STRUCTURE AND SUBSTRATE SPECIFICITY OF GH47 ALPHA 1,2-MANNOSIDASES

by

YONG XIANG

(Under the Direction of Kelley W. Moremen)

ABSTRACT

Family 47 glycohydrolases cleave α1,2-mannose linkages on Asn-linked oligomannose structures by an inverting mechanism that involves an substrate interaction with an enzyme-bound Ca^{2+} ion for substrate distortion into the transition state. Among enzymes in this family, ER α-mannosidase I (ERManI) and Golgi α-mannosidase IA (GMIA) are key enzymes involved in *N*-linked glycan biosynthesis in the endoplasmic reticulum (ER) and Golgi complex. The structures of ERManI and GolgiManIA are similar in overall protein fold, but have complementary activities in their cleavage of the natural $Man_9GlcNAc_2$ substrate. ERManI preferentially cleaves a single α1,2-Man residue from the central B-branch of the tri-branched oligomannose substrate, whereas GMIA preferentially cleaves three α1,2Man residues from branches A and C, but poorly cleaves the B branch α1,2-Man residue. In order to determine the structural basis for the substrate specificity differences, human ERManI and murine GMIA were inactivated by replacing their enzyme-bound Ca^{2+} with La^{3+}, and co-crystallized with the substrate analog $Man_9GlcNAc_2$-PA. The crystal structures were resolved at resolutions of 1.65 Å for ERManI and 1.77 Å for GMIA. The enzyme-substrate complex structures reveal that GMIA binds the non-reducing terminal α1,2-Man of branch A in the -1 enzyme subsite, while the

remainder of the glycan has extended interactions in the enzyme cleft. In contrast, ERManI binds the -1 enzyme subsite with the terminal α1,2-Man of branch B with a distinctive set of interactions between the glycan substrate and the enzyme cleft. Non-conserved residues in the topologically equivalent positions of the glycan binding clefts were swapped between ERManI and GMIA and the mutations led to substrate specificity changes as well as the compromised enzymatic activity and substrate binding affinity in some cases. In order to search for the selective inhibitors towards ERManI as the potential therapeutic agents for protein misfolding disorders, a collection of kifunensine and 1-deoxymannojirimycin derivatives have been screened. Varying degrees of selectivity of inhibition were found among several compounds. The structural and enzymatic studies on ERManI and GMIA in this work provided more insights into the structural basis of substrate specificity for members of this enzyme family.

INDEX WORDS: GH47 α1,2-mannosidases, ERManI, GMIA, $Man_9GlcNAc_2$, substrate specificity, substrate conformation, crystallography, binding affinity, inhibitors.

STRUCTURE AND SUBSTRATE SPECIFICITY OF GH47 ALPHA 1,2-MANNOSIDASES

by

YONG XIANG

B.S. Wuhan University, China, 2000

M.S. Shanghai Pharmaceutical Institute, China, 2002

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2015

STRUCTURE AND SUBSTRATE SPECIFICITY OF GH47 ALPHA 1,2-MANNOSIDASES

by

YONG XIANG

Major Professor:	Kelley W. Moremen
Committee:	Michael Pierce
	Michael Tiemeyer
	Lance Wells

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2015

DEDICATION

To my family, who have never stopped showing me their understanding and love and have never lost confidence in me. You always encourage me to continue exploring, not only in science but also in searching for the meaning of life. Your love has rooted happiness in me. Your extraordinary endurance has calmed me and fueled me with strength in hard times. You are always the greatest inspiration and support for me to go forward with passion.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my major professor, Dr. Kelley W. Moremen, for his support of my Ph.D. study and research. The freedom that Dr. Moremen gave to me in the research allowed me to explore on my own, and thus to develop the abilities to learn and work more independently. I have been and will continuously be benefited by the knowledge and capabilities that I have acquired in Dr. Moremen's lab.

I would like to thank the rest of the committee professors: Dr. Michael Pierce, Dr. Michael Tiemeyer and Dr. Lance Wells for their encouragement, thoughtful understanding, insightful comments as well as their hard questions.

I would like to give my sincere thanks to the organizers and scientists in the CCP4 workshop held in the Argonne National Laboratory in 2011. Their expertise, devotion and unselfish scientific spirits enlightened me and helped me make a crucial breakthrough in my research.

I would like to thank Dr. Bicheng Wang, Dr. John Rose, Dr. Zachary Wood and Dr. William Lanzilotta at the University of Georgia for their advice and generosity in letting me have the access to the crystallography facilities in their labs. I want to thank the staff in Dr. Wang's lab, especially Dr. Hao Xu, Dr. Lirong Chen, Ms. Hua Zhang, and Ms. Zhiqu Li for their assistance. I would like to thank our collaborators Dr. Jessica Cardot and Dr. Geert-Jan Boons for providing the compounds for this research.

My heartfelt gratitude also goes to Patricia Sheahan whose kindness and wisdom have been guiding me through all these years, and all the members in Moremen lab as well as the

colleagues in the Complex Carbohydrate Research Center who have also helped me in their ways.

It's my great pleasure to also thank my friends outside UGA: Olga Stuchlik, Defeng Deng, Yihong Cai, Gerald Rowland, Emily Rowland, Jon Antila, Zhaoxia Liu, Caryn Hale, Conor Hale, Yu Zhang, Pin Fan, Zhenjiang Li and Weiyong Wang. Their genuine friendship made me a better person.

Most importantly, I can never thank my family enough. To finish this dissertation would have been impossible without the love and support of my dearest family to whom this dissertation is dedicated.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

Introduction

This dissertation is focused on the structure-substrate specificity relationship study of two enzymes: human endoplasmic reticulum (ER) mannosidase I and mouse Golgi mannosidase IA. The two enzymes belong to the CAZy Family 47 glycohydrolases (GH47), and are also known as class 1 α1,2-mannosidases [1]. They have been studied extensively regarding to their catalytic activities and mechanisms, substrate specificities, metal ion coordination, structural features, enzyme-related diseases as well as their inhibitors. However, little is known about the structural determinants required for their distinctive substrate specificities. All the acquired knowledge about the GH47 α1,2-mannosidase, together with prior studies on the structural determinants for their function, constitute the foundation for this dissertation study. The introduction of this dissertation aims to present this literature foundation with a summary of the previous studies of the enzyme family, in particular of human ER mannnosidase I (ERManI) and mouse Golgi mannosidas IA (GMIA). The previously well-documented biological functions and significance of the enzymes are briefly described in the introduction while the major parts of the introduction are dedicated to the structural studies as they are the main subjects of this dissertation. To present the structural data coherently and systematically, comparison of the two enzymes is the logical thread in describing the results from the previous studies, with the highlights on the similar aspects and moreover the differences displayed by the enzyme structures. At the end of the introduction, a brief research plan for the next-stage study is outlined as the outcome of the literature review and as a preview for this dissertation.

1.1. General overview of GH47 α-Mannosidases

Mannosidases involved in glycoprotein maturation and catabolism have been divided into three subfamilies. Two of the families are named as class 1 (GH47) and class 2 (GH38) α-mannosidases [2-6], and they are exo-mannosidases. The class 3 (GH99) enzymes are endo-α-mannosidases. The class 1 (GH47) α-mannosidases are the focus of this dissertation.

The GH47 α1,2-mannosidases can be divided into three distinct subgroups. Subgroup 1 and 2 enzymes are required for the maturation of *N*-glycans from high-mannose to complex or hybrid structures [7]. Subgroup 1 includes the yeast and mammalian ER α1,2-mannosidases that preferentially form the $Man_8GlcNAc_2$ isomer B (lacking a mannose on the middle branch B of oligosaccharide) from $Man_9GlcNAc_2$ [8, 9]. Subgroup 2 includes three mammalian Golgi α1,2-mannosidases, as well as fungal enzymes that form $Man_8GlcNAc_2$ isomer A and/or C from $Man_9GlcNAc_2$ [10, 11]. Subgroup 3 proteins, such as the mammalian EDEMs1-3 and yeast Htm1p/Mn11p, take part in ERAD of misfolded glycoproteins [12-15]. Although EDEM1 and EDEM3 have recently been shown to have glycoside hydrolase activity in *vivo* [16, 17], no activity has been demonstrated for mammalian subgroup 3 proteins in *vitro* [13, 14].

The GH47 α1,2-mannosidases have several characteristics that differ from other glycoside hydrolase families : 1) they have considerable sequence similarity within a conserved 440-510 amino acid catalytic domain; 2) they specifically cleave α1,2-mannosidase linkages; 3) Ca^{2+} is required for their catalytic activity; 4) they are inhibited by the pyranose substrate mimics, 1-deoxymannojirimycin (dMNJ) and kifunensine (Kif); 5) they are inverting enzymes, which cleave the glycosidic linkage by inversion of configuration of the released mannose residue [11, 18-20]. In contrast, the class 2 (GH38) mannosidases are present in the ER, Golgi, lysosomes, and cytosol of mammalian cells. Class 2 mannosidases require Zn^{2+} or Co^{2+} for catalytic activity.

They are inhibited by the furanose transition state analogs swainsonine and 1,4-dideoxy-1,4-imino-D-mannitol [20], and cleave glycosidic linkages by retention of anomeric configuration [21].

1.2. Overview of *N*-glycosylation

N-linked glycans are synthesized by enzymes in the ER and Golgi complex, starting from a common oligosaccharide precursor to form an array of complex structures [22]. The synthesis and maturation of *N*-linked oligosaccharides occur in three stages, which are conserved in virtually all eukaryotic organisms [23-25].

1.2.1. Synthesis of a single oligosaccharide precursor and its transfer to polypeptides

In the first stage, a dolichol-linked $Glc_3Man_9GlcNAc_2$ oligosaccharide precursor is synthesized and transferred *en bloc* to the Asn-X-Ser/Thr sequons on newly synthesized polypeptides by oligosaccharide transferase complex [26] (Figure 1).

1.2.2. Early *N*-glycan processing events and ER quality control

In the second stage, glycan trimming occurs after the transfer of $Glc_3Man_9GlcNAc_2$ structures to proteins. Two of the three glucose residues are removed by glucosidases I and II resulting in polypeptides containing $GlcMan_9GlcNAc_2$ structures (Figure 1). Calnexin (CNX) and calreticulin (CRT), the homologous ER luminal and membrane lectin chaperones, recognize *N*-glycans containing $GlcMan_9GlcNAc_2$ structures as the natural ligands [27]. These lectins either associate with the glycan polypeptides directly, or through the recruitment of other proteins such as ERp57 [28]. Correctly folded glycoproteins are released from CNX/CRT and are acted upon by glucosidase II, which removes the last glucose residue and eliminates further interaction with the lectin chaperones. Glycoconjugates are then assembled into anterograde transport vesicles and proceed through the secretory pathway to the Golgi complex.

Incompletely folded proteins remain transiently associated with CNX/CRT until glucosidase II removes the last glucose residue [29, 30]. Unfolded glycoproteins containing $Man_9GlcNAc_2$ structures can be acted upon by UDP-Glc:glycoprotein glucosyltransferase (UGGT) to add the Glc residue back in the same position and linkage where it was removed by glucosidase II, thus allowing the protein to re-bind the CNX/CRT. UGGT works as a "sensor" of the folding status of the newly synthesized glycoproteins by recognizing only unfolded glycoprotein intermediates [31].

In order to achieve effective folding, the newly synthesized glycoproteins go through a multi-round cycle, which consists of binding to CNX/CRT, sugar cleavage, release from the chaperones, sugar addition, and re-binding to the chaperones [27]. Correctly folded glycoproteins are released from the ER chaperones and are transported to the Golgi, whereas glycoproteins that fail to fold after a defined period of time are targeted for ER associated degradation (ERAD) [32, 33].

Mannose trimming to a specific $Man_8GlcNAc_2$ isomer B by ERManI is a key step in targeting unfolded or misfolded proteins for degradation [34, 35]. Disposal of mutant carboxypeptidase Y (CPY) in *S. cerevisiae* was abrogated by blocking the formation of the $Man_8GlcNAc_2$ isomer B structure through gene disruptions in the ERManI. Misfolded CPY accumulated in the ER because of failure to form the $Man_8GlcNAc_2$ isomer B structure [36]. In mammalian cells, the degradation of misfolded forms of a variety of proteins, including α1-AT [37], the T cell receptor subunit CD3-δ [38], pre-pro α-factor [39], and tyrosinase [40] were all blocked by treatment of the cells with the GH47 α-mannosidase inhibitors, deoxymannojirimycin (dMNJ) or kifunensine (Kif), indicating that the formation of the $Man_8GlcNAc_2$ isomer B structure by ERManI was necessary for glycoprotein degradation. Overexpression of ERManI

accelerates the disposal rate of misfolded glycoproteins, or even wild type glycoproteins [34, 41]. Thus, glycan trimming to $Man_8GlcNAc_2$ isomer B structure by ERManI acts as the rate-determining step in generating the signal for glycoprotein disposal.

1.2.3. Further *N*-glycan processing events in Golgi apparatus

Properly folded glycoproteins with $Man_8GlcNAc_2$ *N*-glycans are transported from the ER to the Golgi complex, where the Golgi α1,2-mannosidase subfamily enzymes (Golgi α1,2-mannosidase IA (GMIA [42]), Golgi α1,2-mannosidase IB (GMIB [43]), and Golgi α1,2-mannosidase IC (GMIC [10])) cleave the remaining α1,2-mannosyl linkages to generate $Man_5GlcNAc_2$ structure [42-44]. GlcNAc transferase I catalyzes the transfer of a single GlcNAc from UDP-GlcNAc to the α 1,3-linked core mannose to form $GlcNAcMan_5GlcNAc_2$. Golgi α-mannosidase II then cleaves the terminal α1,3- and α1,6-linked mannose residues from $GlcNAcMan_5GlcNAc_2$ to form $GlcNAcMan_3GlcNAc_2$. The prior addition of the single GlcNAc by GlcNAc transferase I is the prerequisite for mannosidase II function. The branching extension (i.e. addition of fucose, sialic acid, polylactosamine) of the oligosaccharides by a series of Golgi glycosyltransferases constitutes the third and most varied stage of *N*-glycan maturation to generate the extreme diversity of *N*-glycan structures found in eukaryotic glycoproteins (Figure 1).

1.3. The substrate specificities of GH47 α1,2-mannosidases

GH47 α-mannosidases recognize and cleave only α1,2-linked mannose residues from the $Man_9GlcNAc_2$ high-mannose precursor, but each of the subfamilies has unique specificities for recognition of terminal α1,2-mannose residues in high-mannose structures. ERManI is the first mannosidase to act in this pathway, it primarily cleaves a single mannose residue from branch B of $Man_9GlcNAc_2$ to produce a unique $Man_8GlcNAc_2$ isomer (Man8 isomer B; Figure 2).

ERManI shows poor efficiency in cleaving the mannose residues on branches A and C of $Man_9GlcNAc_2$ [45]. *In vitro* digestion of $Man_9GlcNAc_2$ by ERManI resulted in the rapid production of the $Man_8GlcNAc_2$ isomer, but further cleavage of the remaining α1,2-mannose residues to produce the $Man_5GlcNAc_2$ intermediate was far less efficient [45]. The Golgi ManI subfamily of enzymes (GMIA, GMIB and GMIC) each share similar substrate specificities, which are complementary to the action of ERManI [11]. *In vitro* the Golgi ManI isoforms can efficiently trim the mannose residues on branch A and C of $Man_9GlcNAc_2$, but poorly cleave the central branch mannose residue that is the target of ERManI action [11]. In contrast, the *in vivo* cleavage of *N*-glycans is initiated by ERManI action to cleave $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$ isomer B followed by the action of GMIA/GMIB/GMIC to cleave the remaining three α1,2-mannose residues from $Man_8GlcNAc_2$ to form $Man_5GlcNAc_2$ (Figure 2). Thus, ERManI and Golgi ManI enzymes have complementary and non-overlapping specificities for cleavage of all four α1,2-mannose residues from the tri-antennary high mannose glycan precursor structure.

1.4 The structural studies of GH47 α1,2-mannosidases.

The structural studies on GH47 α-mannosidases were first achieved with *Saccharomyces cerevisiae* ERManI [46], followed by human ERManI [47], *Tricoderma reesei* α1,2-mannosidase [48], *Penicillium citrinum* α1,2 mannosidase [49]**,** mouse GMIA [50] and a bacterial GH47 α1,2-mannosidase [51]. The structures were solved not only as apo-enzymes, but also in complex with different enzyme inhibitors and disaccharide analogs and led to an understanding of the enzyme structures, enzyme-inhibitor complexes, and enzyme-disaccharide complexes. The development of structural studies for the GH47 α-mannosidases and the features of the structures were presented in flow chart in Figure 3.

The crystal structures of those enzymes revealed that they all consist of an $(\alpha\alpha)_7$ barrel domain structure comprised of two layers of concentric inner and outer helices, in which the catalytic site is located deep within the barrel above a β-hairpin inserted into one end of the barrel. The β-hairpin contains a Thr residue in the core of the barrel that is involved in coordinating a Ca^{2+} ion assisted by indirect interactions through water molecules. The opposite end of the barrel consists of an open conical cleft leading to a more constricted pocket in the core of the barrel (Figure 4).

Superimposing the structures of GH47 α-mannosidases indicated that the general fold of the proteins is similar, with a high degree of overlap in the enzyme barrel core, and more deviations and variation on the solvent-exposed exterior (Figure 5). Structural comparisons showed that there were a collection of conserved acidic amino acid residues in the core of the active site contributing to Ca^{2+} interactions and that these residues were essentially superimposable for all family members [49, 50, 52]. The structural similarity is consistent with their similarity in primary sequence. In addition, the high structural similarity of the substrate binding -1 and +1 subsites was also observed indicating that the catalytic mechanisms of the mammalian, fungal, and bacterial enzymes are very likely to be identical. The main functional differences between the ERManI and GMIA were the respective branch specificities for glycan cleavage, which is likely to result from their structural difference of the ≥ +1 subsites.

1.4.1. The crystal structure of *S. cerevisiae* ER ManI: a complex of enzyme and product

The crystal structure the *S cerevisiae* ERManI was the first resolved structure of a GH47 α1,2-mannodsidases [46]. This enzyme is a glycoprotein with three *N*-glycans. The crystal structure of this enzyme revealed that a protein molecule interacts with the neighboring protein molecule in the crystal lattice by inserting one $Man_5GlcNAc_2$ glycan into the enzyme active site

cleft of the neighboring enzyme to form a protein-glycan complex (PDB entry: 1DL2). The -1 glycan binding subsite was vacant in this structure while the ≥+1 subsites were occupied with the $Man_5GlcNAc_2$ glycan structure. Since the $Man_5GlcNAc_2$ found in the enzyme cavity is the enzymatic product of an extensively processed glycan structure by the *Saccharomyces cerevisiae* enzyme, it was proposed that the structure reflected enzyme-product complex. (Figure 6)

Although the -1 subsite is unoccupied, mannose residue M7 of branch B was found to be located in the +1 subsite and branches A and C were spread out and extended across the open face of the enzyme active site cleft. The structure and conformation of the processed glycan in the enzyme-product complex suggested the preference of yeast ERManI for binding to the intact glycan structure: residues M10 and M7 of branch B were preferred to bind to the -1 and +1 subsites rather than the terminal α1,2-linked mannosyl residues on the branches A and C. However, direct structural evidence for interactions with the intact $Man_9GlcNAc_2$ glycan substrate are still required to confirm this hypothesis (Figure 7).

1.4.2. The crystal structures of enzyme-inhibitor complexes of human ERManI, *Penicillium citrinum* α1,2-mannosidase, and bacterial GH47 α1,2-mannosidase.

The structural studies of the enzyme-inhibitor complex were first reported for human ERManI in complex with the inhibitors dMNJ and Kif. The structures of co-crystallized enzyme-inhibitor complexes demonstrated that the core of the barrel was the site of glycone interactions [47] . In both the enzyme-dMNJ complex and enzyme-Kif complex the inhibitors bound at the base of the enzyme cleft in the -1 subsite with the equivalent C-1 position of the sugar facing outward towards the opening of the pocket. In contrast to the yeast ERManI crystal structure, in which the -1 subsite is vacant and the ≥+1 subsites were occupied, for the inhibitor complexes

the -1 glycan binding subsite was occupied by dMNJ or Kif, while the ≥+1 subsites remained unoccupied. (Figure 8)

Two unique structural characteristics of GH47 α-mannosidases were found in the enzyme-inhibitor structures. First, the Ca^{2+} ion was found to bind directly with the C-2 and C-3 hydroxyl oxygens of the inhibitors, indicating that the Ca^{2+} ion was involved directly in substrate interactions. Second, the conformation of the inhibitor within the active site was unique. Free mannose in solution, or as part of an oligosaccharide, is mostly found in the $^4C_1$ conformation, with the C-3 and C-4 hydroxyls and the C-5 hydroxymethyl groups in low free energy equatorial positions. In contrast, the six-member ring of both Kif and dMNJ bound to the enzyme in the energetically unfavorable $^1C_4$ conformation, with the C-3 and C-4 hydroxyls and the C-5 hydroxymethyl are all in an axial conformation. This energetically unfavorable $^1C_4$ confirmation is stabilized through direct interactions with the Ca^{2+} ion as well as extensive hydrogen bonding and hydrophobic interactions in the -1 subsite [47].

The results from the enzyme-inhibitor complex studies for the *Penicillium citrinum* α1,2-mannosidases agreed with those of the human ERManI enzyme-inhibitor structure, both dMNJ and Kif were found to bind the active site of the *Penicillium citrinum* α1,2-mannosidases with the $^1C_4$ conformation [49].

The complex structures of *Caulobacter* GH47 α1,2 mannosidase with noeuromycin and mannoimidazole (two other α-D-mannopyranose analog) also showed noeuromycin occupied the enzyme active site with $^1C_4$ conformation [51]. A distinct sugar conformation was observed for the mannoimidazole complex, where it bound the *Caulobacter* GH47 α1,2-mannosidase in a $^3H_4$ conformation, which was previously proposed as the transition state for GH47 α-mannosidase

action. This is the sole inhibitor complex where this proposed transition state conformation was observed among the GH47 α-mannosidases.

The inhibitors bound in the -1 subsites of human ERManI and *Caulobacter* GH47 α1,2-mannosidase are shown in Figure 8. The conformations of inhibitors (dMNJ and Kif in human ERManI; noeuromycin and mannoimidazole in *Caulobacter* GH47 α1,2-mannosidase) are presented in Figure 9.

1.4.3. The crystal structures of enzyme-disaccharide complexes of human ERManI, *Penicillium citrinum* α1,2-mannosidase, and bacterial GH47 α1,2-mannosidase.

The yeast ERManI crystal structure and the complex structures of α1,2-mannosidases with inhibitors revealed either the occupied ≥+1 subsites or the occupied -1 subsite. In order to reveal how the -1 and +1 subsites are simultaneously bound by a Man-α1,2-Man substrate analog, and to understand whether and how the mannose in the +1 subsite assists the nonreducing-end mannose in interaction with the -1 subsite, different disaccharide analogues were introduced into the active site structures of two α1,2-mannosidases. In the crystal structure of *Penicillium citrinum* α1,2-mannosidase complex with methyl-α-D-lyxopyranosyl-(1',2)-α-D-mannopyranose (LM) the reducing end of LM bound the -1 subsite site in a $^{1}C_{4}$ conformation, and the reducing end bound in the +1 subsite with $^{4}C_{1}$ conformation. The crystal structure of a co-complex between an uncleavable thio-disaccharide analogue, Man-α1,2-thio-Man-O-methyl (S-Man_2), and human ERManI revealed that the disaccharide analogue binds across the -1 and +1 subsites, occupying the +1 subsite in the low-energy $^{4}C_{1}$ conformation like LM in the *Penicillium citrinum* enzyme. However, unlike the $^{1}C_{4}$ conformation of LM in the -1 subsite, S-Man_2 bound the -1 subsite with a skew boat $^{3}S_{1}$ conformation [49]. A recently solved co-crystal structure of *Caulobacter* G47 α1,2-mannosidase in a S-Man_2 complex showed that the thiodisaccharide also

bound in the -1 and +1 subsites of this enzyme, with the $^{3}S_{1}$ conformation, in the -1 subsite and $^{4}C_{1}$ conformation in the +1 subsite [51]. The enzyme-bound S-Man_2 and LM in human ERManI and *Penicillium citrinum* α1,2-mannosidase, respectively, are shown in Figure 10, the conformations of the non-reducing end of the two compounds in the -1 subsite of the enzymes were shown in Figure 11.

1.4.4. The crystal structures of Golgi α1,2-mannosidase IA: another enzyme-product complex

Similar to the *S. cerevisiae* ERManI structure, the mouse GMIA structure also revealed an enzyme-product complex in the crystal lattice. The Golgi enzyme expressed in *P. pastoris* contains a single *N*-glycan structure and the crystal structure indicated an electron density corresponding to a $Man_6GlcNAc_2$ *N*-glycan in the glycan binding cleft of GMIA. A $Man_5GlcNAc_2$ *N*-glycan would be expected as the limit enzymatic cleavage product of the α1,2-mannosidase action, yet an extra mannose residue (denoted as *Man* in Figure 12 to differentiate it from the other mannose residues M3-M7 of $Man_5GlcNAc_2$) was added to the M5 residue through α1,6 linkage during the combinational protein expression and secretion in the fungal host *P. pastoris*. The oligosaccharide (glycan A) associated with one protein molecule (molecule A) in the crystal lattice was found to extend into the cleft of the adjoining protein molecule (molecule B) to form an equivalent of an enzyme-product complex [50] (Figure 12). The ≥+1 subsites of GMIA were occupied by the $Man_6GlcNAc_2$ glycan, but the -1 subsite of the GMIA was not occupied in the enzyme-product complex similar to the $Man_5GlcNAc_2$ complex found for the *S. cerevisiae* enzyme.

1.4.5 Comparison of the glycan conformation in the two enzyme-product complexes of *S. cerevisiae* ERManI and mouse GMIA.

Comparison of the two enzyme-product complexes of *S. cerevisiae* ERManI and GMIA revealed the similarities and differences in glycan interactions for the two enzymes. The core GlcNAc residues of the glycan ligands extended out of the binding cleft in an approximately similar direction. However, the positioning of the glycan residues in the ≥+1 glycan binding pockets were quite different. The core of the two glycans displayed a ~180^0 rotation around the M3-M4 linkage, which resulted in the insertion of alternative glycan branch into the +1 subsite. The glycan in GMIA used branch C to extend to the active site (though the -1 subsite was vacant), and the mannose residue M6 was found in the +1 subsite for the Golgi enzyme. In contrast, the glycan in the yeast enzyme used branch B to approach the active site (the -1 subsite is also unoccupied) and mannose M7 was found to reside in the +1 subsite (Figures 11 and 12)

Although there was no residue in the -1 subsite in either of the two co-complexes, the cleaved residue that originally bound in the -1 subsite before the cleavage can be inferred by determining which residues were connected through α1,2 linkage to M6 (for GMIA), or M7 (for yeast ERManI) in the intact $Man_9GlcNAc_2$ substrate. Therefore, the complexes reflect the insertion of the M6-M9 residues (branch C) into the +1 and -1 subsites for GMIA, and the insertion of M7-M10 residues (branch B) into the +1 and -1 subsites for *S. cerevisiae* ERManI [50].

1.5. The coordination of Ca^{2+} in the active site of the GH47 α-mannosidase.

The Ca^{2+} is essential for the catalytic activity and substrate binding by the GH47 α-mannosidases [43, 53, 54]. A Ca^{2+}-centered 8-fold square antiprismatic coordination geometry has been found in all of the structures of wild type GH47 α1,2-mannosidases [47]. In such a coordination geometry, the Ca^{2+} ion interacted with eight oxygen atoms, two were from a conserved Thr residue in the β-hairpin at the bottom of the barrel, the rest were from water

molecules or ligand hydroxyls when bound in the active site. In the crystal structure of GMIA, for example, there was no mannose analog bound in the -1 subsite, therefore the Ca^{2+}-centered eight-fold coordination consisted of six water molecules and two atoms from Thr635 (O-γ and the backbone carboxyl oxygen atom). Two oxygen atoms from Thr635 and two water molecules (W5 and W6) were on one side of the square antiprismatic structure, and the other four water molecules (W1, W2, W3, and W4) were on the opposite side (Figure 13). Similar water and Thr coordination were found in yeast ERManI [55]. However, the enzyme-inhibitor and enzyme-disaccharide complexes differed from the non-ligand bound enzymes by substitution of two water molecules with two oxygen atoms from the respective ligands. For example, in the enzyme-dMNJ complex of human ERManI, there were only four water molecules (W1, W2, W3, and W4) coordinating the Ca^{2+} ion from one side of the square antiprismatic structure [47]. On the other side, instead of two waters W5 and W6 like in GMIA, the C-2 and C3 hydroxyl oxygens from the inhibitor and the two oxygens form the Thr688 residue were directly coordinated with the Ca^{2+} ion [50] (Figure 13). The participation of the enzyme-bound ligands in the metal ion coordination is a main feature of the enzyme-inhibitor and enzyme-disaccharide complexes. The four coordinating water molecules (W1, W2, W3 and W4) found in all GH47 α-mannosidase structures not only interact with the Ca^{2+} but also interact with other amino acid residues in the base of the active site pocket. In the GMIA crystal structure, for example, W1 interacts with Glu429, Thr635 and Glu552; W2 interacts with Glu552, Glu549, Glu611 and Thr635; W3 interacts with Glu549; and W4 interacted with the side chain of Glu419 and Thr635. Those residues and interactions are conserved among the crystal structures of all GH47 α-mannosidase family members.

1.6. The enzyme-ligand interactions in the -1 subsite.

In addition to the coordination between the inhibitor C2 and C3 hydroxyl groups and the enzyme-bound Ca^{2+} ion, the inhibitors in the -1 subsite also interacted directly with amino acid residues within the active site pocket. For example, in the complex of human ERManI and dMNJ, the inhibitor interacted with the amino acid residues Arg597, Glu599, Glu663, and Glu689 directly, and with Glu330, Glu602, Glu467 and Asp463 indirectly via water molecules [47, 49] (Figure 14). The inhibitor also used C-4, C-5 and C-6 to provide hydrophobic interaction with Phe659 in human ERManI. Other observed $^1C_4$ conformers such as Kif in human ERManI, the non-reducing end of LM in the *Penicillium citrinum* α1,2-mannosidase, and noeuromycin in the bacterial GH47 mannosidase have also been found to bind with the enzymes through the identical interactions with equivalent residues in the -1 subsites [49, 51].

The interactions between enzymes and ligands in the -1 subsite were highly similar among the GH47 α-mannosidase family members because the amino acid sequences of the -1 subsites are highly conserved, and the equivalent of the C-2, C-3, C-4, C5, and C-6 atoms of the ligands in the -1 subsites, including the $^1C_4$ (dMNJ, Kif, noeuromycin, and LM), $^3S_1$ (S-Man_2) or $^3H_4$ (mannoimidazole) conformers, were all found to have essentially identical configurations in the complex structures (Figure 15). However, the C-1 and ring nitrogen N-6 (or ring oxygen O-6) in the $^3S_1$ (S-Man_2) or $^3H_4$ (mannoimidazole) conformers were distorted and in altered positions relative to their counterparts in the $^1C_4$ conformer (dMNJ, Kif, noeuromycin, and LM), (Figure 15). The C-1 atoms of the ligands either did not have functional groups (dMNJ), or directed them towards the opening of the catalytic site (C−S bond in S-Man_2; C−O bond in LM and noeuromycin; C−N bond in Kif; the C=N bond in mannoimidazole). Therefore, all hydrogen bonding between the enzymes and bound ligands in the -1 subsite were essentially the same for all of the α1,2-mannosidase structures. In addition, the ligands in the -1 subsites also participated

the hydrophobic interactions using their C-4, C-5 and C-6 atoms, despite the varied configurations of the C-1 atoms, and those hydrophobic interactions were similar for all the ligands, except for LM, which did not contain a C-6 atom.

An example of the conserved enzyme-mannose analog interactions in the -1 subsite was displayed by the complex of human ERManI and S-Man_2. The non-reducing end glycone of S-Man_2 was found in the $^3S_1$ skew boat conformation [54]. Compared to their respective configurations in the $^1C_4$ conformer, the C-2, C-3, C-4, C-5 and C-6 retained the same configurations in the $^3S_1$ conformer, but C-1 and the ring oxygen O-6 were distorted (Figure 15). All the hydrogen bonds between the non-reducing end of S-Man_2 in the -1 subsite with the enzymes were identical to those in the enzyme-dMNJ complex, including hydrogen bonding directly with R334, R597, E599, E663, E689, and hydrogen bonding via water molecules with E330, S464, E599. **(**Figure 16). Like the $^1C_4$ conformer, the $^3S_1$ conformer also interacted with the enzyme through hydrophobic interactions with Phe659 and C-4, C-5, and C-6 of the substrate analog.

1.7 Computational studies of the conformations in the -1 subsite

All the crystal structures of enzyme-inhibitor and enzyme-disaccharide complexes of the GH47 α1,2-mannosidases revealed that the D-α-mannopyranose analogs reside in the -1 subsite with a $^1C_4$, $^3H_4$, or $^3S_1$ conformation, in which the C-3, C-4 hydroxyls and C-5 hydroxylmethyl group all took on axial positions [49, 51, 54, 56]. In contrast, free α-D-mannopyranose in solution favors the $^4C_1$ conformation, which has the lowest free energy among all sugar conformations, with C-3 and C-4 hydroxyls, and C-5 hydroxylmethyl groups in the sterically less-hindered equatorial position (with the exception of the C-2 hydroxyl group, which is axial for mannose in the $^4C_1$ conformation). The experimental results for the bound sugar

conformation raise some questions: why $^4C_1$ was not the bound conformer in the -1 subsite? With multiple possibilities for starting substrate conformation, the conformation of the pseudo-Michaelis complex, the transitional state, and the conformation of the enzymatic product, which conformer is represented in each individual state? A computational study on the conformers in the -1 subsite gave some clues to answer those questions.

A computational study was previously performed in an effort to explain why the $^1C_4$ was one of the preferred conformations in interactions with the -1 subsite in terms of binding energy contributions. The models of dMNJ, Kif, and α-D-mannopyranose in $^1C_4$ conformation (dMNJ-$^1C_4$, KIF-$^1C_4$, and Man-$^1C_4$) and $^4C_1$ conformation (dMNJ-$^4C_1$, Kif-$^4C_1$, and Man-$^4C_1$) were computationally docked in the active site of the *S. cerevisiae* ERManI [57]. The docked Man-$^1C_4$ model fit in the enzyme in a very similar manner as dMNJ/Kif found in the complex crystal structure; the computation results showed that dMNJ-$^1C_4$ and Kif-$^1C_4$ had more negative docking free energy than the $^4C_1$ conformers. For example, the docking energy of dMNJ-$^1C_4$ (-95.00 kcal/mol) is thermodynamically more favorable than that of dMNJ-$^4C_1$ (-77.70 kcal/mol), suggesting the $^1C_4$ conformer bound enzyme more stably. Comparison of the docking energy of Man-$^1C_4$ (-88.51 kcal/mol) and Man-$^4C_1$ (-84.21 kcal/mol) indicated that Man-$^1C_4$ was only slightly more favorable for to binding to the -1 subsite. However, it was also shown in the same study that the enzyme active-site funnel-shaped neck dimension between the -1 and +1 subsites was too constricted for Man-$^4C_1$ to gain access to the -1 subsite, because of the equatorially positioned ring substituents, while the more compact Man-$^1C_4$ with axial hydroxyl positioning was able to more readily access the subsite. Two salt bridges on either side of the opening, Glu132–Arg136 and Glu399–Arg433, stiffen the opening and reduce its flexibility. There is also the possibility that the neck expands to allow entry of the ligand (Figure 15). However, if the

neck of the opening were flexible, its flexibility would be reflected in the temperature factors of the residues of the neck of the active-site funnel in the yeast crystal structure. These residues show very low temperature factors (13–20 $Å^2$) and belong to the helices in the interior of the protein (except for Arg433, which belongs to a loop but also has low temperature factors). In addition, it is anticipated that an expansion of the helical core to accommodate the ligand would be energetically too expensive. Local movement of the side chains of the neck residues seems restricted, as is reflected by their low temperature factors [57].

The docking studies also suggested that the observed $^1C_4$ conformer could represent the starting state of the substrate interactions. The disaccharide substrate model, α-D-Man*p*-(1,2)-α-D-Man*p* (α-Mannobiose) was docked into the yeast ERManI -1 and +1 subsites with its glycone in chair ($^1C_4$, $^4C_1$), half-chair ($^3H_2$, $^3H_4$, $^4H_3$), skew boat ($^OS_2$,$^3S_1$, $^5S_1$), boat ($^{2,5}B$, $^{3,O}B$, $B_{1,4}$, $B_{2,5}$) and envelope (^3E, $_4E$, E_3, E^4) conformations [58]. The calculated docking energies and forces suggested that both $^1C_4$ and $^OS_2$ can be the starting conformers. However, so far, only $^1C_4$, but not the $^OS_2$ conformer, has been observed in the crystal structures of GH47 α-mannosidases.

Computational studies also supported the observed $^3S_1$ conformer as the pseudo-Michaelis complex. The previous computational study on β-D-mannopyranose revealed that the isolated sugar can access a small number of possible conformations in lower free energy states [59]. However, individual glycoside hydrolase families select only one of these possible conformers as the pre-activated intermediate to reach the transition state [60]. For α-D-mannopyranose $B_{2,5}$, $^OS_2$, $^{3,O}B$, $^3S_1$, $B_{1,4}$, $^5S_1$, $^{2,5}B$, and $^4C_1$ are among the most stable conformations, and these conformations were found to be pre-activated for catalysis in terms of energy, anomeric charge, and structural parameters [51]. The observed conformations for pseudo-Michaelis complexes in α-mannosidases ($^OS_2$ and $^3S_1$ in retaining and inverting enzymes, respectively) are among the

most stable ones, and thus represent pre-activated conformations. The quantum and molecular mechanics calculations using α1,2-mannobiose as the disaccharide substrate model on the bacterial G47 α-mannosidase indicated that this enzyme has restricted the energetically accessible conformational landscape of the -1 mannose residue. The undistorted $^{4}C_{1}$ conformer is no longer an energy minimum. The $^{3,O}B/^{3}S_{1}$ conformations are the only stable distorted conformations, defining a clear $^{3,O}B/^{3}S_{1} \rightarrow {}^{3}H_{4} \rightarrow {}^{1}C_{4}$ conformational pathway for the reaction coordinate [61, 62].

The $^{3}S_{1}$ conformer as the pre-activated intermediate is also supported by its own structural features. The generally accepted transition state (TS) model of glycosidases is an oxocarbenium ion-like state that consists of partially formed/broken bonds between the attacking nucleophile and the glycosidic oxygen of the leaving group. This oxocarbenium ion has partial double-bond character between the anomeric carbon and the ring oxygen and results in ring flattening of the C5-O5-C1-C2 bonds. For bond-breaking and bond-making to occur with the proper orbital overlap in the catalytic cleavage, the lone pair orbital of the nucleophilic water oxygen must be located in an anti-periplanar geometry (on opposite side) with respect to the glycosidic bond oxygen (leaving group), and C-5, O-5, C-1 and C-2 at the transition state must be on the same plane [63-65]. Among the two identified conformers in the -1 subsite of the crystal structures of GH47 α-mannosidases, the $^{3}S_{1}$ conformation satisfies the antiperiplanar requirement for formation of the oxocarbenium ion-like TS, and can switch directly to the TS with the minimal distortion of a conformational twist at C-5, O-5, C-1 and C-2. The $^{4}C_{1}$ glycone residue is anticipated to experience a few different conformations in the process of going from its low free energy ground state to the TS in the -1 subsite. Therefore, $^{3}I_{1}$ may present as the conformation of the intermediate in closest proximity to the TS.

1.8. The conformation of the transition state in the -1 subsite: $^{3}H_{4}$

The transition state conformations are often identified either by trapping covalent intermediate analogs or natural substrates in combination with inactive catalytic site mutants. Most enzyme transition states have been difficult to catch directly by experimental methods due to their poor stability and very short lifetime, whereas the pre-activated intermediates are relatively more stable. If a reaction goes through several intermediates, each intermediate needs go up-slope energetically to reach an unstable transition state with high energy, and then go down-slope energetically to form the following intermediate. The transition state could be inferred by the two intermediate conformations that bracket each side of the transition state. Structural data and computational studies indicated the $^{3}S_{1}$ conformation as the pre-activated intermediate. The $^{1}C_{4}$ conformer was indicated to represent the cleaved β-mannose product by both the computational study and identification of conformation in the dMNJ complex associated with human ERManI. The possible TS conformation was proposed to meet the two requirements: first, it should be between $^{3}S_{1}$ and $^{1}C_{4}$ on the conformational itinerary map (Figure 18); second, it must have an oxocarbenium-like structure. Among all of the pseudorotational conformational itineraries for pyranose sugar ring interconversions, there are only four potential conformations where the planarity of C-5, O-5, C-1 and C-2 is satisfied. Those four conformations are $^{2,5}B$, $B_{2,5}$, $^{4}H_{3}$, $^{3}H_{4}$. [63]. One of these potential transition state conformations are employed among all known hydrolase families: GH families 6, 8, 11,67 are proposed to use $^{2,5}B$, GH families 26 and 38 are proposed to use $B_{2,5}$, and GH families 5,7,18 and 20 are proposed to use $^{4}H_{3}$, and the $^{3}H_{4}$ conformation was proposed for the transition state for GH47 α-mannosidases. In support of this hypothesis, mannoimidazole bound in the active site of the bacterial GH47 mannosidase was observed to be in $^{3}H_{4}$ conformation. The C=N bond in mannoimidazole fixes the equivalent C-5,

N-5, C-1, and C-2 of the six member ring in the same plane and as an result the rigid structure of the six-member ring of mannoimidazole is a mimic of the transition state. This compound is the only example of a complex in the proposed TS conformation in crystal structure studies of GH47 α-mannosidases.

1.9. The catalytic mechanism of GH47 α1,2-mannosidases

1.9.1. The general catalytic mechanism of glycosidases

Enzymatic cleavage of glycosides generally requires the presence of two catalytic groups, usually Asp or Glu residues [61, 64, 65]. In an inverting glycosidase, there are two catalytic amino acid residues, one that is deprotonated (basic form) and functions as a base to deprotonate a nearby water molecule. The deprotonated water will then attack the anomeric carbon as a nucleophile. The other catalytic amino acid residue serves as a proton donor by protonating the glycosidic oxygen to assist the glycosidic bond cleavage.

The nucleophilic attack during bond hydrolysis has been proposed to require the formation of an oxocarbenium ion at or near the transition state. The charge of the oxocarbenium ion delocalizes between the anomeric center (C-1) and the ring oxygen (O-5), to form partial double bond character between C-1 and O-5 and results in ring flattening of the C5-O5-C1-C2 bonds. The nucleophilic water attacks the anomeric carbon C-1 from the opposite side of the leaving group that is originally linked to C-1 by the glycosidic bond. Thus, after the glycosidic bond breaks and the new bond formation, the configuration of C-1 is inverted (Figure 19).

1.9.2. Unique catalytic mechanism of human ERManI

The structure of human ERManI indicated that the distance between the proton donor amino acid residue and glycosidic oxygen is too large for direct protonation. Initial studies proposed a through-water protonation mechanism where an intervening water molecule works as

a shuttle to carry the proton from a Glu residue (E330 in human ERManI) to the glycosidic oxygen [54]. The "shuttle" water is first protonated by the Glu side chain, and then goes on to protonate the glycosidic oxygen by transferring the proton. The unique mechanism was illuminated as follows: proton transfer from the E330 general acid through water to the glycosidic oxygen results in partial glycoside bond cleavage and formation of the $^{3}H_{4}$ ring flattened oxocarbenium ion transition state, E599 deprotonates another water molecule to become the nucleophile. The latter deprotonated water attacks the C-1 carbon to produce $^{1}C_{4}$ chair configuration of the enzymatic product containing an inverted β-hydroxyl at the anomeric center. The mechanism is unique in that it employs an extra water molecule compared to the normal mechanism.

1.9.3. Computation studies proposed an alternative mechanism: Asp 463, instead of E330, is the catalytic proton donor in human ERManI.

The standard mechanism for glycoside hydrolase action includes the protonation of the leaving group hydroxyl oxygen usually by a closely positioned amino acid carboxyl side chain. Because of the lack of an appropriately positioned carboxyl side chain in the active site of GH47 α-mannosidases, it has been difficult to identify the proton donor for these enzymes. This led to a more speculative hypothesis for general acid function for the GH47 enzymes where a shuttle-water is protonated by a nearby acidic amino acid residue then the proton is transferred to the glycosidic oxygen making it a good leaving group during the nucleophilic attack on the anomeric carbon atom by another deprotonated water molecule. Automated docking experiments [61] agreed with the previously proposed mechanism [54] for Glu599 as the catalytic base, but came to different conclusions regarding the proton donor residue. The originally proposed proton donor for human ERManI was Glu330, but the results from the docking experiments favored

Asp463 based on comparison of the respective dissociation constants. In ERManI Glu330 and adjoining Arg334 are located so closely that they are likely an ion pair. As a result, Glu330 will be deprotonated, and the protonated Arg334 cannot be a proton donor because of its strong proton binding affinity. Only protonated Asp463 is suitable to assist as a proton donor through a separate water molecule with proper orientation to transfer the proton to the glycosidic oxygen atom.

1.9.4. Mutagenesis analysis revealed the impacts of some residues in human ERManI

In an effort to better understand the catalytic mechanism, a series of site-directed mutations were generated for potential catalytic residues of human ERManI [54, 66]. The influences of those residues on catalysis and substrate binding were examined, and the results are summarized below.

1.9.4.1.Mutation of the potential proton donor E330

E330 has been proposed to work as the general acid in the catalytic mechanism. Mutation of E330 to an isosteric amide (E330Q) resulted in a 44-fold decrease in $k_{cat,}$ and about 100-fold increase of binding affinity for a $Man_9GlcNAc_2$-gycopeptide ligand in surface plasmon resonance studies. The reduced hydrolytic activity but increased binding affinity of the E330Q resulted in a lectin-like activity for the enzyme toward high mannose glycans [54, 66].

1.9.4.2. Mutation of the potential base E599

E599 was proposed to act as the catalytic base in the enzymatic reaction. Mutation of E599 to the isosteric amide (E599Q) resulted in significant decrease in k_{cat} (10^5-fold). The binding affinity of E599Q decreased ~1000-fold for dMNJ, but only 4-fold for a $Man_9GlcNAc_2$-glycopeptide ligand by SPR, suggesting that the loss in catalytic turnover by the absence of the

catalytic base resulted in part from an unfavorable binding of glycone residue in the -1 subsite rather than substrate docking in the ≥+1 subsites [54, 66].

1.9.4.3. Mutation of the D463

D463 has been proposed by the computational docking studies as the potential general acid in the catalytic mechanism, but this residue also provides critical hydrogen bonding to the C3 and C4 hydroxyls of the Man residue in the +1 subsite. Mutation of D463 to an isosteric amide (D463N) resulted in a 1000-fold decrease in k_{cat} and a failure of binding to $Man_9GlcNAc_2$-gycopeptide ligand in the SPR experiments. These data suggest that this residue contributes to both catalysis as well as substrate binding affinity, but the relative contributions to both functions is not clear [54, 66].

1.9.4.4. Double Mutations.

Both the potential proton donor and catalytic base residues were mutated to test the enzyme activity. The catalytic activity of the E330Q/E599Q double mutant decreased to 0.06% of the wild type enzyme, and the combination of D463/E599 exhibited only 0.0003% of the wild type catalytic activity.

1.10. The mutation of T688 and the impact of Ca^{2+} ion

The essential role of the enzyme-bound Ca^{2+} ion was first revealed by chelation with EDTA, which led to the complete loss of activity for human ERManI. Similarly, the binding affinity for dMNJ in the absence of Ca^{2+} was decreased by 340-fold, but there was a minimal effect on binding affinity for $Man_9GlcNAc_2$ (2-fold decrease in binding affinity). A mutation of T688, the key residue involved in Ca^{2+} coordination, to Ala resulted in a reduction in k_{cat} by 62-fold [53], yet the affinity of the enzyme for Ca^{2+} was essentially unchanged. However, the binding affinity for the $Man_9GlcNAc_2$ substrate surprisingly increased by 50-fold and the binding

affinity for dMNJ decreased by 5-fold. These results indicated that the conserved T688 residue was not required for the Ca^{2+} binding, and neither Ca^{2+} nor T688 was required for the binding of the $Man_9GlcNAc_2$ substrate. However, the presence of Ca^{2+} was critical for catalysis and for dMNJ binding in the -1 subsite, which could be explained by altered coordination between the metal ion and dMNJ. The complete loss of catalysis along with residual binding affinity for the $Man_9GlcNAc_2$ substrate in the absence of Ca^{2+} suggested that the coordination of Ca^{2+} is more essential for the enzyme activity than for substrate binding. The fact that loss of metal coordination totally inactivated the enzyme and that the T688A mutant compromised activity indicated that the interaction between the T688 and the Ca^{2+} ion and especially the coordination between the Ca^{2+} ion and the side chain γ-O of the T688 had a significant effect on catalytic activity. Further investigation of the Ca^{2+} coordination in the T688A mutant may provide insights into the importance of proper ion coordination for enzyme function.

The crystal structure of the T688A-thiodisaccharide complex revealed a change in the Ca^{2+} coordination. Compared to the Ca^{2+}-centered 8-fold square antiprismatic coordination geometry of wild type human ERManI in complex with inhibitors and uncleavable substrate analogs, the Ca^{2+} in the T688A mutant in complex with S-Man_2 revealed coordination by only seven oxygen atoms (Karaveg and Moremen, unpublished). The missing point of coordination was due to the lack of a side chain hydroxyl group of in the T688A mutant residue. Even though the positions of the seven coordinating groups in the T668A-S-Man_2 complex are approximately similar to those in the complex of wild type enzyme and S-Man_2, the geometry of the seven coordinating oxygen atoms around the Ca^{2+} is different in the two complexes. In both the wild type and T688A mutant enzymes the Ca^{2+} is coordinated with a similarly positioned substrate analog and carbonyl oxygens from T688/A688 and no protein side chains other than the T688 residue were altered in

the mutant enzyme. These data indicate that the reduction in k_{cat} by the T688A mutant does not result from altered positioning of the substrate in the active site.

In contrast, the absence of the T688 hydroxyl side chain results in an alteration in geometry and positioning of the coordinating water molecules associated with the enzyme bound Ca^{2+} ion. In the wild type and T688A ERManI substrate analog complexes four water molecules coordinate to the enzyme bound Ca^{2+}. Three of the coordinating waters (W1, W2, and W3) are located in essentially superimposable positions between the two structures. However, the fourth water (W4) in the T668A mutant complex was significantly altered in position relative to the equivalent W4 in the wild type enzyme complex. This water was repositioned more closely to the missing Thr side chain γ-O in the T688A mutant. Thus, the square antiprismatic geometry of Ca^{2+} coordination of the wild type enzyme was altered by the removal of one point of Ca^{2+} coordination and repositioning of a second point of coordination as a result of the T668A mutation.

1.11. The outline of the dissertation study.

The goal of this dissertation study was to explore the structural basis for substrate interactions between two GH47 α-mannosidases as a foundation for exploring the distinctions in substrate specificity for glycan trimming by the two enzymes. The study includes three parts:

1. X-ray crystallography studies

The specificity of Golgi mannosidase IA and ER mannosidase I for substrate interactions has been studied through examining co-complexes of the enzymes with uncleaved $Man_9GlcNAc_2$ substrates. Both enzymes were inactivated by cation replacement. The natural substrate analog of the GH47 α1,2-mannosidases, $Man_9GlcNAc_2$-PA, was used

in X-ray crystallography studies with the two enzymes to elucidate the enzyme-substrate complex structures.

2. Mutation, enzyme kinetic studies and substrate binding studies

The structures of the enzyme-bound $Man_9GlcNAc_2$-PA complexes with ERManI and GMIA were compared. The common and distinct residues in both enzymes that were used in substrate interactions were compared and mapped. Enzyme mutation experiments were performed to swap the respective residues in ERManI to the corresponding ones in the GMIA. Both single-site mutants and multiple-site mutants were created, and tested for their catalytic activity and substrate specificity towards $Man_9GlcNAc_2$-PA. Surface plasmon resonance experiments were done using a few mutants to test their binding affinity to $Man_9GlcNAc_2$-PA.

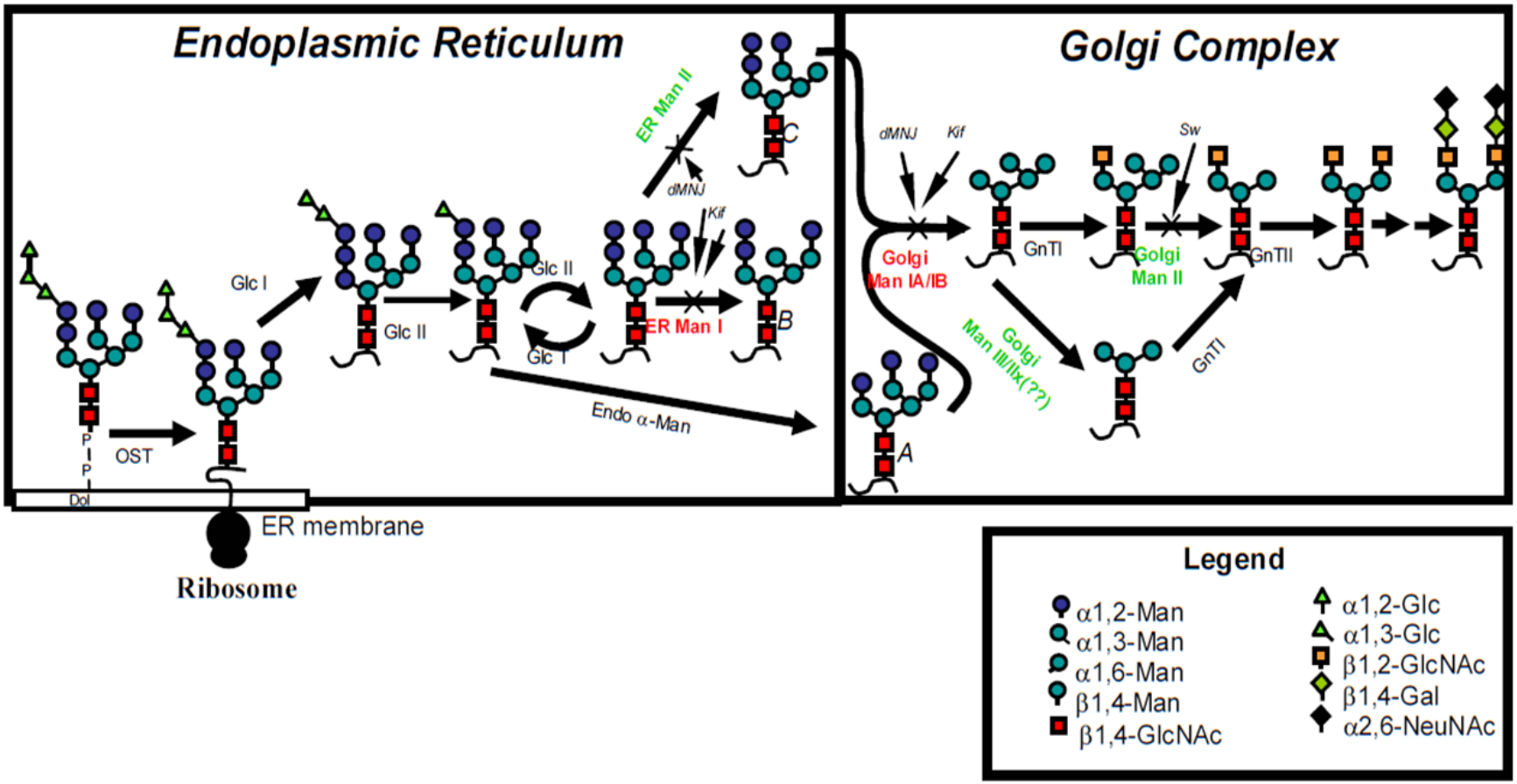
3. Inhibitor studies

The synthesized derivatives of both dMNJ and Kif were used to test their inhibition potency on both ERManI and GMIA in order to find those with selectivity for one enzyme but against the other. The structures of the compounds displaying potential selective inhibition activity were analyzed and compared to draw clues for the designing of more effective and selective inhibitors in the future.

In this dissertation, the combination of crystallography, mutation, enzymatic kinetic analysis, and substrate binding studies will reveal more insights into the structural basis for the substrate specificities of ERManI and GMIA. The findings from this study, as a part of our continuous exploration of the GH α1,2-mannosidases, will not only extend the spectrum of our knowledge and understanding of the enzyme family, but also provide a foundation for further studies.

Figure 1: The biosynthesis of glycoproteins in mammalian cells. The dolichol pyrophosphate oligosaccharide precursor of *N*-linked glycans is transferred to a newly synthesized polypeptide by the oligosaccharyl transferase (OST) as the peptide passes through the ER membrane. The oligosaccharide is then subjected to the modification by a series of enzymes in the ER: glucosidase I (Glc I); glucosidase II (Glc II); UDP-Glc:glycoprotein glucosyltransferase (Glc T); ER mannosidase I (ERManI); ER mannosidase II (ERManII); endo α-mannosidase (Endo α-man). The correctly folded glycoproteins are transferred to the Golgi complex to be processed further by a series of the Golgi enzymes: Golgi mannosidase IA and IB (Golgi ManIA/B); GlcNAc transferase I (GnTI); and either α-mannosidase III or Golgi mannosidase IIx(Golgi Man III/IIx). The *N*-linked glycans are modified to the complex glycan structures by a series of glycosyl transferase in the Golgi complex.

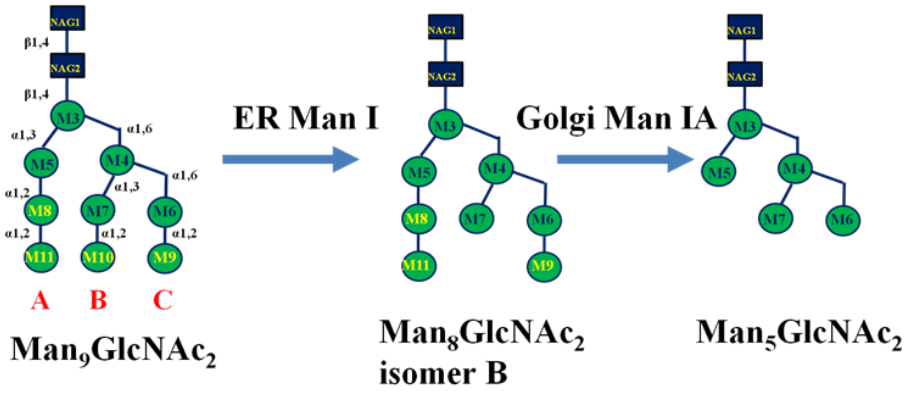
The glycosidases and glycosyltransferases involved in glycoprotein maturation and degradation are indicated by a bold solid arrow with the following enzyme abbreviations: Glc I, glucosidase I; Glc II, glucosidase II; ER Man I, ER mannosidase I; ER Man II, ER mannosidase II (the ER form of the cytosolic/ER mannosidase II); Golgi Man IA, Man IB, Man IC, Golgi mannosidases IA, IB, and IC; GnTI, GlcNAc transferase I; Golgi Man II, Golgi mannosidase II; GnTII, GlcNAc transferase II Positions where processing inhibitors can act to block enzyme reactions are indicated by a thin arrow with the following abbreviations: 1-deoxymannojirimycin (dMNJ); kifunensine (Kif); swainsonine (Sw). The legend for the oligosaccharide structure displayed in the figure is indicated in the lower right.



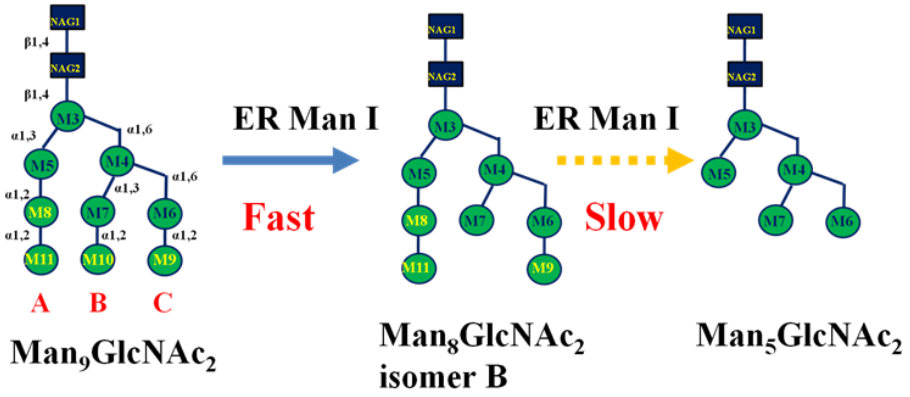
Modified from Mormen, K.W.(2000) in *A comprehensive Handbook: Carbohydrate in Chemistry and Biology*, P.81-117

Figure 2. Schematic diagram of the $Man_9GlcNAc_2$ processing steps catalyzed by ER mannosidase I and Golgi mannosidase IA *in vivo* and *in vitro.* *In vitro,* $Man_9GlcNAc_2$was efficiently processed by ERManI to $Man_8GlcNAc_2$ isomer B, and then slowly down to $Man_5GlcNAc_2$. Trimming of $Man_9GlcNAc_2$ by the Golgi ManI subfamily of enzymes generates $Man_6GlcNAc_2$ efficiently, but cleavage to generate $Man_5GlcNAc_2$ occurs slowly. *In vivo*, ERManI cleaved $Man_9GlcNAc_2$ to produce $Man_8GlcNAc_2$ isomer B, followed by the Golgi enzymes to produce $Man_5GlcNAc_2$. The GlcNAc residues are indicated by red shaded squares. Mannose residues are green shaded circles. The identities of mannose residues are labeled with capital letter M followed by numbers (M3-M11). The GlcNAc residues were labeled with NAG1 and NAG2. The three branches of the $Man_9GlcNAc_2$ are labeled with letter A, B, and C.

In vivo



In vitro by ER Man I



In vitro by Golgi Man IA

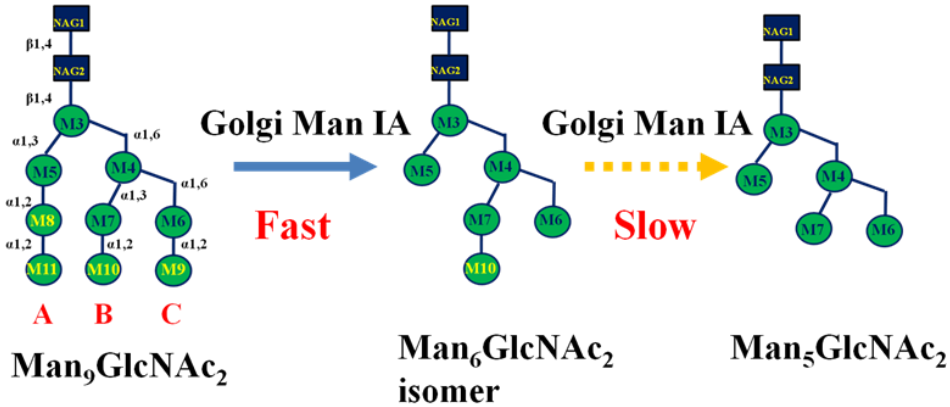


Figure 3: **The development of structural studies of GH47 α1,2-mannosidases.** The history of structural studies on *Saccharomyces cerevisiae* ER mannosidase I, human ERManI, *Tricoderma reesi* α1,2-mannosidase, *Penicillium citrinum* α1,2-mannosidase**,** mouse GMIA and the bacterial GH47 α1,2-mannosidase are presented in the flow chart in the order of year that the structures were published. The PDB identity codes of each structure were listed in the chart in blue color. The ligands (inhibitors, disaccharides, complex glycans) co-crystalized with the enzyme structures and the conformations in the -1 subsites of the enzymes are shown on the right side of the chart. The structures were denoted into four groups by the nature of ligands bound in the enzymes: enzyme-product complex, enzyme-inhibitor complex, enzyme-disaccharide complex and enzyme-substrate complex.

The Development of Crystal Structure Studies of GH47 Mannosidase

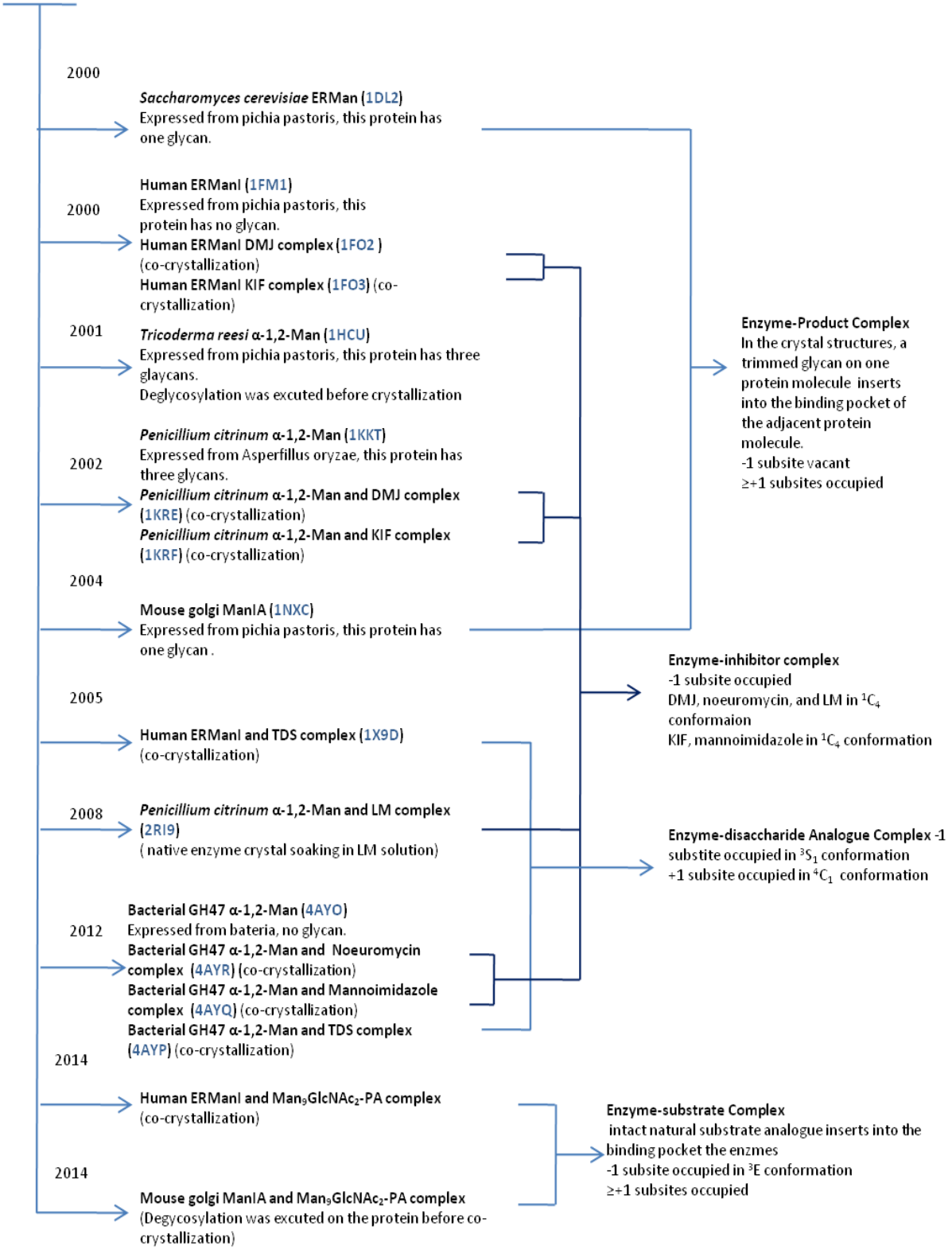
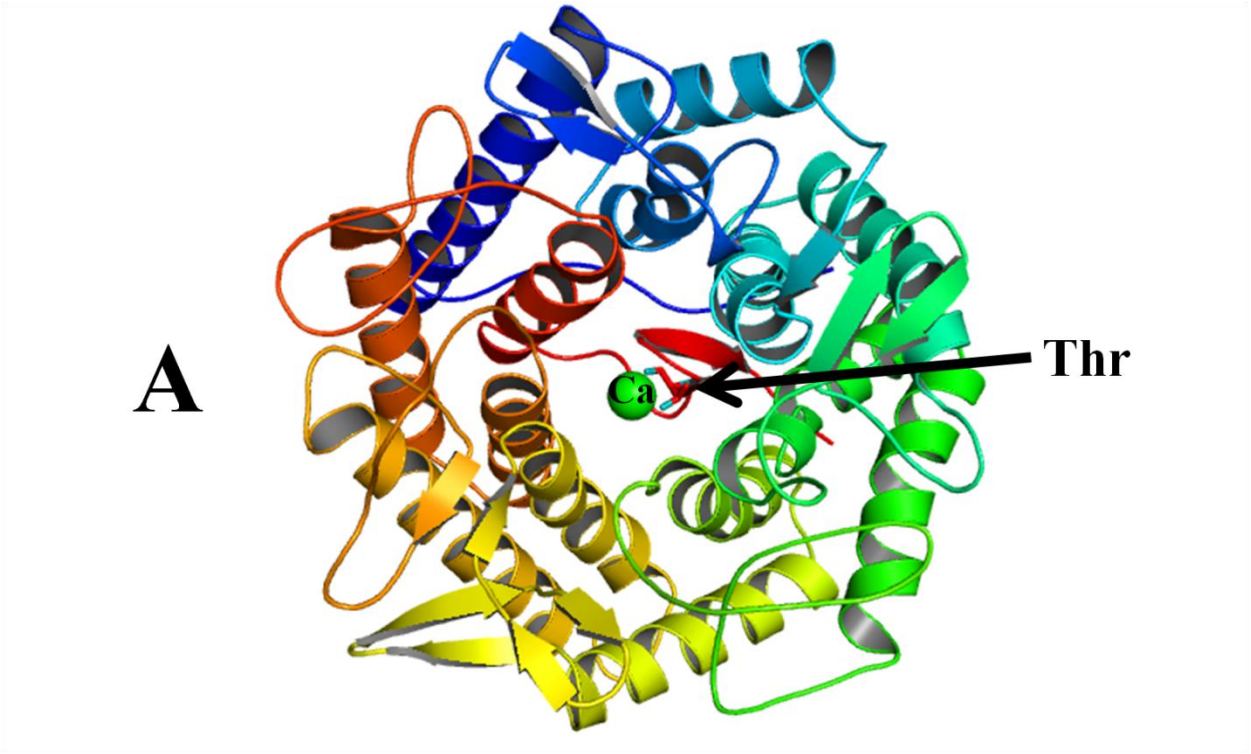


Figure 4: **The general structure of GH47 α1,2-mannosidases:** The crystal structures of GH47 α1,2-mannosidases consist of an $(\alpha\alpha)_7$ barrel structure comprised of two layers of concentric inner and outer helices connecting by loops and β-sheets. The catalytic site is located at the bottom of the barrel. One end of the barrel is plugged with a β-hairpin containing a Thr residue in the core of the barrel. The Thr residue was involved in coordination of the Ca^{2+} ion. The opposite end of the barrel consists of an open conical cleft leading to a more constricted pocket in the core of the barrel. The cartoon representation was generated with Pymol software using the Golgi mannosidase IA structure from Protein Data Bank (1NXC). The Ca^{2+} ion is represented by the green colored sphere. The Thr residue is represented in stick form (with carbon in red color, and oxygen in light blue). Figure A presents the view from the top of the barrel and Figure B presents the view from the side of the barrel.


A
Ca
Thr

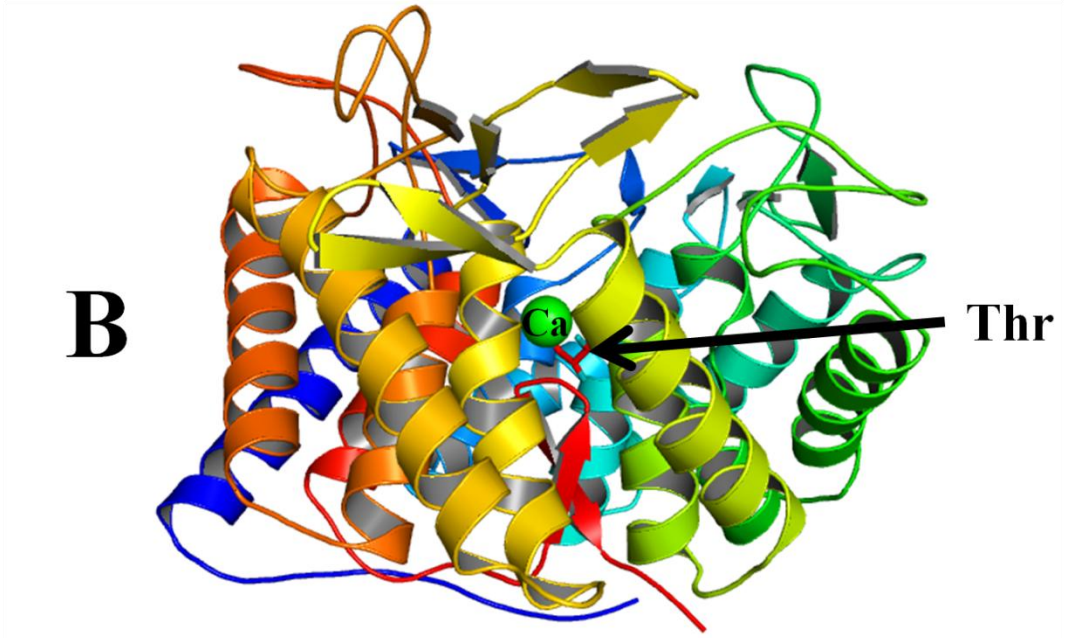

B
Ca
Thr

Figure 5: Comparison of the general structures of GH47 α1,2-mannosidases: The general protein folding of GH47 α1,2-mannosidases are similar. They all consist of an $(\alpha\alpha)_7$ barrel structure comprised of two layers of concentric inner and outer helices connected by loops and β-sheets. The Ca^{2+} ion is bound in the core of the barrel. The alignments of the structures of *Saccharomyces cerevisiae* ERManI (presented in yellow color), human ERManI (red) and mouse GMIA (blue) are shown in cartoon presentation (A) and ribbon presentation (B). The figures in panels A and B were generated with Pymol software using structures of *Saccharomyces cerevisiae* ERManI (1X9D), human ERManI (1DL2) and mouse GMIA (1NXC) in the Protein Data Bank.

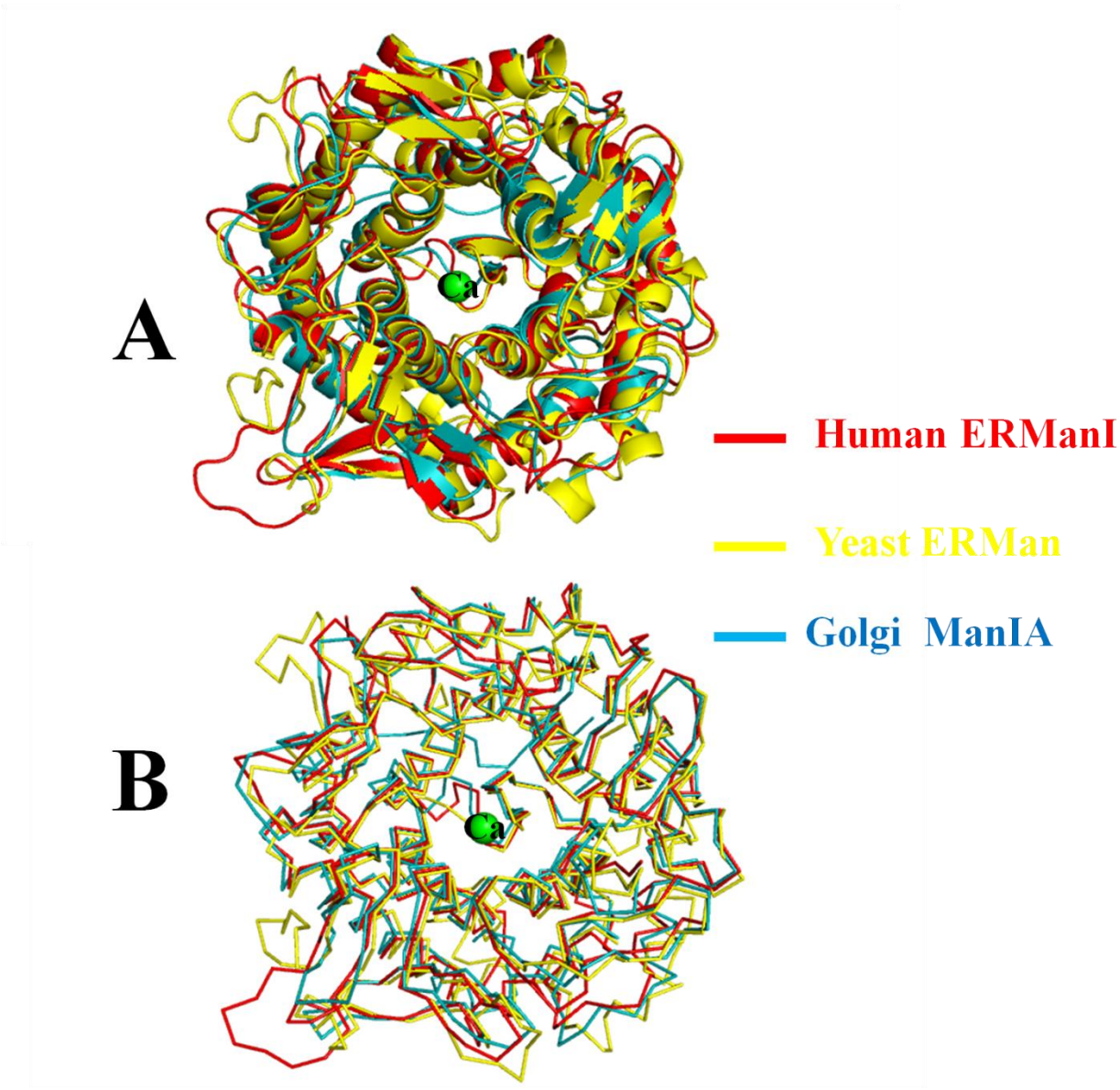
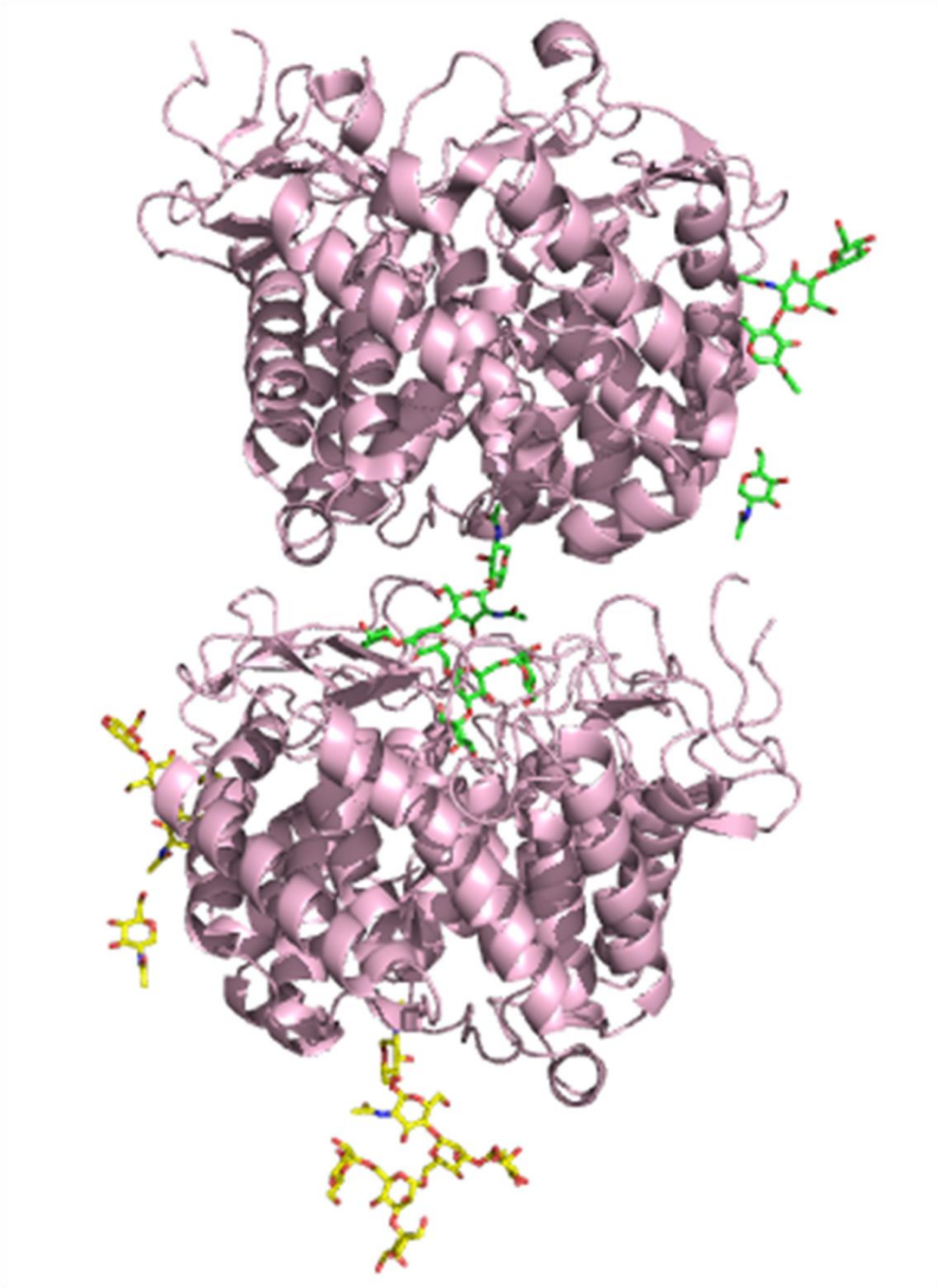

A
B
Ca
Ca
Human ERManI
Yeast ERMan
Golgi ManIA

Figure 6: **The crystal structure of *Saccharomyces cerevisiae* ERManI represents a complex of enzyme and product**: In the crystal lattice of the *S. cerevisiae* ERManI, two adjacent protein molecules A and B interact with each other with molecule A inserting one of its three glycans into the active site cavity of molecule B. The glycans inside the enzyme cavities were found to be $Man_5GlcNAc_2$ structures. The two protein molecules are represented with pale purple colored cartoon. The glycans are presented in sticks and colored in green (for molecule A) and yellow (for molecule B). The figure was generated with Pymol software using the *Saccharomyces cerevisiae* ERManI structure in the Protein Data Bank (1DL2).

Molecule A



Molecule B

Figure 7: **The glycan complex in the active site of *Saccharomyces cerevisiae* ERManI**. The glycan structure in the *S. cerevisiae* ERManI active site binds with the -1 subsite vacant and ≥+1 subsites occupied (A). The glycan was rendered in green sticks, the Ca^{2+} ion in green sphere, and the protein in light purple surface. The structure of the $Man_5GlcNAc_2$ glycan in the active site of *S. cerevisiae* is shown in panel B. The mannose residues are labeled as M3-M7, and GlcNAc residues were labeled as NAG1and NAG2. The M7 residue was found in the +1 subsite. The unoccupied -1 subsite, and occupied +1 subsite were highlighted in red circles. Panel C: The diagram of $Man_5GlcNAc_2$ is labeled as branches A, B and C. The dashed circle and lines represent the mannose residues and the α1,2-Man linkages of $Man_9GlcNAc_2$ that have been cleaved. The figure was generated with Pymol software using the *Saccharomyces cerevisiae* ERManI structure in the Protein Data Bank (1DL2).

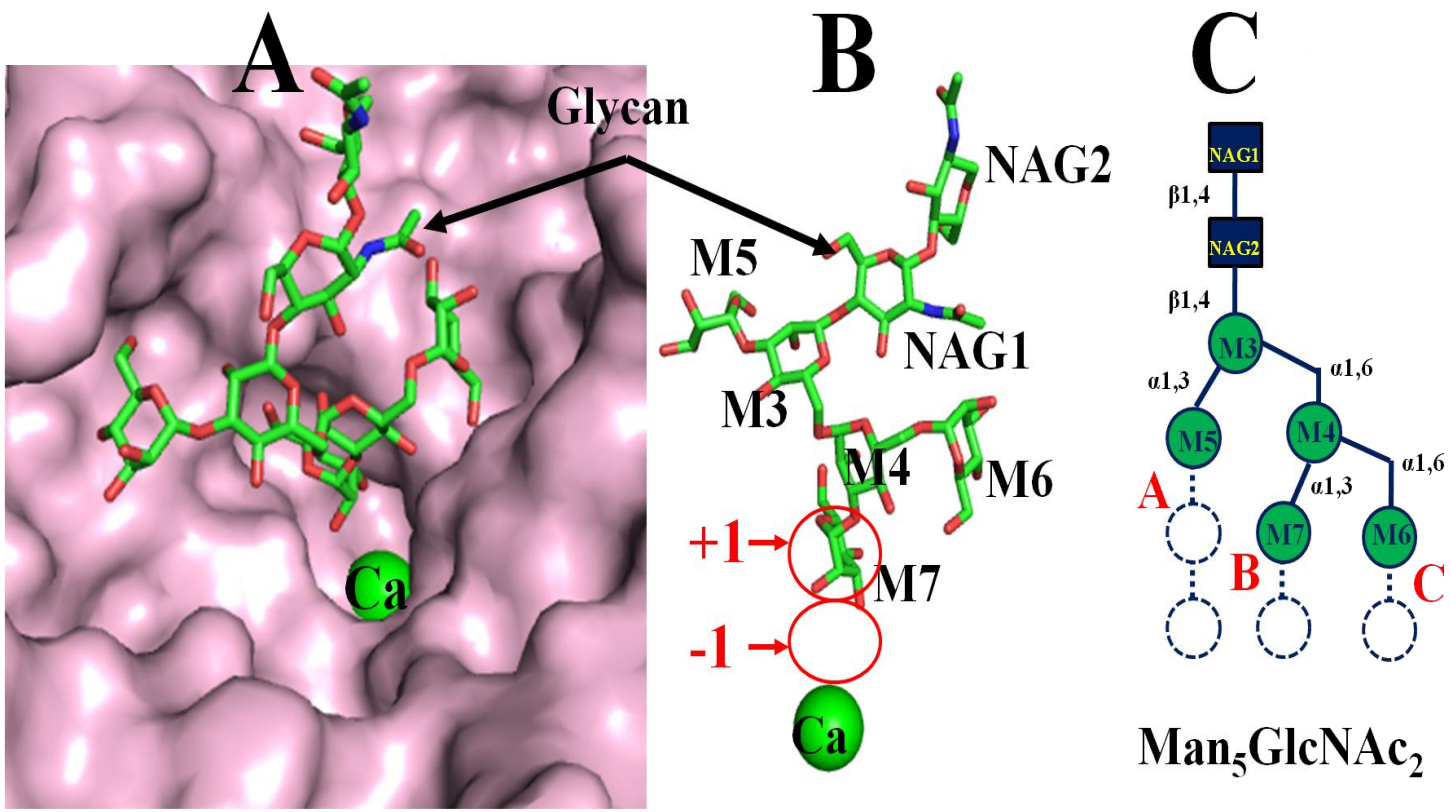
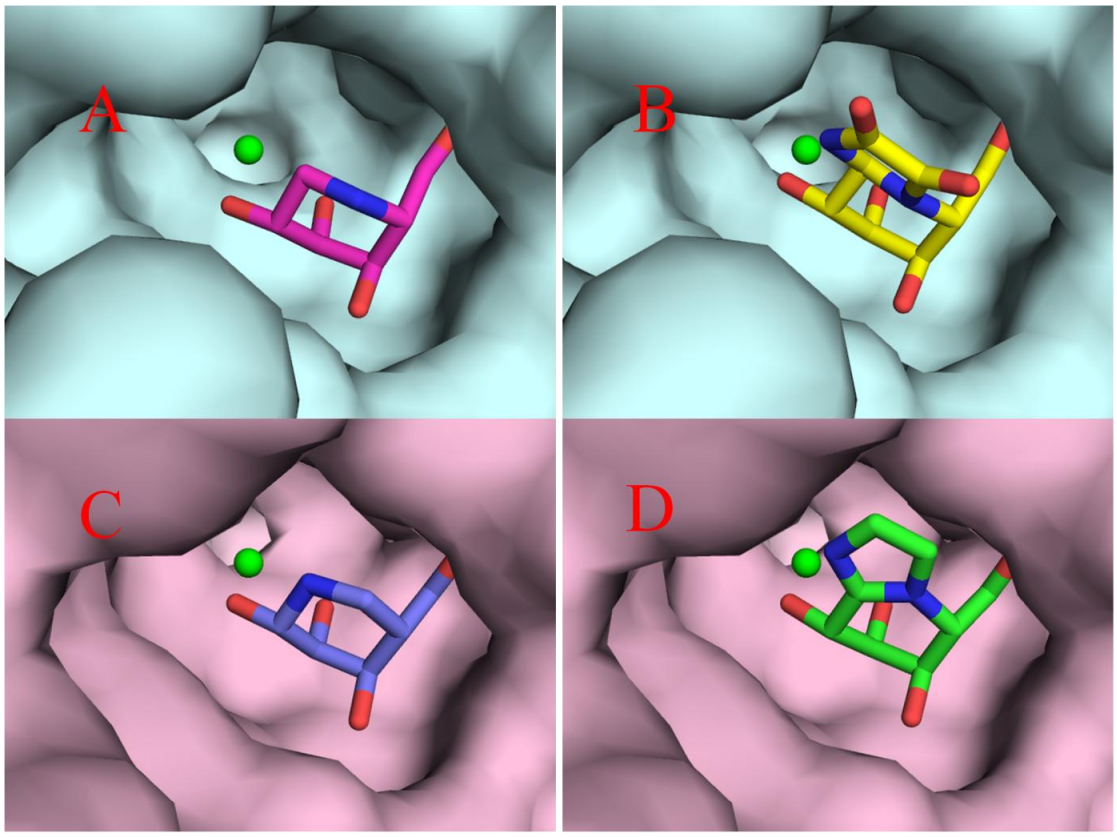
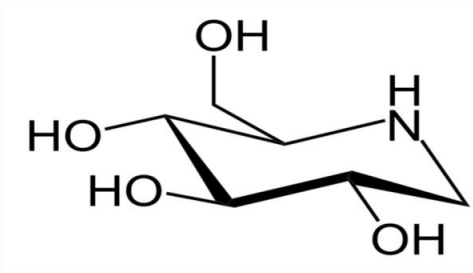
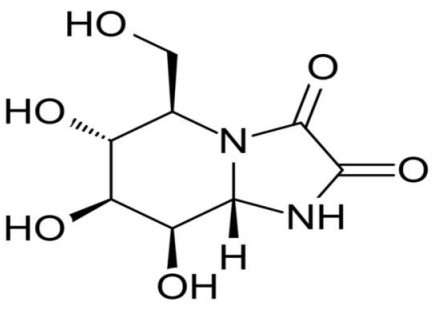

A
Glycan
Ca
B
NAG2
M5
NAG1
M3
M4
M6
+1
M7
-1
Ca
C
NAG1
β1,4
NAG2
β1,4
M3
α1,3
α1,6
M5
M4
α1,6
α1,3
A
M7
M6
B
C
Man5GlcNAc2

Figure 8: Inhibitor bound in the enzyme clefts of human ERManI and bacterial GH47 mannosidase. Panel A: dMNJ bound in human ERManI; Panel B: Kif bound in human ERManI; Panel C: noeuromycin bound in Bacterial GH47 mannosidase; Panel D: mannoimidazole bound in bacterial GH47 mannosidase. The active site pocket of human ERManI was rendered in surface presentation and colored in blue in Panels A and B. The cleft of bacterial GH47 mannosidase was rendered in surface presentation and colored in pink in Panels C and D. The inhibitors are rendered in stick representation and colored in purple for dMNJ, yellow for Kif, blue for noeuromycin, and green for mannoimidazole. Panels A, B, C, D were generated with Pymol software using the entries, 1FO2, 1FO3, 4AYR, and 4AYQ, in Protein Data Bank. The chemical structures of the four inhibitors were shown on the bottom of the figure.

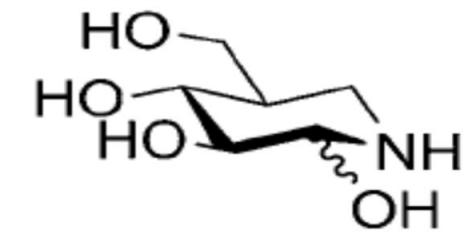




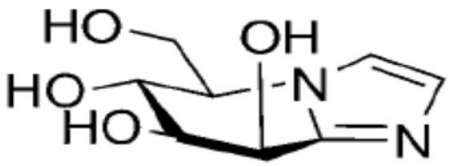
dMNJ



Kif



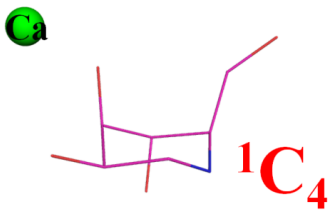
Noeuromycin



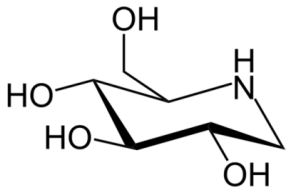
Mannoimidazole

Figure 9: Conformations of inhibitors in the -1 subsites of GH47 α1,2-mannoidases. Panel A: dMNJ in $^{1}C_{4}$ conformation; Panel B: Kif in $^{1}C_{4}$ conformation; Panel C: noeuromycin in $^{1}C_{4}$ conformation; Panel D: mannoimidazole in $^{3}H_{4}$ conformation. The inhibitors are rendered in line presentation and colored in purple for dMNJ, yellow for Kif, blue for noeuromycin, and green for mannoimidazole. Panels A, B, C, D were generated with Pymol software using the entries, 1FO2, 1FO3, 4AYR, and 4AYQ, in Protein Data Bank. The corresponding chemical structures of each inhibitor are shown on the right side of the figure.

A

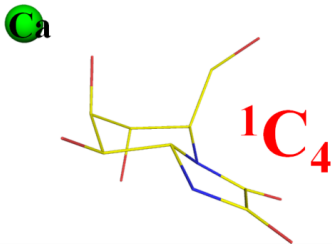


dMNJ

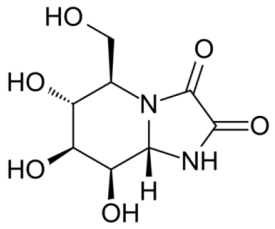


dMNJ

B

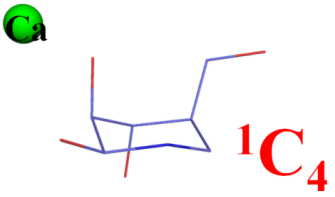


Kif

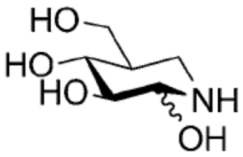


Kif

C

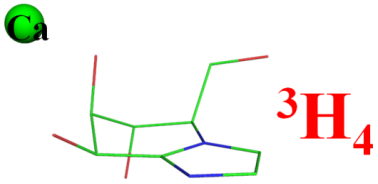


Noeuromycin

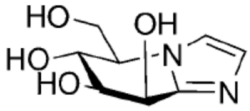


Noeuromycin

D



Mannoimidazole



Mannoimidazole

Figure 10: The structures of bound disaccharide analogs in complex with GH47 α1,2-mannosidases. Panel A: S-Man_2 in human ERManI; Panel B: LM in *Penicillium citrinum* α1,2-mannosidase; Panel C: the non-reducing end of S-Man_2 in the -1 subsite of ERManI in a $^{3}S_{1}$ conformation. Panel D: the non-reducing end of LM in the -1 subsite of *Penicillium citrinum* α1,2-mannosidase in a $^{1}C_{4}$ conformation. Panel E: the chemical structure of S-$Man_{2;}$ Panel F: the chemical structure of LM. The enzyme clefts were rendered in surface presentation with ERManI in light blue and *Penicillium citrinum* α1,2-mannosidase in pink. The disaccharides were rendered in stick representation with S-Man_2 in yellow, and LM in blue. Panels A-D were generated with Pymol software using the entries 1X9D and 2RI9 from Protein Data Bank.

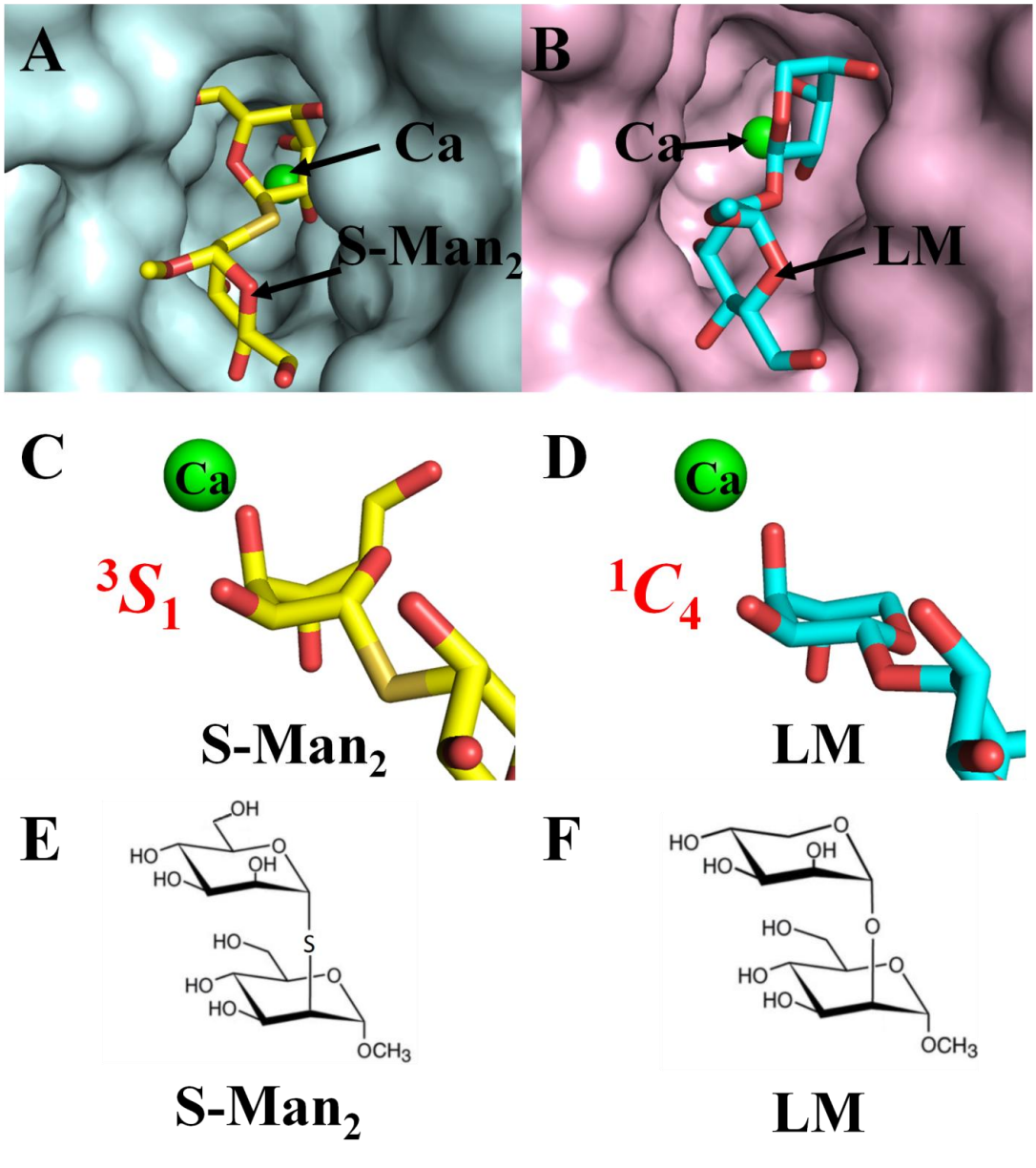

A
Ca
S-Man2
B
Ca
LM
C
Ca
3S1
S-Man2
D
Ca
1C4
LM
E
OH
HO
HO
O
OH
HO
S
HO
HO
O
HO
OCH3
S-Man2
F
HO
HO
O
OH
HO
O
HO
HO
O
HO
OCH3
LM

Figure 11: Golgi α1,2-mannosidase IA crystal structure presented as an enzyme-product complex. Molecule A inserts its glycan (Glycan A) into the cleft of the adjacent Molecule B. The glycans in the molecule active sites were found to be $Man_6GlcNAc_2$ structures. The protein was rendered in cartoon presentation, and the glycan in stick presentation. The figure was generated with Pymol software using the entry 1NXC from Protein Data Base.

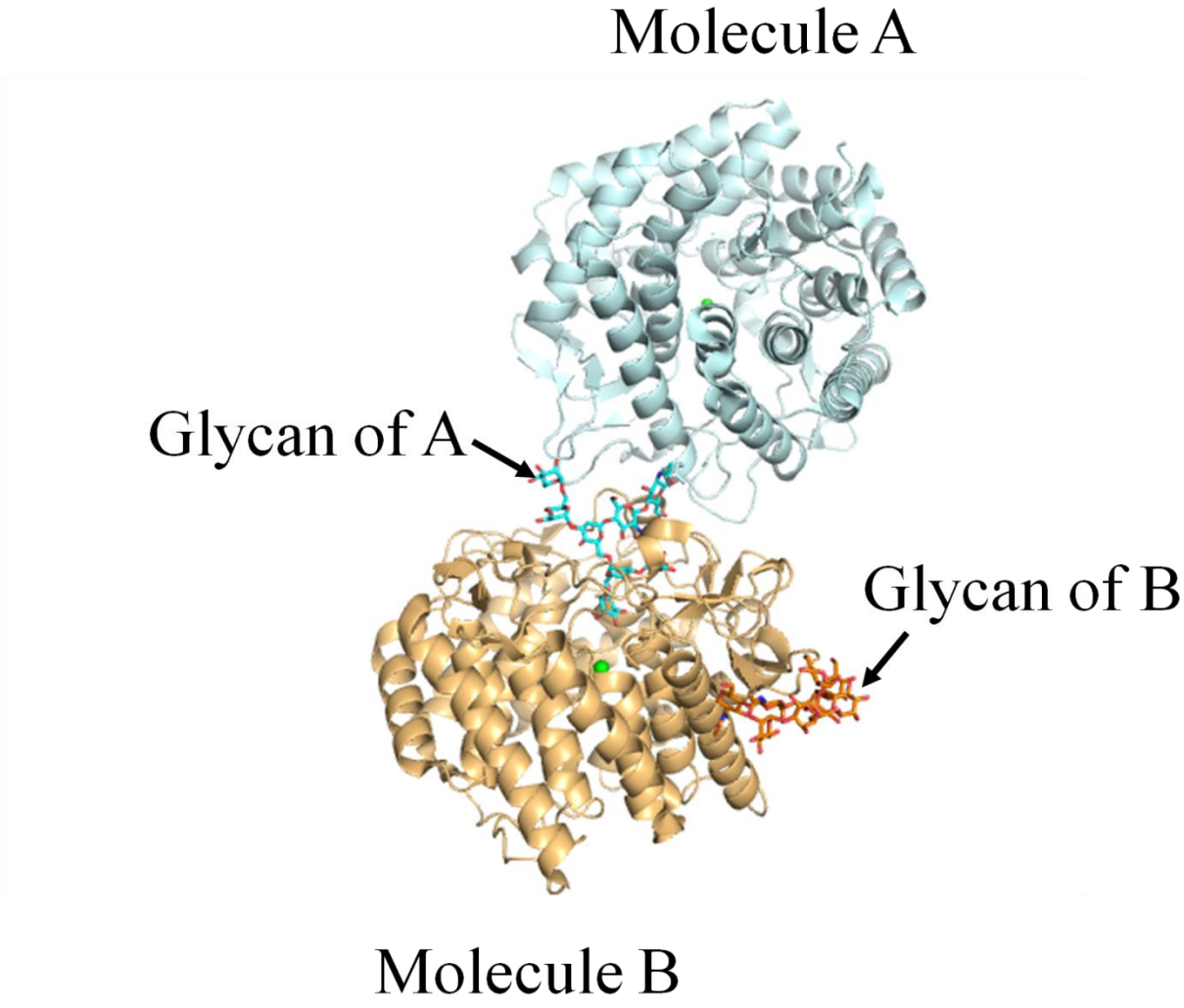

Molecule A
Glycan of A
Glycan of B
Molecule B

Figure 12: The $Man_6GlcNAc_2$ glycan structure in the cleft of GMIA. Panel A: The glycan bound in the active site cleft of mouse GMIA with the -1 subsite vacant and ≥+1 subsites occupied. The glycan was rendered in blue sticks, the Ca^{2+} ion in green sphere, and the protein in light brown surface. Panel B: Diagrammatic representation of the $Man_6GlcNAc_2$ glycan in the mouse GMIA active site. The mannose residues were labeled as M3-M7 and *Man* for the extra mannose residue added by the *P. pastoris* expression host. The GlcNAc residues were labeled as NAG1and NAG2. The mannose residue M6 was found in the +1 subsite. The unoccupied -1 subsite and the occupied +1 subsite were highlighted in red circles. C: The diagram of $Man_5GlcNAc_2$ with branches labeled as A, B and C. The dashed circle and lines represent the mannose residues and the α1,2-Man linkages of $Man_9GlcNAc_2$ that had been cleaved. The figure was generated with Pymol software using the entry 1NXC in Protein Data Bank.

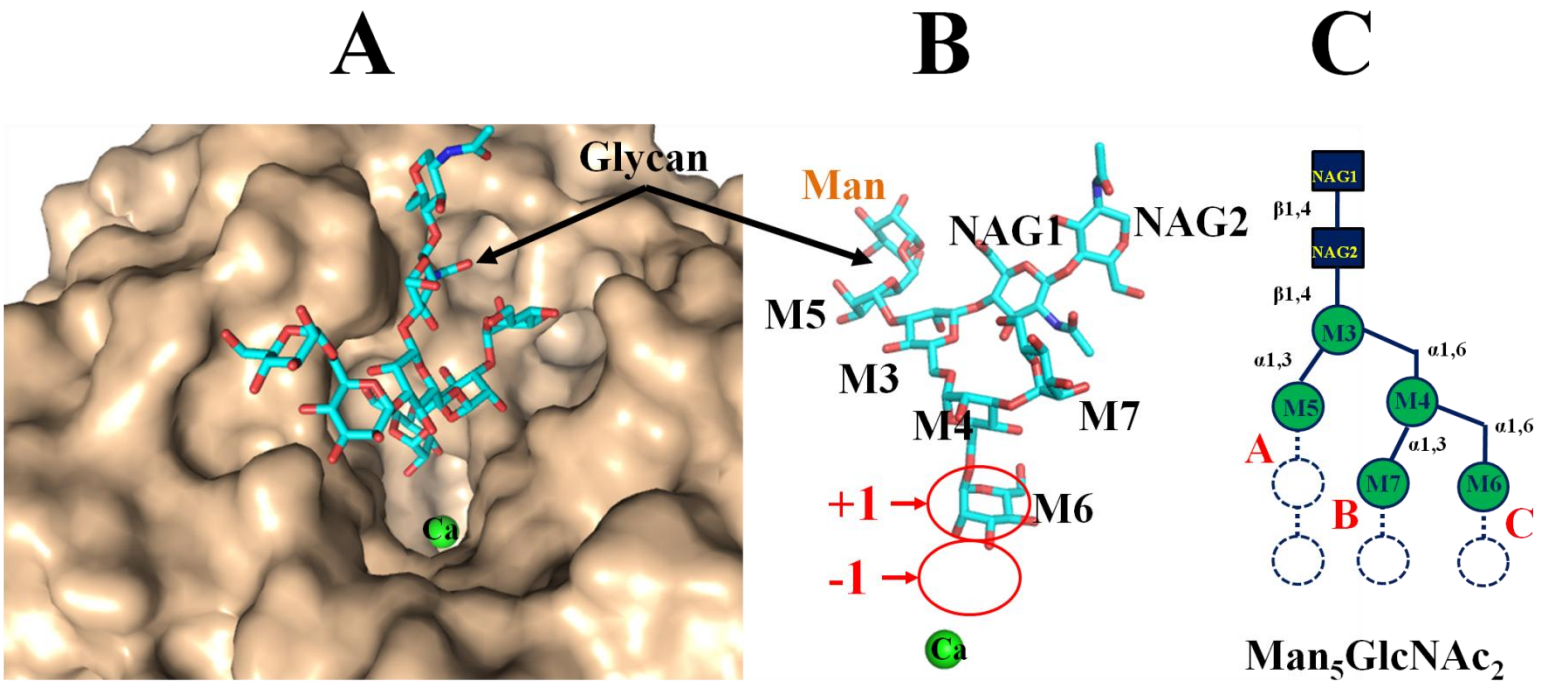
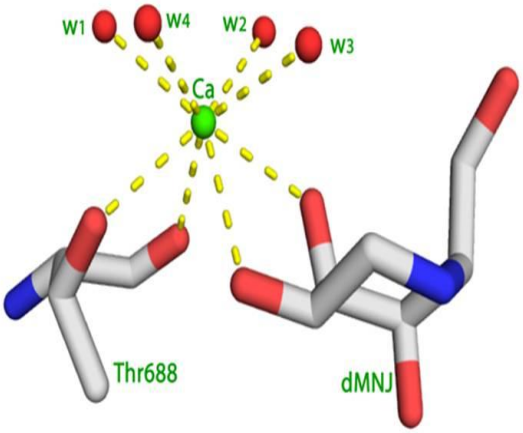

A
B
C
Glycan
Ca
Man
NAG1
NAG2
M5
M3
M4
M7
+1
M6
-1
Ca
NAG1
β1,4
NAG2
β1,4
M3
α1,3
α1,6
M5
M4
α1,6
α1,3
A
M7
M6
B
C
Man5GlcNAc2

Figure 13: The Ca^{2+} coordination in GH47 α1,2-mannosidases. The Ca^{2+}-centered 8-fold square antiprismatic coordination geometry has been found in all the solved structures of wild type GH47 α1,2-mannosidases. In the ERManI-dMNJ complex (left), oxygen atoms from four water molecules (W1, W2, W3, W4) comprised half of the coordination points. The O-δ and the backbone carboxyl oxygen of Thr 688, together with oxygens associated with the C-2 and C-3 hydroxyls of dMNJ comprised the other four points of coordination. In mouse GMIA (right), the -1 subsite was unoccupied and W5 and W6 coordinated with the Ca^{2+} ion. The other six coordinating groups were identical to their counterparts in the ERManI structure. The figure was generated with Pymol software using coordinate files 1FO2 and 1NXC from Protein Data Bank.


W1
W4
W2
W3
Ca
Thr688
dMNJ

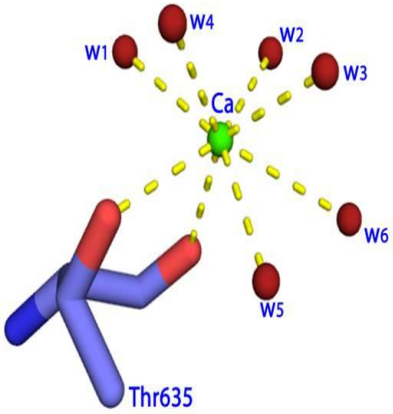
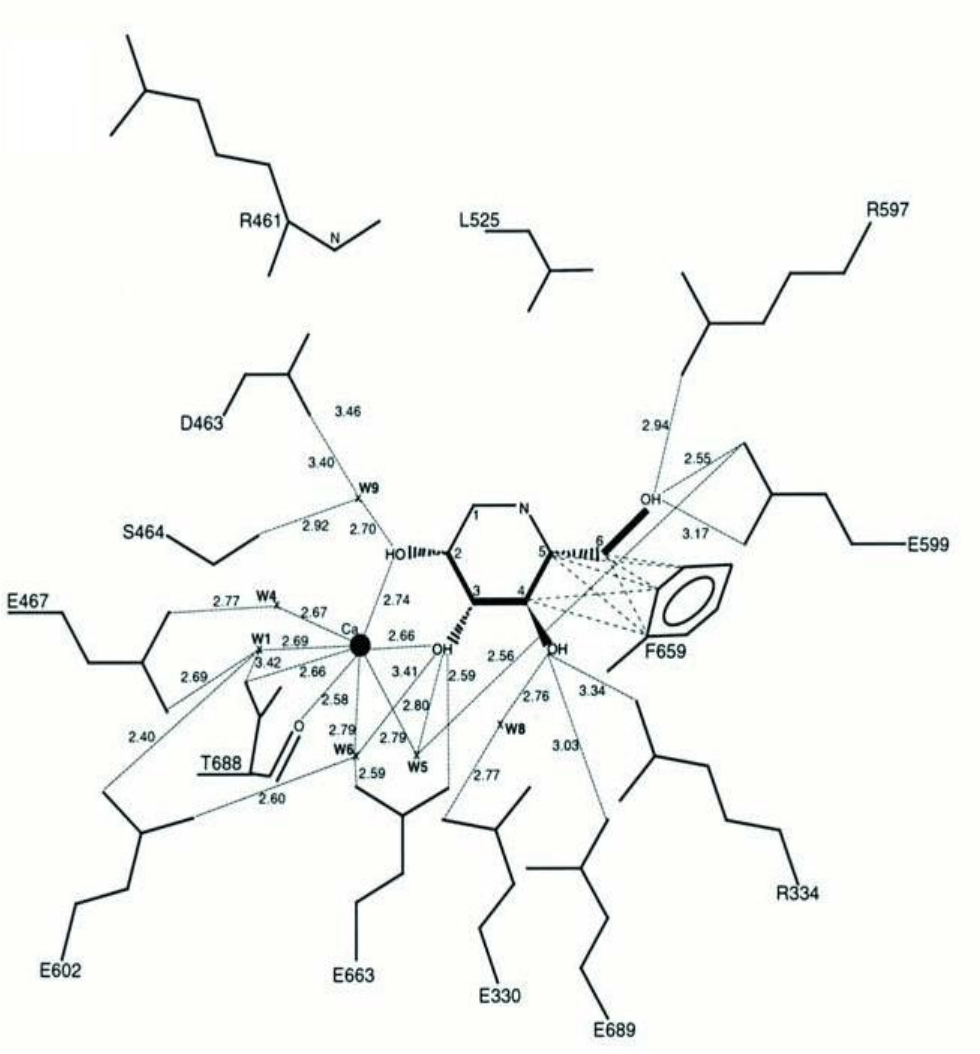

W4
W1
W2
W3
Ca
W6
W5
Thr635

Figure 14: The interactions between human ERManI and dMNJ in the -1 subsite. Short and long dahsed lines represent hydrogen bond interactions and van der Walls contacts, respectively. Only hydrogen bonds between the protein, water, and inhibitor molecules are respresented. Water-water hydrogen bonds are not represented. The amino acid residues are reprented in solid lines the identities are labeled beside each. The distance between the two interacting partners are labeled next to the dashed lines.



From Vallee, F., K. Karaveg, *et al.* (2000) *J.Biol. Chem.* **275**: 41287-98

Figure 15: Comparison of the configurations of atoms of α-D-manopyranose in $^{1}C_{4}$, $^{3}H_{4}$ and $^{3}S_{1}$ Conformers. The three conformers are identical in the configurations at positions C-2, C-3, C-4, and C-6, but different in the configurations of C-1, O5, and C-5. The half part of the six-member ring that is identical among different conformers is clored in green. The other half that distorts among different conformers is colored in black.

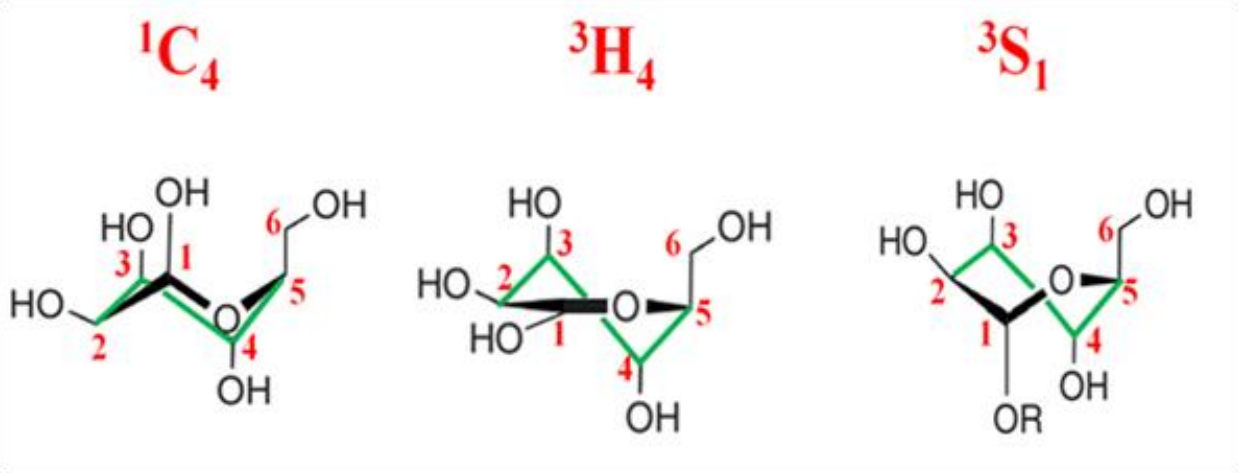
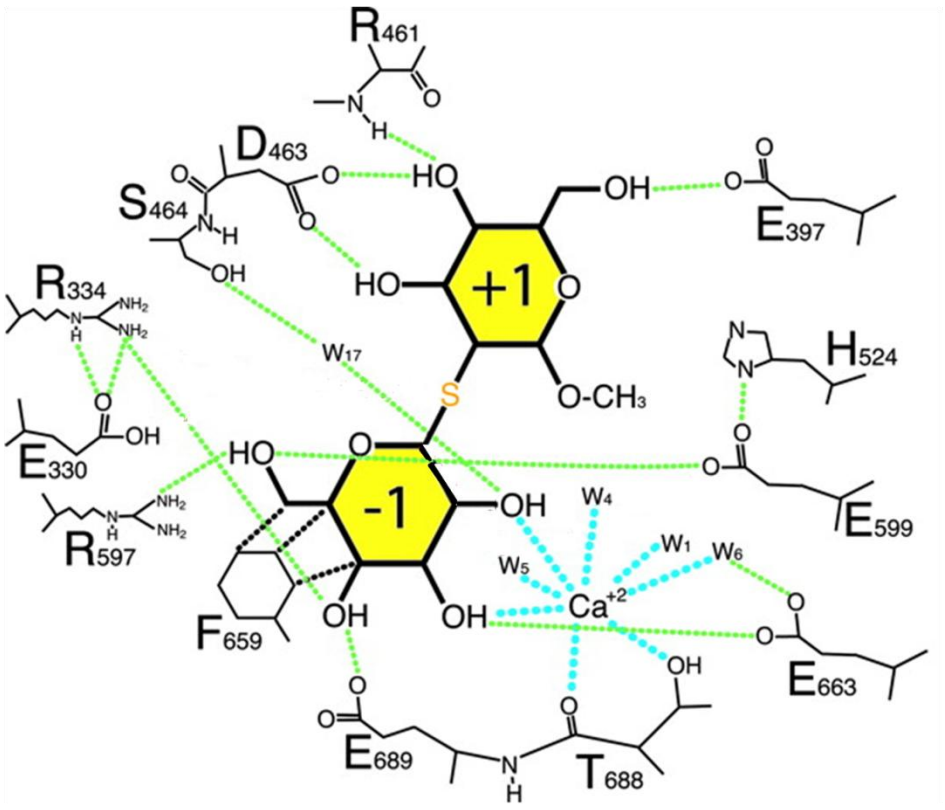
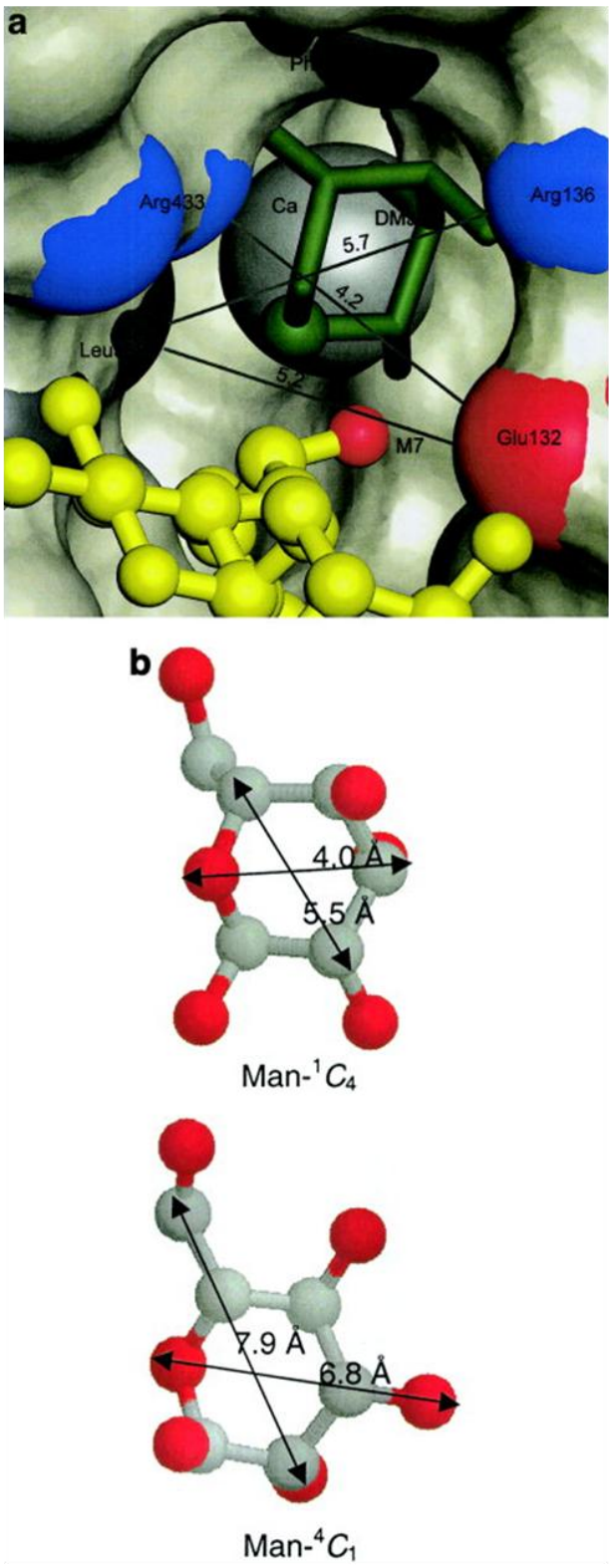

$^{1}C_{4}$
$^{3}H_{4}$
$^{3}S_{1}$
OH
HO
6
OH
3
1
HO
5
O
2
4
OH
HO
3
6
OH
HO
2
O
5
HO
1
4
OH
HO
3
6
OH
HO
2
O
5
1
4
OH
OR

Figure 16: The interactions between human ERManI and S-Man$_2$. Shown is a schematic diagram of the interactions between the thiodisaccharide and ERManI in the –1 and +1 subsites, demonstrating hydrogen bonding interactions (*green dotted lines*), direct coordination of the enzyme-associated Ca^{2+} ion (*blue dotted lines*), hydrophobic stacking of Phe659 with the C-4–C-5–C-6 region of the –1 residue (*black dotted lines*). Residue numbering of amino acid side chains in the respective subsites is indicated.



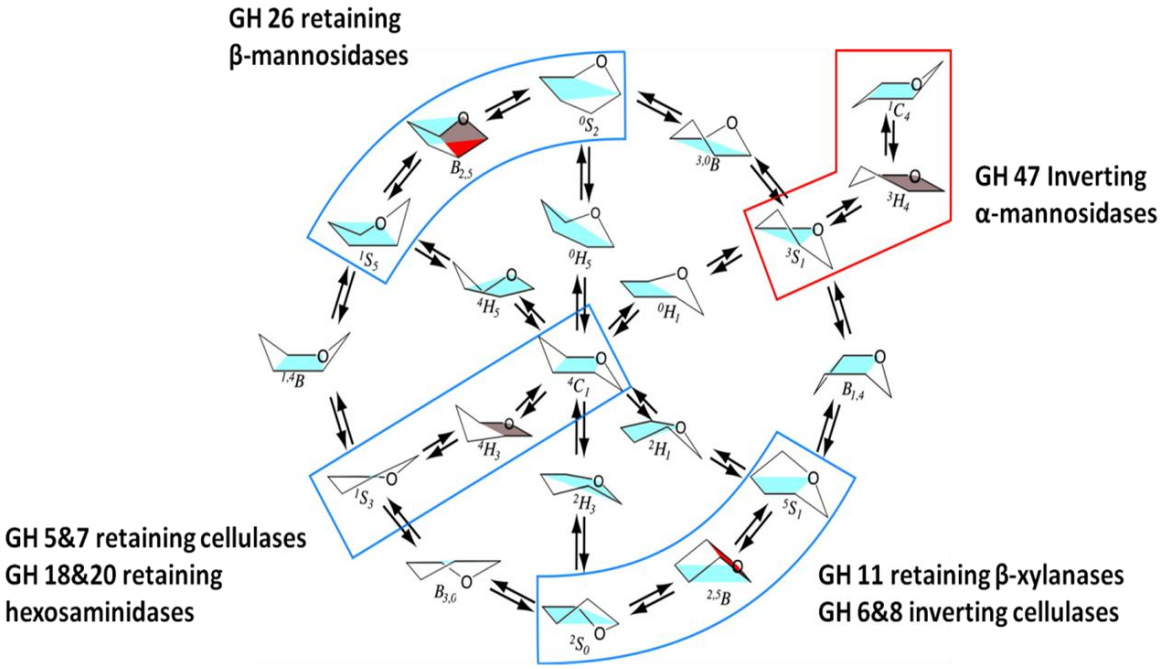
Modified From **Khanita Karaveg et al. J. Biol. Chem. 2005;280:16197-16207**

Figure 17. Comparison of the dimension of the active-site funnel neck of yeast ERManI and the dimensions of Man-$^{1}C_{4}$ and Man-$^{4}C_{1}$. Panel A: Active-site neck dimensions compared to those of Man-$^{1}C_{4}$ and Man-$^{4}C_{1}$ (Panel B). Distances in Å are approximately along the plane of the neck opening. The active-site funnel neck is constricted and is just wide enough for Man-$^{1}C_{4}$ to squeeze through, while Man-$^{4}C_{1}$ is larger and not expected to pass through the constriction. Two salt bridges on either side of the opening, Glu132–Arg136 and Glu399–Arg433, stiffen the opening and reduce its flexibility.



From **Mulakala, C. and P.J. Reilly, *Understanding protein structure-function relationships in Family 47 alpha-1,2-mannosidases through computational docking of ligands.* Proteins, 2002. 49(1): p. 125-34.**

Figure 18: Conformational itinerary of pyranose ring. Pyranose sugar ring conformational interconversions are shown with nomenclature for the reference plane of four ring atoms (shaded blue) as per IUPAC-IUB rules [67]. Routes are noted for the interconversions between conformers from the low free energy $^{4}C_{1}$ chair conformation through various half chair (*H*), boat (*B*), and skew-boat (*S*) conformations. Only four of the conformations that have a co-planarity at C5-O5-C1-C2 are show with red shading in $B_{2,5}$, $^{4}H_{3}$, $^{2,5}B$, and $^{3}H_{4}$ [68]. Overlap of the nomenclature reference plane and the C5-O5-C1-C2 plane is indicated by brown shading. The blue boxed regions illustrate the proposed transition state conformations for $B^{2,5}$, $^{4}H_{3}$, and $^{2,5}B$ bracketed by conformations that have been identified based on the isolation of trapped intermediates in various co-crystal studies with inverting and retaining enzymes. The red boxed region illustrates the potential $^{3}H_{4}$ transition state for GH 47 α-mannosidases bracketed by the $^{4}C_{1}$ chair conformation found for the ERManI-dMNJ complex and the $^{3}S_{1}$ conformation.



From Karaveg, K., et al., *Mechanism of class 1 (glycosylhydrolase family 47α-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. 280(16): p. 16197-207.

Figure 19: General mechanism for inverting glycoside hydrolases. The hydrolysis of glycosidic bonds involves alterations in the conformation of the glycone sugar ring to allow the formation of an oxocarbenium ion followed by a bond cleavage reaction, which is induced by two carboxylate side chains and a water molecule in the enzyme active site. The inverting mechanism is carried out in a single step by protonation of the glycosidic oxygen via a carboxyl group of the general acid catalyst and an attack of the C-1 position by a nucleophilic water residue which is activated by the carboxylate side chain of the general base catalyst. The inverting mechanism results in the formation of the inverted configuration of the C-1 atom. (The figure was generated by ChemDraw software)

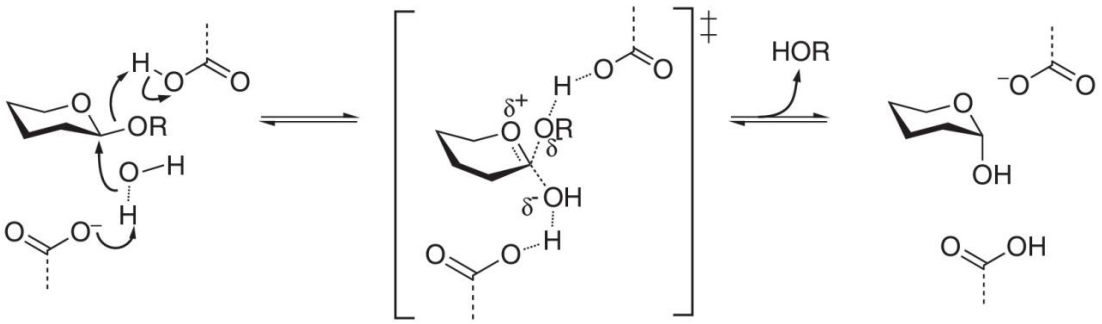
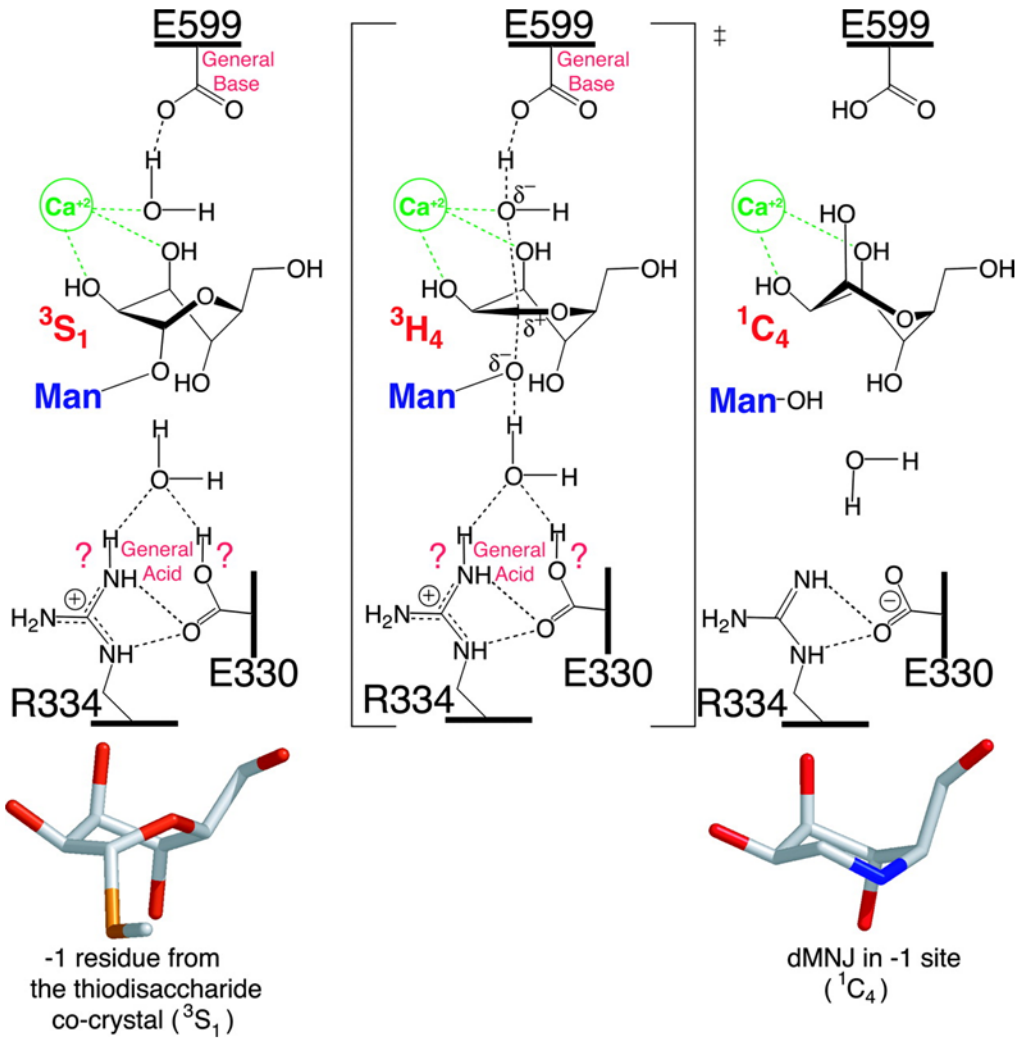

H
O
O
O
OR
O
H
H
O
O⁻
δ+
O
OR
δ
δ⁻
OH
H
O
O
H
O
O
‡
HOR
⁻O
O
O
OH
O
OH

Figure 20: Proposed mechanism for human ERManI: The model for the mechanism of ERManI catalyzed hydrolysis of Manα1,2Man linkages based on kinetic analysis and substrate analog co-crystal structure**.** Substrate binding to the –1 subsite results in distortion of the glycone into a $^{3}S_{1}$ skew boat conformation by binding interactions in both the –1 and +1 subsites. E330 acts as the general acid in a through-water protonation of the glycosidic oxygen leading to the formation of the $^{3}H_{4}$ oxocarbenium ion transition state with ring planarity at C2-C1-O5-C5. Base-catalyzed (E599) proton abstraction from the water nucleophile (W3) results in the attack of the anomeric center and inversion of configuration to form a β-hydroxyl on the enzymatic product in a $^{1}C_{4}$ chair conformation.



From **Khanita Karaveg et al. J. Biol. Chem. 2005;280:16197-16207**

Figure 21: The Ca^{2+} coordination in the T668A mutant of human ERManI. The Thr688 plays an important role in the enzyme acitive site by participating in coordination with the essentical Ca^{2+} ion. Wild type ERManI displays a Ca^{2+} centered 8-fold coordination. The T688A mutation of human ERManI altered the metal coordination to a 7-fold geometry due to the lack of the hydoxyl group on the side chain of A688. Panel A: Ca^{2+} coordination in the wild type ERManI as a S-Man_2 complex; Panel B: Ca^{2+} coordination in the complex of ERManI T668A mutant with S-Man_2. Panel C: Superimpostion of Ca^{2+} coordinations in the enzyme and S-Man_2 complexes for both wilde type and T688A mutant. The panel displaying the metal coordination of the wild type enzyme was generated with Pymol software using the entry 1X9D form Protein Data Bank. The metal coordination of the T688A mutant was generated from the unpublished data from Moremen lab.

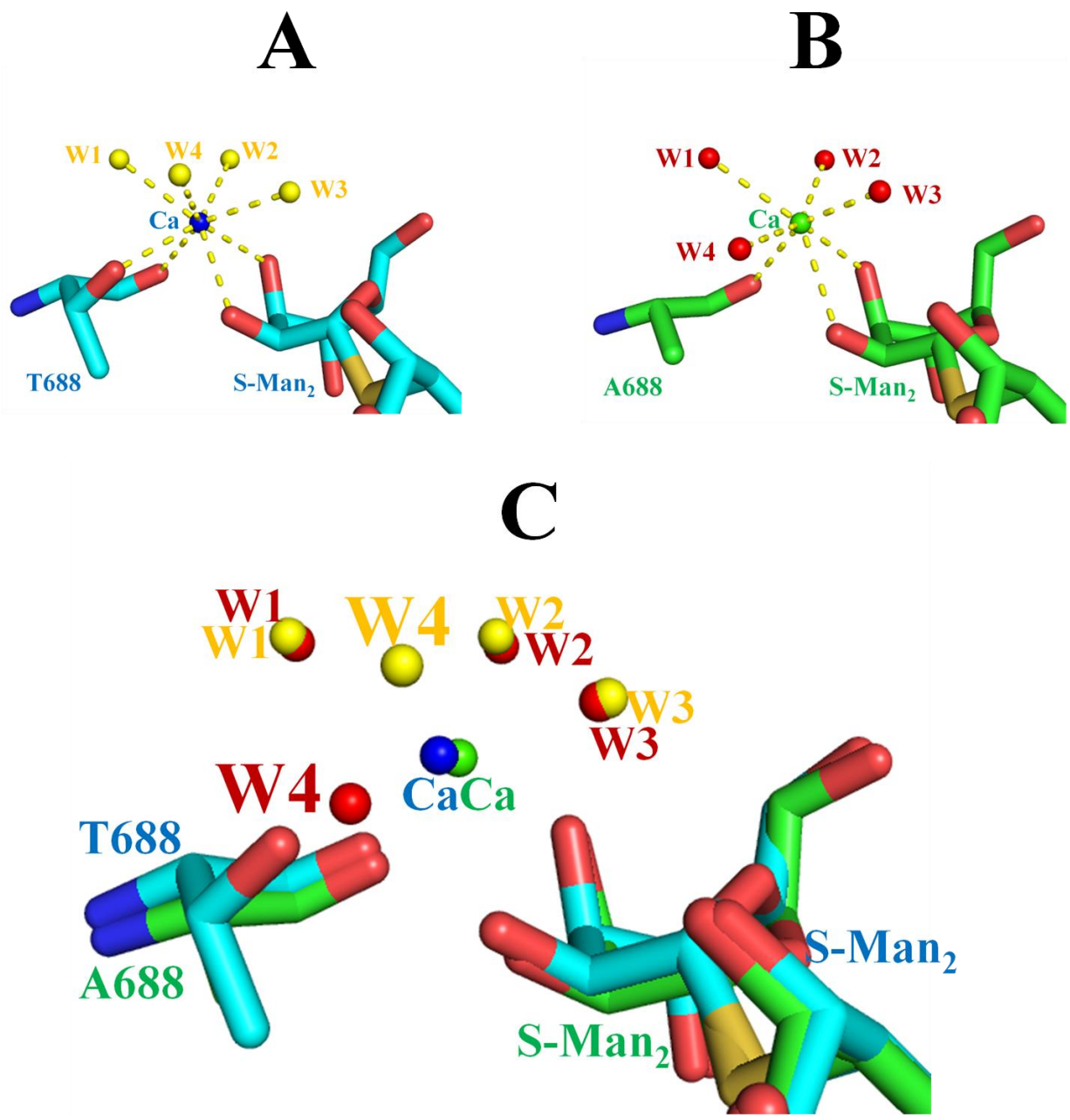
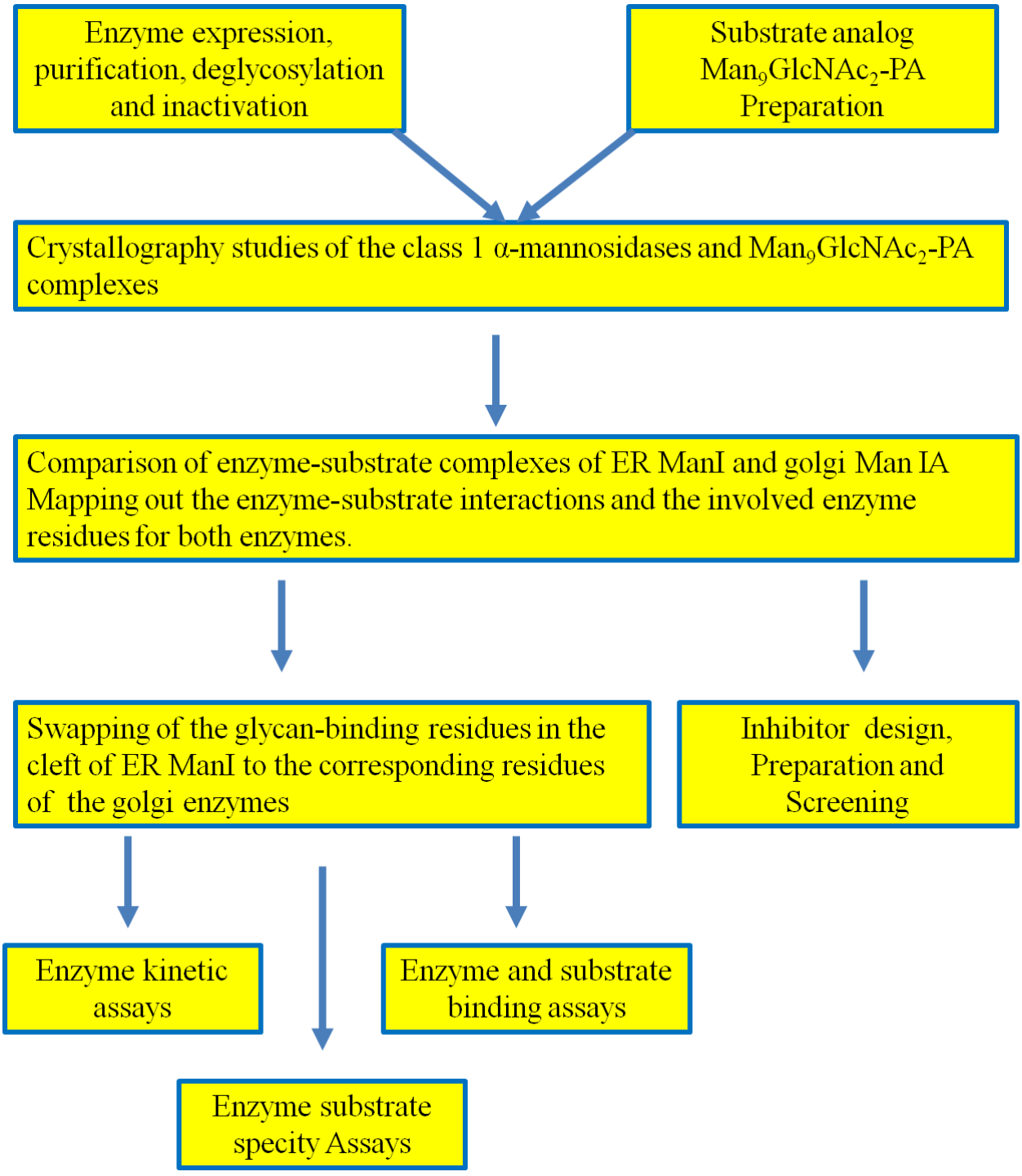

A
W1
W4
W2
W3
Ca
T688
S-Man2
B
W1
W2
W3
Ca
W4
A688
S-Man2
C
W1
W1
W4
W2
W2
W3
W3
W4
CaCa
T688
A688
S-Man2
S-Man2

Figure 22: Outline of the disseration study. The project started with the expression, purification deglycosyaltion and inactivation of Golgi mannosidase IA and the preparation of the substrate analog $Man_9GlcNAc_2$-PA. The enzyme and substrate analog were co-crystalized and the structrue of the complex was determined. The enzyme-substrate complex structures of ERManI and GMIA were compared, and the residues involved in the enzyme-substrate interactions were determined. The mutants containing the glycan-binding residues were swapped between the ER and Golgi enzymes and tested for enzyme kinetics, substrate specificity, and substrate binding studies. The inhibitors designed to have potential selectivities between ERManI and GMIA were screened for their inhibitory activities with the two enzyme preparations.


Enzyme expression, purification, deglycosylation and inactivation
Substrate analog $Man_9GlcNAc_2$-PA Preparation
Crystallography studies of the class 1 α-mannosidases and $Man_9GlcNAc_2$-PA complexes
Comparison of enzyme-substrate complexes of ER ManI and golgi Man IA Mapping out the enzyme-substrate interactions and the involved enzyme residues for both enzymes.
Swapping of the glycan-binding residues in the cleft of ER ManI to the corresponding residues of the golgi enzymes
Inhibitor design, Preparation and Screening
Enzyme kinetic assays
Enzyme and substrate binding assays
Enzyme substrate specity Assays

References

1. Moremen, K., *α-mannosidases in Asparagine-linked oligosaccharide processing and catabolism, Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook.* 2000. **II**: p. 81–117.
2. Henrissat, B. and A. Bairoch, *Updating the sequence-based classification of glycosyl hydrolases.* Biochem J, 1996. **316 (Pt 2)**: p. 695-6.
3. Henrissat, B. and G. Davies, *Structural and sequence-based classification of glycoside hydrolases.* Curr Opin Struct Biol, 1997. **7**(5): p. 637-44.
4. Henrissat, B., *Glycosidase families.* Biochem Soc Trans, 1998. **26**(2): p. 153-6.
5. Bourne, Y. and B. Henrissat, *Glycoside hydrolases and glycosyltransferases: families and functional modules.* Curr Opin Struct Biol, 2001. **11**(5): p. 593-600.
6. Davies, G.J. and B. Henrissat, *Structural enzymology of carbohydrate-active enzymes: implications for the post-genomic era.* Biochem Soc Trans, 2002. **30**(2): p. 291-7.
7. Herscovics, A., *Structure and function of Class I alpha 1,2-mannosidases involved in glycoprotein synthesis and endoplasmic reticulum quality control.* Biochimie, 2001. **83**(8): p. 757-62.
8. Gonzalez, D.S., Y. Kagawa, and K.W. Moremen, *Isolation and characterization of the gene encoding the mouse broad specificity lysosomal alpha-mannosidase1.* Biochim Biophys Acta, 1999. **1445**(1): p. 177-83.
9. Tremblay, L.O. and A. Herscovics, *Cloning and expression of a specific human alpha 1,2-mannosidase that trims Man9GlcNAc2 to Man8GlcNAc2 isomer B during N-glycan biosynthesis.* Glycobiology, 1999. **9**(10): p. 1073-8.

10. Tremblay, L.O. and A. Herscovics, *Characterization of a cDNA encoding a novel human Golgi alpha 1, 2-mannosidase (IC) involved in N-glycan biosynthesis.* J Biol Chem, 2000. **275**(41): p. 31655-60.

11. Lal, A., et al., *Substrate specificities of recombinant murine Golgi alpha1, 2-mannosidases IA and IB and comparison with endoplasmic reticulum and Golgi processing alpha1,2-mannosidases.* Glycobiology, 1998. **8**(10): p. 981-95.

12. Olivari, S. and M. Molinari, *Glycoprotein folding and the role of EDEM1, EDEM2 and EDEM3 in degradation of folding-defective glycoproteins.* FEBS Lett, 2007. **581**(19): p. 3658-64.

13. Hosokawa, N., et al., *A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation.* EMBO Rep, 2001. **2**(5): p. 415-22.

14. Mast, S.W., et al., *Human EDEM2, a novel homolog of family 47 glycosidases, is involved in ER-associated degradation of glycoproteins.* Glycobiology, 2005. **15**(4): p. 421-36.

15. Olivari, S., et al., *A novel stress-induced EDEM variant regulating endoplasmic reticulum-associated glycoprotein degradation.* J Biol Chem, 2005. **280**(4): p. 2424-8.

16. Hirao, K., et al., *EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming.* J Biol Chem, 2006. **281**(14): p. 9650-8.

17. Olivari, S., et al., *EDEM1 regulates ER-associated degradation by accelerating de-mannosylation of folding-defective polypeptides and by inhibiting their covalent aggregation.* Biochem Biophys Res Commun, 2006. **349**(4): p. 1278-84.

18. Moremen, K.W., R.B. Trimble, and A. Herscovics, *Glycosidases of the asparagine-linked oligosaccharide processing pathway.* Glycobiology, 1994. **4**(2): p. 113-25.

19. Lipari, F., B.J. Gour-Salin, and A. Herscovics, *The Saccharomyces cerevisiae processing alpha 1,2-mannosidase is an inverting glycosidase.* Biochem Biophys Res Commun, 1995. **209**(1): p. 322-6.

20. Daniel, P.F., B. Winchester, and C.D. Warren, *Mammalian alpha-mannosidases--multiple forms but a common purpose?* Glycobiology, 1994. **4**(5): p. 551-66.

21. Howard, S., et al., *Human lysosomal and jack bean alpha-mannosidases are retaining glycosidases.* Biochem Biophys Res Commun, 1997. **238**(3): p. 896-8.

22. Kornfeld, R. and S. Kornfeld, *Assembly of asparagine-linked oligosaccharides.* Annu Rev Biochem, 1985. **54**: p. 631-64.

23. Helenius, A. and M. Aebi, *Roles of N-linked glycans in the endoplasmic reticulum.* Annu Rev Biochem, 2004. **73**: p. 1019-49.

24. Helenius, A. and M. Aebi, *Intracellular functions of N-linked glycans.* Science, 2001. **291**(5512): p. 2364-9.

25. Samuelson, J., et al., *The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases.* Proc Natl Acad Sci U S A, 2005. **102**(5): p. 1548-53.

26. Knauer, R. and L. Lehle, *The oligosaccharyltransferase complex from yeast.* Biochim Biophys Acta, 1999. **1426**(2): p. 259-73.

27. Trombetta, E.S. and A. Helenius, *Lectins as chaperones in glycoprotein folding.* Curr Opin Struct Biol, 1998. **8**(5): p. 587-92.

28. Zapun, A., et al., *Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57.* J Biol Chem, 1998. **273**(11): p. 6009-12.

29. Parodi, A.J., *Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation.* Biochem J, 2000. **348 Pt 1**: p. 1-13.

30. Tessier, D.C., et al., *Cloning and characterization of mammalian UDP-glucose glycoprotein: glucosyltransferase and the development of a specific substrate for this enzyme.* Glycobiology, 2000. **10**(4): p. 403-12.

31. Sousa, M.C. and A.J. Parodi, *The interaction of the UDP-GLC:glycoprotein glucosyltransferase with the acceptor glycoprotein.* Cell Mol Biol (Noisy-le-grand), 1996. **42**(5): p. 609-16.

32. Trombetta, E.S. and A.J. Parodi, *Quality control and protein folding in the secretory pathway.* Annu Rev Cell Dev Biol, 2003. **19**: p. 649-76.

33. Jarosch, E., U. Lenk, and T. Sommer, *Endoplasmic reticulum-associated protein degradation.* Int Rev Cytol, 2003. **223**: p. 39-81.

34. Wu, Y., et al., *Elucidation of the molecular logic by which misfolded alpha 1-antitrypsin is preferentially selected for degradation.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8229-34.

35. Cabral, C.M., Y. Liu, and R.N. Sifers, *Dissecting glycoprotein quality control in the secretory pathway.* Trends Biochem Sci, 2001. **26**(10): p. 619-24.

36. Jakob, C.A., et al., *Degradation of misfolded endoplasmic reticulum glycoproteins in Saccharomyces cerevisiae is determined by a specific oligosaccharide structure.* J Cell Biol, 1998. **142**(5): p. 1223-33.

37. Liu, Y., et al., *Oligosaccharide modification in the early secretory pathway directs the selection of a misfolded glycoprotein for degradation by the proteasome.* J Biol Chem, 1999. **274**(9): p. 5861-7.

38. Yang, M., et al., *Novel aspects of degradation of T cell receptor subunits from the endoplasmic reticulum (ER) in T cells: importance of oligosaccharide processing, ubiquitination, and proteasome-dependent removal from ER membranes.* J Exp Med, 1998. **187**(6): p. 835-46.

39. McCracken, A.A. and J.L. Brodsky, *Assembly of ER-associated protein degradation in vitro: dependence on cytosol, calnexin, and ATP.* J Cell Biol, 1996. **132**(3): p. 291-8.

40. Wang, Y. and M.J. Androlewicz, *Oligosaccharide trimming plays a role in the endoplasmic reticulum-associated degradation of tyrosinase.* Biochem Biophys Res Commun, 2000. **271**(1): p. 22-7.

41. Hosokawa, N., et al., *Enhancement of endoplasmic reticulum (ER) degradation of misfolded Null Hong Kong alpha1-antitrypsin by human ER mannosidase I.* J Biol Chem, 2003. **278**(28): p. 26287-94.

42. Lal, A., et al., *Isolation and expression of murine and rabbit cDNAs encoding an alpha 1,2-mannosidase involved in the processing of asparagine-linked oligosaccharides.* J Biol Chem, 1994. **269**(13): p. 9872-81.

43. Herscovics, A., et al., *Isolation of a mouse Golgi mannosidase cDNA, a member of a gene family conserved from yeast to mammals.* J Biol Chem, 1994. **269**(13): p. 9864-71.

44. Igdoura, S.A., et al., *Alpha-mannosidases involved in N-glycan processing show cell specificity and distinct subcompartmentalization within the Golgi apparatus of cells in the testis and epididymis.* Eur J Cell Biol, 1999. **78**(7): p. 441-52.

45. Gonzalez, D.S., et al., *Identification, expression, and characterization of a cDNA encoding human endoplasmic reticulum mannosidase I, the enzyme that catalyzes the first mannose trimming step in mammalian Asn-linked oligosaccharide biosynthesis.* J Biol Chem, 1999. **274**(30): p. 21375-86.

46. Vallee, F., et al., *Crystal structure of a class I alpha1,2-mannosidase involved in N-glycan processing and endoplasmic reticulum quality control.* EMBO J, 2000. **19**(4): p. 581-8.

47. Vallee, F., et al., *Structural basis for catalysis and inhibition of N-glycan processing class I alpha 1,2-mannosidases.* J Biol Chem, 2000. **275**(52): p. 41287-98.

48. Van Petegem, F., et al., *Trichoderma reesei alpha-1,2-mannosidase: structural basis for the cleavage of four consecutive mannose residues.* J Mol Biol, 2001. **312**(1): p. 157-65.

49. Lobsanov, Y.D., et al., *Structure of Penicillium citrinum alpha 1,2-mannosidase reveals the basis for differences in specificity of the endoplasmic reticulum and Golgi class I enzymes.* J Biol Chem, 2002. **277**(7): p. 5620-30.

50. Tempel, W., et al., *Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases.* J Biol Chem, 2004. **279**(28): p. 29774-86.

51. Thompson, A.J., et al., *The reaction coordinate of a bacterial GH47 alpha-mannosidase: a combined quantum mechanical and structural approach.* Angew Chem Int Ed Engl, 2012. **51**(44): p. 10997-1001.

52. Vallee, F., et al., *Purification, crystallization and preliminary X-ray crystallographic analysis of recombinant murine Golgi mannosidase IA, a class I alpha-mannosidase*

involved in Asn-linked oligosaccharide maturation. Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 2): p. 571-3.

53. Karaveg, K. and K.W. Moremen, *Energetics of substrate binding and catalysis by class 1 (glycosylhydrolase family 47) alpha-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(33): p. 29837-48.

54. Karaveg, K., et al., *Mechanism of class 1 (glycosylhydrolase family 47) {alpha}-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(16): p. 16197-207.

55. Romero, P.A., B. Sleno, and A. Herscovics, *Glycoprotein biosynthesis in Saccharomyces cerevisiae. Partial purification of the alpha-1,6-mannosyltransferase that initiates outer chain synthesis.* Glycobiology, 1994. **4**(2): p. 135-40.

56. Karaveg, K., et al., *Crystallization and preliminary X-ray diffraction analysis of lectin-1 from Pseudomonas aeruginosa.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 7): p. 1241-2.

57. Mulakala, C. and P.J. Reilly, *Understanding protein structure-function relationships in Family 47 alpha-1,2-mannosidases through computational docking of ligands.* Proteins, 2002. **49**(1): p. 125-34.

58. Mulakala, C., W. Nerinckx, and P.J. Reilly, *Docking studies on glycoside hydrolase Family 47 endoplasmic reticulum alpha-(1-->2)-mannosidase I to elucidate the pathway to the substrate transition state.* Carbohydr Res, 2006. **341**(13): p. 2233-45.

59. Mulakala, C., W. Nerinckx, and P.J. Reilly, *The fate of beta-D-mannopyranose after its formation by endoplasmic reticulum alpha-(1-->2)-mannosidase I catalysis.* Carbohydr Res, 2007. **342**(2): p. 163-9.

60. Davies, G.J., A. Planas, and C. Rovira, *Conformational analyses of the reaction coordinate of glycosidases.* Acc Chem Res, 2012. **45**(2): p. 308-16.

61. Cantu, D., W. Nerinckx, and P.J. Reilly, *Theory and computation show that Asp463 is the catalytic proton donor in human endoplasmic reticulum alpha-(1-->2)-mannosidase I.* Carbohydr Res, 2008. **343**(13): p. 2235-42.

62. Gregg, K.J., et al., *Analysis of a new family of widely distributed metal-independent alpha-mannosidases provides unique insight into the processing of N-linked glycans.* J Biol Chem, 2011. **286**(17): p. 15586-96.

63. Davies, G.J., et al., *Mapping the conformational itinerary of beta-glycosidases by X-ray crystallography.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 523-7.

64. Rye, C.S. and S.G. Withers, *Glycosidase mechanisms.* Curr Opin Chem Biol, 2000. **4**(5): p. 573-80.

65. Zechel, D.L. and S.G. Withers, *Dissection of nucleophilic and acid-base catalysis in glycosidases.* Curr Opin Chem Biol, 2001. **5**(6): p. 643-9.

66. Karaveg, K. and K.W. Moremen, *Energetics of Substrate Binding and Catalysis by Class 1 (Glycosylhydrolase Family 47) {alpha}-Mannosidases Involved in N-Glycan Processing and Endoplasmic Reticulum Quality Control.* J. Biol. Chem., 2005. **280**(33): p. 29837-29848.

67. Stoddart, J.F., *Stereochemistry of carbohydrates*. 1971, New York,: Wiley-Interscience. xi, 249 p.

68. Davies, G.J., et al., *Snapshots along an enzymatic reaction coordinate: analysis of a retaining beta-glycoside hydrolase.* Biochemistry, 1998. **37**(34): p. 11707-13.

CHAPTER 2

Structural basis for substrate recognition by mammalian GH47 α1,2-mannosidases

2.1 Abstract

CAZy glycoside hydrolase family 47 α-mannosidases [1] cleave α1,2-mannose linkages on Asn-linked oligomannose structures by an inverting mechanism that involves an enzyme-bound Ca^{2+} ion that contributes to substrate distortion of the glycone into a transition state conformation. Among enzymes in this family, ER α-mannosidase I (ERManI) and Golgi α-mannosidase IA (GMIA) are key enzymes involved in *N*-linked glycan biosynthesis in the endoplasmic reticulum (ER) and Golgi complex. The structures of ERManI and GMIA are similar in overall protein fold, but have complementary activities in their cleavage of the natural $Man_9GlcNAc_2$ substrate. ERManI preferentially cleaves a single α1,2-mannose residue from the central B-branch of the tri-branched oligomannose substrate, whereas GMIA preferentially cleaves three α1,2-mannose residues from branches A and C, but poorly cleaves the B branch α1,2-mannose residue. In order to determine the structural basis for their respective substrate specificity differences, human ERManI and murine GMIA were each inactivated by replacing the enzyme-bound Ca^{2+} with La^{3+} and co-crystallized with the substrate analog $Man_9GlcNAc_2$-PA. The crystal structures were resolved by molecular replacement at resolutions of 1.65 Å for ERManI and 1.77 Å for GMIA. The enzyme-substrate complex structures reveal that the two enzymes bind their natural glycan substrates quite differently, consistent with the topology differences in their glycan binding clefts leading to the respective active sites. GMIA binds the non-reducing terminal α1,2-mannose residue of branch A in the -1 enzyme subsite, while the remainder of the glycan has extended interactions in the cleft leading to the active site. In contrast, ERManI binds to the non-reducing terminal α1,2-mannose residue of branch B with a

distinctive set of interactions between the glycan substrate and the active site cleft. The comparison of the crystal structures provides insights into the structural basis of substrate specificity for members of this enzyme family.

2.2 Introduction

Mammalian α1,2-mannosidases are also known as "class 1 α-mannosidases" or glycoside hydrolase family 47 (GH47) enzymes [1]. Members of this family cleave specific α1,2-mannose linkages on Asn-linked (*N*-linked) oligomannose structures by an inverting mechanism [2, 3]. In mammals there are seven GH47 members, including ER α1,2-mannosidase I (ER mannosidase I or ERManI), three Golgi α1,2-mannosidases (Golgi mannosidase IA/IB/IC (GMIA, GMIB, and GMIC)), which play important roles on the *N*-glycan trimming in the ER and Golgi complex, and three EDEM proteins (EDEM1, EDEM2, and EDEM3) that play roles in ER-associated degradation (ERAD) [1]. In the classical model for *N*-glycan processing, newly synthesized glycoproteins, containing $Glc_3Man_9GlcNAc_2$ glycan precursor structures, are trimmed by α-glucosidases in the ER to generate a $Man_9GlcNAc_2$ processing intermediate. Subsequent action by ERManI results in the cleavage of a single α1,2-mannose residue from the middle branch of the $Man_9GlcNAc_2$ structure to generate $Man_8GlcNAc_2$ isomer B (Figure 23). Glycoproteins containing $Man_8GlcNAc_2$-B *N*-glycans are then transported to the Golgi complex, where the Golgi mannosidase I subfamily enzymes cleave the remaining three α1,2-mannosyl residues to generate a $Man_5GlcNAc_2$ glycan intermediate structure [4-6]. The subcellular location of ERManI has recently been called into question with the identification of a post-ER localization for the enzyme either in the Golgi complex [7, 8] or in an intermediate quality control compartment [9] where it acts as a critical timing step for entry of misfolded proteins into ERAD. Glycosyltransferases and hydrolases in the Golgi complex then further modify the

$Man_5GlcNAc_2$ intermediate to yield complex and hybrid glycan structures on cell surface and secreted glycoproteins [10]. *In vitro,* ERManI and the Golgi α-mannosidases display distinct substrate specificities towards the four α1,2-mannose residues present in the $Man_9GlcNAc_2$ structure (Figure 23). ERManI rapidly cleaves the single M10 α1,2-mannose residue on the middle branch (branch B) of $Man_9GlcNAc_2$ to generate $Man_8GlcNAc_2$ isomer B, but inefficiently cleaves the remaining three α1,2-mannose residues on branches A and C to yield $Man_5GlcNAc_2$. GMIA, by contrast, preferentially cleaves α1,2-mannose residue M11 on branch A, first to generate $Man_8GlcNAc_2$ isomer A, and then sequentially cleaves residues M9 and M8 on the branches C and A to generate $Man_6GlcNAc_2$. Cleavage of the final α1,2-mannose residue (M10) on branch B, the preferred substrate for ERManI, is inefficient and GMIA generates the $Man_5GlcNAc_2$ structure relatively slowly [3, 11-13] (Figure 23). Thus, ERManI and GMIA are complementary and minimally overlapping in their specificity toward cleavage of the four α1,2-mannose residues in the $Man_9GlcNAc_2$ precursor structure.

The crystal structures of several GH47 α-mannosidases have been determined and all are comprised of an $(\alpha\alpha)_7$ barrel structure with a broad cleft extending from one end of the barrel leading to an active site in the barrel core. Previous studies have also explored substrate interactions by determining the crystal structures of enzymes in complex with substrate analogs within the -1 subsite [14-18], bridging the -1 and +1 subsites [16, 19], and extending from the ≥+1 subsites [17, 20, 21]. However, these studies were not able to provide a full snapshot of interactions for the intact substrate spanning from the -1 to ≥+1 subsites [14, 17, 19, 20, 22, 23] to map the determinants of glycan branch specificity. In the -1 subsite, co-complex structures of ERManI with monosaccharide mimics, 1-deoxymannojirimycin (dMNJ) and kifunensine (Kif), revealed that the glycone binding subsite was occupied by the inhibitors in a distorted high

energy $^{1}C_4$ sugar conformation, with the equivalent of the C-2' and C-3' hydroxyls being coordinated with an essential enzyme-bound Ca^{2+} ion [12, 17]. Several other enzyme residues contributed interactions that stabilized the energetically unfavorable $^{1}C_4$ conformation indicating that this structure was critical for catalysis. All of the enzyme residues in the -1 subsite coordinating the Ca^{2+} ion and stabilizing inhibitor interactions were conserved among GH47 family members suggesting a common mechanism for glycone interactions [16]. Similarly, a co-complex between ERManI and an uncleavable α1,2-mannobiose thiodisaccharide analog revealed substrate interactions within the -1 and +1 subsites, including a distorted glycone residue in the -1 subsite in a $^{3}S_1$ conformation [19].

Enzyme complexes that occupy the ≥+1 subsites were also obtained for both mouse GMIA and the yeast ERManI [12, 20]. Both of these complexes were obtained as a result of cross-lattice contacts of the glycan structures attached to one enzyme monomer extending into the active site cleft of the adjoining monomer in the crystal lattice. The -1 subsites of these complexes were unoccupied, but the respective +1 subsites were occupied with distinctive residues for each enzyme isoform suggesting that they represented enzyme-product complexes. For GMIA, the M6 mannose residue of branch C occupied the +1 subsite, while ERManI had the M7 mannose residue of branch B in the +1 subsite. In contrast to the residues within the -1 and +1 subsites, which are conserved among the GH47 α-mannosidases, the residues and topologies of the broad clefts leading to the active site cores are quite distinct between the respective GH47 family members [15, 17, 18, 20, 21]. Differences in these cleft residues contribute to the distinctive interactions that facilitate branch-specific cleavage of the $Man_9GlcNAc_2$ substrate. For example, Arg273 in yeast ERManI had extensive interactions with numerous mannose residues in the substrate compared to minimal interactions of Leu413 in the equivalent position of GMIA.

However, mutagenesis of Arg273 in ERManI to Leu led to a low efficiency enzyme with hybrid activity indicating that additional residues in the cleft region are required for efficient branch-specific cleavage [21, 24].

While extended substrate interactions from 1- to >+1 residues can be inferred for both ERManI and GMIA by modeling the combination of partial substrate analog and product complexes into the active site, it is unclear whether this assembly of interactions would reflect the true nature of an intact substrate-bound Michaelis complex. For ERManI, the $Man_5GlcNAc_2$ product complex [17] contains the M6 mannose residue from branch B in the +1 subsite as would be expected for cleavage of the $Man_9GlcNAc_2$ substrate to $Man_8GlcNAc_2$ isomer B. However, for GMIA the enzyme-product complex [20] contains the M6 mannose residue from branch C in the +1 subsite rather than the M8 residue as expected for the preferred primary cleavage of $Man_9GlcNAc_2$ branch A by the enzyme. Thus, a structural understanding of branch specificity of mannose cleavage among the GH47 α-mannosidases remains unclear.

A challenge for the structural analysis of enzyme-substrate complexes for the GH47 α-mannosidases is the rapid cleavage of glycan substrates even with partially inactivating mutants and the requirement for the enzyme-bound Ca^{2+} ion in the active site for both the substrate binding and catalysis. In the present study we tested the replacement of the Ca^{2+} ion with other ions that would maintain substrate affinity, but eliminate enzymatic activity. Through these studies we determined that exchange of the enzyme bound Ca^{2+} ion with La^{3+} resulted in enzyme inactivation but retains high substrate binding affinity. Co-crystallization of the La^{3+}-inactivated GMIA and ERManI with the natural substrate analog $Man_9GlcNAc_2$-PA led to the formation of enzyme-glycan complexes with distinct branches engaged in the -1 subsites (mannose residue M10 from branch B residing in the -1 subsite of ERManI and M11 from branch A for GMIA).

The two structures are proposed to reflect intact enzyme-substrate complexes for GH47 α-mannosidases and provide structural insights into substrate recognition for this family of enzymes.

2.3 Experimental Procedures

2.3.1 Preparation of substrate analog $Man_9GlcNAc_2$-PA.

Soybean agglutinin (SBA) was selected as the starting material for oligosaccharide isolation due to its well characterized $Man_9GlcNAc_2$ glycan structure [25]. SBA is a tetramer, and has a single homogeneous *N*-linked $Man_9GlcNac_2$ glycan on each subunit. Crude SBA was extracted from soybean paste using a pH 4.5 HCl solution, and was then digested with trypsin to generate glycopeptides. The glycopeptides were separated by Sephadex G50 and HW40 size exclusion chromatography and the glycopeptide fraction was pooled. Glycans were cleaved from the glycopeptides by PNGaseF digestion. Phenol/chloroform extraction was then used to remove most of the peptides and the free oligosaccharides were reductively aminated with 2-amino pyridine (PA). Phenol/chloroform extraction was performed to remove most of the free PA and Bio-Gel P-2 size exclusion chromatography was used to further remove the remaining free PA from the tagged glycans. The fluorescent tagged oligosaccharides were further separated by reverse HPLC with either analytical or preparative Hypersil APS-2 amine columns to generate highly enriched substrate structures for enzymatic analysis.

2.3.2 Protein Expression, Deglycosylation and Metal Ion Replacement.

The *Pichia pastoris* expression system was chosen to express human ERManI and GMIA. The plasmid construction, transformation, culture growth, protein expression and purification were carried out as previously described [20]. Recombinant EndoF1 as a cellulose binding module fusion was purified and bound to cellulose beads as previously reported [26].

The purified GMIA was concentrated to ~1 mg/ml and mixed with the EndoF1 cellulose beads in buffer A (100 mM NaMES, pH 6.0). After incubation with GMIA at 37 ℃ for 4 hours, the EndoF1 beads were removed from the reaction solution by centrifugation. The detached glycans were then separated from the de-glycosylated protein by dialysis against buffer A. The deglycosylated GMIA was further purified by Superdex-75 gel filtration chromatography in buffer B (20 mM HEPES, pH 6.0) and was concentrated to 1 mg/ml in buffer B.

EDTA was added to the concentrated deglycosylated GMIA to a final concentration of 1 mM followed by incubation at 4℃ for 1 hr. The EDTA treated sample was then injected onto a Superdex-75 gel filtration column (2 x 65 cm, Amersham Pharmacia Biotech) pre-equilibrated with buffer C (20 mM HEPES, 100 μM EDTA pH 7.0). The gel filtration column was performed in buffer C at a flow rate of 1 ml/min for 200 minutes. The fractions of calcium-free GMIA (48 kDa by SDS-PAGE) were combined and concentrated with stirred concentration pressure cell (Amicon, Inc., Beverly, MA) to a final concentration of 20 mg/ml in buffer D (20 mM HEPES, 5 mM $LaCl_3$, 200 mM NDSB-201, pH 6.5) for crystallization studies.

2.3.3 Crystallization and X-ray diffraction.

Equal volumes of $Man_9GlcNAc_2$-PA (5 mM) and the deglycosylated GMIA protein preparation (10 mg/ml in buffer D) were mixed immediately before crystallization. Microbatch crystallization was used to screen and optimize the crystallization conditions. Crystals began to grow after 2 days at 18℃ in 100 mM MES, 50 mM ammonium sulfate, 15% PEG 4000, pH 6.5, and were mounted and flash-frozen after 8 days. The diffraction data were collected at SER-CAT/APS (Southeast Regional Collaborative Team/Advanced Photon Source at Argonne National Lab) by remote data collection from University of Georgia. The data were collected with beam line 22ID using a MAR300 detector. A total of 360 frames of 1 ° oscillation images

were collected. HKL2000 software was used for data integration and scaling. The structure was resolved by molecular replacement using the CCP4 Program Suite and PDB 1NXC as the search model. The structure model was built with Coot Software and repeatedly refined with Refmac-5 /CCP4 software.

2.4 Results

2.4.1 Deglycosylation of GMIA.

The deglycosylation of GMIA was confirmed by SDS-PAGE by comparison with GMIA samples prepared before and after the EndoF1 treatment. The deglycosylated GMIA ran as a discrete band ~2 kDa smaller than the wild type enzyme. Parallel assays of the deglycosylated GMIA and native GMIA using $Man_9GlcNAc_2$-PA as substrate did not show any differences in enzymatic activity (data not shown). In the presence of NDSB-201(200 mM), both wild type and deglycosylated forms of GMIA can be concentrated to 10 mg/ml, and stored at 4℃ for a week without observed precipitation.

2.4.2 Inactivation of the enzyme by La^{3+} ion

Calcium is essential for catalytic activity of the GH47 α1,2-mannosidase family [5, 11, 17, 21]. The conserved Thr residue (Thr635 in GMIA, Thr688 in ERManI) was found to participate the coordination of the calcium ion [17, 20]. Mutation of the Thr688 in ERManI resulted in a 61-fold decrease in k_{cat} [14]. Depletion of calcium by EDTA leads to a complete inactivation of both GMIA and ERManI towards $Man_9GlcNAc_2$-PA in a 24 hr reaction. Since the glycone mannose residue and the inhibitor in the -1 subsite are both stabilized by interactions with the calcium ion, depletion of the calcium will also remove the stabilization conferred by calcium coodination. Our strategy was to find a substitute metal ion to replace the Ca^{2+} ion while retaining the coordination and its stabilizing effect on substrate binding, and at the same time, inactivation of

the enzyme. Addition of calcium chloride to the EDTA-treated enzymes successfully rescued enzymatic activity for cleavage of $Man_9GlcNAc_2$-PA. However, addition of lanthanum chloride to the EDTA-treated enzymes did not rescue enzyme activity even with a prolonged 24 hr incubation (data not shown).

2.4.3 The crystal structures of mouse GMIA-$Man_9GlcNAc_2$-PA and human ERManI-$Man_9GlcNAc_2$-PA.

Both mouse GMIA and yeast ERManI are *N*-linked glycoproteins and their crystal structures were found to have cross lattice contacts between the glycan structures associated with each protein in the crystal lattice and the active site of an adjoining enzyme monomer in the lattice to form an enzyme-oligosaccharide complex (Figure 24A and 24B) [20, 21]. The oligosaccharides found in the enzyme active site clefts were equivalent to an exhaustive enzymatic cleavage products of α1,2-mannosidase action, therefore the crystal structures of both yeast ERManI and GMIA were proposed to represent enzyme-product complexes. In the present study, the EndoF1 cleavage of GMIA to remove protein-linked glycans circumvented the possibility of such cross lattice protein-glycan crystal contacts and raised the possibility of generating exogenous substrate complexes to probe substrate interactions. As a result, the structure of deglycosylated GMIA, where the Ca^{2+} ion had been swapped for La^{3+} and incubated in the presence of $Man_9GlcNAc_2$-PA, allowed the formation of an uncleaved enzyme-substrate complex, and there were no glycan-based contacts between protein units in the crystal lattice (Figure 24C).

Human ERManI enzyme does not contain a consensus *N*-glycan acceptor site so EndoF1 digestion prior to crystallization was not required. The structure of human ERManI, where the Ca^{2+} ion was swapped for La^{3+} and incubated in the presence of $Man_9GlcNAc_2$-PA, led to the formation of an uncleaved enzyme-substrate complex consistent with prior structures of human

ERManI as inhibitor- (dMNJ/Kif) [17] or uncleaved thiodisaccharide-bound complexes [19] (Figure 24D).

2.4.4 Deglycosylation, $Man_9GlcNAc_2$-PA binding, and metal ion replacement have no apparent influence on the general structure of the enzymes.

Comparison of the previously resolved GH47 α1,2-mannosidase structures [15-18, 20, 21] revealed that they all consist of an $(\alpha\alpha)_7$ barrel structure comprised of two layers of concentric inner and outer helices in which the catalytic site is located deep within the barrel above a β-hairpin that plugs one end of the barrel. A Thr residue at the apex of the β-hairpin in the core of the barrel is involved in coordinating a metal ion (Ca^{2+}) assisted by indirect interactions through water molecules. The opposite end of the barrel consists of an open funnel-shaped conical cleft leading to the more constricted pocket in the core of the barrel containing the -1 glycone-binding subsite. In the GMIA-$Man_9GlcNAc_2$-PA complex and the ERManI-$Man_9GlcNAc_2$-PA complex, a La^{3+} ion is found to replace the Ca^{2+} *in situ*. Clear electron densities for the $Man_9GlcNAc_2$-PA glycan structures are seen extending across the broad cleft and deep into the -1 subsite for both enzymes. The two enzyme-substrate complexes of GMIA and ERManI do not show any apparent changes in protein structure as a result of the loss of the surface glycan for GMIA or as a consequence of the replacement of the bound Ca^{2+} ion with the La^{3+}, or the presence of bound $Man_9GlcNAc_2$-PA in both enzymes, (Figure 24C and 24D). Thus, no induced changes in protein structure are observed upon binding of the intact $Man_9GlcNAc_2$-PA glycan in either enzyme.

In a previous study, the sizes of the glycan-binding pockets of several GH47 α1,2-mannosidases were measured and compared [20]. The ≥+1 subsite cavity of GMIA [20] was found to be more constricted compared to human ERManI and the *P. citrinum* α1,2-mannosidase [17]. In the structures of enzyme-$Man_9GlcNAc_2$-PA substrate analog complexes of mouse GMIA

and human ERManI, the measurements cross the >+1 glycan-binding subsites and the upper portion of the glycan binding cleft were similar to those made in the previous study [20]. In the +1 subsites of the enzyme-substrate complexes for mouse GMIA, the distances between the amino acid residue pairs (GMIA: Glu282 and Gly412, Glu351 and Leu476) are similar to those in the enzyme-$Man_6GlcNAc_2$ product complex of GMIA. In the enzyme-$Man_9GlcNAc_2$-PA substrate complex of human ERManI, the distances between the residue pairs (human ERManI: Glu330 and Ala460, Glu397 and Leu525) are also similar to those in the enzyme-disaccharide complex of human ERManI. In the upper part of the glycan-binding clefts of the enzyme-substrate complexes, distances between the residue pairs (GMIA: Ser233 and Lys465; ERManI: Trp284 and Ala514) show no apparent differences with the binding of $Man_9GlcNAc_2$-PA compared to those in the enzyme-$Man_6GlcNAc_2$ product complex and enzyme-disaccharide complexes respectively. The distance measurement of other residue pairs on the upper cleft region (GMIA: Trp341 and Asn540; human ERManI: Trp389 and Ala590) revealed an apparent size increase in this dimension. Both the residue pairs are farther apart in the enzyme-$Man_9GlcNAc_2$-PA substrate complexes compared to their counterparts in the enzyme-product or enzyme-inhibitor complexes. In the GMIA-$Man_9GlcNAc_2$-PA complex, the distance between Trp341 and Asn540 is 16.9 Å, an increase of 2.4 Å compared to that in the GMIA-$Man_6GlcNAc_2$ product structure. In the human ERManI-$Man_9GlcNAc_2$-PA complex, the distance between Trp339 and Ala590 increased to 23.5 Å compared to 17.6 Å in the ERManI-dMNJ complex structure [17]. The increase of the distances between the residues of those two pairs results from the different positioning of the side chains of Trp341 and Arg338 in the enzyme-substrate complexes. The side chain of Trp341 in the GMIA structure has hydrophobic stacking with the nonpolar face of NAG2 residue of the $Man_6GlcNAc_2$ [20]. However, in the GMIA-

$Man_9GlcNAc_2$-PA complex structure, the side chain of Trp341 flips ~180 °away from its original position in GMIA in order to avoid a clash with the NAG2 residue of $Man_9GlcNAc_2$-PA in the enzyme-substrate complex. There is no stacking interaction between the Trp341 side chain and NAG2 of $Man_9GlcNAc_2$-PA in the GMIA enzyme-substrate complex. The Arg338 residue in human ERManI is located at the edge of the enzyme cavity. The side chain of this residue has more flexibility than those residues inside the crowded glycan binding cleft. The Arg338 side chain in the enzyme-substrate complex extends away from Ala590 compared to the relative orientation in the enzyme-inhibitor complex, and thus increases the distance between the two residues. Despite the enlarged upper portion of the binding clefts for both GMIA and human ERManI in the enzyme-$Man_9GlcNAc_2$-PA substrate complexes, the entire binding cleft of GMIA-$Man_9GlcNAc_2$-PA complex is still more constricted than the cleft of the human ERManI-$Man_9GlcNAc_2$-PA complex.

2.4.5 Coordination of Lanthanum in the Active Site.

In the GMIA crystal structure containing $Man_6GlcNAc_2$ in the active site, the Ca^{2+}-centered 8-fold square anti-prismatic coordination consists of six water molecules along with Thr635 O-γ and the backbone carbonyl oxygen, with the two Thr635 oxygens and two waters on one side and the other four water molecules on the opposite side of the Ca^{2+} ion (Figure 25E). Calcium coordination in the human ERManI-dMNJ complex is similar with that of GMIA structure. The metal ion holds an 8-fold coordination geometry, with four coordinated waters reside on one side of the Ca^{2+} ion and the O-γ and the backbone carbonyl oxygen atoms of Thr688 as well as the C-2 and C-3 hydroxyl groups of dMNJ reside on the other side of the Ca^{2+} ion (Figure 25B). In contrast, the structures of the GMIA-$Man_9GlcNAc_2$-PA complex and the human ERManI-$Man_9GlcNAc_2$-PA complex revealed that La^{3+} binds the enzymes with different coordination

geometries compared to the Ca^{2+}. The La^{3+} ion takes a 9-fold trigonal prismatic (square face tricapped) geometry (Figure 25A and 25D).

One side of the La^{3+} ion has a similar geometry to the Ca^{2+} coordination with four functional groups coordinating with the metal ion: two points of coordination are still contributed by the O-γ and backbone carbonyl oxygen of the conserved Thr residue on the top the β-hairpin (Thr635 in GMIA, Thr688 in human ERManI); the other two points of coordination are contributed by the C-2 and C-3 hydroxyl oxygens of the mannose residue located in the -1 glycone binding subsite (M11 for GMIA, M10 for human ERManI). On the other side of the metal ion, instead of four water molecules, five waters coordinate with the La^{3+} (Figure 25A, 25D).

To better compare the La^{3+} and Ca^{2+} coordination, the structures of enzyme-$Man_9GlcNAc_2$-PA substrate complex and enzyme-$Man_6GlcNAc_2$ product complex of GMIA were superimposed as were the structures of the enzyme-substrate and enzyme-dMNJ complexes of human ERManI. (Figure 25C, Figure 25F). The C-2 and C-3 hydroxyls of the M11 mannose residue in the La^{3+} bound GMIA-$Man_9GlcNAc_2$-PA substrate complex were similar to the positions otherwise occupied by the two waters ($W_{Ca}5$ and $W_{Ca}6$) in the Ca^{2+}-bound GMIA enzyme-product structure (Figure 24F). In the human ERManI-$Man_9GlcNAc_2$-PA substrate complex, the C-2 and C-3 hydroxyl groups of M10 also reside in the similar locations as those of dMNJ in the Ca^{2+}-bound human ERManI-inhibitor complex (Figure 25C). Despite the replacement of the two water molecules ($W_{Ca}5$ and $W_{Ca}6$) by the hydroxyl groups of M11/M10, the four-fold geometries of the coordination on this side of the La^{3+} ion are in very similar positions compared to the Ca^{2+} bound enzymes.

Alignment of the La^{3+}-bound GMIA-$Man_9GlcNAc_2$-PA complex and human ERManI-$Man_9GlcNAc_2$-PA complex indicated that the La^{3+} coordination in the two complex structures

are essentially superimposable (Figure 25G), so the following description of metal coordination will only use GMIA as the structure model to represent the complexes for both GMIA and ERManI. Compared to the Ca^{2+} coordination, the apparent geometric difference of the La^{3+} coordination lies on the side opposite to the Thr635 and the M11 mannose residue. Unlike the coordination geometry of Ca^{2+} with four water molecules ($W_{Ca}1$, $W_{Ca}2$, $W_{Ca}3$ and $W_{Ca}4$), La^{3+}-coordination comprises five water molecules ($W_{La}1$, $W_{La}2$, $W_{La}3$, $W_{La}4$ and $W_{La}5$), with four water molecules located approximately on one plane, while the fifth water molecule falls above the plane and between the other four water molecules with almost equal distance to the four water molecules to form a square-pyramid. Alignment of the La^{3+}-bound GMIA and Ca^{2+}-bound GMIA structures reveals, that $W_{La}1$, $W_{La}2$, and $W_{La}3$ in La^{3+} coordination are located approximately at the equivalent positions relative to their counterparts, $W_{Ca}1$, $W_{Ca}2$, and $W_{Ca}3$, in Ca^{2+} coordination (Figure 25F). Therefore, the three pairs of water molecules ($W_{La}1$ and $W_{Ca}1$, $W_{La}2$ and $W_{Ca}2$, $W_{Ca}3$ and $W_{La}3$) reside approximately on the same plane, as does $W_{La}4$. However, $W_{Ca}4$ falls out of the plane with $W_{La}4$, and also resides in a location distinct from $W_{La}5$. In summary, the coordination geometry change from Ca^{2+} to La^{3+} is consistent between GMIA and human ERManI and presumably accounts for the lack of enzyme activity for both enzymes (Figure 25C, 25F, 25G).

2.4.6 Mannose Conformation in the Enzyme-Substrate Complex.

Mannose in solution, or as part of an oligosaccharide, is mostly found in a $^4C_1$ conformation, with the C-3 and C-4 hydroxyls and the C-5 hydroxymethyl groups in a low free energy equatorial conformation. Previous studies demonstrated that the inhibitor dMNJ, an analog of mannopyranose, bound to human ERManI and *P. citrinum* α1,2-mannosidase in an energetically unfavorable $^1C_4$ conformation with the C-3 and C-4 hydroxyls and the C-5 hydroxymethyl all in

axial conformation (Figure 26A) [15, 17]. Methyl-2-S-(α-D-mannopyranosyl)-2-thio-α-D-mannopyranoside (S-Man_2) is a thiodisaccharide that is resistant to cleavage by α1,2-mannosidases. Crystal structures of the human ERManI-S-Man_2 complex and the GH47 bacterial mannosidase-S-Man_2 complex showed that both the -1 and +1 subsites are occupied respectively by mannosyl units connected to each other by a thio-glycosidic bond spanning from the -1 to the +1 subsite. Instead of the $^1C_4$ or $^4C_1$ conformation, the glycone of the thiodisaccharide in the -1 subsite adopts the $^3S_1$ skew boat conformation, which contains the identical configurations of C-2, C-3, C-4, C5 and C-6, compared to those in the $^1C_4$ conformation, as well as the distorted configurations of C-1 and O-5. It was proposed that the dMNJ represents the conformation of the β-mannose enzymatic product while the glycone of the S-Man_2 in the -1 subsite represent the pre-transition intermediate conformation [17, 21, 27-29].

In the present study, mannose residues M11 and M10 were found in the -1 subsites of GMIA-$Man_9GlcNAc_2$-PA complex and human ER ManI-$Man_9GlcNAc_2$-PA complex, respectively, both in a $^1C_4$ conformation with the C-2 and C-3 hydroxyl groups directly coordinating with the La^{3+} ion. The C-1 atoms are connected to the C-2 atom of the adjoining mannose residue in the +1 subsite through an intact α1,2-glycosidic bond. The results indicate that the $^1C_4$ conformation may be the starting conformation for α-mannopyranose binding in the active site of GH47 α1,2-mannosidases rather than the $^3S_1$ skew boat conformation.

2.4.7 Interactions and conformation of the bound glycans in the enzyme active site clefts

In the crystal structure of yeast ERManI, a $Man_5GlcNAc_2$ glycan was found docked in the active site cleft. This glycan was devoid of any α1,2-mannose residues indicating that it had been acted upon by an α1,2-mannosidase activity, most likely by the ERManI enzyme activity itself, at some point during preparation [21]. This result is consistent with the low, but detectable,

promiscuous α1,2-mannose cleavage activity for ERManI under conditions of high enzyme concentration [30]. As demonstrated for human ERManI in surface plasmon resonance binding studies [14], the enzyme has an appreciable affinity for the enzymatic product glycan structure and the observation that the yeast ERManI-$Man_5GlcNAc_2$ complex was bound to a glycan product with the B branch (M7 residue) in the +1 subsite provided a structural basis for this interaction.

In contrast to the structure of the yeast ERManI-$Man_5GlcNAc_2$ complex, the glycan found in the human ERManI-$Man_9GlcNAc_2$-PA complex is the intact, uncleaved substrate analog as a result of enzyme inactivation with the bound La^{3+} ion. Three of the four α-1,2 linked mannose residues, M8, M9, and M10, have clear electron densities in the structure, while the M11 residue has no observable electron density as a result of its extension into solvent in the outer part of the glycan binding cleft. Both the PA tag and NAG1 are also not visible in the enzyme-substrate structure, likely as a result of extending into solvent at the top of the enzyme cleft (Figure 27B). Among the terminal α1,2-mannose residues in the $Man_9GlcNAc_2$ substrate glycan, ERManI strongly prefers to cleave the terminal M10 residue on branch B. The glycan structure in the enzyme-substrate complex is consistent with the substrate specificity. Residue M10 from branch B occupies the -1 subsite, and residue M7 on the same branch occupies the +1 subsite. Comparison of the structures of the enzyme-bound $Man_5GlcNAc_2$ product complex with enzyme-bound $Man_9GlcNAc_2$-PA substrate complex was performed by aligning the yeast ERManI structure and human ERManI-$Man_9GlcNAc_2$-PA complex. This comparison demonstrated that the glycan conformations are very similar in the active site clefts of the respective enzymes despite the fact that the sizes of the two glycans are quite different (Figure 27C). All of the α1,3- and α1,6-linked mannose residues occupy similar positions in the two

enzymes, and assume similar conformations. In both cases the enzymes extend branch B of the glycan into the +1 subsite. The three additional α1,2-mannose residues in the ERManI-$Man_9GlcNAc_2$-PA complex extend the termini of the three respective glycan branches: one α1,2-mannose extends into the -1 subsite (M10 residue), and one mannose residue each is found at the ends of branches A (M8 residue) and C (M9 residue). These latter two branches are extended in opposite directions across the face of the glycan binding cleft. Additional hydrogen bonding interactions are observed with residues M8 and M9, while the final residue (M11 residue) is not visible. The similarity in positioning of the glycan substrate and product residues in the active site clefts also indicates that the presence or absence of the α-1,2 linked mannose units do not influence the manner by which the enzymes from two very different species sources preferentially engage branch B in the active site cleft. This specificity is clearly conferred by the core $Man_5GlcNAc_2$ structure in its interaction with the unique topology and interacting residues of the ERManI glycan binding cleft.

Similar to the crystal structure of yeast ERManI in complex with the $Man_5GlcNAc_2$ enzymatic product, the original crystal structure of mouse GMIA also represented an enzyme-product complex [20]. In this latter case the enzyme contains a $Man_6GlcNAc_2$ glycan docked into the enzyme glycan binding cleft, with an extra α1,6-mannose residue that was added during protein secretion by the mannan-extending enzymes of the *P. patoris* expression host [20]. All four of the α1,3- and α1,6-linked mannose residues, the β1,4-mannose residue, NAG1, NAG2 and the additional Man residue are visible in the GMIA crystal structure, but all of the α1,2-linked mannose residues are missing. The absence of α1,2-mannose residues likely reflects the cleavage of these residues by the GMIA enzyme itself similar to the cleavage of the same residues by yeast ERManI (Figure 27D).

Similar to the bound $Man_9GlcNAc_2$-PA structure in the human ERManI-$Man_9GlcNAc_2$-PA complex, the $Man_9GlcNAc_2$-PA in the Golgi enzyme complex is intact, but in this case residues M9, NAG1 and PA are not visible due to extension into solvent. All the other residues of the $Man_9GlcNAc_2$-PA substrate analog are visible in the structure. In contrast to the similarity in glycan ligand binding geometry for yeast and human ERManI, $Man_9GlcNAc_2$-PA and $Man_6GlcNAc_2$ bind in the GMIA glycan binding cleft in very different modes of interaction. Comparison of the complexes of GMIA with $Man_9GlcNAc_2$-PA and $Man_6GlcNAc_2$ by the alignment of protein structures revealed that the two glycans are engaged in the active site clefts by different termini (Figure 27F). In the GMIA enzyme-product complex, the -1 subsite of the enzyme is vacant, similar to the yeast ERManI-product complex, since there are no α1,2-linked mannose residue in the $Man_6GlcNAc_2$ glycan. The +1 subsite is occupied by mannose residue M6 from branch C (Figure 27D). In contrast, $Man_9GlcNAc_2$-PA uses residue M11 from branch A to occupy the -1 subsite and the underlying M8 residue occupies the +1 subsite (Figure 27E). The positioning branch A of the $Man_9GlcNAc_2$-PA glycan in GMIA active site is consistent with the substrate specificity of this enzyme, where residue M11 of branch A is the first residue to be cleaved by the enzyme [3]. Unlike ERManI, which cleaves a single M10 residue from branch B and will only cleave additional α1,2-mannose residues after prolonged incubation at high enzyme concentrations, GMIA is more promiscuous in specificity for α1,2-mannose cleavage. Residue M11 (branch A) is the initial target for α1,2-mannose cleavage, followed by cleavage of M9 (branch C) and subsequently M8 (branch A) in a defined cleavage order[3]. Residue M10 (branch B) is only cleaved after prolonged incubation. The sequential order of glycan binding and cleavage is presumably determined by the affinities of the ≥+1 residues, since all of the termini contain identical α1,2-mannose linkages. Binding studies with ERManI have also

indicated that the major contributor to glycan binding interactions is the +1 residue [14]. Thus, if GMIA was presented with a $Man_8GlcNAc_2$ enzymatic product (with residue M11 cleaved) it would be expected to bind to branch A with high affinity and place the M8 in the +1 subsite. However, the enzyme would eventually cleave α1,2-mannose residues M9, M8, and M10 to result in a $Man_5GlcNAc_2$ product that no longer contains the M8 residue from branch A. Of the termini that are still present in the $Man_5GlcNAc_2$ structure, the M6 residue would reflect a cleavage product of residue M9 of branch C, the second highest affinity residue cleaved by GMIA. The other two termini in the $Man_5GlcNAc_2$ structure (residues M5 from branch A and M7 from branch B) would likely be lower affinity ligands for the enzyme since they reflect products of lower ranking reactions in the cleavage series by the enzyme. Thus, the two enzyme-substrate complexes reflect distinct snapshots of the interactions with the respective enzymes that span from the -1 subsite to the glycan core. These structures expand our understanding of the structural basis of substrate recognition and catalysis for a pair of enzymes that employ similar protein folds to cleave identical substrates, through distinct modes of recognition to provide branch-specific cleavage during glycan processing.

2.4.8 Interactions between the glycan residues and the enzyme active site cleft.

The interactions between the enzymes and the respective glycan substrate analogs or products are quite different, particularly for GMIA that has distinct modes of interaction with glycan branches in the two structures. For example, in the GMIA enzyme-substrate complex, all three of the terminal α1,2-linked mannose residues that can be resolved (M11, M10, and M8) interact with the enzyme. Two additional mannose residues (M5, M6) and NAG2 also are involved into the interactions with the enzyme (Figure 28C). In the GMIA enzyme-product complex, with all α1,2-mannose residues trimmed, the enzyme employs residues M3, M4, M5, M6, M7, NAG1

and NAG2 to participate the enzyme interactions (Figure 28 B). In the ERManI-$Man_9GlcNAc_2$-PA complex, the oligosaccharide uses all of the residues (except for M11 that is not visible in the structure) to interact with enzyme (Figure 28C). Similarly, in the yeast ERManI-product complex, all the residues of $Man_5GlcNAc_2$-PA participate the interactions with the enzymes (Figure 27C) [21].

2.4.8 Interactions in the -1 subsite.

The sequence and structural alignments of the GH47 α1,2-mannosidases illustrate that their catalytic sites are highly similar and can generally be superimposed regarding all interactions in the -1 subsite. Substrate interactions in the -1 subsite for human ERManI and mouse GMIA reflect this similarity with identical interactions in the -1 subsite. The hydrogen bonding and hydrophobic stacking between the two enzymes and $Man_9GlcNAc_2$-PA are illustrated in Figure 27E. In addition to the direct coordination of the C-2 and C-3 hydroxyl groups with the enzyme bound La^{3+} ion, an arrays of hydrogen bonds and hydrophobic interactions facilitates the interaction with the α-1,2 linked mannose in a $^1C_4$ conformation. The O-ε2 of Glu611 (Glu663 in human ERManI) interacts with C3' hydroxyl group, the NH2 of Arg286 (Arg334 in human ERManI) and O-ε1 of Glu636 (Glu689 in ERManI) interact with the C4' hydroxyl group, the NH1 of Arg547 (Arg597 in ERManI) and O-ε2 interact with the C6' hydroxyl group. The O-γ and the backbone carbonyl oxygen atom of Thr635 (Thr688 in human ERManI) also interact with C2' and C3' hydroxyl groups. The side chain of Phe607 (Phe659 in human ERManI) has hydrophobic stacking interaction with the C4'-C5'-C6' of the mannose residue. All the interactions work together to distort the mannose residue into the $^1C_4$ conformation in the -1 subsite. Numerous interactions preclude the ability to accommodate a $^4C_1$ conformation, most predominately Phe607 that interacts with the axial C4' and C5' carbon atoms through

hydrophobic interactions that would provide a steric barrier for equatorial positioning of hydroxyl residues in the -1 subsite.

2.4.9 The interactions in the ≥+1 subsite

Human ERManI and mouse GMIA not only share similar interactions with the mannose residue in the -1 subsite, but also share the similar interactions with the mannose residue in the +1 subsite. Residue M8 in the GMIA enzyme-substrate complex uses its C-3 and C-4 hydroxyl groups to interact with the OD1 and OD2 of Asp415, and the C-6 hydroxyl group to interact with OE1 of Glu351, as well as the C-4 hydroxyl group to interact with backbone NH of Leu413. Similarly, M7 in the human ER-substrate complex also uses the C-3, C-4 and C-6 hydroxyl groups to interact respectively with the OD1, OD2 of Asp463 and OE1 of Glu397, as well as C-4 hydroxyl group to interact with the back bone NH of Arg461.

However, the differences in interactions with >+1 residues of $Man_9GlcNAc_2$-PA are what distinguish the enzyme-substrate interactions (Table 2). First, there are more enzyme-oligosaccharide interactions in the human ERManI-$Man_9GlcNAc_2$-PA complex than in the GMIA-$Man_9GlcNAc_2$-PA complex. In GMIA, 7 amino acid residues (His230, Ser232, Ser233, Glu351 Leu413, Asp415, and Arg547) have polar interactions with the oligosaccharide analog as identified using Pymol software. Those seven residues form eleven hydrogen bonds with the $Man_9GlcNAc_2$-PA. In contrast, in the human ERManI-substrate complex, twelve residues (Asn327, Ser375, Asp376, Arg388, Glu397, Asp461, Asp463, Lys521, Asp523, Glu570, Asp591, Arg597) were found to generate 21 hydrogen bonds with $Man_9GlcNAc_2$-PA. In the human ERManI enzyme-substrate complex, Trp389 has hydrophobic stacking with the nonpolar face of the NAG2 residue of $Man_9GlcNAc_2$-PA (Table 2 and Figure 28C). Similarly, in the GMIA enzyme-product complex structure, the NAG2 of $Man_6GlcNAc_2$ and Trp339 also share

hydrophobic stacking interactions (Figure 28B). However, the GMIA enzyme-substrate complex does not appear to employ hydrophobic stacking between Trp389 and the NAG2 because of repositioning of Trp389 to flip away about 180°, otherwise it would clash with NAG2 of $Man_9GlcNAc_2$-PA in the complex.

Second, in the human ER ManI-$Man_9GlcNAc_2$-PA complex there are more glycan residues directly involved in enzyme-substrate interactions than in the GMIA-$Man_9GlcNAc_2$-PA complex (Table 2). In the human ERManI-$Man_9GlcNAc_2$-PA structure, mannose residues on all three branches and NAG2 in the glycan core participate in the interactions with the enzyme. However, in the GMIA enzyme-substrate complex, the residues M10, M7, M3 and NAG2 of $Man_9GlcNAc_2$-PA do not have interactions with the enzyme residues. Branch B does not interact with GMIA and all the hydrogen bonding interactions come from branch A and C of the glycan structure (Figure 28C and 28D).

Third, the amino acid residues in human ERManI have bifurcating hydrogen bonding interactions with multiple oligosaccharide residues, whereas GMIA exhibits only single hydrogen bonding interactions (Table 2). In the GMIA-$Man_9GlcNAc_2$-PA complex, Glu351 Leu413, Asp415, and Arg547 only interact with residues on the branch A, while His230, Ser232 and Ser233 only interact with residues on the branch C. In human ERManI-$Man_9GlcNAc_2$-PA complex, Arg461 interacts not only with the C-4 hydroxyl of M7 on branch B, but also with the C-2 hydroxyl of M3, the C-3 hydroxyl of NAG2, as well as the glycosidic oxygen atom between M3 and NAG2 in the glycan core. Also in the human ERManI enzyme-substrate complex, Arg597 interacts with the C-6 hydroxyl of M10 of branch B, the C-5 hydroxyl of M9, and the C-4 hydroxyl of M4. Residues M10 and M9 are at the termini of branches B and C respectively (M10 is in the -1 subsite) while M4 is near the glycan core. Thus, interactions of Arg597 in

human ERManI extends from the -1 to the >+1 subsites, and from terminal α1,2-mannose linkages to the glycan core.

As described above, the substrate and product complexes with yeast and human ERManI are strikingly similar despite the presence of interactions with the additional α1,2-mannose residues in the enzyme-substrate structure, the differences in occupancy of the -1 subsite between the two enzymes, the open ring structure of NAG1 in the $Man_9GlcNAc_2$-PA complex versus an intact reducing terminus for the enzyme product complex, and the primary sequence differences between the yeast and human enzymes. In contrast, the GMIA enzyme-substrate and enzyme-product are strikingly different as the consequence of three different aspects of the two complexes. First the orientations of the $Man_9GlcNAc_2$-PA and $Man_6GlcNAc_2$ in the enzyme active sites are quite different, with branch A in the -1/+1 subsites for the enzyme-substrate complex versus branch C in the +1 subsite for the product complex. Second, the presence of the extra α1,6-mannose residue and absence of the α1,2-linked mannose residues in $Man_6GlcNAc_2$ product complex is distinct from the extended $Man_9GlcNAc_2$-PA substrate complex. Third, the positioning change of Trp389 alters the interactions with the glycan core. The amino acid residues that participate in the enzyme-oligosaccharide interactions in both the enzyme-substrate and enzyme-product complexes are: Ser233, Glu351 Leu413, and Asp415, but each of them interacts with different oligosaccharide residues in the two complexes. Some amino acid residues and oligosaccharide residues participate uniquely in either the enzyme-product complex or the enzyme-substrate complex, but not both. Trp341 only interacts through hydrophobic stacking with the NAG2 of $Man_6GlcNAc_2$ in the enzyme-product complex, but not the NAG2 of $Man_9GlcNAc_2$-PA in the enzyme-substrate complex. The backbone carbonyl oxygen and the side chain hydroxyl group of Ser455 interact, respectively, with the O-5 and C-6 hydroxyl group of

the extra mannose residue of the $Man_6GlcNAc_2$, but not with any residue of $Man_9GlcNAc_2$-PA. Glu512 interacts with the NAG1 of $Man_6GlcNAc_2$ in the enzyme-product complex, but does not contribute to interactions in the GMIA substrate complex. In contrast, GMIA employs different residues to interact with $Man_9GlcNAc_2$-PA compared with $Man_6GlcNAc_2$. His230 and Ser232 interact with the M6 residue of $Man_9GlcNAc_2$-PA, but not with the $Man_6GlcNAc_2$ structure. Arg547 interacts not only with the C-6 hydroxyl of the M11 in the -1 subsite, but also with the C-3 hydroxyl of the M5, but has interactions with $Man_6GlcNAc_2$ in the enzyme-product complex.

2.5 Discussion

2.5.1 The replacement of Ca^{2+} by La^{3+}

Lanthanides have been extensively used to probe alkaline earth metal-binding sites and Ca^{2+}-binding sites in particular [31-40]. Among the lanthanide family, lanthanum ion (La^{3+}) exhibits similar physical and chemical properties with Ca^{2+}, including a similar ionic radius (1.17 and 1.14 Å, respectively, for six coordinated metal ions, and 1.30 and 1.26 Å, respectively, for eight coordinated metal ions) [41, 42]. La^{3+}, like Ca^{2+}, is also a hard Lewis acid, which prefers to bind hard bases containing oxygen (–COOH, H_2O, -OH) [43]. The bonding of coordinated La^{3+}, also like that of Ca^{2+}, is essentially ionic. La^{3+} appears to be an ideal biomimetic agent for Ca^{2+} in the study of enzyme-bound metal ion replacements. However, La^{3+} also possesses distinct properties from Ca^{2+}. La^{3+} is a stronger Lewis acid than Ca^{2+} [44]. Trivalent La^{3+} has higher affinity towards oxygen-containing biological ligands than divalent Ca^{2+}, therefore can bind the ligands more tightly. The observed coordination number of Ca^{2+} in proteins varies from 6 to 9, while the coordination number of La^{3+} varies from 8 to 9 in Protein Data Bank. Regarding the biological functions of La^{3+}, it has been reported that La^{3+} is capable of competing with the Ca^{2+} for the

same binding site in the Ca^{2+}-dependent guanylate cyclase and inactivates the enzyme [34, 35, 37, 38]. In addition, La^{3+} exhibits dual effects on the activities of some enzymes, such as horseradish peroxidase and acetylcholine acetylhydrolase. At low concentration La^{3+} inactivates those enzymes, but at high concentration and prolonged exposure, will inhibit the enzymes [37, 45].

Two crystal structures representing enzyme-substrate complexes of human ER α-mannosidase I and mouse Golgi α-mannosidase IA are presented in this paper. Both the enzymes are GH47 α-mannosidases, cleave α1,2-mannose linkages with the same inverting mechanism, have similar overall structures, and interact with substrates through almost identical residues in the -1 and +1 subsites. However, ERManI and GMIA exhibit distinct and complementary specificities towards $Man_9GlcNAc_2$. The two enzymes were co-crystallized with the natural substrate analog $Man_9GlcNAc_2$-PA following the replacement of the essential enzyme-bound Ca^{2+} ion with La^{3+}. The structures of the two enzyme-$Man_9GlcNAc_2$-PA complexes were solved by molecular replacement and the important characteristics, including the coordination of the enzyme-bound metal ion, the conformation of the glycone in the -1 subsite, the orientation of the enzyme-bound oligosaccharide, as well as the enzyme-oligosaccharide interactions, were compared between the two enzyme-substrate complexes and the previously determined enzyme-product complexes of the enzymes. Significant insights regarding the structural and functional relationships of the two enzymes were drawn from the comparisons.

The replacement of the enzyme-bound Ca^{2+} ion with La^{3+} not only inactivated the two purified enzymes, but also changed the metal ion coordination in the active site. However, whether the inactivation of the enzyme was related to the change of the coordination of the enzyme-bound metal ion or differences in electrostatics is still unknown. The mechanism of how

La^{3+} influences the enzyme activities of ion-bound complexes has not been systematically studied or reported on at a molecular level. Compared to the 2^+ charge of the calcium ion, lanthanum ion carries a 3^+ charge and may influence the electrostatic environment of the active site differently from Ca^{2+}. Also La^{3+} is a much stronger Lewis acid and would have a stronger affinity toward the coordinated water molecules. According to the proposed catalytic mechanism [19], one Ca^{2+}-coordinated water serves as the nucleophile in the catalytic mechanism, where it attacks the C-1 of the mannose residue in the -1 subsite to result in glycosidic bond cleavage. A tighter interaction between the nucleophilic water and the La^{3+} ion may prevent the water from leaving its coordinated position to fulfill its role as a nucleophile, thus inactivating the enzyme.

Mapping of the conformational itinerary of the glycone in the -1 subsite is important for studies on the catalytic mechanism of the α1,2-mannosidases. Previous studies revealed a partial conformational itinerary map including proposed intermediate, transition state, and enzymatic product structures in the -1 subsite, but not the starting conformer in this subsite [16, 19]. Two enzyme-inhibitor complex structures (ERManI with bound dMNJ or Kif) revealed that the six-membered sugar ring resides in the -1 subsite in a $^1C_4$ conformation, in which the C-3 and C-4 hydroxyl and the C-5 hydroxymethyl groups all take axial positions. It was also proposed that the *in situ* cleaved β-D-mannopyranose product takes on a $^1C_4$ conformation [19]. This proposal was derived from the fact that dMNJ is the analogue of D-mannopyranose and is not attached to any ligand residues in +1 subsite, so the $^1C_4$ conformation of dMNJ in the -1 subsite may reflect the conformation of a β-D-mannopyranose product. The $^3S_1$ conformation of the glycone residue in the thiodisaccharide substrate analog in the human ERManI active site was proposed as the conformation of a pre-transition state intermediate, which would proceed through a short-lived $^3H_4$ transition state to eventually become the $^1C_4$ product [19]. In this proposed conformation

itinerary progressing from $^{3}S_{1}$ pre-transition state to a $^{3}H_{4}$ transition state and a $^{1}C_{4}$ released product, the starting conformer would presumably be close to $^{3}S_{1}$.

In our study, the structures of the GMIA-$Man_9GlcNAc_2$-PA and ERManI-$Man_9GlcNAc_2$-PA complexes provide experimental evidence that the glycone mannose residue can bind within the -1 subsite in a $^{1}C_{4}$ starting conformation. The predicted $^{3}S_{1}$ conformer was not found in the -1 subsite of either enzyme-substrate complex. In comparison with the structure of the dMNJ complex, the α1,2-linkage between the mannose residues in the -1 and + 1 subsites does not appear to influence the conformation of the mannose residues. In contrast, comparison of the enzyme-substrate complex to the ERManI-thiodisaccharide complex suggests that the longer length of the C-S-C bond in the thiodisaccharide complex may contribute to distortion of the glycone into the $^{3}S_{1}$ conformation.

By comparing the structures of the different conformers in the -1 subsite, and analyzing the interactions that the enzyme can apply to those conformers, we gain more understanding of the structural basis for $^{1}C_{4}$ as the starting conformer. The structural difference between the two conformers ($^{1}C_{4}$ and $^{3}S_{1}$) found in the -1 subsite is the angle of C1-O5-C5 linkage as it flips in opposite directions between the two conformers (Fig 26C). Therefore, the relative position of O5 to other atoms in the sugar ring are different in $^{1}C_{4}$ and $^{3}S_{1,}$ but the relative positions of C-1, C-2, C-3, C-4, C-5 and C6 atoms and their substituents in the six-member ring are essentially identical. The enzyme-ligand interactions are identical for $^{1}C_{4}$ and $^{3}S_{1}$. In contrast, the $^{4}C_{1}$ conformation is drastically different from the $^{1}C_{4}$ conformation, with the configurations of all of the ring-atoms altered and a compromising of anchoring interactions within the enzyme active site. The equatorial C2-hydroxyl and axial C-3 hydroxyl of $^{1}C_{4}$ mannose, together with the two oxygens of Thr (Thr688 in ERManI, Thr635 in GMIA) comprise one side of the 8-fold square

antiprismatic coordination with the calcium ion. By comparison, the axial C-2 hydroxyl and equatorial C-3 hydroxyl of $^4C_1$ conformer would not be predicted to interact with the Ca^{2+} ion without the deformation of the conserved coordination geometry. Similarly, hydrophobic interactions between the Phe (Phe659 in human ERManI, Phe607 in mouse GMIA) and the axially positioned C-4 and C-5 in a $^1C_4$ glycone could not be accommodated by the equatorial positions of a $^4C_1$ conformation. In summary, the conserved active site residues of the enzymes are only able to anchor a glycone structure containing the equivalent of $^1C_4$ conformation at C2-C3-C4-C5-C6. The structure of the enzyme-substrate complex favors a $^1C_4$ conformation as an initial binding state in the catalytic reaction, but cannot rule out the possibility of a $^3S_1$ structure during the conformational itinerary of catalysis.

All of the residues that interact with the mannose residue in the -1 subsite are also conserved between ERManI and GMIA. The computational studies by Mulakala et al [27-29] suggested that the conserved +1 subsite of the ERManI is optimal for interaction with a $^4C_1$ conformer of mannose and the structural studies presented here indicate that all of the $\geq$+1 subsite mannose residues assume a standard $^4C_1$ conformation. How and where the mannose conversion of the glycone from $^4C_1$ to $^1C_4$ takes place is still not clear and remains an interesting topic for further speculation. Mulakala et al. [29] pointed out that the funnel-shaped neck of the active site is too narrow to allow the passage of the flat and extended ground-state $^4C_1$ mannose residue with most of its ring substituents in an equatorial position. But the constricted neck region can accommodate the more globular and compact $^1C_4$ mannose with its ring substituents predominantly in the axial position. The funnel-shaped neck is composed of residues that form the inner layer of the $(\alpha\alpha)_7$ barrel domain and is more rigid than the loop region on the outer parts of the enzymes. In our study, comparisons of the structures of ERManI-DMNJ and

ERManI-S-Man_2, with ERManI-$Man_9GlcNAc_2$-PA and GMIA-$Man_9GlcNAc_2$-PA illustrate that the -1 to +1 subsites are superimposable regardless of the binding of the ligands, and no apparent size change can be observed in the lower part of the cleft. For all the structures listed above, the low B-factor values (<25 $Å^2$) of the residues in this area suggest that their side chains have a low range of movement. The funnel-shaped neck is constricted as predicted by the computational study, and its expansion to allow the passage of the substrate would be energetically disfavored.

For mannose in solution, the $^4C_1$ and $^1C_4$ conformers reach equilibrium with 99% of the population in the $^4C_1$ conformation, and <1% in $^1C_4$ conformation [29]. It was suggested that the enzyme active site selectively binds and cleaves the $^1C_4$ α1,2-linked mannose on the non-reducing end of the glycan branches, and continues to push the equilibrium toward the $^1C_4$ conformer [29]. The relaxation time of the conformational change between $^1C_4$ and $^4C_1$ is of the order of microseconds [46], so diffusion of the bulky substrate into the active site would likely be the rate-limiting step in the reaction. The low percentage of the $^1C_4$ conformer might be the reason that α1,2-mannosidase is considered to be a slow-acting timer in glycoprotein degradation [29, 47, 48].

Enzyme Specificity towards α-1,2 mannosyl linkage.

A first level set of studies examining structural and functional relationships on the enzymology of a metabolic process, including that of the α1,2-mannosidases, is to explore how the enzymes discriminate among different mannosyl linkages. Although the sequence similarity of the human ERManI and mouse GMIA is less than 30%, the amino acid residues that interact with the substrate in the -1 and +1 subsites are identical and the enzyme-ligand interactions found in the two subsites are also similar among all the structures found. One scenario can be envisaged where neither Man-α1,3-Man nor Man-α-1,6-Man can fit into the conserved -1 and +1

subsites without disturbing or compromising those observed interactions. If Man-α1,3-Man or Man-α-1,6-Man glycans bind to the active site with the non-reducing end mannose positioned as the $^{1}C_{4}$ conformer of the Man-α1,2-Man structure, then the positioning of the reducing-end mannose in the +1 site will be entirely different than in the Man-α-1,2-Man substrate complex. The α1,2-glycosidic bond is critical for the positioning of the reducing-end mannose in the -1 subsite and forming the stabilizing interactions with conserved residues in the +1 subsite. Prior binding energetics studies demonstrated that >50% of the binding energy for the entire glycan structure is derived from interactions in the +1 subsite [14]. In contrast, mannosyl linkages through α1,3- or α1,6-glycosidic bonds would alter the positioning of the residue in the +1 subsite because of the steric constrains and result in the loss of key stabilizing interactions . This implies that the conserved residues in the -1 and +1 subsites work in concert to anchor the Man-α1,2-Man disaccharide through the observed combination of interactions, including direct coordination with an enzyme-bound Ca^{2+} ion, a collection of hydrogen bonding interactions in the -1 and +1 subsites, and hydrophobic interactions in the -1 subsite. In the past, mutation studies have been focused on the enzyme active site in order to map out the catalytic residues. More mutation studies should be carried out on those conserved residues in the +1 subsite to test their influence on the enzymatic activity and specificity towards Man-α1,2-Man, Man-α1,3-Man, and Man-α1,6-Man substrates. Different isomers of $Man_8GlcNAc_2$ were modeled into the binding cleft of yeast and human ERManI structures and the results showed that no allowed conformation could be found where branch A docked into the either yeast or human ERManI, nor could branch C dock into the yeast enzyme [15]. We expect that efforts to test the computational docking of Man-α1,3-Man, and Man-α1,6-Man into the -1 and +1 subsites of both ERManI and GMIA will shed more light onto whether and how the conserved residues in the

two subsites exclusively accommodate only Man-α1,2-Man, but not Man-α1,3-Man or Man-α1,6-Man structures.

The substrate specificity differences of GMIA and human ERManI.

The different orientations of the substrate analog, $Man_9GlcNAc_2$-PA, in the enzyme active sites provide direct evidence that mouse GMIA and human ERManI have different specificities towards the three branches of $Man_9GlcNAc_2$ glycans on glycoprotein substrates. The second level of structural and functional relationship studies for α1,2-mannosidases is to explore the structural basis for substrate priority. How do the two enzymes recognize and bind the same substrate differently to engage different branches in the respective active sites? The detailed comparison of the enzyme-ligand interactions in the GMIA-$Man_9GlcNAc_2$-PA and ERManI-$Man_9GlcNAc_2$-PA structures reveals that some enzyme residues are different in the equivalent positions within ligand binding cleft even though the protein backbone has a similar protein fold. There are differences in loop positions on the edge of the glycan binding clefts that vary between the two enzymes, and as a result the enzyme-substrate interactions involving those residues are clearly different between the two enzymes. Among those residues, one pair of the residues in the equivalent positions, Leu413 in GMIA and Arg463 in the ERManI, have been particularly interesting. The Arg463 residue in ERManI interacts with three glycan substrate residues: M3 and NAG2 in the glycan core and M7 on branch B in the enzyme-substrate complex. In contrast, Leu413 in GMIA only interacts with M8 on branch A. *S. cerevisiae* ERManI has an equivalent Arg residue at this position (Arg273) and the enzyme has a specificity similar to human ERManI. In contrast, the *T. reesei* and the *P. citrinum* α-mannosidases have less bulky residues in this position (Leu243 and Gly265, respectively) and exhibit abroad substrate specificity similar to GMIA [15]. Human and yeast ERManI, employ the respective Arg residue to interact with M3

and NAG2 in the glycan core and M7 of branch B and result in the narrow specificity for cleavage of the M10 residue in branch B, but poorly cleave the α-mannose residues M11, M9, and M8 on branches A and C [49-51]. Mutation of Arg273 to Leu in the yeast enzyme resulted in an enzyme with low enzymatic activity and a hybrid specificity between ERManI and GMIA [24]. The yeast ERManI mutant R273L gains a broadened specificity that enables it to cleave $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$ by the sequential removal of M10, M11, and M8, and finally M9 [24]. The bulky Arg residue in this position is believed to help guide the middle branch into the active site by anchoring the glycan core M3 and NAG2 with hydrogen bonding in contrast to the smaller Leu or Gly residues that confer greater flexibility to the substrate in its binding process. In order to explore further the Arg *vs* Leu effect on substrate binding, it is necessary to swap the Arg461 and Leu413 in human ERManI and GMIA, and to test whether the human ERManI mutant R461L or/and the GMIA mutant L413R will also demonstrate hybrid specificity, and how the residue mutation will influence substrate binding affinity.

In addition to the differences in the Arg461/Leu413 residues in the binding clefts of ERManI and GMIA, there are also other residues located in nearly equivalent positions in the two enzymes that interact with $Man_9GlcNAc_2$-PA differently. Both Ser375 and Asp376 in human ERManI interact with M5, while the corresponding residues in GMIA (Ala327 and Leu328) have no interactions with the substrate. Asp523 and Asp 591 in ERManI interact with M6 and M9 of $Man_9GlcNAc_2$-PA, respectively, but their counterparts in GMIA (Gly474 and Glu541) have no interactions with $Man_9GlcNAc_2$-PA.

On the edge of the binding cleft, Ala590 in human ERManI and Asn540 in GMIA both interact with $Man_9GlcNAc_2$-PA, but because of the differences in their side chain structures and size, Ala590 uses the backbone carboxyl oxygen to interact with M9, while Asn540 uses the side

chain NH to interact with M10. Compared to the core of the binding site, the enzyme sequences and protein folds on the edge of the cleft are more varied between ERManI and GMIA, and some of the enzyme-substrate interactions are unique to each enzyme. The residues His230, Ser232, and Ser233 of GMIA interact with glycan residues M6 and M9 of $Man_9GlcNAc_2$-PA, whereas the equivalent region of the human ERManI-$Man_9GlcNAc_2$-PA complex has a different loop fold and neither equivalent residues nor any interactions with the bound substrate can be found.

Another enzyme residue likely to contribute to the specific substrate interactions is the Trp389 in human ERManI. The Trp389 side chain stacks on the nonpolar face of NAG2 of $Man_9GlcNAc_2$-PA with hydrophobic interactions in the enzyme-substrate complex structure. In contrast, the GMIA-$Man_9GlcNAc_2$-PA complex does not have this hydrophobic stacking interaction, since the side chain of Trp341 flips away from the NAG2 to avoid a steric clash. NAG2 is located in the core of the glycan structure and interactions between the enzyme and NAG2 can aid in glycan orientation within the active site cleft. However, the Trp-NAG2 interaction is not the determining factor for the glycan interaction. The comparison of the yeast ERManI-glycan product structure and the human ER-$Man_9GlcNAc_2$-PA substrate complex structure shows that the glycan structures have similar orientations and interactions, except for the lack of α1,2-linked mannose residues in the $Man_5GlcNAc_2$ complex. However, the yeast enzyme does not appear to employ the Trp residue in hydrophobic stacking interactions with NAG2 and this does not prevent the yeast enzyme from sharing the same glycan binding pattern with human ERManI. This implies that the Trp-NAG2 hydrophobic stacking is only a contributing factor in anchoring interactions at the glycan core in addition to other interactions such as Arg461 for human ERManI and Arg273 in the yeast enzyme.

In contrast to the yeast ERManI-glycan complexes where the intact substrate complex employs a hydrophobic stacking with NAG2 and the $Man_5GlcNAc_2$ enzyme product does not, the opposite scenario is true in the complexes of GMIA with its respective substrate and product. The GMIA-$Man_9GlcNAc_2$-PA substrate complex structure does not employ Trp341 for interactions at the glycan core, whereas the GMIA enzyme-product complex employs hydrophobic stacking interactions between Trp341 and NAG2. The absence of an anchoring interaction on the glycan core in the GMIA-$Man_9GlcNAc_2$-PA complex might explain the broader specificity of the enzyme toward cleavage of terminal α-mannose linkages, but does not explain the sequential nature of the defined isomer series produced in the cleavage of $Man_9GlcNAc_2$ to $Man_6GlcNAc_2$ and the inefficient cleavage of the terminal branch B (M10 residue) to generate the final $Man_5GlcNAc_2$ structure. In addition, the presence of the anchoring hydrophobic stacking interactions between Trp341 and NAG2 in the GMIA-$Man_5GlcNAc_2$ product complex suggests that interactions with the glycan core may contribute to cleavage of the C branch residue, the second step in trimming of $Man_8GlcNAc_2$ to $Man_7GlcNAc_2$. Thus, a complex series of interactions within the active site cleft of the GH47 α-mannosidases provides a selective pathway for trimming of individual glycan branch termini that depends on the geometry and characteristics of a given enzyme active site cleft and the structure of the trimmed glycan intermediate.

In this work, the crystal structures of the enzyme-substrate complexes of both human ERManI and GMIA have been generated and the comparison of the substrate orientations and the enzyme-substrate interactions have been described in detail. The contributions of individual residues to distinct substrate recognition, binding, and catalysis in their complementary substrate specificities remain an interesting subject for further study.

Table 1. X-ray crystallography diffraction Data collection and refinement statistics

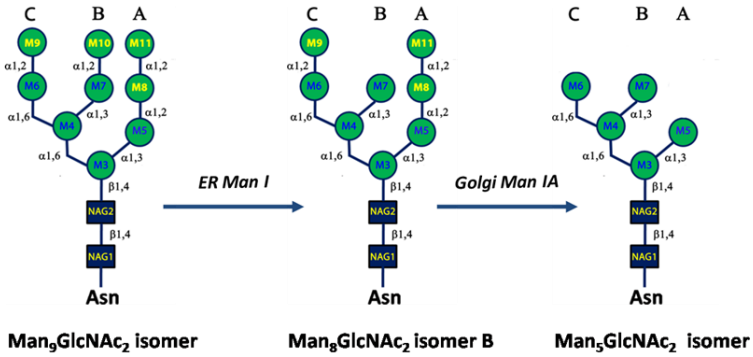
Parameter	Value (human ERManI)	Value (GMIA)
Data resolution (high resolution shell) Å	48.68-1.65	41.76-1.77
Unique observed reflection	50419	96347
Completeness (%)	96.53	94.95
<I>/<σ>>	24.1	26.36
Space group	P1	P 21 21 2
Unit cell lengths *a, b, c* (Å)	50.62, 53.53, 56.12	94.95, 131.48, 87.78
Unit cell angles *α*, β, γ (°)	89.65, 63.60, 62.70	90.00, 90.00, 90.00
Number of reflection in free set	3518	5169
R_{cryst} (%)	19.39	19.72
R_{free} (%)	24.08	23.95
Number of non-hydrogen atoms refined	4302	8264
r.m.s. deviation (Å) Bonds	0.011	0.027
r.m.s. deviation (°) Angles	1.36	2.16
Mean temperature factor (Å^2)	8.99	20.36

Table 2: Comparison of the enzyme-substrate interactions. The arrangement of the $Man_9GlcNAc_2$-PA substrate in the binding sites of ERManI and GMIA are different, as a result the two enzymes interact with the substrate differently. The enzyme-substrate interactions in the ≥+1 substrate-binding subsites of human ERManI and mouse GMIA were mapped respectively by using the "Finding Polar Contacts" function of Pymol software using the respective enzyme-$Man_9GlcNAc_2$-PA complex structures. The interactions were summarized and the amino acid residues of ERManI and GMIA are listed in the left column and the third column, respectively. The atoms of the amino acid residues and substrate residues that are involved in the interactions are listed in the adjacent columns for each enzyme. The topologically equivalent residues in ERManI and GMIA were listed in the same row of the respective columns. The nomenclatures of the residues and atoms are consistent with those in the PDB files of the enzyme-substrate complexes.

ERManI	Interaction in ER enzyme	GMIA	Interaction in Golgi enzyme
Ser282	no interaction	His230	NE2 interacts with O6 of M6
	non-equivalent	Ser232	OG interacts with O5, O6 of M6
	non-equivalent	Ser233	N interact with glycosidic O link between M6 and M9
Asn327	ND2 interacts with O4, O6 of M5	Ser279	no interaction
Ser375	OG interacts with O6 of M5	Ala327	no interaction
Asp376	OD1 interacts with O4, O6 of M5	Leu328	no interaction
Pro387	no interaction	Trp339	no interaction, but clash with M5 if MAN9 fits in like in ER enzyme
Arg388	NH1 interacts with O3 of M8	close to Pro340	no interaction, the loops of the two enzyme do not overlap
Trp389	hydrophobic interaction with NAG	Trp341	no interaction,
Glu397	OE2 interacts with O6 of M7	Glu351	OE1 interacts with O6 of M8
Arg461	NH2 interacts with O2 of M3	Leu413	backbone NH interacts with O4 of M8
	NH2 interacts with O3 of NAG2		
	NH2 interacts with glycosidic O link beteween M3 and NAG2		
	N interacts with O4 of M7		
Asp463	OD1 Interacts with O4 of M7	Asp415	OD1 interacts with O4 of M8,
	OD2 interacts with O3 of M7		OD2 interacts with O3 of M8
Lys521	NZ interacts with O4 of M6	Lys472	no interaction
Asp523	OD1 interacts with O4 of M6	Gly474	no interaction
Glu570	OE1 interacts with O6 of M6	Glu524	no interaction (interact through H2O with M5)
Ala590	backbone O interacts with O4 of M9	Asn540	NH2 interacts with O3 of M10
Asp591	OD1 interacts with O3,O4 of M9	Glu541	no interaction
Arg597	NH2 interacts with O4 of M4	Arg547	NH1 interacts with O6 of M11, NH2 interacts with O3 and O4 of M5

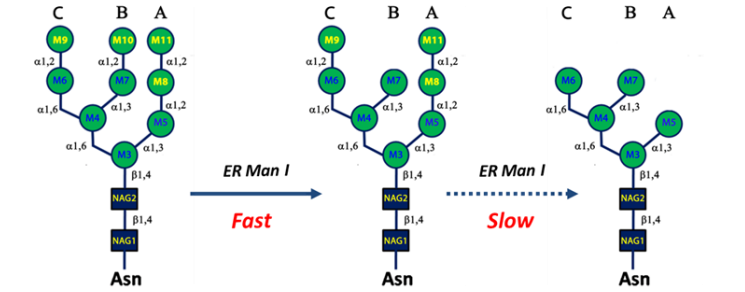
Figure 23. Schematic diagram of the $Man_9GlcNAc_2$ processing steps catalyzed by ER mannosidase I and Golgi mannosidase IA *in vivo* and *in vitro*. *In vitro,* $Man_9GlcNAc_2$ was processed by ERManI to $Man_8GlcNAc_2$ isomer B rapidly, followed by a slow cleavage to $Man_5GlcNAc_2$. Trimming $Man_9GlcNAc_2$ by the Golgi ManI subfamily of enzymes generates $Man_6GlcNAc_2$ rapidly, but further cleavage to $Man_5GlcNAc_2$ occurs slowly. *In vivo,* ERManI cleaves $Man_9GlcNAc_2$ to produce $Man_8GlcNAc_2$ isomer B, followed by the action of the Golgi α-mannosidases (GMIA) to produce $Man_5GlcNAc_2$. The GlcNAc residues are indicated by red shaded squares. Mannose residues are green shaded circles. The identities of mannose residues are labeled with capital letter M followed by numbers (M3-M11). The GlcNAc residues were labeled with NAG1 and NAG2. The three branches of the $Man_9GlcNAc_2$ are labeled as branches A, B, and C.

ER ManI and ***Golgi Man IA*** digestion in *vivo*



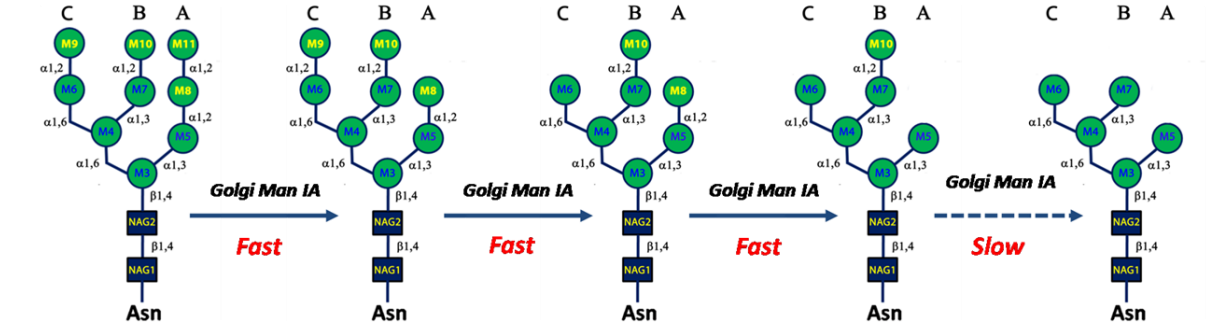
$Man_9GlcNAc_2$ isomer $Man_8GlcNAc_2$ isomer B $Man_5GlcNAc_2$ isomer

ER ManI digestion in *vitro*



$Man_9GlcNAc_2$ isomer $Man_8GlcNAc_2$ isomer B $Man_5GlcNAc_2$ isomer

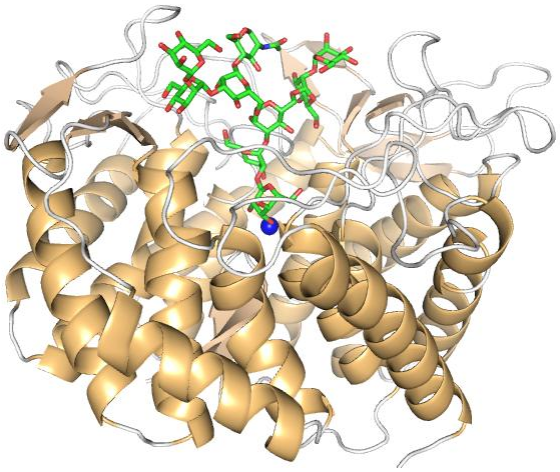
Golgi Man IA digestion in *vitro*



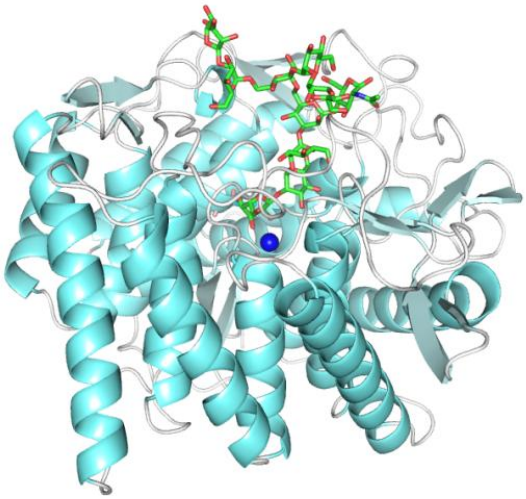
$Man_9GlcNAc_2$ isomer $Man_8GlcNAc_2$ isomer A $Man_7GlcNAc_2$ isomer $Man_6GlcNAc_2$ isomer $Man_5GlcNAc_2$ isomer

Figure 24. Structural models of ERManI and GMIA. Schematic ribbon representations are shown for **Panel A:** human ERManI in complex with $Man_9GlcNAc_2$-PA; **Panel B**: deglycosylated GMIA in complex with $Man_9GlcNAc_2$-PA; **Panel C**: two yeast ERManI protein units in the crystal lattice showing cross-linking of enzyme structures by the docking of the glycan enzymatic product in the active site of the adjoining unit in the crystal lattice; **Panel D**: two murine GMIA protein units in the crystal lattice showing cross-linking of enzyme structures by the docking of the glycan enzymatic product in the active site of the adjoining unit in the crystal lattice. The Ca^{2+} ion bound at the base of the barrel is shown as an olive sphere, and the La^{3+} ion is shown as a blue sphere. All Asn-linked oligosaccharides ($Man_6GlcNAc_2$ and $Man_9GlcNAc_2$–PA) are shown in green stick representations. In the crystal lattice in **Panels C** and **D**, the Asn-linked oligosaccharide structure bridges between protein subunits and docks into the funnel-shaped opening at the end of the barrel of an adjoining enzyme molecule in a presumed enzyme-product complex. Models in **Panels A** and **B** were generated from the X-ray diffraction data directly. Models in **Panels C** and **D** were modified from the PDB entries 1X9D and 1FO2 respectively.

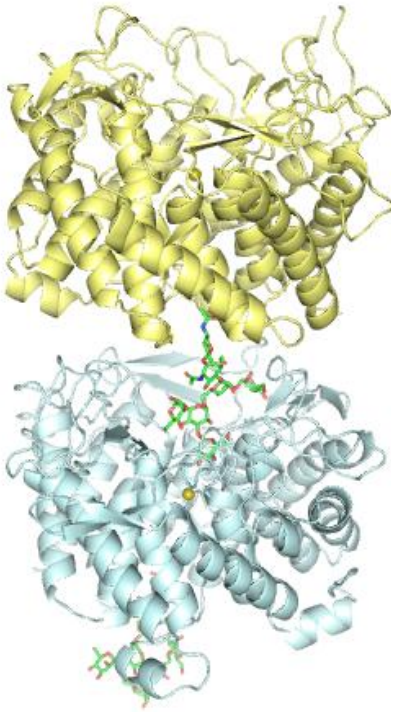
A



B



C



D

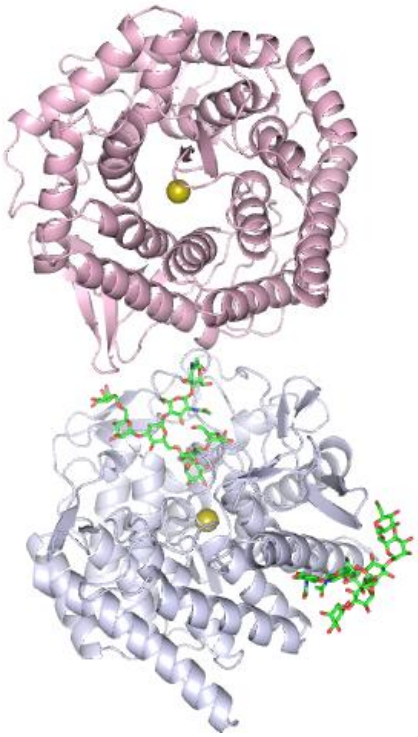
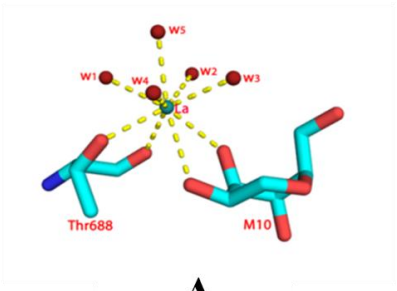
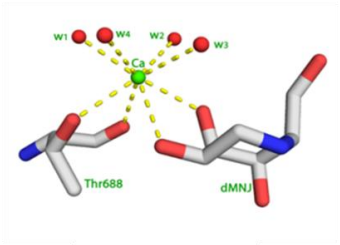


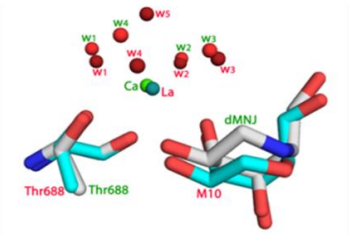
Figure 25. The coordination geometry of the metal ions in the active sites of the GH47 α1,2-mannosidases. The figure shows water molecules (sphere representation, labeled as "w"), Thr residues, dMNJ, and mannopyranose residues (stick representation) that coordinate with the metal ions. La^{3+} or Ca^{2+} (sphere representation), are rendered with Pymol software using the PDB files of the different α1,2-mannosidase structures. **Panel A**: human ER mannosidase I-$Man_9GlcNAc_2$-PA complex, La^{3+} is colored in teal, waters are colored in fire-brick red; **Panel B**: human ERManI in complex with dMNJ, Ca^{2+} is colored in green, waters in TV red; **Panel C**: alignments of the structures shown in **Panels A and B**; **Panel D**: Mouse GMIA-$Man_9GlcNAc_2$-PA complex, La^{3+} is colored in cyan, and waters in TV red; **Panel E**: mouse GMIA, Ca^{2+} is colored in green and waters in fire-brick red; **Panel F**: alignment of structures shown in **Panels D** and **E**; **Panel G**: alignment of structures shown in **Panels A and D**.


W5
W1
W4
W2
W3
La
Thr688
M10

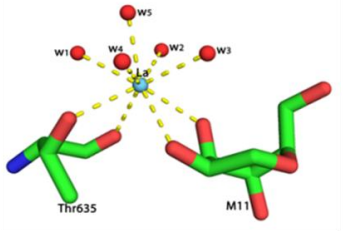
A


W1
W4
W2
W3
Ca
Thr688
dMNJ

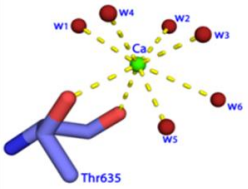
B


W5
W4
W1
W3
W2
Ca
La
dMNJ
Thr688
Thr688
M10

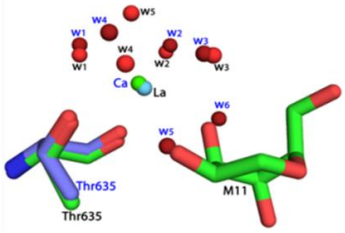
C


W5
W1
W4
W2
W3
La
Thr635
M11

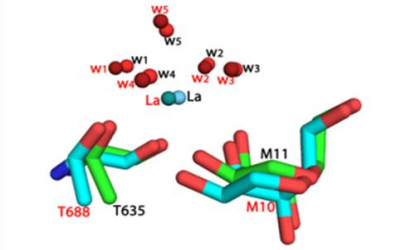
D


W1
W4
W2
W3
Ca
W6
W5
Thr635

E


W5
W4
W1
W2
W3
W1
W4
W2
W3
Ca
La
W6
W5
Thr635
Thr635
M11

F


W5
W5
W1
W2
W3
W1
W4
W2
W3
W4
La
La
M11
T688
T635
M10

G

Figure 26. The conformations of glycone residues (glycone mimics) in the -1 subsites of the GH47 α1,2-mannosidases. **Panel A**: dMNJ bound in the -1 subsite of human ERManI ($^1C_4$); **Panel B**: the mannopyranose structure of the M10 residue in the -1 subsite of human ERManI ($^1C_4$); **Panel C**: the mannopyranose structure of the M11 residue in the -1 subsite of mouse GMIA ($^1C_4$); **Panel D**: the glycone residue of the Man-S-Man thiodisaccharide in the -1 subsite of human ERManI ($^3S_1$); **Panel E**: alignment of the structures in **Panel B** (blue) and **Panel C** (green); **Panel F**: alignment of the structures in **Panel A** (gray), **Panel B** (blue) and **Panel D** (yellow).

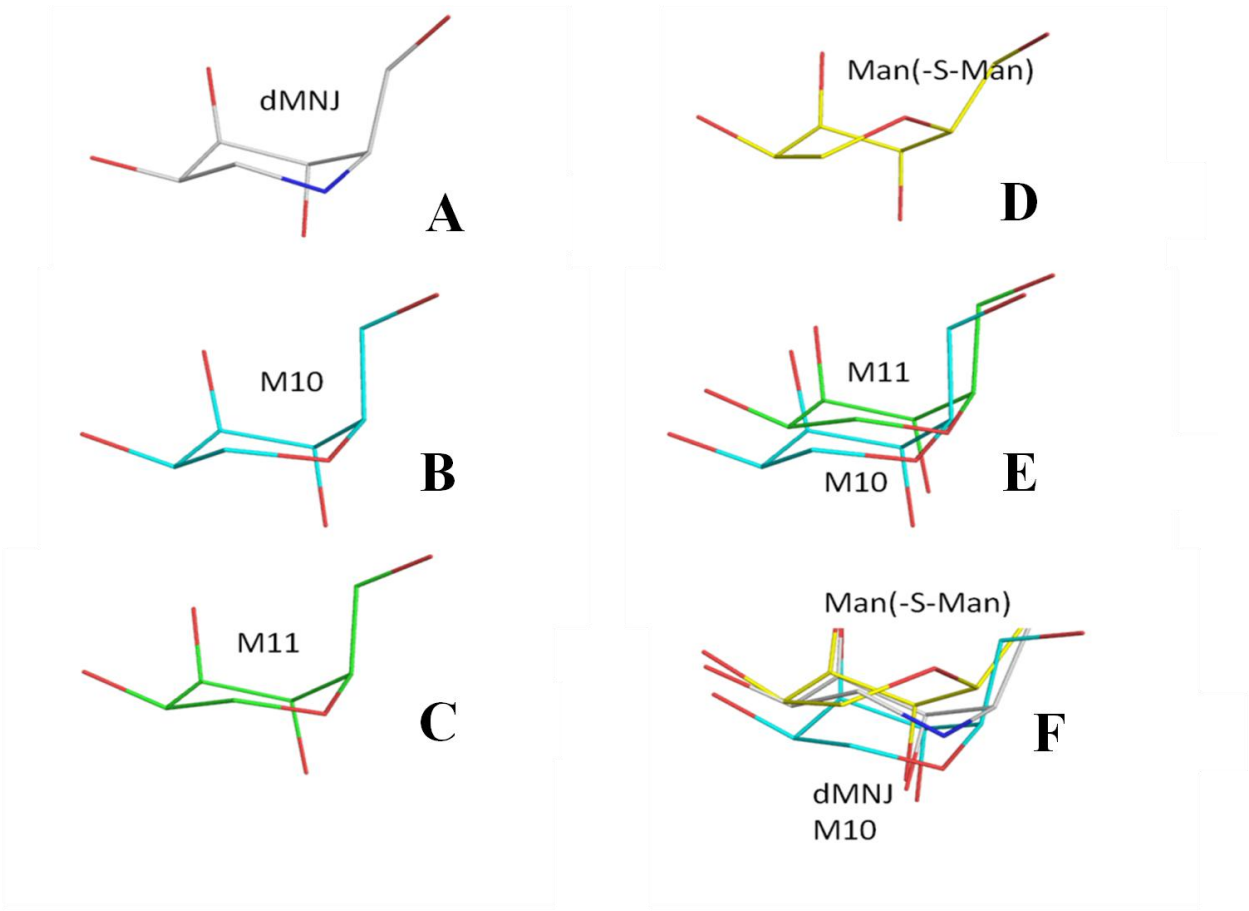

dMNJ
A
M10
B
M11
C
Man(-S-Man)
D
M11
M10
E
Man(-S-Man)
F
dMNJ
M10

Figure 27. The orientation of oligosaccharides in GH47 α1,2-mannodidase enzyme active sites. The metal ions are shown in sphere representations, the oligosaccharides are in stick representations. **Panel A**: $Man_5GlcNAc_2$-PA in the yeast ERManI active site with bound Ca^{2+}; **Panel B**: $Man_9GlcNAc_2$-PA in the human ERManI active site with bound La^{3+}; **Panel C**: alignment of structures shown in **Panel A** (yellow) and **Panel B** (cyan); **Panel D**: $Man_6GlcNAc_2$-PA in the mouse GMIA active site with bound Ca^{2+}; **Panel E**: $Man_9GlcNAc_2$-PA in the mouse GMIA active site with bound La^{3+}; **Panel F**: alignment of the structures shown in **Panel D** (blue) and **Panel E** (green).

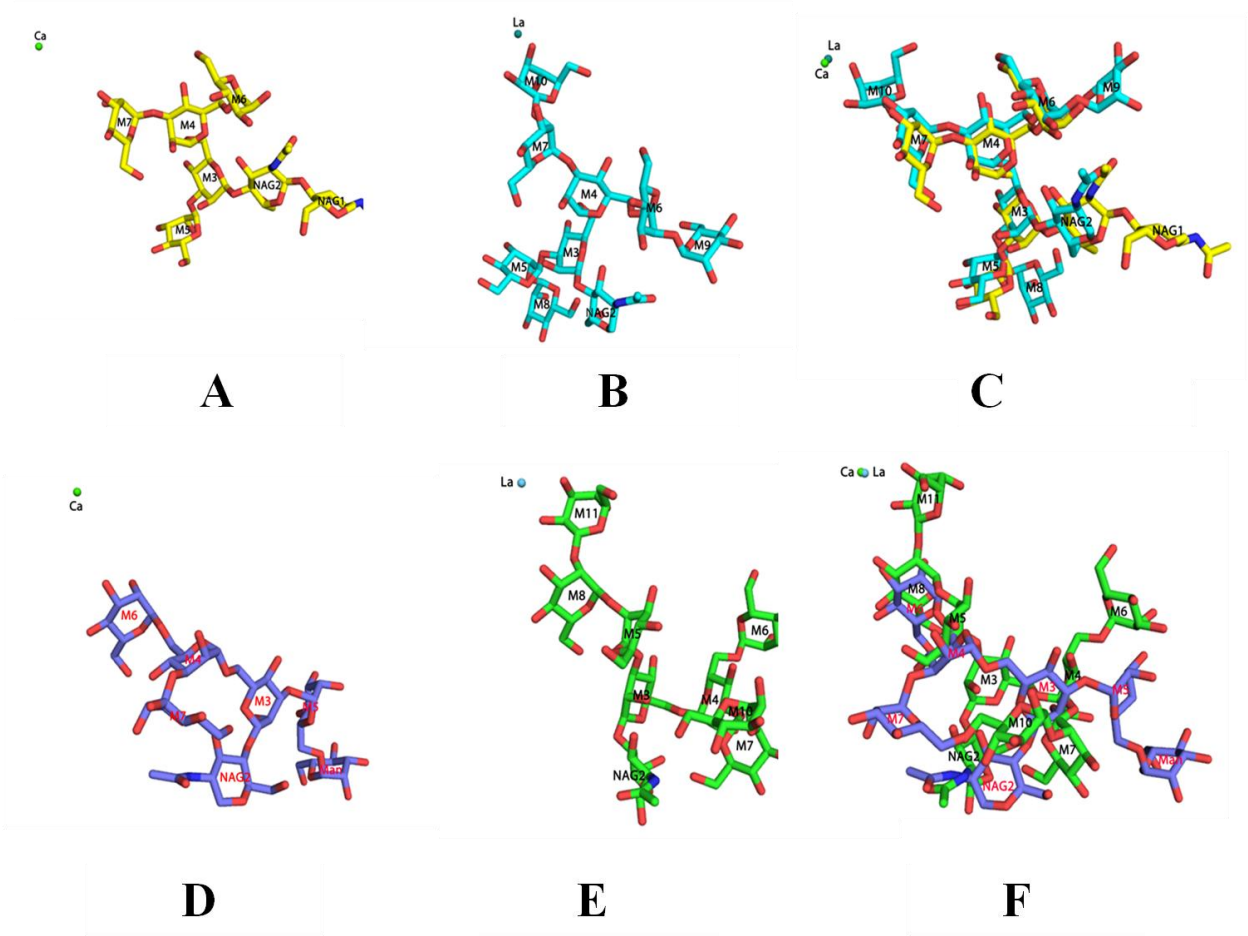
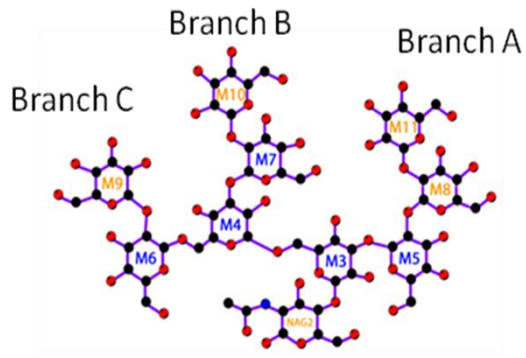
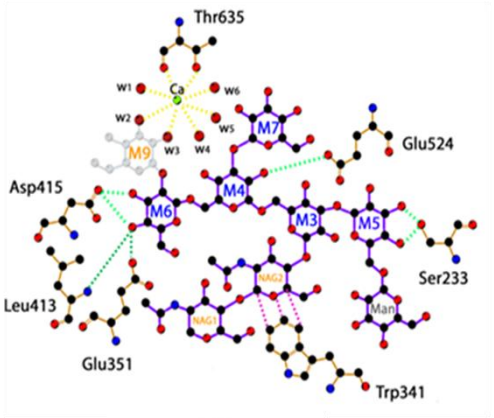

Ca
M6
M7
M4
M3
NAG2
NAG1
M9
A
La
M10
M7
M4
M6
M9
M3
M5
M8
NAG2
B
La
Ca
M10
M6
M9
M7
M4
M3
NAG2
NAG1
M5
M8
C
Ca
M6
M3
NAG2
D
La
M11
M8
M5
M6
M4
M10
M7
NAG2
E
Ca La
M11
M8
M5
M6
M4
M3
M2
M10
M7
NAG2
F

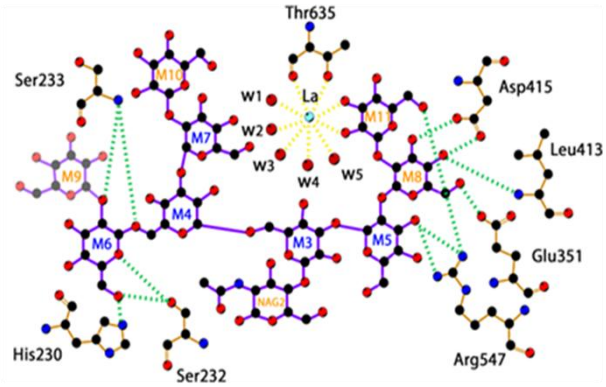
Figure 28. The interactions between oligosaccharides in the active sites of GH47 α1,2-mannosidases. The flattened oligosaccharide and amino acid residue structures are generated by the Ligplot^{+} program from the PDB files of the corresponding structures. The amino acid residues and oligosaccharides are in line representations, metal ions and water molecules are shown as sphere representations. The carbon atoms are in black, oxygen atoms in red, and nitrogen atoms in blue. The bonds in the amino acid residues are shown in orange and the bond in oligosaccharides are shown in purple. The coordination around the metal ions is represented as yellow dashed lines. The polar interactions between the respective amino acid residues and the oligosaccharides are represented as green dashed lines. The hydrophobic interactions are represented as pink dashed lines. The α1,2-mannose residues M8, M9, M10, and M11 and the NAG residues are labeled in orange letters. All other α1,3-, α1,6- and β1,6-linked mannose residues M3, M4, M5, M6 and M7 are labeled in blue letters. The mannose residues cleaved by the enzymes are shown in transparent gray, the mannose residues that are not visible in the crystal structure due to their high flexibility are shown in the same color pattern with the visible residues, but in transparent form. **Panel A**: legend for labeling of the $Man_9GlcNAc_2$ glycan; **Panel B**: $Man_6GlcNAc_2$ bound in mouse GMIA active site; **Panel C**: $Man_9GlcNAc_2$-PA bound in mouse GMIA active site; **Panel D**: $Man_9GlcNAc_2$-PA bound in human ERManI active site; **Panel E**: the interactions of the manno-pyranose residue in the -1 subsite of ERManI and GMIA (M11 for the Golgi enzyme and M10 for the ER enzyme) with the amino acid residues in the GMIA (labeled in black) or ERManI (labeled in blue).


Branch B
Branch A
Branch C
M10
M11
M7
M9
M8
M4
M6
M3
M5
NAG1

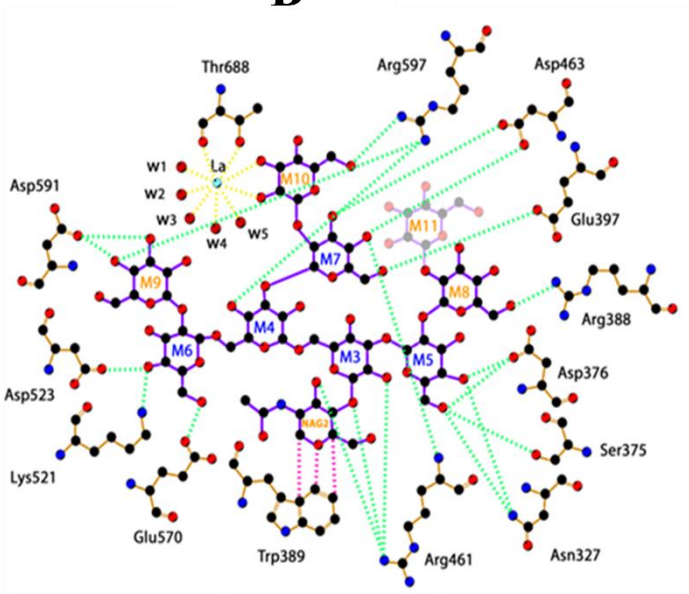
A


Thr635
W1
Ca
W6
W2
W5
M7
M9
W3
W4
Glu524
Asp415
M4
M6
M3
M5
Ser233
NAG2
Leu413
NAG1
Man
Glu351
Trp341

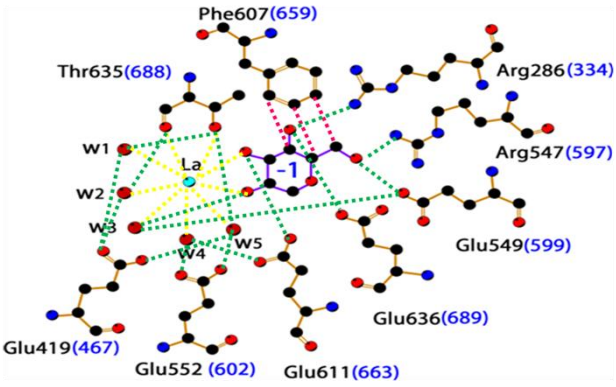
B


Thr635
Ser233
M10
W1
La
Asp415
M7
W2
M11
M9
W3
W4
W5
M8
Leu413
M4
M6
M3
M5
Glu351
NAG2
His230
Ser232
Arg547

C


Thr688
Arg597
Asp463
W1
La
Asp591
M10
W2
W3
W4
W5
Glu397
M11
M7
M9
M8
Arg388
M4
M6
M3
M5
Asp376
Asp523
NAG2
Ser375
Lys521
Glu570
Trp389
Arg461
Asn327

D


Phe607(659)
Thr635(688)
Arg286(334)
W1
La
-1
Arg547(597)
W2
W3
W4
W5
Glu549(599)
Glu636(689)
Glu419(467)
Glu552 (602)
Glu611(663)

E

References

1. Henrissat, B. and A. Bairoch, *Updating the sequence-based classification of glycosyl hydrolases.* Biochem J, 1996. **316 (Pt 2)**: p. 695-6.
2. Daniel, P.F., B. Winchester, and C.D. Warren, *Mammalian alpha-mannosidases--multiple forms but a common purpose?* Glycobiology, 1994. **4**(5): p. 551-66.
3. Lal, A., et al., *Substrate specificities of recombinant murine Golgi alpha1, 2-mannosidases IA and IB and comparison with endoplasmic reticulum and Golgi processing alpha1,2-mannosidases.* Glycobiology, 1998. **8**(10): p. 981-95.
4. Herscovics, A., et al., *Isolation of a mouse Golgi mannosidase cDNA, a member of a gene family conserved from yeast to mammals.* J Biol Chem, 1994. **269**(13): p. 9864-71.
5. Lal, A., et al., *Isolation and expression of murine and rabbit cDNAs encoding an alpha 1,2-mannosidase involved in the processing of asparagine-linked oligosaccharides.* J Biol Chem, 1994. **269**(13): p. 9872-81.
6. Igdoura, S.A., et al., *Alpha-mannosidases involved in N-glycan processing show cell specificity and distinct subcompartmentalization within the Golgi apparatus of cells in the testis and epididymis.* Eur J Cell Biol, 1999. **78**(7): p. 441-52.
7. Pan, S., X. Cheng, and R.N. Sifers, *Golgi-situated endoplasmic reticulum alpha-1, 2-mannosidase contributes to the retrieval of ERAD substrates through a direct interaction with gamma-COP.* Mol Biol Cell, 2013. **24**(8): p. 1111-21.
8. Pan, S., et al., *Golgi localization of ERManI defines spatial separation of the mammalian glycoprotein quality control system.* Mol Biol Cell, 2011. **22**(16): p. 2810-22.
9. Benyair, R., et al., *Mammalian ER mannosidase I resides in quality control vesicles, where it encounters its glycoprotein substrates.* Mol Biol Cell, 2015. **26**(2): p. 172-84.

10. Gallily, R., I. Stain, and S. Kornfeld, *Non-immune non-activated chicken macrophages destroy murine fibroblasts.* Dev Comp Immunol, 1985. **9**(4): p. 691-9.

11. Lipari, F. and A. Herscovics, *Production, purification and characterization of recombinant yeast processing alpha 1,2-mannosidase.* Glycobiology, 1994. **4**(5): p. 697-702.

12. Herscovics, A., *Processing glycosidases of Saccharomyces cerevisiae.* Biochim Biophys Acta, 1999. **1426**(2): p. 275-85.

13. Tremblay, L.O., N. Campbell Dyke, and A. Herscovics, *Molecular cloning, chromosomal mapping and tissue-specific expression of a novel human alpha1,2-mannosidase gene involved in N-glycan maturation.* Glycobiology, 1998. **8**(6): p. 585-95.

14. Karaveg, K. and K.W. Moremen, *Energetics of substrate binding and catalysis by class 1 (glycosylhydrolase family 47) alpha-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(33): p. 29837-48.

15. Lobsanov, Y.D., et al., *Structure of Penicillium citrinum alpha 1,2-mannosidase reveals the basis for differences in specificity of the endoplasmic reticulum and Golgi class I enzymes.* J Biol Chem, 2002. **277**(7): p. 5620-30.

16. Thompson, A.J., et al., *The reaction coordinate of a bacterial GH47 alpha-mannosidase: a combined quantum mechanical and structural approach.* Angew Chem Int Ed Engl, 2012. **51**(44): p. 10997-1001.

17. Vallee, F., et al., *Structural basis for catalysis and inhibition of N-glycan processing class I alpha 1,2-mannosidases.* J Biol Chem, 2000. **275**(52): p. 41287-98.

18. Van Petegem, F., et al., *Trichoderma reesei alpha-1,2-mannosidase: structural basis for the cleavage of four consecutive mannose residues.* J Mol Biol, 2001. **312**(1): p. 157-65.

19. Karaveg, K., et al., *Mechanism of class 1 (glycosylhydrolase family 47) {alpha}-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(16): p. 16197-207.

20. Tempel, W., et al., *Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases.* J Biol Chem, 2004. **279**(28): p. 29774-86.

21. Vallee, F., et al., *Crystal structure of a class I alpha1,2-mannosidase involved in N-glycan processing and endoplasmic reticulum quality control.* EMBO J, 2000. **19**(4): p. 581-8.

22. Karaveg, K., et al., *Crystallization and preliminary X-ray diffraction analysis of lectin-1 from Pseudomonas aeruginosa.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 7): p. 1241-2.

23. Mast, S.W., et al., *Human EDEM2, a novel homolog of family 47 glycosidases, is involved in ER-associated degradation of glycoproteins.* Glycobiology, 2005. **15**(4): p. 421-36.

24. Romero, P.A., et al., *Mutation of Arg(273) to Leu alters the specificity of the yeast N-glycan processing class I alpha1,2-mannosidase.* J Biol Chem, 2000. **275**(15): p. 11071-4.

25. Franco-Fraguas, L., et al., *Preparative purification of soybean agglutinin by affinity chromatography and its immobilization for polysaccharide isolation.* J Chromatogr B Analyt Technol Biomed Life Sci, 2003. **790**(1-2): p. 365-72.

26. Kwan, E.M., et al., *N-Glycosidase-carbohydrate-binding module fusion proteins as immobilized enzymes for protein deglycosylation.* Protein Eng Des Sel, 2005. **18**(10): p. 497-501.

27. Mulakala, C., W. Nerinckx, and P.J. Reilly, *Docking studies on glycoside hydrolase Family 47 endoplasmic reticulum alpha-(1-->2)-mannosidase I to elucidate the pathway to the substrate transition state.* Carbohydr Res, 2006. **341**(13): p. 2233-45.

28. Mulakala, C., W. Nerinckx, and P.J. Reilly, *The fate of beta-D-mannopyranose after its formation by endoplasmic reticulum alpha-(1-->2)-mannosidase I catalysis.* Carbohydr Res, 2007. **342**(2): p. 163-9.

29. Mulakala, C. and P.J. Reilly, *Understanding protein structure-function relationships in Family 47 alpha-1,2-mannosidases through computational docking of ligands.* Proteins, 2002. **49**(1): p. 125-34.

30. Herscovics, A., P.A. Romero, and L.O. Tremblay, *The specificity of the yeast and human class I ER alpha 1,2-mannosidases involved in ER quality control is not as strict previously reported.* Glycobiology, 2002. **12**(4): p. 14G-15G.

31. Dimicoli, J.L. and J. Bieth, *Location of the calcium ion binding site in porcine pancreatic elastase using a lanthanide ion probe.* Biochemistry, 1977. **16**(25): p. 5532-7.

32. Dux, L., et al., *Crystallization of the Ca2+-ATPase of sarcoplasmic reticulum by calcium and lanthanide ions.* J Biol Chem, 1985. **260**(21): p. 11730-43.

33. Fernandez-Belda, F., *Lanthanum as a calcium-substituting ion for binding to sarcoplasmic reticulum ATPase.* Arch Biochem Biophys, 1988. **267**(2): p. 770-5.

34. Fitzpatrick, L.A., *Differences in the actions of calcium versus lanthanum to influence parathyroid hormone release.* Endocrinology, 1990. **127**(2): p. 711-5.

35. Fujimori, T. and W.P. Jencks, *Lanthanum inhibits steady-state turnover of the sarcoplasmic reticulum calcium ATPase by replacing magnesium as the catalytic ion.* J Biol Chem, 1990. **265**(27): p. 16262-70.

36. Gomez, J.E., E.R. Birnbaum, and D.W. Darnall, *The metal ion acceleration of the conversion of trypsinogen to trypsin. Lanthanide ions as calcium ion substitutes.* Biochemistry, 1974. **13**(18): p. 3745-50.

37. Guo, S., et al., *Inhibition mechanism of lanthanum ion on the activity of horseradish peroxidase in vitro.* Spectrochim Acta A Mol Biomol Spectrosc, 2010. **75**(2): p. 936-40.

38. Switzer, M.E., *The lanthanide ions as probes of calcium ion binding sites in biological systems.* Sci Prog, 1978. **65**(257): p. 19-30.

39. Wilkins, A.L., et al., *Metal-binding studies for a de novo designed calcium-binding protein.* Protein Eng, 2002. **15**(7): p. 571-4.

40. Xue, J., et al., *Sugar-metal ion interactions: the coordination behaviors of lanthanum with erythritol.* Carbohydr Res, 2012. **361**: p. 12-8.

41. Shannon, R.D., *Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides.* Acta Cryst., 1976. **A32**: p. 751-767.

42. Horn, J.D.V., *Electronic Table of Shannon Ionic Radii.* 2001.

43. March, J., Advanced Organic Chemistry 1992(4th Ed).

44. Terrier, C., et al., *Density functional theory based molecular dynamics study of hydration and electronic properties of aqueous La(3+).* J Chem Phys, 2010. **133**(4): p. 044509.

45. Wang, L., et al., *Interaction between lanthanum ion and horseradish peroxidase in vitro.* Biochimie, 2010. **92**(1): p. 41-50.

46. Stenger J., C.M., Eggers F., Eyring E.M., Kaatze U., Petrucci S., *Molecular dynamics and kinetics of monosaccharides in solution.A broadband ultrasonic relaxation study.* J Phys Chem B, 2000. **104**: p. 4782–4790.

47. Helenius, A., *How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum.* Mol Biol Cell 1994(5): p. 253–265.

48. Liu Y, C.P., Cabral CM, Sifers RN., *Oligosaccharide modification in the early secretory pathway directs the selection ofa misfolded glycoprotein for degradation by the proteasome.* J Biol Chem 1999(274): p. 5861–5867.

49. Byrd, J.C., et al., *Glycoprotein synthesis in yeast. Identification of Man8GlcNAc2 as an essential intermediate in oligosaccharide processing.* J Biol Chem, 1982. **257**(24): p. 14657-66.

50. Gonzalez, D.S., et al., *Identification, expression, and characterization of a cDNA encoding human endoplasmic reticulum mannosidase I, the enzyme that catalyzes the first mannose trimming step in mammalian Asn-linked oligosaccharide biosynthesis.* J Biol Chem, 1999. **274**(30): p. 21375-86.

51. Jelinek-Kelly, S., et al., *Characterization of a specific alpha-mannosidase involved in oligosaccharide processing in Saccharomyces cerevisiae.* J Biol Chem, 1985. **260**(4): p. 2253-7.

CHAPTER 3

Mutational study of the non-conserved residues involved in the enzyme-substrate interactions in human ERManI and mouse Golgi ManIA

3.1 Abstract

A structural comparison of the complexes between human ERManI and mouse Golgi ManIA (GMIA) with the $Man_9GlcNAc_2$-PA substrate analog revealed several non-conserved residues in topologically equivalent locations have unique interactions with the substrate. The effects of mutations in the non-conserved residues on enzyme substrate specificity were explored. Results from substrate digestion time course studies, kinetic assays, and binding analyses of the mutants showed that deletion of specific enzyme-substrate interactions by mutation altered the enzyme substrate specificity as well as the enzyme catalytic activity and binding affinity.

3.2. Introduction

CAZy Glycoside Hydrolase Family 47 (GH47) α-mannosidases play important roles in the *N*-glycan maturation and ER-associated degradation (ERAD) of *N*-glycoproteins in mammalian cells [1]. Among the seven members of this enzyme family in mammals, two enzyme forms, ER α-mannosidase I (ERManI) and Golgi α-mannosidase IA (GMIA), have been extensively studied in regard to their substrate specificity, catalytic mechanisms and structural features [2-8]. ERManI and GMIA share significant similarity in their amino acid sequences, general protein structures, and cleavage activity towards α1,2-mannosyl linkages. However, ERManI and GMIA also exhibit exclusive and complementary substrate specificities for cleavage of terminal α1,2-mannosyl residues on $Man_9GlcNAc_2$ substrate glycans. *In vitro*, ERManI prefers to cleave the M10 residue from the middle branch (branch B) of the glycan

structure to produce $Man_8GlcNAc_2$ isomer B, and exhibits low cleavage efficiency towards the other α1,2-mannose residues on branches A and C (Figure 23). In contrast, GMIA first cleaves the terminal α1,2-mannose residue on branch A to produce the $Man_8GlcNAc_2$ isomer A and then continues to cleave the terminal α1,2-mannose residues on branch C followed by subsequent cleavage of the underlying residue on branch A to produce the $Man_6GlcNAc_2$ isomer retaining a single α1,2-mannose residues on branch B. However, cleavage of the terminal residue on branch B, the primary linkage cleaved by ERManI, occurs slowly for GMIA indicating that the ER and Golgi enzymes have complementary but non-overlapping specificities for cleavage of all terminal α1,2-mannose residues in the $Man_9GlcNAc_2$ substrate. *In vivo*, the substrate specificity differences between ERManI and GMIA are demonstrated by the initial cleavage of the $Man_9GlcNAc_2$ *N*-glycan precursor first by ERManI to generate $Man_8GlcNAc_2$ isomer B in the ER lumen and then further cleaved in the Golgi apparatus to $Man_5GlcNAc_2$ by Golgi α-mannosidases IA/IB/IC.

Previous structural studies of ERManI and GMIA have shed some light on the structural basis for their similarities in catalytic activity and differences in substrate specificities. The conserved fold of the GH47 α-mannosidases, including ERManI and GMIA, features an $(\alpha\alpha)_7$ anti-parallel barrel structure with one end of the barrel plugged with a β-hairpin and the other end comprised of a broad funnel-shaped cleft extending into a deep cavity containing the active site. This active site cavity is employed for an exo-α-mannosidic cleavage of non-reducing terminal α1,2-mannose linkages by an inverting catalytic mechanism. In this mechanism the glycone and the adjoining residue in the terminal Man-α1,2-Man linkage enter highly conserved -1 and +1 glycan binding subsites to initiate bond cleavage. Interactions involve a combination of hydrophobic interactions, Ca^{2+} ion coordination to two glycone hydroxyl residues, and an

extensive matrix of hydrogen bonding interactions with the -1 and +1 residues. These interactions result in the distortion of the glycone residue into a high free energy $^{1}C_{4}$ sugar conformation that approaches the ring-flattened $^{3}H_{4}$ transition state structure required for bond hydrolysis. The -1 and +1 subsites also contain catalytic residues that allow the nucleophilic attack of the glycosidic carbon atom by an activated water molecule and protonation of the glycosidic oxygen of the mannose leaving group.

While the residues, geometries, and reactivities within the -1 and +1 subsites are identical between ERManI and GMIA, the residues in the less conserved >+1 subsites that extend further out in the glycan binding cleft are quite distinct. These differences in binding cleft residues have been proposed to contribute to the substrate specificity differences between the enzymes. Therefore, a detailed mapping and characterization of the non-conserved residues interacting with the oligosaccharide substrate will be critical for revealing the structural determinants of substrate specificity for each enzyme. Structural comparison of the enzyme-substrate and enzyme product complexes for ERManI and GMIA indicate that there are numerous differences in interacting residues. For example, yeast ERManI contains an Arg residue (Arg273) within the active site cleft that interacts with three glycan substrate residues, including the β1,4-linked mannose and β1,4-linked GlcNAc in the glycan core and an outer α-mannose residue in branch B. In contrast, the equivalent residue in GMIA (Leu413) only interacts with an outer α-mannose residue on branch A. The single R273L mutation in yeast ERManI results in a poor efficiency enzyme that can cleave $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$ *in vitro* [9], suggesting that Arg273 is at least partly responsible for the restricted substrate specificity of ERManI.

The structures of the enzyme-$Man_9GlcNAc_2$-PA complexes for both human ERManI and mouse GMIA were determined by X-ray crystallography (see Chapter 2). These two structures

serve as enzyme-substrate complex models and allowed identification of the enzyme and oligosaccharide residues involved in enzyme-substrate interactions and the nature of those interactions. Comparison of the enzyme-substrate interactions of human ERManI and mouse GMIA revealed that the orientation and arrangements of the glycans in the two enzyme-substrate complexes are different (Figure 29). Besides using the conserved residues in the equivalent position of -1 and +1 subsites that interact with the substrate, the two enzymes use non-conserved amino acid residues to interact with the $Man_9GlcNAc_2$-PA, and the latter is proposed to confer the different substrate specificities for each respective enzyme.

In this study we test the hypothesis that the cleft residues of ERManI and GMIA determine the specificity of glycan trimming by swapping the >+1 subsites residues of human ERManI and mouse GMIA in an attempt to observe their influence on substrate specificity of the respective enzymes. Enzyme kinetic analysis and time-course profiling of the catalytic products were done on the mutant and wild type enzymes. The results showed that mutations of some of the non-conserved and non-catalytic amino acid residues in the substrate binding cleft of ERManI influenced both catalytic activity and branch specificity for glycan cleavage and led to the hybrid substrate specificity. The measurement of substrate binding affinity by SPR on a subset of the mutants indicated non-conserved residues play important roles in the binding of the $Man_9GlcNAc_2$ substrate. This study suggested that the coordinated interactions of the residues in the enzyme in the >+1 subsites act coordinately to provide discrimination in branch specificity for glycan substrate interactions and catalytic efficiency

3.3. Experimental Procedures

3.3.1 Oligosaccharide isolation and tagging.

As reported previously [2, 10], $Man_9GlcNAc_2$ was isolated from crude soybean agglutinin that was purified from soybean powder following acid extraction and ammonium sulfate precipitation. The lectin was digested using a combination of trypsin and elastase, and the resulting glycopeptides were incubated with PNGase to liberate $Man_9GlcNAc_2$ [2]. The reducing end of the released glycan was tagged with 2-amino-pyridine by reductive amination to produce the $Man_9GlcNAc_2$-pyridylamine (PA) substrate for kinetic and binding studies [11-13]. Further purification by HPLC to separate $Man_9GlcNAc_2$-PA from smaller degraded oligosaccharides and free PA was carried out as described in Chapter 2.

3.3.2. Site-directed mutagenesis and protein expression.

The single site-directed mutants were created using expression constructs encoding the truncated catalytic domains of human ERManI or GMIA in the plasmid vector pPICZα as a template [4]. Mutagenesis primers were synthesized by MWG Biotech (Huntsville, AL). The QuikChange® site-directed mutagenesis kit from Stratagene (La Jolla, CA) and the protocol provided by the manufacturer were employed for the mutagenesis. The products of mutagenesis were transformed into Top10 *E. coli* cells after digestion with the restriction enzyme DpnI. The transformants were grown at 37℃ overnight in low salt LB medium with Zeocin as selection antibiotic. Plasmid DNA preparations were carried by using QIAprep spin miniprep kit from Qiagen (Valencia, CA). The multi-site mutants were made also using the same QuikChange® mutagenesis procedure starting with wild type plasmid as template and by generating a sequential set of single mutants. The coding regions from the resulting mutant plasmids were completely sequenced by MWG Biotech to confirm the absence of other undesired mutations before being transformed with the *P. pastoris* strain GS115 for expression.

3.3.3. Protein expression and purification.

All the mutant proteins were expressed in *P. pastoris* in shake flask cultures as reported previously for the wild type enzyme [5]. The yeast transformants were first grown for 48 hours in 1L BMGY media at 30 ℃ with strong agitation to generate cell mass. The recombinant protein production was initiated by addition of methanol at 12-hr intervals to the culture to a final concentration of 0.5% (v/v). The duration of induction varied from 5 to 7 days depending on the expression level of the mutants. Purification of the mutant proteins employed procedures identical to the wild type enzyme. Sequential purification over SP-Sepharose, Phenyl-Sepharose, and Superdex-75 columns (Amersham-Pharmacia Biotech) were employed to generate the final purified proteins as described previously [5]. The final purified enzymes were concentrated to 10 mg/ml in 20 mM MES, 150 mM NaCl, 5 mM $CaCl_2$ and 200 mM NDSB-201 (pH 6.5 for GMIA mutants; pH 7.0 for ERManI mutants) using an Amicon pressure concentrator with a YM-10 membrane and Amicon centrifugal filter devices (30 kDa molecular weight cutoff). The concentrated protein was flash frozen with liquid nitrogen and stored at -80 ℃.

3.3.4. α-Mannosidase kinetic analysis.

Assays were performed using $Man_9GlcNAc_2$-PA as the substrate and purified enzyme preparations in the microfuge tubes with 10 µl of substrate at various concentration in reaction buffer (20 mM MES, 150 mM NaCl, 5 mM $CaCl_2$), 10 µl of enzyme (concentration depending on the activity of respective enzyme preparation) in reaction buffer in a total volume of 50 µl at 37 ℃ for varied periods of time depending on the concentration and activity of the enzyme mutants. The reactions were stopped by addition of 50 µl acetonitrile followed by thorough mixing. Separations of the digested products were performed on a Hypersil APS-NH_2 column (4.6 × 250 mm, Alltech) using isocratic mobile phase elution containing 40% acetonitrile in 50 mM phosphate buffer (pH 4.0) at flow rate 1.0 ml/min. $Man_{9-5}GlcNAc_2$-PA were separated with

an elution time of less than 20 min. Variations in the flow rate (0.8-1.5 ml/min) or the percentage of acetonitrile in the mobile phase (40-55%) were used to shift the elution position of $Man_{9-5}GlcNAc_2$-PA in order to optimize the resolution. Elution of the $Man_{9-5}GlcNAc_2$-PA was monitored by in-line fluorescence detection (excitation wave length, 320 nm; emission wave length, 400 nm). Pure $Man_{9-5}GlcNAc_2$-PA standards were employed under identical separation conditions to identify the oligosaccharides from the enzymatic cleavage reactions. The areas under the resolved peaks were integrated to determine the quantities of the corresponding oligosaccharides in the reaction product mixtures and to determine the decreasing abundance of the $Man_9GlcNAc_2$-PA substrate and appearance of $Man_{8-5}GlcNAc_2$-PA products. The decreasing abundance of $Man_9GlcNAc_2$-PA was used in a time point prior to the appearance of $Man_{7-5}GlcNAc_2$-PA to determine the initial reaction rates for both ERManI and GMIA. Initial rates (v) for the enzyme reactions were determined at various substrate concentrations raging from 10 μM to 1600 μM. Catalytic coefficient (k_{cat}) and Michaelis constant (K_m) values were determined by fitting initial rates to the Michaelis-Menten function by nonlinear regression analysis using SigmaPlot (Jandel Scientific, San Rafael, CA).

3.3.5. Time-course profiling of the catalytic products.

To determine the rates of conversion among trimmed glycan products, time-course assay were performed by mixing 1 nmol $Man_9GlcNAc_2$-PA in 50 μl of 20 mM MES, 150 mM NaCl, 5 mM $CaCl_2$, 2-20 μl of enzyme stock solution (10 μg/ml in the same buffer) depending on the activity of the enzymes in a total volume of 100 μl of reaction buffer. The digestion was incubated at 37 ℃ and 5 μl aliquots of the reaction mix were removed at 5 min intervals and mixed immediately with 5 ml acetonitrile to stop the reaction. The 10 μl samples were separated by HPLC using the same methods as in the kinetic analysis. The peak areas of the Man_{9-}

$_5$GlcNAc$_2$-PA structures were integrated and the percentage ratio of each oligosaccharide in the reaction mix were calculated. Curves of isomer percentage *versus* time for all the oligosaccharides in the reaction mixture were generated in the same plot for a single enzyme reaction.

3.3.6. Glycan binding studies by surface plasmon resonance.

Surface plasmon resonance (SPR) analysis was performed using a Biacore T100 instrument (Biacore AB), with the recombinant enzymes immobilized on the CM5 analytical chip surface (Biacore AB) at 25 ℃ using amine coupling and the Man$_9$GlcNAc$_2$-PA ligand flowing over the chip surface as the analyte. To activate the chip surface, 50 mM N-hydroxysuccinimide and 200 mM 1-ethyl-3-(dimethylaminopropyl) carbodimide were mixed and injected over the chip surface for 7 min at 5 µl/min. The enzyme to be immobilized was diluted to 10 µg/ml in 10 mM sodium succinate (pH 6.0) and passed through a 0.2 µm membrane filter (Millipore Corp.) The immobilization of enzyme was conducted by injecting the enzyme across the cell surface at the flow rate of 5 µl/min with the running HBS-P buffer (10 mM HEPES; 150 mM NaCl; 0.005% v/v P20 surfactant; pH 7.4) at 25 ℃. The remaining active groups on the chip surface were blocked by injection of 1 M ethanolamine-HCl (pH 8.5) at 5 µl/min for 7 min. To strip bound calcium from the immobilized enzyme on the chip surface, HPB-P buffer supplemented with 10 mM EDTA was injected through the flow cell at 5 µl/min for 1 hr, followed by injection of buffer containing 10 mM MES, 150 mM NaCl (pH 6.5) at 5 µl/min for 10 min to remove residual EDTA. To bind La^{3+} to the immobilized enzyme, the running buffer (10 mM MES, 150 mM NaCl, 5 mM $LaCl_3$, pH 6.5) was injected into the flow cell at 5 µl/min for 1 hr and the same running buffer was used for the SPR analysis. For each analysis, a second flow cell was used as reference and treated identically to the enzyme-immobilized cell, except

that the enzyme solution was replaced by the blank HBS-P buffer during the immobilization step. The analyte $Man_9GlcNAc_2$-PA solutions (0.5 to 200 μM) were prepared in the running buffer and injected across both the immobilized and reference flow cells for 1 min, followed by 3-5 min of dissociation phase. Regeneration and equilibrium of the flow cells were conducted after each run with the same running buffer. SPR experiments for each concentration of analyte were conducted twice, the data for all concentrations were collected and globally fit to two-state curve to calculate values the equilibrium dissociation constant (K_D) using BiaEvaluation version 3.2 software. Alternatively, the maximal equilibrium sensorgram values were used to plot a saturation binding curve and calculate the values of K_D directly.

3.3.6. Separation of $Man_8GlcNAc_2$-PA isomers by HPLC.

$Man_8GlcNAc_2$-PA isomer mixtures generated by wild type and mutant enzymes were separated and collected at the certain time points of cleavage for profiling of isomer products using a Hypersil® APS-2 amine column as described in Chapter 2. The collected $Man_8GlcNAc_2$-PA isomer mixtures were dried by using speed-vac and dissolved back into distilled water for further separation. A Cosmosil® C-18 column (6 x 0.5 cm) was used to separate $Man_8GlcNAc_2$-PA isomers on a Shimatzu® HPLC system. The mobile phase contained 20 μm NH_4 acetate and 1% butanol in water (pH 4.0). The flow rate was 0.8 ml/min.

3.4. Results

Comparison of the ERManI and GMIA enzyme-substrate glycan complexes revealed conserved glycan interactions within the enzyme -1 and +1 subsites as well as non-conserved interactions within the >+1 subsites of the glycan binding clefts. The equivalent pairs of the non-conserved amino acid residues in the two enzymes were found to either interact with different residues of the $Man_9GlcNAc_2$-PA substrate or to participate in interactions where the

corresponding residues in the other enzyme do not participate. Mutations were performed to swap the non-conserved residues in the >+1 subsite clefts between the two enzymes to examine their impacts on enzyme substrate specificity.

3.4.1. $Man_9GlcNAc_2$-PA digestion profiles of wild type ERManI and GMIA.

Wild type human ERManI can efficiently cleave the M10 mannose residue from the $Man_9GlcNAc_2$-PA substrate (terminal linkage on branch B) to produce the $Man_8GlcNAc_2$-PA isomer B. The time-course digestion profile of this enzyme reflected this narrow specificity by showing a continuous decrease in $Man_9GlcNAc_2$-PA and a corresponding increase in $Man_8GlcNAc_2$-PA to 100% without further cleavage to smaller structures over the time course (Figure 30).

In contrast to the simple cleavage profile of ERManI, wild type mouse GMIA has a more complicated time-course of glycan cleavage. The time course for mouse GMIA indicates a rapid and sequential cleavage of $Man_9GlcNAc_2$-PA to $Man_8GlcNAc_2$-PA, $Man_7GlcNAc_2$-PA, and progressive accumulation of $Man_6GlcNAc_2$-PA structures. Further conversion to $Man_5GlcNAc_2$-PA was at a much slower rate where ~15% of the products were represented by $Man_5GlcNAc_2$-PA structures when ~80% were comprised of $Man_6GlcNAc_2$-PA at the end of the time course. Similar to the ERManI digestion profile, $Man_8GlcNAc_2$-PA structures initially accumulated as $Man_9GlcNAc_2$-PA decreased in abundance. However, in contrast ERManI, $Man_8GlcNAc_2$-PA accumulation only reached a maximum of ~45% of total glycan structures before decreasing during the conversion to smaller structures. Similarly, $Man_7GlcNAc_2$-PA structures reached a transient maximal abundance of 20% of the total glycan structures. Eventually $Man_6GlcNAc_2$-PA structures accumulated as the dominant glycan structure with a slow and progressive conversion to $Man_5GlcNAc_2$-PA structures (Figure 30).

The cleavage profiles of wild type ERManI and GMIA were consistent with the previously reported substrate specificities for these two enzymes towards $Man_9GlcNAc_2$ [4, 8]. ERManI was shown to cleave the M10 residue on branch B of the tri-antennary substrate structure, but was resistant to cleavage of the three terminal α1,2-mannose residues on the other two antennae (branches A and C). In contrast, GMIA efficiency cleaves the α1,2-mannose residues on branches A and C, but cleaves the M10 residue on branch B with low efficiency.

3.4.2. The influence of non-conserved substrate binding residues on the substrate specificity

3.4.2.1. The influence of R461 and W389 on the substrate specificity of ERManI.

In the structure of the ERManI-$Man_9GlcNAc_2$-PA complex, an Arg residue (R461) interacts with residues M3, M7 and NAG2 of $Man_9GlcNAc_2$-PA (Figure 31), whereas the topologically equivalent residue in GMIA is a Leu residue (L413) that has no polar contacts with the substrate glycan (Figure 31). In the ERManI enzyme-substrate complex a Trp residue (W389) was also found to share hydrophobic stacking interactions with the NAG2 residue of $Man_9GlcNAc_2$-PA, while the equivalent Trp residue (W341) in GMIA was flipped away and did not interact with the glycan substrate (Figure 32). To test the roles of the respective residues in substrate interactions each residue was mutated and a substrate cleavage time course was performed. A R461L mutation was generated to swap the ERManI residue for the equivalent GMIA residue. Digestion of the $Man_9GlcNAc_2$-PA substrate with the ERManI R461L mutant (Figure 31) resulted in a dramatic change in the substrate specificity compared to the wild type enzyme. The enzyme cleaved $Man_9GlcNAc_2$-PA to $Man_8GlcNAc_2$-PA and continued to further digest the substrate to smaller oligosaccharide structures. In comparison to wild type ERManI, the digestion of $Man_9GlcNAc_2$-PA by the R461L mutant to $Man_8GlcNAc_2$-PA proceeded much more slowly, requiring 8 times more enzyme to digest the same amount of $Man_9GlcNAc_2$-PA to

$Man_8GlcNAc_2$-PA within the equivalent time period. The time course indicated that the $Man_8GlcNAc_2$-PA intermediate transiently reached 100% of the total glycan structures before being converted to $Man_7GlcNAc_2$-PA (maximal relative abundance of 20%) prior to conversion to $Man_6GlcNAc_2$-PA and $Man_5GlcNAc_2$-PA with similar abundances in later time points of the time course. These data indicate that the R461L mutant is a poor efficiency α1,2-mannosidase with a hybrid activity that initially cleaves $Man_9GlcNAc_2$-PA to $Man_8GlcNAc_2$-PA and further promiscuously cleaves to $Man_{6-5}GlcNAc_2$-PA structures with even lower efficiency.

To test the contribution of the W389 on the substrate specificity of ERManI, this residue was mutated to Ala (W389A), since both ERManI and GMIA contain a Trp residue at this equivalent position. Similar to the R461L mutation, the W389A mutation resulted in an enzyme with reduced catalytic efficiency (~14-fold more enzyme required to cleave equivalent $Man_9GlcNAc_2$-PA to $Man_8GlcNAc_2$-PA) that was able to further trim the glycan to structures smaller than $Man_8GlcNAc_2$-PA (Figure 32). However, in contrast to the R461L mutant, the W389A mutant did not accumulate $Man_8GlcNAc_2$-PA or $Man_7GlcNAc_2$-PA intermediates to an appreciable degree during the time course (maximal accumulation of 25% and 15%, respectively). Instead, the digestion progressed rapidly to $Man_6GlcNAc_2$-PA and subsequently to $Man_5GlcNAc_2$-PA. By comparison to the wild type GMIA time course, the ERManI W389A mutant accumulated even less $Man_8GlcNAc_2$-PA intermediates (maximal accumulation of 25% versus 45% for wild type GMIA) during the time course. Accumulation of the $Man_7GlcNAc_2$-PA intermediate for the two enzyme forms was comparable (maximal accumulation of 15% for both ERManI W389A and wild type GMIA). Thus, the W389A mutation appears to result in an enzyme that cleaves $Man_9GlcNAc_2$-PA substrates to smaller structures in a time course profile similar to GMIA, albeit with a reduced efficiency compared with either wild type enzyme.

3.4.2.2. The influence of N327, S375 and D376 on the substrate specificity of ERManI

Comparison of the ERManI-$Man_9GlcNAc_2$-PA complex with the equivalent structure of GMIA also revealed that three residues in ERManI, N327, S375 and D376, interact in concert with the M5 residue of $Man_7GlcNAc_2$-PA, while the equivalent residues in GMIA (S279, A327 and L328) do not interact with any residues of the $Man_9GlcNAc_2$-PA substrate (Figure 33A). We tested the influence of these amino residues on glycan cleavage specificity by swapping individual amino acid residues in ERManI to their equivalent GMIA residues (N327S, S375A and D376L, respectively) and performing time-course cleavage studies. The resulting time-course profiles indicated that the mutants changed the glycan cleavage pattern to differing degrees (Figure 33B, 33C and 33D). The digestion profile of mutants S375A and D376L indicated that both enzymes continue to digest the substrate beyond $Man_8GlcNAc_2$-PA to produce smaller oligosaccharides with low efficiency. Cleavage to $Man_8GlcNAc_2$-PA occurred rapidly and transiently accumulated to 100% of the total glycan population similar to wild type ERManI, but was slowly converted to $Man_{7\text{-}5}GlcNAc_2$-PA structures. In contrast, the N327S mutant had a reduced rate of $Man_9GlcNAc_2$-PA cleavage to $Man_8GlcNAc_2$-PA, but the $Man_8GlcNAc_2$-PA did not accumulate to the same extent (maximal accumulation of 80%) before being further trimmed to $Man_7GlcNAc_2$-PA and trace amounts of $Man_{6\text{-}5}GlcNAc_2$-PA.

3.4.2.3. The influence of A590 and D591 on the substrate specificity of ERManI

ERManI was also found to interact through residues A590 and D591 with the M9 residue of $Man_9GlcNAc_2$-PA in the enzyme-substrate complex (Figure 34A). The topologically equivalent residues in GMIA are N540 and E541, and those residues had no interactions with the glycan in the GMIA enzyme-substrate complex. The digestion profile of the ERManI A590N mutant indicated that this mutation had a similar profile as wild type ERManI indicating that the

residue has little influence on substrate specificity (Figure 34B). The D591E of ERManI had a minor change of the substrate specificity (Figure 34C), where $Man_8GlcNAc_2$-PA accumulated to maximum of ~80% of the total oligosaccharide abundance and was very slowly converted to $Man_7GlcNAc_2$-PA and $Man_{6-5}GlcNAc_2$-PA structures.

3.4.2.4 The influence of residue D523 on the substrate specificity of ERManI

The side chain D523 interacts with the M6 residue in the ERManI-$Man_9GlcNAc_2$-PA complex whereas the topologically equivalent residue in GMIA is G474 does not interact with the glycan substrate in the GMIAI-$Man_9GlcNAc_2$-PA complex (Figure 35A). The D523G mutant of ERManI displayed a minor change in substrate specificity compared to wild type ERManI (Figure 35B). Similar to the D591E mutant, $Man_8GlcNAc_2$-PA accumulated in the digestion time-course for the D523G mutant to maximum of ~80% of the total oligosaccharide abundance and was very slowly converted to $Man_7GlcNAc_2$-PA and $Man_{6-5}GlcNAc_2$-PA structures.

3.4.2.5. The influence of R388 on the substrate specificity of ERManI

The side chain NH_2 group of R388 interacts with the M8 residue of $Man_9GlcNAc_2$-PA in the ERManI-substrate complex whereas the topologically equivalent P340 residue in GMIA does not interact with the substrate oligosaccharide (Figure 36A). The R388P mutant of ERManI had a similar digestion profile to wild type ERManI indicating that the residue has little influence on substrate specificity of the enzyme (Figure 36B). Conversion of $Man_9GlcNAc_2$-PA to $Man_8GlcNAc_2$-PA increased up to 100% of the total glycan abundance and remained at that plateau without being cleaved to the smaller structures during the duration of the time course.

3.4.2.6. The influence of multiple point mutants on the substrate specificity of ERManI

To compare the impact of swapping the full collection of glycan interacting residues between the two GH47 α1,2-mannosidases we generated a multi-site combinatorial mutant of

ERManI to the corresponding residues in GMIA. The combination of all of the individual mutations includes N327S, S375A, D376L, R388P, R461L, D523G, A590N, D591E, but not W389A. The latter Trp residue was conserved in both enzymes and had been mutated to Ala in the single mutants to test its contribution to ERManI activity. The combinatorial mutant of ERManI to GMIA interacting residues was examined in a substrate cleavage time course study. In contrast to wild type ERManI, the combinatorial ERManI mutant resembles the cleavage time course of wild type GMIA (Figure 37A). Cleavage of $Man_9GlcNAc_2$-PA resulted in an increase in $Man_8GlcNAc_2$-PA to a maximum of ~35% of total glycan abundance followed by conversion to $Man_7GlcNAc_2$-PA (maximum of ~25% of total glycan abundance) and subsequently $Man_6GlcNAc_2$-PA (maximum of 50% of total glycan abundance), before the slow conversion to a $Man_5GlcNAc_2$-PA structure. The slow conversion from $Man_6GlcNAc_2$-PA to $Man_5GlcNAc_2$-PA is similar to the wild type GMIA where inefficient cleavage of the M10 residue on branch B leads to the accumulation of the $Man_6GlcNAc_2$-PA structure. Thus, the conversion of all of the interacting residues in the >+1 glycan binding cleft of ERManI to the equivalent residues in GMIA also converts the apparent specificity for glycan cleavage to approach the specificity of wild type GMIA. However, the catalytic efficiency of the multiple residue mutant was considerably worse than wild type GMIA (~46-fold lower k_{cat} and a 16-fold higher K_m) indicating that additional constraints contribute to the efficient cleavage of high mannose glycans equivalent to the wild type GMIA activity (Table 3 Figure 38, and Figure 39, see below)

3.4.2.7. The influence of multiple point mutants on the substrate specificity of GMIA

While the combinatorial multi-mutant form of ERManI resulted in an activity resembling GMIA, a combinatorial mutant was also generated to swap the collection of glycan interacting residues in GMIA to the equivalent residues in ERManI. The combinatorial GMIA mutant

included S279N, A327S, L328D, P340R, L413R, G474D, N540A, and E541D. The time-course study of $Man_9GlcNAc_2$-PA glycan cleavage for the combinatorial mutant revealed alterations in the GMIA cleavage profile, but did not convert the profile to one resembling wild type ERManI (Figure 30A). Cleavage of $Man_9GlcNAc_2$-PA resulted in an increase in $Man_8GlcNAc_2$-PA to a maximum of ~45% of total glycan abundance followed by conversion to $Man_7GlcNAc_2$-PA (maximum of ~35% of total glycan abundance) and subsequently to $Man_6GlcNAc_2$-PA and $Man_5GlcNAc_2$-PA(Figure 37B). Compared to wild type GMIA (Figure 30B), the cleavage profile of the sequential mutant exhibited a slower cleavage of $Man_8GlcNAc_2$-PA, a slower increase of $Man_6GlcNAc_2$-PA, and a more rapid relative conversion to $Man_5GlcNAc_2$-PA. These differences in time course of cleavage are consistent with a conversion in specificity from GMIA to a hybrid activity that partially approximates the activity of wild type ERManI.

3.4.3. Isolation of $Man_8GlcNAc_2$-PA isomers generated by the mutant enzymes

To provide greater insight into the effect of the conversion in specificity for ERManI following mutagenesis, the $Man_8GlcNAc_2$-PA structure was isolated from an intermediate time point in the glycan cleavage time-course and resolved by HPLC to determine the isomer structure that was generated during the cleavage reaction. The HPLC profiles indicated that the compositions of the $Man_8GlcNAc_2$ isomers generated by the combinatorial mutants are different from those generated by the wild type enzymes (Figure 38). As predicted from prior substrate specificity studies [4, 8], $Man_8GlcNAc_2$-PA isomer B is the only isomer generated by wild type ERManI. The $Man_8GlcNAc_2$-PA isomer A (90%) and isomer C (10%) are both generated by wild type GMIA. In addition, the $Man_8GlcNAc_2$-PA isomer B is also the primary isomer produced by the ERManI W389A mutant.

The composition of the $Man_8GlcNAc_2$-PA isomers generated by the combinatorial point mutant of ERManI includes both $Man_8GlcNAc_2$-PA isomer B and isomer A (60% : 40%), but not isomer C. In combination with the greater cleavage to glycan structures smaller than $Man_8GlcNAc_2$-PA, these results indicate that the mutant ER enzyme gained broader substrate specificity more closely resembling GMIA, but the is clearly not identical to the activity of the wild type Golgi enzyme since the B isomer is still a significant component of the $Man_8GlcNAc_2$-PA isomer composition. Thus, the mutant enzyme displayed an activity that was a hybrid between the activities of ERManI and GMIA. Similarly, the combinatorial point mutant of GMIA was converted from a wild type activity that generated predominately $Man_8GlcNAc_2$-PA isomer A (90%) to an activity that produced a mixture of all three $Man_8GlcNAc_2$-PA isomers A (13%), B (39%) and C (48%) (Figure 38). Surprisingly, this isomer mixture (predominately isomer C, slightly lower amount of isomer B, and relatively little isomer A) was completely different from wild type GMIA (predominately isomer A) and also quite different from wild type ERManI (exclusively isomer B). These results indicate that the combinatorial mutant of GMIA generated an enzyme with hybrid specificity partially resembling ERManI, but also exhibiting a novel specificity for production of $Man_8GlcNAc_2$-PA isomer C unlike either of the wild type enzymes.

3.4.4. Kinetic analysis of the GH47 α1,2-mannosidase mutants

Although the mutated residues for both enzymes were all located within the ≥+1 subsites of the glycan binding cleft and were not expected to participate directly in glycosidic bond cleavage, kinetic analysis for initial rates of cleavage of $Man_9GlcNAc_2$-PA to $Man_8GlcNAc_2$-PA showed that the mutants generally compromised the catalytic activities of the resulting enzymes.

The catalytic activities were measured on the mutants that displayed substrate specificity change and kinetic constants are shown in Table 3.

Among the mutants, the combinatorial mutants exhibited the most compromised catalytic activity. The ERManI and GMIA multiple point mutants showed respective 120- and 62-fold decreases in k_{cat}, while K_m values were increased by 22- and 14-fold compared to their wild type counterparts. The single W389A ERManI mutant displayed a ~15-fold decrease in k_{cat} and a 9-fold increase in K_m. Similarly, the ERManI R461L mutant exhibited an 8-fold decrease in k_{cat} and a 6.5-fold increase in K_m. The other single mutants also showed a certain degree of compromised catalytic activity, but none were as compromised as the W389A and R461L mutants (Table 3). The apparent changes of the substrate specificities, along with the decreases of the k_{cat} values and increased K_m values for the R461L, W389A, and combinatorial mutants indicate that those altered residues not only changed the binding pattern of the substrate in the enzyme glycan binding clefts, but also weakened their enzymatic activities. The ratios of the relative k_{cat} and K_m values of the mutants compared to the wild type ERManI are presented in Figure 39.

3.4.5. Binding interactions to GH47 α1,2-mannosidase mutants studied by SPR.

To further study the influence of the glycan binding cleft mutations on the affinity of oligosaccharide interactions with ERManI, SPR experiments were performed using $Man_9GlcNAc_2$-PA as the binding analyte and the endogenous enzyme-bound Ca^{2+} ion swapped for La^{3+}. For wild type ERManI, and the W389A and W461A mutants a two-state conformation model was used to fit the data and calculate the binding constants since the curves fit the data more effectively than in a rigid 1:1 model. In contrast (Figure 41A, 41B and 41C), prior studies of glycan or inhibitor binding to wild type and catalytic -1 subsite mutants of ERManI

(containing bound Ca^{2+}) indicated an effective fit using the 1:1 model for the binding sensorgrams [2]. The two-state model for the La^{3+}-bound enzyme in the SPR studies may reflect the conformational change of the glycone residue from a $^4C_1$ solution conformation to the $^1C_4$ conformation required for entrance into the -1 subsite during binding of the uncleaved glycan in the present studies. For the combinatorial mutant the equilibrium binding values for the sensorgrams were employed (Figure 41D) since the rapid on- and off-rated precluded effective calculation of K_D values for the mutant enzyme. Wild type ERManI displayed a K_D of 14.5 μM for binding of the $Man_9GlcNAc_2$-PA ligand. In contrast, the K_D values of the R461L and W389A mutants increased to 198 μM (13-fold increase) and 474 μM (33-fold increase), respectively. These results indicate that alterations in interactions between the ERManI R461 and W389 residues and the glycan substrate compromises binding affinity and thus reduces catalysis by the mutant enzymes. The K_D for the combinatorial mutant increased to 2 mM (148-fold increase over wild type ERManI), suggesting that substrate binding affinity for the mutant enzyme was seriously compromised (Table 4).

3.5 Discussion

The time-course substrate digestion profiles of the ERManI and GMIA mutants revealed the influence of the glycan binding site residues on the substrate specificities of the enzymes. The structures of enzyme-$Man_9GlcNAc_2$-PA complexes demonstrated that the oligosaccharides bind in the respective enzyme glycan binding clefts with unique sets of interactions for each enzyme. It is reasonable to hypothesize that those interactions are the determinants for the different arrangement of the oligosaccharides in the ERManI and GMIA active site clefts and result in the respective differences in substrate specificity. The present studies support this hypothesis in that mutations of some of the cleft residues in ERManI altered its substrate

specificity to reflect an activity more resembling GMIA. Two of the single site mutants that stood out were R461L and W389A, where each mutant resulted in a hybrid substrate specificity between the wild type ERManI and GMIA. The structural features that distinguished those two residues from others were that they both contributed multiple interactions with the glycan core of the $Man_9GlcNAc_2$-PA substrate. The side chain NH_2 of R461 interacts with O-2 of M3, O-3 of NAG2 and the glycosidic oxygen atom between M3 and NAG2. The side chain aromatic group of W389 interacts with ring of NAG2 by hydrophobic stacking interactions, which are likely stronger contributions than single hydrogen bond. The results suggest that deletion of those interactions by mutation to the equivalent residues in GMIA weakened the anchoring interactions that hold the substrate in a fixed orientation in the ERManI active site and resulted in an enzyme with loosened specificity for selective branch recognition, but also reduced substrate binding affinity.

The results of the time-course substrate cleavage profiling indicated that the glycan specificity for the combinatorial mutants that swap residues between ERManI and GMIA were altered to result in hybrid activities that were intermediate between the two enzyme specificities. The ERManI combinatorial mutant gained more flexibility in substrate interactions to result in cleavage beyond $Man_8GlcNAc_2$-PA along with a partial generation of a $Man_8GlcNAc_2$ isomer A intermediate, both hallmarks of GMIA activity. In contrast, the GMIA combinatorial mutant did not display the narrow substrate specificity of wild type ERManI, but the enzyme did have altered specificity for the $Man_8GlcNAc_2$ isomer produced, partially generating the $Man_8GlcNAc_2$ isomer B intermediate in addition to the novel production of the isomer C unlike either wild type enzyme. The mutant GMIA enzyme also more effectively cleaves $Man_6GlcNAc_2$-PA to $Man_5GlcNAc_2$-PA. Thus, the mutant exhibits some characteistics of ERManI enzyme specificity.

The results from the enzyme kinetic assays and binding affinity studies are also consistent with the substrate specificity results. With the broadened substrate specificity, the R461L, W389A and combinatorial mutant enzymes all displayed decreased substrate binding affinities and compromised enzyme activities. Thus, the systematic characterization of the substrate specificity, enzyme kinetics, and binding affinity for mutants of ERManI and GMIA have provided complementary insights into the structural and enzymatic contributions to substrate recognition by the GH47 α1,2-mannosidases.

Figure 29. The orientation of $Man_9GlcNAc_2$-PA in human ERManI and mouse GMIA active sites. **Panel A**: Stick presentation of $Man_9GlcNAc_2$-PA in the ERManI active site indicates that residue M10 of branch B was in the +1 subsite of the enzyme; **Panel B**: Stick presentation of $Man_9GlcNAc_2$-PA in the GMIA active site indicates that residue M11 of branch A was in the +1 subsite of the enzyme; **Panel C**: Stick presentation of the superimposition of $Man_9GlcNAc_2$-PA glycan structure (green) found in the active site of ERManI or GMIA (teal); **Panel D**: A cartoon depiction of $Man_9GlcNAc_2$ in coloring according to the nomenclature of the Consortium for Functional Glycomics. The enzymes in **Panels A, B,** and **C** are in gray cartoon representation, and the La^{3+} metal ion is in blue sphere representation. **Panels A, B**, and **C** were generated using the Pymol program with the PDB files of human ERManI-$Man_9GlcNAc_2$ complex and GMIA-$Man_9GlcNAc_2$ complex.

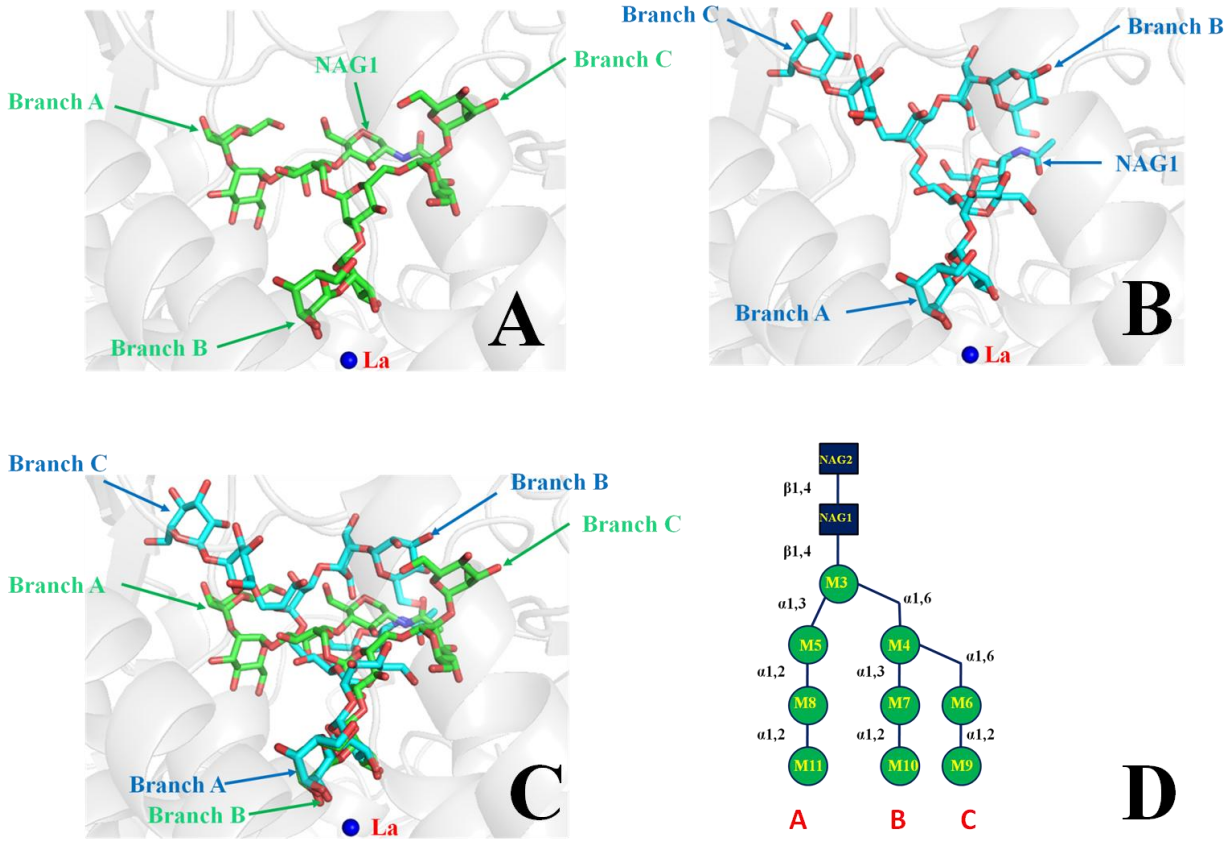

Branch A
NAG1
Branch C
Branch B
La
A
Branch C
Branch B
NAG1
Branch A
La
B
Branch C
Branch B
Branch C
Branch A
Branch A
Branch B
La
C
NAG2
β1,4
NAG1
β1,4
M3
α1,3
α1,6
M5
M4
α1,6
α1,2
α1,3
M8
M7
M6
α1,2
α1,2
α1,2
M11
M10
M9
A
B
C
D

Figure 30. Time-course digestion profiles of $Man_9GlcNAc_2$-PA by human ERManI (left) and GMIA (right). The profile of digestion of $Man_9GlcNAc_2$-PA *in vitro* by ERManI shows the relative abundance curves for $Man_9GlcNAc_2$-PA and $Man_8GlcNAc_2$-PA in the time-course. The $Man_8GlcNAc_2$-PA was generated efficiently with the corresponding disappearance of $Man_9GlcNAc_2$-PA, while the smaller oligosaccharides, $Man_{7-5}GlcNAc_2$-PA, were generated very slowly and could not be observed by HPLC analysis in the given duration and enzyme concentration in the experiment. The digestion profile of $Man_9GlcNAc_2$-PA *in vitro* by GMIA indicates a conversion from $Man_9GlcNAc_2$-PA to $Man_5GlcNAc_2$-PA. The $Man_9GlcNAc_2$-PA substrate is rapidly converted to $Man_6GlcNAc_2$-PA and slowly cleaved further to $Man_5GlcNAc_2$-PA.

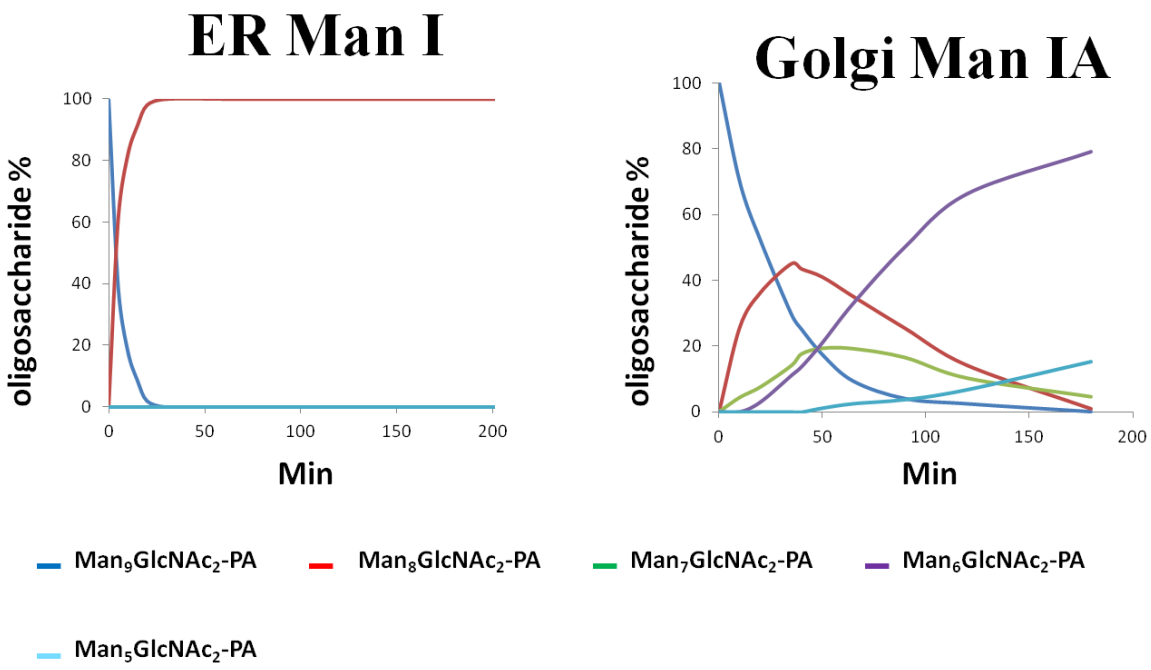
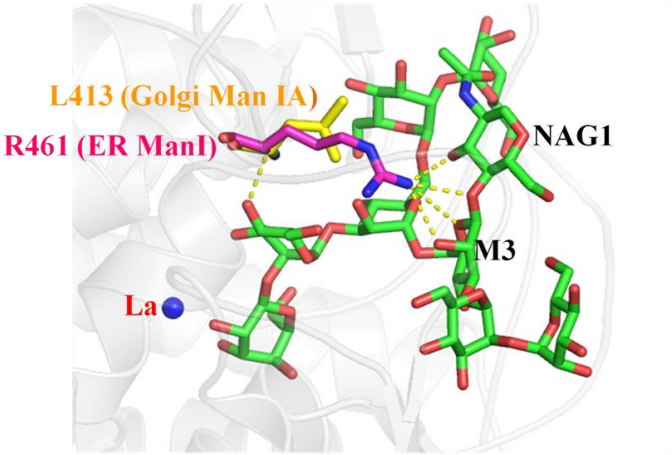

ER Man I
oligosaccharide %
100
80
60
40
20
0
0
50
100
150
200
Min
Golgi Man IA
oligosaccharide %
100
80
60
40
20
0
0
50
100
150
200
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 31. Time course of $Man_9GlcNAc_2$-PA digestion by the R461L mutant of ERManI changes the substrate specificity profile to an activity resembling GMIA. Panel A: The cartoon presentation of the ERManI active site (grey) showing the side chain of R461 in ERManI (magenta stick form) *versus* the corresponding side chain of L413 in GMIA (yellow stick form) interacting with the $Man_9GlcNAc_2$-PA substrate glycan (green stick form). Hydrogen bonding interactions with the glycan substrate are shown in yellow dotted line. The hydrogen bonding interactions with the R461 side chain in ERManI are not present in the corresponding GMIA substrate complex. **Panel B**: The time-course digestion profile of $Man_9GlcNAc_2$-PA by the ERManI R461L mutant. The profile indicates a hybrid substrate specificity for the mutant enzyme between those of wild type ERManI and wild type GMIA.

A


L413 (Golgi Man IA)
R461 (ER ManI)
NAG1
M3
La

B

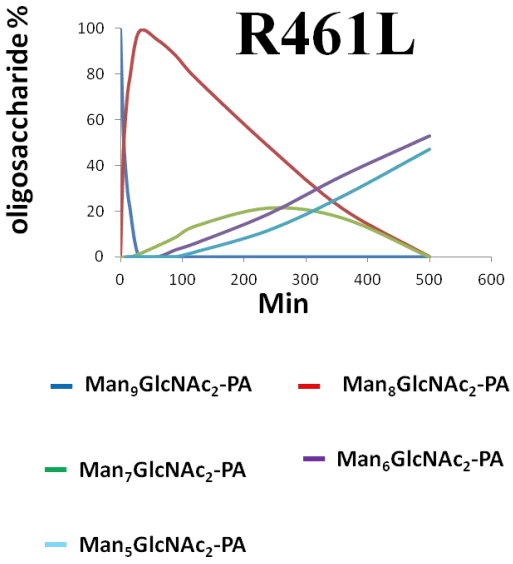

R461L
oligosaccharide %
100
80
60
40
20
0
0
100
200
300
400
500
600
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 32. Mutation of Trp389 to Ala in the ERManI glycan binding site cleft changes the substrate specificity of the ER enzyme. Panel A: The cartoon representation of the ERManI active site (grey) showing the side chain of W389 in ERManI (magenta stick form) *versus* the corresponding side chain of W341 in GMIA (yellow stick form) interacting with the $Man_9GlcNAc_2$-PA substrate glycan (green stick form). Hydrophobic interactions with the glycan substrate are shown with yellow dotted lines. The hydrophobic bonding interactions with the W389 side chain in ERManI are not present in the corresponding GMIA substrate complex. **Panel B**: The time course digestion profile of $Man_9GlcNAc_2$-PA by the ERManI W389A mutant. The profile indicates a hybrid substrate specificity for the mutant enzyme similar to the specificity of wild type GMIA.

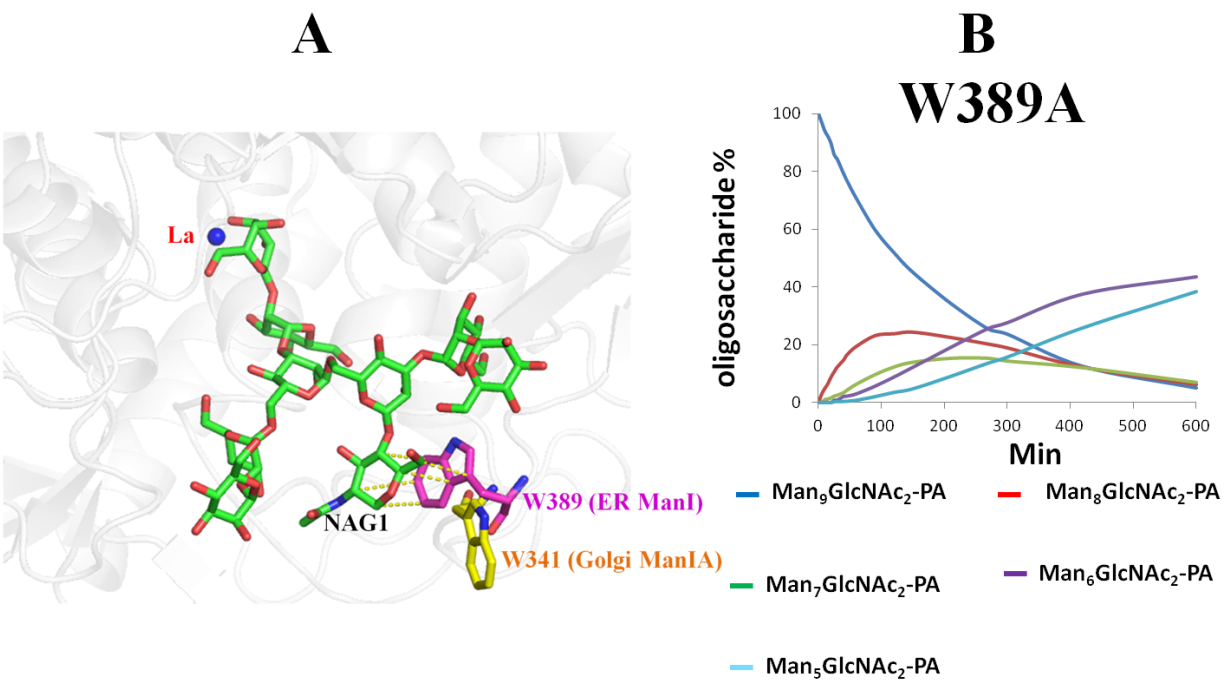

A
B
W389A
La
NAG1
W389 (ER ManI)
W341 (Golgi ManIA)
oligosaccharide %
100
80
60
40
20
0
0
100
200
300
400
500
600
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 33. Mutations of residues N327, S375, and D376 of the ERManI enzyme to the equivalent residues in the GMIA change the substrate specificity of the ER enzyme. **Panel A**: The cartoon representation of the ERManI active site (grey) showing the side chains of N327, S375, and D376 in ERManI (magenta stick form) *versus* the corresponding side chains of S279, A327, L328 in GMIA (yellow stick form) interacting with the $Man_9GlcNAc_2$-PA substrate glycan (green stick form). Hydrogen bonding interactions with the glycan substrate are shown with yellow dotted lines. The hydrogen bonding interactions with the substrate in ERManI are not present in the corresponding GMIA substrate complex. The time course digestion profiles of $Man_9GlcNAc_2$-PA by the ERManI D376L (**Panel B**), S375A (**Panel C**) and N327S (**Panel D**) mutants are shown in the panels at the right. The profiles indicate partial conversion in substrate specificity for the mutant enzymes between those of wild type ERManI and wild type GMIA.

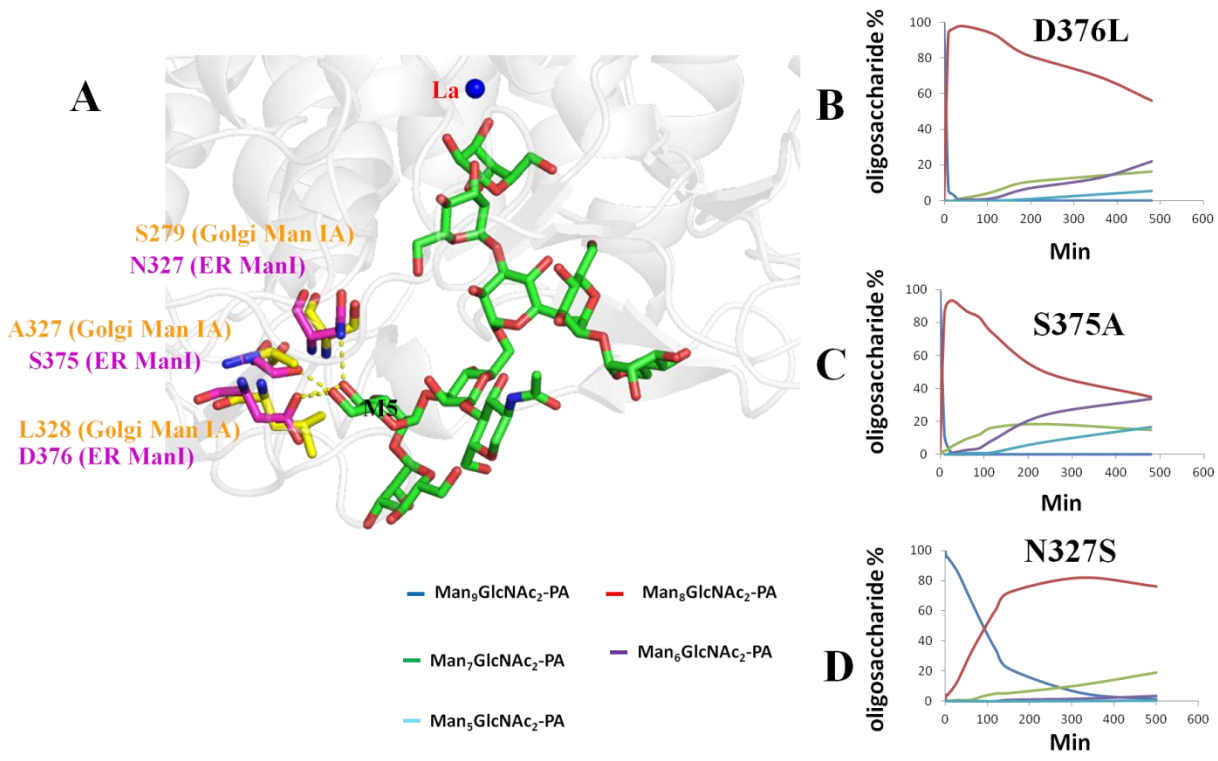

A
La
S279 (Golgi Man IA)
N327 (ER ManI)
A327 (Golgi Man IA)
S375 (ER ManI)
L328 (Golgi Man IA)
D376 (ER ManI)
M5
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA
B
D376L
C
S375A
D
N327S
oligosaccharide %
Min
0
20
40
60
80
100
0
100
200
300
400
500
600

Figure 34. Mutation of A590 to Asn in the ERManI glycan binding site cleft does not affect the substrate specificity of the enzyme, while mutation of D591 to Gln in the ERManI glycan binding site cleft changes the substrate specificity of the ER enzyme. **Panel A**: The cartoon presentation of the ERManI active site (grey) showing the side chains of A590 and D591 in ERManI (magenta stick form) *versus* the corresponding side chains of N540 and E541 in GMIA (yellow stick form) interacting with the $Man_9GlcNAc_2$-PA substrate glycan (green stick form). Hydrogen bonding interactions with the glycan substrate are shown with yellow dotted lines. The hydrogen bonding interactions with glycan substrate are not present in the corresponding GMIA substrate complex. The time course digestion profiles of $Man_9GlcNAc_2$-PA by the ERManI A590N (**Panel B**) and D591E (**Panel C**) mutants are shown at the right of the figure. The profile of D591E indicates the minor change of the substrate specificity for the mutants with a hybrid character between wild type ERManI and GMIA.

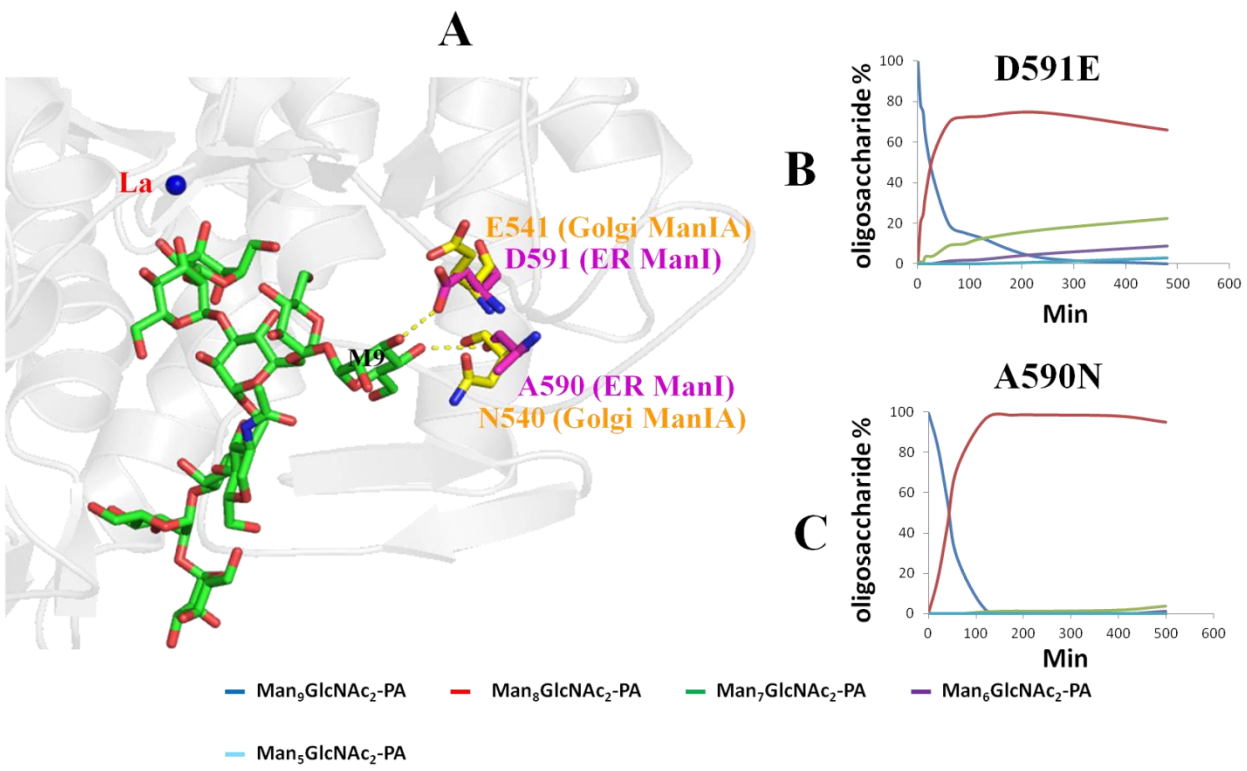

A
La
E541 (Golgi ManIA)
D591 (ER ManI)
M9
A590 (ER ManI)
N540 (Golgi ManIA)
B
D591E
oligosaccharide %
Min
C
A590N
oligosaccharide %
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 35. Mutation of D523 to Gly in the ERManI glycan binding site cleft influences the substrate specificity of the ER enzyme. **Panel A**: The cartoon presentation of the ERManI active site (grey) showing the side chain of D523 in ERManI (magenta stick form) *versus* the corresponding side chain of G474 in GMIA (yellow stick form) interacting with the $Man_9GlcNAc_2$-PA substrate glycan (green stick form). Hydrogen bonding interactions with the glycan substrate are shown with yellow dotted lines. The hydrogen bonding interactions with the D523 side chain in ERManI are not present in the corresponding GMIA substrate complex. **Panel B**: The time course digestion profile of $Man_9GlcNAc_2$-PA by the ERManI D523G mutant. The profile indicates a minor change in substrate specificity similar to the specificity of wild type ERManI.

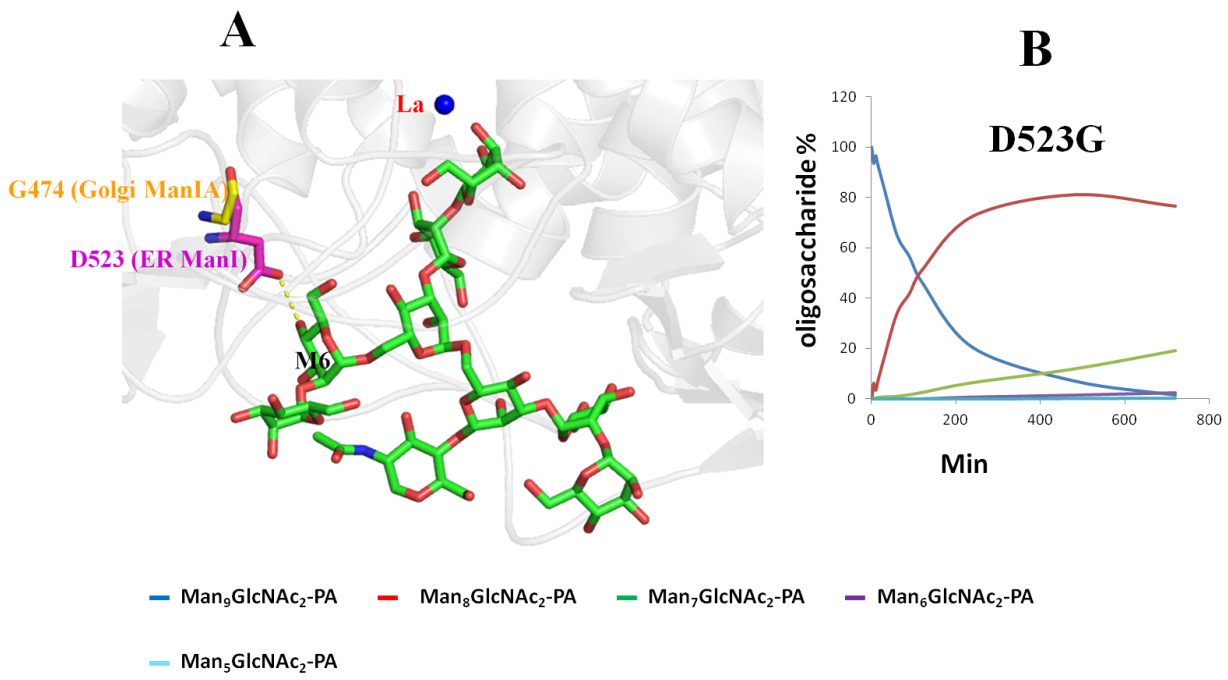

A
La
G474 (Golgi ManIA)
D523 (ER ManI)
M6
B
D523G
oligosaccharide %
120
100
80
60
40
20
0
0
200
400
600
800
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 36. Mutation of R388 to Pro in the ER Man I glycan binding site cleft has no influence on the substrate specificity of the ER enzyme. **Panel A**: The cartoon presentation of the ERManI active site (grey) showing the side chain of R388 in ERManI (magenta stick form) *versus* the corresponding side chain of P340 in GMIA (yellow stick form) interacting with the $Man_9GlcNAc_2$-PA substrate glycan (green stick form). Hydrogen bonding interactions with the glycan substrate are shown with yellow dotted lines. The hydrogen bonding interactions with the R388 side chain in ERManI are not present in the corresponding GMIA substrate complex. **Panel B**: The time course digestion profile of $Man_9GlcNAc_2$-PA by the ERManI R388P mutant. The profile indicates a substrate specificity for the mutant enzyme similar to the specificity of wild type ERManI.

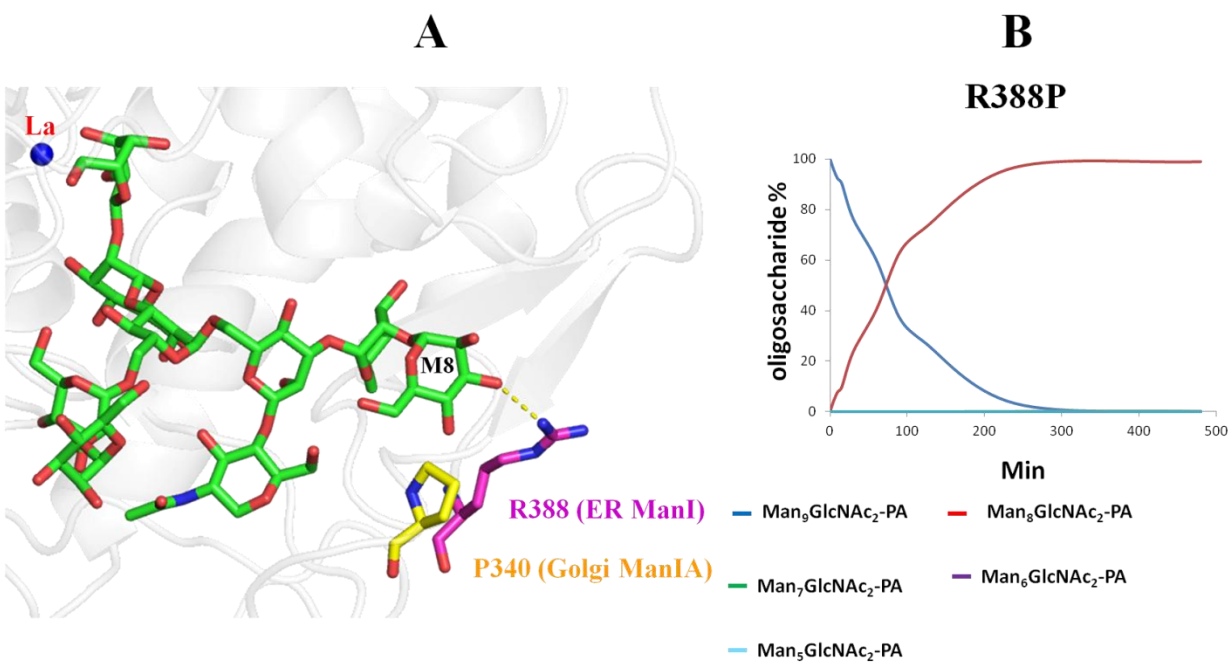

A
B
La
M8
R388 (ER ManI)
P340 (Golgi ManIA)
R388P
oligosaccharide %
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 37. Time course digestion profiles of $Man_9GlcNAc_2$-PA by the combinatorial point mutant of human ERManI (Panel A) and combinatorial point mutant of GMIA (Panel B). Both mutants display hybrid substrate specificities between those of the wild type ERManI and GMIA. The digestion profiles by the two mutant enzymes showed the appearance of $Man_{8-5}GlcNAc_2$-PA with the disappearance of $Man_9GlcNAc_2$-PA. In the glycan digestion by the ERManI combinatorial mutant, the $Man_8GlcNAc_2$-PA product was further digested to $Man_{7-5}GlcNAc_2$-PA, unlike the wild type ERManI enzyme activity. In the digestion by the GMIA combinatorial mutant, the $Man_5GlcNAc_2$-PA was readily generated from $Man_6GlcNAc_2$-PA, which was not observed in the digestion by wild type GMIA.

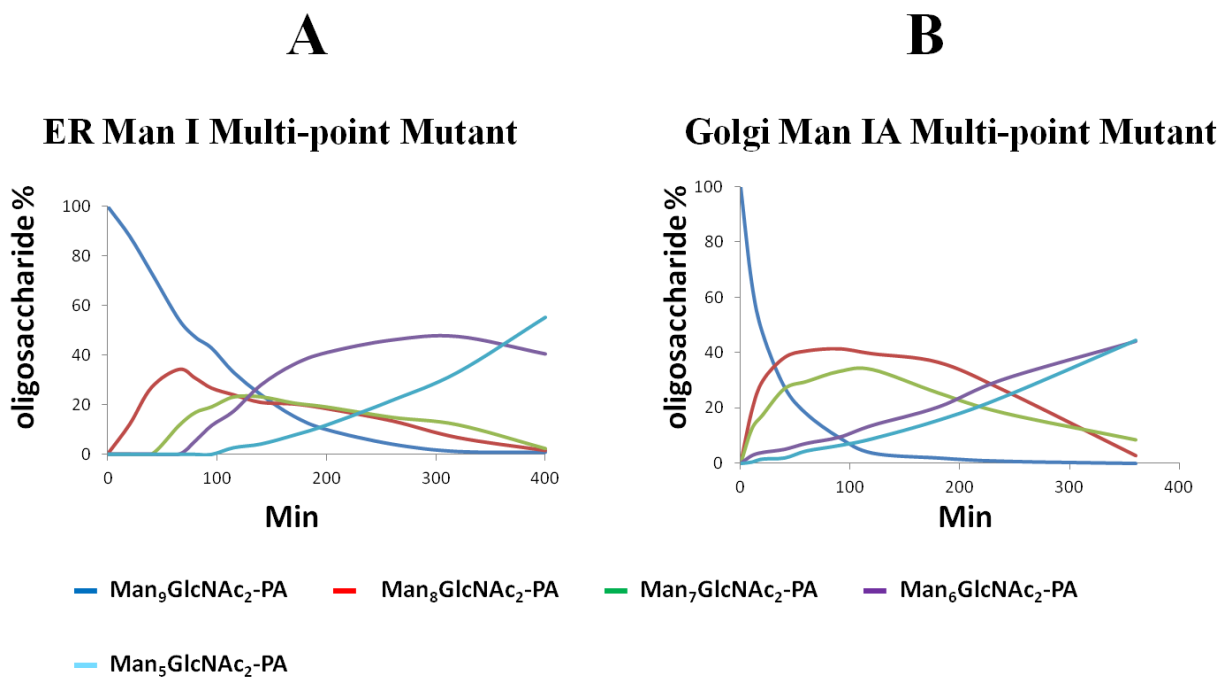

A
B
ER Man I Multi-point Mutant
Golgi Man IA Multi-point Mutant
oligosaccharide %
Min
Man9GlcNAc2-PA
Man8GlcNAc2-PA
Man7GlcNAc2-PA
Man6GlcNAc2-PA
Man5GlcNAc2-PA

Figure 38. The $Man_8GlcNAc_2$-PA isomers generated by the GH47 α1,2-mannosidases. Chromatographic C-18 HPLC profiles of the $Man_8GlcNAc_2$-PA isomers generated by wild type and mutant enzymes are indicated for (from top to bottom) ERManI, wild type GMIA, ERManI W389A mutant, ERManI combinatorial mutant, and GMIA combinatorial mutant. The retention time for the $Man_8GlcNAc_2$-PA isomers A, B, C are ~6.5 min, ~5.5min and ~8.5 min, respectively. The percentages of each isomer were calculated by their relative peak areas: wild type ERManI produced isomer B (100%); wild type GMIA produced isomer A (90%) and isomer C (10%); ERManI W389A mutant produced isomer B (100%); ERManI combinatorial mutant produced isomer A (60%), isomer B (40%); GMIA combinatorial mutant produced isomer A (13%), isomer B (39%) and isomer C (48%).

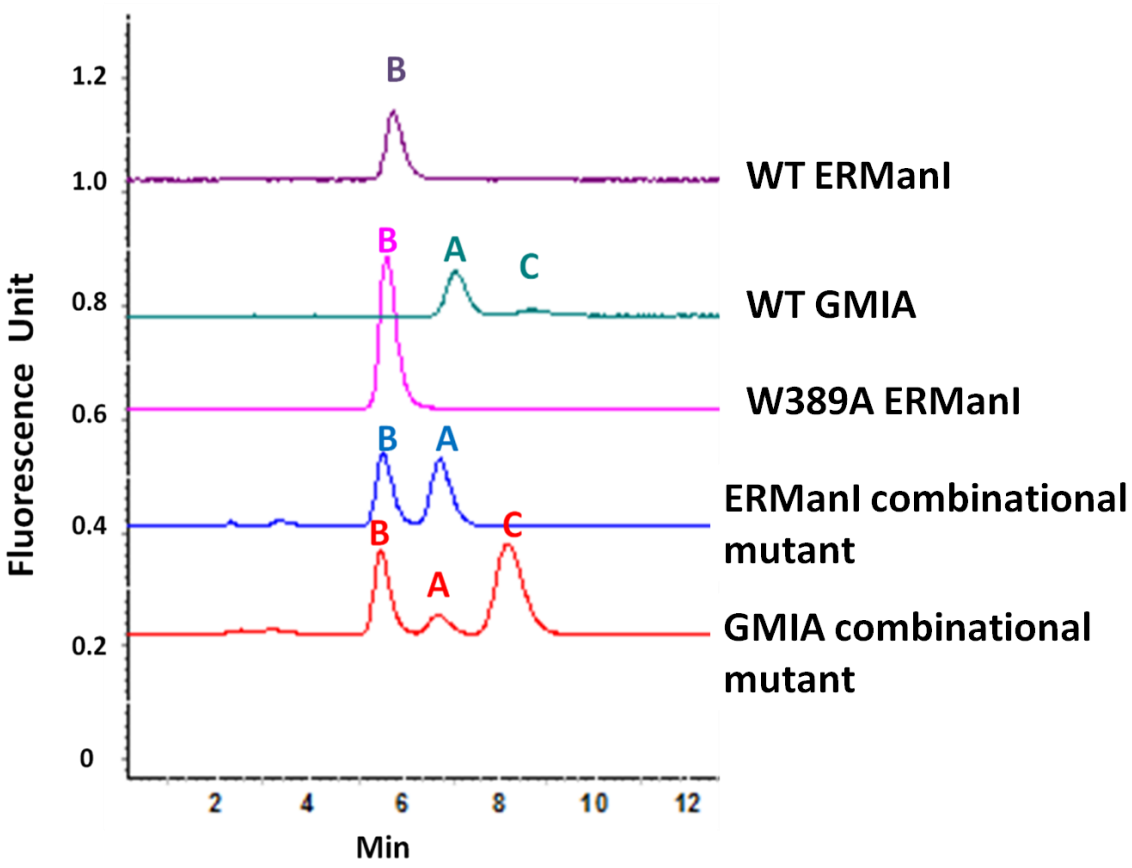
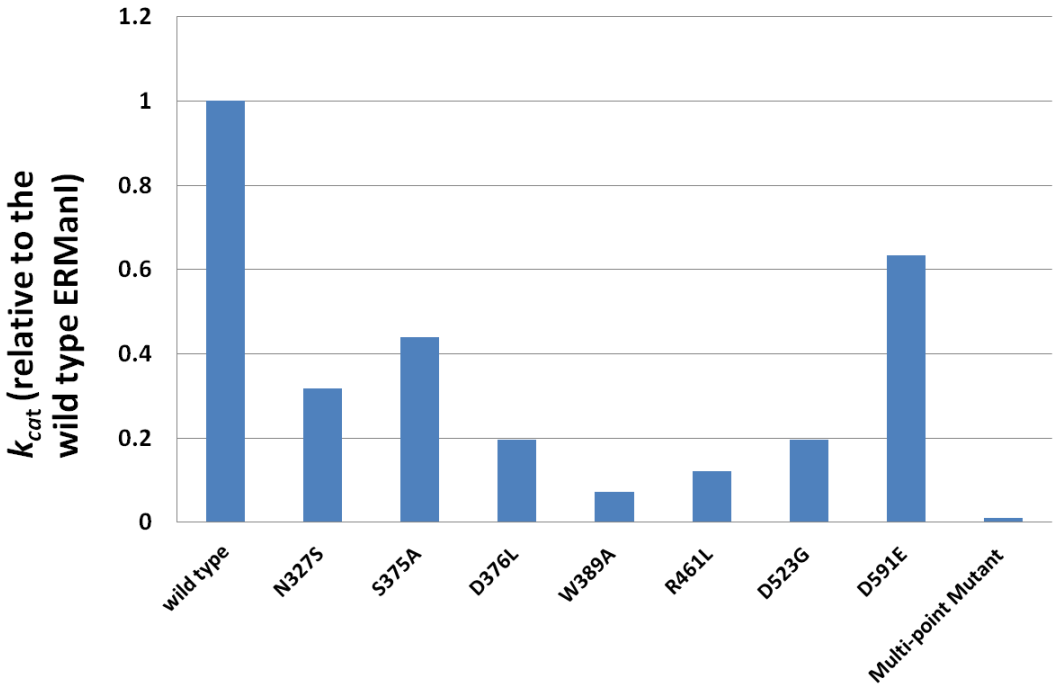

Fluorescence Unit
1.2
1.0
0.8
0.6
0.4
0.2
0
2
4
6
8
10
12
Min
B
B
A
C
B
A
B
C
A
WT ERManI
WT GMIA
W389A ERManI
ERManI combinational mutant
GMIA combinational mutant

Table 3. Kinetic constants for wild type and mutant human ERManI and mouse GMIA using $Man_9GlcNAc_2$-PA as substrate.

	Km (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}/μM)	$(k_{cat}/K_m)/(k_{cat}/K_m)_{wt}$
ERManI(wt)	104 ± 9	4.12± 0.06	0.039	1.00
N327S	359 ± 31	1.26 ± 0.10	0.0036	0.092
S375A	455 ± 62	1.80 ± 0.30	0.0039	0.10
D376L	371 ± 22	0.79 ± 0.10	0.0021	0.054
W389A	970 ± 45	0.28 ± 0.09	0.00031	0.0078
R461L	680 ± 80	0.52 ± 0.19	0.00073	0.018
D523G	204 ± 10	0.83 ± 0.20	0.0039	0.099
D591E	277 ± 14	2.58 ± 0.05	0.0093	0.024
Multi-point Mutant	2257 ± 390	0.035 ±0.08	0.000016	0.0039
GMIA (wt)	142 ± 15	1.6 ± 0.3	0.011	1.0
Multi-point Mutant	1985 ± 202	0.026	0.000013	0.0012

Figure 39. The plots of relative k_{cat} and K_m values for wild type and mutant forms of ERManI . The k_{cat} and K_m values for wild type and mutant forms of ERManI in Table 3 were converted to values relative to wild type ERManI and plotted as relative k_{cat} values (**Panel A**) and relative K_m values (**Panel B**).

A


1.2
1
0.8
0.6
0.4
0.2
0
k_{cat} (relative to the wild type ERManI)
wild type
N327S
S375A
D376L
W389A
R461L
D523G
D591E
Multi-point Mutant

B

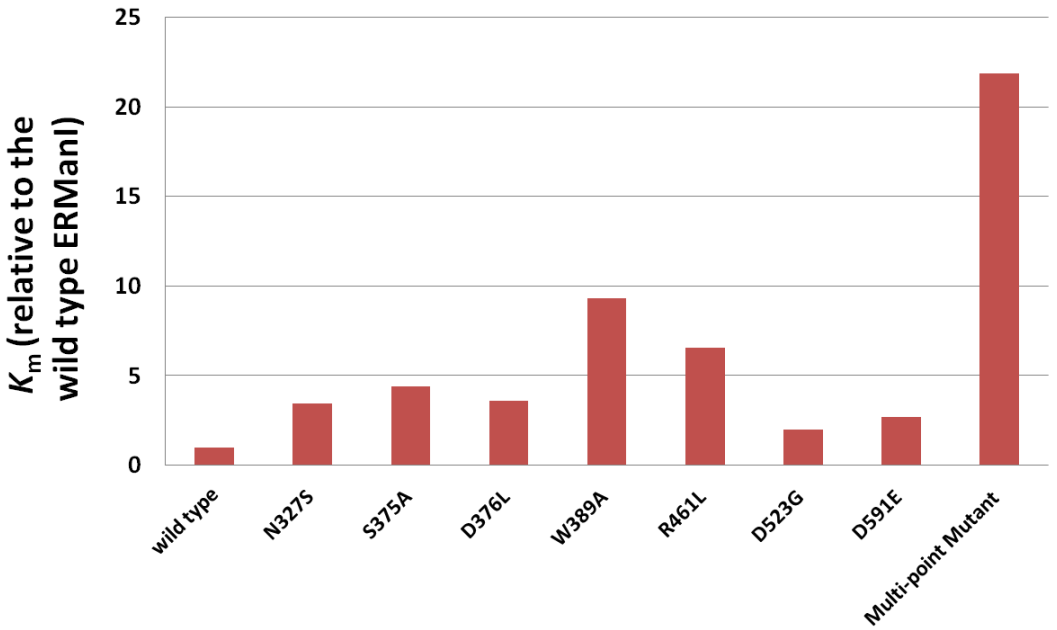

25
20
15
10
5
0
K_m (relative to the wild type ERManI)
wild type
N327S
S375A
D376L
W389A
R461L
D523G
D591E
Multi-point Mutant

Figure 40. Kinetic analysis for the digestion of $Man_9GlcNAc_2$-PA by the wild type and mutant forms of ERManI. The enzymatic reactions were performed using $Man_9GlcNAc_2$-PA as the substrate. The samples taken from the enzymatic reactions at different time points were used for HPLC separtation. The HPLC peak areas of the corresponding oligosaccharides in the reaction mixtures were used to calculate the abundance of the $Man_9GlcNAc_2$-PA substrate and the $Man_{8-5}GlcNAc_2$-PA products. The decreasing abundance of $Man_9GlcNAc_2$-PA was used in a time point prior to the appearance of $Man_{7-5}GlcNAc_2$-PA to determine the initial reaction rates for both wild type and mutant ERManI. Initial rates (v) for the enzyme reactions were determined at various substrate concentrations raging from 10 μM to 1600 μM. Catalytic coefficient (k_{cat}) and Michaelis constant (K_m) values were determined by fitting initial rates to the Michaelis-Menten function by nonlinear regression analysis using SigmaPlot (Jandel Scientific, San Rafael, CA). The enzymes characterized in this assay include: the wild type ERManI (**Panel A**), the mutant N327S (**Panel B**), the mutant S375A (**Panel C**), the mutant D376L (**Panel D**), the mutant W389A (**Panel E**), the mutant R461L (**Panel F**), the mutant D523G (**Panel G**), the mutant D591E (**Panel H**), and the combinatorial mutant of ERManI (**Panel I**).

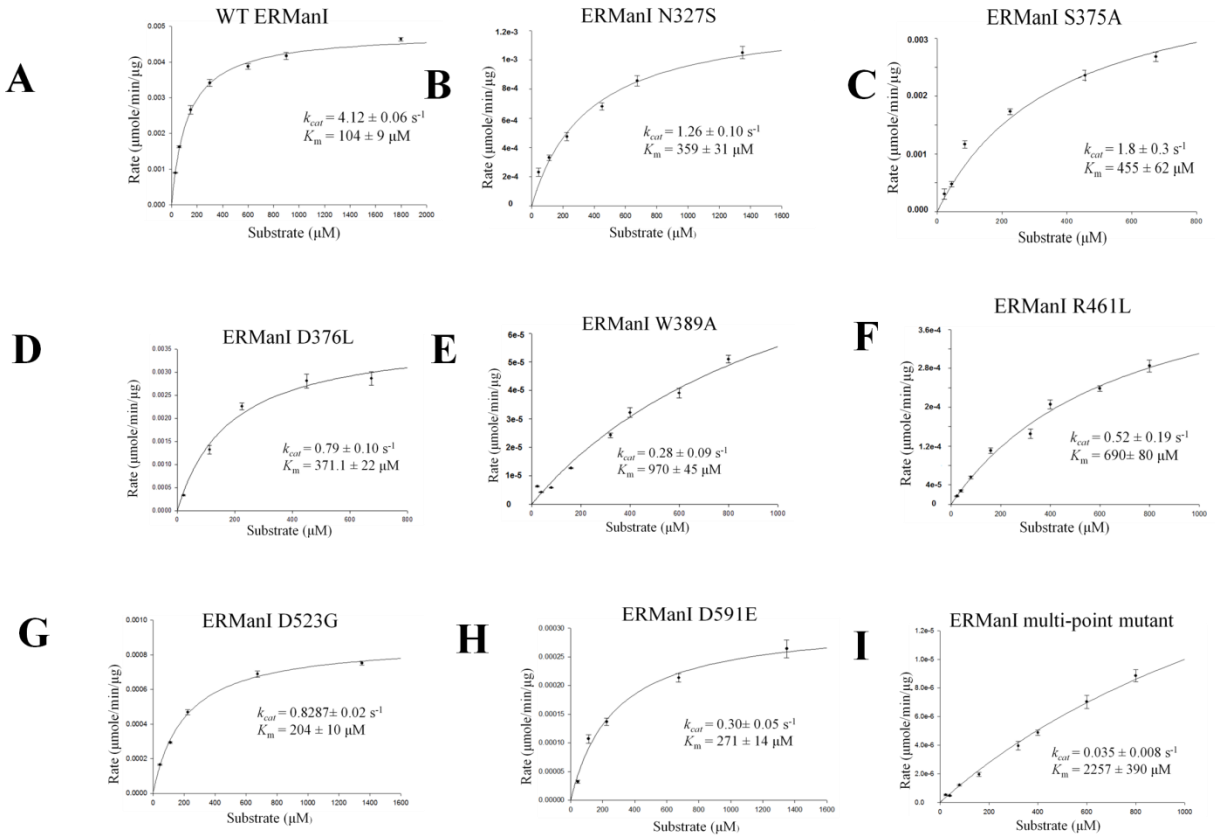

A
WT ERManI
k_{cat} = 4.12 ± 0.06 s^{-1}
K_m = 104 ± 9 µM
Rate (µmole/min/µg)
Substrate (µM)
B
ERManI N327S
k_{cat} = 1.26 ± 0.10 s^{-1}
K_m = 359 ± 31 µM
Rate (µmole/min/µg)
Substrate (µM)
C
ERManI S375A
k_{cat} = 1.8 ± 0.3 s^{-1}
K_m = 455 ± 62 µM
Rate (µmole/min/µg)
Substrate (µM)
D
ERManI D376L
k_{cat} = 0.79 ± 0.10 s^{-1}
K_m = 371.1 ± 22 µM
Rate (µmole/min/µg)
Substrate (µM)
E
ERManI W389A
k_{cat} = 0.28 ± 0.09 s^{-1}
K_m = 970 ± 45 µM
Rate (µmole/min/µg)
Substrate (µM)
F
ERManI R461L
k_{cat} = 0.52 ± 0.19 s^{-1}
K_m = 690± 80 µM
Rate (µmole/min/µg)
Substrate (µM)
G
ERManI D523G
k_{cat} = 0.8287± 0.02 s^{-1}
K_m = 204 ± 10 µM
Rate (µmole/min/µg)
Substrate (µM)
H
ERManI D591E
k_{cat} = 0.30± 0.05 s^{-1}
K_m = 271 ± 14 µM
Rate (µmole/min/µg)
Substrate (µM)
I
ERManI multi-point mutant
k_{cat} = 0.035 ± 0.008 s^{-1}
K_m = 2257 ± 390 µM
Rate (µmole/min/µg)
Substrate (µM)

Figure 41. SPR analyses of $Man_9GlcNAc_2$-PA binding to immobilized wild type or forms of mutant ERManI. Wild type or mutant ERManI was immobilized on a Biacore CM-5 chip as described in "Materials and Methods". Various concentration of $Man_9GlcNAc_2$-PA were tested for interactions with immobilized enzymes: wild type ERManI (**Panel A**), ERManI R46IL mutant (**Panel B**), ERManI W389A mutant (**Panel C**) and the combinatorial ERManI mutant (**Panel D**). The data were fit with the two-state kinetic model for wild type ERManI, R461L, and W389. The data for the combinatorial mutant was fit using the equilibrium sensorgram values and plotted as a saturation curve to determine the equilibrium binding constant, K_D. All curve fitting models were generated using software provided with the Biaocore® instrument (GE Healthcare).

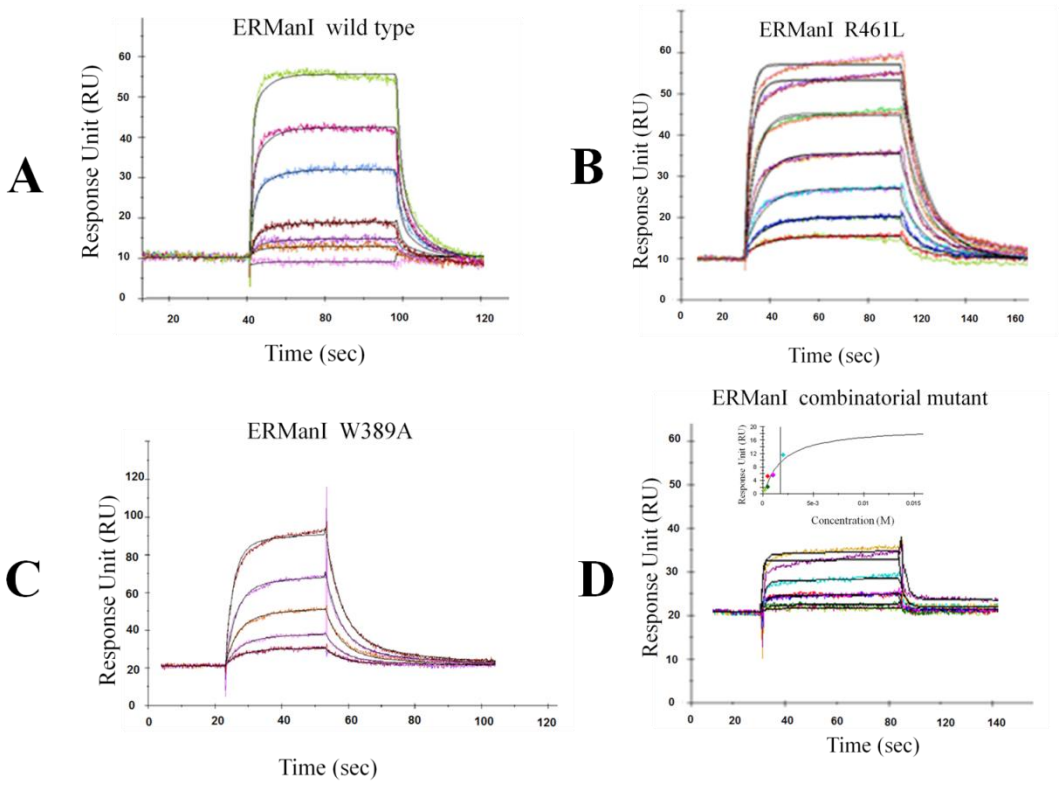

A
ERManI wild type
Response Unit (RU)
Time (sec)
B
ERManI R461L
Response Unit (RU)
Time (sec)
C
ERManI W389A
Response Unit (RU)
Time (sec)
D
ERManI combinatorial mutant
Response Unit (RU)
Concentration (M)
Time (sec)

Table 4. Binding affinity values determined by surface plasmon resonance for wild type and mutant forms of human ERManI

Enzyme	ligand	K_D (μM)	$K_D/K_{D\,(wt)}$
Wild type	$Man_9GlcNAc_2$-PA	14.5	1.00
R461L	$Man_9GlcNAc_2$-PA	198.8	13.8
W389A	$Man_9GlcNAc_2$-PA	474.2	32.7
Combinatorial mutant	$Man_9GlcNAc_2$-PA	2156	148.6

References:

1. Henrissat, B., *A classification of glycosyl hydrolases based on amino acid sequence similarities.* Biochem J, 1991. **280 (Pt 2)**: p. 309-16.
2. Karaveg, K. and K.W. Moremen, *Energetics of substrate binding and catalysis by class 1 (glycosylhydrolase family 47) alpha-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(33): p. 29837-48.
3. Karaveg, K., et al., *Mechanism of class 1 (glycosylhydrolase family 47) {alpha}-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(16): p. 16197-207.
4. Lal, A., et al., *Substrate specificities of recombinant murine Golgi alpha1, 2-mannosidases IA and IB and comparison with endoplasmic reticulum and Golgi processing alpha1,2-mannosidases.* Glycobiology, 1998. **8**(10): p. 981-95.
5. Tempel, W., et al., *Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases.* J Biol Chem, 2004. **279**(28): p. 29774-86.
6. Vallee, F., et al., *Structural basis for catalysis and inhibition of N-glycan processing class I alpha 1,2-mannosidases.* J Biol Chem, 2000. **275**(52): p. 41287-98.
7. Vallee, F., et al., *Crystal structure of a class I alpha1,2-mannosidase involved in N-glycan processing and endoplasmic reticulum quality control.* EMBO J, 2000. **19**(4): p. 581-8.
8. Gonzalez, D.S., et al., *Identification, expression, and characterization of a cDNA encoding human endoplasmic reticulum mannosidase I, the enzyme that catalyzes the*

first mannose trimming step in mammalian Asn-linked oligosaccharide biosynthesis. J Biol Chem, 1999. **274**(30): p. 21375-86.

9. Romero, P.A., et al., *Mutation of Arg(273) to Leu alters the specificity of the yeast N-glycan processing class I alpha1,2-mannosidase.* J Biol Chem, 2000. **275**(15): p. 11071-4.

10. Franco-Fraguas, L., et al., *Preparative purification of soybean agglutinin by affinity chromatography and its immobilization for polysaccharide isolation.* J Chromatogr B Analyt Technol Biomed Life Sci, 2003. **790**(1-2): p. 365-72.

11. Hase, S., *Analysis of sugar chains by pyridylamination.* Methods Mol Biol, 1993. **14**: p. 69-80.

12. Hase, S., *High-performance liquid chromatography of pyridylaminated saccharides.* Methods Enzymol, 1994. **230**: p. 225-37.

13. Tokugawa, K., S. Oguri, and M. Takeuchi, *Large scale preparation of PA-oligosaccharides from glycoproteins using an improved extraction method.* Glycoconj J, 1996. **13**(4): p. iv.

CHAPTER 4

Profiling of selective inhibitors for GH47 α1,2-mannosidases

4.1. Abstract

Potent inhibitors for GH47 α1,2-mannosidases have been synthesized and used for profiling of selectivity of enzyme inhibition based on the structural framework 1-deoxymannojirimycin (dMNJ) and kifunensine (Kif) as archetypes of high affinity inhibitors. In general these compounds display poor selectivity in inhibition of endoplasmic reticulum (ER) α-mannosidase I (ERManI) *versus* Golgi α-mannosidase IA (GMIA). Inhibitors with potential for clinical use in the treatment of unfolded protein disorders should show strong inhibition towards ERManI as a result of its role in ER-associated glycoprotein degradation (ERAD), but minimal inhibition towards GMIA based on its role in glycoprotein maturation. To address this challenge libraries of chemically modified Kif and dMNJ derivatives were synthesized and screened for inhibition of ERManI and GMIA. Varying degrees of selectivity of inhibition were found among several Kif and dMNJ derivatives, among which a pyridine functionalized Kif derivative (**21**) and a 3-methyl-thiophene functionalized dMNJ derivative (**49**) inhibited ERManI over GMIA by 10- and 32.5-fold, respectively. Comparison of the structures and the inhibition selectivities of the derivatives provide clues to the further structural modifications of Kif and DMNJ to achieve better inhibition selectivity towards ERManI.

4.2. Introduction

Polyhydroxylated iminosugars are widely known as glycosidase inhibitors and have demonstrated promising potential as therapeutic agents [1-3]. Among them are the α1,2-mannosidase inhibitors, dMNJ and Kif [4], that show potential in the treatment of ERAD-related diseases such as lysosomal storage disorders (LSDs) [5, 6] and sarcoglycannopathies [7-9].

Several lines of evidence have demonstrated that the initial steps in ER-associated degradation (ERAD) of misfolded glycoproteins are catalyzed by enzymes that cleave specific monosaccharide residues from *N*-glycan structures. The enzymes catalyzing these reactions are members of a multigene family of α1,2-mannosidases (CAZy family 47 glycoside hydrolases (GH47) [10-12], which play pivotal roles in both *N*-glycan maturation in the ER and Golgi complex as well as decisions for glycoprotein disposal in the ER. Recognition of misfolded glycoproteins is initiated through a timing mechanism that involves trimming of glycan structures by a subset of the GH47 α-mannosidases, including ERManI and a collection of EDEM proteins (EDEM1-3 in mammals). The trimming steps are followed by recognition of the trimmed glycan structure by the ER luminal lectins (Yos9 or XTP3-B), transfer to a retrotranslation complex in the ER membrane (termed the HRD complex) with the mechanical assistance of a cytoplasmic AAA-ATPase complex, and finally degradation by the cytoplasmic proteasome ([13-16]). The proximal step in this process is the glycan cleavage by ERManI, the presumed rate-limiting timing step for initiation of ERAD for glycosylated proteins [17]). Inhibition of one of ERManI has been shown to delay the disposal of incompletely folded intermediates in glycoprotein misfolding disorders and can potentially aid in the secretion of functional mutant glycoproteins. However, all known potent ERManI inhibitors also strongly inhibit other Golgi GH47 α-mannosidases, including Golgi α-mannosidases IA/IB/IC, leading to a blockage in the normal maturation of glycans to complex type structures on cell surface and secreted glycoproteins.

LSDs are a group of genetic diseases that are caused by genetic deficiencies in lysosomal enzymes and consequently result in the abnormal systemic accumulation of undegraded metabolites such as glycosphingolipids, glycogen, mucopolysaccharides and glycoproteins in

patient tissues [18]. The most notable pathologies are in neural tissues where cognitive and developmental deficits can often lead to perinatal lethality [8, 19-22]. Additional pathological consequences of the enzymatic defects include connective tissue and muscular disorders, seizures, dementia, deafness blindness, hepatomegaly, splenomegaly, pulmonary and cardiac problems, abnormal bone growth and death in early infancy depending on the severity of the enzymatic defect and the nature of the defective enzyme [3, 19-21, 23, 24]. The diseases generally result from recessive loss-of-function genetic defects in the production of enzymes involved in the catabolism of glycoconjugate structures that are normally degraded in the cellular lysosomal compartment. Genetic lesions leading to loss of function can range from large chromosomal abnormalities leading to ablated gene expression to point mutations leading to frameshifts, nonsense, or missense mutations. In the latter case the respective protein is often expressed and the degree of enzymatic defect is dependent on the severity of the mutation. In many cases the missense mutations lead to hypomorphic rather than true null enzymatic phenotypes as a result of the reduced stability or rapid disposal of the nascent enzymes.

Lysosomal enzymes are initially synthesized on membrane-bound ribosomes associated with the ER and are co-translationally translocated into the ER lumen where protein folding is initiated. Wild type enzymes generally fold with the assistance of classical ATP-driven chaperones, protein disulfide isomerases, and the glycan-mediated chaperones, calnexin and calreticulin [14] An ER quality-control system ensures that only properly folded proteins are transported to the Golgi apparatus for further maturation, whereas the improperly folded mutant enzymes will not be transported to the Golgi and are retained in ER for targeted for ER disposal. The efficient recognition of misfolded or incompletely folded polypeptides for disposal can compete with the kinetics of protein folding, particularly for slowly folding enzymes in the

lumen of the ER. As a result, some missense mutations, as well as even wild type enzymes, may be prematurely degraded despite having the capacity to generate a functionally active enzyme forms if provided an additional kinetic lifetime to complete the folding process.

The failure for the disease-related mutant enzymes to complete their folding itinerary and be transported to lysosomes is a contributing factor for the incomplete degradation and the aberrant accumulation of the corresponding metabolites in the tissues of LSD patients. Theoretically, if the disposal arm of the ER quality control machinery was delayed via ERAD inhibition, it could lead to extension of the folding time frame and potentially lead to rescue of function for some of the hypomorphic missense mutant enzymes that have compromised folding kinetics. This approach as been demonstrated in two clinically distinct LSDs, Gaucher and Tay-Sachs disease, where ERAD prevents native folding of mutant lysosomal enzymes [6, 25, 26]. In this case the GH47 α-mannosidase inhibitor, Kif, given at very low concentration (50 nM), can successfully prolong the ER retention of those mutant enzymes via inhibition of ERManI. This inhibition results in enhanced native folding, trafficking, and activity of these mutant enzymes in lysosomes without causing misfolded proteins to accumulate or inducing apoptosis in the patient-derived cell lines.

Sarcoglycans (SG) are a group of plasma membrane glycoproteins that associate with dystrophin to form the dystrophin glycoprotein complex. Some sarcoglycan mutations can cause sarcoglycannopathies, which are autosomal recessive muscular disorders of the Limb–girdle Muscular Dystrophy (LMGD) group. The sarcoglycan mutations associated with sarcoglycanopathies can produce misfolded glycoproteins that still display capabilities to be assembled into the dystrophin-glycoprotein complex. However these misfolded proteins, are

efficiently degraded in ER via ERAD before being translocated to their native destinations. Partial inhibition of the ERAD machinery can be achieved by inhibition of ERManI, thus preventing mutant sarcoglycan proteins from degradation [7, 8]. The undegraded mutant sarcoglycans can then be rescued for functional transport to form the distrophin-sarcoglycan complex. In studies using cell-based assays and animal models, Kif was found to be a potential lead compound for the treatment of some forms of sarcoglycannopathies, especially LGMD 2D, which results from the R77C substitution in the α-sarcoglycan gene [7-9]. Treatment of sarcoglycanopathies using Kif has been patented by Genethon [9]. The potential of use of GH47 α1,2-mannosidase inhibitors as therapeutic strategies for treatment of a broader collection of protein folding disorders also appears promising. However, one of the limiting factors in their general therapeutic application is their lack of discrimination in inhibition among the GH47 α1,2-mannosidases, including ERManI and the Golgi α1,2-mannosidases. Therefore, development of second-generation inhibitors that display selective potent inhibition of ERManI and weak inhibition of the Golgi α1,2-mannosidases will be critical for their broad application as selective ERAD inhibitors in the treatment of LSDs, LMGDs, and other glycoprotein misfolding disorders.

Previous studies efforts to increase selectivity of enzyme inhibition while maintaining inhibitory potency was attempted through the alkylation of Kif at the N-1 position. The resulting compounds lost all inhibitory capacity for the GH47 enzymes [27] . A crystal structure of human ERManI in complex with Kif reveals why the N-1 alkylated derivatives were not favorable. The enzyme-inhibitor complex structure demonstrated that the N-1 amine group is positioned along the wall of the narrow neck of the binding cleft, and as a result leaves very little room for extension without a steric clash in the active site. Instead structural modification and extension

on the amide carbonyl group (C-2) may provide an effective position for structural modification since this site extends outward toward the solvent-exposed opening of the funnel-shaped glycan binding cleft (Figure 41) [28]. The orientation of the N-6 atom of dMNJ also suggests that structural modification at this position will extend towards the ≥+1 subsites of the glycan binding cleft. In this study, libraries of Kif derivatives containing modifications at the C-2 position of Kif and the N-6 positon of dMNJ were used in inhibition assays for both ERManI and GMIA in order to probe enzyme selective preferences for ERManI *versus* GMIA. The results provide insights into the design of more potent and selective inhibitors towards ERManI as potential therapeutics for protein folding disorders.

4.3. Methods and Experiments

The catalytic domains of human ERManI and GMIA in pPICZα vectors were expressed in *Pichia pastoris*, purified through cation exchange, hydrophobic interaction, and size exclusion chromatography methods [29, 30] and concentrated to the final concentration of 5 mg/ml in the buffer with 20 mM MES, 100 mM NaCl, 5 mM $CaCl_2$, pH 6.5. The synthesized inhibitor compounds (obtained from J. Cardot and G.J. Boons, unpublished) were prepared as stock solutions in the same buffer at 100 μM - 1 mM concentration depending on the solubility of the individual compounds. Assays were performed in 96-well microplates. The assay format was carried out by mixing enzymes and inhibitors in each well, with the final concentrations of inhibitor at 50 nM, 100 nM, 250 nM, 500 nM, 1 μM, 2 μM, and 5 μM, and final concentration of enzyme at 100 μg/ml. Following mixing of the enzyme and inhibitor, the mannobiose disaccharide substrate, Man-α1,2-Man, was added to a final concentration of 500 nM, 1 μM, 10 μM, 30 μM, 50 μM, 100 μM, 250 μM or 500 μM in a final reaction volume of 100 μl. The reaction mixtures were incubated at 37℃ for 30 min to 1 hr and the reaction was stopped by

heating to 90 ℃ for 5 minutes on a heat block. The precipitated proteins and salts were removed by centrifugation and the supernatant solutions were transferred to a new microplate.

The spectrophotometric assays for released mannose in the microplate wells were carried out as reported previously [31]. Briefly, 100 μl of mannose assay solution containing glucose oxidase (55 U/ml), horseradish peroxidase (1 purpurogallin unit/ml) and *o*-dianisidine dihydrochloride (40 mg/ml) in 1 M Tris-HCI, pH 7.5, was added to each well. The microplates were kept in an incubator at 37 ℃ for 30 min to 1 hr until there was no further increase in absorbance. Absorbances were read at 450 nm in a BioRad microplate reader. Absorbance values were converted to nmol of mannose released using a mannose standard curve, obtained as described below. K_i values were obtained by fitting the rate data to the Michaelis-Menten equation, using nonlinear regression analysis employing the Marquardt-Levenberg algorithm (SigmaPlot, Jandel Scientific).

A standard curve for mannose quantification was generated before each enzyme assay by adding D-mannose to microplate wells to the final concentration range of 500 nM - 500 μM in 50 mM MES buffer, pH 6.5, and 100 μl of the mannose assay solution was added into each well. Assay blanks containing only buffer and mannose assay solution were subtracted from the absorbance readings and the absorbance values were plotted *versus* mannose concentrations with Microsoft Excel. The linear fit equation was generated by Excel.

4.4. Results and Discussion

Two libraries of compounds comprised of Kif derivatives and dMNJ derivatives were generated in the Boons lab (J. Cardot and G.-J. Boons, personal communication) and employed in the inhibitor screening of human ERManI and GMIA. The Kif derivative library was generated using Kif-hydrazone (**15**) as the lead compound. The remaining compounds in this

library were synthesized by covalent extension of the primary amine group of the hydrazone. The dMNJ derivatives were synthesized by covalent extension of the 6-N atom of the dMNJ ring.

4.4.1 Inhibitor profiling of Kif derivatives

The initial chemical modification of Kif into the hydrazone derivative **15** resulted in an ~8-fold and ~12-fold decrease in inhibition towards GMIA and ERManI respectively compared to underivatized Kif, without displaying an apparent change in selectivity between the two enzymes. However, the amine group in the Kif hydrazone derivative **15** provides a convenient linker for further chemical modifications. In addition, the crystal structure of ERManI in complex derivative **15** revealed that the hydrazone moiety does extend into the spacious glycan binding cleft of the enzyme as expected, and the configuration of the remainder of the inhibitor complex was identical in structure to the ERManI-Kif complex (Y. Xiang and K. W. Moremen, unpublished). Therefore, the interactions of the hydrazone derivative with enzyme in the -1 subsite were identical as those of Kif (Figure 42).

A collection of aromatic and alkyl groups were covalently coupled to the hydrazone amine group and produce a series of derivatives, **16-40 (Table 5)**. The derivatives were expected to keep the same interactions with the enzyme -1 subsite residues as Kif, but also potentially gain additional interactions within the enzyme in the ≥+1 subsites. The aromatic and long alkyl substituents coupled to Kif through the hydrazone linker may influence binding affinity either through hydrophobic interactions with the enzyme or possibly decreased binding affinity through steric clashes within the active site. Compared to Kif, derivative **16** with a benzyl group as the extension exhibited a 32-fold and 38-fold increase in K_i towards ERManI and GMIA,

respectively, indicating the benzyl group attached to the hydrazone NH_2 group compromises binding in the active site.

To further probe the influence of substituted benzyl moieties on enzyme inhibition, additional derivatives with electron withdrawing or donating substituents were introduced into the benzyl ring at different positions and were tested for inhibition of both ERManI and GMIA (Table 5). Aromatic substituents with the electron withdrawing -Br and -F substituents on the benzyl group, **17** and **18,** respectively, showed distinct inhibition profiles relative to **16**. The -Br derivative showed mildly decreased inhibition of ERManI (K_i = 2.31 μM versus 1.94 μM for the hydrazone), but significantly increased inhibition of GMIA (K_i = 1.3 μM versus 3.1 μM for the hydrazone). In contrast, the -F derivative had slightly improved inhibition for both enzymes (K_i = 0.93 μM for ERManI and 2.32 μM for GMIA). Neither compound was as effective an inhibitor as Kif. A derivative with a benzyl moiety containing a strong electron donating hydroxyl substituent and a mild electron donating methoxyl substituent, **19,** exhibited a minor increase in inhibition compared to the benzyl derivative **16** for ERManI, but had an improved inhibition of GMIA to result in a minor increase in selectivity toward GMIA. The benzenediol derivative **20,** which has a strong electron donating effect from the two hydroxyl groups**,** exhibited a 4-fold increase in inhibition towards ERManI compared to the hydrazone derivative **15**, along with a 3.5-fold increase in selectivity towards ERManI. These results indicate that the extension of an electron-rich aromatic moiety, which is linked via the hydrazone NH_2 group at the C-2 position of Kif, may favor interactions between the inhibitor and the ERManI, and consequently result in increased inhibition and selectivity.

The Kif derivatives with various heterocyclic aromatic moieties at the C-2 position have also been investigated for their inhibition and selectivity between ERManI and GMIA (Table 5). The

derivative **21** with a pyridine moiety, compared to the hydrazone derivative **15**, showed a significant increase in inhibition of ERManI, with a K_i of 0.07 μM, which is comparable to underivatized Kif. However, the derivative **21** did not exhibit increased inhibition towards GMIA**.** The K_i for this compound toward GMIA was 0.66 μM, similar to the hydrazone derivative **15**. The significant difference in K_i values of compound **21** towards ERManI and GMIA demonstrates a 9.5-fold stronger inhibition towards ERManI *versus* GMIA. The Kif derivative **21** stands out as a selective inhibitor for ERManI that maintained potency (Table 5).

Comparison of the structures of derivatives **16** and **21** in their inhibition towards both ERManI and GMIA provides clues on the influence of substituents attached to the aromatic moiety on inhibition activity. Introduction of the ring nitrogen in the derivative **21** led to a 28-fold increase in inhibition towards the ERManI and only a ~4.5-fold increase in inhibition towards the GMIA. Even though the moieties attached to the C-2 position of the two Kif derivatives have comparable configurations, the electron density distributions on the attached aromatic rings are different. In contrast to the electron distribution of the aromatic moiety in **16**, the electron density of pyridine moiety in **21** is not evenly distributed because of the electronegativity difference of nitrogen and carbon atoms, and a dipole formation with electron density extending more towards the nitrogen. In addition, the presence of the nitrogen makes the pyridine more basic than benzene. One or both of these properties may contribute to its increased capacity for inhibition. A series of Kif derivatives containing heterocyclic moieties including imidazole (**22**), furan (**23**), and thiophene (**24**) moieties were also investigated. Of these derivatives, compound **22** with an imidazole moiety displayed compromised potency comparable to Kif, with a 10-fold decreased inhibition towards ERManI and 30-fold decreased inhibition towards GMIA. However, compound **22** showed a 3-fold selectivity towards the ERManI

compared to Kif. These results indicate that the derivatives at the Kif C-2 position can play a key role in providing selectivity for enzyme isoform inhibition while maintaining the high binding affinity of the Kif lead compound. The Kif derivatives, with long aliphatic chains as extensions, **25**, **26**, and **27**, exhibited low potency and no selectivity in the inhibition towards ER and Golgi enzymes (Table 5).

Since the Kif pyridine derivative **21** maintained potency of inhibition and showed selectivity towards ERManI, a small library of Kif derivatives that have substituted pyridines or moieties similar to pyridine attached to Kif via the C-2 hydrazone linker were investigated (Table 6). Eight derivatives, **31**, **32**, **33**, **35**, **36**, **38**, **39**, and **40**, demonstrated 4-fold to 6.5-fold selectivity towards the ERManI (Table 6). However, the potencies of inhibition of these derivatives, except for compound **31**, were generally less than Kif or the pyridine derivative **21**. The K_i values of those derivatives towards GMIA dramatically increased relative to Kif, suggesting a significant loss of binding affinity towards GMIA. For the latter compound series the improved selectivity towards ERManI were caused more by the unfavorable interactions of the bulkier substituents with the GMIA rather than the favored interactions with the ER enzyme. An exception within this group was the 4-chloropyridine derivative **31**, which demonstrated a 4.5-fold selectivity and low nanomolar inhibition toward ERManI comparable to that of compound **16**. The inhibition tests with this set of derivatives indicates that the Golgi enzyme cannot effectively accommodate the bulky moieties that likely to interact with the $\geq$+1 subsite residues, but these substituents are more effectively tolerated by the ER enzyme.

4.3.2. dMNJ derivatives

A small library of dMNJ derivatives, which contain various moieties attached to the nitrogen of the dMNJ via a hydrazone linkage, were investigated for their inhibition and selectivity

towards ERManI and GMIA (Table 7). The lead compound, dMNJ, is a less potent α-mannosidase inhibitor relative to Kif with little selectivity between the ER and Golgi enzymes.

All five hydrazone DMJ derivatives demonstrated increased selectivity towards ERManI over GMIA. Derivatives **47** and **49** had the highest degree of selectivity with 21- and 32.5-fold selectivity toward ERManI, respectively. The 3-methyl-thiophene derivative **49** not only had a very high degree of selectivity, but also had a remarkably increased potency. The *K*i for **49** is in the high nanomolar range, with a ~70-fold increased potency for ERManI relative to dMNJ and a 43-fold increase in selectivity toward ERManI over GMIA compared to dMNJ (Table 7).

In summary, a novel collection of hydrazine functionalized Kif and dMNJ analogues were synthesized and utilized to screen for potency and selectivity for inhibition of GH47 α1,2-mannosidases. Most derivatives had reduced affinities and modestly improved selectivity toward ERManI. However, a pyridine derivative at the C-2 position of Kif **21** retained high binding affinity (K_i = 70 nM) and improved selectivity (9.5-fold better inhibition of ERManI over GMIA) relative to the lead Kif compound. In addition, a 3-methyl-thiophene functionalized DMJ analogue **49** had significantly improved inhibition of ERManI (70-fold decrease in K_i to 300 nM) and a greatly improved selectivity (32.5-fold better inhibition of ERManI over GMIA). Thus, the screening of Kif and dMNJ derivatives have identified the first selective and potent GH 47 α1,2-mannosidase inhibitors derived from dMNJ or Kif.

Figure 42. Chemical structures of Kif and Kif-hydrazone and the X-ray crystal structures of the two compounds in the enzyme cleft of human ERManI. **Panel A**: chemical structure of Kif, the ring atoms are numbered according to IUPAC nomenclature rules for hetero bicyclic compounds; **Panel B**: chemical structure of Kif-hydrazone with the ring atoms numbered in the same way as Kif; **Panel C**: The stick presentation of Kif bound in the active site of human ERManI (surface presentation), with N-1 along the wall of the cleft and C-2 carbonyl group pointing towards the spacious glycan binding cleft; **Panel D**: The stick presentation of Kif-hydrazone bound in the active site of human ERManI (surface presentation) with the hydrazone amine group extending into the spacious part of the substrate binding cleft. In both **Panels C** and **D**, the calcium ion is represented as a blue sphere.

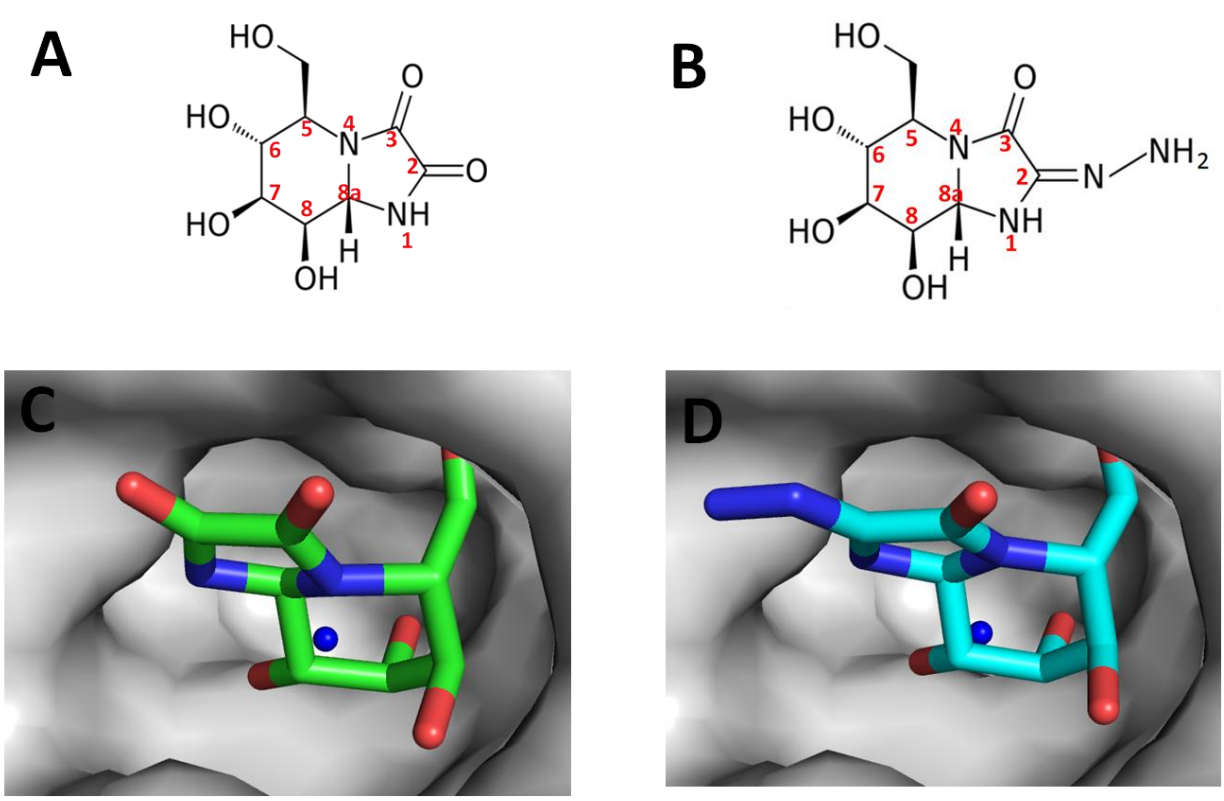
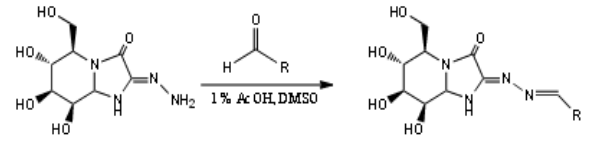
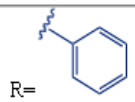

A
HO
HO
HO
OH
O
O
N
NH
H
B
HO
HO
HO
OH
O
N
NH2
N
NH
H
C
D

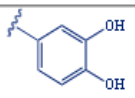
Table 5. Inhibition data (K_i) of Kif derivative compounds 15-27 toward ERManI and GMIA.

Inhibitor	K_i (ERManI)	K_i (GMIA)	K_i (GMIA) /K_i (ERManI)
Kifunensine	0.06	0.08	1.3
Hydrazone (15)	0.83	0.68	0.8
16	1.94	3.15	1.5
17	2.31	1.30	0.6
18	0.93	2.32	2.5
19	1.70	0.86	0.5
20	0.27	0.98	3.6
21	0.07	0.66	9.5
22	0.69	2.10	3.0
23	1.60	2.06	1.3
24	1.05	2.43	2.4
25	1.62	1.91	1.2
26	1.01	1.94	2.0

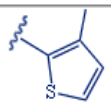




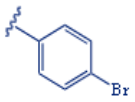
16



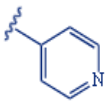
20



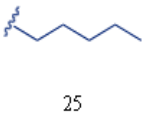
24



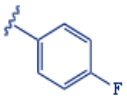
17



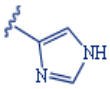
21



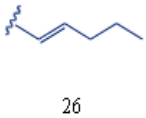
25



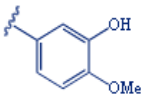
18



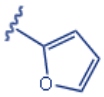
22



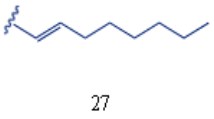
26



19



23



27

Table 6. Inhibition data (K_i) of Kif derivative compounds 28-40 toward ERManI and GMIA

Inhibitor	K_i (ERManI)	K_i (GMIA)	K_i (GMIA) /K_i (ERManI)
Kifunensine	0.06	0.08	1.3
Pyridine-Kif (21)	0.07	0.66	9.5
28	1.29	2.67	2.0
29	0.37	0.91	2.5
30	1.48	3.55	2.4
31	0.18	0.80	4.4
32	1.08	5.00	4.6
33	2.25	11.23	5.0
34	0.86	3.06	3.5
35	1.27	4.73	3.9
36	0.62	3.22	5.2
37	0.81	2.79	3.4
38	1.41	5.20	3.8
39	1.29	8.55	6.6
40	0.39	1.68	4.5

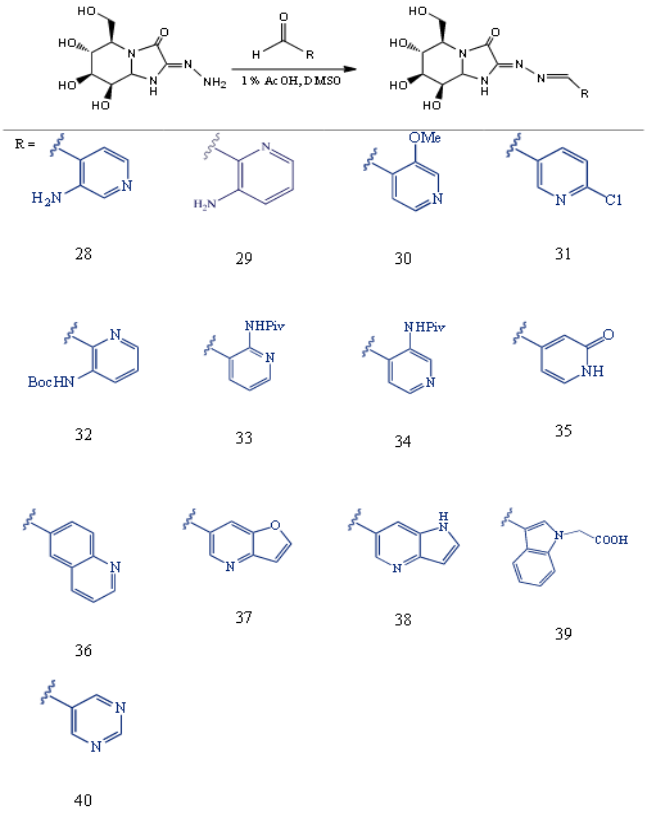
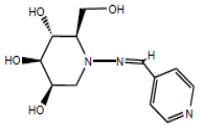
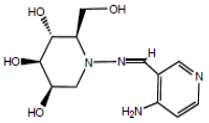


Table 7. Inhibition data (K_i) of dMNJ derivative compounds 46-50 toward ERManI and GMIA

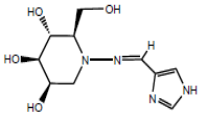
Inhibitor	K_i (ERManI)	K_i (GMIA)	K_i (GMIA) /K_i (ERManI)
DMJ	21.1	15.4	0.73
46	39.3	286.2	7.4
47	24.1	496.9	21
48	9.0	65.5	7.3
49	0.3	9.7	32.3
50	2.2	8.6	3.9



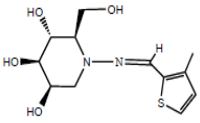
46



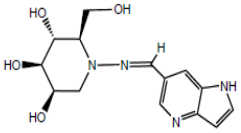
47



48



49



50

References

1. Stutz, A.E., Ed., *Iminosugars as Glycosidase Inhibitors;*. 1999.
2. Winchester, B. and G.W. Fleet, *Amino-sugar glycosidase inhibitors: versatile tools for glycobiologists.* Glycobiology, 1992. **2**(3): p. 199-210.
3. Zitzmann, N., et al., *Glycosylation: disease targets and therapy.* Adv Exp Med Biol, 2005. **564**: p. 1-2.
4. Elbein, A.D., et al., *Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I.* J Biol Chem, 1990. **265**(26): p. 15599-605.
5. Fan, S.L.W., *Methods for treating lysosomal storage diseases using L-type Ca^{2+} Channel blockers with a 1,4 dihydropyridien structure and inhibitors of ER-associated degradation.* 2012, Univ. Rice; William M; Segatori Laura; Wang Fan: US.
6. Wang, F., et al., *Inhibition of endoplasmic reticulum-associated degradation rescues native folding in loss of function protein misfolding diseases.* J Biol Chem, 2011. **286**(50): p. 43454-64.
7. Soheili, T., et al., *Rescue of sarcoglycan mutations by inhibition of endoplasmic reticulum quality control is associated with minimal structural modifications.* Hum Mutat, 2012. **33**(2): p. 429-39.
8. Bartoli, M., et al., *Mannosidase I inhibition rescues the human alpha-sarcoglycan R77C recurrent mutation.* Hum Mol Genet, 2008. **17**(9): p. 1214-21.
9. Richard, I., *Drugs for the treatment of sarcoglycanopathies* Jul 31, 2012, Centre National De La Recherche Scientifique, Genethon: US.

10. Moremen, K., *α-mannosidases in Asparagine-linked oligosaccharide processing and catabolism, Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook.* 2000. **II**: p. 81–117.

11. Henrissat B. Henrissat, A.B., *Updating the sequence-based classification of glycosyl hydrolases.* Biochem. J, 1996(316): p. 695-696.

12. Moremen, K.W., R.B. Trimble, and A. Herscovics, *Glycosidases of the asparagine-linked oligosaccharide processing pathway.* Glycobiology, 1994. **4**(2): p. 113-25.

13. Ellgaard, L. and A. Helenius, *ER quality control: towards an understanding at the molecular level.* Curr Opin Cell Biol, 2001. **13**(4): p. 431-7.

14. Helenius, A., *Quality control in the secretory assembly line.* Philos Trans R Soc Lond B Biol Sci, 2001. **356**(1406): p. 147-50.

15. Jarosch, E., U. Lenk, and T. Sommer, *Endoplasmic reticulum-associated protein degradation.* Int Rev Cytol, 2003. **223**: p. 39-81.

16. Trombetta, E.S. and A.J. Parodi, *Quality control and protein folding in the secretory pathway.* Annu Rev Cell Dev Biol, 2003. **19**: p. 649-76.

17. Wu, Y., et al., *Elucidation of the molecular logic by which misfolded alpha 1-antitrypsin is preferentially selected for degradation.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8229-34.

18. Winchester, B.S., C., Markman, O., *The Development of Iminosugars as Drugs. In Glycobiology*. 2006. pp 308–324.

19. Winchester, B., A. Vellodi, and E. Young, *The molecular basis of lysosomal storage diseases and their treatment.* Biochem Soc Trans, 2000. **28**(2): p. 150-4.

20. Johnson, B.A., A. Dajnoki, and O. Bodamer, *Diagnosis of lysosomal storage disorders: Gaucher disease.* Curr Protoc Hum Genet, 2014. **82**: p. 17 15 1-6.

21. Nagueh, S.F., *Anderson-Fabry disease and other lysosomal storage disorders.* Circulation, 2014. **130**(13): p. 1081-90.

22. Thomas, A.S., A. Mehta, and D.A. Hughes, *Gaucher disease: haematological presentations and complications.* Br J Haematol, 2014. **165**(4): p. 427-40.

23. Aronovich, E.L. and P.B. Hackett, *Lysosomal storage disease: gene therapy on both sides of the blood-brain barrier.* Mol Genet Metab, 2015. **114**(2): p. 83-93.

24. Gimovsky, A.C., P. Luzi, and V. Berghella, *Lysosomal storage disease as an etiology of nonimmune hydrops.* Am J Obstet Gynecol, 2015. **212**(3): p. 281-90.

25. Segatori, L.W., Fan, *"Methods for treating lysosomal storage diseases using L-type Ca2+ channel blockers with a 1,4 dihydropyridine structure and inhibitors of ER-associated degradation"*. 2012, William Marsh; Rice University.

26. Elfrink, H.L., et al., *Inhibition of endoplasmic reticulum associated degradation reduces endoplasmic reticulum stress and alters lysosomal morphology and distribution.* Mol Cells, 2013. **35**(4): p. 291-7.

27. Hering, K.W., et al., *A practical synthesis of kifunensine analogues as inhibitors of endoplasmic reticulum alpha-mannosidase I.* J Org Chem, 2005. **70**(24): p. 9892-904.

28. Vallee, F., et al., *Structural basis for catalysis and inhibition of N-glycan processing class I alpha 1,2-mannosidases.* J Biol Chem, 2000. **275**(52): p. 41287-98.

29. Tempel, W., et al., *Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases.* J Biol Chem, 2004. **279**(28): p. 29774-86.

30. Vallee, F., et al., *Crystal structure of a class I alpha1,2-mannosidase involved in N-glycan processing and endoplasmic reticulum quality control.* EMBO J, 2000. **19**(4): p. 581-8.

31. Scaman, C.H., F. Lipari, and A. Herscovics, *A spectrophotometric assay for alpha-mannosidase activity.* Glycobiology, 1996. **6**(3): p. 265-70.

CHAPTER 5

Summary and prospect of studies on GH47 α1,2-mannosidases

The structural basis for the substrate specificity of human endoplasmic reticulum (ER) α-mannosidase I (ERManI) and Golgi α-mannosidase IA (GMIA) have been explored in this dissertation. The enzyme-substrate complexes of ERManI and GMIA reveal the amino acid residues and glycan residues that participate in the enzyme-substrate interactions as well as providing insights into the reaction mechanisms for both enzymes. The roles of active site residues in ERManI have been studied by measuring the impact on substrate specificity that results from the swapping of residues from structurally equivalent positions from GMIA. The results indicated that elimination of the specific enzyme-substrate interactions by the swapping individual residues between ERManI and GMIA generally had minor effects on substrate specificity. However, comprehensive swapping of glycan binding cleft residues or single residues interacting with the glycan core can broaden the specificity of ERManI to resemble the activity of GMIA. Our systematic approach that employs X-ray structural studies, site-directed mutagenesis, kinetic analysis, and detailed time course studies of glycan digestion have now resulted in a more comprehensive view of selective substrate recognition and catalysis by this family of enzymes.

5.1. X-ray crystallography studies of Golgi Mannosidase IA and $Man_9GlcNAc_2$-PA complexes.

The enzyme-$Man_9GlcNAc_2$-PA substrate complexes of human ERManI and mouse GMIA reveal how the intact substrates bind in the respective enzyme active site clefts. In combination with prior studies on enzyme-product complexes for GMIA and yeast ERManI [1-3], the structures reveal many aspects of how enzyme-substrate recognition and cleavage

proceeds from initial $Man_9GlcNAc_2$ substrate binding to cleavage down to $Man_5GlcNAc_2$ final cleavage products. The individual structures represent snapshots of the glycan orientations and interactions for both substrate and final cleavage products[1-3]. For example, cleavage of $Man_9GlcNAc_2$ glycans by GMIA is a sequential reaction process, starting with initial substrate cleavage of residue M11 from branch A to form $Man_8GlcNAc_2$. Subsequent cleavage of the M9 residue from branch C forms a single $Man_7GlcNAc_2$ isomer and final cleavage of the M8 residue from branch A generates a $Man_6GlcNAc_2$ isomer containing a single α1,2-mannose linkage on branch B. This remaining α1,2-mannose residue is cleaved with very poor efficiency *in vitro* to generate a $Man_5GlcNAc_2$ limit digestion product, but this reaction is unlikely to occur *in vivo* through GMIA action since ERManI would be anticipated to act on this linkage. Thus, the product generated in each glycan cleavage step is subsequently re-oriented in the active site and becomes the substrate employed in the following step in the cleavage series. Snapshots of the bound substrate complexes with $Man_9GlcNAc_2$-PA and final product of $Man_{5+1}GlcNAc_2$-PA (containing an additional α1,6-mannose residue added during secretion in the *P. pastoris* recombinant host used for enzyme production [1]) reveals enzyme-oligosacchride interactions at two points of the cleavage series, while the binding patterns for other sequential products in the reaction series remain unknown. Thus, the structural studies clearly indicate that the patterns of binding interactions for the sequential products within the glycan binding site change at each stage of the cleavage series, illustrating how glycan cleavage is a dynamic and selective process for the ordered and sequential mannose trimming process catalyzed by GMIA [4].

Examination of the enzyme-substrate and enzyme-product complexes of ERManI and GMIA resulted in some interesting comparisons. The enzyme-substrate and enzyme-product complexes for ERManI ($Man_9GlcNAc_2$ *versus* $Man_5GlcNAc_2$ bound in the active sites for human

and yeast ERManI, respectively) both contain branch B extending to the +1 and -1 subsites of the enzyme and the arrangement of $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$-PA in the active site cleft were essentially superimposable with the exception of the additional α1,2-mannose residues in the $Man_9GlcNAc_2$-PA complex. These similarities in substrate orientations occur despite the significant differences in species sources and protein sequences for the two enzymes. In contrast, the arrangements of the oligosaccharides in the enzyme-substrate and enzyme-product complexes for GMIA were quite different. As expected for the initial cleavage of the M11 residue from branch A identified in substrate specificity studies [4], the $Man_9GlcNAc_2$-PA substrate was bound by GMIA using the branch A in the -1 subsite. However, the $Man_{5+1}GlcNAc_2$ enzyme-product complex for GMIA inserted branch C into the +1 subsite. The limit cleavage product for GMIA is $Man_5GlcNAc_2$ containing no α1,2-mannose residues and the last residue to be cleaved should be the low affinity α1,2-mannose residue on branch B. Thus, the GMIA enzyme-product complex containing branch C in the +1 subsite does not reflect the final cleavage product of the digestion series. However, it is not entirely surprising that branch C is engaged in the active site for the GMIA enzyme-product complex. The ordered and sequential cleavage of the four α1,2-mannose residues by GMIA reflects the relative hierarchical affinities for the respective glycan termini in the enzyme active site. Initial binding of branch A in the GMIA-$Man_9GlcNAc_2$-PA complex presumably results from a highest affinity for this branch in the glycan binding cleft. Following cleavage of the M11 residue, the next highest affinity terminus for binding to the active site must be branch C, since residue M9 is exclusively cleaved next based on substrate specificity studies [4]. Likewise, the next-highest binding affinity would be the underlying α1,2-mannose on branch A (M8 residue). The α1,2-mannose residue on branch B (residue M10) would have considerably lower affinity, since the rate of cleavage of this

residue is >10-fold lower than the other three α1,2-mannose residues. During expression, purification, and concentration of glycosylated GMIA the enzyme presumably cleaved all four α1,2-mannose residues and during crystallization it formed a complex with one of the glycan termini. The $Man_5GlcNAc_2$-PA terminal digestion product does not contain either residue M11 or M8 on branch A, thus the enzyme does not have the opportunity to engage branch A as it would for the product of cleavage of residue M11. Residue M9 from branch C is the second cleavage in the sequential reaction hierarchy and the underlying residue (M6) is still present in the $Man_5GlcNAc_2$-PA terminal digestion product. As a result residue M6 of branch C was inserted into the +1 subsite as an enzyme-product complex. The other termini on branches A and B were lower on the hierarchal cleavage series as enzymatic products and are thus not represented in the crystal structure as the enzyme-product complex..

5.2. Profling of enzyme substrate specificity, enzyme kinetics and enzyme-substrate binding studies

Structural studies on enzyme-substrate complexes for ERManI and GMIA revealed conserved and non-conserved residues involved in glycan interactions between the two enzymes. These data along with the high degree of similarity in protein fold between the two enzymes led to the possibility that the differences in substrate specificity could be probed by swapping amino acid residues between the two enzyme active site clefts. The ERManI structure was chosen as the template for the mutagenesis studies and acted as a recipient for equivalent residues from the respective GMIA structure. This was the preferred enzyme choice, since it is easier to study the expansion of substrate specificity for cleavage beyond $Man_8GlcNAc_2$ by mutant forms of ERManI in time course digestion studies than restriction in glycan cleavage specificity in mutants of GMIA. The time course of substrate cleavage for many of the single site-directed

mutants had minimal effects on glycan cleavage or resulted in some additional cleavage to smaller species reflecting hybrid activities between ERManI and GMIA for the resulting mutants. In addition, many of the mutants had reduced enzymatic activity indicating that altered specificity occurred in conjunction with reduced catalytic capacity and presumably reduced substrate binding affinity. Surprisingly, the two residues in ERManI that appeared to have the greatest impact on the extent of glycan cleavage were R461 and W389. These residues interact with the glycan core (R461 interacts with residues M3, M7 and NAG2 of $Man_9GlcNAc_2$-PA while W389 has hydrophobic stacking interactions with the NAG2) rather than peripheral glycan branches. Both the R461L and the W389A mutations significantly expanded the specificity for glycan cleavage by ERManI indicating that interactions with the glycan core play a significant role in defining the restricted substrate specificity for the enzyme. Surprisingly, the GMIA-$Man_9GlcNAc_2$-PA substrate complex showed minimal enzyme interactions with the glycan core, including an absence of hydrophobic interactions with W341 in GMIA, the equivalent of W389 in ERManI. This residue was rotated away from the glycan core in the GMIA enzyme-substrate complex, but was replaced by a hydrophobic interaction between W399 and residue M3 in the glycan core. However, in the GMIA enzyme-product complex W341 was rotated back to form clear hydrophobic stacking interactions with NAG2 analogous to the interaction between ERManI and $Man_9GlcNAc_2$-PA.

In comparison to the single residue mutants of ERManI, a comprehensive swap of glycan binding site residues from GMIA to ERManI resulted in an enzyme with a hybrid activity more closely resembling GMIA. These results indicate that it is the composite collection of residues that influences the broadened specificity of the Golgi enzyme. However, the inverse residue swap, where ERManI active site cleft residues replaced equivalent residues in GMIA, resulted in

a hybrid enzyme that could still cleave efficiently beyond $Man_8GlcNAc_2$ and did not exhibit the restricted substrate specificity of ERManI. However, the mutant enzyme could now more efficiently cleave the M10 residue from the B branch of $Man_6GlcNAc_2$-PA suggesting a partial conversion to ERManI-like specificity.

The results of the kinetics assays showed that the mutants with apparent substrate specificity change also have more compromised catalytic activity compared to those mutants with minor substrate specificity change. The substrate binding assays by SPR indicated that the binding affinity of the mutants towards $Man_9GlcNAc_2$-PA was also decreased compared to that of the wild type enzyme. The results from enzyme kinetics and substrate binding assays were consistent with the results from the substrate specicicity profiling experiments.

5.3. Characterization of the enzyme cleavage products and isomers generated by the mutant forms of the GH47 α-mannosidases. Prior studies demonstrated that wild type ERManI cleaves $Man_9GlcNAc_2$ exclusively to the $Man_8GlcNAc_2$ B isomer, while GMIA produces a mixture of 90% $Man_8GlcNAc_2$ isomer A and 10% isomer C. In the present study, the $Man_8GlcNAc_2$ isomers produced by the W389A mutant, the combinatorial ERManI mutant, and the combinatorial GMIA mutant were further characterized. Similar to wild type ERManI, the W389A mutant exclusively produced $Man_8GlcNAc_2$ isomer B, despite the ability of the mutant to cleave more promiscuously down to $Man_{6\text{-}5}GlcNAc_2$ structures similar to GMIA. Thus, the anchoring of the glycan core structure via hydrophobic interaction with W389 influenced the ability of the enzyme to cleave beyond $Man_8GlcNAc_2$, but had no impact on the position of initial cleavage. In contrast, the combinatorial ERManI mutant and combinatorial GMIA both generated additional $Man_8GlcNAc_2$ isomers compared to their wild type counter parts suggesting that that it is the broader collection of residues in the glycan binding cleft that determines the

specificity for initial branch insertion into the enzyme +1 and -1 subsites for glycoside bond cleavage. The combinatorial ERManI mutant exhibited a hybrid of ERManI (isomer B) and GMIA (isomer A) specificities producing similar levels of the two isomers and essentially no isomer C. Surprisingly, the combinatorial GMIA mutant produced $Man_8GlcNAc_2$ isomers with little resemblance to wild type GMIA. The predominant isomer produced was $Man_8GlcNAc_2$ isomer C, with slightly lower levels of isomer B and very little isomer A. This reflects a novel specificity for the mutant enzyme unlike either parent enzyme. Thus, substrate interactions in the active site clefts of the GH47 α-mannosidases are more complex than anticipated. Distinct sets of substrate interactions occur at each step in the cleavage series from $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$. However, these interactions are difficult to predict based on a limited number of snapshots from static crystal structures from two processing intermediates.

5.4. Metal ion coordination in the GH47 enzyme active site

Prior structural studies in the Moremen lab revealed that the Ca^{2+} coordination in the ERManI-thiodisaccharide complex of a T688A mutant exhibited a 7-fold coordination in comparison to the eight-fold coordination for the wild type enzyme (Karaveg and Moremen, personal communication) [2, 5, 6]. The absence of the side-chain hydroxyl group from T688 in the Ca^{2+} coordination sphere did not result in its replacement by a water molecule. Instead, the position of a coordinated water molecule was altered to result in a change in Ca^{2+} coordination. However, no alteration in the position of the bound glycone hydroxyl residues or the water nucleophile was observed. The T688A mutant was reduced 30-fold in enzyme activity towards $Man_9GlcNAc_2$-PA likely resulting from altered electrostatics or binding affinity for the substrate with the enzyme bound Ca^{2+} ion [6]. In the present studies, swapping of the enzyme-bound Ca^{2+} ion for La^{3+} resulted in a 9-fold ion coordination with the insertion of an extra water molecule

into the coordination sphere. The result was a complete inactivation of the enzyme; yet again the positions of the glycone C-2 and C-3 hydroxyls and the water nucleophile in the substrate-bound complex were unaltered. These observations suggest that the 8-fold coordination of the enzyme-bound Ca^{2+} is essential for the enzyme activity. The impact of ion coordination on enzyme activity was also implied from additional ion swap experiments previously performed in the Moremen lab (Karaveg and Moremen, personal communication), where other divalent cations predicted to have >8-fold coordination also resulted in reduced enzyme activity. Additional structural and glycan binding affinity studies with alternative ion complexes using wild type or the T688A mutant of ERManI could lead to additional insights into the role of ion coordination on the electrostatics and substrate binding affinity for the GH47 α-mannosidases.

5.5. Future studies

The enzyme-$Man_9GlcNAc_2$-PA complex structures of human ER Man I and mouse GMIA revealed how the intact substrate bind in the enzyme cleft. The yeast ER mannosidase structure and mouse GMIA enzyme structure revealed how the final enzyme cleavage product $Man_5GlcNAc_2$ ($Man_{5+1}GlcNAc_2$ in GMIA because an extra mannose was added to the glycan structure by the expression host *Pichia* pastoris) bind in the enzyme cleft. Those structures provided the snapshots of the binding patterns of both intact substrate and final cleavage products. However, the cleavage of high mannose $Man_9GlcNAc_2$ to $Man_6GlcNAc_2$ by GMIA is a sequential chain reaction, starting from the initial substrate $Man_9GlcNAc_2$-PA, via $Man_8GlcNAc_2$-PA isomer A and isomer C, $Man_7GlcNAc_2$-PA isomers, $Man_6GlcNAc_2$-PA and eventually slowly to the final product $Man_5GlcNAc_2$-PA. The product generated in one step is the substrate of the immediate following step in the cleavage stream. The snapshots of the binding pattern of $Man_9GlcNAc_2$-PA and final product $Man_{5+1}GlcNAc_2$-PA only revealed the

enzyme-glycan interactions in the very beginning and end of the cleavage stream, while the binding patterns of other sequential products in the middle of the chain reaction have remained unknown. The revelation of the static binding patterns of the sequential products ($Man_{8-5}GlcNAc_2$-PA), and the more important the enzyme-oligosaccharide interactions of each stage. By comparing interactions between the enzyme and $Man_{9-5}GlcNAc_2$-PA in each structures and by putting the static snapshots of the layouts of $Man_{9-5}GlcNAc_2$-PA together, a dynamic scenario of sequential cleavage process will be illustrated and how the enzyme (GMIA) use the interactions to send a specific mannose residue into the active site in every step of the sequential cleavage will be derived.

To co-crystallize the GMIA with substrate analogs $Man_{8-5}GlcNAc_2$-PAs and to resolve the crystal structures of those enzyme-oligosacchride complexes will elucidate the binding pattern of those oligosaccharides in the enzyme The experimental methods employed in this dissertation provide validated experimental approaches for the structural studies on GMIA-$Man_{8-5}GlcNAc_2$-PA complexes. HPLC methods with C-18 columns have been used to isolate $Man_{8-6}GlcNAc_2$-PA isomers. Workflows to generate deglycosylated GMIA by EndoF1 digestion have also been optimized and successfully used for oligosaccharide interaction studies and crystallography without the complications of competing glycan structures covalently attached to the enzyme Asn side chains. The co-crystallization methods for generating enzyme-$Man_9GlcNAc_2$-PA complexes and additional crystal screening methods can be readily applied to other oligosaccharide complexes with the enzymes. Finally, the molecular replacement methods developed here can also be applied to analysis of additional X-ray diffraction data sets to determine the structures of other enzyme-oligosaccharide complexes using GMIA-$Man_9GlcNAc_2$-PA complex as the molecular replacement template.

Additional experiments to identify the $Man_8GlcNAc_2$ isomers produced by other mutants still need to be performed to determine if the hybrid activities that were identified in the digestion time course studies are also indicated by the isomers produced at each stage of trimming. The HPLC separation methods with amine and C-18 columns were developed as a part of these dissertation studies and can be used to isolate and quantify the isomers of cleavage products of the mutants.

Another type of experimental study was performed previously to probe substrate binding affinities in the enzyme active site of ERManI [6]. The energetics of glycan interactions were examined using SPR to identify the relative contributions of individual glycan residues within the enzyme active site. In light of present dissertation studies, we now have the ability to study the energetic contributions of individual enzyme residues and glycan structures in a more comprehensive manner. Using the enzyme inactivated by La^{3+} bound in the active site will now allowe to perform detailed SPR studies on multiple potential glycan structures without concern about potential cleavage of the glycan ligand during the binding studies. Comparison of the binding energetics for wild type and mutant enzyme forms could allow the calculations of contributions of individual protein residues in the active site.

In summary, this dissertation has greatly expanded our the previous understanding of the GH47 α1,2-mannosidases, especially in revealing the structural contributions to the glycan cleavage specificity by human ERManI and mouse GMIA. The results of this dissertation shed additional light on mechanisms of glycan recognition and cleavage by these enzymes and will provide an experimental framework for further studies on this family of enzymes in the future.

References

1. Tempel, W., et al., *Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases.* J Biol Chem, 2004. **279**(28): p. 29774-86.
2. Vallee, F., et al., *Structural basis for catalysis and inhibition of N-glycan processing class I alpha 1,2-mannosidases.* J Biol Chem, 2000. **275**(52): p. 41287-98.
3. Vallee, F., et al., *Crystal structure of a class I alpha1,2-mannosidase involved in N-glycan processing and endoplasmic reticulum quality control.* EMBO J, 2000. **19**(4): p. 581-8.
4. Lal, A., et al., *Substrate specificities of recombinant murine Golgi alpha1, 2-mannosidases IA and IB and comparison with endoplasmic reticulum and Golgi processing alpha1,2-mannosidases.* Glycobiology, 1998. **8**(10): p. 981-95.
5. Karaveg, K., et al., *Mechanism of class 1 (glycosylhydrolase family 47) {alpha}-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(16): p. 16197-207.
6. Karaveg, K. and K.W. Moremen, *Energetics of substrate binding and catalysis by class 1 (glycosylhydrolase family 47) alpha-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control.* J Biol Chem, 2005. **280**(33): p. 29837-48.