ELUCIDATION OF THE CATALYTIC MECHANISM OF GOLGI α-MANNOSIDASE II

by

Niket Shah

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Medical Biophysics
University of Toronto

Copyright © MMVIII by Niket Shah
Abstract

Elucidation of the catalytic mechanism of Golgi $\alpha$-mannosidase II

Niket Shah

Doctor of Philosophy

Graduate Department of Medical Biophysics

University of Toronto

MMVIII

The central dogma of molecular biology outlines the process of information transfer from a DNA sequence, to a protein chain. Beyond the step of protein synthesis, there are a variety of post-translational modifications that can take place, one of which is addition of carbohydrate chains to nascent proteins, known as glycosylation.

The N-linked glycosylation pathway is responsible for the covalent attachment of multifunctional carbohydrate chains on asparagine residues of nascent proteins at Asn-X-Ser/Thr consensus sequences. These carbohydrate chains are thought to aid in cell signaling, immune recognition, and other processes.

Golgi $\alpha$-mannosidase II (GMII) is the enzyme in the N-glycosylation pathway that is responsible for cleaving two mannose linkages in the oligosaccharide $\text{GnMan}_5\text{Gn}_2$ (where Gn is N-acetylglucosamine and Man is mannose), thereby producing $\text{GnMan}_3\text{Gn}_2$, which is the committed step in complex N-glycan synthesis. It has been speculated that GMII is an excellent therapeutic target for cancer treatment, as the unusual distribution of carbohydrates on the surface of tumour cells has been characterized in many cancers. In addition, swainsonine—a strong, yet nonspecific inhibitor of GMII—has been shown to block metastasis and improve the clinical outcome of patients with certain cancers, including those of the colon, breast and skin.

This thesis examines Golgi $\alpha$-mannosidase II from *Drosophila melanogaster* (dGMII) as a model for all GMII enzymes. First, a 1.80 Å resolution crystal structure of a weak inhibitor, kifunensine, binding to dGMII provides mechanistic insights into the substrate distortion in the GMII reaction. It is hypothesized that the GMII reaction proceeds via a $^1S_5$ intermedi-
ate. Second, a 1.40 Å resolution structure of a mutant dGMII bound to its natural substrate, GnMan$_5$Gn, identifies key substrate binding and catalytic residues, as well as expanding the definition of the GMII active site to include two distant sugar−binding subsites. Finally, the results are taken together, with knowledge of other related enzymes to synthesize a plausible itinerary for the GMII reaction.
Acknowledgements

I am grateful, first and foremost, to David Rose for inviting me into his lab and providing an environment in which curiosity, discovery and collegiality flourish. My committee members—Lynne Howell, Gil Privé and Harry Schachter—have enriched my graduate school experience with their support and enthusiasm for the project. I have had the pleasure of working with several exceptional individuals over the years, and their contribution to my journey has been immense. The present and past members of the Rose Lab bear particular mention, particularly Doug Kuntz, from whom I have learned a great deal, and who has been a pleasure to work with. The faculty and students of the Department of Medical Biophysics have been inspiring to work with and alongside.

I have been very fortunate to have a strong network of people around me for the past several years. My family has been a constant source of support, and I would not have been able to succeed without them. Having the right friends has led to years of quality pints, good bottles of wine and late nights filled with discussions of a scientific, philosophical, or culinary nature. My most profound thanks to my friends, particularly Eden, Kellie, Larissa, Lyann, Tina, Janet, Wes, Costin, Apeksha, Craig, Amanda and Chris.

Lastly—and most importantly—I am indebted to my dearest Caroline. She always knows how to get a laugh out of me.

I long to accomplish great and noble tasks, but it is my chief duty to accomplish humble tasks as though they were great and noble. The world is moved along, not only by the mighty shoves of its heroes, but also by the aggregate of the tiny pushes of each honest worker.

Helen Keller
# Contents

1 Introduction ............................................................. 1
   1.1 Biological roles of oligosaccharides ......................... 1
      1.1.1 Information transport .................................. 2
      1.1.2 Cellular adhesion ...................................... 4
      1.1.3 Self/nonself recognition ............................... 5
      1.1.4 Molecular trafficking and clearance .................. 6
   1.2 Asparagine-linked glycosylation ............................. 7
      1.2.1 Synthesis of sugar-nucleotide donors ................. 7
      1.2.2 N-linked glycosylation pathway ....................... 8
   1.3 Glycosylation and cancer ..................................... 11
   1.4 Glycoside Hydrolases ........................................ 13
      1.4.1 Golgi α-mannosidase II ............................... 13
      1.4.2 Related glycoside hydrolases ......................... 14
   1.5 Golgi α-mannosidase II inhibition and cancer ............. 16
   1.6 Structural study of Golgi α-mannosidase II ............... 19
   1.7 Dissertation overview ...................................... 21

2 Substrate distortion in the dGMII active site .......... 22
   2.1 Chapter overview ............................................ 22
   2.2 Materials and Methods ...................................... 23
      2.2.1 Protein expression and purification ................. 23
      2.2.2 Inhibition assay ...................................... 23
3.3 Results ......................................................... 44
  3.3.1 Structure of substrate–dGMII (D204A) complex .......................... 44
  3.3.2 dGMII(D204A)–GnManGn interactions ................................. 45
  3.3.3 Identification of dGMII subsites ....................................... 46
  3.3.4 Multiple sequence alignment .................................... 48
3.4 Discussion ......................................................... 49
  3.4.1 Conserved mode of substrate binding and catalysis .................... 54
  3.4.2 Comparison of substrate binding to inhibitor binding ................. 55

4 A novel catalytic mechanism for the action of dGMII .............................. 57
  4.1 Chapter overview ............................................. 57
  4.2 Materials and Methods ........................................ 58
    4.2.1 dGMII (D204A) crystallization .................................. 58
    4.2.2 Data collection ............................................. 58
    4.2.3 Structure determination ...................................... 58
    4.2.4 Generation of intermediate substrate models ......................... 58
    4.2.5 Lysosomal α-mannosidase sequence and structural analysis ........ 59
  4.3 Results ......................................................... 59
    4.3.1 Structure of dGMII (D204A) .................................... 59
    4.3.2 Intermediate substrate models .................................. 60
    4.3.3 Lysosomal α-mannosidase sequence and structural analysis ........ 61
  4.4 Discussion ....................................................... 61
    4.4.1 Comparison of substrate–dGMII to unliganded dGMII ............... 61
    4.4.2 Comparison of substrate–dGMII to a covalent reaction intermediate ... 62
    4.4.3 Comparison of substrate–dGMII to a synthetic mannose tetra-saccharide 62
    4.4.4 Putative catalytic mechanism .................................... 63
    4.4.5 Internal communication between dGMII subsites ....................... 67
    4.4.6 Comparison of dGMII to lysosomal α-mannosidase ..................... 69
5 Future Studies

5.1 Nuclear magnetic resonance analysis of the GMII reaction ........................................ 75
5.2 Molecular dynamics simulation of GMII reaction .......................................................... 76
5.3 Novel inhibitor development ......................................................................................... 77
  5.3.1 Increasing inhibitor specificity: GMII over LM ......................................................... 77
  5.3.2 Avidity-based approach ......................................................................................... 79
  5.3.3 Fragment-based approach ..................................................................................... 79
5.4 Fluorescence-based inhibitor screening ........................................................................ 80
  5.4.1 Fluorescence polarization ...................................................................................... 80
  5.4.2 Development of labeling protocol ......................................................................... 81
5.5 Expression profiling of glycosidases .......................................................................... 82
5.6 Further structural study of GMII ................................................................................ 83
  5.6.1 Mutagenesis studies .............................................................................................. 83

6 Conclusions ......................................................................................................................... 85

Bibliography ............................................................................................................................ 87
List of Tables

1.1 Information-carrying potential of biological polymers ............................................. 3

2.1 Data collection and refinement statistics for the kifunensine—dGMII complex ............... 25
2.2 Kifunensine conformational energies ................................................................. 27
2.3 dGMII—inhibitor atomic interactions ....................................................................... 31
2.4 Comparison of class I and class II α-mannosidases .................................................... 34

3.1 Data collection and refinement statistics for the GnMan₅Gn—dGMII complex ............... 45
3.2 GnMan₅Gn—dGMII interactions ............................................................................... 51
3.3 GnMan₅Gn crystallographic temperature factors ....................................................... 52

4.1 Data collection and refinement statistics for the dGMII (D204A) mutant .................... 60
4.2 Key residues in the dGMII catalytic site with their bLM equivalents .......................... 72
# List of Figures

1.1 N-linked glycosylation pathway ................................. 10
1.2 GMII reaction schematic ........................................ 15
1.3 GnMn$_5$Gn$_2$—Asn: The GMII substrate ......................... 16
1.4 Glycosyl hydrolase retaining mechanism ....................... 16

2.1 Colourimeric mannosidase assay ................................. 23
2.2 Electron density of kifunensine in the dGMII active site ........ 26
2.3 Chemical structure of inhibitors .................................. 28
2.4 Kifunensine in the dGMII active site ............................ 29
2.5 Kifunensine−dGMII interactions ............................... 30
2.6 Kifunensine and 1-deoxymannojirimycin in two $\alpha$-mannosidase active sites 31
2.7 Kifunensine, swainsonine, and 1-deoxymannojirimycin in the dGMII active site 33
2.8 Glycosyl hydrolase inverting mechanism ....................... 34
2.9 Pyranoside ring interconversions .............................. 35

3.1 Deglycosylation of RNaseB ....................................... 40
3.2 Separation of oligosaccharides from deglycosylated RNaseB .................. 41
3.3 Separation of GnMan$_5$Gn from GnT I reaction mixture ........ 42
3.4 Electron density of GnMan$_5$Gn in the dGMII active site ........ 46
3.5 Electron density of GnMan$_5$Gn in stereo ...................... 47
3.6 GnMan$_5$Gn in the dGMII active site ............................ 48
3.7 GnMan$_5$Gn in the dGMII active site (stereo) ................. 49
3.8 GnMan$_5$Gn—dGMII interactions ............................... 50
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Ångström, $1 \times 10^{-10}$ m</td>
</tr>
<tr>
<td>ASGP-R</td>
<td>asialoglycoprotein receptor</td>
</tr>
<tr>
<td>CDG</td>
<td>Congenital disorder of glycosylation</td>
</tr>
<tr>
<td>CMP</td>
<td>cytidine monophosphate</td>
</tr>
<tr>
<td>CRD</td>
<td>carbohydrate recognition domain</td>
</tr>
<tr>
<td>DMJ</td>
<td>1-deoxymannojirimycin</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>ERK</td>
<td>extracellular signal-regulated kinase</td>
</tr>
<tr>
<td>(h)ERMI</td>
<td><em>Homo sapiens</em> endoplasmic reticulum α-mannosidase</td>
</tr>
<tr>
<td>G2</td>
<td>β1,4-linked core N-acetylglucosamine</td>
</tr>
<tr>
<td>G3</td>
<td>terminal β1,2-linked branching N-acetylglucosamine</td>
</tr>
<tr>
<td>GDP</td>
<td>guanosine diphosphate</td>
</tr>
<tr>
<td>(d)GMII</td>
<td><em>Drosophila melanogaster</em> Golgi α-mannosidase II</td>
</tr>
<tr>
<td>Gn</td>
<td>N-acetylglucosamine</td>
</tr>
<tr>
<td>GnT I</td>
<td>N-acetylglucosaminyltransferase I</td>
</tr>
<tr>
<td>JNK</td>
<td>c-Jun N-terminal kinase</td>
</tr>
<tr>
<td>KIF</td>
<td>Kifunensine</td>
</tr>
<tr>
<td>(b)LM</td>
<td><em>Bos taurus</em> lysosomal α-mannosidase</td>
</tr>
<tr>
<td>M1</td>
<td>β1,4-linked core mannose</td>
</tr>
<tr>
<td>M2</td>
<td>α1,3-linked core mannose</td>
</tr>
<tr>
<td>M3</td>
<td>α1,6-linked core mannose</td>
</tr>
<tr>
<td>M4</td>
<td>α1,3-linked terminal mannose</td>
</tr>
<tr>
<td>M5</td>
<td>α1,6-linked terminal mannose</td>
</tr>
<tr>
<td>Man</td>
<td>mannose</td>
</tr>
<tr>
<td>MAPK</td>
<td>mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MTD</td>
<td>maximum tolerated dose</td>
</tr>
<tr>
<td>NF−κB</td>
<td>nuclear factor−κB</td>
</tr>
<tr>
<td>NK</td>
<td>natural killer</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>PSGL-1</td>
<td>P-selectin glycoprotein ligand-1</td>
</tr>
<tr>
<td>Sia</td>
<td>sialic acid</td>
</tr>
<tr>
<td>sLex</td>
<td>sialyl Lewis x</td>
</tr>
<tr>
<td>STD-NMR</td>
<td>Saturation transfer difference nuclear magnetic resonance</td>
</tr>
<tr>
<td>SWA</td>
<td>Swainsonine</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>UDP</td>
<td>Uridine diphosphate</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Glycosylation is one of the most widespread, and yet poorly understood modifications of proteins. Every cell in nature is coated with a dense layer of carbohydrates at its surface. Decorating intra- and extracellular proteins with large carbohydrate chains is a costly endeavour for a cell to undertake, yet it is seen in ubiquity, indicating that it has great importance. Carbohydrates comprise significant amount of the mass in biological systems, yet were long considered to be useful in little more than energy storage. The landmark discovery that the molecular determinant of blood type was a carbohydrate began the notion that sugars played a functional role in biological systems [1]. Research in recent decades has highlighted many detailed and important functions of these molecules.

It is now understood that functional carbohydrates have myriad functions in the cell and are synthesized and functionalized by highly conserved multi-enzymatic pathways. Furthermore, disruption of these intricate pathways can have broad functional effects.

1.1 Biological roles of oligosaccharides

The precise biological function of glycoconjugates is a topic which has been under intense study for many years. For example, the social interactions between a cell and the substratum, other cells, and functional molecules is intimately entwined with molecules at the cell surface. In many cases, cell surface carbohydrates, which are covalently linked to proteins
and lipids to form glycoconjugates, represent the initial point of the cell to make contact in an encounter. For this, and other reasons, glycoconjugates have been shown to participate in a wide variety of cellular functions. Recent advances have shown that covalently associated oligosaccharides are involved in vital cellular processes, and it follows that the disruption of their correct synthesis can result in a variety of disease states [2, 3, 4]. A dizzying array of functions ranging from cell adhesion, molecular trafficking, to immunological effects have all been studied in the context of covalently-linked carbohydrates. A brief survey of the roles of oligosaccharide-protein interactions in information encoding, cell adhesion, self/nonself recognition and molecular trafficking will be explored in sections 1.1.1, 1.1.2, 1.1.3 and 1.1.4, respectively.

1.1.1 Information transport

Recent work on the structures of biological oligosaccharides has been able to ascertain that, like proteins and nucleic acids, biological oligosaccharides can also act as information carriers, thus being the transfer mechanism of the "sugar code" (reviewed in [5]).

Protein-conjugated oligosaccharides represent non-template guided determinants which confer functional outcomes. They are uniquely positioned to be variable in three dimensions: oligosaccharide sequence, oligosaccharide branching, and conformational flexibility. They are also spatially available, being positioned at the protein surface with minimal protein–glycan interactions and often reaching some 3 nm from the glycosylation site.

Also, due to their propensity for branching and consequent variability leading to a large number of structural isomers, oligosaccharides can carry information with a very high degree of efficiency, and in a highly compact package [6]. When comparing the information carrying potential of biological polymers, the ability of oligosaccharides to branch affords a much higher degree of isomerization and, consequently, information potential (Table 1.1). This ability has been termed a "high-density coding system" by Gabius et al [5].

Furthermore, if uniquely branched oligosaccharides are to be understood to be information carriers, then it bears mention that lectins would the ‘readers’ of the glycan code [7]. Lectins are highly abundant in genomes, with nearly 200 C-type lectins in 20 000 open reading frames
Chapter 1. Introduction

<table>
<thead>
<tr>
<th>Polymer</th>
<th>Monomer</th>
<th>Topology</th>
<th>Template</th>
<th># of hexamers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotides</td>
<td>4 nucleic acid bases (A,C,G,T/U)</td>
<td>linear</td>
<td>DNA/RNA</td>
<td>4096</td>
</tr>
<tr>
<td>Oligosaccharides</td>
<td>9 monosaccharides (Gal, Man, Fuc, Sia, GalNAc, GlcNAc, Glc, GlcA, Xyl)</td>
<td>linear, branched</td>
<td>n/a</td>
<td>$1.44 \times 10^{15}$</td>
</tr>
</tbody>
</table>

Table 1.1: Information-carrying potential of biological polymers. The branching potential of oligosaccharides allows for a greater number of unique hexamers from a similar number of monomers as compared to nucleotides and proteins.

in the nematode, placing them seventh overall in that organism [5].

Despite the apparent ubiquity of oligosaccharide ‘signaling,’ there have been significant challenges in pursuing in-depth study. The lack of a template-driven mechanism of synthesis makes the production of large amounts of pure oligosaccharide difficult. Moreover, purification from natural sources poses challenges due to the heterogeneity of oligosaccharides, as well as their high degree of isomerization.

The heterogeneity seen in biological oligosaccharides is due to several coincident factors. First, the expression of glycosyltransferase genes required for oligosaccharide synthesis are regulated in a tissue-specific manner. Second, the sugar-nucleotide substrates used by glycosyltransferases to synthesize oligosaccharides are not universally available, thus limiting their addition by certain enzymes. Third, there is competition between glycosyltransferase enzymes for acceptor substrates (nascent oligosaccharides) in the synthesis pathway.

There are significant efforts to improve chemical synthesis techniques and solid-phase synthesis of oligosaccharides is an example of a method that has shown promise [8]. It is hoped that a method to produce pure populations of oligosaccharides in large quantities will assist in structural and functional studies of these important molecules, and the details of the sugar code can be explored.
1.1.2 Cellular adhesion

Leukocyte trafficking is a vital part of the innate and adaptive immune response. Leukocyte function depends on specific targeting from the vascular environment to extravascular sites of inflammation. It has been shown that protein–oligosaccharide interactions are intimately involved in this complex process (recently reviewed in [9]). Leukocytes typically flow through the vasculature with a multitude of other cells. For specific targeting to occur, these cells must be slowed and stopped at specific sites in the face of such challenges of high shear flow, as well as in the presence of a myriad of other cells. Three members of the selectin family— P-selectin (expressed by platelets and the endothelium), E-selectin (expressed by the endothelium), and L-selectin (expressed by leukocytes)—and their counter-receptors mediate this process.

Selectins are Type I glycoproteins that contain large extracellular carbohydrate-recognition domains (CRDs). E- and P-selectins are similar in that they are expressed primarily on the endothelium, with their glycan counter receptors decorating the leukocyte cell surface. They make numerous weak interactions with circulating leukocytes, causing the cells to roll along the endothelial surface of vascular structures, and cause slowing near sites of extravascular inflammation. L-selectin, on the other hand, is expressed by most circulating leukocytes and recognizes counter-receptors directly on the surface of endothelial cells (termed primary capture), or on the surface of leukocytes previously adhered to the endothelium (termed secondary capture).

In the cases mentioned above, the counter-receptors for the selectins are strongly believed to be oligosaccharide structures, and the function of the CRDs are to specifically recognize these carbohydrates from amidst a multitude of cells and molecules. The 'weak interactions' between CRD and counter-receptor allow the leukocytes to slowly roll across the endothelium, thereby causing a slow accumulation at the needed site of action. The rolling action is understood to be a necessary first step in the cascade of leukocyte recruitment to inflammatory sites.

Early experiments performed in vitro previously identified the E-selectin counter-receptor to be the sialyl Lewis x (sLex) tetrasaccharide and its isomers [10]. However, a glycoprotein
or glycolipid bearing this ligand could not be identified in leukocytes. Further experiments showed that the sLex tetrasaccharide is not expressed by leukocytes, thus indicating that it is not a physiological E-selectin ligand in leukocyte trafficking. Later work identified specific glycoproteins with counter-receptor activity towards E-selectin on a host of cells. A prominent glycoprotein in this group proved to be P-selectin glycoprotein ligand-1 (PSGL-1). Mice that have a PSGL-1 gene deletion show defects in E-selectin-dependent leukocyte adhesion. More recent work has shown that the oligosaccharide component of PSGL-1 that contributes to P-selectin binding is dependent upon core-2-type O-linked glycans [11].

Selectins are also implicated in anti-tumourigenic effects stemming directly from natural killer (NK) cell recruitment [12]. The original discovery of NK cells noted that they have the ability to kill certain tumour cells \textit{in vitro} without prior sensitization. The link between NK cells and anti-tumourigenic activity has been explored for some time. Briefly, \textit{in vivo} stimulation of lymph nodes by complete Freund’s adjuvant or by metastatic B16 melanoma cells is shown to result in a recruitment of a subset of NK cells. An analysis of these cells shows that they express L-selectin on the cell surface. Further, this selective trafficking of NK cells is facilitated by L-selectin expressed on the cell surface, and the L-selectin counter-receptor on the endothelial surface. Defects in this process result in aggressive tumour formation in lymph nodes.

1.1.3 Self/nonsel relationship

A key role of the immune system is the inflammatory response to foreign particles, which can be intricately linked to saccharide recognition by lectins. One of the earliest steps in mounting an effective immunological defense is the proper recognition of infectious agents to trigger the correct cellular response. The Toll-like receptors (TLRs) are a family of cell surface receptors that mediate recognition of various of microbial products and initiate the signaling pathway in the innate immune system [13]. Despite the diversity of TLR ligands, the culmination of all TLR stimulation is the activation of a common set of signaling pathways including nuclear factor–κB (NF–κB) transcription factors, mitogen-activated protein kinases (MAPKs), extracellular signal–regulated kinase (ERK), p38, and c-Jun N-terminal kinase (JNK). Since
many components of the bacterial cell wall are composed of carbohydrates ($\beta$-glucans, mannan, mannan, mannoproteins and chitin), it follows that certain TLRs exhibit lectin behaviour leading to stimulation, or employ cell-surface lectins in recognizing certain microbial products [14]. For example, the recognition of bacterial lipopolysaccharides by the immune system is brokered by TLR2 in concert with dectin-1, a C-type lectin expressed on the surface of monocytes, macrophages and dendritic cells [15].

In addition to being involved in the recognition of foreign particles, carbohydrates are also linked to autommunity. The $\alpha$-mannosidase II enzyme resides in the Golgi apparatus and cleaves two mannose residues from an oligosaccharide, thus catalyzing the committed step of complex N-glycan synthesis. Aged mice lacking the $\alpha$-mannosidase II gene show systemic autoimmune disease reminiscent of human systemic lupus erythematosus, including hematological, renal and immunological disorders [16]. Since lymphocytes from aged mutant mice showed complex N-glycans at nearly wild-type levels, it has been hypothesized that there is an age-dependent effect. Furthermore, it has also been proposed that altered N-glycan branching has deleterious effects on the self-tolerance ability of the immune system.

These two roles of carbohydrates in self/nonself recognition may not be as separate as originally thought. There is recent evidence that upregulation of TLR4, the Toll-like receptor for bacterial lipopolysaccharides, results in autoimmune disease in a renal cell model [17]. This result provides a link between detection of foreign carbohydrate particles in mounting an immune response to autoimmunity due to misidentification of self cells.

### 1.1.4 Molecular trafficking and clearance

An intriguing example of molecular trafficking and clearance as modulated by glycosylation status is that of the hepatic asialoglycoprotein receptor (ASGP-R). ASGP-R is a cell surface receptor in hepatic cells which has a high binding affinity for terminal $\beta$-linked galactose (Gal) or N-acetylgalactosamine (GalNAc) residues. ASGP-R was hypothesized to utilize its high affinity for certain glycans to assist in the clearance of glycoproteins from the serum, thus contributing to the clearance of proteins destined for degradation. However, the vast majority of glycoproteins in serum are terminated by an $\alpha$2,3- or $\alpha$2,6-linked sialic acid (Sia), which seems con-
tradictory to the glycoprotein clearance role of ASGP-R considering its high affinity for Gal and GalNAc. However, recent work using wild-type and ASGP-R deficient mice has shown that ASGP-R clears glycoconjugates whose terminal glycan sequence is Siaα2,6GalNAc. ASGP-R is also able to bind core-substituted glycans with the terminus and Siaα2,6Galβ1,4GlcNAc but not those whose glycan terminus is Siaα2,3Gal-β1,4GlcNAc. The authors further hypothesize that the endogenous ligands for ASGP-R are Siaα2,6Gal and Siaα2,6GalNAc [18]. The ASGP-R case is a striking example of how small perturbations in the chemical linkage of an oligosaccharide can have drastically different biological effects.

1.2 Asparagine-linked glycosylation

The Asparagine-linked (N-linked) glycosylation pathway in eukaryotes consists of a set of conserved enzymes in the endoplasmic reticulum (ER) and Golgi apparatus whose function is to attach and modify oligosaccharides at the γ-amido position of asparagine side-chains at the asparagine-X-serine/threonine (N-X-S/T; where X can be any amino acid except proline) sequon in nascent proteins (Figure 1.1 and [19, 20]). The attachment of these saccharide groups is achieved by the use of ‘activated’ sugars, also known as sugar-nucleotide donors. It has been recently estimated that the process of eukaryotic N-glycosylation requires the action of more than two hundred gene products, and defects in many of these have been shown to have a variety of disease-related outcomes [21]. Congenital disorders of glycosylation (CDG) are a disease subtype that refer specifically to patients lacking a functional enzyme component of the N-glycosylation pathway [22, 23]. CDG patients often present with broad clinical symptoms, indicating that proper glycosylation is vital for a number of systems.

1.2.1 Synthesis of sugar-nucleotide donors

The intricate function of glycoconjugates begins with their synthesis. First, monosaccharide precursors are ingested from dietary sources. The nine monosaccharides used in mammalian glycosylation—fructose, galactose, N-acetylgalactosamine, glucose, N-acetylglucosamine, glucuronic acid, mannose, neuraminic acid and xylose—are ingested directly by cells or synthe-
sized from dietary precursors. The monosaccharide precursors are enzymatically activated—primarily in the cytosol—to form sugar-nucleotide donors (mostly UDP, but also GDP and CMP). These sugar-nucleotide donors are utilized by glycosyltransferase enzymes in the extension of sugar chains. The majority of glycosyltransferase action occurs in the Golgi lumen while most sugar-nucleotide donors are synthesized in the cytosol. This necessitates the use of transporters that transport sugar-nucleotide donors across the Golgi membrane for use in N-linked glycosylation [24, 25, 26]. Vital nucleotide-sugar uptake by the Golgi is elegantly achieved by the use of antiporters, which simultaneously expel electroneutral mononucleotides, thus enriching the Golgi lumen in sugar-nucleotide donors, while removing transferase breakdown products.

1.2.2 N-linked glycosylation pathway

Briefly, the process of asparagine-linked glycosylation begins with synthesis of a lipid-linked oligosaccharide precursor (Man$_5$Gn$_2$—dolichol) by enzymes on the cytoplasmic face of the ER. The partially formed dolichol-linked Man$_5$Gn$_2$ oligosaccharide is flipped across the ER membrane into the lumen, where synthesis of Glc$_3$Man$_9$Gn$_2$—dolichol is completed. The oligosaccharide is transferred to a nascent protein en bloc by a multisubunit complex known as the oligosaccharyltransferase complex (Figure 1.1, leftmost oligosaccharide structure). Once transferred, the terminal α1,2-linked glucose is removed by α-glucosidase I (Figure 1.1, step 1). α-glucosidase II removes two terminal α1,3-linked glucoses, but removes the second one much more slowly (step 2). This allows for the singly glucosylated precursor sugar to interact with the protein chaperones calnexin and calreticulin, which assist in protein folding. In addition, a glucosyltransferase is able to add a single glucose to the Man$_9$Gn$_2$ oligosaccharide to allow the nascent protein to interact with the chaperones, in the case that protein folding was not successful (step 3). Many cycles of glucosylation and deglucosylation may occur over several minutes to ensure proper protein folding. An alternative pathway for deglucosylation exists by endomannosidase, which removes a glucose and mannose from Glc$_1$Man$_9$Gn$_2$ to yield a specific isomer of Man$_8$Gn$_2$ (step 4). ER-resident mannosidases remove a single mannose from Man$_9$Gn$_2$ to produce different isomers of Man$_8$Gn$_2$ (steps 5
and 6). Man$_8$Gn$_2$ is then transported to the Golgi apparatus for further processing. Once in the Golgi, the Golgi-resident mannosidases IA, IB and IC trim mannoses from Man$_8$Gn$_2$ to yield Man$_5$Gn$_2$ (step 7), which is followed by the action of N-acetylglucosaminyltransferase I (GnT I), adding a single N-acetylglucosamine to produce GnMan$_5$Gn$_2$ (step 8). There is enzymatic competition for the GnMan$_5$Gn$_2$ substrate. Galactosyltransferases (GalT’s; steps 9 and 10), N-acetylglucosaminyltransferase III (GnT III; step 11), Golgi $\alpha$-mannosidase II (GMII; step 12) and fucosyltransferase (FucT; step 13) are all able to utilize GnMan$_5$Gn$_2$. The actions of the GalT’s and GnT III commit the saccharide to the hybrid type because the sugar residues inserted by these enzymes prevent the action of Golgi $\alpha$-mannosidase II (GMII) [27]. On the other hand, GMII advances the saccharide towards complex structures if it acts on GnMan$_5$Gn$_2$ before either GalT or GnT III has acted. GMII acts with high specificity to remove two mannoses that are $\alpha_{1,3}$- and $\alpha_{1,6}$-linked to the N-glycan core producing GnMan$_3$Gn$_2$. GnMan$_3$Gn$_2$ is a substrate for N-acetylglucosaminyltransferase II (GnT II; step 14) whose action converts the saccharide to the complex type. Fucosyltransferase can also act after GMII to modify the complex saccharide. The actions of various N-acetylglucosaminyltransferases in the Golgi further add branch points to the complex oligosaccharide. As mentioned, N-acetylglucosaminyltransferases I and II add a $\beta_{1,2}$-linked Gn to the $\alpha_{1,3}$- and $\alpha_{1,6}$-linked mannoses respectively (GnT I; a & GnT II; d). Furthermore, GnT IV adds a $\beta_{1,4}$-linked Gn (position b) to the $\alpha_{1,3}$-linked arm of the mannosyl core, while GnT III (position c) adds a $\beta_{1,4}$-linked Gn to the $\beta_{1,4}$-linked mannose of the core. GnT V and VI add $\beta_{1,6}$ and $\beta_{1,4}$-linked Gns to the $\alpha_{1,6}$-linked arm of the trimannosyl core (positions e & f, respectively). These positions (a-f) can be extended by other glycosyltransferases in the medial and trans Golgi to produce large complex oligosaccharide structures. Lastly, the addition of phosphate groups to the carbon-6 of the mannose (steps 15 & 16) signals the mannose-6-phosphate receptor to export the glycosylated protein to the lysosome.
Figure 1.1: An overview of the eukaryotic N-linked glycosylation pathway as adapted from [20] and [28]. The oligosaccharide structures are all linked to an asparagine in the Asn-X-Ser/Thr sequon. Glu = glucose, Man = mannose, GlcNAc = N-acetylglucosamine, Gal = galactose, Fuc = fucose and P = 6-phosphate. See text—particularly Section 1.2.2—for details.
The sequon associated with N-linked glycosylation has been found to occur in nearly 65% of proteins in the SWISS-PROT database, with a sequon-containing protein containing an average of 3.1 N-X-S/T sequons per polypeptide chain. The annotation of the SWISS-PROT database allows one to examine the data in detail and to assess the actual glycosylation state of those proteins containing the N-glycosylation sequon. When considering the measured glycosylation state of sequon-containing proteins and the relatively small contribution of O-linked glycosylation, it has been concluded that more than half of proteins expressed are glycoproteins [29].

Interestingly, there is growing evidence that the evolution of multicellular organisms from unicellular ancestors is closely tied to N-linked glycosylation. As it stands, multicellular organisms decorate their cell surfaces with oligomannose, hybrid and complex N-glycan chains while their unicellular cousins lack the gene encoding GnT I and therefore cannot synthesize hybrid and complex N-glycans. This evolutionary boundary has been correlated with the appearance of multicellular organisms [3]. If this is true, glycosylation would be a striking example of an essential biological process contributing to multicellularity.

N-linked glycosylation has been shown to occur in both prokaryotes and eukaryotes. Though showing some similar elements, the two pathways are quite divergent and contribute to different biological functions (recently reviewed in [30]). For the purpose of this dissertation, the focus will be on N-linked glycosylation in eukaryotes.

1.3 Glycosylation and cancer

Malignant transformation has been shown to be associated with changes in cell surface glycosylation in several cases. When comparing the N-linked oligosaccharides on murine tumour cells by lectin labeling, it has been shown that there is a distinct population of oligosaccharides on metastatic cells vs. non-metastatic cells of the same subtype [31]. Specifically, the loss of sialylated oligosaccharides was shown to correlate with a decrease in metastasis, which was hypothesized to result from increased adhesive properties with components of the extracellular matrix, specifically fibronectin.
Dennis and colleagues showed that $\beta_{1,6}$-branched complex oligosaccharides (Figure 1.1, position e) are directly associated with increased metastasis [32]. These conclusions were based on results seen in a metastatic tumour cell line that showed an increase in $\beta_{1,6}$-branched complex oligosaccharides, as well as a murine nonmetastatic carcinoma cell line whose acquisition of metastatic potential correlated with similar oligosaccharide branching. Dennis et al. were also able to show that mutant tumour cells lacking $\beta_{1,6}$ N-acetylglucosaminyltransferase V activity lacked metastatic potential \textit{in vivo}. MDAY-D2 tumour cell lines lacking $\beta_{1,6}$ N-acetylglucosaminyl-transferase V produced 95% fewer metastases in the liver, and exhibited an overall decrease of 50% in their solid tumour growth rate. These findings demonstrated a clear link between an enzyme in the N-glycosylation pathway and metastatic potential \textit{in vivo}. In addition to oligosaccharide branching, terminal glycan sequences have a role in malignancy as well. A mutant cell line that overexpresses $\alpha_{2,6}$ sialic acid transferase showed a large decrease in metastasis, as well as greater than 50% slower tumour growth [33]. These cells exhibited a high degree of $\alpha_{2,6}$-linked sialic acid on their cell surface, as opposed to the wild-type $\alpha_{2,3}$-linkage.

Further to that finding, it was shown by Dennis and Laferté that NK cells are able to recognize Asn-linked oligosaccharides on tumour cell surfaces when they have certain characteristics [34]. As mentioned earlier, NK cells were first discovered due to their ability to specifically target and kill specific tumour cells. By using Concanavalin A—a lectin derived from plants that is specific for high-mannose oligosaccharides—they were able to show that NK cells were able to target the tumour cells when the cell surface oligosaccharides exhibited high-mannose characteristics, or were incompletely complex. The ability of immune system cells—the NK cells—to detect and clear tumour cells based on their glycosylation status encouraged research interest in pursuing inhibitors of glycosylation events as potential anticancer drugs.

It is theorized that aberrant glycosylation on the tumour cell surface is a result of a ‘high-jack’ of the glycosylation machinery. This glycosylation state (including $\beta_{1,6}$-branching as above), allows the tumour cells to evade the immune system. Inhibiting the production of complex glycans via the use of glycosylation inhibitors allows for immune cells, such as the NK cells, to detect and clear tumour cells from the body (reviewed in [35, 36]).
1.4 Glycoside Hydrolases

Glycosidic linkages have been shown to be remarkably stable, with an estimated half-life of approximately 5 million years for a single glycosidic linkage in cellulose [37]. Glycoside hydrolases catalyze the hydrolytic cleavage of these linkages, making them some of the most proficient enzymes, with rate enhancements on the order of $10^{17}$ [37].

Classification of glycoside hydrolases, or glycosidases, based on amino acid sequences was first performed in the early 1990’s [38]. It was determined that the 301 known glycoside hydrolase sequences could be grouped into 35 families. The substrate specificities and reaction mechanisms of the classified enzymes corresponded to their family placement. Currently, the carbohydrate-active enzyme database (CAZy; www.cazy.org) has classified thousands of glycoside hydrolases into more than one hundred families [39]. The member enzymes of a given family come from a wide variety of organisms, but retain similarity in the types of reactions catalyzed, and the mechanism employed. This indicates that, despite the increase in sequence data, the classification scheme remains robust, as it has been validated by biochemical means. Furthermore, CAZy also classifies other carbohydrate-active proteins, including glycoside transferases, polysaccharide lyases, carbohydrate esterases and carbohydrate-binding modules. In addition, numerous structural studies of glycoside hydrolases from a variety of species have shown that this enzyme class encompasses a large part of the genome, and contains a wide diversity of protein folds [39].

1.4.1 Golgi α-mannosidase II

Golgi α-mannosidase II (GMII; mannosyl oligosaccharide 1,3-1,6-α-mannosidase II; EC 3.2.1.114) is a Golgi-resident glycosyl hydrolase in the N-linked glycosylation pathway that catalyzes the committed step of complex N-linked glycan synthesis. It is a type II membrane protein with a short N-terminal cytoplasmic tail, followed by a single transmembrane helix and luminal linker, followed by a large luminal C-terminal catalytic domain. GMII was first purified and characterized from rat liver Golgi membranes in the Touster lab, and was determined to catalyze the hydrolysis of two mannose-mannose linkages from the oligosaccharide substrate
GnMan$_5$Gn$_2$–Asn to yield GnMan$_3$Gn$_2$–Asn (Figures 1.2 and 1.3) [40, 41]. The enzyme was thereafter discovered to be strongly inhibited by the alkaloid swainsonine (Section 1.5) [42]. GMII has significant sequence similarity across species, and is classified as a member of glycosyl hydrolase family 38 and is also a class II mannosidase [38, 43, 44]. The GMII reaction takes place via a retaining mechanism, whose definitive feature is the conservation of the stereocentre between the substrate and the product (Figure 1.4). Interestingly, the two linkages removed are chemically distinct—$\alpha$1,3 and $\alpha$1,6—but are thought to be cleaved in a single GMII active site and early NMR data indicated that it was likely that the $\alpha$1,6 linkage was the first to be cleaved [45]. This thesis provides strong evidence that this is indeed the case.

1.4.2 Related glycoside hydrolases

Biochemical studies have shown that there is catalytic activity similar to that of GMII in the cell. Several enzymes have been shown to be related by sequence to GMII, thus making them members of the same glycoside hydrolase family. In-depth biochemical and genetic study has shown that there are distinctions between the enzymes, and the reactions they catalyze. It is somewhat unclear whether the substrate specificity of these Family 38 glycoside hydrolase members is fully accurate, though ongoing work will clarify these issues.

**Lysosomal $\alpha$-mannosidase**

Lysosomal $\alpha$-mannosidase (LM) is a GMII-related enzyme that catalyzes the catabolic breakdown of oligosaccharides in the lysosome [46]. It too employs a retaining mechanism, but has a much broader substrate specificity than GMII, cleaving all $\alpha$-linked mannoses from high mannose oligosaccharides. In addition, due to its subcellular localization in the lysosome, it is more highly active at lower pH values [46]. Inhibition of lysosomal $\alpha$-mannosidase produces a phenocopy of $\alpha$-mannosidosis, a lysosomal storage disorder [47].

**Cytosolic $\alpha$-mannosidase**

There is a cytosolic $\alpha$-mannosidase that does not seem to contain a transmembrane helix and has a unique metal dependency for this class of enzyme. It seems to be Co$^{2+}$ activated, and
likely functions to trim mannoses from proteins that have been incompletely glycosylated yet transported to the cytosol [48].

**Golgi α-mannosidase IIx**

Golgi α-mannosidase IIx (GMIIx) is a Golgi-resident enzyme that is similar to GMII. A recent GMII/GMIIx double knockout mouse has been shown to lack complex-type N-glycans [49]. Recombinant mouse GMII and GMIIx enzymes showed identical substrate specificities toward N-glycan substrates (conversion of GnMan$_3$Gn$_2$ to GnMan$_3$Gn$_2$), suggesting that GMIIx is an isozyme of GMII [49].

**Golgi α-mannosidase III**

Golgi α-mannosidase III (GMIII) is a class II cobalt-dependent α-mannosidase from insect (Sf9) cells with amino acid sequence and biochemical similarities to mammalian Golgi α-mannosidase II [50]. GMIII catalyzes the cleavage of two mannosyl linkages, converting Man$_3$Gn$_2$ to Man$_3$Gn$_2$ but it cannot act on GnMan$_3$Gn$_2$, making it quite distinct from GMII. GMIII appears to provide an alternate pathway to complex oligosaccharide synthesis that does not rely on GnT I or GMII.

![Figure 1.2: Golgi α-mannosidase II catalyzes the cleavage of two mannosyl linkages, an α1,3-linkage between M3 and M4 and an α1,6-linkage between M3 and M5, converting GnMan$_3$Gn$_2$ to GnMan$_3$Gn$_2$. The reaction proceeds via a retaining mechanism and a glycosyl-enzyme intermediate.](image-url)
1.5 Golgi α-mannosidase II inhibition and cancer

Inhibition of Golgi α-mannosidase II by the alkaloid swainsonine, also known as 8αβ-indolizidine-1α,2α,8β-triol, has shown some success in \textit{in vivo} experiments and clinical settings. It has been a focus of anticancer work due to its low toxicity, nanomolar inhibition of GMII, and ease of administration. A brief scientific and clinical overview of the role of swainsonine in GMII inhibition and cancer will be presented here.

Swainsonine was first isolated from the Australian plant \textit{Swainsona canescens}, and upon its discovery found to inhibit α-mannosidases. It has subsequently been found and isolated from North American plants of the genera \textit{Astragulus} and \textit{Oxytropis}, and the fungus \textit{Rhizoc-
tonia leguminocola. Treatment by swainsonine was found to produce symptoms similar to those seen in lysosomal storage disorder, an inability to catabolize oligosaccharides in the lysosome. This was believed to be due to the inhibitory effect of swainsonine against lysosomal α-mannosidase [47]. Swainsonine has also been shown to inhibit other Family 38 glycoside hydrolases as well.

It is now understood that swainsonine also inhibits GMII (a class II mannosidase) at a nanomolar level, while exhibiting little inhibition against class I mannosidases. This feature allows the glycans in swainsonine-treated cells to reach GnMan$_3$Gn in the pathway before swainsonine has an effect. Inhibition of GMII prevents the action of subsequent enzymes in the pathway, including N-glucosaminyltransferases II and V, thus diverting the pathway away from complex glycans towards the hybrid state. This has the effect of diverting cell surface glycans away from $\beta_1,6$-branched structures, such as those seen in highly malignant samples.

It has been shown that two cancer cell lines—MDAY-D2 (lymphoid tumour) and B16-F10 (melanoma)—showed reduced metastatic ability when grown in swainsonine for 48 hours prior to their intravenous injection in mice [51, 52]. When swainsonine was added to the drinking water of the mice prior to cell injection, lung colonization by the B16-F10 melanoma cells was further reduced. Furthermore, though swainsonine itself did not inhibit tumour cell growth in vitro, its application assisted in the antiproliferative effect of interferon, a widely-utilized anticancer treatment.

A mouse xenograft model of melanoma was also studied in the context of swainsonine treatment [53]. Briefly, athymic nude mice which had subcutaneously implanted human MeWo melanoma cells were given swainsonine orally or by miniosmotic pumps. Both delivery mechanisms showed a strong effect in reducing the growth rate of the tumours (~ 50%). Interestingly, a glycosylation mutant of the MeWo melanoma was unaffected by swainsonine treatment, both in vitro and in vivo. These results suggest that the mechanism of swainsonine action is via the effect on the glycan pool of these tumour cells.

In vitro effects of swainsonine and 1-deoxymannojirimycin, another GMII inhibitor, have also been linked to the expression of extracellular matrix remodeling enzymes [54]. The decreased metastatic potential noted above was also seen, with a 25-33% decrease in cellular ad-
hesion to reconstituted basement membranes or human umbilical vein endothelial cell monolayers. These effects were correlated to a decrease in human type IV collagenase mRNA and protein levels. Type IV collagenase has been closely tied to the invasive and metastatic ability of tumours, and its loss seems to be an outcome of swainsonine treatment in this study.

The encouraging nature of the above in vitro and in vivo experiments fueled the movement to test swainsonine as an anticancer agent in the clinic. A crucial method was developed to accurately test the amount of swainsonine in the serum of cancer patients participating in the clinical trial [55]. It was shown that the serum concentrations of swainsonine were in the 3–11.8 mg/L range, which represented a concentration one hundred to four hundred times that of the IC$_{50}$ against GMII and lysosomal $\alpha$-mannosidase.

A Phase I study of swainsonine involving nineteen cancer patients at the advanced stage of their disease was undertaken to assess the dose-ranging and pharmacokinetics of swainsonine [56]. The drug was administered intravenously over a five days, with a repetition at 28-day intervals. The maximum tolerated dose (MTD) was determined to be 550 $\mu$g/kg/day, and the MTD–1 level was determined to be 450 $\mu$g/kg/day. Swainsonine seemed to improve the clinical condition of several patients, with one patient who had head and neck cancer showing a $>50\%$ shrinkage in tumour mass. Through the trial, swainsonine was confirmed to have minimal toxicity, even at the MTD. There were some side-effects present, and a high level of oligomannosides in the patient urine indicated that lysosomal $\alpha$-mannosidase was being inhibited as well. A follow-up study further refined the MTD for swainsonine, and also noted side-effects including pulmonary edema [57]. Though, it is unclear whether this is a drug-related or disease-related effect. If the former, a lower MTD is expected, and if the latter, a higher MTD can be tolerated.

A Phase II study of the effect of swainsonine in seventeen patients with renal cell carcinoma was discontinued due to disease progression or toxicity with no anti-tumour activity detected [58]. More than half the patients on the study displayed gastrointestinal side effects, including nausea (53%) and diahrrhea (59%), though not to any severe degree. One of the reasons for the trial failure may have been due to the amount of swainsonine administered, as this study employed a dosage of 37.5 $\mu$g/kg/day for three weeks, followed by a week off in each cycle.
Swainsonine has inhibitory selectivity against Golgi $\alpha$-mannosidase II and lysosomal $\alpha$-mannosidase. The \textit{in vitro}, \textit{in vivo} and clinical work summarized above indicates that GMII inhibition is correlated with a positive clinical outcome in some cancers, while lysosomal $\alpha$-mannosidase inhibition results in symptoms consistent with lysosomal storage disorder and $\alpha$-mannosidosis [59, 60, 61, 62]. These clinical symptoms include accumulation of oligomannosides in the lysosome, as they cannot be cleared. This buildup has a deleterious effect on cellular function, though the exact mechanism of the impairment is not known. Clinically, one can monitor the urine for oligomannosides, as they tend to buildup in tissues and body fluids. However, this is rarely seen in the brain, leading to the hypothesis that $\alpha$-mannosidosis by LM inhibition has little neurological effect.

It should be noted that the clinical trial work with swainsonine does not adequately explain what the specific anticancer effect of this molecule is nor the exact mechanism of clinical improvement. Whether it is direct GMII inhibition, or a combination of actions against other Family 38 glycoside hydrolases remains to be seen. Specific inhibitors against GMII will assist in the study of the action of swainsonine in the cell.

In addition to the potential usefulness of swainsonine in anticancer efforts, it has also been implicated as a possible adjuvant. Swainsonine has been shown to protect and stimulate both the mouse and human hematopoietic systems during treatment with cyclophosphamide and also with 3'-azido-3'-deoxythymidine, thus protecting from chemotherapeutic toxicity [63]. In addition to this finding, very recent work has shown that swainsonine reduces drug resistance of 5-fluorouracil (5-FU) in a colorectal cancer cell line with no exhibited toxicity [64]. This is a significant finding in that glycosylation may play a larger role than originally thought in 5-FU resistance.

### 1.6 Structural study of Golgi $\alpha$-mannosidase II

The GMII gene from \textit{Drosophila melanogaster} was first cloned as a related protein to murine GMII, and was later characterized and determined to be a suitably analogous protein to the mammalian GMII [65, 66]. More specifically, the substrate specificity, pH activity profile and
other biochemical characteristics, as well as sequence alignment (identity & similarity to human GMII: 41% and 61%, respectively), showed dGMII to resemble the human GMII enzyme quite closely.

The dGMII enzyme was expressed in S2 cells, and purified using standard chromatographic techniques. The 1.86 Å resolution atomic structure of the lumenal catalytic domain from GMII from *Drosophila melanogaster* was solved in 2001 by van den Elsen et al [67]. The protein exhibited a single-domain architecture, and populated a new protein fold. The region determined to be the active site by biochemical studied was very highly conserved, and contained a zinc ion, as determined by atomic absorption spectroscopy. The apo-structure of the enzyme was presented alongside a swainsonine–dGMII complex, as well as a crystallographic complex of 1-deoxymannojirimycin–dGMII. In addition, the first detailed view of the structure for this enzyme allowed for the speculation that there might be an N-acetylglucosamine binding site some distance from the catalytic center, and that substrate rearrangement may be a part of the catalytic mechanism.

Continued structural study of dGMII, along with advanced techniques in synthetic organic chemistry have produced many fascinating structures of dGMII bound to inhibitors, sugar analogs, and trapped covalent intermediates. Several analogs of mannostatin A have been studied bound to dGMII in the hopes of developing a strong, specific inhibitor against this enzyme [68]. In addition, recent techniques in developing S-linked sugar analogs, in conjunction with crystal soaking of dGMII and crystallographic structure solution have allowed for the study of the tolerance of the dGMII active site for unusual sugar-like molecules [69, 70]. Most impressively, the development of fluorinated sugar analogs has assited in ‘trapping’ of a covalent intermediate of the dGMII reaction and allowed for detailed speculation regarding the mode of action of the dGMII reaction mechanism [71]. More detailed aspects of these results will be discussed further in the context of the results in Chapters 2-4.

In the space of a few short years, dGMII has been under intense structural and enzymatic study in the hopes that the development of a novel inhibitor against the enzyme would prove to have clinical merit in the treatment of cancer.
1.7 Dissertation overview

The work presented in this dissertation is focused on the catalytic mechanism of GMII. By using the *Drosophila melanogaster* Golgi α-mannosidase II (dGMII) as a model for broader GMII function, a variety of biochemical techniques and X-ray crystallography are utilized to explore inhibitor and substrate binding and to formulate a plausible model for the conserved catalytic mechanism of GMII.

First, the differences between Class I and Class II mannosidases are explored structurally with regards to inhibitor binding and active site distortion. By comparing a variety of high resolution structures of small-molecule inhibitors bound to dGMII, structural and energetic comparisons are made both within and across mannosidase classes. These can be cast within the framework of conformational itineraries for glycoside hydrolases allowing for a detailed examination of the conformational states likely present in the GMII mechanism.

Second, the atomic structure of dGMII binding a natural substrate oligosaccharide (GnMan$_5$Gn, Figure 1.3) allows for the detailed examination of enzyme–substrate interactions and identification of key residues involved in substrate binding. Utilizing multiple sequence alignments of GMII allows for the examination of the mode of substrate binding as a conserved feature of GMII. The method by which GMII binds its natural substrate is explored in atomic detail, allowing for the identification of multiple highly conserved binding subsites that play key roles in the substrate specificity and overall catalysis.

Third, an in-depth analysis of the substrate distortion, as well as natural substrate binding, allows for the formulation of the complete catalytic mechanism of GMII. The proposed multistep mechanism is explored in detail, and with the knowledge of the unliganded structure of GMII, as well as lysosomal alpha-mannosidase (a close enzymatic cousin of GMII), it is believed that the mechanism elegantly serves its purpose and is unique among this class of enzyme.
Chapter 2

Substrate distortion in the dGMII active site

2.1 Chapter overview

This chapter examines the distorting ability of the GMII active site. This is achieved by the close structural and biochemical examination of several enzyme—inhibitor structures, and their comparison to another class I mannosidase. This examination proves useful in hypothesizing the conformational itinerary in the catalytic mechanism of GMII.

Briefly, wild-type dGMII enzyme was complexed with kifunensine, a weak inhibitor of retaining mannosidases. The crystallographic complex was solved, and the structure examined, particularly the conformation of kifunensine. Kifunensine and 1-deoxymannojirimycin were examined from the standpoint of dGMII binding vs. ER mannosidase I (ERMI) binding.

Enzyme kinetics experiments were performed using dGMII in conjunction with swainsonine, 1-deoxymannojirimycin and kifunensine, in order to confirm that the $K_i$ values corresponded with the measured IC$_{50}$ values.

The distorting ability of the dGMII active site was examined vis-à-vis strong vs. weak inhibition, and catalytic mechanism.

The material presented in this chapter is modified from a published work by Shah et al [72].
2.2 Materials and Methods

2.2.1 Protein expression and purification

Wild-type dGMII was expressed and purified in *Drosophila melanogaster* S2 cells as described previously [67]. Briefly, the cDNA for the lumenal catalytic domain of dGMII was cloned into an inducible vector and used to stably transform *Drosophila melanogaster* S2 cells. Single-cell clones that expressed dGMII highly, as determined by the pNP-mannose assay, were selected and grown in serum-free medium. Once scaled up to Fernbach flasks, the dGMII was purified from the culture supernatant by standard chromatographic techniques using blue agarose and Ni-NTA. The protein was dialyzed into storage buffer, and concentrated to 20–25 mg/mL prior to storage at −80°C.

2.2.2 Inhibition assay

![Chemical structure](image)

Figure 2.1: *p*-nitrophenyl mannopyranoside is hydrolyzed by dGMII to yield mannose and a charged *p*-nitrophenyl moiety. When in a basic solution (such as sodium carbonate), the freed *p*-nitrophenol moiety shows increased absorption at λ = 405 nm.

Inhibition of α-mannosidase activity was carried out in microtitre plates in a final volume of 50 µL. Inhibitors were dissolved in water to a final concentration of 200 mM. The reaction mixture consisted of 25 µL of varying concentrations (1-10 mM) *p*-nitrophenol-α mannopyranoside (*p*NP-mannose, Sigma-Aldrich Canada Ltd., Oakville, ON), 10 µL of 200 mM buffer and 10 µL of water or inhibitor (Figure 2.1). The buffer used was MES pH 5.75 which is optimal for this enzyme [66]. The reaction mixture was pre-warmed to 37°C and 5 µL of α-mannosidase in 10 mM sodium phosphate pH 6.8, 100 mM NaCl was added to initiate the reaction. The
amount of enzyme added was that which is necessary to keep the reaction in the linear range. This represented approximately 350 ng of protein for a 15-minute reaction. At the endpoint, the reaction was stopped using 50 µL of 0.5 M sodium carbonate. The absorbance of the reaction was measured at 405 nm with 520 nm background correction on a micotitre plate reader. The measurements were taken in triplicate in the absence and presence of each inhibitor to determine Michaelis-Menten kinetics. $K_i$ values were derived from the $x$-intercepts of the double reciprocal plots.

2.2.3 Crystallization

Crystallization of dGMII was carried out as described in [67]. Briefly, 2 µL of concentrated protein solution (20 mg/mL in 10 mM Tris pH 8, 100 mM NaCl) was combined with 2 µL of reservoir buffer (100 mM Tris pH 7, 8.5% w/v PEG 6000, 2.5% v/v 2-methyl-2,4-pentanediol) to form the crystallization drop on a siliconized glass cover slip. The coverslip was sealed atop a well of a 24-well plate containing 0.6 mL reservoir buffer (as above). After 16-18 hours of growth, the crystals were fully grown (0.4—0.7 mm) and ready to be exposed to inhibitor.

2.2.4 Inhibitor soaking and crystal freezing

The crystals were extracted from the crystallization drop using a CryoLoop (Hampton Research, Aliso Viejo, CA, USA) and soaked in phosphate reservoir buffer: 100 mM sodium phosphate pH 7, 8.5% w/v PEG 6000, 2.5% v/v 2-methyl-2,4-pentanediol. Kifunensine was added in phosphate reservoir buffer at 10 mM concentration. Three 2 µL additions of kifunensine were added to the drop, each one followed by a removal of 2 µL of solution from the drop. Each addition and removal step was followed by a soak time of approximately 30 minutes. The overall effect of this technique was to reduce the Tris concentration in the drop while simultaneously increasing the kifunensine concentration.

The crystals were then removed from the phosphate and kifunensine drop using a 0.4 mm CryoLoop and exposed to phosphate reservoir buffer & kifunensine with increasing concentrations of 2-methyl-2,4-pentanediol (5 → 10 → 15 → 20% v/v) as a cryoprotectant and flash-frozen in a nitrogen stream. They were tested for preliminary diffraction, and then stored in
liquid nitrogen for transport to the synchrotron source.

## 2.2.5 Data collection and reduction

The crystals were exposed to X-ray radiation of wavelength of 0.9504 Å and at a temperature of 100 K at beam-line F1 at the Cornell High Energy Synchrotron Source (CHESS) at Cornell University in Ithaca, NY, USA. Data were collected using dual ADSC Quantum-4 CCD detectors. Typically, 300 diffraction images were collected per crystal with a 0.5° oscillation per frame. *DENZO* and *SCALEPACK* were used to process the data [73]. Final reduction statistics are presented in Table 2.1.

### Data collection

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>Resolution (Å) 1.80</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>Reflections 93117</td>
</tr>
<tr>
<td>a,b,c (Å)</td>
<td>R&lt;sub&gt;work&lt;/sub&gt; / R&lt;sub&gt;free&lt;/sub&gt; 0.200 / 0.222</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>No. Atoms</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>Protein 8181</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;</td>
<td>Ligand 16</td>
</tr>
<tr>
<td>&lt;I&gt; / σ &lt;I&gt;</td>
<td>Water 991</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>B-factors</td>
</tr>
<tr>
<td>Redundancy</td>
<td>Protein 13.66</td>
</tr>
<tr>
<td></td>
<td>Ligand 17.86</td>
</tr>
<tr>
<td></td>
<td>Water 24.06</td>
</tr>
<tr>
<td>R.M.S. deviations</td>
<td>Bond lengths (Å) 0.0058</td>
</tr>
<tr>
<td></td>
<td>Bond angles (°) 1.34</td>
</tr>
<tr>
<td>Ramachandran analysis</td>
<td>Favoured (%) 96.8</td>
</tr>
<tr>
<td></td>
<td>Allowed (%) 99.8</td>
</tr>
<tr>
<td></td>
<td>&lt;µ&gt; by Luzzati (Å) 0.172</td>
</tr>
</tbody>
</table>

Table 2.1: Data reduction and refinement statistics for the kifunensine−dGMII complex.

## 2.2.6 Atomic structure determination

The previously solved structure of dGMII was used in conjunction with the collected data to generate σ<sub>A</sub>-weighted difference maps (Fo −Fc) using CNS [74]. These maps unambiguously showed the position and orientation of the kifunensine in the active site (Figure 2.2). The calculated electron density maps were verified from a second, independently grown and treated

Crystal. Kifunensine parameter and topology files were generated using the HIC-Up server [75].

Figure 2.2: Electron density of kifunensine in the dGMII active site. The $\sigma_A$-weighted unbiased difference density $(2F_o - F_c)$ is presented to 1.80 Å resolution and was contoured to 1.4$\sigma$ (0.62 electrons/Å$^3$). This density was calculated before inclusion of the inhibitor in the calculations.

2.2.7 Kifunensine–dGMII interactions

To determine the protein residues directly involved in kifunensine binding, an analysis was performed using HBPLUS/LIGPLOT [76, 77]. Briefly, hydrogens were added to the inhibitor and protein using HBPLUS, which then evaluated KIF–dGMII interactions for distance, geometry and chemical relevance. The calculated hydrogen bonds and hydrophobic interactions were visualized using LIGPLOT. dGMII-KIF interactions are presented in Figure 2.5 and Table 2.3. A similar analysis was repeated for 1-deoxymannojirimycin and swainsonine in the dGMII active site (Table 2.3).

2.2.8 Conformational energy calculations

The coordinates of kifunensine were isolated from both dGMII and Homo sapiens endoplasmic reticulum mannosidase I from (hERMI) and were subjected to a molecular dynamics simulation using the DISCOVER suite in InsightII [78]. The parameters employed for this analysis
were $n_{\text{equilibration}}=100$, $n_{\text{steps}}=1000$, $n_{\text{history}}=10$ and InsightII performed its conformational calculations in vacuo. Since the molecular dynamics simulation performed is quite ‘short,’ the selection of an in vacuo environment should not have a significant effect. The resulting energy values are reported in Table 2.2.

<table>
<thead>
<tr>
<th></th>
<th>Isolated from hERMI</th>
<th>Isolated from dGMII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>97.590 ± 2.060</td>
<td>101.457 ± 0.987</td>
</tr>
<tr>
<td>Potential Energy</td>
<td>73.073 ± 3.000</td>
<td>76.413 ± 2.881</td>
</tr>
<tr>
<td>Kinetic Energy</td>
<td>24.518 ± 2.767</td>
<td>25.044 ± 2.758</td>
</tr>
</tbody>
</table>

Table 2.2: Kifunensine conformational energies (kcal/mol)

2.3 Results

2.3.1 Inhibition of dGMII

It was determined that kifunensine was a weak inhibitor of dGMII, with an $K_i$ value of 5.2 mM. Swainsonine, by comparison, was confirmed as a very strong inhibitor of dGMII with a $K_i$ of 10.5 nM under identical conditions. 1-deoxymannojirimycin was determined to inhibit dGMII at an intermediate level ($K_i = 610 \mu$M). Together, these three inhibitors represent an inhibition range of nearly six orders of magnitude against dGMII. Interestingly, by crystal structure determination, it is possible to clearly see these three inhibitors in the dGMII active site.

2.3.2 Obtaining a crystal complex

Initial co-crystal trials failed to show the presence of kifunensine in the active site. The electron density seen in the active site subsequent to typical crystal soaking conditions was attributable to a single 2-amino-2-hydroxymethyl-1,3-propanediol (Tris) molecule. Tris is a weak inhibitor of class II mannosidases and has been seen in the active site of dGMII ([67], PDB Structure ID:1HTY) and in the lysosomal $\alpha$-mannosidase structure ([79], PDB Structure ID:1O7D). Due to the presence of this interfering Tris molecule, an attempt was made to grow crystals in the presence of phosphate. Although this strategy could produce large, prismatic crystals in
this manner, they failed to diffract beyond 7 Å. A method was therefore developed to remove the bound Tris by soaking the molecules in a solution based on the mother liquor with the Tris replaced by sodium phosphate. Initial crystallographic studies of these phosphate-soaked crystals indicated the absence of clear Tris density in the active site, so kifunensine was added to phosphate-washed crystals. Kifunensine was also added to the solutions used for cryoprotection prior to freezing to prevent the loss of this weakly bound substrate. Kifunensine-soaked dGMII crystals were assessed for quality on an home X-ray source, and those with the best diffraction and lowest mosaicity were analyzed on the F1 beamline at CHESS. Data sets to a conservative cutoff of 1.8 Å resolution were collected from two crystals, though some weak diffraction was observable beyond that limit.

2.3.3 Structure of kifunensine–dGMII complex

![Chemical structures of the inhibitors examined in this chapter. (A) Kifunensine, (B) swainsonine, and (C) 1-deoxymannojirimycin; abbreviated KIF, SWA, and DMJ, respectively.](image)

Kifunensine binds to dGMII in the active site and is stabilized by interactions with a variety of residues (Figure 2.4 and PDB Structure ID: 1PS3). The 2’ and 3’ hydroxyl groups on the ‘pyranosyl-like’ portion of kifunensine coordinate the zinc ion at the active site (Figure 2.3A). The six-membered ring of kifunensine stacks against W95 of the enzyme, with W415 stacking against the C5-C6-C7 portion of the inhibitor. There are many inhibitor-enzyme interactions, including hydrogen bonds with putative catalytic residues D204, D341 and D472. These and other interactions are summarized in Table 2.3 and Figure 2.5. The mode of binding is reminiscent of the extremely efficacious class II α-mannosidase inhibitor swainsonine, whose 7- and 8-hydroxyl groups on the five-membered ring also coordinate the zinc ion, and similar stacking
and hydrogen bonding interactions are also seen (Figure 2.3B). It is also similar to the binding of 1-deoxymannojirimycin to dGMII as that ‘pyranosyl-like’ inhibitor makes similar contacts and coordinates the zinc with its 2’- and 3’- hydroxyls as well (Figure 2.3C) [67]. The pyranose portion of kifunensine adopts a $1,4^\beta$ boat conformation in the dGMII active site and has a prominent kink along the C$_1$–N bond. The kifunensine ring hydroxyls adopt an all-equatorial conformation, in contrast to the all-axial conformation seen in the kifunensine–hERMI crystal complex structure [80].

Figure 2.4: Kifunensine in the dGMII active site. There are many KIF-residue interactions, including hydrogen bonding with protein and water (yellow and cyan dashes, respectively), metal coordination (green dashes) and stacking with aromatic side chain of W95. See text and Table 2.3 for details.

2.3.4 Kifunensine conformational energy

The total energy of kifunensine (the sum of potential and kinetic energies) was calculated to be 101.46 kcal/mol in dGMII as compared to 97.59 kcal/mol in hERMI (Table 2.2). This difference of 4 kcal/mol may be attributed to the kinked conformation that kifunensine adopts in the dGMII active site, as opposed to the planar conformation in hERMI. The energy difference cannot be attributed to protein-inhibitor interactions, as the energy calculation was performed on kifunensine after it had been isolated from the active sites of the two enzymes.
2.4 Discussion

2.4.1 Kifunensine conformations in hERMI and dGMII

Kifunensine forms similar types of protein–inhibitor interactions in both dGMII and hERMI as a considerable number of hydrogen bonds are seen between amino acid residues and ring hydroxyls, as well as stacking interactions with aromatic side chains. These types of interactions are very reminiscent of those seen in many complexes of protein and carbohydrates [81]. However, biochemical assays have shown that kifunensine is a strong inhibitor of class I α-mannosidases, but is rather ineffective against class II α-mannosidases, including dGMII. The details of the kifunensine in the dGMII active site may indicate what features make kifunensine a poor inhibitor in class II α-mannosidases. Also, an examination of the interacting residues of dGMII might help to understand the elements of active site architecture which lead
Table 2.3: Atomic interactions of dGMII with kifunensine, swainsonine and 1-deoxymannojirimycin (KIF, SWA and DMJ respectively). All distances shown are in Å.

<table>
<thead>
<tr>
<th>dGMII</th>
<th>KIF Distance</th>
<th>SWA Distance</th>
<th>DMJ Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>D204 OD1</td>
<td>O₂H 2.92</td>
<td>N₄ 2.88</td>
<td>O₂H 2.75</td>
</tr>
<tr>
<td>D341 OD2</td>
<td>N₉ 2.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D472 OD1</td>
<td>O₄H 2.61</td>
<td>O₁H 2.51</td>
<td>O₄H 2.80</td>
</tr>
<tr>
<td>D472 OD2</td>
<td>O₃H 2.37</td>
<td>O₇H 2.61</td>
<td></td>
</tr>
<tr>
<td>D92 OD1</td>
<td>O₂H 2.75</td>
<td>O₈H 2.91</td>
<td>O₃H 2.90</td>
</tr>
<tr>
<td>R876 O</td>
<td>O₆H 2.76</td>
<td></td>
<td>O₆H 2.60</td>
</tr>
<tr>
<td>Y269 OH</td>
<td>N₉ 3.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y727 OH</td>
<td>O₄H 2.72</td>
<td>O₃H 2.69</td>
<td></td>
</tr>
<tr>
<td>Zn</td>
<td>O₂H 2.41</td>
<td>O₇H 2.31</td>
<td>O₂H 2.41</td>
</tr>
<tr>
<td></td>
<td>O₃H 2.37</td>
<td>O₈H 2.30</td>
<td>O₃H 2.33</td>
</tr>
<tr>
<td>Kᵢ</td>
<td>5.2 mM</td>
<td>10.5 nM</td>
<td>610 µM</td>
</tr>
</tbody>
</table>

Figure 2.6: Superposition of kifunensine and 1-deoxymannojirimycin from two α-mannosidases: dGMII (yellow) and hERMI (red). (A) Kifunensine binds to the enzymes in two distinct conformations in each enzyme: $^{1,4}B$ in dGMII and $^1C_4$ in hERMI. In addition, there is an all-equatorial conformation for the pyranose ring hydroxyls in dGMII, as contrasted to an all-axial conformation in hERMI. (B) 1-deoxymannojirimycin also adopts two distinct chair conformations in each enzyme: $^4C_1$ in dGMII and $^1C_4$ in hERMI.

Kifunensine binds to both dGMII and hERMI in non-standard conformations as shown in Figure 2.6A. The piperdine ring of kifunensine adopts a $^1C_4$ chair conformation in the hERMI active site, with an all-axial configuration of ring hydroxyls [80]. It is hypothesized that this conformation exists due to the presence of the fused five-membered ring to the piperdine portion of kifunensine. Neither the $^1C_4$ conformation in hERMI nor the $^{1,4}B$ conformation in dGMII correspond to low energy $^4C_1$ conformation of mannose derivatives [82]. However, the
\(^1C_4\) conformation has been shown to be of a lower energy than \(^1,^4B\), and thus closer to the expected global minimum of \(^4C_1\) [83]. This is supported by the energy calculations, which show a 4 kcal/mol increase in energy between the hERMI- and dGMII-bound kifunensine forms. The two distinct conformations of the kifunensine six-membered ring also reflect the different reaction mechanisms of these two classes of enzymes, in particular the distinct conformational itineraries taken by their substrates during catalysis.

### 2.4.2 Comparison of dGMII inhibitors

High-resolution crystal structures of dGMII complexes with kifunensine, swainsonine and 1-deoxymannojirimycin have been solved (Figure 2.3) [67]. These three molecules represent a span of dGMII inhibitor constants over six orders of magnitude (Table 2.3).

The structural comparison of kifunensine, swainsonine, and 1-deoxymannojirimycin in dGMII demonstrates the similarity of metal coordination (Figure 2.7). Kifunensine coordinates the active site zinc with its O2 and O3 positions at respective distances of 2.41 and 2.37 Å. Swainsonine is somewhat closer to the zinc atom, as its O7 and O8 coordination distances are 2.31 and 2.30 Å, respectively. Lastly, 1-deoxymannojirimycin has coordination distances very similar to that of kifunensine, with its O2 and O3 positions at 2.41 and 2.33 Å from the zinc. All three of these inhibitors make very significant stacking interactions with W95, as well as many hydrogen bonds with acidic residues in the vicinity of the active site (Table 2.3). It is interesting to note that the weakest inhibitor, kifunensine, is oriented so that it occupies a region of the active site shown to be unoccupied in dGMII–DMJ and dGMII–swainsonine. It could initially be assumed that this indicates that there are unfavourable interactions taking place in that part of the active site, reducing the efficacy of kifunensine as an inhibitor. However, closer examination of the KIF–dGMII complex structure indicates that the O7 and O8 carbonyls of kifunensine point away from the active site pocket. This indicates that the differences in inhibition may stem from conformational mimicry of the transition state.
Figure 2.7: Superposition of inhibitors in dGMII active site. The molecules shown are kifunensine (yellow), 1-deoxymannojirimycin (green), swainsonine (salmon), and the dGMII active site zinc (green sphere). These molecules represent a range of dGMII inhibition over five orders of magnitude.

2.4.3 Inhibitor conformations in hERMI and dGMII

The inhibitors kifunensine and 1-deoxymannojirimycin have been shown in the active sites of both dGMII and hERMI. As mentioned, the six-membered ring of kifunensine adopts a $^{1,4}B$ conformation with a prominent C$_1$–N kink in dGMII and a $^{1}C_4$ conformation in hERMI.

1-deoxymannojirimycin is a smaller inhibitor which contains a single piperidine ring (Figure 2.3C). This inhibitor is forced into distinct chair conformations in each active site, as seen by a $^{4}C_1$ conformation in dGMII and $^{1}C_4$ conformation in hERMI (Figure 2.6B).

2.4.4 Implications in the context of substrate binding and the catalytic mechanism

The examination of competitive inhibitors in the dGMII active site allows the mode of substrate binding to be predicted. All three of the inhibitors studied in this chapter show significant similarity to saccharide moieties, particularly in their mode of binding. The residues that are important for inhibitor binding likely play a key role in coordinating the natural substrate of dGMII. Interestingly, the size of these inhibitors is quite small as compared to the full substrate, so it is expected that the binding of the full dGMII substrate will require the contribution of other amino acids in addition to the ones mentioned here. This is explored further in Chapters 3 and 4.
There has been significant progress in recent years in defining the conformational states adopted by substrates, transition state analogues and inhibitors on binding to glycosyl hydrolase catalytic sites, with emphasis primarily (but not exclusively) on retaining β-glycosidases. Recently, Davies et al. have cast these observations into the context of conformational itineraries that a saccharide ring undergoes in the process of hydrolysis, and have begun to assign the particular itineraries undertaken by various glycosyl hydrolase families (Figure 2.9) [84, 85].

<table>
<thead>
<tr>
<th>Glycosyl hydrolase classification</th>
<th>Class I</th>
<th>Class II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal ion dependance</td>
<td>Family 47</td>
<td>Family 38</td>
</tr>
<tr>
<td>Topology</td>
<td>Ca(^{2+})</td>
<td>Zn(^{2+}), Co(^{2+})</td>
</tr>
<tr>
<td>Mannose linkages hydrolyzed</td>
<td>Type II TM</td>
<td>Type II TM</td>
</tr>
<tr>
<td>Stereochemistry of product(^{1})</td>
<td>α(_{1,2})</td>
<td>α(<em>{1,2}; α</em>{1,3}; α_{1,6})</td>
</tr>
<tr>
<td>Optimal inhibitors</td>
<td>Inverted</td>
<td>Retained</td>
</tr>
<tr>
<td></td>
<td>Pyranose analogs</td>
<td>Furanose analogs</td>
</tr>
</tbody>
</table>

Table 2.4: Comparison of class I and class II α-mannosidases [72]. \(^{1}\)The glycosyl hydrolase inverting and retaining mechanisms are depicted in Figures 2.8 and 1.4 respectively.

Figure 2.8: General mechanism of the inverting glycosyl hydrolase reaction. The inversion of stereochemistry at the nucleophilic site without the formation of a glycosyl-enzyme intermediate are defining features of this reaction.

As Davies’ review points out, the \(^{1}C_{4}\) kifunensine conformation observed in the family 47 inverting hERMI is consistent with this enzyme following a series of conformations passing through a \(^{3}H_{4}\) transition state during the course of the reaction (Figure 2.8 and Table 2.4). The data presented here on the \(^{1,4}B\) structure for kifunensine in dGMI suggest that this family 38 retaining enzyme proceeds via a \(B_{2,5}\) transition state. This proposal supports the recent obser-
viation of covalent intermediates in dGMII, which suggests that the glycosyl hydrolase family 38 mechanism follows an itinerary opposite to that of the family 26 retaining β-mannanase [71]. Although the glycosylation steps for both pass through a $B_{2.5}$ transition state, in family 38 the covalent intermediate adopts a $1S_5$ conformation, while in the case of family 26 this conformation is seen for the Michaelis complex.

As pointed out by Davies et al., the early supposition that glycosyl hydrolases pass through a transition state with a $4H_3$ configuration has been modified with the accumulation of more data. The results presented here show that the conformations of common inhibitors in enzymes of different families can shed light on the pathways followed by those enzymes, and on their likely transition state configurations. This information is particularly valuable in support of efforts to derive specific inhibitors of glycosidases with medical importance.

Finally, the inhibitor distortion seen in the dGMII active site is likely related to substrate catalysis. The oligosaccharide substrate of dGMII is a large molecule and is highly flexible. The contribution of active site residues to substrate distortion is explored in following chapters.
Chapter 3

Crystallographic structure of the substrate—dGMII complex

3.1 Chapter overview

The structure of the GMII—substrate complex is presented in this chapter, along with a detailed examination of the GMII enzyme family. The mode of substrate binding and the role of several key portions of the substrate oligosaccharide are examined in detail.

Briefly, a nucleophile mutant of dGMII—dGMII(D204A)—was crystallized using wild-type dGMII seeds and complexed with purified GnMan₅Gn. The nucleophile mutant was necessary for the prevention of hydrolysis prior to the crystallographic study. The substrate-enzyme complex was examined closely, particularly regarding the key residues in substrate interaction. Three subsites were discovered which provides a basis for the discussion of the catalytic mechanism in the following chapter. Furthermore, the examination of the sequence conservation in the three subsites, as well as the substrate requirements in this portion of the N-glycosylation pathway help to shed more light on the importance of this enzymatic step.

Some material presented in this chapter forms the basis of the work presented in Shah et al [86].
3.2 Materials and Methods

3.2.1 Protein expression and purification

Wild-type dGMII and the D204A mutant were expressed and purified in *Drosophila melanogaster* S2 cells as described previously [67]. Briefly, the cDNA for the lumenal catalytic domain of dGMII was cloned into an inducible vector and used to stably transform *Drosophila melanogaster* S2 cells. Single-cell clones that expressed dGMII highly, as determined by the pNP-mannose assay, were selected and grown in serum-free medium. Once scaled up to Fernbach flasks, the dGMII was purified from the culture supernatant by standard chromatographic techniques using blue agarose and Ni-NTA. The protein was dialyzed into storage buffer, and concentrated to 20–25 mg/mL prior to storage at −80°C. In the case of the D204A mutant, site-directed mutagenesis was performed using the QuickChange method.

3.2.2 Crystallization

Typically, mutant dGMII crystallization is dependent on the use of wild-type dGMII crystal microseeds in order to overcome the crystallization nucleation barrier. Due to the high velocity of the glycosyl hydrolase reaction, and the relative scarcity of the natural substrate, it was necessary to eliminate or greatly minimize the amount of wild-type dGMII in the crystallization experiment. Attempts to grow mutant crystals directly from mutant microseeds were unsuccessful. In order to grow high-quality D204A crystals, while reducing the amount of wild-type enzyme in the final crystal, the following strategy was employed: the initial use of dGMII wild-type microseeds, followed by successive rounds of mutant seed preparation to minimize the amount of wild-type dGMII in the crystallization experiment.

Wild-type dGMII crystals were grown using the Nextal Crystallization Tool (Qiagen Inc., Hilden, Germany) by hanging-drop vapour diffusion. Concentrated protein solution (20 mg/mL) was mixed in a 1:1 ratio with reservoir solution containing 0.1 M Tris pH 7, 8.5% w/v PEG 8000 and 2.5% v/v MPD and suspended over reservoir (0.6 mL) for 18 hours. The resulting single crystals were used to prepare wild-type dGMII microseeds using the SeedBead protocol (Hampton Research, Aliso Viejo, CA, USA). These microseeds were used to seed crystal
growth of D204A dGMII nucleophile mutants in the above conditions, though starting with concentrated (12 mg/mL) D204A protein solution in place of wild-type dGMII. The resulting D204A crystals were once again subjected to the SeedBead protocol to make mutant crystal seeds and minimize the amount of wild-type enzyme in the crystals. The final D204A microseeds were used to grow crystals of the D204A mutant.

D204A protein crystals were grown using hanging-drop vapour diffusion in the above-mentioned conditions for 16-18 hours.

3.2.3 Substrate isolation and purification

The GMII substrate (GnMan$_5$Gn$_2$, Figures 1.2 and 1.3) is an intermediate in the N-glycosylation pathway and exhibits extremely low bioavailability as it is usually acted upon by GMII very quickly after the GnT I reaction [87]. Due to its low abundance in biological systems, and the low production amounts available by chemical synthesis methods, the decision was made to attempt to isolate and purify a precursor (Man$_5$Gn$_2$) which could then be reacted with GnT I in vitro to produce the required substrate.

Chicken egg ovalbumin

The first target for purification was chicken egg ovalbumin. Ovalbumin is a 44 kD protein with a single Asn-X-Ser/Thr sequence that is glycosylated with a heterogeneous population of carbohydrates. In 1972, Huang and colleagues outlined a glycopeptide fractionation procedure that they conducted on ovalbumin using pronase to exhaustively digest the protein into glycopeptides, most of which carried only a single amino acid (Asn); pronase does not digest glycans [88]. These larger glycopeptides were separated from the smaller non-glycosylated amino acids and peptides using gel filtration, followed by further purification of the glycopeptides on an ion-exchange column. One of the peaks eluting from the ion-exchange column contained approximately 90% Man$_5$Gn$_2$–Asn.

Chicken egg ovalbumin was obtained from commercial sources (Sigma-Aldrich Canada Ltd., Oakville, ON) and was treated with pronase exhaustively. The resulting peptides and glycopeptides were separated on a Sephadex G-25 column. Carbohydrate-containing frac-
tions as determined by the phenol-sulfuric acid assay [89] were pooled and lyophilized and the remaining residue was dissolved in 0.5 mL H$_2$O. The glycopeptides were applied to a 2 metre-long Dowex AG-50 column that was held at pH 3.5 with 1 mM sodium acetate. The ionic strength of the buffer was not sufficient to hold the pH at 3.5, and separation of the glycopeptides could not be achieved. Previous work indicated that the low ionic strength of the buffer was crucial for oligosaccharide separation, and indeed an increase the buffer concentration resulted in a lack of separation.

**Genetically engineered *Pichia pastoris***

*Pichia pastoris* is a methyltropic yeast that has emerged recently as a strong candidate system for heterologous protein expression. Contreras and colleagues were able to elegantly engineer a strain of *P. pastoris* so that its glycosylation pattern was overwhelmingly Man$_5$Gn$_2$ [90, 91]. This modified strain of yeast was grown to saturation and its cell-wall mannoproteins were isolated by treatment with citrate under high heat and pressure. While the isolation of mannoproteins from the non-protein cell-wall constituents was successful, the procedure was brought to a halt there. Unfortunately a key chemical component required for the deglycosylation step was unavailable due to Government of Canada security restrictions.

**Bovine pancreatic ribonuclease B**

Ribonuclease B from the pancreas of *Bos taurus* is a commercially available product with an intriguing glycosylation pattern. It is almost universally decorated at its single glycosylation site with high-mannose oligosaccharides ranging from Man$_{5–9}$Gn$_2$. As determined by Fu and colleagues, Man$_5$Gn$_2$ is an enriched glycan on this protein [92]. It was envisioned that the high-mannose oligosaccharides could be isolated in a quantity suitable for structural study, and could then be treated by GnT I, which would be specific for Man$_5$Gn$_2$, creating the product GnMan$_5$Gn$_2$.

Bovine pancreatic ribonuclease B (Sigma-Aldrich Canada Ltd., Oakville, ON) was dissolved in a denaturing solution of 0.5% w/v SDS and 40 mM DTT to a concentration of 12 mg/mL and incubated at 95°C for 30 minutes to denature the protein. The denatured RNaseB
was treated with sodium citrate to a final concentration of 50 mM pH 5.5 at which point 10 000 units of Endo H$_f$ (New England Biolabs, Ipswich, MA) were added, and the mixture was incubated at 37°C for 18 hours to release the high-mannose oligosaccharides. Endo H$_f$ catalyzes the hydrolysis between the first and second N-acetylglucosamine of high-mannose oligosaccharides. The oligosaccharide release was monitored by SDS-PAGE (Figure 3.1).

![Figure 3.1: Deglycosylation of bovine pancreatic ribonuclease B. The lanes marked 'M' denote Mark12 protein standard, while the time point of the deglycosylation reaction is indicated above each lane in hours. The gel was stained using Coomassie Brilliant Blue G-250.](image)

The mixture was applied to a BioGel P-10 column (15 × 200 mm) equilibrated with 50 mM sodium citrate pH 5.5 and run by gravity to separate the released oligosaccharides from the digested protein. Fractions (1.0 mL) were collected and tested for protein content by the BCA Protein Assay (Pierce Biotechnology Inc., Rockford, IL, USA), and for carbohydrate content by the phenol sulfuric acid assay (Figure 3.2) [89]. Carbohydrate-containing fractions were pooled and lyophilized. The resulting residue was dissolved in 1.0 mL H$_2$O and applied to a PD-10 column equilibrated with H$_2$O for the purposes of buffer exchange. The carbohydrate-containing fractions—as determined by the aforementioned phenol sulfuric acid assay—were
pooled, lyophilized, and dissolved in 0.6 mL H₂O.

![Figure 3.2: Separation of oligosaccharides from deglycosylated bovine pancreatic ribonuclease B. The blue trace represents protein content by the A₅₉₅ reading using the Bradford assay (primary y-axis) and the red trace shows carbohydrate content by the A₄₉₀ reading using the phenol sulfuric assay (secondary y-axis).](image)

In order to add the GlcNAc required for GMII function (G3 in Figure 1.2), the RNaseB oligosaccharides were treated with N-acetylglucosaminyltransferase I in a reaction similar to the one performed by Chen et al [93]. Briefly, the oligosaccharides (which acted as the acceptor substrate in the reaction) were dissolved to a concentration of 1 mM in a mixture containing 50 mM MES pH 6.5, 3 mM AMP, 15 mM MnCl₂, 60 mM GlcNAc and 2 mM UDP-GlcNAc. The MES was effective in maintaining the Golgi-like pH, while Mn²⁺ is known to be a necessary cofactor for the GnT I enzyme. AMP was added to inhibit phosphatases that may be in the GnT I-containing SF9 culture supernatant and may degrade the donor substrate UDP-GlcNAc. The high concentration of GlcNAc was added because the GnT I-containing SF9 culture supernatant may contain a small amount of hexosaminidase that could remove the added GlcNAc residue from the GMII substrate, thus reversing the GnT I reaction. N-acetylglucosamine is known to be inhibitory against hexosaminidase, so the high concentration of GlcNAc was used as an inhibitor against this enzymatic activity. The UDP-GlcNAc is a co-substrate in the
reaction. The mixture was pre-warmed to 37°C, then SF9 culture supernatant containing secreted GnT I was added to a total volume of 10%. The reaction was incubated at 37°C for 24 hours, and the reaction was stopped by the addition of an equal volume of 200 mM glycine, pH 10.

The reaction mixture was then applied to a BioGel P-10 column (15 × 200 mm) equilibrated with 200 mM glycine pH 10 to separate both mono- and oligosaccharides from the other components of the reaction mixture. The carbohydrate containing fractions (1.0 mL) were pooled, lyophilized and re-dissolved in 0.5 mL H₂O. The carbohydrates were then applied to a BioGel P-4 column (15 × 200 mm) to obtain the GnMan₅Gn-containing fractions (Figure 3.3). The fractions (0.5 mL) were pooled, lyophilized and dissolved in 0.2 mL H₂O. Separation of individual oligosaccharides was attempted by long column purification, and postponed upon successful crystal complex formation with the enriched GnMan₅Gn oligosaccharide pool.

Figure 3.3: Final step of the gel filtration purification of GnMan₅Gn from the GnT I reaction mixture. The leading peak is composed mainly of GnMan₅Gn.
3.2.4 Crystal complex formation

Cocrystallization of mutant dGMII and GnMan₅Gn ultimately proved to be unsuccessful. The cocrystals diffracted very poorly or not at all. As an alternative, crystal soaking experiments were attempted. Mutant dGMII crystals were soaked in substrate as follows: freshly grown D204A crystals were transferred to a solution of GnMan₅Gn in reservoir buffer for 18-24 hours. The molar ratio of GnMan₅Gn to dGMII was calculated to be 9:1.

3.2.5 Data collection

Soaked D204A crystals were cryoprotected by soaking in increasing concentrations of 2-methyl-2,4-pentanediol in reservoir buffer (5→10→15→20% v/v). The cryoprotected crystals were mounted on a 10 µm MicroMesh (MiTeGen Inc., Ithaca, NY, USA) or 0.4 mm CryoLoops (Hampton Research, Aliso Viejo, CA, USA) followed by flash freezing in an Oxford Cryosystems nitrogen coldstream set at 100 K (Oxford Cryosystems, Devens, MA, USA). The crystals were tested for preliminary diffraction on a Bruker AXS X8 PROTEUM (Bruker AXS GmbH, Karlsruhe, Germany) home source before transport to the synchrotron source.

The crystals were exposed to X-rays at Station A1 of the Cornell High Energy Synchrotron Source (CHESS) at Cornell University in Ithaca, NY. Data were collected at a wavelength of 0.9760 Å at 100 K and diffraction patterns were collected on an ADSC Quantum-210 CCD detector. Typically, 400 images were collected per crystal with an oscillation of 0.5° per frame.

3.2.6 Structure determination

The data obtained from the diffraction experiment were integrated and scaled using DENZO and SCALEPACK, respectively, within the HKL2000 graphical interface [73]. Model building, refinement and visualization were accomplished using REFMAC within the CCP4i suite, and Coot [94, 95]. In the case of the substrate complex, the $F_o - F_c \sigma_A$-weighted generated electron density was used to fit GnMan₅Gn into the model (Figures 3.4 and 3.5). Molecular topology for the substrate oligosaccharide were generated using the PRODRG server [96]. The images were prepared using the PyMOL Molecular Graphics System [97].
3.2.7 Substrate–dGMII interactions

To determine the protein residues directly involved in substrate binding, an analysis was performed using HBPLUS/LIGPLOT [76, 77]. Briefly, hydrogens were added to the substrate and protein using HBPLUS, then hydrogen bonds and hydrophobic interactions were visualized using LIGPLOT. dGMII−GnMan$_5$Gn interactions are presented in Figure 3.8 and summarized in Table 3.2.

3.2.8 Multiple sequence alignment and structural mapping

Amino acid sequences for Golgi α-mannosidase II from several organisms (D. melanogaster, A. gambia, H. sapiens, M. musculus, X. laevis, A. thaliana and C. elegans) were obtained from the National Center for Biotechnology Information (NCBI). MULTALIN was used to align these sequences (Blosum62 comparison table; gap open def=12, gap ext def=2), and the corresponding multiple sequence alignment was used in combination with ConSurf to produce an atomic coordinate file highlighting conserved residues on the dGMII structure [98, 99].

3.3 Results

3.3.1 Structure of substrate−dGMII (D204A) complex

The $F_o - F_c$ σ$_A$-weighted electron density map showed continuous electron density in the active site, and allowed for the unambiguous placement of all atoms of the substrate oligosaccharide GnMan$_5$Gn (Figures 3.4 and 3.5). Hydrogen bonds and hydrophobic interactions as determined by HBPLUS/LIGPLOT are presented in Figure 3.8 and Table 3.2.

The GnMan$_5$Gn oligosaccharide binds in a large groove on the surface of the enzyme. This groove contains the site of the nucleophile (D204), acid/base catalyst (D341) and zinc ion, and represents a region of extremely high degree of amino acid conservation (Figures 3.9, 3.10 and 3.11).
Data collection

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>$P_{2_1}2_1$</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>$a,b,c (\text{Å})$ 68.93, 109.58, 138.96</td>
</tr>
<tr>
<td></td>
<td>$\alpha,\beta,\gamma (^\circ)$ 90, 90, 90</td>
</tr>
<tr>
<td>Resolution ($\text{Å}$)</td>
<td>30–1.40 (1.44–1.40)</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.10 (0.49)</td>
</tr>
<tr>
<td>$&lt;I&gt;/\sigma &lt;I&gt;$</td>
<td>14.7 (4.9)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.94 (99.55)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Refinement

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution ($\text{Å}$)</td>
<td>1.40</td>
</tr>
<tr>
<td>Reflections</td>
<td>194389</td>
</tr>
<tr>
<td>$R_{work}$ / $R_{free}$</td>
<td>0.178 / 0.197</td>
</tr>
<tr>
<td>No. Atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>8434</td>
</tr>
<tr>
<td>Ligand</td>
<td>84</td>
</tr>
<tr>
<td>Water</td>
<td>1216</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>15.51</td>
</tr>
<tr>
<td>Ligand</td>
<td>19.11</td>
</tr>
<tr>
<td>Water</td>
<td>29.51</td>
</tr>
<tr>
<td>R.M.S. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond lengths ($\text{Å}$)</td>
<td>0.008</td>
</tr>
<tr>
<td>Bond angles ($^\circ$)</td>
<td>1.207</td>
</tr>
<tr>
<td>Ramachandran analysis</td>
<td></td>
</tr>
<tr>
<td>Favoured (%)</td>
<td>97.6</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>99.7</td>
</tr>
<tr>
<td>$&lt;\mu&gt;$ by Luzzati ($\text{Å}$)</td>
<td>0.145</td>
</tr>
</tbody>
</table>

Table 3.1: Data collection and refinement statistics for the $\text{dGMII(D204A)}$–$\text{GnMan}_5\text{Gn}$ complex.

3.3.2 $\text{dGMII(D204A)}$–$\text{GnMan}_5\text{Gn}$ interactions

Interestingly, almost all of the key protein-carbohydrate interactions take place at three saccharide positions: M5, M4 and G3 (Figure 3.8 and Table 3.2). Additionally, the ligand B-factors as determined by the crystallographic structure solution indicate that M5, M4 and G3 are the most stably bound in the crystal structure (Table 3.3). M1 and G2—the saccharide positions most distant from M5, M4 and G3—have the highest B-factors and lack any interactions with the enzyme. M5, M4 and G3 represent the saccharides most distal to the nascent protein. Substrate recognition by sensing the distal positions of the oligosaccharide is similar to what is seen with lectins [7]. In addition, the B-factors for the protein residues that interact with the substrate oligosaccharide are all lower than the average for the protein. Many are significantly lower, meaning that in binding the substrate, there is reduction in thermal motion and concomitant increase in rigidity for these residues.
3.3.3 Identification of dGMII subsites

The GnMan₅Gn–dGMII complex has shown that the active site of dGMII is made up of three sugar binding subsites. These have been termed the catalytic, holding and anchor subsites.

Catalytic site

The α1,6-linked mannose (M5) is tightly bound in the catalytic site forming many hydrogen bonds and stacking interactions with highly conserved residues in this region of the active site. Hydroxyls at position 2, 3 and 4 hydrogen bond to two conserved tyrosines (Y269 and Y727), a histidine (H471) and an aspartic acid (D472). In addition the hydrophobic face of the saccharide ring forms a stacking interaction with W95 while there is some hydrophobic contribution
from Y727 as well. The oxygens at positions 2 and 3 coordinate the stably bound zinc ion in the active site, which is consistent with many GMII inhibitors that have been studied (Figure 2.7 and Table 2.3) [67, 72, 69]. The site of nucleophilic attack (C1) is positioned between the nucleophile and acid/base catalyst, indicating that the hydrolysis reaction occurs at this site exclusively.

**Holding site**

The α1,3-linked mannose (M4)—which will also be cleaved by the GMII reaction—appears to be in a ‘holding’ site some 9 Å from the D204 nucleophile initially, and is multiply coordinated by a conserved arginine (R343) in this site. The architecture of this holding site is fairly open, with no aromatic stacking interactions contributing to binding whatsoever. Interestingly, previous work with this enzyme has not revealed any inhibitor binding in this site.

**Anchor site**

The N-acetylglucosamine (G3) required for mannosidase action is stably bound in a tight pocket 13 Å and 14 Å from M5 and M4, respectively. It forms strong stacking interactions with a conserved tyrosine residue (Y267), its acetyl position is buried in a hydrophobic pocket formed by a conserved tryptophan (W299) and proline (P298), and it forms a hydrogen bond with a conserved histidine (H273). The requirement of this non-hydrolyzed anchor is clear when considering the nature of the substrate. Oligosaccharide molecules are highly flexible.
both about their glycosidic bonds and in the conformation of their saccharide rings [100]. The presence of a stabilizing anchor as observed for G3 likely assists in binding and orienting the substrate for the hydrolysis reaction as well as increasing the local concentration of the linkages to be cleaved by GMII. An additional role for the anchor position in priming substrate catalysis is explored in Section 4.4.5.

### 3.3.4 Multiple sequence alignment

The regions of high sequence identity in the multiple sequence alignment are consistent with key substrate-binding regions of the active site cleft (Figures 3.9, 3.10 and 3.11). In particular, those residues involved in substrate binding are universally conserved across the GMII enzymes from a wide variety of species.
3.4 Discussion

Substrate recognition by dGMII

The question of how glycoside hydrolases in the N-glycosylation pathway are able to catalyze highly specific reactions has been a topic of discussion for many years. In the case of GMII, this enzyme is able to cleave two specific linkages on a particular oligosaccharide substrate. This is noteworthy when one considers the context in which the reaction is taking place. The Golgi apparatus during protein glycosylation is awash with many different types of protein-bound oligosaccharides, which makes the need for high specificity and proficiency all the greater.

Many of the dGMII side chains involved in substrate binding are both invariant as well as hydrophobic, which seems inconsistent with the hydrophilic nature of oligosaccharides in general. However, the contribution of hydrophobic side chains in the saccharide binding sites is consistent with the observation that oligosaccharide chains contain a significant amount of hydrophobic character, particularly over their ring elements [101]. Indeed, the presence of aromatic side chains has been seen in other structures of saccharide–protein interactions as well [102, 103, 104].
Figure 3.8: dGMII-GnMan$_5$Gn interactions as determined by HBPLUS/LIGPLOT. The substrate is depicted in red, and protein side chains in black. The half-circle depiction of residues refers to hydrophobic interactions. See text and Table 3.2 for details. The figure is based upon an orientation determined using LIGPLOT [76, 77].

Substrate conformation

The observation that the $\alpha$1,6-linked arm of the oligosaccharide folds back on itself was first postulated by Brisson and Carver in the early 1980s [105, 106, 107]. They used NOE measurements using pure oligosaccharides to investigate the solution conformation of these molecules. Remarkably, there was evidence that there was an interaction between the $\alpha$1,6-linked arm and the core N-acetylglucosamine of the oligosaccharide. It is noteworthy that this arrangement of the oligosaccharide is seen when bound to dGMII.

When one considers the substrate dependencies in the N-linked glycosylation pathway, the
### Table 3.2: dGMII-GnMan$_5$Gn interactions as determined by HBPLUS. The position of the dGMII residue on the sequence alignment in Figures 3.9 and 3.10 is indicated.

<table>
<thead>
<tr>
<th>GnMan$_5$Gn</th>
<th>dGMII</th>
<th>Residue</th>
<th>Atom</th>
<th>Residue</th>
<th>Atom</th>
<th>Interaction Type</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5</td>
<td>O$_2$H</td>
<td>Y269</td>
<td>OH</td>
<td>H471</td>
<td>NE2</td>
<td>hydrogen bond</td>
<td>369</td>
</tr>
<tr>
<td>M5</td>
<td>O$_3$H</td>
<td>D472</td>
<td>OD2</td>
<td>Y727</td>
<td>OH</td>
<td>hydrogen bond</td>
<td>864</td>
</tr>
<tr>
<td>M5</td>
<td>O$_4$H</td>
<td>D472</td>
<td>OD1</td>
<td>Y727</td>
<td>OH</td>
<td>hydrogen bond</td>
<td>864</td>
</tr>
<tr>
<td>M5</td>
<td>O$_5$H</td>
<td>Zn</td>
<td></td>
<td></td>
<td></td>
<td>metal coordination</td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>W95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>193</td>
</tr>
<tr>
<td>M5</td>
<td>Y727</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>864</td>
</tr>
<tr>
<td>M4</td>
<td>O$_3$H</td>
<td>R343</td>
<td>NE</td>
<td>R343</td>
<td>NE</td>
<td>hydrogen bond</td>
<td>444</td>
</tr>
<tr>
<td>M4</td>
<td>O$_4$H</td>
<td>R343</td>
<td>NE</td>
<td>R343</td>
<td>NH$_2$</td>
<td>hydrogen bond</td>
<td>444</td>
</tr>
<tr>
<td>M4</td>
<td>D340</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>441</td>
</tr>
<tr>
<td>M2</td>
<td>Q64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>162</td>
</tr>
<tr>
<td>M1</td>
<td>O$_1$H</td>
<td>Y267</td>
<td>OH</td>
<td></td>
<td></td>
<td>hydrogen bond</td>
<td>367</td>
</tr>
<tr>
<td>G3</td>
<td>O$_3$H</td>
<td>H273</td>
<td>NE2</td>
<td></td>
<td></td>
<td>hydrogen bond</td>
<td>373</td>
</tr>
<tr>
<td>G3</td>
<td>Y267</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>367</td>
</tr>
<tr>
<td>G3</td>
<td>P298</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>399</td>
</tr>
<tr>
<td>G3</td>
<td>W299</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>stacking</td>
<td>400</td>
</tr>
</tbody>
</table>

structural mechanism with regards to this enzyme becomes clearer. $\beta$1,3-galactosyltransferase, $\beta$1,4-galactosyltransferase and N-acetylglucosaminyltransferase III all act on the dGMII substrate and commit the oligosaccharide to the hybrid pathway, preventing dGMII action. Until now, the structural mechanism for this process was unknown. Though, it was postulated that the addition of a bisecting N-acetylglucosamine would prevent the proper substrate conformation and binding [106, 107]. The action of $\beta$1,3-GalT or $\beta$1,4-GalT on the dGMII substrate prevents further action of dGMII. The 3- and 4-position of the terminal N-acetylglucosamine face the enzyme, and substitution here would prevent substrate binding, and indeed catalysis from proceeding.

Furthermore, the action of GnT III adds a $\beta$1,4-linked Gn to the $\beta$1,4-linked mannose of the core of the dGMII substrate. The 4-position of the $\beta$1,4-linked core mannose faces into the active site cleft, and there is no discernible access for an enzyme, and the addition of an N-acetylglucosamine at this position would prevent proper substrate binding.
Table 3.3: Crystallographic temperature factors of the saccharide positions of GnMan₅Gn. See Figure 1.2 for a corresponding substrate schematic.

Figure 3.9: A multiple sequence alignment comparing GMII sequences from a variety of species. Residues that are implicated in substrate binding and catalysis are indicated. The specific roles of each of these residues can be found in Table 3.2.

Indeed, the examination of the dGMII–substrate complex allows one to construct a structural basis for the oligosaccharide dependencies seen at this step of the N-linked glycosylation pathway.
Figure 3.10: A multiple sequence alignment comparing GMII sequences from a variety of species. Residues that are implicated in substrate binding and catalysis are indicated. The specific roles of each of these residues can be found in Table 3.2.

Comparison with dGMII−Man₅ structure

A structure of the dGMII(D204A)−Man₅ structure [86] indicates that the G3 position of the GnMan₅Gn₂ oligosaccharide helps to properly position the flexible oligosaccharide in the dGMII active site. The Man₅ oligosaccharide bound to dGMII demonstrates the necessity of the anchor site in proper substrate recognition, as the α₁,6- and α₁,3-linked sugars bind to the catalytic and holding sites respectively in a manner nearly identical to that seen in GnMan₅Gn, but, due to the lack of the anchor saccharide, the other mannose positions are quite variant, with the M1 and M2-containing oligosaccharide 'tail' extending out of the active site cleft with O4 of the M1 saccharide binding D873 (Figure 3.12). Thus, the enzyme is able to bind the Man₅ substrate in the partially correct orientation even without the G3, and yet Man₅ is not cleaved. The necessity of the G3 position for M4 and M5 cleavage has been explored in past
Figure 3.11: A surface model of dGMII depicting the substrate binding site and coloured according to the multiple sequence alignment in Figures 3.9 and 3.10. Higher degrees of red shading indicate increased sequence conservation. Note the high degree of conservation in the cleft where the substrate (yellow) binds. The full view of the dGMII surface is shown in the inset, with the substrate depicted in yellow and indicated by an arrow.

biochemical characterization of GMII [40, 41]. This indicates that the G3 residue is critical not only in binding to the M4 and M5 sites, but also in the catalytic mechanism itself.

3.4.1 Conserved mode of substrate binding and catalysis

It has been known for some time that GMII enzymes from a wide variety of species are conserved, but these results indicate some of the reasons that certain highly conserved amino acids are crucial for substrate binding and catalysis. Non-binding residues in the catalytic, holding, and anchor positions all retain a high degree of sequence conservation among GMII enzymes. The residues involved in the putative internal communication are also highly conserved. In addition, it is proposed in Section 4.4.4 that the M3 position plays a key role in intermediate substrate rearrangement. Residues surrounding the M3 ‘swivel’ sugar are also highly conserved. It follows that the sequence of catalytic events and intermediates are also
Figure 3.12: A comparison of GnMan₅Gn (yellow) and Man₅ (green) bound to dGMII. The absence of the G3 anchor position in Man₅ results in a very different mode of binding distal from the catalytic and holding sites.

Conserved across species. This allows for greater confidence when using model systems such as Drosophila melanogaster GMII in human drug discovery efforts. This conservation speaks to the high degree of precision required to orchestrate the series of events as described in Chapter 4, as well as the physiological importance of this complicated function. It should also be noted that the substrate-bound structure of dGMII was solved in a catalytically inactive nucleophile mutant. Due to the significant binding contribution of non-nucleophile residues (Section 3.3.2), the almost invariant structure of dGMII(D204A) as compared to dGMII and the orientation of the substrate as being primed for catalysis (Section 4.4.4), one can conclude that the results presented here are relevant when considering the catalytically active enzyme.

3.4.2 Comparison of substrate binding to inhibitor binding

Swainsonine occupies the same site as the M5 position of the substrate oligosaccharide, and shares many residue contacts as well (Figure 3.13A). The hydrophobic face of swainsonine is positioned precisely for stacking with W95, and many of the hydroxyls are well positioned for
hydrogen bonding with other catalytic site residues.

Figure 3.13: Swainsonine compared to (A) undistorted and (B) distorted M5 in the dGMII catalytic site. The swainsonine (salmon) shows excellent superposition with the M5 hydroxyls in the distorted transition state (yellow).

Interestingly, when the M5 position is distorted into its $B_{2,5}$ conformation, swainsonine mimics the M5 position even more closely (Figure 3.13B). In particular, M5 positions O2, O3 and O4 superimpose very favourably with O7, O8 and O1 of swainsonine, respectively. In addition the N4 position of swainsonine is well-situated when compared with the ring oxygen of M5.

Based on the positions of the 7’ and 8’ hydroxyls, one can be reasonably certain that swainsonine’s high efficacy as a GMII inhibitor stems from its mimicry of the reaction transition state (Figure 3.13B).
Chapter 4

A novel catalytic mechanism for the sequential action of dGMII

4.1 Chapter overview

This chapter describes the formulation of the complete catalytic mechanism of GMII. This catalytic itinerary is explored in detail, and offers the ability to compare the potential differences between the catalytic mechanisms of GMII and LM, based on several structural and biochemical differences.

Briefly, the structure of the dGMII(D204A) catalytic nucleophile mutant is solved and presented here. It is noteworthy that there is little structural change upon substrate binding to dGMII. The full catalytic itinerary of dGMII is synthesized, with knowledge of the three sugar-binding subsites, the substrate specificity and requirements of this enzyme, the structure of the covalent intermediate, and the catalytic itinerary proposed for the ring interconversion and structure of the oxocarbenium-ion transition state. The key features of the mechanism involve hydrolysis at the catalytic site, followed by substrate rearrangement, a second hydrolysis event, and product release. The stark differences between dGMII and lysosomal α-mannosidase are explored via the subsites, and a strategy for selective inhibition is presented.

Some material presented in this chapter forms the basis of the work presented in Shah et al [86].
4.2 Materials and Methods

4.2.1 dGMII (D204A) crystallization

D204A crystals were grown as described in Chapter 3. Instead of performing crystal soaking with oligosaccharides, the crystals were prepared for data collection.

4.2.2 Data collection

D204A crystals were cryoprotected by soaking in increasing concentrations of 2-methyl-2,4-pentanediol in reservoir buffer (5→10→15→20% v/v). The cryoprotected crystals were mounted on a 10 µm MicroMesh or 0.4 mm CryoLoops followed by flash freezing in an Oxford Cryosystems nitrogen coldstream set at 100 K (Oxford Cryosystems, Devens, MA, USA). The crystals were tested for preliminary diffraction on an X-ray home source before transport to the synchrotron source.

The crystals were exposed to X-rays at Station A1 of the Cornell High Energy Synchrotron Source (CHESS) in Ithaca, NY. Data were collected at an X-ray beam wavelength of 0.9760 Å at 100 K and diffraction patterns were collected on an ADSC Quantum-210 CCD detector. Typically, 400 images were collected per crystal with an oscillation of 0.5° per frame.

4.2.3 Structure determination

The data obtained from the diffraction experiment were integrated and scaled using DENZO and SCALEPACK, respectively, within the HKL2000 graphical interface [73]. SADABS and XPREP were also used to convert and scale the data. Model building, refinement and visualization were accomplished using REFMAC within the CCP4i suite, and Coot [94, 95]. Images were prepared for publication using the PyMOL Molecular Graphics System [97].

4.2.4 Generation of intermediate substrate models

The partially cleaved GnMan₅Gn oligosaccharide was modeled using knowledge of the structural flexibility of oligosaccharides as well as biochemical data concerning the GMII reaction. Firstly, the ‘pre-rearrangement’ intermediate (GnMan₄Gn) was modelled by removing
the M5 saccharide from the structure of dGMII(D204A)−GnMan₅Gn. The model of the ‘post-rearrangement’ intermediate was built monosaccharide by monosaccharide with knowledge of the M4 (catalytic) and M3 (swivel) position from the dGMII complex with the synthetic tetrasaccharide (Zhong et al., submitted), the electron density from the GnMan₅Gn−dGMII structure, and knowledge of the precise shape of the dGMII active site cleft. The precise position and orientation of the M4 saccharide in the catalytic site and the M3 saccharide were based on the aforementioned synthetic tetrasaccharide. Geometric restraints for the intermediate substrate were obtained from the PRODRG server, and utilized to minimize geometric strain in the model.

4.2.5 Lysosomal α-mannosidase sequence and structural analysis

In order to compare LM and GMII, the consensus sequence between GMII and LM—as determined by MULTALIN (Blosum62 comparison table; gap open def=12, gap ext def=2)—was mapped to the bLM structure using ConSurf [98, 99]. The atomic structure of bLM was structurally aligned to the atomic structure of dGMII to compare differences in molecular architecture using PyMOL [97].

4.3 Results

4.3.1 Structure of dGMII (D204A)

A cryoprotected and frozen dGMII (D204A) crystal is shown mounted in a 0.4 mm CryoLoop in Figure 4.1. Data collection and refinement statistics are presented in Table 4.1, and a structural representation of the mutant enzyme is shown in Figure 4.2. The dGMII (D204A) is nearly identical to the GnMan₅Gn−bound structure, with a 0.12 Å RMSD over protein backbone atoms, and a 0.30 Å RMSD over all atoms in the structure. Residues in the catalytic, holding and anchor sites remain in nearly identical positions in the presence and absence of GnMan₅Gn binding.
Figure 4.1: A protein crystal of dGMII (D204A) frozen and mounted in a CryoLoop at the Cornell High Energy Synchrotron Source. The scale bar is 100 \( \mu \text{m} \) in length.

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>Resolution (Å) 1.30</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>Reflections 238288</td>
</tr>
<tr>
<td>( a,b,c ) (Å)</td>
<td>( R_{work} / R_{free} ) 0.192 / 0.216</td>
</tr>
<tr>
<td>( \alpha, \beta, \gamma ) (°)</td>
<td>No. Atoms</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>Protein 8415</td>
</tr>
<tr>
<td>( R_{merge} )</td>
<td>Water 1499</td>
</tr>
<tr>
<td>( &lt; I &gt; / \sigma &lt; I &gt; )</td>
<td>B-factors</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>Protein 11.25</td>
</tr>
<tr>
<td>Redundancy</td>
<td>Water 28.92</td>
</tr>
<tr>
<td></td>
<td>R.M.S. deviations</td>
</tr>
<tr>
<td></td>
<td>Bond lengths (Å) 0.007</td>
</tr>
<tr>
<td></td>
<td>Bond angles (°) 1.21</td>
</tr>
<tr>
<td>Ramachandran analysis</td>
<td>Favoured (%) 97.7</td>
</tr>
<tr>
<td></td>
<td>Allowed (%) 99.8</td>
</tr>
<tr>
<td></td>
<td>( &lt; \mu &gt; ) by Luzzati (Å) 0.154</td>
</tr>
</tbody>
</table>

Table 4.1: Data collection and refinement statistics for the dGMII (D204A) mutant.

### 4.3.2 Intermediate substrate models

The generation of an intermediate substrate model with M4 in the catalytic site shows two striking features. First, the G3 (anchor) position undergoes a significant (1.8 Å) shift from its original position, but this shift is confined to one half of the saccharide, as the N-acetyl group at position 2 remains bound in its hydrophobic pocket. Second, the M3 (swivel) position is largely perturbed from its original orientation, but remains unbound to any protein sidechains. These observations and their implications upon the catalytic mechanism are discussed in greater detail in Section 4.4.4.
Figure 4.2: The crystal structure of the catalytically inactive nucleophile mutant of dGMII, D204A.

4.3.3 Lysosomal \( \alpha \)-mannosidase sequence and structural analysis

A sequence alignment of lysosomal \textit{Bos taurus} \( \alpha \)-mannosidase (bLM) and \textit{Drosophila melanogaster} Golgi \( \alpha \)-mannosidase II (dGMII) was performed using the MULTALIN server (Blosum62 comparison table; gap open def=12, gap ext def=2). The resulting alignment was mapped to the bLM structure using the ConSurf server [98, 99].

The structures of bLM and dGMII were compared using the DaliLite server [67, 79, 108]. The alignment result had a Z-score of 35.4 over 780 equivalent residues and the root-mean-square deviation of the C\[^\alpha\] positions was 2.1 Å.

4.4 Discussion

4.4.1 Comparison of substrate–dGMII to unliganded dGMII

The positions of the active site residues is almost invariant between the apo- and substrate-bound forms of dGMII (D204A). This information, in combination with the examination of the active site temperature factors, indicates that there is little movement of protein side chains in the dGMII active site upon substrate binding. The B-factors in this region are also quite low, indicating that there is minimal dynamic motion both in the presence and absence of
substrate. One would surmise that major molecular motion by active site residues might have a deleterious impact on performing the reaction at a high velocity.

The correlation between catalytic power and enzyme thermostability has recently been explored by Roca et al [109]. Briefly, by using a dihydrofolate model enzyme in conjunction with simulation studies, Roca et al. were able to counter the long-held notion that structural flexibility is correlated with increased efficacy of catalysis [110]. Though some efficient enzymes display significant molecular motion, it was concluded that this motion is perpendicular to the reaction coordinate, and does not have an effect on catalysis. The reduction of the activation barrier remains the chief determinant of enzyme processivity. Roca and colleagues were able to conclude that in terms of enzyme-catalyzed reactions, a better catalyst is one which displays a smaller structural displacement while reducing the activation barrier. Their use of a model enzyme system is not universally applicable to all enzymes, but the conclusions bear mention when considering the catalytic trajectory of GMII.

### 4.4.2 Comparison of substrate–dGMII to a covalent reaction intermediate

High-resolution crystal structures of synthetic monosaccharides covalently linked to the dGMII nucleophile have been presented [71]. 5-fluoro-gulopyranosyl-fluoride and 2-deoxy-2-fluoro-α-D-mannopyranosyl fluoride (5FGulF and 2FManF, respectively) are able to covalently trap dGMII and indicate the preferred transition state of the reaction as proceeding via a $^1S_5$ intermediate. Both molecules are covalently linked in a $^1S_5$ conformation, indicating that both α- and β-retaining mannosidases employ a similar reaction coordinate.

As shown in Figure 4.3, 5FGulF bound to the wild-type dGMII active site closely resembles the M5 position in the substrate complex. It is noteworthy that in order for M5 to superimpose fully, the C1 position must traverse a distance of approximately 2 Å. Such a translation could be accomplished with a distortion from the M5 $^4C_1$ conformation to a distorted $^1S_5$ skew boat.

### 4.4.3 Comparison of substrate–dGMII to a synthetic mannose tetrasaccharide

The structure of a synthetic mannose tetrasaccharide (Figure 4.4 and PDB structure ID: 3BVV) bound to the active site of dGMII suggests that the catalytic site of dGMII can accommodate
both α1,6- and α1,3-linked mannoses. An α1,6- and α1,3-containing tetrasaccharide binds in the active site with the α1,3-linked mannose in the catalytic site because of the presence of a thiol linkage to the α1,6-linked sugar, which is not tolerated in the catalytic site (Zhong et al., submitted). Importantly, the saccharide in the swivel position is shifted significantly from its position in the substrate complex. The α1,3-linked mannose in this structure adopts precisely the same conformation and makes identical interactions with catalytic site residues as does the α1,6-linked mannose of the full substrate (Figure 4.5). Therefore, it is suggested that the mode of mannose binding in the catalytic site is identical for the α1,6 and α1,3-linked mannoses in the GnMan₅Gn₂ substrate during catalysis.

### 4.4.4 Putative catalytic mechanism

**Initial substrate orientation**

It has been assumed previously that either of the terminal mannoses (M4 or M5) could enter the catalytic site upon initial substrate binding with equal probability [41]. However, these results clearly show that the catalytic site exclusively binds the α1,6-linked mannose and the
α1,3-linked mannose resides in a holding site. In the context of the full GMII substrate, the holding site cannot accommodate the longer α1,6-linked arm while maintaining G3 in the anchor site, as highly conserved residues pack against and make significant contacts with the α1,3-linked mannose in that subsite (Figure 4.6). It appears that the steric restrictions in the dGMII holding site favour binding of the α1,6-linked sugar in the catalytic site initially. Thus, the chief determinant of bond cleavage order appears to be the holding site. This seems consistent with the notion that the catalytic site requires a certain degree of promiscuity in order to cleave both the α1,6 and α1,3-linked mannose residues while the holding site, in its sole binding of the α1,3 linked mannose, does not.

**First hydrolysis reaction**

It is hypothesized that substrate distortion occurs during the nucleophilic attack of the hydrolysis reaction. As mentioned in Section 2.4.4, previous work with this enzyme and a trapped covalent intermediate suggests that the catalytic mechanism of GMII proceeds via a $B_{2,5}$ intermediate as the saccharide ring of 5-fluoro-gulopyranosyl-fluoride is covalently bound to
CHAPTER 4. A NOVEL CATALYTIC MECHANISM FOR THE ACTION OF dGMII

Figure 4.5: A superposition of the mannosyl moieties in the catalytic site of dGMII from the full substrate (yellow), and the tetrasaccharide (green). The $\alpha_1,6$-linked mannose from the substrate superposes quite closely to the $\alpha_1,3$-linked mannose from the trisaccharide. This shows that the catalytic site is able to accommodate both linkage types, and provides a basis for the hypothetical substrate rearrangement and binding in the intermediate steps of the GMII reaction pathway.

The OE2 in a $^1S_5$ conformation in the catalytic site (Figure 4.3 and Section 4.4.2) [72, 71]. This indicates that when D204 is present in the wild-type enzyme the M5 saccharide undergoes a distortion, from a $^4C_1$ conformation seen in the crystal structure to a distorted $^1S_5$ concurrently with nucleophilic attack by OE2 of D204. This attack forms the glycosyl-enzyme intermediate, which is disrupted when D341 abstracts a proton from a water molecule, leading to nucleophilic attack by a hydroxyl ion at position C$_1$ of the M5 saccharide. This regenerates the enzyme and is followed by release of the M5 mannose.

Intermediate substrate rearrangement

After release of the M5 saccharide, M4 remains in the holding site 9 Å from the catalytic site. Substrate rearrangement follows which results in the positioning of the M4 saccharide in the catalytic site for hydrolytic cleavage. Such rearrangements are reminiscent of those reported in the context of processing endoglycanases [111, 112]. In previous dGMII structures, mannose binding has been seen at the M3 position to be quite variant. Thus, the M3 position has
been termed the ‘swivel’ position owing to this conformational variability as well as the open character of the active site cleft in this region. Hence, the swivel position allows for significant motion about its mannosyl linkages. A model of a rearranged intermediate, based on the solved crystal structure as well as the tetrasaccharide structure mentioned above, shows that it is straightforward to superpose the M4 saccharide in the intermediate in the same position as the M5 mannose from the original substrate (Figure 4.7A). However, in doing so, the M3 saccharide is displaced from its original position in the swivel site. A consequence of intermediate substrate rearrangement is that of the change in G3–anchor interactions. Due to its shorter length, the α1,3 linked mannose (M4) enters the catalytic site following M5 removal. The entire oligosaccharide must then undergo a 1.8 Å shift which weakens the G3–anchor site interaction (Figure 4.7B). Interestingly, the N-acetyl moiety of the N-acetylglucosamine at position G3 is able to retain its hydrophobic contacts, while the pyranose ring shifts 1.8 Å. It appears that this putative substrate rearrangement serves two purposes: i) to position M4 in the catalytic site, and ii) to facilitate product release. No side chain flexibility is required in the catalytic site to accommodate this rearrangement, in accordance with the stability observed in this site in many dGMII complex structures.
Second hydrolysis reaction and product release

It is hypothesized that the second hydrolysis reaction likely takes place much as the first, with nucleophilic attack being concurrent with distortion of the M4 saccharide ring from a $^4C_1$ to a $B_{2,5}$, followed by proton abstraction of water by D341, and deglycosylation restoring the enzyme to its initial state.

An important consequence of the substrate rearrangement is the removal of G3 from its tight binding pocket allowing for diffusion of the product from the enzyme.

4.4.5 Internal communication between dGMII subsites

An important issue in the mechanism is to what extent the occupancy of each of the sites affects the characteristics of the other sites. There are numerous hydrogen bonds between the conserved residues D270 and H273 (Figures 3.8 and 3.6). Although it does not make direct contacts with the substrate, D270 acts to coordinate H273 in the proper orientation to form a hydrogen bond with $O_3H$ of the G3 N-acetylglucosamine. In addition, D270 lies in a re-
CHAPTER 4. A NOVEL CATALYTIC MECHANISM FOR THE ACTION OF dGMII

Figure 4.8: Schematic of the proposed GMII reaction. The reaction begins with the binding of the GnMan₅Gn₂ oligosaccharide to GMII, with the M5 position binding in the catalytic site, M4 occupying the holding site and G3 in the anchor site (1). After the hydrolysis of M5 by the enzyme, the catalytic site is free (2). Oligosaccharide rearrangement ensues to allow M4 to enter the catalytic site (3) which is hypothesized to be concurrent with a weakening of G3-anchor site interactions. M4 is subsequently hydrolyzed in the catalytic site by a similar mechanism to M5, and the M4 release is concurrent with product release (4).

The region of the active site that makes it equidistant from both the G3 binding site, and the initial M4 holding site. This may indicate that, through proton transfer, the D270 position could facilitate communication between the anchor and holding subsites, perhaps in the intermediate substrate rearrangement step, or in assisting in the ejection of the G3 anchor from its site during rearrangement. It is hypothesized that the departure of M4 from the holding site during substrate rearrangement facilitates G3 release, which is required for release of the final doubly-cleaved product.

Additionally, the presence of a histidine (H273) at a key position of the catalytic cleft could suggest that the communication proceeds via a proton shuttle [113, 114, 115, 116]. Histidine residues are well-suited to be both donors and acceptors of protons at physiological pH and the presence of H273 in coordinating the anchor site and making contact with D270 may indi-
cate such a role. It is hypothesized that in the absence of a substrate moiety in the anchor site, H273 is able to use the basic nitrogen of its imidazole ring to abstract a proton from the water molecules in the anchor region, which is followed by a donation of a proton from the acidic nitrogen of the imidazole to the well-positioned side chain of D270. When the anchor site is occupied, Y269 is properly positioned which, in effect, primes the catalytic site for binding and catalysis. Therefore, the enzyme is able to translate information from two of the key sites in the active site cleft, the anchor site, and catalytic site. This signal transfer must be precisely orchestrated such that the catalytic site remains competent until the second linkage is cleaved, at which time the product is released.

4.4.6 Comparison of dGMII to lysosomal α-mannosidase

Knowledge of the structure of a lysosomal α-mannosidase from Bos taurus (bLM) solved by Heikinheimo et al. [79] in conjunction with the data presented in this thesis permits detailed structural comparison and will assist in the development of specific GMII inhibitors.

Overall characteristics

The active site cleft of bLM is much wider than that of dGMII (12 Å vs. 20Å) but still contains a tightly bound zinc in the catalytic position as shown in Figures 4.9 and 4.10. There is a striking amount of sequence similarity between the two enzymes in some regions, and almost none in others. There is a lack of sequence conservation at both the holding and anchor sites, and an increased width of the active site cleft in bLM (Figure 4.9B). In examining the surface representation of the two enzymes, it seems that the size and shape of the dGMII active site cleft is well contoured to the size and shape of the GnMan\(_5\)Gn\(_2\) oligosaccharide, while the bLM site is much wider.

Catalytic site

Interestingly, the catalytic sites of both enzymes are quite similar one another. Almost all the key residues from dGMII which are involved in M5 binding and Zn coordination are present in the bLM active site, in strikingly similar positions (Figure 4.10 and Table 4.2).
A comparison of the dGMII and bLM active site clefts coloured according to sequence conservation with respectively related mannosidases. (A) The dGMII active site cleft with bound substrate. The catalytic (M5), holding (M4) and anchor (G3) sites are indicated as shown. (B) The bLM active site aligned in the same orientation as (A) above. The dGMII substrate is presented in transparent yellow.

The residues involved in M5 hydrogen bonding (Y269/261, H471/446, D472/447 and Y727/660) are present in both enzymes, and adopt similar conformations and would ostensibly coordinate saccharide moieties in a nearly identical manner. The position of the Zn$^{2+}$ ion is almost invariant, with key residues coordinating the ion (H90/72, D92/74, H470/445 and H471/446) being conserved in both sequence and atomic position. The W95 position in dGMII that is believed to be extremely important for saccharide binding has an equivalent residue in bLM (W77) which adopts a nearly identical conformation, and likely contributes a great deal to saccharide binding in this enzyme. Lastly, the nucleophile (D204/196) adopts a similar position in both enzymes, and likely performs a very similar role.
In two catalytic sites this similar, one can surmise that the mechanism of nucleophilic attack, regeneration by the acid/base catalyst, and product release are quite similar to one another. This lends credibility to the notion that one must look beyond the catalytic site as an inhibitor target, if one wishes to distinguish between the enzymes by small molecule inhibition.

**Holding site**

Instead of a conserved arginine in the holding site, the bovine enzyme contains a two weakly conserved residues, a glutamine (Q321) and a leucine (L272). Due to their arrangement, it is difficult to envision a mode by which they would be able to coordinate the hydroxyl groups of the M4 saccharide (Figure 4.11). In addition, the GMII-conserved aspartic acid that provides some saccharide binding ability by its hydrophobic face in dGMII (D340) is not present in bLM. Instead, there is a serine some distance away, that likely would not contribute to binding. Overall, there do not seem to be residues that would be suited to coordinating saccharide moieties, and there is little sequence conservation in this area.
Chapter 4. A novel catalytic mechanism for the action of dGMII

<table>
<thead>
<tr>
<th>dGMII</th>
<th>bLM</th>
<th>role</th>
</tr>
</thead>
<tbody>
<tr>
<td>W95</td>
<td>W77</td>
<td>M5 stacking</td>
</tr>
<tr>
<td>Y269</td>
<td>Y261</td>
<td>M5 H-bonding</td>
</tr>
<tr>
<td>H471</td>
<td>H446</td>
<td>M5 H-bonding</td>
</tr>
<tr>
<td>D472</td>
<td>D447</td>
<td>M5 H-bonding</td>
</tr>
<tr>
<td>Y727</td>
<td>Y660</td>
<td>M5 H-bonding</td>
</tr>
<tr>
<td>Zn</td>
<td>Zn</td>
<td>M5 coordination</td>
</tr>
<tr>
<td>H90</td>
<td>H72</td>
<td>Zn coordination</td>
</tr>
<tr>
<td>D92</td>
<td>D74</td>
<td>Zn coordination</td>
</tr>
<tr>
<td>H470</td>
<td>H445</td>
<td>Zn coordination</td>
</tr>
<tr>
<td>H471</td>
<td>H446</td>
<td>Zn coordination</td>
</tr>
<tr>
<td>D204</td>
<td>D196</td>
<td>nucleophile</td>
</tr>
</tbody>
</table>

Table 4.2: Key residues in the dGMII catalytic site with their structurally-aligned bLM equivalents.

Anchor site

The anchor site does not seem to exist in bLM. There is no saccharide binding pocket to speak of that would correspond to the dGMII anchor site. There is instead a pocket of moderately conserved residues, none of which possess the acidic or aromatic character present in the dGMII anchor site. The formation of the anchor site in dGMII depends on the extension of two loops from the main body of the protein. The first loop carries Y267, which makes a crucial stacking interaction with the G3 position and a hydrogen bond with O$_4$H of the M1 saccharide (Figure 4.12A). The tyrosine exists on a loop that is stabilized in position by a short $\alpha$-helical segment immediately preceding it. This tyrosine, along with the loop and helical segment, are completely absent in bLM. The second loop carries a proline (P298) and a tryptophan (W299) which are uncharacteristically at the protein surface. It seems that they form a small hydrophobic patch which interacts with the N-acetyl group of G3 (Figure 4.12B). The loop extension is completely absent in the bLM structure. In fact, there are no bLM residues nearby that could conceivably form the same hydrophobic patch seen in dGMII. The fact that bLM lacks these two key loops, and subsequently the residues necessary to form the anchor site, indicates that there likely is no anchor binding taking place.

The expanded active site cleft and limited sequence conservation is consistent with the sub-
Figure 4.11: A comparison of the atomic structures of the dGMII (blue) holding site and the equivalent region in bLM (magenta). The M4 position as solved from the dGMII–substrate structure is presented in transparent yellow.

The two unconserved regions, the holding site and, in particular, the G3 anchor site, are appealing targets for deriving specificity against GMII over the lysosomal counterpart.
Figure 4.12: A comparison of the atomic structures of the dGMII (blue) anchor site and the equivalent region in bLM (magenta). The G3 position as solved from the dGMII–substrate structure is presented in transparent yellow. The anchor site is shown in two orientations to highlight the (A) aromatic stacking and (B) the hydrophobic interactions.
5.1 Nuclear magnetic resonance analysis of the GMII reaction

Nuclear magnetic resonance (NMR) spectroscopy-based characterization of the glycosyl hydrolase reaction is a method that can reveal in-depth mechanistic insights concerning the catalytic mechanism. As recently shown, Cex—a β-(1,4)-glycosidase from Cellulomonas fimi—which is a popular model enzyme for various xylanases and cellulases—shows an increase in local and global stability over the course of its reaction mechanism. This suggests that the binding of the substrate, the formation of the glycosyl-enzyme intermediate, and the release of the product are thermodynamically favourable events [118]. NMR relaxation studies of Bacillus circulans xylanase also showed an increase of stability upon substrate binding and the formation of the glycosyl-enzyme intermediate [119].

A more recent NMR technique, called saturation transfer difference nuclear magnetic resonance spectroscopy (STD-NMR), is well-suited to the study of ligand-target interactions in biological systems. It monitors the transfer of magnetization from a protein to a bound ligand. It is effective in the determination of dissociation constants from the nM to mM range, and can be performed on protein amounts as low as the nmol level, and even hundreds of pmol with the use of cryo-probes. In fact, due to the particulars of the technique, it allows for the acquisition of ligand-target information even when the size of the complex is too high for traditional NMR techniques (>100 kDa) [120]. STD-NMR has been successfully applied in the interro-
gation of the α-L-fucosidase reaction mechanism. The use of STD-NMR on α-L-fucosidase in the presence of inhibitors as well as fucosylated oligosaccharide substrates provides detailed information regarding the catalytic mechanism of the enzyme. Of particular note is the use of low temperature studies in the oligosaccharide case in order to slow the reaction to an observable velocity [121].

If the GMII mechanism could be interrogated by NMR, it would reveal mechanistic insights of the reaction as it proceeds. Oligosaccharides could conceivably be monitored by $^1$H NMR, without the use of labels, though that would require testing.

### 5.2 Molecular dynamics simulation of GMII reaction

Based on the work presented in the preceding chapters, it is possible to envision a still-frame trajectory of the oligosaccharide substrate in the GMII reaction. By using a molecular dynamics approach, one can use these intermediate structures and link them by simulating their motion in the active site. The use of targeted molecular dynamics (TMD) allows for the specification of an ‘end-state’ of the simulation, and energies during the simulation trajectory can be examined. Though not a pure, open-ended molecular dynamics method, studying what may be a real biological motion with TMD may be instructive.

The GLYCAM06 force field has recently been developed by the Woods group in order to more accurately study the molecular dynamics of oligosaccharides [122]. The utilization of this force field in a molecular dynamics simulation—in conjunction with the crystallographic studies presented here—would provide useful information regarding the energetics of substrate catalysis, rearrangement, and release in GMII.

A full energetic model of substrate rearrangement in the GMII active site would be an interesting undertaking, thus helping to answer some questions about the energies involved in that transition. Further work could assist in assessing details of substrate binding or product release.

Furthermore, *in silico* determination of the transition state structure would be useful for future inhibitor design. For example, several *ab initio* studies of the glycosyltransferase mech-
anism have been undertaken using high-level quantum mechanical calculations in order to
determine the precise structure of the transition state and formulate powerful inhibitors (re-
cently reviewed in [123]).

5.3 Novel inhibitor development

Novel inhibitor development against GMII has relied thus far on expanding the functionality
of known inhibitors. Work presented in this thesis provides the opportunity to approach this
problem in a novel way, specifically relying on the novel architecture of the GMII active site to
conferr high specificity.

5.3.1 Increasing inhibitor specificity: GMII over LM

As discussed in Section 4.4.6, GMII and LM have many similarities, which is not particularly
surprising, but several of the differences could be of use in increasing the selectivity of inhi-
bition. The strategy one would employ would hopefully avoid targeting enzyme similarities
while simultaneously taking advantage of the differences between the enzymes.

Targeting the holding site

Current GMII inhibitors have exhibited no binding in the holding site, and there is significant
dissimilarity between GMII and LM in this region (Figure 4.11). This indicates that the holding
site could be a useful target for GMII-specific inhibition.

Recent structural work with Escherichia coli aspartate transcarbamoylase (ATCase) and
two phosphonates—phosphonoactamid and N-(phosphonoacetyl)-l-aspartate—has demon-
strated the ability of phosphonates to coordinate arginine residues in enzyme active sites
[124]. Furthermore, the electrostatic strength of the phosphate—arginine interaction has re-
cently been touted as being able to withstand even the collision effects of mass spectrometry,
which speaks to its tight binding [125].

Based on the above, an attempt to target the holding site using phosphates or phosphonates
is an attractive approach. Due to the ease of synthesis and higher stability of phosphonates,
they will be initially tested for holding site binding ability. This work will be undertaken in collaboration with Professor Mario Pinto of Simon Fraser University. The initial steps will involve enzymatic and structural studies of dGMII in complex with various phosphonates. Models of phosphonic acid binding in the dGMII holding site are shown in Figure 5.1.

**Figure 5.1:** A model of three phosphonic acid molecules binding to R343 in the dGMII holding site. Phosphonates can be conjugated to a variety of groups, thus making them a very versatile chemical group, as shown in the lower right panel. Ethylphosphonic acid (A), a carbohydrate-like phosphonic acid (B) and cinnamylphosphonic acid (C) are depicted here.

**Targeting the anchor site**

The absence of the anchor site in LM, along with its high level of sequence conservation in GMII makes it a strong candidate for inhibition targeting. Though not directly involved in binding scissile residues of the substrate, its importance has been discussed. N-acetylglucosamine or close derivatives are immediately available and would serve well to bind the anchor site, as
they can access the strong stacking interactions with Y267, and the small hydrophobic pocket formed by P298 and W299.

5.3.2 Avidity-based approach

Designing an inhibitor that would be able to access both the holding and anchor sites by linking together efficacious binders of each would be a novel approach toward GMII inhibition. Combining this two-headed molecule to a moiety that could target the catalytic site—such as a proven inhibitor of GMII—would add a third ‘head’ that could provide an orientation element for the inhibitor and increase its effectiveness. If this approach were to be undertaken, it would likely be beneficial to use a weak inhibitor of GMII for the catalytic-targeting element. If too strong a GMII inhibitor were used, such as swainsonine, LM cross-inhibition would still be a problem.

Increasing inhibitor specificity through the use of distal binding elements has been seen in recent work with IMP dehydrogenase (IMPDH). IMPDH catalyzes the rate-limiting step in guanine synthesis and is selectively inhibited by mycophenolic acid (MPA), which traps the covalent intermediate of the enzyme and binds in the nicotinamide half of the active site. Digits et al performed experiments with two inhibitors, one that is nicotinamide-specific (tiazofurin), and the other adenosine-specific (ADP), in order to test inhibitor coupling, and determine species-specific effects. They were able to demonstrate the phenomenon of distal binding site effects on catalytic function [126].

5.3.3 Fragment-based approach

Though it is tempting to infer that an ideal GMII inhibitor will be composed of molecules that are reminiscent of saccharides, it is conceivable that the chemical composition of an ideal inhibitor will be quite divergent from carbohydrate structures. To that end, escaping the possible ‘local minimum’ of saccharide-based inhibitors requires the use of a technique that approaches the problem of inhibitor development from an unbiased perspective. Fragment-based methods have emerged recently and provide a fascinating platform upon which to base the development of novel de novo GMII inhibitors.
The multi-subsite architecture of the GMII active site lends the enzyme very well to fragment-based approaches of inhibitor discovery. The fragment-based approach utilizes libraries of low molecular weight (< 250 Da) compounds that are suitable for screening at high concentrations. These fragments are tested for binding to enzymes by a variety of methods, the most applicable to GMII being enzymatic assays and X-ray crystallography. It is envisioned that, in contrast to standard high-throughput drug screening techniques, fragment-based screening will result in binding that is highly productive for 'hot spots' in an enzyme's active site. The discovery that the active site of GMII is composed of three subsites (catalytic, holding and anchor), the availability of a well-characterized enzyme assay, and its facile crystallization and structure solution, makes dGMII an excellent candidate for fragment-based screening. Upon discovery of fragments that bind strongly to GMII subsites, it is typical to tether these molecules together to form highly specific inhibitors that make multiple productive interactions with active site elements. Fragment-based screening has recently been reviewed by Carr and colleagues [127].

5.4 Fluorescence-based inhibitor screening

Fluorescent moieties linked to small molecules are both highly quantitative and sensitive, and can be scaled to high throughput applications. The multiple binding sites of GMII have unique characteristics, and it is an enticing prospect to develop multiple small molecules—each uniquely fluorescently labeled—to access and interrogate the individual GMII subsites. A high-throughput approach would allow for the rapid testing of a large amount of small molecules in order to determine efficacious inhibitors more quickly.

5.4.1 Fluorescence polarization

Fluorescence polarization uses fluorescence anisotropy to determine dissociation constants of small fluorescent molecules bound to larger ones. In the case of GMII, an appropriate arrangement would be the use of small fluorescently-labeled ligands with an unlabeled enzyme.

Fluorescence polarization and anisotropy are routinely used in high-throughput screening efforts. FP/FA is particularly well-suited to high-throughput screening due to the lack of
Chapter 5. Future Studies

separation steps needed. [128]. Furthermore, these methods employ a single-label, are fast and sensitive. An FP assay for GMII would employ a fluorescently labeled oligosaccharide ligand and a library of unlabeled ligands utilized in competition according to the scheme below:

\[
F - \text{Ligand} + \text{Receptor} \rightleftharpoons F - \text{Ligand} : \text{Receptor}
\]

\[
I + \text{Receptor} \rightleftharpoons I : \text{Receptor}
\]

This method would allow for the screening of many inhibitors while keeping all other conditions identical, a key feature for high-throughput study. A similar fluorescence polarization assay has been developed for peptidyl prolyl cis/trans isomerase for use in high-throughput systems [129]. Furthermore, fluorescence polarization has recently been used in glycan-protein binding experiments using bovine prion protein and heparin [130].

One can envision a system subsite-specific oligosaccharide ligands are fluorescently-labeled and used in competition with small molecule libraries to find inhibitors of specific subsites of GMII. Typically, fluorescence polarization and anisotropy experiments are performed in a singly-labeled manner. Though, if the fluorescent spectra of each site were unique, it may even be possible to perform the experiments simultaneously in an orthogonal fashion, and test binding at three subsites in a single reaction. Minituarization to a 96-well format would keep material consumption low, and increase the speed of the assay by utilizing a fluorescence-equipped plate-reader.

5.4.2 Development of labeling protocol

In order to perform the fluorescence-based experiments described here, it will be necessary to have glycans fluorescently labeled in a nonselective and nondestructive manner. Conventional methods of fluorescent conjugation employ aminopyridine-based fluors, which tend to desialylate oligosaccharides, and are somewhat nonselective. Recent techniques employing aniline-based fluors have shown much more promise.

Protein-derived glycans can be labeled at the reducing terminus 2-amino benzamide (2-AB) and 2-anthranilic acid (2-AA) [131]. The method is robust as the labeling has high mo-
lar efficiency and is independent of glycan structure. Moreover, The 2-AB label is compatible with chromatographic, enzymatic and mass spectrometric techniques that are used to analyze and separate glycans, while 2-AA’s intrinsic negative charge makes it a poor choice for ion-exchange chromatography and reverse-phase HPLC, but an excellent choice for electrophoretic analyses, particularly those requiring sharp resolution. Fluorescent labeling and detection is a desirable method as it does not require radioactive labeling of oligosaccharides, while providing high sensitivity and compatibility with many analytical techniques.

5.5 Expression profiling of glycosidases

The recent development of a highly-annotated microarray (glycogene-chip v1) based on the Affymetrix platform have allowed for insights into the transcriptional regulation of genes involved in glycosylation. The focused gene microarray contains families of 436 human and 285 murine genes that code for the proteins responsible for glycan synthesis and recognition. Newer microarrays (glycogene-chip v2) expand the gene set to contain even more glyco-related genes [132].

Based on this technology, the expression status of the gene encoding GMII as well as other glycoside hydrolases in the N-glycosylation pathway can be assessed. Wild-type and disease cell lines, as well as patient samples could be examined for up- or down-regulation of specific genes. This information, in conjunction with information of the cell surface glycans, would be an important piece of information for determining the precise effect of the GMII protein on multicellular systems.

A very recent study using qRT-PCR to quantitatively monitor the levels of mRNA transcripts in a variety of mouse tissues has demonstrated this expression profiling strategy [28]. The authors utilized a list of more than 700 glyco-related genes and monitored their mRNA levels in the mouse. They were able to correlate this expression data with detailed study of the glycan pool of the cells under study. They have concluded that this is a robust application of array technology, and look forward to newer array chips, specifically the glycogene-chip v2, as a tool for more detailed experimentation.
5.6 Further structural study of GMII

Recent work detailing the cloning and expression of Golgi α-mannosidase II from *Arabidopsis thaliana* (AtGMII) offers the possibility of performing structural analyses of a GMII enzyme from a completely different kingdom of life [133]. The work describes the enzymatic contribution to plant cells, focusing on the oligosaccharide profile in AtGMII knockouts, as well as deducing the region of the amino acid sequence responsible for Golgi targeting of the enzyme by the use of a green fluorescent protein fusion construct of the N-terminal portion of AtGMII.

A long-term goal would be the structure solution of the *Homo sapiens* Golgi α-mannosidase II. Though it has been established that the Drosophila enzyme is a suitable structural equivalent, it would be interesting to study the human enzyme in-depth to examine precise similarities and differences, and see if the unique features of the human enzyme could be useful in inhibition studies.

5.6.1 Mutagenesis studies

The work described in this thesis expands the definition of the GMII active site, and suggests a number of residues that are likely involved in the catalytic mechanism.

**Binding residues**

Residues such as R343 (holding site), Y267, P298 and W299 (anchor site) form key interactions with saccharide molecules in substrate binding. Targeting these residues for mutagenesis and testing substrate binding could provide some information regarding their contribution. Also, extensive mutagenic study could be pursued to determine if it would be possible to engineer and alter the substrate specificity of GMII by mutagenesis. A clearer understanding of how this enzyme is able to bind oligosaccharide substrates will be useful in the detailed study of other saccharide-binding proteins, including hydrolases, transferases and lectins.
Inter-subsite communication

As mentioned in Section 4.4.5, two key residues are suggested for inter-subsite communication in the GMII mechanism. D270 and H273 seem to be key residues in a putative proton transfer mechanism that allows communication between the anchor site and the catalytic site. Mutating either or both of these residues and studying the catalytic mechanism and substrate binding properties would assist in determining whether this hypothesis is correct. Understanding the details of the GMII mechanism in terms of inter-subsite communication may assist in the understanding of other enzymes that may utilize similar communication methods.

Molecular evolution of GMII

Extensive sequence analysis against other eukaryotic GMII’s might shed some light on the molecular evolution of GMII, and could lead mutagenesis efforts to explore the enzymatic character of ancient GMII’s. This would be an undertaking that would have implications related to the rise of glycosylation as an evolutionary event, and possibly might help explore the unicellular/multicellular barrier in which the formation of complex glycans has already been implicated [3].
Chapter 6

Conclusions

This dissertation provides insights into the catalytic mechanism of Drosophila melanogaster Golgi $\alpha$-mannosidase II and provides a model for the broader enzymatic function of GMII in biological systems. The work presented by using structural and biochemical methods helps in the formulation of a complex, yet plausible, model of GMII action, and opens the door for many future avenues of research.

The study of kifunensine binding to dGMII resulted in a greater understanding of the conformational distortion in mannosidase active sites. The solution of a high-resolution crystal structure (PDB Structure ID: 1PS3), along with the accompanying enzymatic and energetic study, proved to be instructive across the entire mannosidase class of enzymes. The substrate distortion in the dGMII active site allows for inter-family comparison of mannosidases, and provides key information in the energetic and conformational itinerary of this enzyme.

The 1.4 Å resolution structure of GnMan$_5$Gn bound in the dGMII active site cleft (PDB Structure ID: 3CZN) is a key contribution of this thesis, and provides the framework for much of the discussion surrounding the catalytic mechanism. The atomic details in this high-resolution structure are particularly intriguing, and telling of the mechanism itself. The identification of key subsites that are GMII-conserved indicates that the substrate binding and enzymatic mechanism are highly conserved and the key substrate binding and catalytic residues identified in this structure provide a similar function in other GMII enzymes.

Further analysis of the GnMan$_5$Gn–dGMII structure, along with comparisons against a
high-resolution unliganded enzyme structure (PDB Structure ID: 3CZS) and that of the related lysosomal $\alpha$-mannosidase provide key insights into the full mechanistic itinerary of the double cleavage performed by this enzyme. The contribution of substrate binding residues in intersite communication is explored by the suggestion of a novel proton shuttle mechanism, similar to that seen in carbonic anhydrase. Furthermore, key structural differences between GMII and LM are explored in detail, providing key leads in increasing inhibitor specificity against GMII while avoiding LM inhibition, the outcome of which could be the design of improved chemotherapeutic strategies against certain cancers.

Like so much of the work concerning GMII, arriving at an answer often raises new questions. Intricate dissection of the catalytic mechanism by dynamic techniques such as NMR or MD are suggested as possible future courses of study, along with further structural study to provide greater insights into the roles of several key residues mentioned in preceding chapters. This work has renewed the enthusiasm in the search for a highly specific inhibitor against GMII. Instead of following strategies that exhaustively probe chemical analogues of saccharides, a guided approach against three individual subsites is suggested, along with the incorporation of newer techniques such as fragment-based screening, for which this enzyme is an ideal candidate.
Bibliography


[105] J R Brisson and J P Carver. Solution conformation of asparagine-linked oligosaccharides: α1,2-, α1,3-, β1,2-, and β1,4-linked units. *Biochemistry*, 22(15):3671–3680, 1983.


[127] R A Carr, M Congreve, C W Murray, and D C Rees. Fragment-based lead discovery: leads by


polarization-based assay for peptidyl prolyl cis/trans isomerase cyclophilin A. *Anal Biochem*,


and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal

[132] E M Comelli, S R Head, T Gilmartin, T Whisenant, S M Haslam, S J North, N K Wong, T Kudo,
H Narimatsu, J D Esko, K Drickamer, A Dell, and J C Paulson. A focused microarray approach to
2006.

[133] R Strasser, J Schoberer, C Jin, J Glössl, L Mach, and H Steinkellner. Molecular cloning and char-
acterization of *Arabidopsis thaliana* Golgi α-mannosidase II, a key enzyme in the formation of

This thesis was prepared using the LATEX typesetting language.