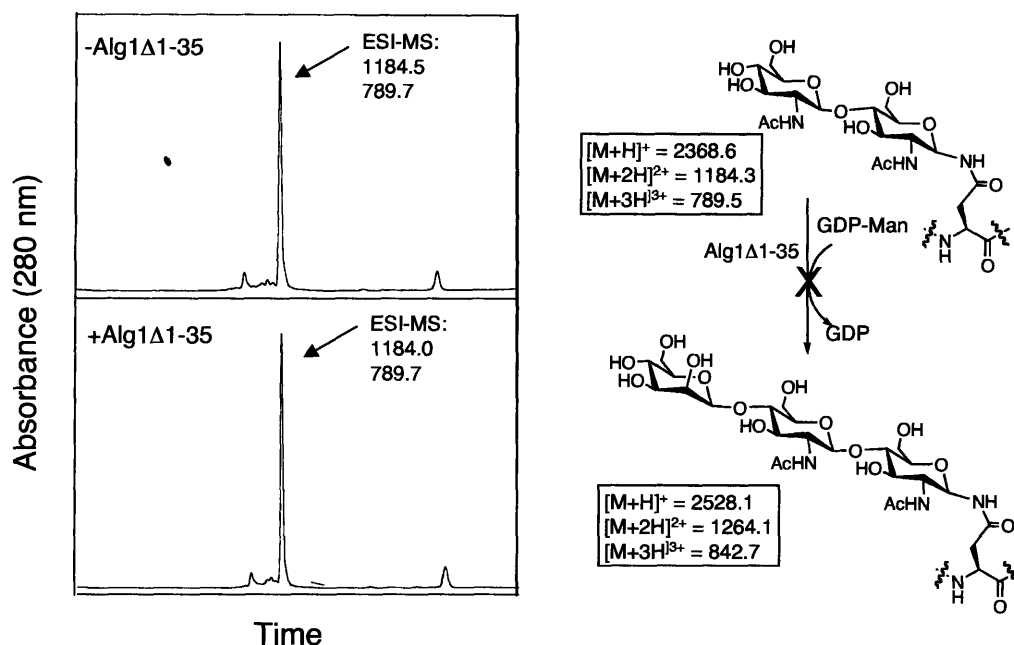
Figure III-9. HPLC and ESI-MS results of Alg1 glycopeptide assays. HPLC traces show starting material and Alg1Δ1-35-treated glycopeptide. The expected masses are indicated below the starting material and product.

Characterization of TRX-Alg1 from *E. coli* expression

The Alg1 fusion protein, with thioredoxin (TRX) at the N-terminus, and a His-tag and V5 epitope at the C-terminus, demonstrated activity in the same radioactive assay described in the previous section (Figure III-10). The results of these assays, performed in the presence of either $MgCl_2$ or $MnCl_2$, suggest that TRX-Alg1 has a preference for Mg^{2+} over Mn^{2+} under these conditions. Thus, $MgCl_2$ was used in all subsequent Alg1 assays, and from the data shown in Figure III-10 with Mg^{2+} as the divalent metal, a specific activity of 1.2 nmol × hr^{-1} × mg^{-1} for this Alg1 preparation was estimated with

GDP-Man at a concentration of 0.18 µM. It should be noted that in these experiments, the enzyme concentration is in excess of the GDP-[^{3}H]Man concentration.
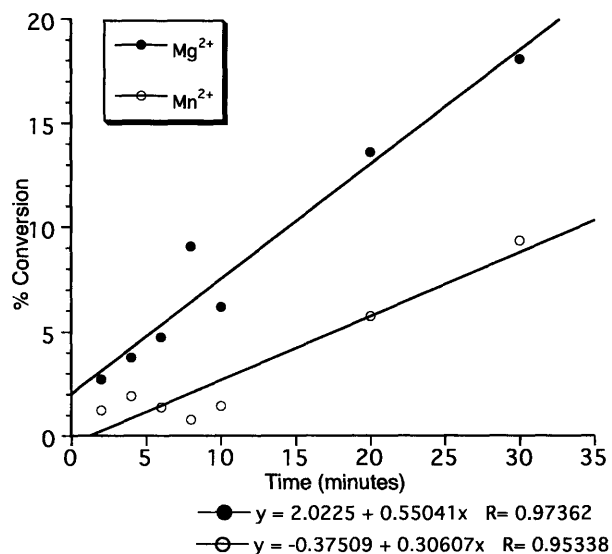
Figure III-10. Activity assays of TRX-Alg1 (0.8 µM) with GlcNAc$_2$-PP-Dol, GDP-(^{3}H)Man (0.2 µM), and 5 mM of either MgCl$_2$ or MnCl$_2$.

TRX-Alg1 (4.5 µM) was also assayed with the preceding intermediate in the pathway, GlcNAc-PP-Dol, to examine the specificity of TRX-Alg1 for the glycosyl acceptor substrate. This truncated version of the native substrate of Alg1 was found to be a substrate of TRX-Alg1 in the presence of GDP-(^{3}H)Man, but not when UDP-(^{3}H)GlcNAc was used in place of GDP-(^{3}H)Man (Figure III-11). As shown with the native substrate, Alg1 is specific for the nucleotide-sugar donor, GDP-Man, and the observed activity is therefore not an artifact of the assay. Thus, it appears that TRX-Alg1 is capable of accepting a truncated version of the native substrate, GlcNAc$_2$-PP-Dol.

89
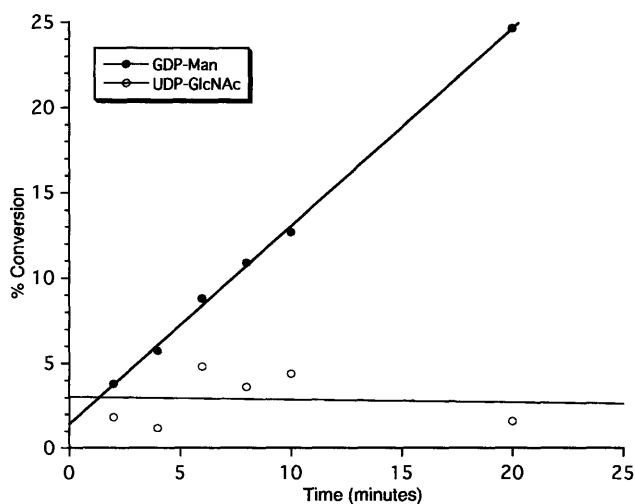
Figure III-11. Activity assays of TRX-Alg1 (4.5 μM) with GlcNAc-PP-Dol (approx. 0.3 mM) and either GDP-(^{3}H)Man (0.18 μM) or UDP-(^{3}H)GlcNAc (0.17 μM).

The rates for TRX-Alg1 reactions containing GlcNAc$_2$-PP-Dol or GlcNAc-PP-Dol were measured as shown in Figure III-12. From these data, an activity of 0.44 nmol × hr^{-1} × mg^{-1} was estimated for TRX-Alg1 in the non-specific reaction with GlcNAc-PP-Dol and 0.18 μM GDP-Man, while the activity for TRX-Alg1 with the native substrate, GlcNAc$_2$-PP-Dol in the presence of 0.23 μM GDP-Man was 2.8 nmol × hr^{-1} × mg^{-1}. A direct comparison between the two dolichylpyrophosphate-linked substrates should be interpreted carefully, due to the difficulty in quantifying these substrates accurately, and the unknown physical properties of the substrates in detergent solution. For instance, the critical micellar concentration (CMC) values are not known for either of the substrates, and may differ.

90
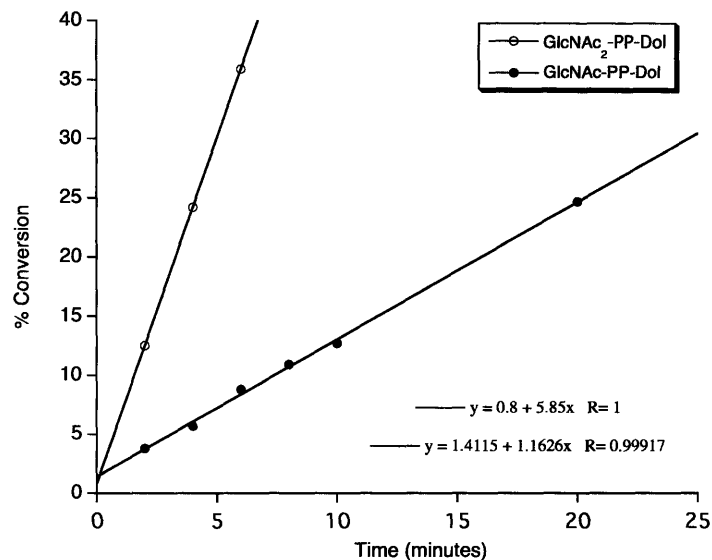
Figure III-12. GlcNAc-PP-Dol vs. GlcNAc$_2$-PP-Dol in the TRX-Alg1 reaction.

Activity assays included TRX-Alg1 (4.5 μM) and either GlcNAc-PP-Dol (closed circles)

or GlcNAc$_2$-PP-Dol (open circles).

In order to identify the product of this non-specific Alg1 reaction, gel filtration

analysis was employed, which separates glycan molecules on the basis of the number of

monosaccharide units they comprise. The radiolabeled glycan product was hydrolyzed

from dolichylpyrophosphate and compared to (^{3}H)GlcNAc, (^{3}H)GlcNAc$_2$, and

(^{3}H)ManGlcNAc$_2$ standards by gel filtration chromatography on BioGel P-4 (BioRad)

(Figure III-13). This study indicated that the product of the non-specific Alg1 reaction to

be a trisaccharide, suggesting that two mannose residues are transferred to GlcNAc-PP-

Dol to afford the non-native intermediate, Man$_2$GlcNAc-PP-Dol. This product could not

be confirmed by mass spectrometry due to the lack of sufficient material.
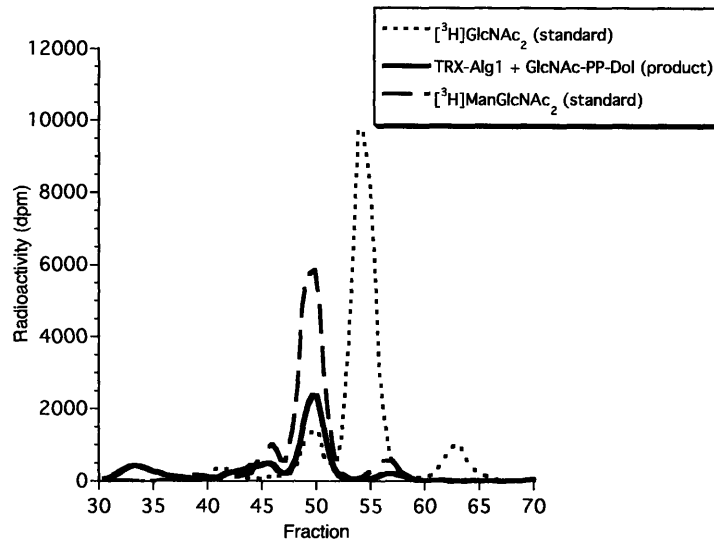
Figure III-13. Gel filtration analysis of the non-native TRX-Alg1 product (solid line) after incubation with GlcNAc-PP-Dol and GDP-(^{3}H)Man.

There are several possible explanations for the observed non-specific behavior of TRX-Alg1. The enzyme concentrations in these assays were in the micromolar range, and the presence of the TRX domain may have influenced the structure enough to affect the function. Similarly, the C-terminal tags of this construct may interfere with homo-oligomerization, considering that the removal of the final 16 residues of the Alg1 sequence was found to abrogate this complexation.[19] Finally, the enzyme has been removed from the membrane and any possible interacting partners. Therefore, we sought to test this non-specific behavior in a more native-like environment, specifically in a yeast microsomal preparation. Microsomes are prepared by differential centrifugation of yeast cell lysate, and consist of non-solubilized ER and Golgi membranes. To simplify the experiment, a crude microsomal preparation from the temperature-sensitive yeast strain, *alg2*, was employed, which accumulates the trisaccharide intermediate *in vitro*.

Thus, to evaluate native Alg1 specificity in this system, GlcNAc-PP-Dol was incubated with the microsomal preparation in the presence of both UDP-(^{3}H)GlcNAc and GDP-(^{3}H)Man, as a positive control, or in the presence of only GDP-(^{3}H)Man. The expected product in the positive control was ManGlcNAc$_2$-PP-Dol. The products of these reactions, extracted into the organic phase and hydrolyzed from dolichol, were analyzed by gel filtration on Bio Gel P-4 (BioRad). As shown in Figure III-14, in the presence of both nucleotide-activated sugars, GlcNAc-PP-Dol is elongated through the sequential action of Enzyme II, using UDP-(^{3}H)GlcNAc to produce (^{3}H)GlcNAc$_2$-PP-Dol, and Alg1, using GDP-(^{3}H)Man to produce (^{3}H)ManGlcNAc$_2$-PP-Dol. In contrast, no elongation of GlcNAc-PP-Dol was detected upon incubation with microsomes in the presence of only GDP-(^{3}H)Man. Therefore, Alg1 in the native system does not accept GlcNAc-PP-Dol to the extent that any product could be detected in this experiment.
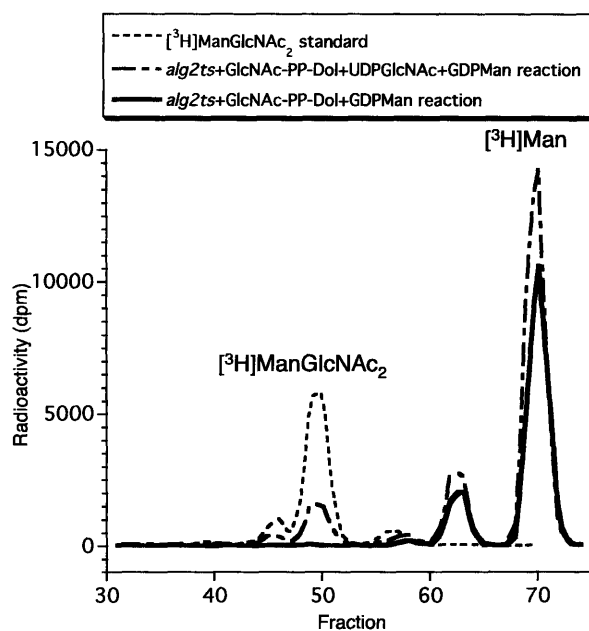
Figure III-14. Gel filtration analysis of *alg2* assay products. The Man peak is from Man-P-Dol synthase activity in the microsomes with residual Dol-P, and provides an internal standard for overall activity of the microsome preparation. The peak at 62-64 minutes is a shoulder of the Man peak, which is a recurring feature of all peaks on this column, but not further investigated. Shoulder and peak sizes were always proportional.

Characterization of Alg1-V5-His from *S. cerevisiae* expression

While the isolated yield of Alg1-V5-His from *S. cerevisiae* expression was not much greater in yeast than in bacteria, the activity was improved by three orders of magnitude in the presence of the native substrate $GlcNAc_2$-PP-Dol. From the data shown in Figure III-15 for the reaction with $GlcNAc_2$-PP-Dol as well as three separate assays, the specific activity of this preparation of Alg1 from *S. cerevisiae* was 1.8 μmol × hr ×

mg (+/- 0.99 μmol × hr × mg standard deviation based on four assays). Two of the possible sources of the non-specific behavior of TRX-Alg1, specifically high enzyme concentration and the presence of the TRX, could now be tested. Thus, Alg1 (0.025 μM) was also assayed with GlcNAc-PP-Dol in the presence of 2 μM GDP-Man. As shown in Figure III-15, turnover of the non-native substrate, GlcNAc-PP-Dol, also occurs, with an activity of 0.11 μmol × hr^{-1} × mg^{-1}. In this case also, using GlcNAc-PP-Dol as the acceptor substrate, the activity of the Alg1 preparation from *S. cerevisiae* was approximately three orders of magnitude higher than that of TRX-Alg1. These results suggest that the observed non-specific activity was not a result of using high concentrations of enzyme or a perturbation of the structure due to the presence of the TRX domain.
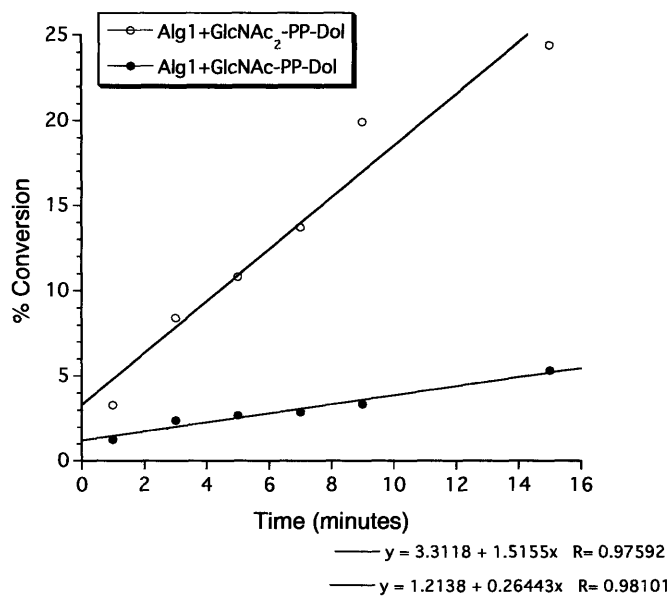
Figure III-15. Alg1 (0.025 μM) assays with GlcNAc-PP-Dol or GlcNAc$_2$-PP-Dol in the presence of 2 μM GDP-Man.

Chemoenzymatic preparation of ManGlcNAc$_2$-PP-Dol using Alg1

The Alg1 preparation from *S. cerevisiae*, stored in 50% glycerol at –20 °C, is extremely stable, with no detectable loss of activity after nearly two years. Therefore, this preparation was chosen to be used as a tool for the chemoenzymatic preparation of Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol. As the putative substrate of Alg2, this intermediate was crucial for characterization of the later steps in the dolichol pathway. The synthesis began with the enzymatic degradation of chitin and subsequent modification to yield peracetylated chitobiose.[23] The remaining steps leading to dolichyl-linked chitobiose were carried out as described previously.[24]

To monitor mannosylation of GlcNAc$_2$-PP-Dol, a new non-radioactive assay was implemented, in which hydrolyzed glycan products were labeled by reductive amination with the fluorophore, 2-aminobenzamide (2-AB) (Figure III-16),[25] and then separated by normal-phase HPLC and observed by fluorescence detection (λ_{Ex} = 330 nm, λ_{Em} = 420 nm).[26]
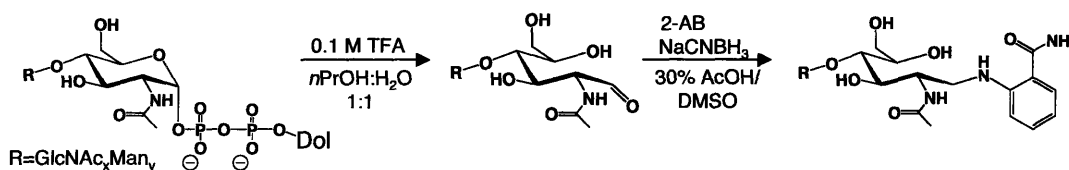
Figure III-16. Labeling of reducing sugars with 2-aminobenzamide.

The Alg1-catalyzed conversion of GlcNAc$_2$-PP-Dol to ManGlcNAc$_2$-PP-Dol was detected by this method, and was found to occur in greater than 95% yield, as estimated by the relative HPLC peak sizes (Figure III-17). The peak with retention time of 40

minutes was characterized by MALDI-MS, and the $\beta1,4$-linkage was verified by sensitivity to a $\beta1,4$-mannosidase. The expected mass of the 2-AB labeled ManGlcNAc$_2$ [M+H]$^+$ is 707.7, and the observed masses in the MALDI spectrum of the 40-minute peak are [M + H]$^+$ = 707.5 and [M + Na]$^+$ = 729.4.
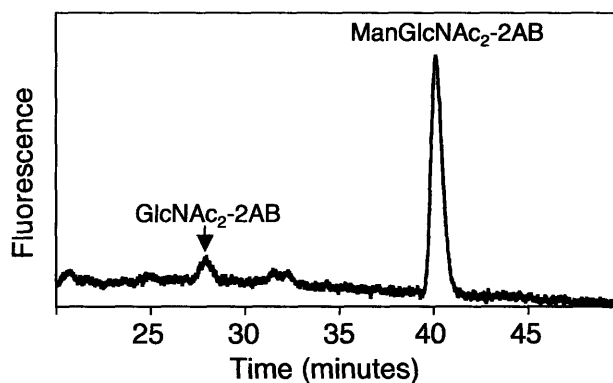
Figure III-17. HPLC trace of ManGlcNAc$_2$-PP-Dol preparation using Alg1.

Using the HPLC assay to test the non-specific behavior of Alg1 from *S. cerevisiae* expression, no product formation was detected when Alg1 was incubated with GlcNAc-PP-Dol in the presence of 1 mM GDP-Man. This result suggested that the highly sensitive radiolabeled assay is necessary to observe the non-specific behavior, and that while this behavior does appear to be real, it does not translate to a high degree of turnover, and is likely why it has not been observed previously. Importantly, this activity does not interfere with the HPLC assay, which was subsequently used to annotate the correct activity of Alg2 (Chapter IV) and Alg11 (Chapter V).

Conclusion

Through the cloning and expression of various Alg1 constructs, some of the details of the inherent substrate specificity of this mannosyltransferase, in the absence of possible participation from the membrane and/or neighboring dolichol pathway glycosyltransferases, have been described. Despite several attempts to mannosylate an Asn(GlcNAc$_2$)-containing synthetic glycopeptide, the specificity of Alg1, even in the absence of the predicted transmembrane domain, does not allow this transformation to occur at a detectable level. It is likely that the pyrophosphate moiety is an important binding determinant, and while it is known that the dolichyl group can be replaced by a phytanyl group, the limits of tolerance for further truncation and/or completely divergent saccharide-presenting scaffolds is still unknown. It is possible that optimization for recognition may be achieved by modifying the peptide sequence, however, such manipulation would limit the general application of Alg1 for the chemoenzymatic synthesis of glycoproteins.

Alg1, isolated from the native membrane environment as well as the rest of the enzymes of the pathway, appears to accept GlcNAc-PP-Dol to carry out two non-specific reactions, resulting in the formation of a non-native trisaccharide intermediate. A specific activity of 0.44 nmol $\times$ hr^{-1} $\times$ mg^{-1} was estimated for TRX-Alg1, with respect to the turnover of the GDP-Man substrate. If indeed two Man units are transferred to the GlcNAc-PP-Dol acceptor, then turnover of the GlcNAc-PP-Dol substrate would be half of this value. Using an Alg1 preparation that exhibited approximately three orders of magnitude greater activity with the native substrate, a similar increase in activity was

observed toward the non-native substrate, for which a specific activity of 0.11 μmol × hr^{-1} × mg^{-1} was estimated. The possible formation of Man$_2$GlcNAc-PP-Dol from the GlcNAc-PP-Dol intermediate in the non-specific Alg1 reaction is understandable, considering the chemistry of the two corresponding steps in the native system. Each of the two steps requires the formation of a β1,4-linkage, and the ability of Alg1 to accept the non-native disaccharide is plausible since the Man and GlcNAc residues only differ from each other at the C-2 position (highlighted in gray, Figure III-18), which is far from where the chemistry occurs.
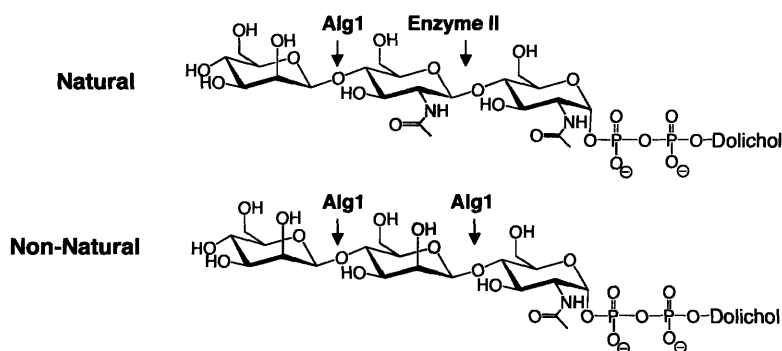
Figure III-18. Native and proposed non-native trisaccharide intermediates.

Therefore, in the first non-specific step, Alg1 needs only accept a slightly truncated version of the native substrate, and in the second step, needs only to accept a substrate with a modified C-2 substituent on the residue that is being acted upon.

While this non-specific behavior appears to be significant when compared to reaction with the native substrate, and occurs even at nanomolar concentrations of Alg1 and in the absence of the TRX domain, it is restricted to detection only by the sensitive radiolabeled assay. Such activity could not be detected in the fluorescence-based HPLC

assay, which monitors absolute turnover of the dolichylpyrophosphate-linked substrate. Also, this non-specific behavior was not observed in a yeast microsomal preparation. These results suggest that the membrane may play a key role, positioning the enzyme and dolichyl-linked substrate in the correct position relative to one another, or perhaps complexation with other glycosyltransferases ensures the proper order.

Finally, expression and purification of Alg1 from the native source, *S. cerevisiae*, yielded a stable enzyme preparation with activity comparable to that of detergent solubilized native Alg1 from yeast microsomes. This preparation has proved invaluable as a tool to gain access to the trisaccharide intermediate for studies of Alg2 (Chapter IV) and, ultimately, Alg11 (Chapter V).

Acknowledgements

Experimental Procedures

Media and buffers

Minimal Media (1 L): 6.7 g yeast nitrogen base (Difco) (- amino acids, + salts), 0.6 g

CSL-His-Leu-Trp (Bio101, Inc.), 50 mg His, 100 mg Leu, 100 mg Trp, 2% D-glucose.

Induction Media (1 L): 6.7 g yeast nitrogen base (Difco) (- amino acids, + salts), 0.6 g

CSL-His-Leu-Trp (Bio101, Inc.), 50 mg His, 100 mg Leu, 100 mg Trp, 2% galactose and

1% raffinose.

Buffer A: 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 10 mM imidazole, pH to 8.0 with

NaOH.

Buffer B: 50 mM Tris-Cl pH 7.5, 2.5 mM $MgCl_2$

Buffer C: 50 mM Tris, pH 7, 0.25 M sucrose, 3 mM DTT, 5 mM $MgCl_2$, 0.13 % NP-40.

2× Buffer D: 38 mM Tris-Cl, pH 7.2, 1.8 mM DTT, 0.28 mM EDTA, 0.26% NP-40.

TUP: $CHCl_3$: MeOH: 4 mM $MgCl_2$ (2.75:44:53.25)

Expression and purification of Alg1Δ1-35

Competent BL21(DE3) cells transformed with *ALGΔ1-105* (pET-15b) were

grown overnight in 6 mL of LB culture containing 0.1 mg/mL ampicillin. The overnight

culture was used to inoculate 100 mL of LB containing 0.1 mg/mL ampicillin, and this

culture was shaken at 37 °C until the optical density reached 2.9. Expression was

induced with 1 mM IPTG and allowed to proceed at 30 °C for 4 hr. Cells were then

harvested by centrifugation (20 min, 3836 × g, 4 °C), and the cell pellet stored at –80 °C.

For the purification, cells were incubated on ice in 10 mL of Buffer A containing 1% Triton X-100 and 1 mg/mL lysozyme for 30 min, followed by sonication (6 pulses, 10 sec/ea with 10-sec cooling intervals). Lysed cells were then centrifuged to pellet insoluble debris (30 min, 3.800 × g, 4 °C). The supernatant from this spin was added to 2.5 mL of pre-equilibrated agarose-bound Ni-NTA (Pierce, provided as a 50% slurry) for 1 mL resin slurry/4 mL lysate. This mixture was incubated for 1 hr at 4 °C with rocking to allow binding to occur, then poured into a plastic 1-cm diameter Econo-column (Pierce). Flow-through was collected at a rate of less than 1 mL/min and the resin was washed twice with 10 mL/ea of wash buffer (Buffer A containing 20 mM imidazole and 1% Triton X-100). Protein was then eluted from the column by adding 5 mL of elution buffer (Buffer A containing 250 mM imidazole and 1% Triton X-100) and collecting 1-mL fractions. Fractions containing purified Alg1Δ1-35 were pooled and dialyzed against 50 mM Tris-OAc, pH 7.0, 0.25 M sucrose, 5 mM $MgCl_2$ for 8 hours total with 2 buffer changes. Protein concentration was determined by the micro-BCA assay (Pierce).

Fl-Ahx-Arg-Gly-Trp-Asn(GlcNAc₂)-Glu-Thr-Ile-Val-Glu-Asn-Leu-Lys-NH₂

The synthesis began with Fmoc-Glu-Thr-Ile-Val-Glu-Asn-Leu-Lys-PAL-PEG-PS, which had been prepared previously by Carlos Bosques. The PAL-PEG-PS resin was purchased from PerSeptive Biosystems. The Fmoc-Asn[GlcNAc₂(TBDMS₅)]-OH building block[22] (1.2 equivalents) was activated with the coupling reagent 7-azabenzotriazol-1-yloxytris(pyrrolidino)phosphonium hexafluorophosphate (PyAOP) (2 equivalents), dissolved in 2 mL of CH_2Cl_2, and added to the resin-bound peptide. To initiate the coupling reaction, 4 equivalents of collidine were added and the mixture was

allowed to shake overnight at room temperature. The extent of coupling was analyzed by a small-scale cleavage/deprotection and HPLC/ESI-MS (see Experimental section of Chapter VI for general procedures) revealed approximately 60% conversion by HPLC. Unreacted peptide was capped with acetic anhydride to prevent further elongation. Acetic anhydride (30 equivalents) and collidine (60 equivalents) were added to the resin and allowed to shake for 10 minutes at room temperature. Tryptophan, glycine, and arginine were coupled using standard solid phase peptide synthesis methods (Chapter VI), and the peptide was then capped using Fmoc-aminohexanoic acid (2 equivalents) with PyAOP (2 equivalents) and collidine (4 equivalents) in DMF for two hours. Capping of the N-terminus with fluorescein was accomplished by adding 2 equivalents of fluorescein isothiocyanate to the resin-bound peptide in the presence of 4 equivalents of collidine. The reaction proceeded overnight in the dark at room temperature. These conditions were repeated twice more before nearly complete coupling was achieved. The peptide was released from the resin as described in Chapter VI (General procedures for peptide release and deprotection) using a cocktail of 92% trifluoroacetic acid (TFA), 2% triisopropylsilane (TIS), 3% H_2O, and 3% ethanedithiol (EDT). Following preparative HPLC purification of the glycopeptide, analytical HPLC revealed the presence of an additional more-polar product with a mass of +16 compared to the desired peptide. This modification was found to be reversible with the addition of 10 mM dithiothreitol (DTT). Therefore, the glycopeptide was henceforth stored in 40 mM Mes-NaOH, pH 6.6 containing 10 mM DTT. Quantification of the fluorescent glycopeptide took advantage of the extinction coefficient of fluorescein (78,000 M^{-1} cm^{-1}).

HPLC : t_R=21.8 min (20-60% CH_3CN/0.1%TFA in 30 min)

LRMS calcd for **1** $[C_{108}H_{157}N_{24}O_{35}S]^+$ requires *m/z* 2382.1. Found 1184.5 (+2), 789.7 (+3). (ESI+)

Alg1Δ1-35 glycopeptide assays

To a mixture of 3.0 – 4.5 μM glycopeptide and 3 - 30 μM of GDP-Man in Buffer C containing 0 or 0.25% Triton X-100, 2.8 to 112 μM of Alg1Δ1-35 was added, for a final volume of 100 μL. Reaction mixtures were incubated overnight at 30°C or 37°C. Following the addition of 100 μL of acetonitrile, samples were centrifuged at high speed to remove any insoluble material, and then analyzed on a Waters analytical C_{18} HPLC column. Peptide was eluted with a linear gradient of 20-60% acetonitrile (0.1% TFA) in water (0.1% TFA) over 30 min at a flow rate of 1 mL/min. Peaks were collected, concentrated on a Sc110 Speed-Vac (Savant), and characterized by ESI-MS.

Cloning of full-length *ALG1*

Genomic DNA was extracted from the *S. cerevisiae* strain PRY46 using the Yeast DNA Extraction Reagent (Y-DER) from Pierce, Inc. *ALG1* was amplified from yeast genomic DNA by PCR using the following primers and Vent polymerase for the extension: 5'-CAC CAT GTT TTT GGA AAT TCC T-3' and 5'-TCA ATG AAT TAG CTT CAA-3'. The *ALG1* PCR product was first cloned into pENTR/SD/D-TOPO entry vectors (Invitrogen) using the Gateway System to yield constructs bearing stop codons. This entry vector was then transferred by homologous recombination using LR Clonase enzymes (Invitrogen) to either of the destination vectors, pBAD-DEST49 or pYES-DEST52, for expression in *E. coli* or *S. cerevisiae*, respectively. To enable the addition

of the His$_6$ tag and the V5 epitope as C-terminal tags in the expressed product, the stop codon in the *ALG1* entry vector was later changed to a Gly residue by site-directed mutagenesis using the QuikChange kit (Stratagene), using the following primers, in which the underlined sequence indicates the STOP to Gly mutation.

5'-ATT TGA AGC TAA TTC ATG GAA AGG GTG GGC GCG C-3'

5'-GCG CGC CCA CCC TTT CCA TGA ATT AGC TTC AAA T-3'.

The *ALG1*-containing entry vector was verified by DNA sequencing at the MIT Biopolymers Laboratory using M13 sequencing primers (Invitrogen):

M13 Forward: 5'-GTAAAACGACGGCCAG-3'

M13 Reverse: 5'-CAGGAAACAGCTATGAC-3'

Expression in *E. coli* and purification of TRX-Alg1

One Shot Top 10 competent cells (Invitrogen) were transformed with 1 µL/ea of *ALG1*-containing pBAD-DEST49 *E. coli* expression vector as well as plasmid isolated from BL21(DE3) codon-plus RIL cells, both of which had been prepared from mini-preparation of 3 mL of overgrown culture using the Eppendorf Perfect Prep Plasmid Mini Kit. Since this vector uses an *ara*BAD (P$_{BAD}$) promoter for tightly controlled expression, an *E. coli* strain deficient in the *ara* operon, such as the TOP10 strain, was necessary. Plasmids were selected for based on the carbenicillin resistance of pBAD-DEST49 and chloramphenicol resistance of the codon-plus plasmid. Cells were shaken at 37 °C in 1 L of Terrific Broth (Invitrogen) until growth had reached the stationary phase (A$_{600}$ > 1). Expression was induced with the addition of arabinose to a final concentration of 0.02% (w/v) followed by shaking at 37 °C for two hours. Cells were then harvested by

105

centrifugation (10 min, 15,344 × g, 4 °C), and pellets were stored at −80 °C until purification.

The expressed TRX-Alg1-V5-His tag construct was purified by Ni-NTA affinity chromatography. All steps were performed at 4 °C. Cell pellets from 1-L expressions were resuspended in 40 mL of Buffer A containing 1% Triton X-100, 1 × protease inhibitor cocktail III (Calbiochem), and 1 mg/ mL lysozyme (Sigma). The cell mixture was allowed to incubate on ice for 30 minutes and then lysed by sonication for 2 minutes or until the solution clarified. Cell debris was pelleted by centrifugation (30 min, 11,625 × g, 4 °C). The supernatant from this spin was filtered through a 0.2-μm syringe-filter and was added to Ni-NTA slurry (2 or 4 mL of a 50% slurry, Qiagen). This mixture was rocked at 4 °C for 1 hr, then poured into a 20-mL plastic Econo-column (BioRad). After collecting the flow-through and washing the resin twice with 20 mL of Buffer A containing 20 mM imidazole and 1% Triton X-100, bound protein was eluted using a step-wise gradient of Buffer A containing 40 mM, 80 mM, and 160 mM imidazole, respectively. In the first step, four 5-mL fractions were collected from the 40 mM imidazole buffer, then three 3-mL fractions were collected from the 80 mM imidazole buffer, and ten 1-mL fractions were collected from the 160 mM imidazole buffer.

Fractions containing purified TRX-Alg1 were concentrated by dialyzing pooled fractions against Buffer C, followed by incubation with 0.5 mL of Ni-NTA slurry at 4 °C for 1 hr. Protein was eluted with two 0.5-mL aliquots of Buffer A containing 200 mM imidazole and 1% Triton X-100. Elutions were then dialyzed as above to remove imidazole and protein concentrations for each fraction were determined using the BioRad Protein Assay to be 0.25 and 0.48 mg/mL, for a yield of 0.37 mg from a 1-L expression.

Expression in *S. cerevisiae* and purification of Alg1

For *S. cerevisiae* expression, the yeast strain INV*Sc*1 was made competent and transformed with pYES-DEST52 constructs using the *S. cerevisiae* EasyComp Transformation Kit (Invitrogen) according to the manufacturer's instructions. Starting from glycerol stocks of pYES-DEST52 clones in Inv*Sc*1 *S. cerevisiae* yeast strains, *ALG1*-containing cells were grown in minimal media, starting with 10-mL cultures and diluting to 500 mL of minimal media. The 500-mL cultures were grown to an O.D. of 4-5, and then four equal aliquots were removed, the volumes of which were chosen to enable resuspension of the cells from each aliquot in 1-L batches of induction media to achieve an O.D. of 0.4. Aliquots were pelleted by centrifugation (2,500 × g, 5 minutes, 4°C), the supernatant discarded, and pellets were then resuspended in 3-5 mL of induction media and added to 1-L batches of induction media. Expression cultures were shaken at 30°C for 24 hours, and then harvested by centrifugation (2,500 × g, 5 minutes, 4°C), washing once with cold Buffer B, and stored at –80°C or left on ice in the cold room (for use within 16 hours). All subsequent steps were performed at 4°C. To prepare solubilized microsomes, the cell pellet was resuspended in 200 mL of Buffer B containing the protease inhibitors AEBSF (0.1 mM), pepstatin A (0.5 µg/mL), and leupeptin (0.5 µg/mL), and lysed using glass beads in a bead beater (16 × 20 sec, with 40-sec cooling intervals). Cell debris was pelleted at 7K rpm for 8 minutes. Supernatant was centrifuged again as in the last step to remove residual debris. Supernatant from the second spin was centrifuged at 117734 × g for 1 hour to pellet the crude microsomes. Crude microsomes were homogenized in 50 mL of Buffer A containing 1% NP-40 using

a homogenizer with a Teflon stopper, shaken on ice for 15 minutes, and then centrifuged

for 1 hour 117734 × g. Supernatant was saved as solubilized microsomes, which were

subsequently filtered through a 0.2 μm syringe filter and incubated with Ni-NTA for 1

hour prior to affinity purification. Following incubation, the flow-through was collected

by pouring the mixture into a 20-mL plastic Econo column (BioRad) with the outlet

open, and then the resin was washed twice with 20 mL of Buffer A containing 20 mM

imidazole and 0.1% NP-40. Protein was eluted with a step-wise gradient using 5-mL

aliquots of 40, 120, and 240 mM imidazole in Buffer A containing 0.1% NP-40. Two of

the most concentrated pure protein-containing fractions were pooled and dialyzed against

Buffer C, then brought to 50% glycerol and stored at –20 °C. Protein concentration was

estimated using the BioRad Protein Assay Reagent, using BSA to generate a standard

curve, to be 0.13 mg/mL, for a yield of 0.26 mg of protein from a 4-L yeast expression.

Radioactive activity assays

For each reaction, GlcNAc-PP-Dol or $GlcNAc_2$-PP-Dol (approximately 60 μg)

was added to a 0.5 mL eppendorf tube in $CHCl_3$: MeOH (3:2) and dried under a stream of

nitrogen. Next, 10-20 μL of GDP-(^{3}H)Man (2 μM, approx. 1,000 dpm/μL) was added

and dried under nitrogen. Buffer C containing 1% NP-40 or Triton X-100 was added to

the tube and vortexed well to disperse substrates. (For the metal comparison assay, this

buffer was prepared without $MgCl_2$ and to it was added either no metal, $MgCl_2$ (5 mM

final), or $MnCl_2$ (5 mM final). Reactions were initiated with the addition of enzyme

(amounts as indicated in "Results and Discussion") for a final volume of 100 μL and

proceeded at room temperature. Aliquots of 10 μL were removed at intervals of time and

quenched by adding to 600 µL of CHCl$_3$:MeOH:4 mM MgCl$_2$ (3:2:1) and vortexing. In endpoint assays (for gel filtration), the entire 100-µL reaction mixture was quenched after 30 minutes by adding it to CHCl$_3$:MeOH (3:2) (0.5 mL) and vortexing. Extractions and scintillation counting were completed as described in Chapter II (Enzyme II assay).

Gel filtration of radiolabeled saccharides

Bio-Gel P2 resin (BioRad) (29.3 g) was added to degassed 50 mM acetic acid (200 mL) and allowed to swell overnight. Half of the buffer was poured off, and the resin slurry was transferred to an Ehrlenmeyer flask to be degassed for 10-15 minutes. To remove gel fines, fresh buffer was added and the resin was swirled. When most of the resin had settled, the buffer was poured off. This step was repeated three additional times. All of the gel slurry was poured into a 1-cm diameter glass Econo-column (BioRad) to a height of 117 cm. Once the gel had settled, the column was equilibrated with degassed 50 mM HOAc containing 150 mM NaCl. To determine the exclusion volume of the column, 250 µL of an 8 mg/mL solution of blue dextran in the same buffer was loaded on the column and eluted with the same buffer. Fractions were collected at 44 drops (1.05 mL)/tube, and the blue dextran came out in fractions 34-36.

Bio-Gel P4 resin was prepared in the same way as Bio-Gel P2, except that 18.3 g of resin was used, and the buffer used to remove gel fines also contained 150 mM NaCl. The height of the settled column was 107 cm, and blue dextran (loaded as 250 µL of an 8 mg/ml solution) eluted in fractions 29-32.

Dolichol-linked saccharides to be analyzed by gel filtration were dissolved in 0.5 mL of n-propanol, followed by addition of 0.5 mL of 2 N HCl. This mixture was heated

to 50°C for 15 minutes, and then concentrated under a stream of nitrogen. The remaining acid was neutralized with NaOH and further concentrated to approximately 250 μL before loading the sample onto the column. The products were eluted with 50 mM HOAc containing 150 mM NaCl. The first 30 fractions (31.5 mL) were collected in a graduated cylinder before connecting the column to a fraction collector and collecting 44 more fractions (40 drops, 1.05 mL) in 7-mL scintillation vials. To each of these fractions, 5 mL of EcoLite scintillation cocktail (ICN Biomedicals, Inc.) were added and the vials counted on a scintillation counter (Beckman, Inc.) (5 minutes/vial). Monosaccharide (GlcNAc) and disaccharide (GlcNAc$_2$) standards were run on the column after hydrolyzing UDP-(^{3}H)GlcNAc and (^{3}H)GlcNAc-GlcNAc-PP-Dol, respectively.

Preparation of ManGlcNAc$_2$-PP-Dol with Alg1 expressed in *S. cerevisiae*.

To a dried aliquot of approximately 15-20 nmol of GlcNAc$_2$-PP-Dol, 500 μL of 2× Buffer D was added and sonicated in a water bath sonicator for 1 min to disperse substrate. To this solution was added 10 μL of 1 M MgCl$_2$, 40 μL of 50 mM GDP-Man, 440 μL of dH$_2$O, and 10 μL of Alg1 (2.5 μM). The reaction mixture was incubated at 37 °C for 1 hr, then quenched by addition to 11 mL of CHCl$_3$: MeOH: 4 mM MgCl$_2$ (6:4:1). The dolichylpyrophosphate-linked product partitions into the organic phase, which was washed twice with TUP, then concentrated to dryness, redissolved in 1 mL of CHCl$_3$: MeOH (3:2) and aliquoted 10 μL/tube for assays. Aliquots were stored dry at –80 °C.

Hydrolysis of glycans from dolichylpyrophosphate for HPLC assay

Dried aliquots of samples to be hydrolyzed were resuspended in 500 µL of 0.1 M

trifluoroacetic acid in dH_2O: n-propanol (1:1) and incubated at 50 °C for 15 min.

Samples were then concentrated to dryness on a Speed-Vac (Sc110, Savant, Inc.) in

preparation for labeling.

2-AB labeling of saccharides[25]

R=GlcNAc$_x$Man$_y$
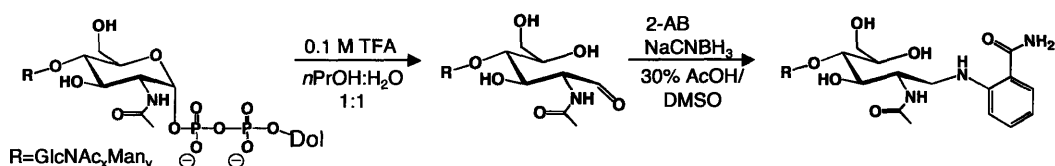
To prepare the dye-labeling reagent, a 150-µL aliquot of acetic acid was added to

a 350-µL aliquot of DMSO, and then a 100-µL aliquot of this mixture was used to

dissolve 5 mg of 2-aminobenzamide (2-AB). The entire dye solution was added to 6 mg

of sodium cyanoborohydride and mixed well to dissolve the reductant. Aliquots of 5 µL

of this reagent were then added to dried samples of hydrolyzed, dried glycans, and heated

to 65°C for 2-4 hours. Post-labeling cleanup was accomplished by using GlykoClean S

cartridges (ProZyme Inc.).

Normal-phase HPLC[26]

The normal phase analytical HPLC column (GlykoSepN) was purchased from

ProZyme, Inc. The solvents used were 50 mM ammonium formate, pH 4.4 (Solvent A)

and acetonitrile (Solvent B). A linear gradient of 20-52% Solvent A in Solvent B over 80

minutes was used to elute the 2AB-labeled glycans. Standardization of the column was done first with commercial standards (ProZyme, Inc.), including a glucose homopolymer and the pentasaccharide core, $Man_3GlcNAc_2$, each labeled with the 2-aminobenzamide (2-AB) fluorophore. Additionally, every dolichyl-linked intermediate from GlcNAc to $Man_5GlcNAc_2$ was prepared as the hydrolyzed and derivatized species, analyzed by HPLC, and confirmed by MALDI-MS. Mono- and disaccharides were hydrolyzed from synthetic preparations, the trisaccharide from the chemoenzymatic Alg1 preparation, and tetra- through heptasaccharides were obtained by elongating the trisaccharide intermediate using microsomes from wild-type yeast (INVSc1, Invitrogen) as well as a temperature-sensitive $alg11$ yeast strain (NDY 13.4). Typically 5-20 pmol of material was loaded for HPLC analysis. Up to 10× more was loaded if material was being collected for MALDI-MS or for mannosidase cleavage analysis.

MALDI-MS

Mass spectrometric analysis was performed on a PE Biosystems Voyager System 4028 using the reflector, positive mode as described.[27] 2, 5-Dihydroxybenzoic acid was used as the matrix. A solution of dihydroxybenzoic acid (DHB) (10 mg) in acetone (0.5 mL) was prepared, and applied to the target as a thin film. The glycans from HPLC fractions were dried on a Sc110 Speed-Vac (Savant, Inc.), then redissolved in 5-10 μL of dH_2O. A 1-μL drop of each glycan sample was spotted onto the target and dried under a slight airstream. A second solution of DHB (10 mg) was prepared by dissolving the matrix in 333 μL of acetonitrile, 167 μL dH_2O, 10 μL 5% perfluorinated Nafion (Sigma-Aldrich), and 0.5 μL TFA. A 0.5-μL drop of this solution was added to each of the dried

glycan spots and dried under a slight airstream. Calibration mixture #1 from the

Sequazyme mass standards kit (Applied Biosystems) was used for standardization.

References

1.   Huffaker, T. C.; Robbins, P. W., Temperature-sensitive yeast mutants deficient in asparagine-linked glycosylation. *J. Biol. Chem.* **1982,** 257, (6), 3203-3210.
2.   Couto, J. R.; Huffaker, T. C.; Robbins, P. W., Cloning and expression in *Escherichia coli* of a yeast mannosyltransferase from the asparagine-linked glycosylation pathway. *J. Biol. Chem.* **1984,** 259, (1), 378-382.
3.   Albright, C. F.; Robbins, P. W., The sequence and transcript heterogeneity of the yeast gene ALG1, an essential mannosyltransferase involved in *N*-glycosylation. *J. Biol. Chem.* **1990,** 265, (12), 7042-7049.
4.   Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.
5.   Abeijon, C.; Hirschberg, C. B., Topography of Glycosylation Reactions in the Endoplasmic-Reticulum. *Trends Biochem. Sci.* **1992,** 17, (1), 32-36.
6.   Bateman, A.; Coin, L.; Durbin, R.; Finn, R. D.; Hollich, V.; Griffiths-Jones, S.; Khanna, A.; Marshall, M.; Moxon, S.; Sonnhammer, E. L. L.; Studholme, D. J.; Yeats, C.; Eddy, S. R., The Pfam protein families database. *Nucleic Acids Res.* **2004,** 32, D138-D141.
7.   Takahashi, T.; Honda, R.; Nishikawa, Y., Cloning of the human cDNA which can complement the defect of the yeast mannosyltransferase I-deficient mutant *alg1*. *Glycobiology* **2000,** 10, (3), 321-327.
8.   Grubenmann, C. E.; Frank, C. G.; Hulsmeier, A. J.; Schollen, E.; Matthijs, G.; Mayatepek, E.; Berger, E. G.; Aebi, M.; Hennet, T., Deficiency of the first mannosylation step in the N-glycosylation pathway causes congenital disorder of glycosylation type Ik. *Hum. Mol. Genet.* **2004,** 13, (5), 535-542.
9.   Schwarz, M.; Thiel, C.; Lubbehusen, J.; Dorland, B.; de Koning, T.; von Figura, K.; Lehle, L.; Korner, C., Deficiency of GDP-Man:GlcNAc$_2$-PP-dolichol mannosyltransferase causes congenital disorder of glycosylation type Ik. *Am. J. Hum. Genet.* **2004,** 74, 472-481.
10.   Kranz, C.; Denecke, J.; Lehle, L.; Sohlbach, K.; Jeske, S.; Meinhardt, F.; Rossi, R.; Gudowius, S.; Marquardt, T., Congenital disorder of glycosylation type Ik (CDGIk): a defect on mannosyltransferase I. *Am. J. Hum. Genet.* **2004,** 74, 545-551.
11.   Janik, A.; Sosnowska, M.; Kruszewska, J.; Krotkiewski, H.; Lehle, L.; Palamarczyk, G., Overexpression of GDP-mannose pyrophosphorylase in *Saccharomyces cerevisiae* corrects defects in dolichol-linked saccharide formation and protein glycosylation. *Biochim. Biophys. Acta* **2003,** 1621, 22-30.

12.   Sharma, C. B.; Lehle, L.; Tanner, W., Solubilization and characterization of the initial enzymes of the dolichol pathway from yeast. *Eur. J. Biochem.* **1982,** 126, 319-325.

13.   Flitsch, S. L.; Goodridge, D. M.; Guilbert, B.; Revers, L.; Webberley, M. C.; Wilson, I. B. H., The chemoenzymatic synthesis of neoglycolipids and lipid-linked oligosaccharides using glycosyltransferases. *Bioorg. Med. Chem.* **1994,** 2, (11), 1243-1250.

14.   Wilson, I. B.; Webberley, M. C.; Revers, L.; Flitsch, S. L., Dolichol is not a necessary moiety for lipid-linked oligosaccharide substrates of the mannosyltransferases involved in in vitro N-linked-oligosaccharide assembly. *Biochem. J.* **1995,** 310 (Pt 3), 909-16.

15.   Revers, L.; Wilson, I. B.; Webberley, M. C.; Flitsch, S. L., The potential dolichol recognition sequence of beta-1,4-mannosyltransferase is not required for enzymic activity using phytanyl-pyrophosphoryl-alpha-N,N'- diacetylchitobioside as acceptor. *Biochem. J.* **1994,** 299 (Pt 1), 23-7.

16.   Watt, G. M.; Revers, L.; Webberley, M. C.; Wilson, I. B. H.; Flisch, S. L., The chemoenzymatic synthesis of the core trisaccharide of *N*-linked oligosaccharides using a recombinant β-mannosyltransferase. *Carb. Res.* **1998,** 305, 533-541.

17.   Watt, G. M.; Revers, L.; Webberley, M. C.; Wilson, I. B. H.; Flitsch, S. L., Efficient enzymatic synthesis of the core trisaccharide of N-glycans with a recombinant beta-mannosyltransferase. *Angew. Chem., Int. Ed. Eng.* **1997,** 36, (21), 2354-2356.

18.   Burda, P.; Jakob, C. A.; Beinhauer, J.; Hegemann, J. H.; Aebi, M., Ordered assembly of the asymmetrically branched lipid-linked oligosaccharide in the endoplasmic reticulum is ensured by the substrate specificity of the individual glycosyltransferases. *Glycobiology* **1999,** 9, (6), 617-625.

19.   Gao, X.-D.; Nishikawa, A.; Dean, N., Physical interactions between the Alg1, Alg2, and Alg11 mannosyltransferases of the endoplasmic reticulum. *Glycobiology* **2004,** 14, (6), 559-570.

20.   LaVallie, E. R.; DiBlasio-Smith, E. A.; Collins-Racie, L. A.; Lu, Z.; McCoy, J. M., Thioredoxin and related proteins as multifunctional fusion tags for soluble expression in E. coli. *Methods Mol. Biol.* **2003,** 205, 119-40.

21.   Southern, J. A.; Young, D. F.; Heaney, F.; Baumgartner, W. K.; Randall, R. E., Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *J. Gen. Virol.* **1991,** 72 (Pt 7), 1551-7.

22.   Bosques, C. J.; Tai, V. W. F.; Imperiali, B., Stereoselective synthesis of beta-linked TBDMS-protected chitobiose-asparagine: a versatile building block for amyloidogenic glycopeptides. *Tetrahedron Lett.* **2001,** 42, (41), 7207-7210.

23.   Terayama, H.; Takahashi, S.; Kuzuhara, H., Large-Scale Preparation of N,N'-Diacetylchitobiose by Enzymatic Degradation of Chitin and Its Chemical Modifications. *J. Carbohydr. Chem.* **1993,** 12, (1), 81-93.

24.   Lee, J.; Coward, J. K., Enzyme-catalyzed glycosylation of peptides using a synthetic lipid disaccharide substrate. *J. Org. Chem.* **1992,** 57, 4126-4135.

25.   Bigge, J. C.; Patel, T. P.; Bruce, J. A.; Goulding, P. N.; Charles, S. M.; Parekh, R. B., Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal. Biochem.* **1995,** 230, (2), 229-38.

26.   Guile, G. R.; Rudd, P. M.; Wing, D. R.; Prime, S. B.; Dwek, R. A., A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. *Anal. Biochem.* **1996,** 240, (2), 210-26.

27.   Hanbke, B.; Thiel, C.; Lübke, T.; Hasilik, M.; Höning, S.; Peters, V.; Heidemann, P. H.; Hoffmann, G. F.; Berger, E. G.; von Figura, K.; Korner, C., Deficiency of UDP-galactose: *N*-acetylglucosamine β-1,4-galactosyltransferase I causes the congenital disorder of glycosylation type IId. *J. Clin. Invest.* **2002,** 109, 725-733.

Chapter IV: Studies toward defining the function of Alg2

Introduction

Although *ALG2* mutants have been identified and characterized in *Saccharomyces cerevisiae*,[1] the zygomycete fungus *Rhizomucor pusillus*,[2] and humans,[3] the results have not been entirely consistent. It has been suggested that Alg2 is responsible for either the second, third, or both the second and third mannosylation steps of the dolichol pathway (Scheme IV-1).[1-5]
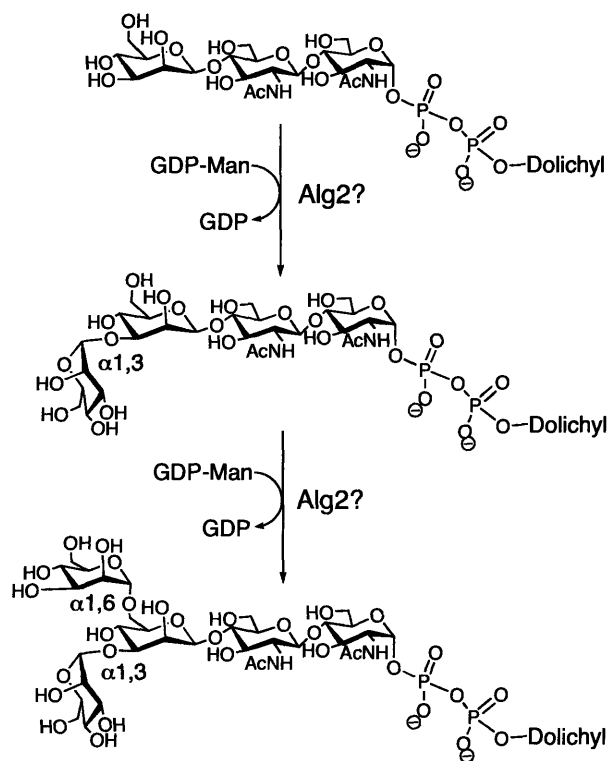
Scheme IV-1. α1,3- and α1,6-Mannosylation reactions putatively catalyzed by Alg2.

The *alg2-1* yeast mutant was isolated by the mannose suicide selection screen described in Chapter III, and was characterized by a temperature-sensitive phenotype, in which an accumulation of both ManGlcNAc$_2$-PP-Dol and Man$_2$GlcNAc$_2$-PP-Dol was

detected.[1] The genetic basis for this phenotype was later mapped to two point mutations, G377R and Q386K. By complementing *alg2-1* with the two corresponding single mutants of the *R. pusillus ALG2* homolog, one of the mutations (G377R in the yeast Alg2) was found to be sufficient to support the phenotype (Figure IV-1).[5] From these results, which must take into account that temperature-sensitive mutants can retain low levels of activity even at the non-permissive temperature, it could only be concluded that Alg2 is a mannosyltransferase that acts early on the dolichol pathway.
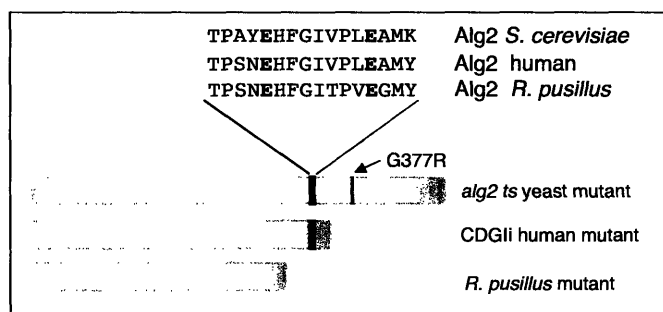
Figure IV-1. Summary of Alg2 mutants characterized to date, highlighting the position of the conserved EX₇E motif, which is shown for each organism with the Glu residues in bold.

Interestingly, *ALG2* in *R. pusillus* is not an essential gene, and when a natural mutant was isolated with nearly the entire C-terminal half of the sequence deleted (Figure IV-1), the only accumulated intermediate was ManGlcNAc$_2$-PP-Dol.[2] This result suggested that Alg2 is responsible for the second mannosylation step, although it does not address any possible involvement of Alg2 in the third mannosylation.

A newly discovered congenital disorder of glycosylation (CDGIi) resulting from mutated human *ALG2* has provided an additional example to consider.[3] In this case, the mutation causes a frame shift that introduces a premature stop codon, leading to a significant deletion of the C-terminal portion of the protein, but not the loss of a conserved EX_7E motif (Figure IV-1). By metabolically labeling CDGIi patient fibroblasts with radiolabeled mannose, a phenotype similar to that found in the yeast mutant was revealed. Specifically, accumulation of both $ManGlcNAc_2$-PP-Dol and $Man_2GlcNAc_2$-PP-Dol was observed.

From these results it was concluded that Alg2 carries out only the second mannosylation to form $Man\alpha1,3$-$ManGlcNAc_2$-PP-Dol. The tetrasaccharide product formed in the patient fibroblasts was proposed to be the non-native $Man\alpha1,6$-$ManGlcNAc_2$-PP-Dol, as opposed to the native intermediate, $Man\alpha1,3$-$ManGlcNAc_2$-PP-Dol, and that the non-specific step to get to this non-native product was carried out by the unidentified mannosyltransferase following Alg2. In the absence of the $\alpha1,3$-linked mannose residue, the non-native intermediate would fail to be further elongated (Figure IV-2). However, no structural characterization of this tetrasaccharide was reported to support the hypothesis that the accumulated tetrasaccharide was the $\alpha1,6$-mannosylated tetrasaccharide. Taken together, the genetic data are consistent with a role for Alg2 as a mannosyltransferase early in the pathway, however the precise function of the enzyme required *in vitro* biochemical validation.
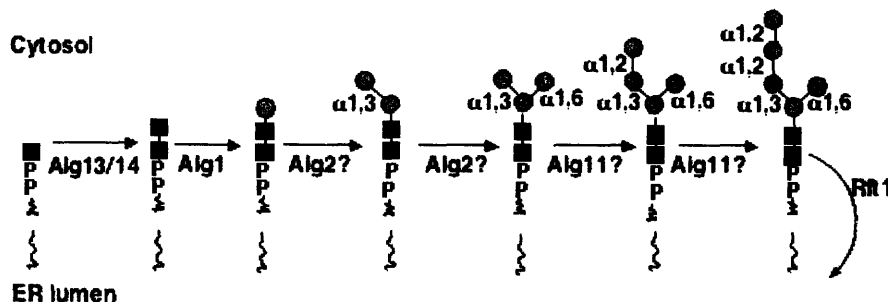
Figure IV-2. Sequence of glycosylation steps on the cytosolic face of the ER membrane. GlcNAc residues are depicted as black squares, Man residues as gray circles, phosphate groups as **P**, and dolichyl moieties as wavy lines.

Alg2 from *S. cerevisiae* is a 503-residue, 58-kDa mannosyltransferase containing four predicted transmembrane domains, as predicted by the TMHMM server, v. 2.0 (Figure IV-3).[6] These domains are grouped into two pairs of closely-spaced domains, with one pair at the extreme C-terminus, and the other pair located near the N-terminus. Our goal for the studies described in this chapter was to find an expression system for active Alg2 that would enable an *in vitro* assay with defined Dol-PP-linked substrates to unambiguously define the function of this enzyme. Alg2 mutants were also used to probe the importance of the Glu residues within the conserved EX_7E motif.
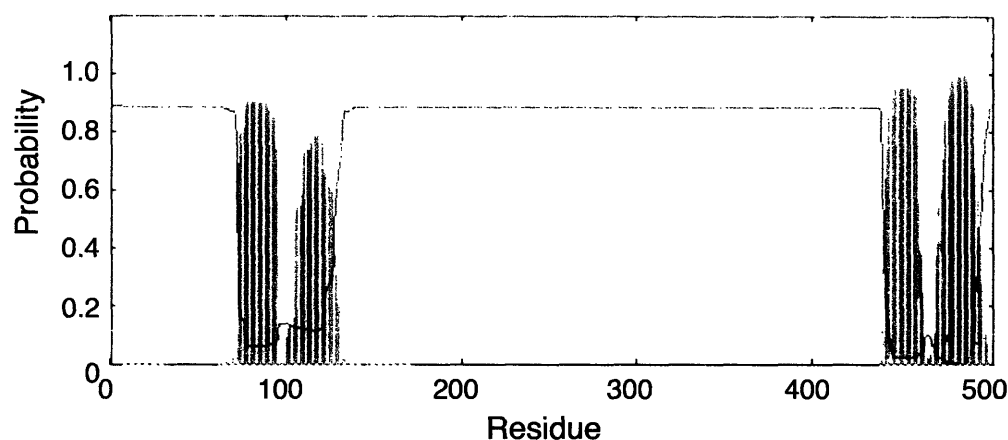
Figure IV-3. Predicted topology of Alg2. Plot of the probability that Alg2 residues are in a transmembrane-spanning region.

Results and Discussion

Cloning and expression of TRX-Alg2-V5-His in *E. coli*

Initially, *ALG2* was amplified from *Saccharomyces cerevisiae* genomic DNA and cloned into the pET-15b vector (Novagen) for a construct encoding an N-terminal His tag. However, no significant expression of this construct was achieved, even using cells that provide extra copies of tRNA for certain codons to compensate for differences in codon usage between *E. coli* and *S. cerevisiae*. To improve expression, *ALG2* was cloned into the pBAD-DEST49 expression vector for *E. coli* expression of a construct that encodes an N-terminal thioredoxin (TRX) tag. This construct also included a C-terminal V5 epitope for detection and C-terminal His tag for purification. Cloning and expression was carried out using the Gateway® Cloning Technology (Invitrogen), as described for Alg1 in Chapter III. Cells were lysed in the presence of detergent, and the clarified lysate

120

was purified by Ni-NTA chromatography to obtain the TRX-Alg2 fusion protein with a predicted molecular weight of 70 kDa as shown in Figure IV-4.
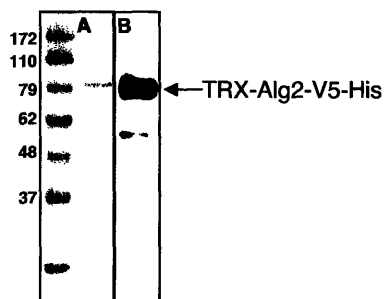
Figure IV-4. Purified TRX-Alg2-V5-His from *E. coli* expression. Molecular weight markers are labeled on the left in kDa. A) Coomassie stain B) Western blot analysis with α-Thio antibody (Invitrogen) for detection of the TRX domain.

TRX-Alg2-V5-His mannosyltransferase activity assays

HPLC assays were carried out on this TRX-Alg2-V5-His preparation, as described in Chapter III, whereby the organic-soluble extracts of reaction mixtures were hydrolyzed and labeled with 2-aminobenzamide prior to HPLC analysis on a normal-phase column with fluorescence detection (λ_{Ex} = 330 nm, λ_{Em} = 420 nm). Assays were carried out with ManGlcNAc$_2$-PP-Dol prepared using Alg1, as described in Chapter III, or with GlcNAc$_2$-PP-Dol in combination with Alg1. Conditions varied in detergent and phospholipid composition and concentration, as well as temperature, GDP-Man concentration, and time. In all cases, however, no activity was detected under any of the conditions tested. Phospholipids used included phosphatidylcholine and phosphatidylethanolamine, and detergents tested included the nonionic detergents Triton

X-100, Nonidet P-40, TWEEN 20, and dodecylmaltoside, as well as steroid-based zwitterionic detergents such as CHAPS and CHAPSO. Detergent concentrations tested ranged from 0.02% to 1% (w/v) or (v/v), but were used in concentrations above the critical micelle concentration (CMC). While the CMC values of the dolichol substrates are not known, Alg1 activity had been observed to be sensitive to high concentrations of GlcNAc$_2$-PP-Dol, and therefore a wide range of ManGlcNAc$_2$-PP-Dol concentrations were tested with Alg2 to avoid the possibility that the concentrations used were inhibitory.

There are numerous possible reasons for the failure to demonstrate activity. ManGlcNAc$_2$-PP-Dol may not be the native substrate of Alg2, an accessory protein may be necessary for activity, or Alg2 itself may only be an accessory protein. Therefore, a separate *in vitro* assay was used whereby expressed Alg2 was added to microsomes prepared from the *alg2* yeast strain, and complementation of the *alg2* mutant defect was tested. This complementation assay would be expected to report Alg2 activity by elongation of the GlcNAc$_2$-PP-Dol substrate to the heptasaccharide intermediate. No further elongation would be expected since the subsequent mannosyltransferases require Man-P-Dol as the glycosyl donor, which was not added. The advantage of this assay is that it does not require prior knowledge of the native Alg2 substrate, since the only component missing is active Alg2. Even if Alg2 were only an accessory protein, addition of an intact Alg2 preparation should complement the defect.

As shown by the HPLC trace in Figure IV-5, the microsomal preparation from *alg2* mutant yeast elongates GlcNAc$_2$-PP-Dol to ManGlcNAc$_2$-PP-Dol in the presence of GDP-Man, with no evidence for the accumulation of larger intermediates. Unfortunately,

122

all of the Alg2 preparations tested were unable to complement the defect, though many of

the same strategies used in the original Alg2 assays to optimize the conditions were also

employed in these complementation assays.
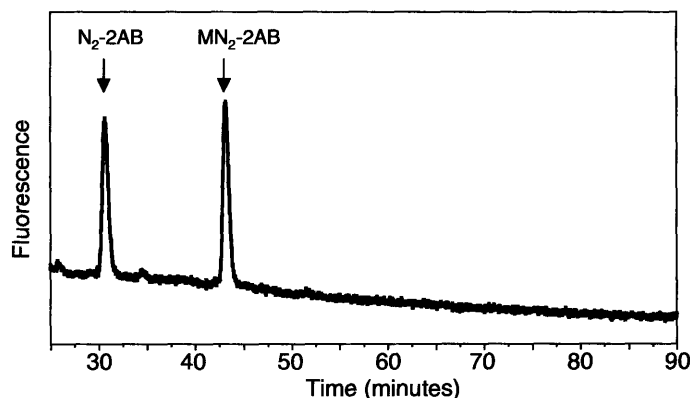
Figure IV-5. Elongation of GlcNAc$_2$-PP-Dol with *alg2* microsomes used in the

complementation assay for Alg2 activity. A product with the retention time of authentic

ManGlcNAc$_2$-2AB appears. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

Cloning and expression of Strep-Alg2-V5-His in *S. cerevisiae*

Since expression of Alg1-V5-His in *S. cerevisiae* yielded a preparation with

improved activity, a similar approach was taken for Alg2. The Alg2 construct, which

was prepared from the same vector used for Alg1 expression in yeast (Chapter III), also

included the C-terminal V5 epitope and His tag. At the N-terminus, a Strep tag (Qiagen)

was included, which is a short peptide that binds with micromolar affinity to an

engineered Streptavidin (Qiagen). This tag was incorporated into the construct by using a

PCR primer encoding the tag in the cloning step. The use of this N-terminal tag was

chosen to allow an additional purification step following Ni-NTA chromatography. For

Alg2 expression in the native host, *S. cerevisiae*, purity is crucial to preclude the detection of activity from other dolichol pathway enzymes, which would complicate the interpretation of the results. The *ALG2* construct was under the control of a galactose-inducible promoter, and expression was alternatively repressed by glucose. Therefore, for tightly regulated expression, cells were grown using glucose as the primary carbon source, and then expression was induced by exchanging the growth media with galactose-containing induction media. Following expression, crude microsomes were prepared by fractionation of the yeast cell lysate and then detergent-solubilized. Affinity purification with Ni-NTA followed by resin-bound Streptactin in detergent solution yielded the preparation shown in Figure IV-6.
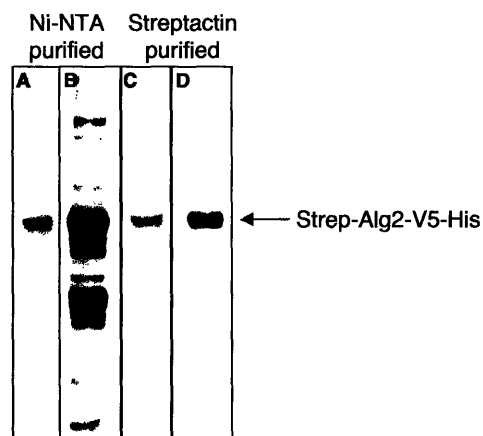
Figure IV-6. Purified Strep-Alg2-V5-His from *S. cerevisiae* expression. A) Coomassie stain and B) Western blot analysis of Ni-NTA purified Strep-Alg2-V5-His. C) Coomassie stain and D) Western blot analysis of sequential Ni-NTA and Streptactin-purified Strep-Alg2-V5-His. Western blots were probed with α-V5 antibody (Invitrogen).

124

Strep-Alg2-V5-His mannosyltransferase activity assays

Enzyme activity of crude microsomes from the Alg2 purification was tested with ManGlcNAc$_2$-PP-Dol, which was prepared using Alg1 as described in Chapter III. The HPLC assay was used to detect elongation of this intermediate in the presence of GDP-Man and microsomes. Indeed, activity of all of the cytosolic mannosyltransferases following Alg1 could be detected at this stage based on the elongation of ManGlcNAc$_2$-PP-Dol to Man$_3$GlcNAc$_2$-PP-Dol, with an observed accumulation of Man$_3$GlcNAc$_2$-PP-Dol (Figure IV-7). In contrast, no elongation was observed under the same conditions using microsomes from yeast that was not exposed to galactose, and thus did not express the *ALG2* construct.
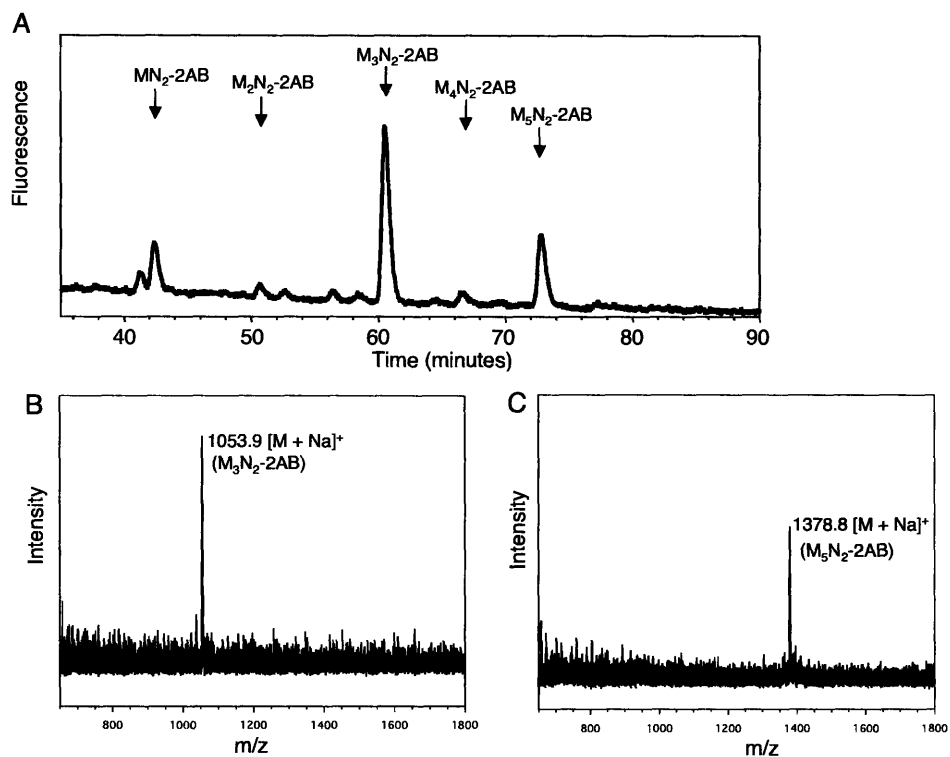
Figure IV-7. Elongation by crude microsomes from Strep-Alg2 preparation A) Fluorescence HPLC trace showing elongation products of ManGlcNAc$_2$-PP-Dol with microsomes from the Strep-Alg2-V5-His expression. B) MALDI-MS of pentasaccharide product corresponding to Man$_3$GlcNAc$_2$-2AB. C) MALDI-MS of heptasaccharide product corresponding to Man$_5$GlcNAc$_2$-2AB. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

Upon purification of Alg2, however, no activity was detected under any of the assay conditions described above. It should be noted that due to the presence of native dolichol pathway enzymes in microsomal preparations, the observed activity in the induced preparation could not be ruled out as native Alg2 rather than expressed Strep-

126

Alg2-V5-His. It also cannot be ruled out that the lack of observed activity in the uninduced microsomes may be due to the alternate growth conditions of the yeast (ie. glucose rather than galactose as the main carbon source), or loss of activity in the preparation.

E. coli membrane fraction preparation of TRX-Alg2

Since numerous attempts at purification of a variety of Alg2 constructs consistently yielded inactive forms of the enzyme, it was postulated that the two remote regions of predicted transmembrane domain pairs (Figure IV-3) may require the membrane to maintain the proper structure and function of the soluble portion of the enzyme. One approach to address the problem of membrane requirement is through the use of proteoliposomes, in which proteins are reconstituted into artificial phospholipid bilayers. This method relies on the proper refolding of proteins upon reintroduction into the membrane. Another approach is to isolate the protein in native intact membranes. For our purposes, this method would require the use of a heterologous system, to avoid the previously mentioned problem of background activity in assigning enzyme function. Isolation of Alg2 in the *E. coli* cell membrane fraction was thus used as a way to preserve the membrane environment of Alg2 while offering a heterologous system that could enable the assignment of function. *E. coli* provides an ideal platform, as dolichol pathway homologs do not exist in this organism.

TRX-Alg2 was expressed in *E. coli* as described above, except that by excluding detergent from the lysis buffer, and using differential centrifugation, this construct could be isolated in the cell membrane fraction. Since this procedure does not include affinity

purification, the construct was simplified by incorporating a stop codon at the 3' end of

the *ALG2* sequence to eliminate the C-terminal tags. In order to assess the requirement of

the N-terminal TRX domain for expression, expression and isolation of Alg2 constructs

either with or without inclusion of the N-terminal TRX domain were carried out. For this

experiment, a His-Alg2 construct and the TRX-Alg2-V5-His construct (without the stop

codon) were used in order to directly compare expression levels using the α-His tag

antibody. Figure IV-8 shows that TRX-Alg2-V5-His was successfully expressed and

isolated in the cell membrane fraction (Lane 3), as detected by an antibody to the His tag.

However, in the absence of the TRX domain (Lane 4), no full-length Alg2 was detected,

and only a smaller product, presumably a truncation product, was observed.
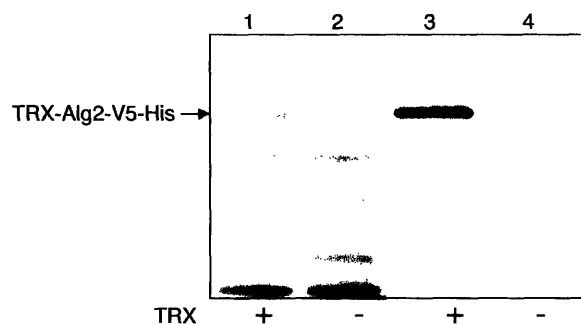
Figure IV-8. Isolated TRX-Alg2 and His-Alg2 from *E. coli* expression. Equivalent

amounts of cell membrane fraction were used for SDS-PAGE and Western blot analysis.

Gelcode protein stain (Pierce Co.) of TRX-Alg2-V5-His (lane 1) and His-Alg2 (lane 2),

Western blot analysis of TRX-Alg2-V5-His (lane 3) and His-Alg2 (lane 4). The Western

blot was visualized using α-His tag antibody (GE Healthcare Bio-Sciences Co.).

TRX-Alg2 elongates ManGlcNAc$_2$-PP-Dol to Man$_3$GlcNAc$_2$-PP-Dol

The membrane fraction containing TRX-Alg2 was assayed for activity with ManGlcNAc$_2$-PP-Dol in the presence of GDP-Man. Incubation with these substrates led to the appearance of two new peaks in the fluorescence HPLC trace, corresponding to the retention times of authentic Manα1,3-ManGlcNAc$_2$-2AB (minor) and Manα1,3-(Manα1,6)-ManGlcNAc$_2$-2AB (major) (Figure IV-9A). These standards represent the tetrasaccharide and pentasaccharide dolichol pathway intermediates, respectively.
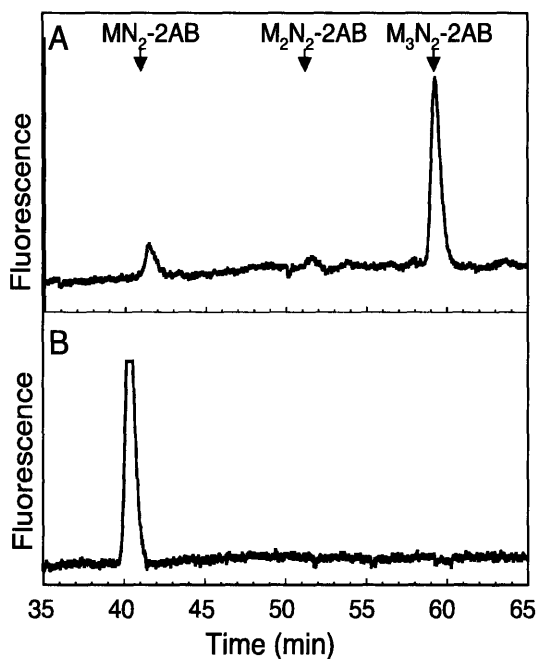
Figure IV-9. HPLC of TRX-Alg2 products. Fluorescence HPLC traces show labeled products from reaction mixtures containing ManGlcNAc$_2$-PP-Dol and cell membrane fraction containing A) TRX-Alg2 or B) TRX-Alg11. The retention times of the TRX-Alg2 products (A) match those of Man$_2$GlcNAc$_2$-2AB and Man$_3$GlcNAc$_2$-2AB authentic standards. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

129

In order to preclude the possibility that the products observed were the result of background *E. coli* enzymes, the experiment was repeated in a manner that is similar to an empty vector control. Instead of removing the gene, however, *ALG2* was replaced with *ALG11*, which encodes a later mannosyltransferase in the pathway and would not be expected to accept the trisaccharide intermediate. Indeed, a membrane fraction containing TRX-Alg11 did not show any reactivity with the trisaccharide intermediate (Figure IV-9B).

Subjecting these products to analysis by MALDI-MS confirmed the identities as the tetrasaccharide, $Man_2GlcNAc_2$-2AB (expected $[M+Na]^+ = 891.3$ and found $[M+Na]^+ = 891.4$), and the pentasaccharide, $Man_3GlcNAc_2$-2AB (expected $[M+Na]^+ = 1053.4$ and found $[M+Na]^+ = 1053.3$), suggesting that Alg2 catalyzes the addition of two mannose residues (Figure IV-10).
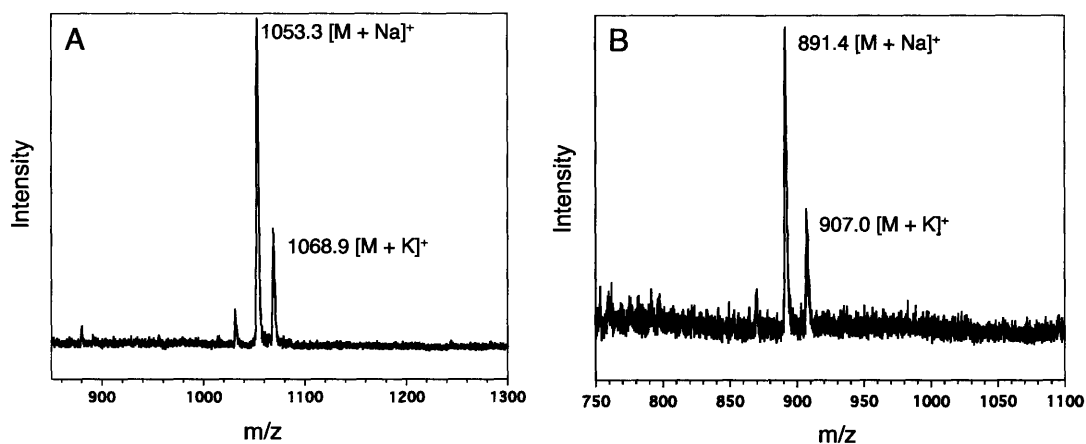
Figure IV-10. MALDI-MS of TRX-Alg2 products. A) Pentasaccharide product (major) of TRX-Alg2, corresponding to $Man_3GlcNAc_2$-2AB B) Tetrasaccharide product (minor) of TRX-Alg2, corresponding to $Man_2GlcNAc_2$-2AB.

To confirm that these saccharides have the same structures as the dolichol pathway intermediates, the glycosidic linkages were mapped by specific mannosidase treatment using an α1-2,3-mannosidase (*Xanthomonas manihotus*, New England Biolabs) and an α1,6-mannosidase (*Xanthomonas* sp., Calbiochem).[7] Treatment of the pentasaccharide with the α1-2,3-mannosidase yielded a tetrasaccharide, which was sensitive to an α1,6-mannosidase (Figure IV-11A-C). These cleavage products were also confirmed by MALDI-MS (Figure IV-11D,E).
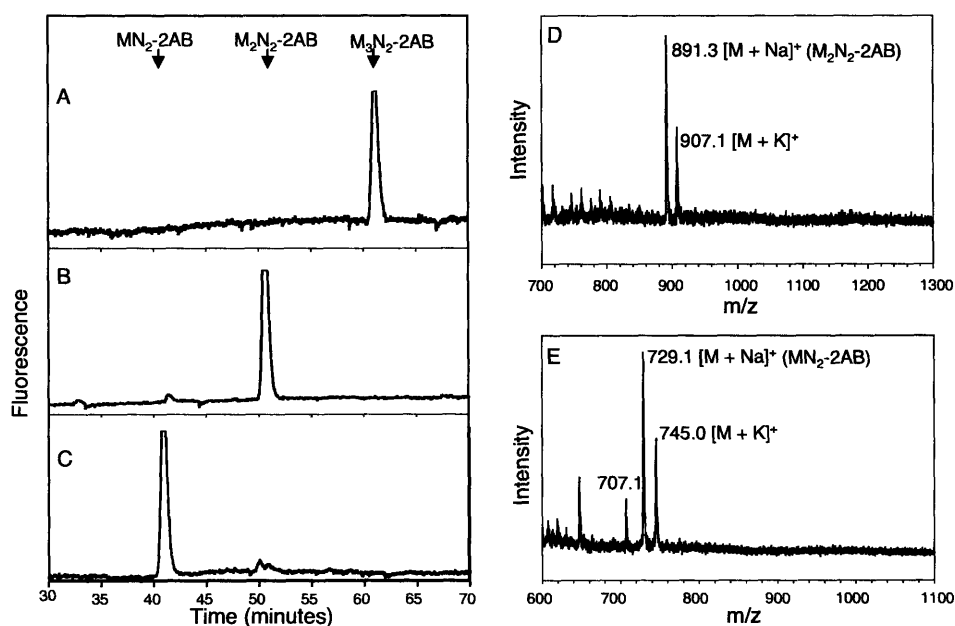
Figure IV-11. Mannosidase mapping of $Man_3GlcNAc_2$-2AB from TRX-Alg2. Fluorescence HPLC traces of A) the pentasaccharide product of TRX-Alg2, $Man_3GlcNAc_2$-2AB, B) the product from α1-2,3-mannosidase treatment of $Man_3GlcNAc_2$-2AB, and C) the product from α1,6-mannosidase treatment of the tetrasaccharide cleavage product of $Man_3GlcNAc_2$-2AB. MALDI-MS results of D) the tetrasaccharide cleavage product (from Panel B) and E) the trisaccharide cleavage product (from Panel C). (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide).

The tetrasaccharide intermediate formed in trace amounts in the TRX-Alg2 reaction was shown to be sensitive to an α1-2,3-mannosidase to yield a trisaccharide (Figure IV-12).
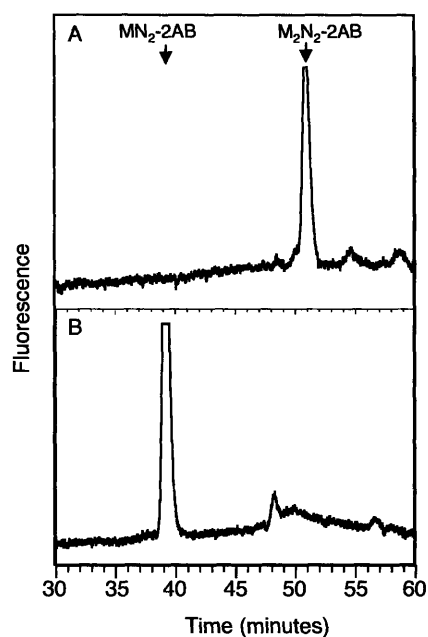
Figure IV-12. Mannosidase mapping of $Man_2GlcNAc_2$-2AB from TRX-Alg2. Fluorescence HPLC traces of A) the tetrasaccharide intermediate from TRX -Alg2 reaction, $Man_2GlcNAc_2$-2AB and B) the cleavage product from treatment of the tetrasaccharide with α1-2,3-mannosidase. (M, Man; N, GlcNAc; 2-AB, 2-aminobenzamide).

These glycoside linkage mapping results are consistent with previous biochemical analysis of accumulated dolichylpyrophosphate-linked intermediates in wild-type and mutant yeast strains, which suggested that the order of addition of mannose residues proceeds by α1,3-mannosylation of the trisaccharide intermediate, followed by α1,6-

132

mannosylation of the β1,4-linked mannose residue to yield the branched core pentasaccharide product.[3, 8]

The pentasaccharide, $Man_3GlcNAc_2$-2AB, was resistant to the α1,6-mannosidase (Figure IV-13), which is consistent with the documented inability of this mannosidase to cleave branched structures.[7] This result supports the branched structure of the pentasaccharide, in accord with the natural pentasaccharide intermediate.
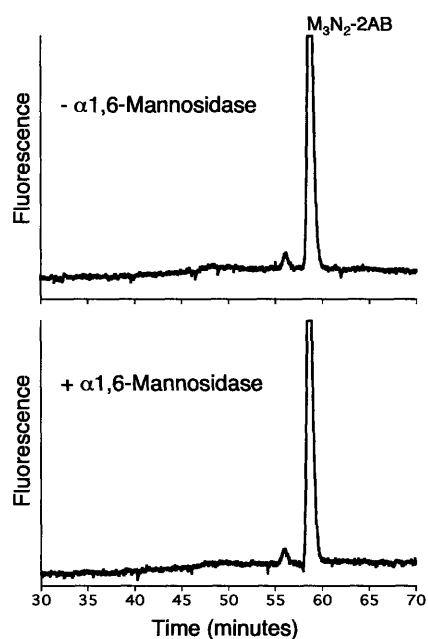
Figure IV-13. Treatment of the pentasaccharide product with α1,6-mannosidase. Equal portions of $Man_3GlcNAc_2$-2AB were incubated overnight in the presence or absence of the α1,6-mannosidase, as described in "Experimental Procedures", except that incubation proceeded for 16 hours.

The only other possible structure based on the pentasaccharide linkage analysis would be a linear structure in which the α1,6-linkage is closer to the reducing end and the terminal mannose is α1,3-(or α1,2-)linked. However, the demonstration that the

133

tetrasaccharide intermediate that bears a terminal α1,3-(or α1,2-) linkage negates this possibility. In addition, while the α1-2,3-mannosidase does not necessarily distinguish between α1,2- and α1,3-linkages, prior knowledge of the structure of the pentasaccharide intermediate in the dolichol pathway as well as the established role of Alg2 in the early steps of the tetradecasaccharide biosynthesis support the presence of an α1,3-linked mannose and a branching of the α1,6-linked mannose from the β1,4-linked mannose. Taken together, these results suggest that TRX-Alg2 catalyzes the mannosylation of Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol to produce Manα1,3-(Manα1,6)-Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol *via* the intermediate, Manα1,3-Manβ1,4-GlcNAcβ1,4-GlcNAc-PP-Dol.

In order to provide evidence for enzyme dependence as well as to quantitate substrate turnover, it is necessary to measure enzyme activity over a range of TRX-Alg2 concentrations. These types of experiments have proven difficult for a number of reasons. Typically, the cell membrane fraction containing TRX-Alg2 was used in assays at 2-5% (v/v) with respect to the detergent-containing assay buffer. At these concentrations, complete turnover of the dolichylpyrophosphate-linked substrate was observed after one hour. In contrast, when 0.2 % (v/v) cell membrane fraction was used, no product formation was observed. This dramatic difference in product formation may not reflect the difference in enzyme concentration, but rather the dilution of stabilizing phospholipids present in the membrane fraction. In assays containing 50% (v/v) cell membrane fraction, product formation corresponded to less than 50% conversion. Therefore, the mannosyltransferase activity also appears to be sensitive to higher concentrations of the membrane fraction in the assay. These results suggest that there is

an additional parameter affecting activity, specifically the concentration of the membrane fraction, which complicates any attempt to correlate TRX-Alg2 concentration with activity. Possible ways of addressing this limitation include titration of the membrane fraction containing TRX-Alg2 into a blank membrane fraction, for instance from uninduced cells, or from the TRX-Alg2 double mutant. However, it would be difficult to assess how well the composition of the blank membranes matches that of the enzyme-containing membranes. In other words, the ability of blank membranes to stabilize TRX-Alg2 to the same extent, or conversely, the extent of the apparent inhibitory behavior of the membranes at higher concentrations, may not be comparable. Therefore, the studies described within were designed to include a variety of controls to relate the observed product formation specifically with the presence of intact TRX-Alg2.

Mannosyltransferase activity of site-directed TRX-Alg2 mutants

The single site-directed mutants, TRX-Alg2 E335A and TRX-Alg2 E343A, as well as the double mutant, TRX-Alg2 E335A/E343A were constructed to determine the importance of the Glu residues in the conserved EX_7E motif (Figure IV-1), and also as further controls to support the specific role of Alg2 in the observed activity. The relative expression levels of each Alg2 mutant were compared to that of the wild-type by analyzing equivalent amounts of each cell membrane fraction by Western blot analysis, as shown in Figure IV-14A. It has been proposed that the first Glu residue of this motif may be involved in catalysis.[9] In the case of one member of this family, mutation of the first Glu residue of the EX_7E motif resulted in loss of all detectable activity, while residual activity could be seen upon mutation of the second Glu residue.[10] Thus, the

TRX-Alg2 E335A mutant was expected to affect activity, and, indeed, a significantly

lower level of product formation was observed (Figure IV-14B) compared to that seen for

wild-type TRX-Alg2 (Figure IV-9). More notable, however, is that the tetrasaccharide

intermediate formed ($[M+Na]^+ = 890.9$, $[M+K]^+ = 906.7$) (Figure IV-15A) is present in

greater proportion than the pentasaccharide product, suggesting that perhaps mutation of

the first Glu residue impairs the second step more severely than the first.
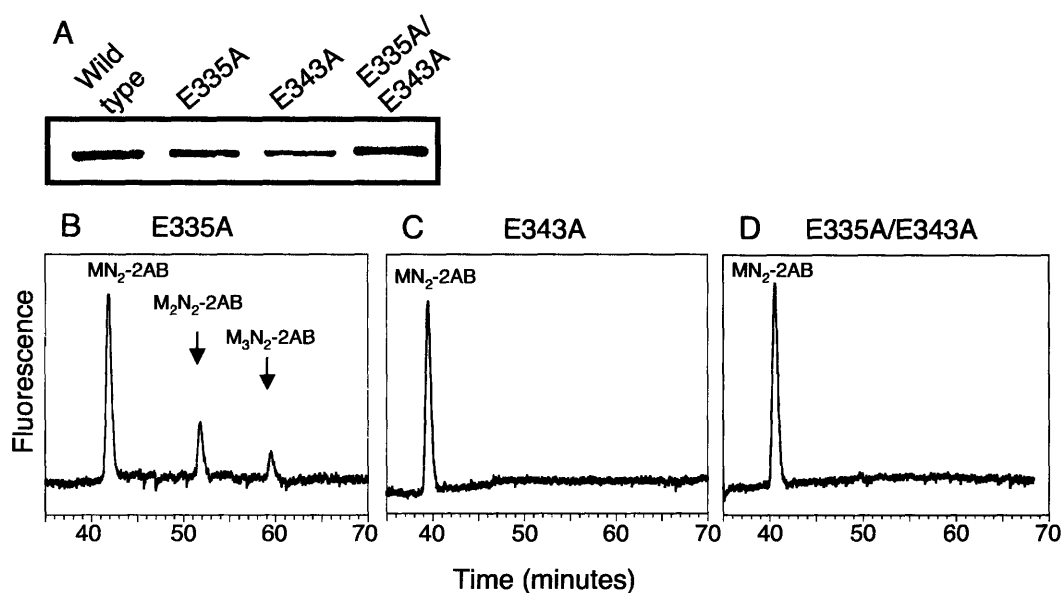
Figure IV-14. Analysis of expression and activity of TRX-Alg2 mutants. A) Western

blot analysis showing relative amounts of wild type and mutants of TRX-Alg2 in

equivalent amounts of cell membrane fraction from expressions. Fluorescence HPLC

traces of elongation assays with B) TRX-Alg E335A, C) E343A, or D) the double

mutant, E335A/E343A. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

To ensure that the tetrasaccharide intermediate formed by the TRX-Alg2 E335A

mutant is the natural intermediate in the pathway, and that the mutation did not cause the

α1,6-mannosylation to occur first, the Man$_2$GlcNAc$_2$-2AB was shown to be sensitive to

the α1-2,3-mannosidase and resistant to the α1,6-mannosidase (Figure IV-15B,C).
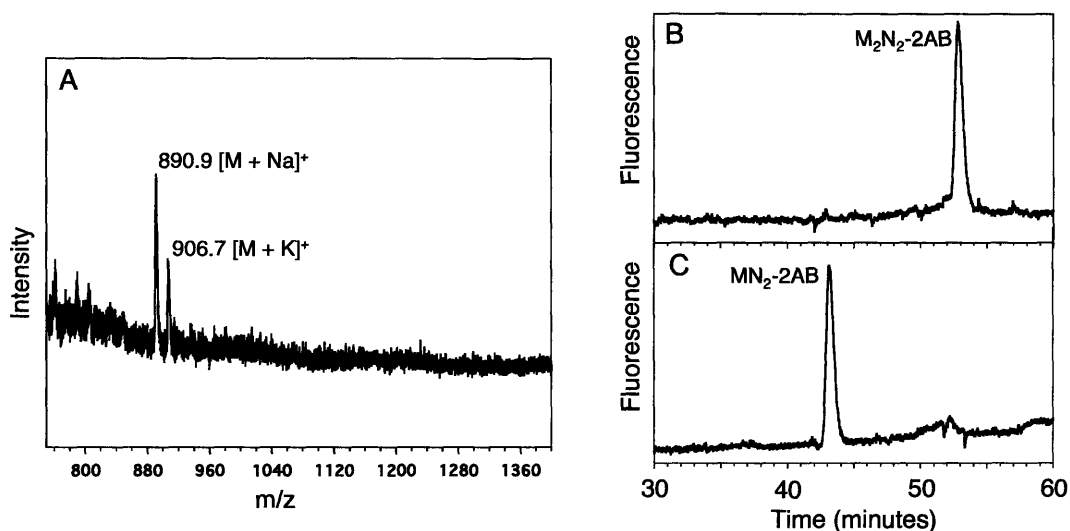
Figure IV-15. Characterization of Man$_2$GlcNAc$_2$-2AB from TRX-Alg2 E335A. A) MALDI-MS of tetrasaccharide product of TRX-Alg2 E335A, corresponding to Man$_2$GlcNAc$_2$-2AB. Fluorescence HPLC traces show that this product is B) resistant to α1,6-mannosidase treatment and C) sensitive to α1-2,3-mannosidase treatment. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

The second Glu mutant, E343A, did not show any detectable activity (Figure IV-14C). From this assay, it is impossible to know whether this mutation affects both steps, or only the first mannosylation. In an attempt to address this question, the E335A and E343A single mutants were assayed together. If the latter mutant was able to carry out the α1,6-mannosylation, then the combination of these mutants should yield the pentasaccharide product. However, the only activity observed was that of the E335A

mutant (data not shown). As expected, based on the inactivity of the second single mutant, the double mutant does not exhibit any detectable level of activity (Figure IV-14D). These mutants, in addition to providing preliminary insight into the possible role of this motif, also offer additional controls to confirm that the observed activity can indeed be attributed to TRX-Alg2, and not to background enzymes in the bacterial membrane fraction.

Specificity of Alg1 and Alg2 for the dolichyl moiety

In vitro studies of Alg1 have revealed that the dolichyl moiety can be replaced with the C20 branched alkyl chain, phytanyl (Figure III-2) with very little loss in efficiency of the enzyme.[11] Bacteria such as *E. coli* use a fully-unsaturated and truncated (C55) version of dolichylpyrophosphate known as undecaprenylpyrophosphate (Und-PP) (Figure IV-16) as a carrier in cell wall biosynthesis.

Undecaprenylpyrophosphate (Und-PP)   Dolichylpyrophosphate (Dol-PP)

Figure IV-16. Comparison of the structures of Und-PP and Dol-PP.

In order to test the ability of Alg1 and Alg2 to accept substrates in which the dolichyl group has been replaced by an undecaprenyl group, $GlcNAc_2$-PP-Und was assayed with these enzymes in "one pot" under the same conditions used for the native

substrate. As shown in Figure IV-17, the disaccharide substrate was elongated to the pentasaccharide intermediate, with accumulation of the trisaccharide and tetrasaccharide intermediates. While there is a preference for Dol-PP, both Alg1 and Alg2 are capable of accepting the Und-PP derivatives of the native substrates.
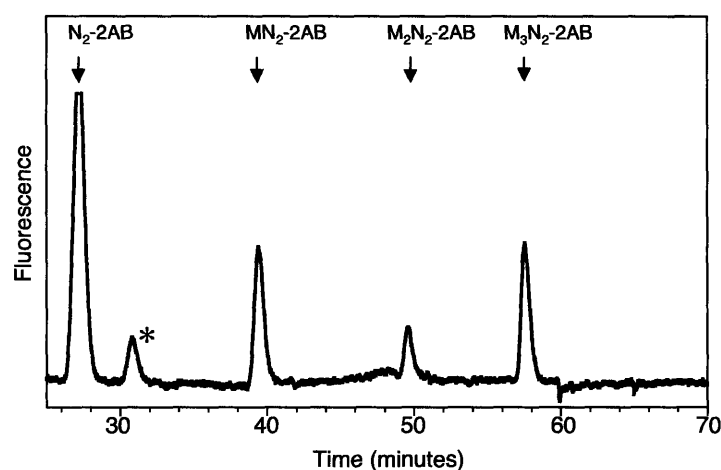
Figure IV-17. Alg1 and TRX-Alg2 accept Und-PP-linked substrates. Fluorescence HPLC showing elongation of $GlcNAc_2$-PP-undecaprenyl by Alg1 and TRX-Alg2. *This peak was not characterized, but has the retention time of peracetylated $GlcNAc_2$, and may be residual product from the synthesis of the acetyl-protected $GlcNAc_2$-PP-Und.

Conclusion

The results presented here provide evidence for the dual function of Alg2, specifically the ability of this enzyme to catalyze an $\alpha 1,3$-mannosylation followed by an $\alpha 1,6$-mannosylation to convert the trisaccharide intermediate, $ManGlcNAc_2$-PP-Dol to the pentasaccharide intermediate, $Man_3GlcNAc_2$-PP-Dol. Other examples of bifunctional

glycosyltransferases that catalyze the formation of more than one type of linkage and/or transfer more than one type of monosaccharide unit include the chondroitin synthases, which display both β1,3-N-acetylgalactosamine transferase and β1,4-glucuronic acid transferase activities,[12-16] and hyaluronan synthase,[16, 17] having β1,3-N-acetylglucosamine transferase and a β1,4-glucuronic acid transferase activity. In addition, the enzyme FT85 from *Dictyostelium* displays both β1,3-galactosytransferase and α1,2-fucosyltransferase activity,[18] and the *E. coli* enzyme, KfiC, polymerizes a repeat structure of a bacterial capsular polysaccharide by alternating α1,4- and β1,4-additions of N-acetylglucosamine and glucuronic acid, respectively.[19] A common feature of all of these bifunctional glycosyltransferases, in addition to being larger than Alg2, is the apparent presence of two separate glycosyltransferase domains that isolate the two distinct activities. This characteristic may be necessary due to the fact that, unlike Alg2, two different types of glycosyl donors must be recognized. The results of the Alg2 mutants presented here suggest that only one active site is used for both transformations, since the first Glu residue of the EX$_7$E motif seems to be important for the α1,6-mannosylation, and the second Glu residue affects (at least) the α1,3-mannosylation. To our knowledge there is no other documented example of a glycosyltransferase that uses one active site to catalyze the formation of glycosidic linkages with different regioselectivity.

As described in Chapter I, nearly all known glycosyltransferase structures can be characterized by one of two major folds, the GT-A or the GT-B fold. The GT-A fold is characterized by two dissimilar subdomains separated by a hinge region, one of which is responsible for binding the nucleotide sugar, while the other binds the acceptor.[9] The structure of the *E. coli* glycosyltransferase, MurG, a β1,4-N-acetylglucosaminyl

transferase, has been solved in the presence and absence of the glycosyl donor, UDP-GlcNAc, and is a member of the GT-B superfamily.[20, 21] This enzyme, like those on the dolichol pathway, is membrane-associated and the glycosyl acceptor is activated by a long-chain polyisoprene pyrophosphate. In this case also, two domains separated by a deep cleft comprise the architecture of this glycosyltransferase. As increasing numbers of glycosyltransferase structures are available, there is mounting evidence that nucleotide-sugar donor utilizing glycosyltransferases have a flexible loop above the donor binding site that closes upon donor binding, creating the binding site of the acceptor.[22] These results are consistent with an ordered binding of the glycosyl donor followed by the glycosyl acceptor. Additional studies of glycosyltransferases have demonstrated that constraining glycosyl acceptors into a variety of conformations can have a significant impact on the kinetic parameters of the glycosyltransferases that recognize them.[23] These results, as well as observed conformational changes and the presence of a flexible linker between acceptor and donor binding sites of glycosyltransferases suggests that a certain amount of plasticity exists in the active sites of these enzymes.

Just as the chemoenzymatic preparation of the core trisaccharide has been accomplished,[24] it is now possible to prepare the core pentasaccharide in a chemoenzymatic fashion. This structure is a valuable intermediate in the preparation of homogenous glycoproteins, since it is found as the core of every N-linked glycan, regardless of the Golgi-mediated elaboration. The chemical synthesis of these oligosaccharides is extremely challenging,[25] and while efforts toward the automated synthesis of oligosaccharides are yielding improvements,[26] glycosyltransferases are often

employed in chemoenzymatic syntheses due to the precise control over the stereo- and regioselectivity of glycosidic bond formation.[27]

Acknowledgements

Experimental Procedures

Media and buffers

Minimal Media (1 L): 6.7 g yeast nitrogen base (Difco) (- amino acids, + salts), 0.6 g

CSL-His-Leu-Trp (Bio101, Inc.), 50 mg His, 100 mg Leu, 100 mg Trp, 2% D-glucose.

Induction Media (1 L): 6.7 g yeast nitrogen base (Difco) (- amino acids, + salts), 0.6 g

CSL-His-Leu-Trp (Bio101, Inc.), 50 mg His, 100 mg Leu, 100 mg Trp, 2% galactose and

1% raffinose.

Buffer A: 50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.1% Nonidet P-

40.

Buffer B: 50 mM Tris-Cl pH 7.5, 0.25 M sucrose, 2.5 mM $MgCl_2$

Buffer C: 50 mM Tris-Cl pH 7.5, 2.5 mM $MgCl_2$

Buffer D (2×): 38 mM Tris-Cl pH 7.2, 1.8 mM DTT, 0.28 mM EDTA, 0.26% NP-40.

Buffer E: 50 mM Tris-acetate pH 8.5, 1 mM EDTA.

TUP: $CHCl_3$: MeOH: 4 mM $MgCl_2$ (2.75:44:53.25)

Cloning of ALG2 constructs

(See Chapter V for cloning of ALG11)

Genomic DNA was extracted from the *S. cerevisiae* strain PRY46 using the Yeast

DNA Extraction Reagent (Y-DER) (Pierce, Inc.). *ALG2* was amplified by PCR from the

yeast genomic DNA by PCR using Vent polymerase and the following primers:

Forward 5'-CAC CAT GAT TGA AAA GGA TAA A-3'

Reverse 5'-TAT TTC TTC ATA AGG GTA-3' (reverse)

The underlined sequence is recognized by topoisomerize, which mediates the ligation of the PCR product into the vector. The rest of the primer sequence corresponds to native *ALG2* sequence. Ligations were carried out using the Gateway Cloning Technology (Invitrogen) that was described in Chapter III, according to the manufacturer's instructions. A stop codon was later inserted into this construct for cell membrane preparations by site-directed mutagenesis, and the mutant was verified by DNA sequencing as above. Primers for site-directed mutagenesis were the following, in which underlined sequence indicates the change.

Forward primer 5'-GAA GAA ATA <u>TAG</u> GGT GGG CGC GCC G-3'

Reverse primer 5'-CGG CGC GCC CAC C<u>CT A</u>TA TTT CTT C-3'.

Cloning and subsequent transfer by homologous recombination to an appropriate destination vector for expression in either *E. coli* or *S. cerevisiae* was accomplished using the Gateway® Cloning Technology (Invitrogen) as described in Chapter III.

Both *ALG2* constructs, including or omitting the stop codon, were recombined with pBAD-DEST49 for *E. coli* expression. The Strep tag for yeast expression was introduced into the *ALG2* entry vector by amplifying the gene from pBAD-*ALG2*-DEST49 using Pfu Turbo DNA polymerase with a forward primer designed to include the sequence for the Strep tag followed by the first nine codons of the *ALG2* coding sequence (as above), and the reverse primer used for amplification of *ALG2* without the stop codon (above). No spacer was incorporated between the Strep tag and native *ALG2* sequence. The reverse primer was the same as indicated above. Cloning and homologous recombination with pYES-DEST52 for expression in *S. cerevisiae* were carried out as for the *ALG1* yeast expression vector (Chapter III).

144

All mutants were prepared at the entry vector stage (pENTR-*ALG2*), using the construct with a stop codon, by site-directed mutagenesis using the QuikChange Kit (Stratagene). The following primers were used for mutagenesis, with underlined nucleotides indicating the site of mutation:

E335Afor: 5'-CACCAGCATATGCGCACTTTGGTATTGTTCC-3'

E335Arev: 5'-GGAACAATACCAAAGTGCGCATATGCTGGTG-3'

E343Afor: 5'-GGTATTGTTCCTTTAGCAGCCATGAAATTAGG-3'

E343Arev: 5'-CCTAATTTCATGGCTGCTAAAGGAACAATACC-3'

The *ALG2* double mutant was prepared by changing the codon corresponding to E343A in the E335A mutant background.

All plasmids in the pENTR/SD/D-TOPO vector were verified by sequencing at the MIT Biopolymers Laboratory using the forward and reverse M13 primers as for *ALG1* constructs (Chapter III).

Expression in *E. coli* and purification of TRX-Alg2-V5-His

One Shot Top 10 competent cells were transformed with pBAD-*ALG2*-DEST49 *E. coli* expression vector as well as plasmid isolated from BL21(DE3) codon-plus RIL cells. Cells were shaken at 37 °C in 1 L of Terrific Broth (Invitrogen) until growth had nearly reached the stationary phase (A_{600} = 0.9-1.0), and the culture was then induced with 0.2% L-arabinose. Upon induction, cells were moved to a 15 °C water bath shaker, and expression was allowed to proceed overnight at this temperature. Cells were then harvested by centrifugation (10 min, 15,344 × g, 4 °C), and pellets were stored at –80 °C until purification.

The expressed TRX-Alg2-V5-His tag construct was purified by Ni-NTA affinity chromatography. A cell pellet from 250 mL of expression culture was resuspended in 35 mL of Buffer A containing 1% Dodecyl-D-maltoside (DDM), 1× protease inhibitor cocktail III (Calbiochem), and 1 mg/ mL lysozyme and allowed to incubate on ice for 30 minutes. Cells were then lysed by sonication for 2 minutes or until the solution clarified using a Branson 250 Sonifier with a 3/8" diameter tip (Branson Ultrasonics Co.), and then centrifuged (30 min, 11,625 × g, 4 °C). The supernatant from this spin was filtered through a 0.2-μm syringe-filter and was added to Ni-NTA slurry (5 mL of a 50% slurry, Qiagen). This mixture was rocked at 4 °C for 1 hr. After collecting the flow-through and washing the resin twice with 20 mL of Buffer A containing 20 mM imidazole and 1% DDM, bound protein was eluted using a step-wise gradient of 5-mL Buffer A aliquots containing 1% DDM and 40 mM, 80 mM, 120 mM, 160 mM, and 200 mM imidazole, respectively. Fractions of 1 mL were collected, and protein-containing fractions were dialyzed against Buffer B and concentrated by incubation with 0.5 mL of Ni-NTA slurry at 4 °C for 1 hr. Protein was eluted with two 0.5-mL aliquots of Buffer A containing 200 mM imidazole and 1% Triton X-100. Elutions were then dialyzed as above to remove imidazole.

Expression in *S. cerevisiae* and purification of Strep-Alg2-V5-His

The yeast strain INV*Sc*1 was transformed with pYES-(*StrepALG2*)-DEST52 using the *S. cerevisiae* EasyComp Transformation Kit (Invitrogen). Transformants were selected based on uracil prototrophy using SC-U plates lacking uracil. A 10-mL overnight culture in minimal media (containing 2% raffinose instead of glucose) was

added to 500 mL of minimal media. Once the 500-mL cultures had grown at to an O.D. of 7 with shaking at 30 °C, four equal aliquots were added to four flasks, each containing 1 L of induction media, for a final O.D. of 0.4 in each culture. Cultures were shaken at 30 °C for 16 hours, and then harvested by centrifugation (2455 × g, 5 minutes, 4°C), washing once with cold Buffer C. All subsequent steps were performed at 4°C. To prepare crude microsomes, the cell pellet was resuspended in 200 mL of Buffer C containing the protease inhibitors 0.1 mM AEBSF, 0.5 µg/mL pepstatin A, 0.5 µg/mL leupeptin (added immediately before use) and lysed using glass beads in a bead beater (16 × 20 sec, with 40-sec cooling intervals). Cell lysate was centrifuged for 8 minutes at 7520 × g, and this step was repeated with the consequent supernatant. Crude microsomes were pelleted by centrifugation of the supernatant from the second spin for 1 hr at 186,000 × g using a type 45Ti rotor in an ultracentrifuge (Beckman). The pellet was homogenized in 50 mL of Buffer A containing 1% Nonidet P-40, shaken on ice for 15 minutes, and then centrifuged for 1 hour at 186,000 × g. The supernatant was saved as solubilized microsomes, which were subsequently filtered through a 0.2 µm syringe filter and incubated with 1 mL of Ni-NTA resin (Qiagen) for 1 hour prior to column purification. The flow-through was collected, and then the resin washed twice with 20 mL of Buffer A. Protein was eluted with Buffer A containing 100 mM imidazole and 0.1% Nonidet P-40. Alg2-containing fractions were added to 1 mL of Streptactin resin (Stratagene) and incubated overnight at 4 °C with shaking. After collecting the flow-through and washing the resin twice with 10 mL of Buffer A containing 1 mM DTT, Alg2 was eluted with 5 mL of Buffer A containing 1 mM DTT and 2.5 mM desthiobiotin,

collecting 0.5-mL fractions. Purified Alg2-containing fractions were pooled and dialyzed against Buffer B.

Preparation and assays of *alg2* mutant and induced and uninduced wild-type Alg2 microsomes

Yeast cultures were grown, and the wild-type pYES-(*StrepALG2*)-DEST52 strain expressed, as described in the previous section. For the uninduced pYES-(*StrepALG2*)-DEST52, galactose and raffinose were replaced by glucose, and otherwise growth was the same. Crude microsomes were prepared from these cultures, also as described for the Strep-Alg2-V5-His preparation, but at the crude microsome stage, rather than homogenizing in detergent solution, glycerol was added to 30% and microsomes were stored at −80 °C. Elongation assays were carried out as described below, using 5-μL aliquots of microsomes and 1× Buffer D for each assay.

Preparation of crude microsomes from the *alg2* yeast strain

Starting from a single colony of the *alg2* strain[19, 20] on a YPAD plate, cells were grown in YPAD by fermentation in a 10-L BioFlow fermentor (New Brunswick Scientific), harvested by centrifugation (3836 × g, 20 min, 4 °C), and then stored at −80 °C until use. Crude microsomes were prepared from a 100-g cell pellet as described in the previous section. This preparation yielded 9.5 mL of crude microsomes, which were mixed well with 2.5 mL of glycerol, aliquoted, and stored at −80 °C until use.

For complementation assays with the *alg2* mutant strain, a range of concentrations (nanomolar to micromolar) of purified TRX-Alg2-V5-His or Strep-Alg2-V5-His were included in the assays.

Expression and isolation of TRX-Alg2 in *E. coli* membranes

For *E. coli* expression, One Shot Top 10 competent cells were transformed with pBAD(*ALG2*)-DEST49 (containing the stop codon) together with the plasmid from BL21(DE3)RIL codon-plus competent cells (Stratagene). Aliquots of 10 mL from overnight cultures were added to 2.5-L batches of Terrific Broth (Invitrogen) and grown at 37 °C with shaking to an OD_{600} of 0.8-1.0. Cultures were then adjusted to 16 °C prior to induction of protein expression with the addition of L-arabinose to 0.2 % (w/v). After 24 hours, cells were harvested at 4 °C by centrifugation (25 minutes, 5,000 × g), washing once with 0.9 % (w/v) NaCl. The resulting pellet was then either used directly for a cell membrane preparation, or stored at –80 °C until use. To prepare the membrane fraction, all steps were performed at 4 °C. Cell pellets from 2.5-L expression cultures were resuspended in 50 mL of Buffer E containing 1× protease inhibitor cocktail III (Calbiochem). Cells were lysed by sonication using a Branson 250 Sonifier with a 3/8" diameter tip (Branson Ultrasonics Co.). Cells were subjected to three 15-sec pulses at the highest output control setting, with 3-min cooling intervals in between pulses. The lysate was then centrifuged for 30 minutes at 5,697 × g. The membrane fraction was pelleted by centrifugation of the supernatant for 1 hr at 142,414 × g using a 45Ti rotor in an ultracentrifuge (Beckman). The resulting brown, translucent pellet was rinsed with

Buffer E, then homogenized in 0.5 mL of Buffer E, and stored at –20 °C in 30% (v/v) glycerol.

Assay for elongation of dolichyl-linked saccharide intermediates

Reaction mixtures contained approximately 40 μM dolichyl-linked starting material, 2 mM GDP-Man, and 5 μL of the cell membrane fraction containing TRX-Alg2 (195 μg total protein). For assays with the mutants, equivalent amounts (5 μL) were used. Starting with a dried aliquot of the dolichyl-linked starting material, a 50-μL aliquot of Buffer D (2×) was added and the mixture sonicated for 1 min to resuspend the substrate. GDP-Man (2 mM final), $MgCl_2$ (10 mM final), dH_2O, and enzyme were added to a final volume of 100 μL. Incubation proceeded for 1 hour at room temperature, followed by quenching into a biphasic mixture of $CHCl_3$: MeOH: 4 mM $MgCl_2$ (6:4:1). (Purified Alg2 assays were also attempted using a buffer that had been commonly used for Enzyme II assays, which was composed of 50 mM Tris, pH 7, 0.25 M sucrose, 3 mM DTT, 5 mM $MgCl_2$, 0.13 % NP-40.) The dolichyl-linked saccharides, partitioned into the organic phase, were subsequently washed twice with 200-μL aliquots of TUP. The extracted organic phase was then concentrated to dryness.

Hydrolysis, fluorescent labeling, HPLC, and MALDI were carried out as described in Chapter III.

Mannosidase cleavage reactions

2-AB labeled glycans were collected from HPLC separation, taking advantage of the fluorescence detection, and then concentrated to dryness using a Sc110 Speed-Vac

(Savant, Inc.). Dried glycans were redissolved in 5 μL of dH₂O, and 1 μL-aliquots were treated with 2 Units (1 μL) of α1-2,3-mannosidase (New England BioLabs, Inc.) in 50 mM sodium citrate, pH 6.0, 5 mM CaCl₂, 100 mg/mL BSA, or with 1 milliUnit (1 μL) of α1,6-mannosidase (Calbiochem) in 50 mM sodium phosphate, pH 5.0 at 37 °C for 5 hours. For the α1-2,3-mannosidase reactions, 0.5-μL aliquots each of commercial buffer (10×) and BSA (10×) were used, which were supplied with the enzyme, in a final reaction volume of 5 μL. α1,6-Mannosidase reactions were performed in a final volume of 10 μL, using 1× buffer.

References

1.   Huffaker, T. C.; Robbins, P. W., Yeast mutants deficient in protein glycosylation. *Proc. Natl. Acad. Sci. USA* **1983,** 80, 7466-7470.

2.   Takeuchi, K.; Yamazaki, H.; Shiraishi, N.; Ohnishi, Y.; Nishikawa, Y.; Horinouchi, S., Characterization of an *alg2* mutant of the zygomycete fungus *Rhizomucor pusillus*. *Glycobiology* **1999,** 9, (12), 1287-1293.

3.   Thiel, C.; Schwarz, M.; Peng, J.; Grzmil, M.; Hasilik, M.; Braulke, T.; Kohlschutter, A.; von Figura, K.; Lehle, L.; Korner, C., A new type of congenital disorder of glycosylation (CDG-Ii) provides new insights into the early steps of dolichol-linked oligosaccharide biosynthesis. *J. Biol. Chem.* **2003,** 278, (25), 22498-22505.

4.   Jackson, B. J.; Kukuruzinska, M. A.; Robbins, P. W., Biosynthesis of asparagine-linked oligosaccharides in *Saccharomyces cerevisiae*: the *alg2* mutation. *Glycobiology* **1993,** 3, (4), 357-364.

5.   Yamakazi, H.; Shiraishi, N.; Takauchi, K.; Ohnishi, Y.; Horinouchi, S., Characterization of ALG2 encoding a mannosyltransferase in the zygomycete fungus Rhizomucor pusillus. *Gene* **1998,** 221, 179-184.

6.   Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.

7.   Wong-Madden, S. T.; Landry, D., Purification and characterization of novel glycosidases from the bacterial genus Xanthomonas. *Glycobiology* **1995,** 5, (1), 19-28.

8.   Chapman, A.; Li, E.; Kornfeld, S., The Biosynthesis of the Major Lipid-Linked Oligosaccharide of Chinese Hamster Ovary Cells Occurs by the Ordered Addition of Mannose Residues. *J. Biol. Chem.* **1979,** 254, (20), 10243-10249.

9.   Bourne, Y.; Henrissat, B., Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr. Opin. Struct. Biol.* **2001,** 11, (5), 593-600.

10.   Abdian, P. L.; Lellouch, A. C.; Gautier, C.; Ielpi, L.; Geremia, R. A., Identification of essential amino acids in the bacterial α-mannosyltransferase AceA. *J. Biol. Chem.* **2000,** 275, (51), 40568-75.

11.   Wilson, I. B.; Webberley, M. C.; Revers, L.; Flitsch, S. L., Dolichol is not a necessary moiety for lipid-linked oligosaccharide substrates of the mannosyltransferases involved in *in vitro* N-linked-oligosaccharide assembly. *Biochem. J.* **1995,** 310 (Pt 3), 909-16.

12.   Hwang, H. Y.; Olson, S. K.; Esko, J. D.; Horvitz, H. R., Caenorhabditis elegans early embryogenesis and vulval morphogenesis require chondroitin biosynthesis. *Nature* **2003,** 423, (6938), 439-43.

13.   Yada, T.; Gotoh, M.; Sato, T.; Shionyu, M.; Go, M.; Kaseyama, H.; Iwasaki, H.; Kikuchi, N.; Kwon, Y. D.; Togayachi, A.; Kudo, T.; Watanabe, H.; Narimatsu, H.; Kimata, K., Chondroitin sulfate synthase-2. Molecular cloning and characterization of a novel human glycosyltransferase homologous to chondroitin sulfate glucuronyltransferase, which has dual enzymatic activities. *J. Biol. Chem.* **2003,** 278, (32), 30235-47.

14.   Yada, T.; Sato, T.; Kaseyama, H.; Gotoh, M.; Iwasaki, H.; Kikuchi, N.; Kwon, Y. D.; Togayachi, A.; Kudo, T.; Watanabe, H.; Narimatsu, H.; Kimata, K., Chondroitin sulfate synthase-3. Molecular cloning and characterization. *J. Biol. Chem.* **2003,** 278, (41), 39711-25.

15.   Ninomiya, T.; Sugiura, N.; Tawada, A.; Sugimoto, K.; Watanabe, H.; Kimata, K., Molecular cloning and characterization of chondroitin polymerase from *Escherichia coli* strain K4. *J. Biol. Chem.* **2002,** 277, (24), 21567-75.

16.   Jing, W.; DeAngelis, P. L., Analysis of the two active sites of the hyaluronan synthase and the chondroitin synthase of *Pasteurella multocida. Glycobiology* **2003,** 13, (10), 661-71.

17.   Jing, W.; DeAngelis, P. L., Dissection of the two transferase activities of the *Pasteurella multocida* hyaluronan synthase: two active sites exist in one polypeptide. *Glycobiology* **2000,** 10, (9), 883-9.

18.   Van Der Wel, H.; Fisher, S. Z.; West, C. M., A bifunctional diglycosyltransferase forms the Fucα1,2Galβ1,3-disaccharide on Skp1 in the cytoplasm of *dictyostelium. J. Biol. Chem.* **2002,** 277, (48), 46527-34.

19.   Griffiths, G.; Cook, N. J.; Gottfridson, E.; Lind, T.; Lidholt, K.; Roberts, I. S., Characterization of the glycosyltransferase enzyme from the *Escherichia coli* K5 capsule gene cluster and identification and characterization of the glucuronyl active site. *J. Biol. Chem.* **1998,** 273, (19), 11752-7.

20.   Ha, S.; Walker, D.; Shi, Y.; Walker, S., The 1.9 Å crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. *Protein Sci.* **2000,** 9, (6), 1045-52.

21.   Hu, Y.; Chen, L.; Ha, S.; Gross, B.; Falcone, B.; Walker, D.; Mokhtarzadeh, M.; Walker, S., Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc. Natl. Acad. Sci. USA* **2003,** 100, (3), 845-9.

22.   Qasba, P. K.; Ramakrishnan, B.; Boeggeman, E., Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem. Sci.* **2005,** 30, (1), 53-62.

23.   Galan, M. C.; Venot, A. P.; Boons, G. J., Glycosyltransferase activity can be modulated by small conformational changes of acceptor substrates. *Biochemistry* **2003,** 42, (28), 8522-9.

24.   Watt, G. M.; Revers, L.; Webberley, M. C.; Wilson, I. B. H.; Flitsch, S. L., Efficient enzymatic synthesis of the core trisaccharide of N-glycans with a recombinant β-mannosyltransferase. *Angew. Chem., Int. Ed. Eng.* **1997,** 36, (21), 2354-2356.

25.   Takatani, M.; Nakama, T.; Kubo, K.; Manabe, S.; Nakahara, Y.; Ito, Y.; Nakahara, Y., Synthesis of N-linked pentasaccharides with isomeric glycosidic linkage. *Glycoconj. J.* **2000,** 17, (6), 361-75.

26.   Ratner, D. M.; Swanson, E. R.; Seeberger, P. H., Automated synthesis of a protected N-linked glycoprotein core pentasaccharide. *Org. Lett.* **2003,** 5, (24), 4717-20.

27.   Hanson, S.; Best, M.; Bryan, M. C.; Wong, C. H., Chemoenzymatic synthesis of oligosaccharides and glycoproteins. *Trends Biochem Sci.* **2004,** 29, (12), 656-63.

Chapter V: Studies toward defining the function of Alg11

Introduction

The results from yeast genetic studies have assisted in narrowing down the possible functions of Alg11 to either the fourth, fifth, or both the fourth and fifth mannosylation steps of the dolichol pathway (Scheme V-1).[1-3] This assignment has been primarily based on the characterization of *alg11* yeast mutants, the first of which was isolated based on resistance to sodium vanadate, which is a common, though not fully understood, feature of mutants with glycosylation defects.[4, 5]
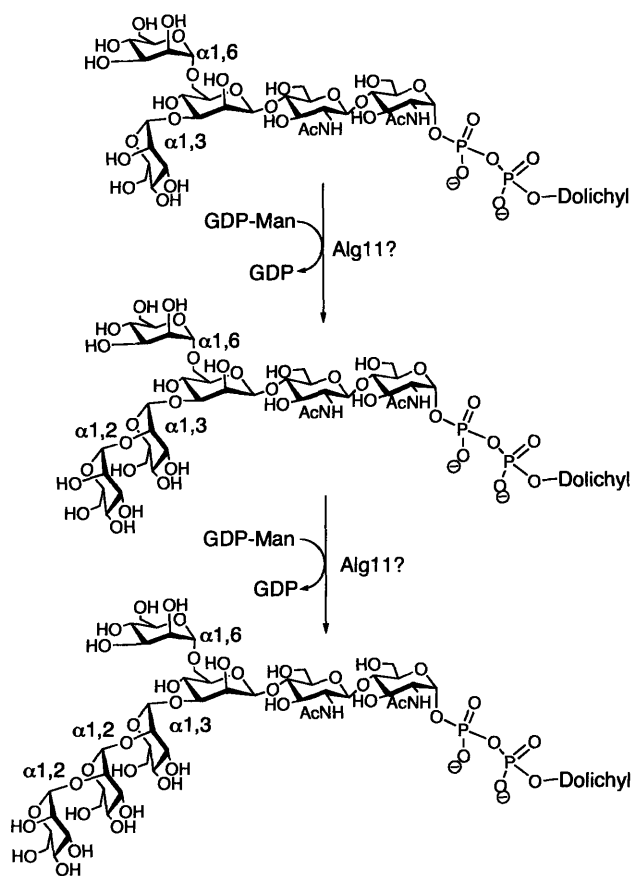
Scheme V-1. α1,2-Mannosylation reactions putatively catalyzed by Alg11.

Similar to the study of other *alg* mutants, oligosaccharide structural studies of dolichylpyrophosphate-linked saccharide intermediates extracted from a yeast strain that lacks Alg11 function have helped to narrow down the possible roles of Alg11.[1] First, *ALG11* was shown to be essential for viability. Since it is known that the lumen-oriented glycosyltransferases are not essential, this feature suggests that Alg11 is a cytosolic-oriented enzyme in the pathway. Due to the essential nature of *ALG11*, a complete deletion strain would be inviable, therefore a diploid strain, *alg11Δ,* heterozygous for the *ALG11* deletion, was constructed and used for oligosaccharide lumenal structural studies. However, unlike the characterization of *alg1* and *alg2* mutants, the results of these studies were greatly complicated by the ability of lumenal transferases to elongate truncated saccharide intermediates containing the minimal core pentasaccharide. Although the *alg11Δ* mutant accumulates $Man_3GlcNAc_2$-PP-Dol as a significant intermediate, approximately 80% of the extracted lipid corresponded to $Hex_7GlcNAc_2$-PP-Dol, where Hex (hexose) could refer to Man or Glc.[1] Using mannosidase mapping in conjunction with known rules concerning the order of saccharide addition, it was proposed that this intermediate is $Man_7GlcNAc_2$-PP-Dol, although the precise structure could not be distinguished between forms missing one or both of the mannose residues putatively added by Alg11 (Scheme V-I). Taken together, these results suggested a role for Alg11 in the addition of the fourth mannose. However, the additional accumulation of a $Man_4GlcNAc_2$-PP-Dol intermediate as well as a presumed elongation product of this hexasaccharide that contained eight mannose residues suggests that the fifth mannosylation is also impaired in the *alg11Δ* mutant. From exhaustive saccharide structural studies of every extracted intermediate, as well as of glycans released from

glycoproteins in this strain, it is apparent that Alg11 is involved in formation of Man$_4$GlcNAc$_2$-PP-Dol and/or Man$_5$GlcNAc$_2$-PP-Dol.[1]

The contribution of activity from lumenal transferases in this study as well as the inability to distinguish between different isoforms in some cases greatly complicated the interpretation of these results, and there has been no biochemical support to establish the validity of these proposals. The fact that extensive genetics screens and bioinformatics searches have not yielded any additional candidates for cytosolic dolichol pathway mannosyltransferases raised the question of whether Alg11, like Alg2 and Alg9,[6] may be responsible for more than one step in the pathway.

Alg11 from *S. cerevisiae* is a 548-residue, 63-kDa protein with one predicted transmembrane domain at the N-terminus and a large soluble domain, as predicted by the TMHMM server, v. 2.0 (Figure V-1).[7]
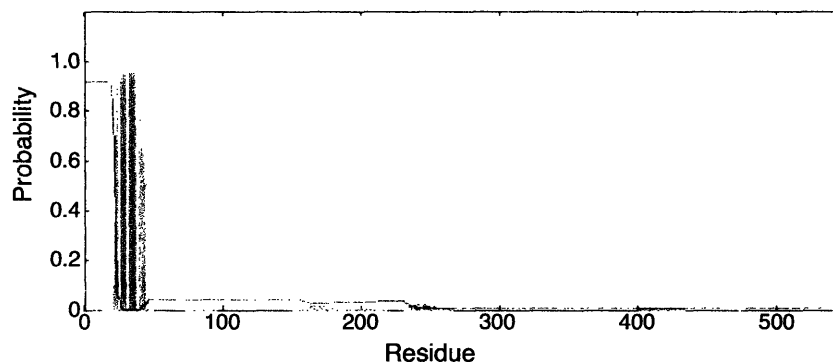
Figure V-1. Predicted topology of Alg11. Plot of the probability that Alg11 residues are in a transmembrane-spanning region.

From the topology prediction, this enzyme is not expected to be as challenging to express and analyze *in vitro* as Alg2 has proven to be. The difficulty in studying Alg11

158

has been that the putative substrate, $Man_3GlcNAc_2$-PP-Dol, is not readily available.

Chemical synthesis is prohibitively complicated, and no mutant strain exists that

accumulates this intermediate predominantly, since it is sufficient for elongation by

lumenal glycosyltransferases. Fortunately, the discovery that Alg2 catalyzes elongation

of the $ManGlcNAc_2$-PP-Dol to $Man_3GlcNAc_2$-PP-Dol (Chapter IV) provided a

chemoenzymatic method to access the putative Alg11 substrate. Thus, having the means

to prepare the pentasaccharide intermediate using TRX-Alg2 presented the opportunity to

examine the function of Alg11. Our goal with respect to Alg11 was to unambiguously

define the function using the same methods described in Chapter IV, for the validation of

Alg2 function. The sensitivity of this preparation to mutations in the conserved EX_7E

motif (Figure V-2) was also examined, primarily to eliminate the possibility of

mistakenly assigning background activity to Alg11 activity, but also to gain some

preliminary insights into the postulated importance of selected conserved residues. TRX-

Alg11 was isolated in a cell membrane fraction, based on the positive results afforded by

this method in the case of TRX-Alg2. Future studies will include purification of this

enzyme.

		H	F	G	I	A	V	V		Y	M	A	S	G	L	I	P
S. Cerevisiae (404-421)	N	H	F	G	I	A	V	V		Y	M	A	S	G	L	I	P
S. Pombe (370-387)	N	H	F	G	I	G	V	V		Y	M	A	A	G	L	I	P
C. Elegans (376-393)	N	H	F	G	I	S	V	V		A	M	A	A	S	T	I	I
Drosophila 1 (379-396)	N	H	F	G	I	G	I	V		S	M	A	A	G	L	I	M
Arabidopsis (375-392)	D	H	F	G	I	S	V	V		Y	M	A	A	G	A	I	P
Leishmania (520-537)	D	H	F	G	I	V	L	L		Y	L	A	A	G	C	I	P

Figure V-2. Alignment of a conserved sequence in Alg11 from six different organisms.

Conserved residues are highlighted with light gray background, and the Glu residues of

the EX_7E motif are highlighted with dark gray background. Adapted from Cipollo et al.[1]

Results and Discussion

TRX-Alg11 cloning, expression, and cell membrane fraction preparation

TRX-Alg11, with a free C-terminus, was cloned, expressed, and isolated in the membrane fraction of *E. coli* (Figure V-3), repeating exactly the methods used for TRX-Alg2 as described in Chapter IV.
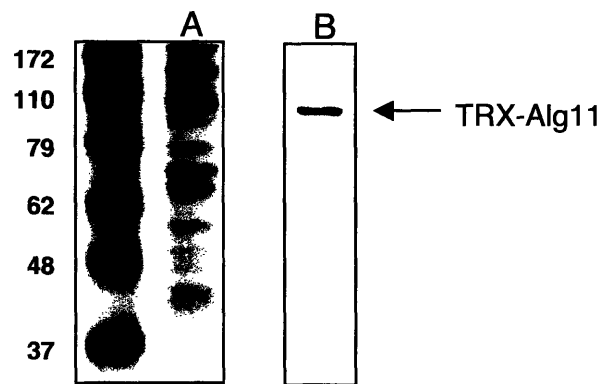
Figure V-3. Isolated TRX-Alg11 from *E. coli* expression. A) Gelcode stain (Pierce Co.) and B) Western blot analysis of the cell membrane fraction from cells overexpressing TRX-Alg11. The Western blot was probed with the α-Thio antibody (Invitrogen), which detects the TRX domain. Molecular weight markers are labeled to the left in kDa.

TRX-Alg11 forms Man$_5$GlcNAc$_2$-PP-Dol from Man$_3$GlcNAc$_2$-PP-Dol

The products of the TRX-Alg2 reaction were used as starting material for assays with TRX-Alg11 in the presence of GDP-Man. The starting material was composed primarily of the pentasaccharide, Man$_3$GlcNAc$_2$-PP-Dol, but also contained minor amounts of trisaccharide and tetrasaccharide (Figure V-4A). Incubation of this mixture with TRX-Alg11 resulted in the disappearance of the pentasaccharide with a concomitant

160

appearance of a heptasaccharide (Figure V-4B). The apparent loss of material in the

conversion of pentasaccharide to heptasaccharide is presumably due to the relatively poor

organic-partitioning properties of the latter compared to the former.
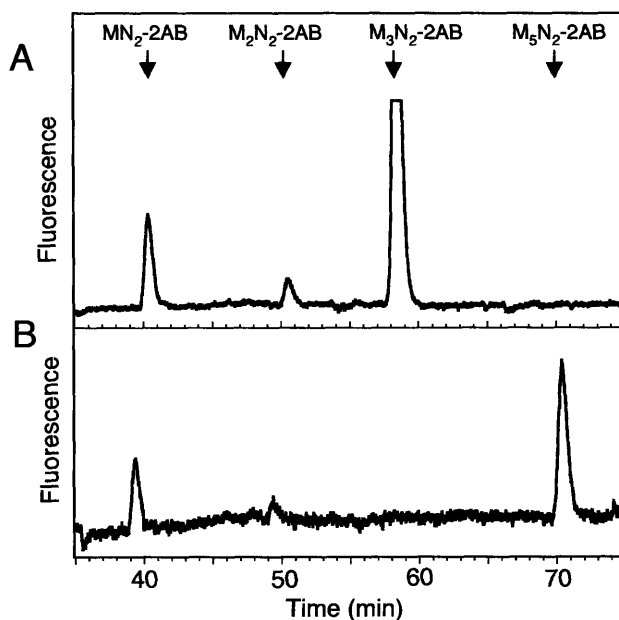
Figure V-4. HPLC of TRX-Alg11 product. Fluorescence HPLC traces show A) TRX-

Alg2 products (starting material for TRX-Alg11 reaction) and B) TRX-Alg11 products.

(M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

The TRX-Alg11 product (Figure V-5A) was characterized by cleavage with an

α1,2-mannosidase to yield a pentasaccharide (Figure V-5B). These results, in addition to

the MALDI-MS of the heptasaccharide product (expected $[M+Na]^+$ = 1377.5, found

$[M+Na]^+$ = 1377.5) (Figure V-5C), are consistent with the expected structure, Manα1,2-

Manα1,2-Manα1,3-(Manα1,6)-Manβ1,4-GlcNAcβ1,4-GlcNAc-2AB (Figure V-5D).
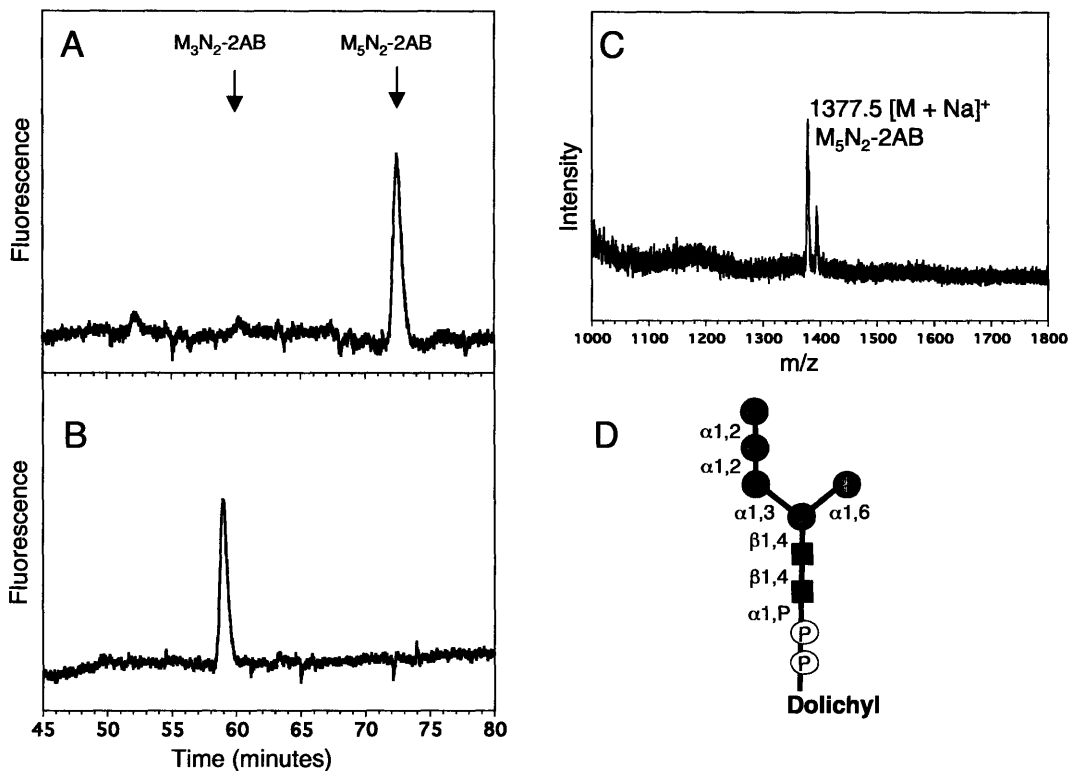
Figure V-5. Characterization of the TRX-Alg11 product. Fluorescence HPLC traces A) before and B) after treatment with an α1,2-mannosidase (Prozyme, Inc) (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide). C) MALDI-MS of the hydrolyzed and labeled heptasaccharide product, $Man_5GlcNAc_2$-2AB. D) Structure of the TRX-Alg11 product based on MALDI and mannosidase characterization. (Black squares, GlcNAc; gray circles, Man; P, phosphate)

Thus, TRX-Alg11 was shown to catalyze the elongation of $Man_3GlcNAc_2$-PP-Dol to $Man_5GlcNAc_2$-PP-Dol with the appropriate linkages corresponding to those of the dolichol pathway heptasaccharide intermediate. The residual trisaccharide and tetrasaccharide intermediates remained unreacted, highlighting the specificity of TRX-

Alg11 and providing more convincing evidence that the pentasaccharide is the native substrate of Alg11. Similar to the use of TRX-Alg11 as a negative control for ManGlcNAc$_2$-PP-Dol elongation in Chapter IV, the fact that the cell membrane fraction containing TRX-Alg2 does not elongate beyond the pentasaccharide product (Chapter IV) argues against the possibility that background *E. coli* enzymes are capable of completing the heptasaccharide biosynthesis.

Mannosyltransferase activities of TRX-Alg11 mutants

Further evidence that the observed activity comes from TRX-Alg11 was sought by mutating the Glu residues in the EX$_7$E motif. The relative expression levels of the three mutants were evaluated by Western blot analysis (Figure V-6A)). For reasons that are unclear, the double mutant is consistently expressed (or isolated) at a significantly higher level compared to the single mutants. This characteristic, however, does not complicate the interpretation of the elongation assay results.
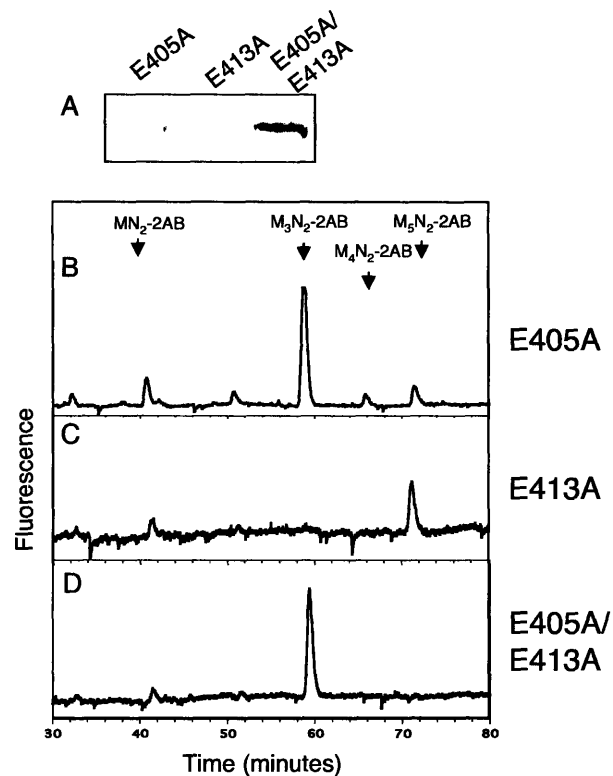
Figure V-6. Analysis of expression and activity of TRX-Alg11 mutants. A) Western blot analysis of equivalent amounts of cell membrane fraction from the overexpression of all TRX-Alg11 mutants to show relative expression levels, with α-Thio (Invitrogen) detection of the TRX domain. Fluorescence HPLC traces show the results of incubating $Man_3GlcNAc_2$-PP-Dol with the TRX-Alg11 mutants B) E405A, C) E413A, or D) the double mutant E405A/E413A.

Mutation of the first Glu to yield the TRX-Alg11 E405A mutant resulted in only trace amounts of product formation (Figure V-6B). These trace products were subjected to α1,2-mannosidase cleavage in order to identify them as the expected hexasaccharide

164

and heptasaccharide products, $Man_4GlcNAc_2$-PP-Dol and $Man_5GlcNAc_2$-PP-Dol, respectively. The cleavage products are shown in Figure V-7.
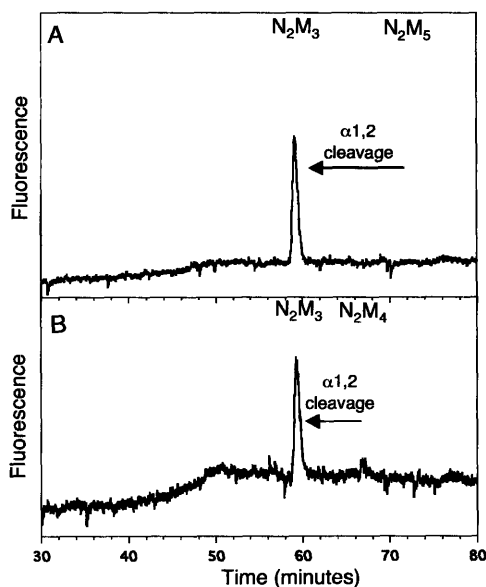
Figure V-7. Characterization of the TRX-Alg11E405A products. α1,2-Mannosidase treatment of the products of TRX-Alg11 E405A. Cleavage products of A) putative heptasaccharide and B) hexasaccharide.

In contrast, the TRX-Alg11 E413A mutant suffered no detectable loss in activity (Figure V-6C), which is consistent with another example in which the first Glu residue of the EX_7E motif in a retaining glycosyltransferase was found to be more critical for catalysis than the second.[8] The complete loss of activity in the double mutant, TRX-Alg11 E405A/E413A (Figure V-6D), suggests that either the E413 is critical for the residual activity displayed by the E405A mutant, or perhaps this perturbation in local structure is enough to abolish the remaining activity. More importantly, in terms of the current goal of defining the precise function of Alg11, the sensitivity of the observed

activity to point mutations in TRX-Alg11, and particularly the complete loss of activity in the double mutant, despite robust expression, argues against any involvement of enzymes in the *E. coli* cell membrane fraction background.

Specificity of TRX-Alg2 and TRX-Alg11

In the interest of unambiguously assigning the function of Alg2 and Alg11 from these studies, it was necessary to show that each is specific for the proposed substrate. To demonstrate that Alg2 and Alg11 are capable of distinguishing the correct substrate from preceding intermediates in the pathway, and to support the results presented here as biologically relevant, different combinations of Alg1 from *S. cerevisiae* expression (Chapter III), TRX-Alg2 (Chapter IV), and TRX-Alg11 were incubated with the Alg1 substrate, $GlcNAc_2$-PP-Dol. In the absence of Alg1, there was no evidence of elongation of $GlcNAc_2$-PP-Dol, demonstrating that neither TRX-Alg2 nor TRX-Alg11 accepts this intermediate (Figure V-8A). In the absence of TRX-Alg2, $ManGlcNAc_2$-PP-Dol accumulates, showing that TRX-Alg11 also does not accept the trisaccharide intermediate (Figure V-8B). The complete elongation to $Man_5GlcNAc_2$-PP-Dol in the presence of Alg1, TRX-Alg2, and TRX-Alg11 confirmed that all components were active (Figure V-8C).
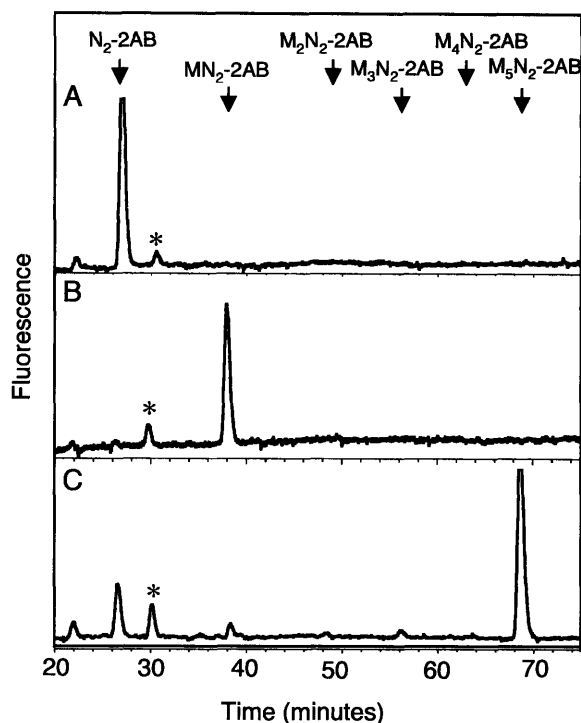
Figure V-8. Specificity of TRX-Alg2 and TRX-Alg11. Fluorescence HPLC traces of GlcNAc$_2$-PP-Dol elongation with different permutations of Alg1, TRX-Alg2, and TRX-Alg11. A) TRX-Alg2 + TRX-Alg11 B) Alg1 + TRX-Alg11 C) Alg1 + TRX-Alg2 + TRX-Alg11. *uncharacterized impurity has the retention time of peracetylated GlcNAc$_2$-2AB. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

The specificity of Alg11 for the isoprenoid carrier was also addressed. As shown in Chapter IV, Alg1 and TRX-Alg2 were each able to accept the undecaprenyl-PP (Und-PP) derivative of their native substrates, though with reduced turnover under comparable conditions. This experiment was repeated with TRX-Alg11 as an additional component, and the formation of the heptasaccharide intermediate demonstrated the ability of TRX-Alg11 to accept Man$_5$GlcNAc$_2$-PP-Und as a substrate (Figure V-9). Interestingly, every

intermediate, with the exception of the hexasaccharide, is observed to accumulate to some extent in this series of transformations.
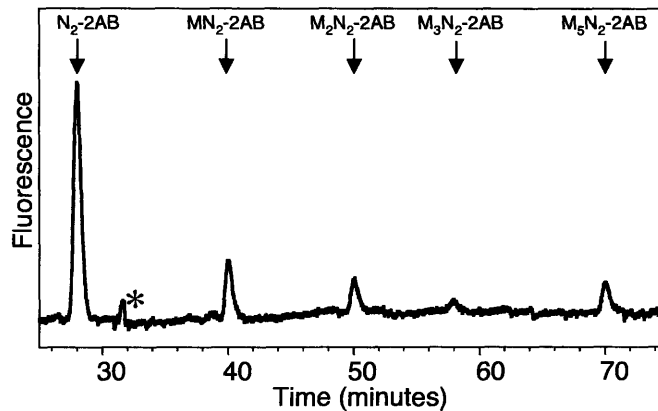
Figure V-9. TRX-Alg11 accepts Und-PP-linked substrate. Fluorescence HPLC of GlcNAc$_2$-PP-Und elongation with Alg1, TRX-Alg2, and TRX-Alg11. *uncharacterized impurity has the retention time of peracetylated GlcNAc$_2$-2AB. (M, Man; N, GlcNAc; 2AB, 2-aminobenzamide)

Conclusion

The results presented in this chapter demonstrate the dual function of Alg11, which complements the recent evidence for the dual function of Alg9,[6] as well as that of Alg2. These assignments complete the identification of the entire ensemble of glycosyltransferases that comprise the dolichol pathway. These findings explain the absence of candidates for any remaining cytosolic mannosyltransferases, despite numerous genetic screens and bioinformatics searches. Similar examples of this paradox, in which too few candidate glycosyltransferases are available to account for all of the

168

steps in a biosynthetic pathway, have recently been found in *Streptomyces cyanogenus* and in *Campylobacter jejuni*.[9, 10] In *S. cyanogenus*, Landomycin A biosynthesis involves the assembly of a hexasaccharide moiety by two monofunctional and two bifunctional glycosyltransferases.[9] The Pgl pathway of *N*-linked protein glycosylation in *C. jejuni* has many similarities to the dolichol pathway, including assembly of a polyisoprenoid pyrophosphate-linked oligosaccharide as the donor for transfer to protein.[11] The α1,4-*N*-acetylgalactosaminyltransferase, PglH, acts iteratively to transfer three GalNac residues to the growing oligosaccharide chain.[10] In these two examples, as in the case of Alg11, the enzymes are transferring the same sugar in the same configuration. Thus, it does not seem likely that more than one active site would be necessary. This hypothesis is supported by the results from the TRX-Alg11 double mutant, showing that a complete loss of detectable turnover can be achieved by point mutations within the confines of a 9-residue stretch of sequence. Due to the lower degree of membrane association compared to Alg2, Alg11 is a promising candidate for efficient expression and purification, which would facilitate more quantitative studies.

Acknowledgements

Experimental procedures

Cloning of pBAD-*ALG11*-DEST49

Genomic DNA was extracted from the *S. cerevisiae* strain PRY46 using the Yeast

DNA Extraction Reagent (Y-DER) (Pierce, Inc.). *ALG11* was amplified from the yeast

genomic DNA by PCR using Vent polymerase with the forward primer 5'-CAC CAT

GGG CAG TGC TTG GAC A-3' and reverse primer 5'-TCA GCC CCT TTC CTC CTC-

3' to yield a construct with a stop codon. This gene was then cloned into the

pENTR/SD/D-TOPO vector, from which it was transferred to the pBAD-DEST49

destination vector by homologous recombination using the Gateway® Cloning

Technology (Invitrogen) for *E. coli* expression of TRX-Alg11.

All mutants were prepared in the entry vector (pENTR/SD/D-TOPO) containing

the *ALG11* gene by site-directed mutagenesis using the QuikChange Kit (Stratagene)

using the following primers. Wild-type and mutant plasmids were all verified by DNA

sequencing at the MIT Biopolymers laboratory. The *ALG11* double mutant was prepared

by incorporating both mutations simultaneously into the wild-type background.

Underlined codons denote sites of mutation.

E405A forward: 5'- GCCATGTGGAAT<u>GCG</u>CACTTTGGAATT-3'

E405A reverse: 5'-AATTCCAAAGTG<u>CGC</u>ATTCCACATGGC-3'

E413A forward: 5'-ATTGCAGTTGTA<u>GCG</u>TATATGGCTTCCGG-3'

E413A reverse: 5'-CCGGAAGCCATATA<u>CGC</u>TACAACTGCAAT-3'

E405A/E413A forward:

5'-GCCATGTGGAAT<u>GCG</u>CACTTTGGAATTGCAGTTGTA<u>GCG</u>TATATGGCTTCCGG-3'

E405A/E413A reverse:

170

5'-CCGGAAGCCATATA<u>CGC</u>TACAACTGCAATTCCAAAGTG<u>CGC</u>ATTCCACATGGC-3'

TRX-Alg11 Expression and elongation assays

Using the construct, pBAD-*ALG11*-DEST49, expression of TRX-Alg11 and preparation of the cell membrane fraction were carried out exactly as described in Chapter IV.

The pentasaccharide substrate for TRX-Alg11 was prepared by scaling up the elongation reaction containing ManGlcNAc$_2$-PP-Dol, GDP-Man, and TRX-Alg2 that was described in Chapter IV. Wild-type TRX-Alg11 reactions contained 135 μg of total protein. Mutant TRX-Alg11 reactions were carried out with equivalent amounts of cell membrane fraction, and relative amounts of TRX-Alg11 in each are indicated by the Western blot analysis (Figure V-6A). In the extraction steps, the organic layer was washed with CHCl$_3$: MeOH: 100 mM KCl (3:48:47).

Glycan analysis by hydrolysis, labeling, HPLC, and MALDI were carried out as described in Chapter III, and the α1,2-mannosidase reactions were performed according to the manufacturer's instructions (Prozyme, Inc.).

GlcNAc$_2$-PP-Dol and GlcNAc$_2$-PP-Und elongation assays

These assays were carried out as described for TRX-Alg2 assays (Chapter IV). The enzymes, used as indicated for individual assays, were used in the following amounts: 0.1 μg of Alg1 (from expression in *S. cerevisiae*, Chapter III), 195 μg of total protein from a cell membrane fraction preparation containing wild-type TRX-Alg2, and

171

50 µg of total protein from a cell membrane fraction preparation containing wild-type

TRX-Alg11.

References

1.   Cipollo, J. F.; Trimble, R. B.; Chi, J. H.; Yan, Q.; Dean, N., The yeast ALG11
gene specifies addition of the terminal alpha 1,2-Man to the Man5GlcNAc2-PP-dolichol
N-glycosylation intermediate formed on the cytosolic side of the endoplasmic reticulum.
J. Biol. Chem. **2001,** 276, (24), 21828-40.
2.   Gao, X.-D.; Nishikawa, A.; Dean, N., Physical interactions between the Alg1,
Alg2, and Alg11 mannosyltransferases of the endoplasmic reticulum. *Glycobiology* **2004,**
14, (6), 559-570.
3.   Kelleher, D. J.; Gilmore, R., An evolving view of the eukaryotic
oligosaccharyltransferase. *Glycobiology* **2005.**
4.   Ballou, L.; Hitzeman, R. A.; Lewis, M. S.; Ballou, C. E., Vanadate-resisitant yeast
mutants are defective in protein glycosylation. *Proc. Natl. Acad. Sci. USA* **1991,** 88,
3209-3212.
5.   Dean, N., Yeast glycosylation mutants are sensitive to aminoglycosides. *Proc.
Natl. Acad. Sci. USA* **1995,** 92, 1287-1291.
6.   Frank, C. G.; Aebi, M., ALG9 mannosyltransferase is involved in two different
steps of lipid-linked oligosaccharide biosynthesis. *Glycobiology* **2005,** 15, (11), 1156-63.
7.   Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting
transmembrane protein topology with a hidden Markov model: application to complete
genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.
8.   Abdian, P. L.; Lellouch, A. C.; Gautier, C.; Ielpi, L.; Geremia, R. A.,
Identification of essential amino acids in the bacterial alpha -mannosyltransferase AceA.
J. Biol. Chem. **2000,** 275, (51), 40568-75.
9.   Luzhetskyy, A.; Fedoryshyn, M.; Durr, C.; Taguchi, T.; Novikov, V.; Bechthold,
A., Iteratively acting glycosyltransferases involved in the hexasaccharide biosynthesis of
landomycin A. *Chem. Biol.* **2005,** 12, (7), 725-9.
10.   Glover, K. J.; Weerapana, E.; Imperiali, B., In vitro assembly of the
undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked
glycosylation. *Proc. Natl. Acad. Sci. USA* **2005,** 102, (40), 14255-9.
11.   Szymanski, C. M.; Logan, S. M.; Linton, D.; Wren, B. W., Campylobacter--a tale
of two protein glycosylation systems. *Trends Microbiol.* **2003,** 11, (5), 233-8.

Chapter VI: Peptidyl inhibitors of Oligosaccharyl Transferase and Peptide: *N*-Glycanase

Introduction

The final enzymatic step in the biosynthesis of asparagine-linked glycoproteins is the co-translational transfer of the conserved tetradecasaccharide donor to asparagine residues in nascent polypeptides within the consensus sequence, Asn-Xaa-Ser/Thr, where Xaa is any amino acid other than proline (Scheme VI-1).[1,2] This glycosyl transfer occurs on the lumenal face of the endoplasmic reticulum membrane and is catalyzed by the multi-subunit, membrane-bound enzyme oligosaccharyl transferase (OT).[3,4]

Scheme VI-1. *N*-linked glycosylation reaction catalyzed by oligosaccharyl transferase

In this transformation, OT appears to simultaneously accommodate both polypeptide and oligosaccharide substrates in a ternary complex. At the current time, due to the complexity of this membrane-bound multimeric enzyme, little is known about the active site of OT and how the enzyme may negotiate these large substrates and still turn over efficiently at a rate that is consistent with the rate of protein synthesis.[5,6] Prior work

in the Imperiali group had provided insight into the conformational requirements of the nascent polypeptide at the site of glycosylation. Specifically, using conformationally-constrained substrates, NMR studies suggested that an Asx-type turn was crucial for glycosylation to occur (Figure VI-2).[7] In contrast, examination of X-ray structures of native glycoproteins revealed a preference for β-turns at sites of glycosylation.[8] These results suggested the possibility of a conformational switch of the peptide structure upon glycosylation. Indeed, both fluorescence resonance energy transfer (FRET)[9] and NMR[10] studies have shown that peptides representing natively glycosylated sequences adopt an extended Asx-turn motif in the unglycosylated form, while adopting a tighter β-turn structure in the presence of a β-linked chitobiosyl moiety on the asparagine residue.

Taking into account the potential requirement of an Asx-turn motif for catalysis, a mechanism was proposed that involves protonation of the Asn side chain carbonyl with concomitant deprotonation of the amide nitrogen by an active site base, leading to tautomerization and generation of a more nucleophilic species that is capable of reacting with the activated saccharide donor (Figure VI-1)[11]. Other proposed mechanisms involve either deprotonation of the Asn side chain amide nitrogen by an active-site base to generate a negatively-charged species,[12] or deprotonation of the hydroxyamino acid side chain by an active-site base to facilitate deprotonation of the amide-nitrogen of the Asn side chain by the consequent alkoxide.[13] The latter mechanism is inconsistent with the need for an Asx turn, as the carbonyl of the Asn side chain would be directed away from the core of the turn. Neither of these mechanisms involve tautomerization of the carboxamide, but rather invoke negatively-charged intermediates. The observation that

replacement of the Asn residue with an Asp residue abrogates binding argues against

such an intermediate.[11]

"Asx-turn"   Dol-PP-Oligosaccharide
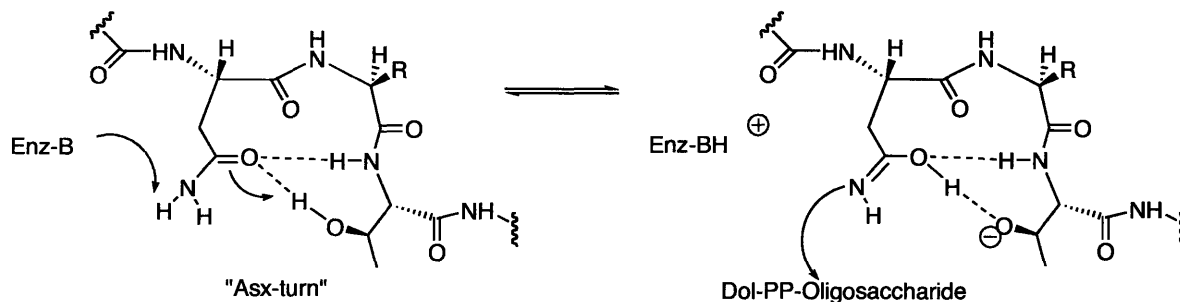
Figure VI-1. Proposed mechanism of OT-catalyzed glycosylation.

As part of an effort to understand the mechanism of asparagine activation, a

number of peptides with modified asparagine surrogates have been prepared.[11, 14, 15] The

tripeptide Bz-Dab-Leu-Thr-NHMe (where Dab = 1,4-diaminobutyric acid) was identified

as a competitive inhibitor of OT with an inhibition constant (K_i) of 1 mM in porcine liver

microsomes, using the corresponding asparagine tripeptide as substrate.[11] Building upon

the discovery that the Asx turn may be important for catalysis, side-chain to main-chain

cyclization was carried out to constrain the tripeptide inhibitor to an Asx turn, leading to

an improvement of the inhibition constant by an order of magnitude $(K_i = 100 \mu M)$, and

at the same time highlighting the need for an Asx turn for catalysis.[16] Extension of the

tripeptide at the C-terminus using the dipeptide Val-Thr greatly improved potency of the

inhibitor,[16, 17] as predicted by the examination of the statistical frequency of amino acids

at these sites in native glycoproteins.[18] The preference of the enzyme for these residues

was later confirmed by evaluating a range of different amino acid combinations at these

sites *in vitro*; none of the dipeptides tested were preferred over Val-Thr.[17] Taking into

account all of these improvements, and including a *p*-nitrophenylalanine residue for

quantitation of peptide concentration, a potent (K_i = 37 nM in yeast) and specific inhibitor, **1** (Figure VI-2), was developed.[16]
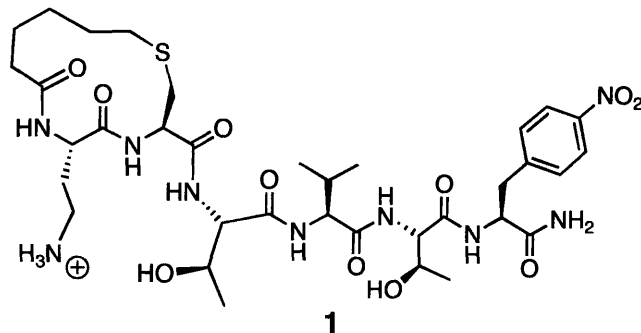
Figure VI-2. Cyclized hexapeptide optimized for OT inhibition (K_i = 37 nM in yeast).

A variety of approaches were envisioned for building upon this progress, including improving potency through the design of neoglycoconjugates and optimization of the Xaa residue of the Asn-Xaa-Ser/Thr consensus sequence, as well as probing the nature of binding and catalysis through isosteric modifications to existing inhibitors. The synthesis of neoglycoconjugates has received considerable attention recently with the development of C- and S-glycoside analogs as well as targets prepared by chemoselective ligation techniques.[19, 20] With the goal of evaluating this class of molecules as product analog inhibitors of OT, a series of neoglycoconjugates was designed based on the use of alanine-β-hydroxylamine (Aβx),[21] alanine β-hydrazide (Aβz),[21] and Dab[11] as asparagine surrogates (Figure VI-3). These residues have been chosen because they react chemoselectively with reducing sugars, without the need for protecting groups or activating agents, to afford N-linked neoglycoconjugates in highly convergent synthetic routes. Also explored was an isosteric C-glycoside that allowed the importance of heteroatoms to be probed.
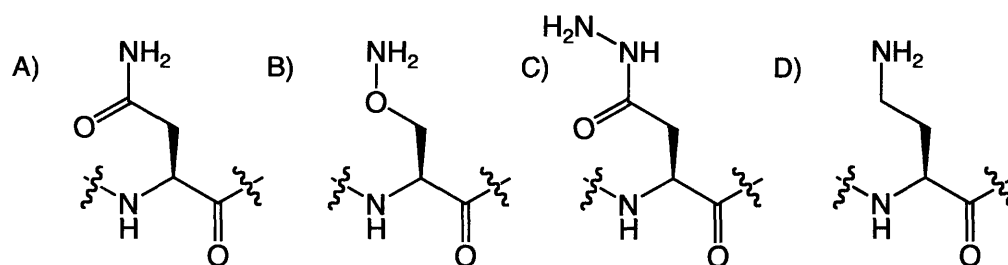
Figure VI-3. Structures of A) Asn, B) Aβx, C) Aβz, and D) Dab side chains.

In addition to the pursuit of more potent inhibitors, there has also been interest in understanding the mechanisms of catalysis and inhibition. Because the γ-amine of Dab is much more likely to be protonated at physiological pH than that of the asparagine side chain, it has been proposed that it is the positive charge that confers the inhibitory activity, by engaging in a salt bridge with a negatively-charged active site residue, and that this interaction compensates for the loss of the Asx-turn stabilization.[11, 14] In other cases of asparagine-mimetic peptides, such as the β-hydroxyasparagine mimetic, in which the amide group is substituted with a hydroxamate group, inhibition may be accompanied by the inhibitor acting as a weak substrate.[14] Our goal was to probe such questions by expanding the available pool of asparagine-surrogate containing peptides modeled after the consensus site for N-linked glycosylation.

OT is tolerant of a variety of substitutions at the central amino acid of the tripeptide consensus sequence Asn-Xaa-Ser/Thr, with proline being the exception. Main-chain to Xaa side-chain cyclization, for example, was not only tolerated, but improved affinity, and in another study, benzophenone was successfully introduced at this position as a photoactivated cross-linking reagent for potental identification of the peptide-binding residues.[22, 23] A more comprehensive study was carried out in which the central residue of

a glycosylation sequon of the rabies virus glycoprotein was varied by site-directed mutagenesis, and glycosylation efficiency was measured.[24] These results suggested that OT is sensitive to the identity of this residue, with effects ranging from inefficient to full glycosylation. As expected, replacement with proline resulted in a complete loss of glycosylation. Additionally, particularly detrimental changes included Trp, Leu, and Asp. These effects, however, may be due to structural perturbations due to other contacts within the protein substrate beyond the sequon, and not simply OT preferences; therefore these results may not be general for all sequons. The use of short peptidyl inhibitors may provide an improved method for probing the discrete preference of OT for the residue at this site, including both natural and non-natural amino acids in an effort to optimize binding affinity.

Having a series of neoglycoconjugate product-based inhibitors of OT at hand enabled an additional study involving an enzyme that uses the product of OT as a substrate. Specifically, the cytoplasmic enzyme, peptide:N-glycanase, or PNGase, catalyzes the hydrolysis of the β-aspartylglycosamine bond in the degradative pathway of misfolded glycoproteins (Scheme VI-2).[25] Interestingly, the bond broken by PNGase is different from the bond formed by the OT-catalyzed reaction. Nascent glycoproteins fold in the ER with the help of an intricate quality-control system that includes lectin-based chaperones and a glucosyltransferase that senses misfolded proteins.[26] When glycoproteins fail to become properly folded, they are exported from the ER and degraded by PNGase.
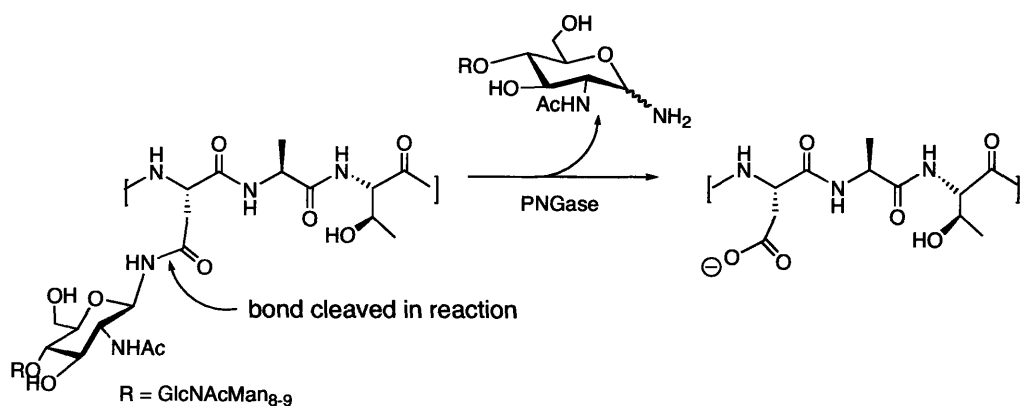
Scheme VI-2. PNGase-catalyzed hydrolysis of *N*-linked glycans catalyzed by PNGase.

PNGase was discovered in *Saccharomyces cerevisiae* (*PNG1*) by the isolation of

a temperature-sensitive mutant, *png1-1*, and glycanase function was validated following

cloning and heterologous expression of the gene in *E. coli*.[27] This soluble and well-

behaved protein is a valuable target for the development of novel inhibitors, as the

perturbation of its function could provide insight into the regulation of quality control in

protein folding. Short glycopeptides are ideal substrates for the screening of putative

inhibitors of PNGase due to their ease of synthesis, provided a minimal saccharide unit is

used, as well as the facility for developing an assay based on the separation of substrates

from products by HPLC. The K_m value of the fetuin-derived glycopeptide Leu-

Asn(NeuAc$_3$Gal$_3$Man$_3$GlcNAc$_5$)-Asp-Ser-Arg for PNGase from mouse fibroblasts was

found to be 114 μM, and this substrate, as well as the asialo form, has been used

extensively for PNGase characterization.[27-31] Previous studies had also shown that in the

case of the bacterial PNGase F and the plant seed PNGase A, the minimum glycan

requirement for hydrolysis was a chitobiosyl unit, which was removed with the same

relative rate as the native glycan from a fetuin-derived glycopeptide.[28] These minimal

binding determinants suggested that synthetically tractable substrate analogs may provide a starting point for development of specific and potent inhibitors.

PNGase has been shown to be weakly inhibited by a variety of glycan structures, including the minimal $GlcNAc_2$, which is a millimolar inhibitor of bacterial, plant, and animal PNGases.[28] Additionally, a known caspase inhibitor, carbobenzyloxy-Val-Ala-Asp-α-fluoromethylketone (Z-VAD-fmk), was found to be a potent, specific, and cell-permeable inhibitor of yeast and mammalian PNGase, having an *in vivo* apparent IC_{50} value of 7 μM.[32] However, it has been pointed out that ER stress, which is linked to PNGase function, can also be associated with caspase activation; thus, results from PNGase inhibition studies that involve the use of a caspase inhibitor should be interpreted carefully.[33] The development of more specific inhibitors would therefore be useful for *in vivo* studies of PNGase function.

Results and Discussion

Effects of conformational preferences in neoglycoconjugate binding to OT

The substrate peptide Bz-Asn-Ala-Thr-Val-Thr-Nph-NH_2 (2) (Table VI-1) was synthesized by standard solid-phase peptide synthesis methods, and the K_m for yeast OT was determined to be 0.31 μM. Comparison of 2 to the corresponding Asn(GlcNAc)-containing peptide (2a) (Table VI-1), which has a K_i value of 100 μM for yeast OT, reveals a loss in binding affinity of over two orders of magnitude when the saccharide is included. Such a drastic change in binding affinity is consistent with the observations that glycosylated peptides tend to favor β-turns,[10] and that the Asx turn is a necessary

motif for binding and catalysis.[7] This situation may suggest a mechanism for favorable

product release, which is needed to ensure that the rate of glycosylation keeps up with the

rate of protein synthesis.
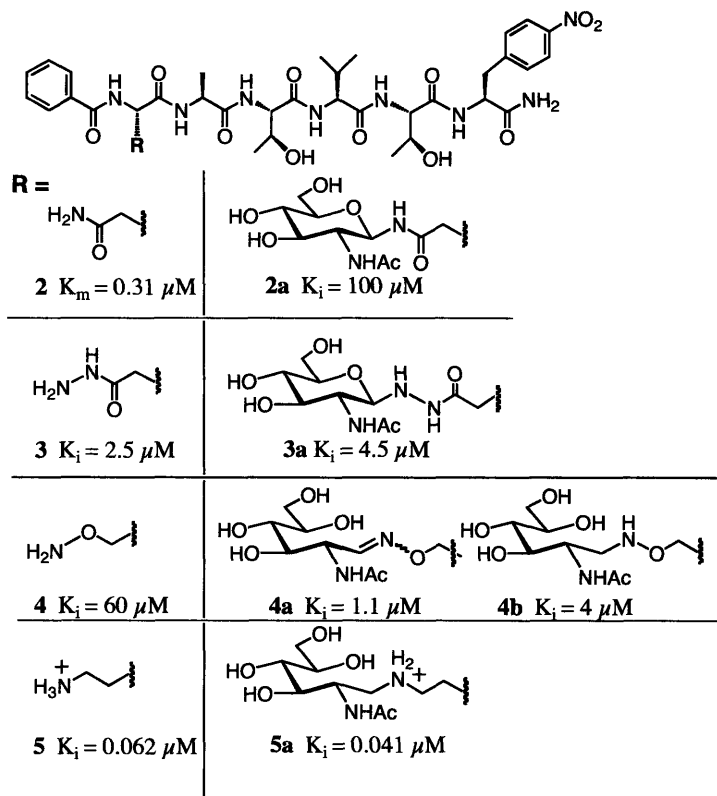
Table VI-1. Structures and kinetic data for neoglycoconjugates and precursors.

Isosteric molecules mimic the geometries and spatial locations of natural

compounds. As such they represent new compounds that can be used to probe the effects

of certain atoms or electronic effects on biological activity. A series of

neoglycoconjugates containing amide isosteres were then prepared and the inhibitory

activity of each was compared to each non-glycosylated analogue. This study was

carried out in collaboration with Drs. Maria de L. Ufret and Stephane Peluso, and the

details of the synthesis and kinetic evaluation can be found in Peluso, et al.[21, 34] These

182

results are summarized in Table VI-1. Glycoconjugates in which the amide was replaced with more flexible linkers such as hydrazide (**3a**), hydroxylamine (**4a-b**), or amine (**5a**) groups did not suffer the same loss in affinity that was observed for the native amide bond glycoconjugate (**2a**). In the case of neoglycoconjugate **4a**, significant improvement in inhibitory activity was even achieved compared to the parent peptide **4**.[34]

Previous studies have shown that N-δ-alkylation of the Dab side chain with a bulky, aromatic naphthyl group improves on already potent inhibitory activity.[35] This effect is in stark contrast with modification of the native asparagine side chain of OT substrates, where a GlcNAc residue drastically reduces binding to the enzyme relative to the unsubstituted form. Considering this observation, and in light of the results described, a hypothesis was proposed that amide bond cis-trans isomerization of a substituted asparagine side chain about the C-N bond might result in a steric clash, disrupting the Asx turn and favoring product release (Figure VI-4).
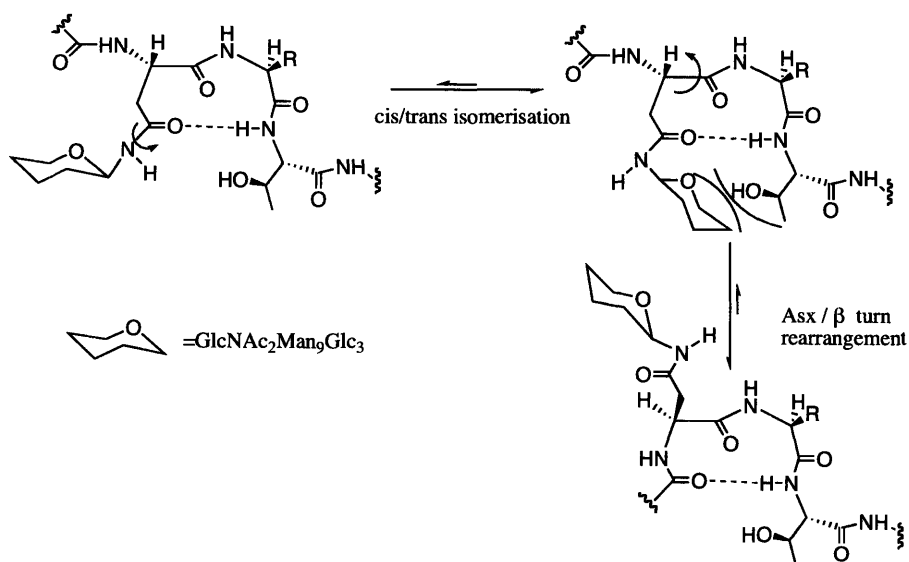
Figure VI-4. Proposed model of Asn amide bond cis-trans isomerization.

183

This model could explain how the absence of the carboxamide carbonyl group with restricted rotation enables accommodation of a naphthyl moiety in a position that does not tolerate substitutents when the carboxamide is present. It may also explain how glycosylated proteins are readily released from the active site, to keep up with the rate of protein synthesis, and is in agreement with NMR evidence that glycopeptides do not adopt an Asx-turn, but rather more typically adopt a β-turn.

Based on the proposed model of amide bond cis/trans isomerization, a C-glycoside was expected to have favorable properties as a product-based inhibitor. The advantages of this analog over the previously described neoglycoconjugates include the replacement of the asparagine amide with the more flexible isosteric methylene spacer, and the preservation of the closed-ring (pyranose) form of the sugar. The C-glycoside residue was incorporated as an Fmoc-protected free acid[36] into both a linear (6) and a constrained (7) hexapeptide (Figure VI-5). Inhibition assays revealed K_i values of 9.2 μM and 6.7 μM, respectively.

Figure VI-5. Structures and inhibition constants for linear (6) and cyclized (7) C-glycoside.

A constrained version of the hexapeptide containing a Dab(GlcNAc) surrogate, c[Hex-Dab-Cys]-Thr-Val-Thr-Nph-NH$_2$ (**8**) (Figure VI-6), was prepared by reductive amination for a more direct comparison with the cyclized C-glycoside peptide. In this case, a K$_i$ value of 0.026 μM was found for **8**.

$$K_i = 0.026 \ \mu M$$
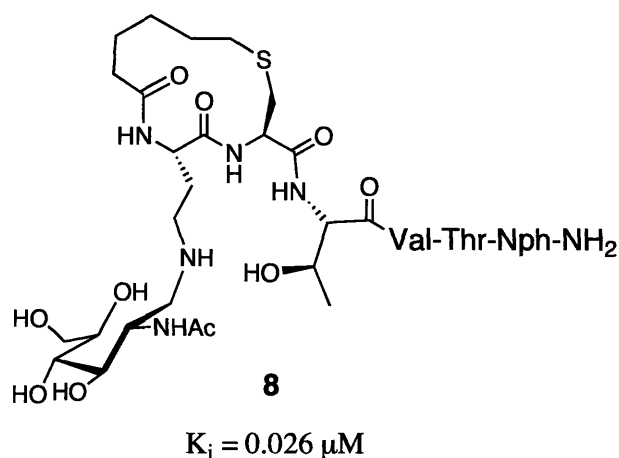
Figure VI-6. Structure and inhibition constant of cyclized Dab-GlcNAc conjugate (**8**).

These results indicate that incorporation of a more flexible, isosteric, all-carbon linker and preservation of the closed form of the sugar confers no improvement over any of the neoglycoconjugates described here. The linear (**5a**) and cyclized (**8**) Dab(GlcNAc) peptides are two orders of magnitude more potent than the linear (**6**) and cyclized (**7**) C-glycoside counterparts, respectively. A direct comparison is precluded by the fact that the saccharide moiety of the Dab(GlcNAc) side chain is locked into the open-ring form, although an even lower inhibition constant may be expected for the Dab(GlcNAc) inhibitors if the saccharide were locked into the closed-ring (pyranose) form. These

results highlight the dramatic effect that the positive charge of the Dab side-chain amine has on inhibition, especially considering that the binding of the C-glycoside is so similar to that of the hydrazide and hydroxylamine neoglycoconjugates, which bear the δ-nitrogen, but are more likely to be neutral under physiological conditions.

Role of side-chain charge in OT inhibition by asparagine-mimetic peptides

In order to test the hypothesis that the positive charge on the Dab side chain is critical in conferring potent inhibitory activity, design of an isosteric side chain that is neutral was envisioned. The functionality chosen was a hydrazine side chain, which is isosteric with Dab, but perturbs the pK_a of the terminal amine such that it is a neutral species under physiological conditions (Figure VI-7).
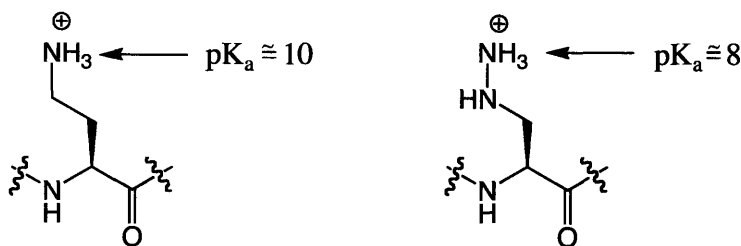
Figure VI-7. Comparison of pK_a values for Dab and hydrazine side chains.

The β-hydrazine analogue would also enable the future synthesis of neoglycoconjugates bearing a stable hydrazone linkage[37] through chemoselective ligation.[21] Such a conjugate would be an improvement over the less stable amine linkage of Dab glycoconjugates, which must be reduced in order to create a stable linkage to the saccharide moiety. The necessary building block for such a surrogate is an amino acid with a β-hydrazine side chain bearing protecting groups on each of the three nitrogens.

The β-hydrazine amino acid (9) was prepared by a four-step synthesis starting from L-serine (10) (Scheme VI-3). Identical protecting groups could be installed at the two nitrogen atoms of hydrazine, but the α-amine required an orthogonal protecting group. Therefore, the first step was to protect the α-amine of L-serine with Alloc.[21] This reaction was done under Schotten-Bauman conditions with allylchloroformate and 4 N NaOH to afford 11. Following this step, Alloc-serine (11) was cyclized by a Mitsunobu-type reaction to provide the β-lactone (12),[38] and the 4-membered lactone was then opened with *t*-butylcarbazate to give N^1-Alloc, N^3-Boc protected β-hydrazino amino acid (13a).[39] Selective attack of the ring methylene carbon can be achieved by using specific solvent polarity and order of addition.[38] However, a minor product (13b) can also be formed from opening of the β-lactone at the carbonyl. Unfortunately these compounds are difficult to distinguish by ^1H and ^{13}C NMR spectra, and therefore further characterization was necessary to ensure that 13a was the major product.
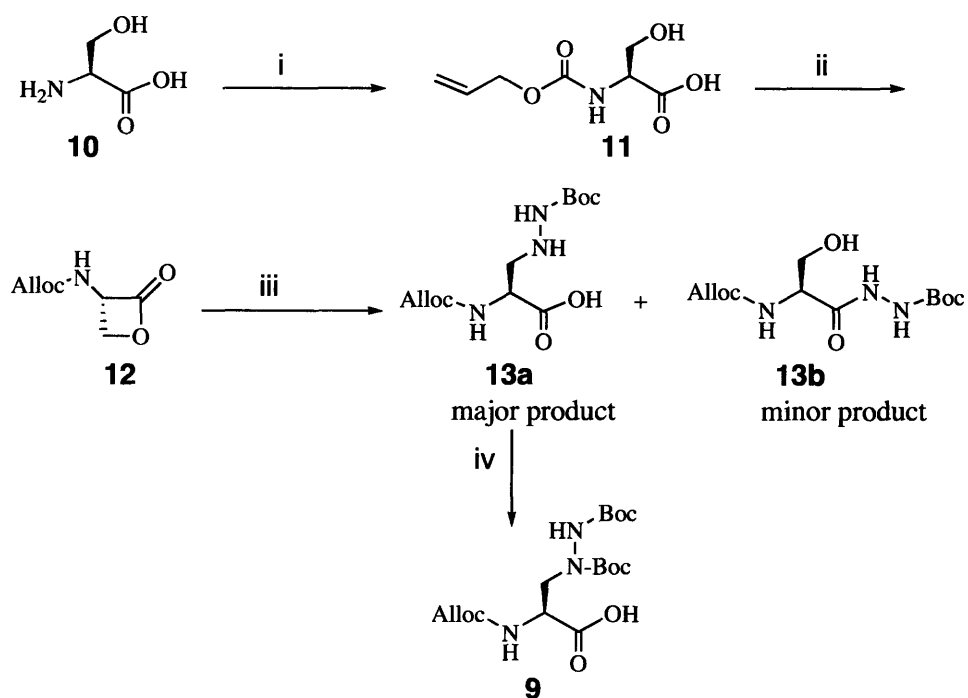
Scheme VI-3: Synthesis of the β-hydrazine amino acid[a]

[a]Reagents and conditions: (i) allylchloroformate, 4 N NaOH, added simultaneously at

0°C under Schotten-Bauman conditions, then warmed to room temperature over 1 hr,

43% yield; (ii) 1) PPh$_3$, 2) DEAD, 3) **11**, -75°C to room temperature over 2 hr in THF,

29%; (iii) *t*-butylcarbazate, anhydrous CH$_3$CN, room temperature, 20 hr, 39% yield major

product; (iv) Boc anhydride, DIEA, H$_2$O: dioxane (2:1), room temperature, 18 hr, 64%

yield.

In order to confirm the presence of the free carboxylic acid, a common coupling

reagent in solid phase peptide synthesis, HBTU, was reacted with the major product **13a**.

This coupling reagent would not react with **13b**, but would only activate **13a**. Indeed, a

new product was formed, suggesting that the activation had occurred. Further evidence

was later provided when the isolated major product was successfully incorporated into a peptide.

Finally, Boc-anhydride was used to protect the proximal amine of the hydrazine. The resulting amino acid (9) was analyzed for racemization using Marfey's reagent (1-fluoro-2,3-dinitrophenyl-5-L-alanine amide), from which only one product from the HPLC separation was found to have the mass of the adduct. Compound 9 was thus used as a building block for the synthesis of 14 (Figure VI-8), using standard solid-phase peptide synthesis techniques. The only exception involved the presence of an Alloc protecting group instead of the standard Fmoc protecting group on the main chain amine of the building block. The procedure for removal of the Alloc group was used as described in "Experimental Procedures" for deprotection.

$K_i = 30 \ \mu M$

14
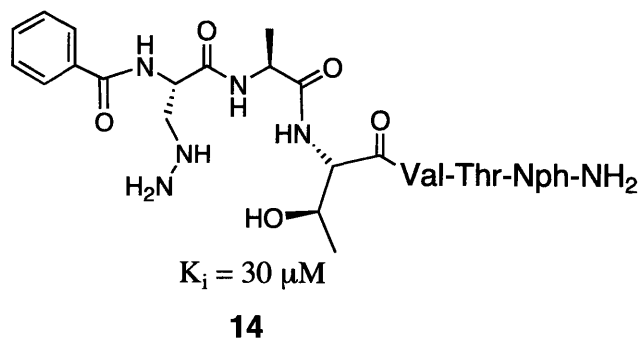
Figure VI-8. Structure and K_i value of the hydrazinopeptide inhibitor 14

Evaluation of the inhibitory activity against yeast OT revealed a K_i of 30 μM, which is three orders of magnitude lower in potency compared to the corresponding Dab-

containing hexapeptide inhibitor (5a) (Table VI-1). Thus, as postulated, charge does seem to have an important role in binding.

Comparison of Asn amide *N*-hydroxylation to *N*-amination in OT inhibition

To further investigate the inhibition of OT by the hydrazide asparagine surrogate peptide (3) (Table VI-1), an analog was designed for comparison that bears a hydroxamate side chain in place of the hydrazine, effectively exchanging the primary amine of the hydrazide with a primary alcohol (Figure VI-9A). A tripeptide version of the hydroxamate analog had previously been shown to be a weak substrate of OT,[14] and therefore one goal with the hydrazide and hydroxamate inhibitors was to evaluate them as OT substrates. Hydroxamate peptide 15 (Figure VI-9B) was prepared by first building the benzoyl-capped hexapeptide with Fmoc-Asp(OAlloc)-OH installed as a precursor to the hydroxamate residue. Selective deprotection of the Alloc group followed by the coupling of hydroxylamine with PyAOP and DIEA yielded the final peptide 15.
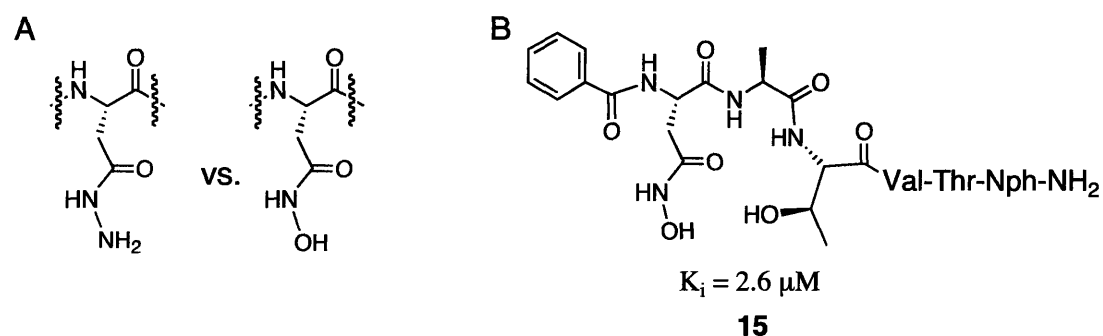
Figure VI-9. A) Structures of hydrazide and hydroxamate side-chain functionalities. B) Structure and inhibition constant of hydroxamate surrogate peptide 15.

Evaluation of the inhibitory activity of **15** revealed a K_i value of 2.6 μM, which is strikingly similar to that of the hydrazide analog (**3**) (2.5 μM). The substrate capabilities of these two inhibitors were found by using a modification of the radioactive assay described in the previous section. The substrate peptide, Bz-Asn-Ala-Thr-Val-Thr-Nph-NH$_2$ (**2**) (Table VI-1) was also included for comparison. Following incubation with OT, peptides were separated from (^{3}H)GlcNAc-GlcNAc-PP-Dol by aqueous/organic extractions and analyzed by HPLC, using the radiolabel to detect the glycopeptide products and the UV absorbance at 228 nm to monitor elution of the parent peptides (Figure VI-10). The UV trace is not sensitive enough to detect the low turnover of peptide to glycopeptide, and therefore fractions were collected and subjected to scintillation counting. The shift in retention time of the glycopeptide product relative to the starting material peak in the UV trace is consistent with the addition of the polar saccharide. Minor impurities observed in the UV trace were expected due to the yeast microsomes used in the assay. The results of these assays reveal that both peptides **3** and **15** exhibit weak substrate capability.
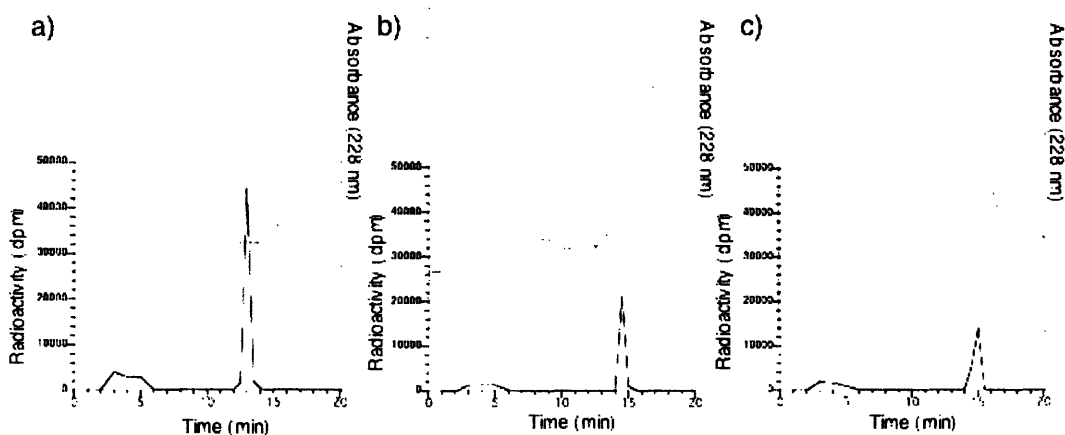
Figure VI-10. UV (top) and radioactive (bottom) HPLC traces of OT substrate assays. A) natural substrate (**2**), B) hydrazide inhibitor (**3**), and C) hydroxamate inhibitor (**15**) as substrates.

Optimization of the central residue of the Asn-Xaa-Ser/Thr consensus sequence

Alanine is a commonly used residue in the central position of the tripeptide consensus sequence, Asn-Xaa-Ser/Thr, of many inhibitors and substrates. A study was carried out to ascertain whether minor modifications to this side chain would have any impact on the binding affinity of OT inhibitors. Bz-Dab-Val-Thr-Val-Thr-Nph-NH$_2$ (**16**) and Bz-Dab-Nva-Thr-Val-Thr-Nph-NH$_2$ (**17**) (Nva = norvaline) (Figure VI-11) were synthesized by standard solid phase peptide synthesis using Fmoc-protected amino acids and were evaluated as inhibitors of OT. Compared to the analogous hexapeptide **5a** bearing an alanine residue at the central position, which has a K$_i$ of 62 nM (Table VI-1),[34] both of these modified peptides have similarly improved potency, with K$_i$ values of 14 nM and 11 nM for **17** and **18**, respectively. The tolerance of OT for a variety of residues

at this position suggests that there is considerable space available for binding, and thus filling this space with larger aliphatic groups may increase binding energy.
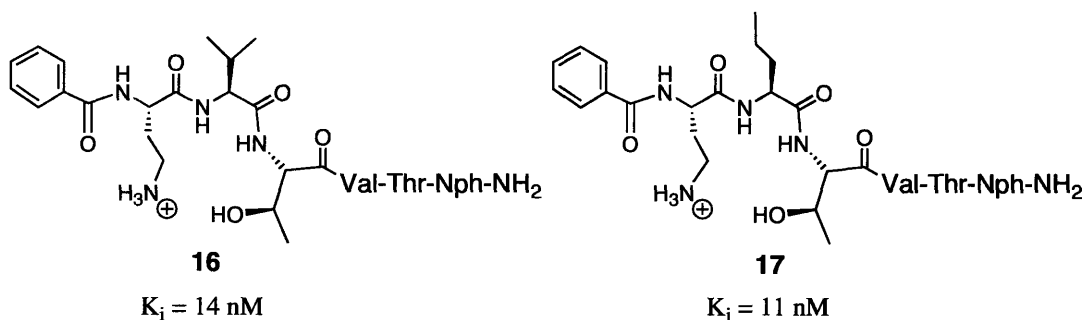
Figure VI-11. Structures and inhibition constants of peptides **16** and **17**.

Activity and inhibitor assays of PNGase1

The product of OT is also the substrate of the deglycosylating enzyme, PNGase, and therefore product-based OT inhibitors such as the neoglycoconjugates described here can also be valuable tools for the study of PNGase. Also facilitating this study was the availability of the fluorescein-labeled interferon fragment glycopeptide (Figure VI-12) that had been prepared to test the substrate specificity of Alg1 (Chapter III), and here could be used as a substrate for screening PNGase inhibitors. The fluorophore enables a sensitive HPLC assay based on the shift in retention time upon deglycosylation.

193
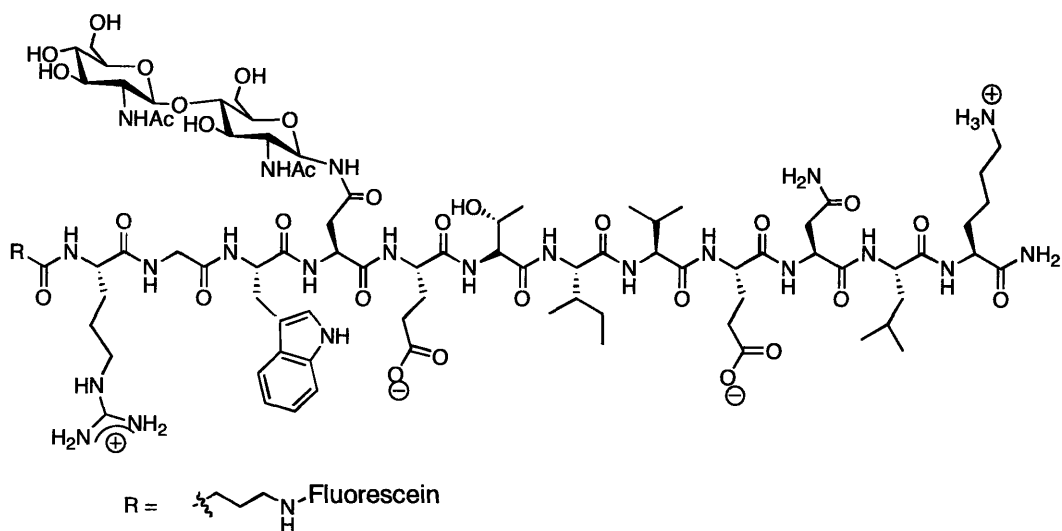
Figure VI-12. Synthetic glycopeptidyl Png1 substrate.

Hexapeptide neoglycoconjugates **18, 19**, and **20** (Figure VI-13), prepared by reductive amination with GlcNAc or GlcNAc$_2$, were synthesized as OT inhibitors by Dr. Stephane Peluso using standard solid-phase peptide synthesis in conjunction with chemoselective ligation methods described elsewhere.[21] These product-mimetic OT inhibitors were postulated to be excellent candidates for substrate-mimetic inhibitors of PNGase, considering the presence of both the peptide and glycan binding determinants, as well as the fact that they are coupled via non-hydrolyzable linkers.
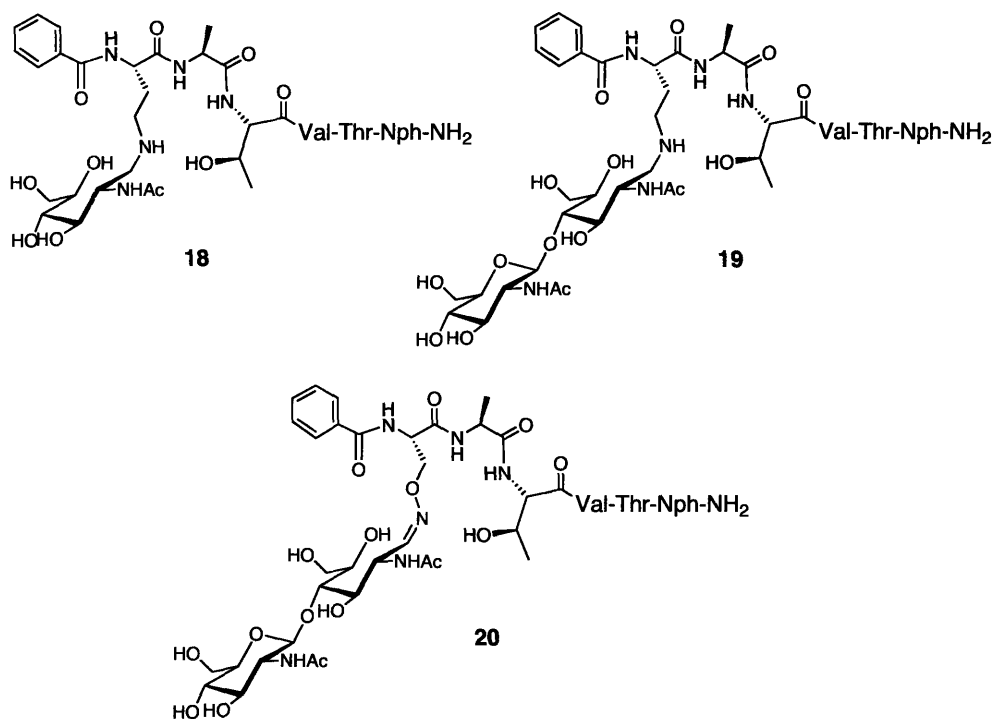
Figure VI-13. Structures of neoglycoconjugate Png1 inhibitors (**18-20**).

PNG1 from *S. cerevisiae* had been cloned into a bacterial expression vector[27] and was provided by the laboratory of Professor William Lennarz. This enzyme was overexpressed and purified to homogeneity by Talon™ affinity chromatography, with a yield of approximately 6 mg per liter of cell culture. The purified enzyme is stable at 4 °C for several months. This preparation was used in hydrolysis assays with the GlcNAc$_2$-linked interferon fragment peptide (Figure VI-12). Reaction progress was monitored by analytical C$_{18}$ HPLC with UV detection, which showed a change in retention time of approximately 2 minutes upon deglycosylation. Percent conversion was measured by dividing the area under the product peak by the sum of the areas under the starting material and product peaks. Mass spectrometry (ESI-MS) was employed to confirm the identity of the starting material and product peaks. The K$_m$ of the

195

fluorescein-labeled glycopeptide used in this study for Png1 is not known, because the range of concentrations necessary to determine that value would be undetectable in this preliminary assay.

Incubation of the fluorescein-labeled glycopeptide (100 μM) in the presence of 3.3 pM Png1 resulted in 15% conversion to the cleaved product after 12 hours at 30 °C. Using these same conditions, but preincubating the enzyme with **18, 19,** or **20** (10 μM) for 30 minutes resulted in a drop in percent conversion to 5%, 9%, and 3%, respectively. From these preliminary data, it appears that all of these neoglycoconjugates have the ability to bind and inhibit Png1.

In order to obtain quantitative data on these inhibitors, the use of a fluorescence detector is necessary. Because fluorescein fluorescence is pH-sensitive, and the acidic conditions used in HPLC (0.1% TFA) are enough to significantly quench the fluorescence, a new substrate was designed. Rhodamine was chosen as the fluorophore, which is not sensitive to lower pH, and is commercially available as the *N*-hydroxysuccinimide ester, making it amenable to solid-phase peptide synthesis. The peptide substrate chosen was based on a fetuin-derived glycopeptide Leu-Asn(NeuAc$_3$Gal$_3$Man$_3$GlcNAc$_5$)-Asp-Ser-Arg. Peptide **21** (Figure VI-14) was synthesized by standard solid-phase peptide synthesis, using the same Asn-GlcNAc$_2$ building block described in Chapter III. Prior to cleavage from the solid support, rhodamine-NHS was coupled to the aminohexanoic acid spacer. This peptide, in the absence of the spacer and rhodamine, was previously shown to be hydrolyzed by both bacterial and plant seed PNGase with the same relative rate as the native fetuin-derived glycopeptide.[28] Given this precedent in conjunction with the demonstration that yeast

Png1 accepts the $GlcNAc_2$-interferon fragment glycopeptide, $GlcNAc_2$-fetGPI was postulated to be an excellent substrate for PNG1 inhibition assays. Also, since PNGase is meant to act on a wide variety of peptide sequences, and a 6-carbon spacer has been included, the rhodamine is not suspected to greatly affect binding and amidase activity. The availability of this substrate will enable future quantitative evaluation of neoglycoconjugates as inhibitors of Png1.
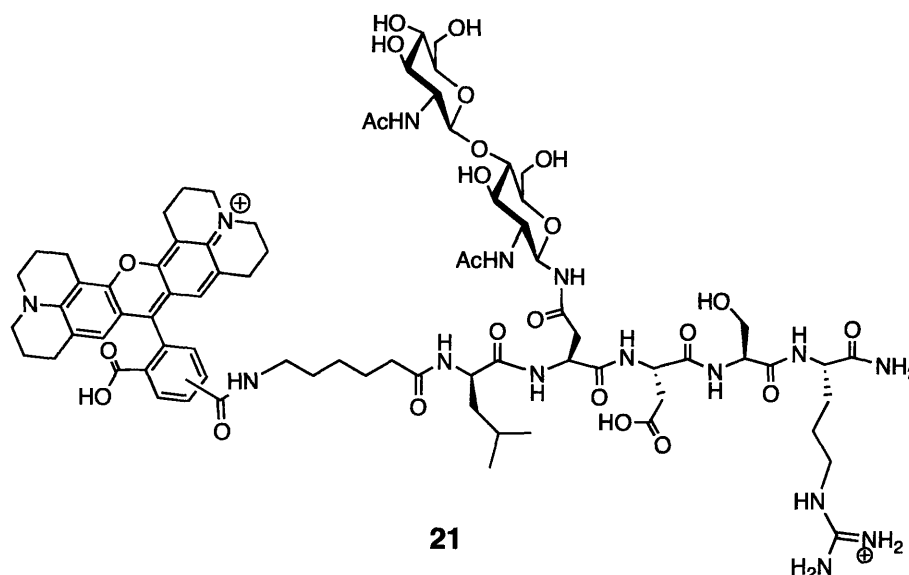

21

Figure VI-14. Structure of the rhodamine-labeled fetuin-based Png1 substrate.

Conclusion

Appending saccharide moieties to asparagine-mimetic side chains in order to make product-based inhibitors did not dramatically improve potency in general, although this study did lead to insights into the possible involvement of cis-trans isomerization in

product release. In the OT inhibitors described, flexible non-amide linkers can tolerate functionalization of the side chain without significant changes in binding affinity. This result is in contrast to the native Asn side chain, which drastically loses binding affinity upon addition of a GlcNAc moiety to the amide nitrogen. However, these findings are reminiscient of the discrepancy between the tolerance of OT for a naphthyl moiety on a Dab side chain and the sensitivity of OT binding to even minor functionalization of the Asn side chain.

As had been previously hypothesized, the positive charge of the side-chain amine on Dab appears to play an important role in the binding affinity of OT inhibitors. This effect was demonstrated in the comparison of amine and hydrazine side chains as asparagine mimics, and in the comparison of the C-glycoside to the Dab-glycoside.

It has been postulated that weak substrate behavior may be the source of inhibitory activity for OT substrate mimics bearing either a hydroxamate and hydrazide side chain in place of the asparagine amine. Indeed, both analogues were found to be low micromolar inhibitors as well as substrates of OT. The reduced substrate behavior of these modified asparagine analogs may be due to impairment of the deprotonation of the amide nitrogen, or, considering the cis/trans isomerization model, reduced binding affinity resulting from disruption of the Asx turn by an additional substituent directed toward the interior of the turn.

Improvements to the binding affinity of Dab-containing inhibitors with modifications to the central site of the consensus sequence suggests that there is much more opportunity for optimization of inhibition by varying the substitution of this site.

Neoglycoconjugates were also found to be inhibitors of the deglycosylating enzyme, Png1. Future directions may include a more quantitative description of this inhibitory activity, as well as improvements to potency and bioavailability. Such molecules could be extremely valuable in probing the intricate pathways of protein folding quality control and degradation.

Taken together, these results show that isosteric surrogates of natural substrates and products can be useful tools for understanding catalysis, inhibition, and product release. Seemingly minor changes in substrates and inhibitors can have dramatic consequences, and have led to a more thorough understanding of the active site of OT, which still remains elusive due to the complicated nature of this enzyme complex.

Acknowledgements

Experimental procedures

General procedure for solid phase peptide synthesis

All peptides were prepared manually using a glass reaction vessel fitted with a sintered glass frit. Fmoc-PAL-PEG-PS resin was washed and swollen in CH_2Cl_2 (2 x 10 mL/g resin x 15 min) and dimethylformamide (DMF) (1 x 10 mL/g resin x 15 min). N-α-Fmoc was removed by treatment with piperidine : DMF 1:5 (3 x 10 mL/g resin x 5 min), followed by washing with DMF (5 x 10 mL/g resin x 1 min). The number of equivalents was determined by quantification of Fmoc release. This was accomplished by measuring the UV absorbance at 300 nm of the pooled deprotection rinses and DMF washes. Coupling reactions were performed using 1.5 equivalents of N-α-Fmoc protected amino acids activated *in situ* with 1.5 equivalents of PyBOP and 3 equivalents of diisopropylethylamine (DIEA) in DMF (10 mL/g resin) for 1 hour.

General procedure for O-Allyl removal on solid support

For the removal of the Allyl protecting group from the aspartic acid side chain, the resin was treated with 25 equivalents of $PhSiH_3$ in CH_2Cl_2 (1 mL/g resin) under nitrogen for 2 min, followed by the addition of 0.25 equivalents of $(PPh_3)_4Pd(0)$ in CH_2Cl_2 (1 mL/g resin). The reaction proceeded for 30 min under nitrogen and then was washed with CH_2Cl_2 (2 x 5 mL x 1 min). This procedure was repeated twice more, then followed by washing with CH_2Cl_2 (2 x 10 mL/g resin x 1 min).

General procedure for N-terminal capping

Capping was accomplished by reacting the resin with 10 equivalents of benzoic anhydride and 10 equivalents of pyridine in DMF (2.5 mL/g resin) for 1 hour. Resin was then washed with DMF (5 x 10mL/g resin x 1 min) and CH_2Cl_2 (2 x 10mL/g resin x 1 min).

General procedure for side chain to main chain cyclization
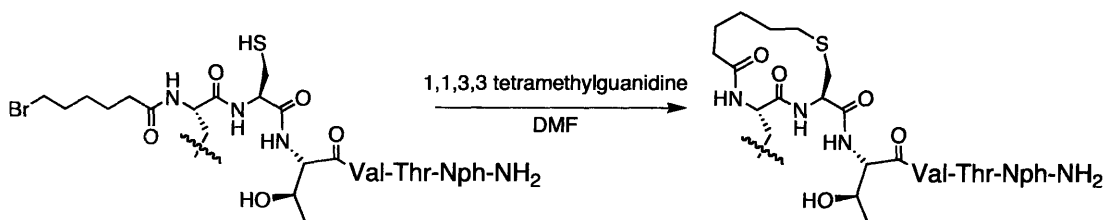
Following the coupling of the last amino acid and removal of Fmoc, 10 equivalents of 6-bromohexanoic acid and 10 equivalents of PyBOP were added to the resin in DMF. The coupling was initiated by the addition of 20 equivalents of diisopropylethylamine, followed by orbital shaking for 12 hours under nitrogen. The monomethoxytrityl (Mmt) protecting group was removed from cysteine by treatment of the resin with 1% trifluoroacetic acid in CH_2Cl_2 (5 X 1 min). DMF was added to the resin (1 mL/1 g resin) and degassed. 5 equivalents of freshly distilled 1,1,3,3-tetramethylguanidine was added and the reaction shaken orbitally under nitrogen. After 12 hours, the resin was washed with DMF (5 X 1 min) and CH_2Cl_2 (2 X 1 min).

General procedure for peptide release and deprotection

Cleavage of peptides from resin with concomitant removal of side chain protecting groups was performed using 10 mL/g resin of trifluoroacetic acid (TFA):

201

CH_2Cl_2: triisopropylsilane: H_2O (90:5:2.5:2.5). The cleavage reaction was allowed to proceed for 1 hour. The mixture was filtered through a pasteur pipet equipped with a glass wool plug, and resin was rinsed with TFA (2 x 1 mL). Filtrate was concentrated under a stream of nitrogen to 1 mL, then triturated with 14 mL of diethyl ether. The white solid obtained was isolated by centrifugation, washed twice with ether, and allowed to dry. Dab-containing peptides were redissolved in acetonitrile: H_2O (1:1), and all others in acetonitrile: DMSO: acetic acid (5:3:2). All peptides were purified by preparative HPLC.

Bz-AsnAlaThrValThrNph-NH$_2$ (2)
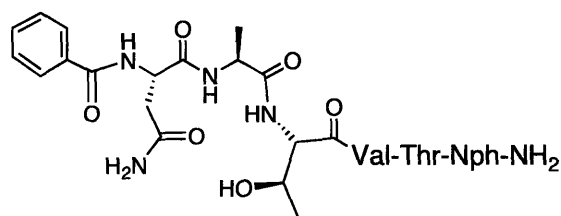
2

The immobilized peptide was assembled according to the general procedures described above.

HPLC : t_R=19.9min (7-100% CH_3CN/0.1%TFA in 30 min)

LRMS calcd for **2** [$C_{36}H_{49}N_9O_{12}$ + H]$^+$ requires m/z 800.8. Found 800.3 (ESI+)

Bz-C-glycoside-Ala-Thr-Val-Thr-Nph-NH$_2$ (6)
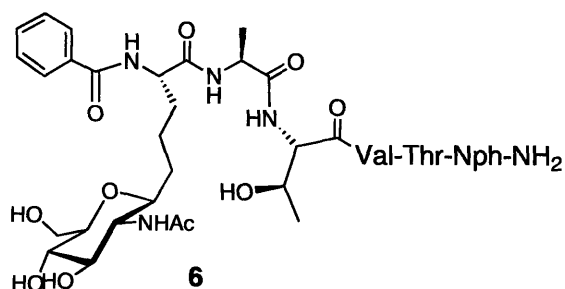
Fmoc-Ala1-Thr-Val-The-Nph-PALPEG-PS was prepared as described, and 150 mg of resin-bound peptide was used for incorporation of the C-glycoside. The exact amount of peptide was determined to be 0.017 mmol by quantifying the amount of Fmoc released upon deprotection as described above. Following deprotection, 11 mg of the C-glycoside was dissolved in 0.5 mL of DMF, and the purity was confirmed by TLC, which was developed with chloroform: methanol: water (65: 25: 4) and stained with PMA. PyAOP (18 mg) was dissolved in 0.5 mL of DMF and added to the resin in addition to the Fmoc-protected C-glycoside in DMF. A 10-μL portion of collidine was added and the reaction proceeded with shaking overnight at room temperature. N-terminal benzyl capping and cleavage from the resin with concomitant deprotection of side chain protecting groups were as described.

Prior to HPLC purification, deprotection of the peracetylated GlcNAc was performed. To a solution of methanol/0.3% (final) NaOMe (2.7 mL), 0.3 mL of cleaved peptide in DMSO was added dropwise with constant stirring in a 10-mL pear-shaped flask. The reaction was allowed to stir at room temperature overnight. The reaction

mixture was brought to pH 4 with glacial acetic acid (10%), and solvent was removed on a rotary evaporator.

HPLC : $t_R=20.7$ min (7-100% $CH_3CN/0.1\%TFA$ in 30 min)

LRMS calcd for **6** $[C_{45}H_{65}N_9O_{16} + H]^+$ requires m/z 988.5. Found 988.3 (ESI+)

(cyc)-C-glycoside-Cys-Thr-Val-Thr-Nph-NH₂ (7)
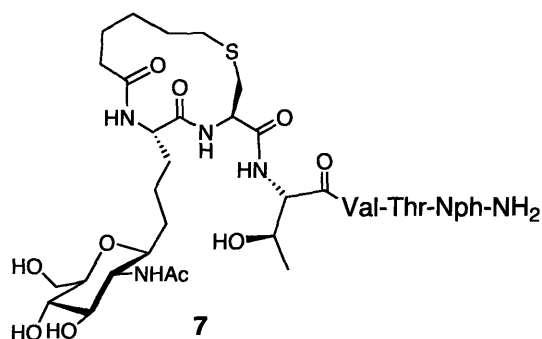
Synthesis was as described for linear C-glycoside, with the exceptions that the Ala is replaced by a Cys(MMT), and the peptide is capped with 6-bromohexanoic acid and cyclized as described above.

HPLC : $t_R=$ 20.0 min (C_{18}, 7-100% $CH_3CN/0.1\%TFA$ in 30 min)

LRMS calcd for **7** $[C_{44}H_{69}N_9O_{16}S + H]^+$ requires m/z 1012.5. Found 1012.5 (ESI+)

(cyc)-Dab(GlcNAc)-Cys-Thr-Val-Thr-Nph-NH₂ (8)

8

Dab(Boc)-Cys(Mmt)-Thr(OtBu)-Val-Thr(OtBu)-Nph-PAL-PEG-PS was prepared

by standard methods, and capping and cyclization followed as described above.

Following cleavage from the resin crude peptide was isolated. To a 63 mg/mL solution

of the crude peptide in DMF (11 mg total) was added 8 μL of acetic acid (4% final) and

13 μL of a 1 M solution of N-acetylglucosamine in water. This mixture was stirred at

70 °C for two hours, then 80 μL of NaBH₃CN (0.1 mg/mL) were added. The cyclized

glycopeptide was purified by standard HPLC procedures.

$C_{44}H_{69}N_9O_{16}S$ (MW= 1013.12, exact mass = 1012.45)

LRMS calcd for **8** $[C_{44}H_{69}N_9O_{16}S]^+$ requires m/z 1015.5. Found 1015.5 (ESI+)

N-Alloc-serine (11)

A solution of L-serine (1.1 g, 10.0 mmol) in 4 N NaOH (2.5 mL) was cooled to

0 °C in an ice water bath. With rapid stirring, allylchloroformate (1.1 mL, 10.0 mmol)

and 4 N NaOH (2.5 mL) were simultaneously added dropwise to serine. The reaction

mixture was removed from the ice bath and allowed to warm to room temperature with continuous stirring. An hour after removing from the ice bath, the reaction mixture was diluted with H_2O (50 mL), and washed with diethyl ether (3 x 50 mL). The aqueous layer was adjusted to pH=3 with 5N HCl, then washed with ethyl acetate (5 x 50 mL). Pooled organic layers were dried over sodium sulfate, filtered, and dried under vacuum for 24 hours, removing residual ethyl acetate to yield 820 mg (2.9 mmol) of compound **11** as a clear oil (43% yield). ^1H NMR (400 MHz, $CDCl_3$): δ 3.88 (d, *J*=8.8 Hz, 1H) CH*H*-OH, 4.06 (d, *J*=10.1 Hz, 1H) C*H*H-OH, 4.40 (m, 1H) αH, 4.59 (m, 2H) CH_2=CH-C*H*$_2$-O-, 5.22 (d, *J*=10.3 Hz, 1H) C*H*H=, 5.32 (d, *J*=17.2 Hz, 1H) CH*H*=, 5.90 (m, 1H) CH_2=C*H*-. R_f = 0.52 in $CHCl_3$:MeOH:H_2O (65:25:4)

N-Alloc-serine-β-lactone (12)

11   **12**
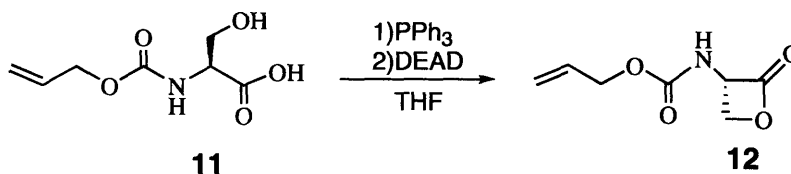
PPh_3 (2.1g, 7.9 mmol) was dissolved in anhydrous THF (50 mL) at −70 °C under nitrogen in a tri-neck round bottom flask. DEAD (1.2 mL, 7.9 mmol) was added over 10 minutes. The solution was then stirred for an additional 10 minutes. One septum on the flask was quickly replaced with a dropping funnel containing *N*-Alloc-serine (1.5 g, 7.9 mmol) in anhydrous THF (25 mL). This solution was added to the reaction mixture over 30 minutes at −70 °C. The reaction was stirred for 20 minutes more at -70° C, then brought to room temperature slowly over 1 hour with stirring. The reaction mixture was concentrated on a rotary evaporator and purified by flash chromatography (5-cm

diameter, 130 g silica). 1 L of hexane: ethyl acetate (85:15) was used to pack the column

and rinse impurities. The product was eluted with 1 L of hexane: ethyl acetate (7:3)

followed by 500 mL of hexane: ethyl acetate (65:35). Compound **12** was recovered as a

white solid in 29% yield. ^{1}H NMR (400 MHz, CDCl$_3$): δ 4.45 (d, J=6 Hz, 2H) -CH_2-,

4.57 (d, J=5.6 Hz, 2H) CH$_2$=CH-CH_2-, 5.10 (m, 1H) ring CH, d (5.22 (d, J=10.4, 1H)

CHH=, 5.29 (d, J=17.2 Hz, 1H) CHH=, 5.88 (m, 1H) CH$_2$=CH-.

R$_f$=0.31 in hexane: ethyl acetate (65:35)

N^1-Alloc-Ala[βHz(N ^{3}-Boc)]-OH (13a)
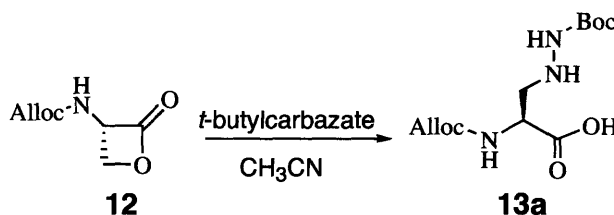
To a solution of t-butylcarbazate (2.2 g, 17 mmol) in anhydrous acetonitrile (20

mL) was cannulated a solution of N-Alloc-serine-β-lactone (**12**) (120 mg, 0.67 mmol) in

anhydrous acetonitrile (10 mL) over 30 minutes. After stirring under nitrogen for 14

hours at room temperature, the reaction was complete. The reaction mi.;ture was

concentrated on a rotary evaporator to a clear oil, to which 20 mL of ethyl acetate were

added. This solution was extracted with 0.1 N HCl (4 x 10 mL). The ethyl acetate layer

was dried over sodium sulfate, filtered, and concentrated on a rotary evaporator. The

product was purified by flash chromatography (2-cm diameter, 40 g silica). The column

was packed in ethyl acetate: hexane (9:1). After loading the crude product, 200 mL of

this solvent system was used to remove impurities, then the product was eluted with

CH$_2$Cl$_2$: MeOH: AcOH (10:1:0.1). After concentrating the product-containing fractions,

207

the product was washed 3X with toluene, then allowed to dry under vacuum overnight to yield compound **13a** as a white solid in 39% yield.

^{1}H NMR (400 MHz, CDCl$_3$): δ 1.44 (s, 9H) -C(CH_3)$_3$, 3.23 (m, 1H) βH$_1$, 3.32 (m, 1H) βH$_2$), 4.44 (m, 1H) αH, 4.57 (d, J=5.4 Hz, 2H) CH$_2$=CH-CH_2-, 5.21 (d, J=10.4 Hz, 1H) CHH=, 5.31 (d, J=17.2 Hz, 1H) CHH=, 5.87 (m, 1H) CH$_2$=CH.

R$_f$ = 0.31 in CH$_2$Cl$_2$:MeOH:AcOH (10:1:0.1)

N^1-Alloc-Ala[βHz(N ^{2,3}-bisBoc)]-OH (9)

HN–Boc
NH
Alloc–N–H ... OH O
13a

Boc$_2$O, DIEA
————————→
H$_2$O/ dioxane

HN–Boc
N-Boc
Alloc–N–H ... OH O
9
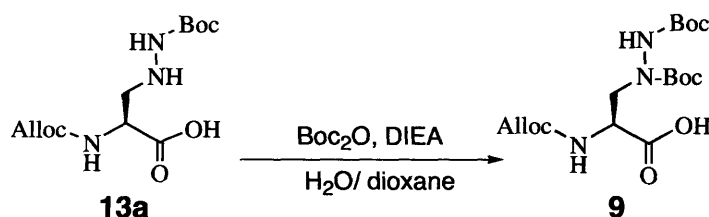
To a solution of N^1-Alloc-Ala[βHz(N^3-Boc)]-OH (**13a**) (0.080 g, 0.26 mmol) in 2:1 H$_2$O: dioxane (1 mL) was added Boc$_2$O (0.011 g, 0.52 mmol) and diisopropyl ethylamine (DIEA) (140 μL, 0.78 mmol). The reaction mixture was stirred at room temperature for 36 hours. The reaction mixture was diluted with ethyl acetate (2 mL) and washed with 0.1 N HCl (4 x 2 mL). The organic layer was dried over sodium sulfate, filtered, and concentrated on a rotary evaporator to a clear oil. The crude product was purified by flash chromatography (2 cm diameter, 40 g silica). Impurities were eluted with 9:1 chloroform: methanol (300 mL), the pure product was eluted with 100:5:1 CH$_2$Cl$_2$: MeOH: AcOH. Product-containing fractions were concentrated on a rotary evaporator, then washed three times with toluene and dried under vacuum to afford a clear oil in 64% yield.

^{1}H NMR (400 MHz, CDCl$_3$): δ 1.46 (s, 9H) -C(CH_3)$_3$, 1.49 (s, 9H) -C(CH_3)$_3$, 3.71 (m, 1H) βH$_1$, 3.87 (m, 1H) βH$_2$, 4.44 (m, 1H) αH, 4.57 (m, 2H) CH$_2$=CH-CH_2-, 5.21 (m, J=8.2 Hz, 1H) CHH=, 5.31 (d, J=17.1 Hz, 1H) CHH=, 5.91 (m, 1H) CH$_2$=CH-.

R$_f$ = 0.53 in CH$_2$Cl$_2$: MeOH: AcOH (10:1:0.1)

Marfey's test for racemization

In order to react Marfey's reagent (1-fluoro-2,3-dinitrophenyl-5-L-alanine amide) with the amino acid, the Alloc protecting group needed to be removed from the main-chain amine. To a solution of N^1-Alloc-Ala[βHz($N^{2,3}$-Bisboc)]-OH (1 mg, 0.0025 mmol) in CH$_2$Cl$_2$ (100 μL) was added phenylsilane (7.7 μL, 0.063 mmol), followed by a solution of (PPh$_3$)$_4$Pd(0) (0.6 mg, 0.0005 mmol) in dichloromethane (100 μL). The reaction mixture was stirred at room temperature for 20 minutes, then dried under a stream of nitrogen to remove CH$_2$Cl$_2$ and under vacuum for 1 hour to remove phenylsilane.

To a solution of the deprotected amino acid in 50 μL of H$_2$O was added 100 μL of Marfey's reagent (10 mg/mL), followed by 20 μL of NaHCO$_3$ (1 M). The reaction mixture was left at 37°C for 1 hour, then was quenched with 20 μL of 1 N HCl. For HPLC analysis, the reaction mixture was diluted with 810 μL of acetonitrile, then filtered through a 0.45-μm syringe filter. The filtrate was passed through a sep-pack, doubling the volume of the sample, then filtered once more. A 20-μL aliquot of sample was injected onto a Waters C$_{18}$ analytical HPLC column and eluted with a linear gradient of 0-70% or 30-70% CH$_3$CN/0.1% TFA over 30 minutes while monitoring the absorbance at both 254 nm and 280 nm. ESI-MS was performed on each of three peaks, which had retention times of 21.8, 25.5, and 26.6 min, respectively, for the 0-70% gradient, and 6.8,

15.8, and 20.0 min, respectively, for the 30-70% gradient. No signal was seen for the first peak, which is likely Marfey's reagent. This molecule may not ionize very easily. The second peak, with a mass of [M+H$^+$ = 279.1 and M+H$^+$+Na$^+$ =301.1], did not have an absorbance at 280 nm, and therefore was not an adduct of Marfey's reagent. The third peak from the HPLC had a mass of [M+H$^+$+Na$^+$ = 594.2], which corresponds to the dipeptide adduct formed in the reaction.

Bz-Ala(βHz)-Ala-Thr-Val-Thr-Nph-NH$_2$ (14)

14
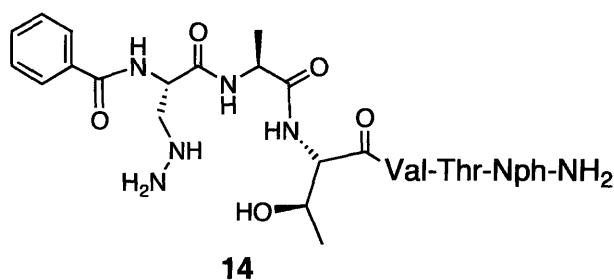
A 200-mg portion of Fmoc-Ala-Thr-Val-Thr-Nph-PAL-PEG-PS (see above for general procedure for peptide synthesis) was transferred to a fritted glass reaction vessel and the Fmoc group was removed with 20% piperidine in DMF (3 x 5 mL x 5 minutes). Following deprotection, resin was rinsed with DMF (5 x 5 mL x 1 minute) and CH$_2$Cl$_2$ (2 x 5 mL x 1 minute). Resin was transferred from the reaction flask to a 10-mL pear-shaped vial containing 16 mg of dry N^1-Alloc-Ala[βHz($N^{2,3}$-Bisboc)]-OH. DMF (2 mL), PyAOP (41 mg, 0.080 mmol), and collidine (21 μL, 0.16 mmol) were added, and the coupling reaction was allowed to shake overnight at room temperature. Contents of the vial were then transferred back to the fritted glass reaction vessel, reagents were drained, and the resin was rinsed with DMF (5 x 5 mL x 1 min) and CH$_2$Cl$_2$ (2 x 5mL x 1 min).

Alloc group removal, capping, cleavage and purification were done as described in general procedures.

HPLC t_R= 19.6 min (C_{18}, 7-100% B in 30 min)

LRMS calcd for **14** $[C_{35}H_{50}N_{10}O_{11} + H]^+$ requires m/z 787.4. Found 787.3 (ESI+)

Bz-Ala(β-hydroxylamide)-AlaThrValThrNph-NH$_2$ (15)
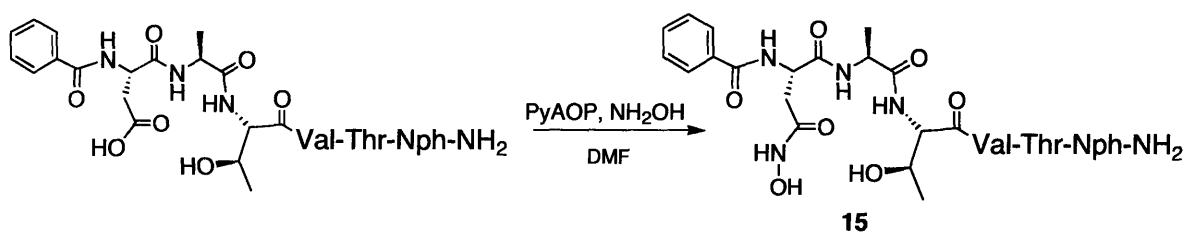
The immobilized peptide, Ala-Thr-Val-Thr-Nph-PAL-PEG-PS, was assembled according to the general procedures described above. Activation of N-α-Fmoc-Asp(OAll) was performed *in situ* with 2 equivalents of PyAOP and 4 equivalents of collidine to prevent epimerization. Following the coupling of this residue to the resin, removal of Fmoc from this residue was completed using 20% piperidine in DMF containing 0.1 M dinitrophenol (3 x 5 mL/g resin x 20 min), then washed with CH_2Cl_2, DMF, and methanol until yellow color disappeared. The hydroxylamide functionality was prepared by deprotection of O-Allyl protected Asp residue as described above, followed by coupling of hydroxylamine (activated *in situ* with 1.5 equivalents of PyAOP and 3 equivalents of DIEA) for 1 hour in DMF (10 mL/g resin).

HPLC : t_R=18.0 min (7-100% CH_3CN/0.1%TFA in 30 min)

LRMS calcd for **15** $[C_{36}H_{49}N_9O_{13} + H]^+$ requires m/z 816.8. Found 816.2 (ESI+)

Bz-DabValThrValThrNph-NH₂ (16)
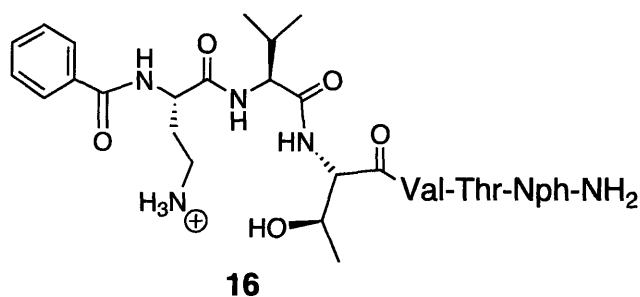
16

The immobilized peptide was assembled according to the general procedures described above.

HPLC : t_R=19.0 min (7-100% CH_3CN/0.1%TFA in 30 min)

LRMS calcd for **16** $[C_{38}H_{56}N_9O_{11}]^+$ requires m/z 814.4. Found 814.3 (ESI+)

Bz-DabNvaThrValThrNph-NH₂ (17)

17
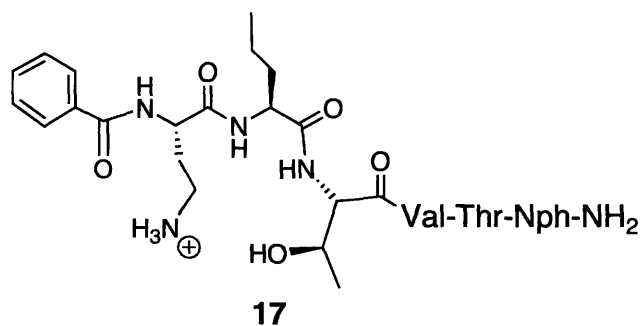
The immobilized peptide was assembled according to the general procedures described above.

HPLC : t_R=20.8 min (7-100% CH_3CN/0.1%TFA in 30 min)

LRMS calcd for **17** $[C_{38}H_{56}N_9O_{11}]^+$ requires m/z 814.4. Found 814.3 (ESI+)

Rho-Ahx-Leu-Asn(GlcNAc₂)-Asp-Ser-Arg-NH₂ (21)

21
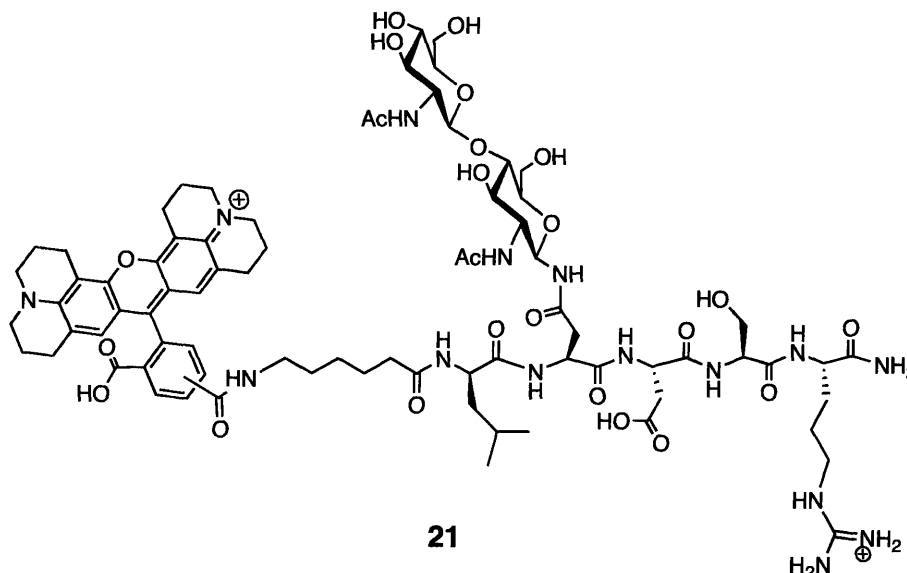
Preparation of this substrate began with the previously described de-allylation of 44 mg of the Fmoc-Asn(GlcNAc$_2$-TBS$_5$)-OAll building block (Chapter III), which had been prepared previously for a putative Alg1 substrate. The peptide NH$_2$-Ahx-Asn(GlcNAc$_2$)-Asp-Ser-Arg-PAL-PEG-PS (where Ahx = aminohexanoic acid) was prepared by standard SPPS methods, incorporating the deprotected Asn(GlcNAc$_2$) building block. One equivalent of the succinimidyl ester of the rhodamine derivative carboxy-X-rhodamine (ROX) was coupled in DMF overnight in the presence of eight equivalents of DIPEA. Peptide **22** was cleaved from the resin in a cocktail consisting of 95% TFA, 2.5% H$_2$O, and 2.5% triisopropylsilane at room temperature for 1 hour. HPLC purification and ESI-MS analysis followed.

HPLC : t$_R$=21.5 min (7-100% CH$_3$CN/0.1%TFA in 30 min)

LRMS calcd for **22** [C$_{78}$H$_{109}$N$_{15}$O$_{24}$$^+$]$^+$ requires m/z 1639.8. [C$_{78}$H$_{109}$N$_{15}$O$_{24}$$^+$ + H]$^{2+}$ requires m/z 819.9. Found 820.1 (ESI+).

213

OT Activity Assays

Dolichylpyrophosphate-linked (^{3}H)GlcNAc-GlcNAc-PP-Dol was prepared using porcine liver microsomes as described.[40] A 50,000-dpm aliquot of the dolichylpyrophosphate-linked substrate was resuspended in 10 μl of dimethylsulfoxide (DMSO) and vortexed to resuspend the sugar. Assay buffer was prepared by adding 30 μL of 1 M MnCl$_2$ to 1.5 mL of the following buffer: 140 mM sucrose, 50 mM Hepes pH 7.5, 1.2% Triton X-100, and 0.5 mg/ml PC. Assay buffer was added to the sugar substrate to achieve a final volume of 200 μL after addition of enzyme and peptide substrate. OT-containing solubulized microsomes, prepared from *S. cerevisiae* as previously described,[41] were then added and vortexed at a low setting. Finally, the reaction was initiated by the addition of 10 μl of Bz-NLT-NHMe (5 mM), followed by vortexing at a low setting and then shaking at 180 rpm. Four 7-mL scintillation vials, each containing 1 mL of chloroform:methanol (3:2) and 200 μL of 4 mM MgCl$_2$, were used to quench 40-μL aliquots from the reaction at exactly two minute intervals until eight minutes had passed. The vials were vortexed at a low setting simultaneously with quenching. Once all four time points had been taken, and the organic and aqueous layers were separated and clear, the aqueous layer from each vial was extracted and transferred to four separate 7-mL scintillation vials. A 600-μL aliquot of theoretical upper layer, or TUP (2.75% chloroform, 44% methanol, 53.25% aqueous magnesium chloride (1.55 mM)), was then added to each of the organic layers and vortexed at a low setting. Once the organic and aqueous layers had separated, the aqueous layer was removed from each vial and added to the previously removed aqueous layer. The TUP extraction was repeated once again, pooling the aqueous layers for each time point separately. Aliquots

214

of 5-mL of Eco-Lite scintillation cocktail (ICN Biomedicals, Inc.) were added to each of the pooled aqueous layers, then the vials were capped and vortexed at high speed to clarify solutions. The solutions were counted by a scintillation counter for five minutes per vial. Disintegrations per minute (dpm) were plotted as a function of time.

K_m of Bz-Asn-Ala-Thr-Val-Thr-Nph-NH$_2$

OT assays were performed as described above, except that instead of Bz-NLT-NHMe, Bz-Asn-Ala-Thr-Val-Thr-Nph-NH$_2$ was used as the disaccharide acceptor. The K_m of the tripeptide substrate, 0.25 μM, is similar to that of the hexapeptide containing asparagine that is used for comparison in this study, which has a K_m of 0.31 μM. The following concentrations were assayed: 5 nM, 50 nM, 125 nM, 250 nM, 500nM, 1 μM, 2 μM, 3 μM, and 4 μM. From the slope of the dpm vs. time plots, velocity (nmol/sec) was calculated. Velocity vs. substrate concentration were plotted to confirm saturation, and a Hanes plot ([S]/V vs. [S]) was used to determine the K_m, with a linear regression of 0.99447. From the Hanes plot, the x-intercept is equal to (-)K_m.

OT Inhibition assays

Inhibition assays proceeded as described in the preceding section with the following changes: A 10-μL aliquot of inhibitor in DMSO was used in place of the 10-μL DMSO aliquot added first to the dried sugar substrate. (^{3}H)GlcNAc-GlcNAc-PP-Dol of approximate specific activity 20 Ci/mmol was used as well as a final concentration of 0.1 mM Bz-AsnLeuThrNH$_2$. Reaction mixtures contained 10 μL of DMSO or inhibitor, 140

μL of assay buffer, 40 μL of solubilized OT microsomes, and to initiate the reaction after a 30 minute incubation with orbital shaking on ice, 10 μL of 2 mM Bz-NLT-NH$_2$.

The disintegrations per minute (dpm) were plotted as a function of time for the control and 3 different inhibitor concentrations. The percentage of inhibition was determined from this plot in order to estimate the IC$_{50}$. Three concentrations were then selected to give between 30% and 70% inhibition. All experiments were run a minimum of three times. In each case, the approximate K$_i$ was determined using the following equation.[42]

$$K_i = \frac{[I] \times (1-i)}{i + \left(\frac{[S]}{K_M} \times i\right)}$$

HPLC Assay for inhibitors as substrates

Peptides to be assayed were prepared as 2 mM solutions in DMSO. A 10-μL aliquot of 2 mM peptide was then added to 100,000 dpm (9 Ci/mmol) of dried (^{3}H)GlcNAc-GlcNAc-PP-Dol and vortexed at high speed. To this mixture was added 140 μL of standard OT assay buffer and vortexed again. To initiate the reaction, 50 μL of solubilized OT microsomes were added with gentle mixing. After fifteen hours of incubation at room temperature with orbital shaking, the reactions were quenched by adding the entire amount to 5 mL of chloroform: methanol (3:2) containing 1 mL of 2.5 mM MgCl$_2$ and vortexing to mix the layers. Once the organic and aqueous layers separated, the aqueous layer was transferred to a separate vial. The organic phase was extracted twice with 4-mL aliquots of TUP without salt [chloroform: methanol: water (2.75 : 44 : 53.25 v/v)], and combined with the original aqueous layer. Pooled aqueous

layers were left to dry overnight under a stream of nitrogen. Once dry, the peptide was redissolved in 880 μL of water : DMSO (10:1) and filtered through a 0.45 micron syringe filter. The entire sample was then injected onto a Waters analytical C_{18} HPLC column equilibrated with 20% acetonitrile, 80% water, 0.1% trifluoroacetic acid (TFA) at a rate of 1 mL/minute. Following injection, this solvent ratio was maintained for 5 minutes, then a gradient of 20-80% acetonitrile/0.1%TFA was completed over 20 minutes. Fractions of 1 mL were collected every minute for the first ten minutes, then 0.5-mL fractions were collected every thirty seconds for ten minutes, and finally 1-mL fractions were collected for the last five minutes. Aliquots of 5 mL of Eco-Lite were added to each fraction and counted on a Beckman LS-5000TD scintillation counter to give the amount of (^{3}H)-GlcNAc incorporation into the peptide.

Png1 expression and purification

The laboratory of Professor W. Lennarz provided the PNG1 gene from *S. cerevisiae* cloned into an expression vector (pET-28b) for *E. coli* expression of this construct with a C-terminal His tag. Cells from a glycerol stock were used to grow a 50-mL overnight culture in LB containing 34 μg/mL chloramphenicol. Following overnight growth, the 50-mL culture was added to 950 mL of LB-chloramphenicol, and grown to an O.D. of 0.79. Protein expression was then induced with the addition of IPTG to 1 mM and growth was continued for 4 hours. Cells were harvested by centrifugation at 3,000 rpm for 30 minutes at 4 °C and stored in four equal-sized batches at –80 °C.

Purification began by thawing one aliquot of cells (from 250 mL of expression culture) on ice, then resuspending them in 25 mL of lysis buffer (20 mM Tris-HCl, pH

7.5, 1% Triton X-100, and 1 mM PMSF). All steps were carried out at 4 °C. Cell lysis was achieved by three 10-second pulses of sonication, separated by 10-second cooling periods. The cells were kept on ice during the entire process. The resulting lysate was clarified by centrifugation at 8,000 rpm for 45 minutes. Supernatant was added to 2 mL of Talon™ resin pre-equilibrated in binding buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl). This mixture was shaken at 4 °C for 2 hr to allow binding of His-tagged protein, then poured into a column with a plugged outlet. The plug was then removed and the flow-through was collected. Resin-bound protein was washed twice with 20 mL of binding buffer containing 10 mM imidazole, and eluted with 12 mL of binding buffer containing 100 mM imidazole, collecting 2-mL elution fractions. Fractions were stored at –20 °C. Png1-containing fractions were thawed, brought to 10 mM DTT, combined, and concentrated to 3 mL with concomitant buffer-exchange into B88 buffer (20 mM HEPES-KOH, pH 6.8, 5 mM potassium acetate, 150 mM magnesium acetate, 250 mM D-sorbitol, 10 mM DTT) using a Millipore Ultrafree-15 centrifugal filter device by centrifuging at 3,000 rpm, 4 °C. Protein was quantified using the BioRad protein assay, according to the manufacterer's instructions. A standard curve was plotted with A_{595} vs. [BSA](mg/mL) using only the values for the BSA standards. The equation from fitting the data to a line was $y=0.029+0.5985x$ $R=0.99933$. From this equation, the concentration of Png1 was determined to be 0.53 mg/mL. Prior to assays, the enzyme was diluted to 14 µg/mL in B88 buffer.

Png1 assays

An aliquot of diluted Png1 (either 14 ng or 140 ng) was added to a solution of the FITC-labeled glycopeptide substrate (450 pmol) in 40 mM Mes-NaOH, pH 6.6, 10 mM DTT for a final volume of 100 μL. Reaction mixtures were heated to 30 °C and incubated overnight. HPLC analysis was performed by removing 30 μL of reaction mixture, adding 70 μL of water and 100 μL of acetonitrile, centrifuging to remove any insoluble material, and injecting the supernatant onto a Waters C_{18} analytical HPLC column. Products were eluted with a gradient of 20-60% acetonitrile/0.1% TFA in water /0.1% TFA. For identification, peaks were collected, dried on a Speed-vac, redissolved in methanol: water (50: 50) containing 1% acetic acid, and analyzed by ESI-MS.

In the first round of inhibitor assays, 1 nmol of inhibitor **19**, **20**, or **21** was added to PNGase1 (14 ng) in Mes-NaOH buffer (above) for a final volume of 19.3 μL and incubated at 30°C for 30 minutes. The reaction was then initiated by the addition of fluorescein-labeled glycopeptide substrate (0.4 nmol), bringing the final volume to 20 μL. Incubation was continued for 12 hours before the reactions were quenched with the addition of TFA to 1%, and the products were analyzed as described above. The second round of inhibition assays was completed in the same manner as the first with the exception that 2 nmol of substrate, 0.2 nmol of inhibitor, and 140 ng of enzyme were used. Percent conversion was measured by dividing the area under the product peak by the sum of the areas under the starting material and product peaks.

References

1.   Kornfeld, R.; Kornfeld, S., Assembly of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* **1985,** 54, 631-664.

2.   Gavel, Y.; von Heijne, G., Sequence Differences Between Glycosylated and Non-glycosylated Asn-X-Thr/Ser Acceptor Sites: Implications for Protein Engineering. *Protein Eng.* **1990,** 3, 433-442.

3.   Hirschberg, C. B.; Snider, M. D., Topography of glycosylation in the rough endoplasmic reticulum and golgi apparatus. *Ann. Rev. Biochem.* **1987,** 56, 63-87.

4.   Dempski, R. E.; Imperiali, B., Oligosaccharyl transferase: gatekeeper to the secretory pathway. *Curr. Opin. Chem. Biol.* **2002,** 6, 844-850.

5.   Knauer, R.; Lehle, L., The Oligosaccharyltransferase Complex from Yeast. *Biochim. Biophys. Acta* **1999,** 1426, (2), 259-273.

6.   Imperiali, B., Protein glycosylation: The clash of the titans. *Acc. Chem. Res.* **1997,** 30, (11), 452-459.

7.   Imperiali, B.; Shannon, K. L., Differences between Asn-Xaa-Thr-containing peptides: a comparison of solution conformation and substrate behavior with oligosaccharyltransferase. *Biochemistry* **1991,** 30, (18), 4374-80.

8.   Beintema, J. J., Do asparagine-linked carbohydrate chains in glycoproteins have a preference for beta-bends? *Biosci. Rep.* **1986,** 6, (8), 709-14.

9.   Imperiali, B.; Rickert, K. W., Conformational implications of asparagine-linked glycosylation. *Proc. Natl. Acad. Sci. USA* **1995,** 92, (1), 97-101.

10.   O'Connor, S. E.; Imperiali, B., Conformational switching by asparagine-linked glycosylation. *J. Am. Chem. Soc.* **1997,** 119, (9), 2295-2296.

11.   Imperiali, B.; Shannon, K. L.; Unno, M.; Rickert, K. W., A Mechanistic Proposal for Asparagine-Linked Glycosylation. *J. Am. Chem. Soc.* **1992,** 114, (20), 7944-7945.

12.   Marshall, R. D., The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochem. Soc. Symp.* **1974,** (40), 17-26.

13.   Bause, E., Model Studies on N-Glycosylation of Proteins. *Biochem. Soc. Trans.* **1984,** 12, (3), 514-517.

14.   Bause, E.; Breuer, W.; Peters, S., Investigation of the active site of oligosaccharyltransferase from pig liver using synthetic tripeptides as tools. *Biochem. J.* **1995,** 312, 979-985.

15.   Welply, J. K.; Shenbagamurthi, P.; Lennarz, W. J.; Naider, F., Substrate recognition by oligosaccharyltransferase. Studies on glycosylation of modified Asn-X-Thr/Ser tripeptides. *J. Biol. Chem.* **1983,** 258, (19), 11856-63.

16.   Hendrickson, T. L.; Spencer, J. R.; Kato, M.; Imperiali, B., Design and evaluation of potent inhibitors of asparagine-linked protein glycosylation. *J. Am. Chem. Soc.* **1996,** 118, (32), 7636-7637.

17.   Kellenberger, C.; Hendrickson, T. L.; Imperiali, B., Structural and functional analysis of peptidyl oligosaccharyl transferase inhibitors. *Biochemistry* **1997,** 36, (41), 12554-9.

18. Gavel, Y.; Vonheijne, G., Sequence Differences between Glycosylated and Nonglycosylated Asn-X-Thr Ser Acceptor Sites - Implications for Protein Engineering. *Protein Eng.* **1990,** 3, (5), 433-442.

19. Marcaurelle, L. A.; Bertozzi, C. R., New directions in the synthesis of glycopeptide mimetics. *Chem.-Eur. J.* **1999,** 5, (5), 1384-1390.

20. Peri, F., Extending chemoselective ligation to sugar chemistry: convergent assembly of bioactive neoglycoconjugates. *Mini Rev. Med. Chem.* **2003,** 3, (7), 651-8.

21. Peluso, S.; Imperiali, B., Asparagine surrogates for the assembly of N-linked glycopeptide mimetics by chemoselective ligation. *Tetrahedron Lett.* **2001,** 42, (11), 2085-2087.

22. Yan, Q.; Lennarz, W. J., Oligosaccharyltransferase: A complex multisubunit enzyme of the endoplasmic reticulum. *Biochem. Biophys. Res. Comm.* **1999,** 266, (3), 684-689.

23. Yan, Q.; Prestwich, G. D.; Lennarz, W. J., The Ost1p subunit of yeast oligosaccharyl transferase recognizes the peptide glycosylation site sequence, -Asn-X-Ser/Thr-. *J. Biol. Chem.* **1999,** 274, (8), 5021-5025.

24. Shakin-Eshleman, S. H.; Spitalnik, S. L.; Kasturi, L., 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.* **1996,** 271, (11), 6363-6.

25. Suzuki, T.; Park, H.; Lennarz, W. J., Cytoplasmic Peptide: N-Glycanase (PNGase) in Eukaryotic Cells: Occurrence, Primary Structure, and Potential Functions. *FASEB J.* **2002,** 16, (7), 635-641.

26. Ellgaard, L.; Helenius, A., Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* **2003,** 4, (3), 181-91.

27. Suzuki, T.; Park, H.; Hollingsworth, N. M.; Sternglanz, R.; Lennarz, W. J., PNG1, a yeast gene encoding a highly conserved peptide: N-glycanase. *J. Cell. Biol.* **2000,** 149, (5), 1039-1051.

28. Suzuki, T.; Kitajima, K.; Inoue, Y.; Inoue, S., Carbohydrate-binding property of peptide: N-glycanase from mouse fibroblast L-929 cells as evaluated by inhibition and binding experiments using various oligosaccharides. *J. Biol. Chem.* **1995,** 270, (25), 15181-6.

29. Li, G.; Zhou, X.; Zhao, G.; Schindelin, H.; Lennarz, W. J., Multiple modes of interaction of the deglycosylation enzyme, mouse peptide N-glycanase, with the proteasome. *Proc. Natl. Acad. Sci. USA* **2005,** 102, (44), 15809-14.

30. Suzuki, T.; Park, H.; Kitajima, K.; Lennarz, W. J., Peptides glycosylated in the endoplasmic reticulum of yeast are subsequently deglycosylated by a soluble peptide: N-glycanase activity. *J. Biol. Chem.* **1998,** 273, (34), 21526-30.

31. Suzuki, T.; Park, H.; Kwofie, M. A.; Lennarz, W. J., Rad23 provides a link between the Png1 deglycosylating enzyme and the 26 S proteasome in yeast. *J. Biol. Chem.* **2001,** 276, (24), 21601-7.

32. Misaghi, S.; Pacold, M. E.; Blom, D.; Ploegh, H. L.; Korbel, G. A., Using a small molecule inhibitor of peptide: N-glycanase to probe its role in glycoprotein turnover. *Chem. Biol.* **2004,** 11, (12), 1677-87.

33. Misaghi, S.; Korbel, G. A.; Kessler, B.; Spooner, E.; Ploegh, H. L., z-VAD-fmk inhibits peptide:N-glycanase and may result in ER stress. *Cell Death Differ.* **2005.**

221

34.	Peluso, S.; de, L. U. M.; O'Reilly, M. K.; Imperiali, B., Neoglycopeptides as inhibitors of oligosaccharyl transferase: insight into negotiating product inhibition. *Chem. Biol.* **2002,** 9, (12), 1323-8.

35.	Ufret, M.L.; Imperiali, B., Probing the extended binding determinants of oligosaccharyl transferase with synthetic inhibitors of asparagine-linked glycosylation. *Bioorg. Med. Chem. Lett.* **2000,** 10, (3), 281-4.

36.	Dondoni, A.; Mariotti, G.; Marra, A., Synthesis of alpha- and beta-glycosyl asparagine ethylene isosteres (C-glycosyl asparagines) via sugar acetylenes and Garner aldehyde coupling. *J. Org. Chem.* **2002,** 67, (13), 4475-86.

37.	Zhang, R. E.; Cao, Y. L.; Hearn, M. W., Synthesis and application of Fmoc-hydrazine for the quantitative determination of saccharides by reversed-phase high-performance liquid chromatography in the low and subpicomole range. *Anal. Biochem.* **1991,** 195, (1), 160-7.

38.	Pansare, S. V.; Huyer, G.; Arnold, L. D.; Vederas, J. C., N-tert-butyloxycarbonyl-L-serine beta-lactone and (S)-3-amino-2-oxetanone p-toluenesulfonic acid salt. *Org. Synth.* **1991,** 70, 1-17.

39.	Kucharczyk, N.; Badet, B.; Legoffic, F., Quantitative Synthesis of L or D N-2-Tert-Butoxycarbonyl-2,3-Diaminopropanoic Acid from Protected L-Serine-Beta-Lactone or D-Serine-Beta-Lactone. *Syn. Comm.* **1989,** 19, (9-10), 1603-1609.

40.	Imperiali, B.; Zimmerman, J. W., Synthesis of Dolichylpyrophosphate-Linked Oligosaccharides. *Tetrahedron Lett.* **1990,** 31, (45), 6485-6488.

41.	Pathak, R.; Hendrickson, T. L.; Imperiali, B., Sulfhydryl modification of the yeast Wbp1p inhibits oligosaccharyl transferase activity. *Biochemistry* **1995,** 34, (13), 4179-85.

42.	Segel, I. H., *Enzyme Kinetics*. John WIley and sons, Inc.: New York, 1975.

MARY KATHERINE O'REILLY

Department of Chemistry 18-544
77 Massachusetts Avenue
Massachusetts Institute of Technology
Cambridge, MA 02139 USA
617-253-0206 (telephone)
oreilly@mit.edu

ACADEMIC BACKGROUND

2006	Ph.D. Biological Chemistry Massachusetts Institute of Technology, Cambridge, MA
1999	B.S. Chemistry Purdue University, West Lafayette, IN

PROFESSIONAL BACKGROUND

2000-2006	Research Assistant, Department of Chemistry Advisor: Professor Barbara Imperiali Massachusetts Institute of Technology
2003	Two-month Industrial Internship Advisor: Dr. Francine Perler New England Biolabs, Inc.
1997-1999	Research Assistant, Department of Chemistry Advisor: Professor Jean Chmielewski Purdue University
1997	Research Assistant, Department of Chemistry Advisor: Professor Bruce Morimoto Purdue University

PUBLICATIONS

1. O'Reilly, M.K.; Imperiali, B. *In vitro* evidence for the dual function of Alg2 and Alg11: Essential mannosyltransferases on the dolichol pathway. *Manuscript in preparation*
2. O'Reilly, M.K.; Imperiali, B. Alg1 accepts GlcNAc-PP-Dol as a substrate *in vitro. Manuscript in preparation*
3. Zhang, G.; O'Reilly, M.K.; Imperiali, B. A Bi-component Protein Complex is the Minimal Requirement for the Second Glycosyl Transfer Step in the Dolichol Pathway – *In Vitro* Reconstitution of Enzyme II Activity from the Yeast YGL047w and YBR070c Proteins, *Manuscript in preparation*
4. Hackenberger, C.P.; O'Reilly, M.K.; Imperiali, B. Improving Glycopeptide Synthesis: a Convenient Protocol for the Preparation of beta-

Glycosylamines and the Synthesis of Glycopeptides. *J. Org. Chem.* **2005** *70(9)*, 3574.

5. Zhao, L.; O'Reilly, M.K.; Shultz, M.D.; Chmielewski, J. Interfacial Peptide Inhibitors of HIV-1 Integrase Activity and Dimerization. *Bioorg. Med. Chem. Lett.* **2003** *13(6)*, 1175.

6. Peluso, S.; Ufret, M.D.; O'Reilly, M.K.; Imperiali, B. Neoglycopeptides as Inhibitors of Oligosaccharyl Transferase: Insight into Negotiating Product Inhibition. *Chem. Biol.* **2002** *9(12)*, 1323.

7. Tai, W.-F.V.; O'Reilly, M.K.; Imperiali, B. Substrate Specificity of *N*-Acetylglucosaminyl Diphosphodolichyl *N*-Acetylglucosaminyl Transferase, a Key Enzyme in the Dolichol Pathway. *Bioorg. Med. Chem.* **2001** *9 (5)*, 1173.

POSTER PRESENTATIONS

O'Reilly, M.K.; Imperiali, B. Investigation of a Series of Mannosyltransferases on the Dolichol Pathway of Asparagine-Linked Glycosylation. *Gordon Research Conference on Carbohydrates* **June, 2005**

O'Reilly, M.K.; Liu, Z.Y.; Wang, P.G.; Imperiali, B. Investigation of Glycosyltransferases on the Dolichol Pathway. *Abstr. Pap. Am. Chem. Soc.* **2002** *224*:151-BIOL part 1.

AWARDS AND SCHOLARSHIPS

2001- 2004	Biotechnology Training Grant National Institutes of Health
Fall 2000	Chemistry department award for outstanding teaching by a graduate student Massachusetts Institute of Technology
Spring 1999	Dale W. Margerum Undergraduate Research Scholarship Purdue University
Spring 1999	Iota Sigma Pi chemistry department outstanding senior woman award Purdue University