MOLECULAR MODELING INSIGHTS INTO THE INHIBITION AND CATALYTIC MECHANISM OF GOLGI α-MANNOSIDASE II

by

SAMEER PRALHAD KAWATKAR

(Under the Direction of Professor Geert-Jan Boons)

ABSTRACT

In this dissertation, we have employed a computational methodology in conjunction with experimental techniques to understand the mechanism of inhibition of Golgi a-Mannosidase II (GMI) by mannostatin A, and to determine the mechanism of the hydrolysis reaction catalyzed by GMII.

*Ab initio* calculations and molecular docking studies were employed to rationalize the inhibition data of mannostatin A analogues and to identify the mode of binding of mannostatin A. It was found that Mannostatin and aminocyclopentitretrol could bind to GMII in a similar mode as that of the known inhibitor Swainsonine. However, due to the flexibility of the five membered rings of these compounds, additional low energy binding modes could be adopted. The thiomethyl moiety of mannostatin was predicted to make favorable hydrophobic interactions with Arg228 and Tyr727, possibly accounting for its greater inhibitory activity.

In order to validate the docking predictions, the X-ray crystal structures of *Drosophila* Golgi α–mannosidase II (dGMII) complexed with the inhibitors mannostatin A and its *N*-benzyl analog have been determined. The X-ray structures were in excellent agreement with the predicted binding modes for mannostatin A analogues. Molecular dynamics simulations and

NMR studies have shown that the five-membered ring of mannostatin A is rather flexible. In the bound state, mannostatin A adopts a $^{2}T_{1}$ twist envelope conformation, which is not significantly populated in solution. Possible conformations of the mannosyl oxacarbenium ion and an enzyme-linked intermediate have been compared to the conformation of mannostatin A in the co-crystal structure with dGMII. It has been found that mannostatin A best mimics the covalent linked mannosyl intermediate. The thiomethyl group is able to make a number of additional polar and non-polar interactions increasing the affinity for dGMII.

Finally, a quantum mechanical investigation of the catalytic mechanism of the hydrolysis reaction catalyzed by GMII has been conducted. These calculations have revealed that in the transition state, the mannopyranosyl ring adopts a conformation that lies between an ideal $B_{2,5}$ and $^{1}S_{5}$ conformation. Superimposition of mannostatin A and swainsonine in their bound conformations onto the optimized transition state structure revealed that both inhibitors mimic the transition state.

INDEX WORDS: Mannosidases, Mannosidases mechanism, Glycosidase Inhibitors, Oxacarbenium ion, Mannostatin, Thiomethyl, X-ray, QM and QM/MM calculations

# MOLECULAR MODELING INSIGHTS INTO THE INHIBITION AND CATALYTIC MECHANISM OF GOLGI α-MANNOSIDASE II

by

SAMEER PRALHAD KAWATKAR

B. Sc., University of Mumbai, India 1996

B. Sc. (Tech), University of Mumbai, Institute of Chemical Technology, India 1999

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

2006

MOLECULAR MODELING INSIGHTS INTO THE INHIBITION AND CATALYTIC MECHANISM OF GOLGI α-MANNOSIDASE II

by

SAMEER PRALHAD KAWATKAR

Major Professor: Geert-Jan Boons

Committee: Robert J. Woods
Kelly W. Moremen
James H. Prestegard

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2006

# DEDICATION

To my wonderful wife, Aarti. Without your support this endeavor was not possible.

To Aai (my mother) and Baba (my father), for all your hardwork to take me to this stage.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my wife, Aarti for all her love and encouragement during my tenure as a graduate student. I know, I can count on her for anything. I also want to thank our daughter, Pooja, as I always enjoyed relaxed moments while being with her.

I would like to thank Professor Geert-Jan Boons for giving me such a wonderful opportunity of working on a multidisciplinary project, and allowing me to pursue my goals. Thank you for being considerate throughout my presence here in the University of Georgia.

My special thanks go to Professor Robert J. Woods who trained me to be a computational chemist. Your encouraging support and mentorship transformed me into an independent scientist. Thank you for giving me the freedom of using the laboratory resources without any reservations.

I wish to thank the administrative staff in the CCRC, for guiding me through different issues. I also wish to thank Carolyn and Shielah for being wonderful administrative assistantants and helping me out on various occasions. I would like to thank members from Woods' group and Boons' group. In particular, I thank Austin, for all those "lunch time" discussions on possibly every subject on the earth.

Finally, many thanks go to the Department of Chemistry and Complex Carbohydrate Research Center for this opportunity.

TABLE OF CONTENTS

Page

# CHAPTER 1

# INTRODUCTION

Carbohydrates, also known as saccharides (Greek: sakcharon, sugar), are essential components of all living organisms and the most abundant class of biomolecules. It has been long known that carbohydrates serve as an energy source. However, increasing research in the field of carbohydrates has revealed that these structures are involved in diverse biological processes including cell-cell adhesion, protein folding, protein trafficking and signal transduction.[1] Despite the involvement of carbohydrates in several biological processes, there are relatively few carbohydrate-based therapeutics,[2, 3] likely due to fears that such agents will suffer from poor oral bioavailability and high clearance rates (which may be of benefit in some applications) and, most notably, the lack of a detailed understanding of the biological phenomena in which carbohydrates are involved.[4] For example, an increasing number of carbohydrate processing enzymes are being identified and validated as targets for therapeutic intervention.[4, 5] Inhibitors of these enzymes have garnered much attention and are currently used for the treatment of diabetes, viral infections and Gauchers disease.[6-10] However, efforts to design more potent and/or selective inhibitors often are limited due to a dearth of knowledge pertaining to the catalytic mechanisms and the mechanism of inhibition by natural inhibitors of these targets.

The research presented in this dissertation focuses on one such enzyme that has been validated as a target for cancer therapy namely Golgi α- mannosidase II (GMII). We have employed a computational approach in conjunction with experimental techniques to understand the mechanism of inhibition of GMII by inhibitors containing in particular cyclopentitol moieties, such as mannostatin A (chapter 2 and 3), and to determine the mechanism of the hydrolysis reaction catalyzed by GMII (chapter 4).

## 1.1 Carbohydrate Processing Enzymes

The biological processes in which carbohydrates are involved are often mediated through biological recognition of specific oligosaccharide structures that are synthesized in the form of glycoconjugates via several biosynthesis pathways. A glycoconjugate is a compound in which one or more monosaccharide or oligosaccharide units (the glycone) are covalently linked to a noncarbohydrate moiety. These glycoconjugates include glycolipids, glycoproteins and glycopeptides. Glycoconjugate biosynthesis requires conversion of monosaccharide building blocks with the help of carbohydrate processing enzymes. Such carbohydrate processing enzymes can be classified into two major classes based on the function they perform specifically, Glycosyl hydrolases (glycosidases) and Glycosyl transferases. Glycosidases are involved in cleaving a linkage between two specific monosaccharide units, whereas glycosyl transferases are involved in formation of such linkages. Huge numbers of these processing enzymes exist, and each enzyme is usually involved in the cleavage or formation of a specific linkage. The following sections present a more detailed review of the general mechanism, mechanistic studies that have been reported in recent years, and the clinical importance of glycosidases. As the subject of this research is glycosidases, the ensuing discussion is largely restricted to glycosidases.

## 1.2 Glycosyl Hydrolases

Glycosidases are a group of enzymes that catalyze the hydrolysis of specific glycosidic linkages through a nucelophilic substitution reaction at the anomeric carbon of a donor sugar moiety (known as glycone), which is characterized by displacement of the

leaving group (the aglycone, generally also a sugar moiety) by an acceptor molecule (water).

Currently, there are about 6000 known glycosidases, which have been classified into 89 families based on their primary structure. Updated information on the classification is available on the Internet at http://afmb.cnrs-mrs.fr/CAZY. In addition to the classification based on the primary structure, glycosidases can also be broadly classified according to the stereochemistry at the anomeric carbon atom in the substrate as well as the corresponding hydrolyzed product. For example, GMII is an α-retaining enzyme, which indicates that the configuration at the anomeric carbon in the natural substrate of GMII is α. This configuration is retained in the hydrolyzed product.

The glycosidic bond, particularly that between two glucose residues, is the most stable of all linkages within naturally occurring biopolymers, with half-lives for the spontaneous hydrolysis of cellulose and starch in the range of 5 million years. [11] Glycosidases, which are responsible for the hydrolyses of these materials therefore face a challenging task, yet they accomplish this with rate constants of up to 1000 $s^{-1}$ (or enhancement of $10^{17}$ fold in the rate of the reaction), earning them a reputation as some of the most proficient catalysts.[12] Hydrolysis of the glycosidic bond can occur with one of two possible stereochemical outcomes, specifically an inversion or retention of anomeric configuration (Figure 1.1). This is achieved by two different mechanisms, originally proposed by Koshland[13] based on chemical mechanisms of similar uncatalyzed reactions. While the general concept of these mechanisms remains acceptable, several details have been revealed and refined in recent years.[12, 14-16]

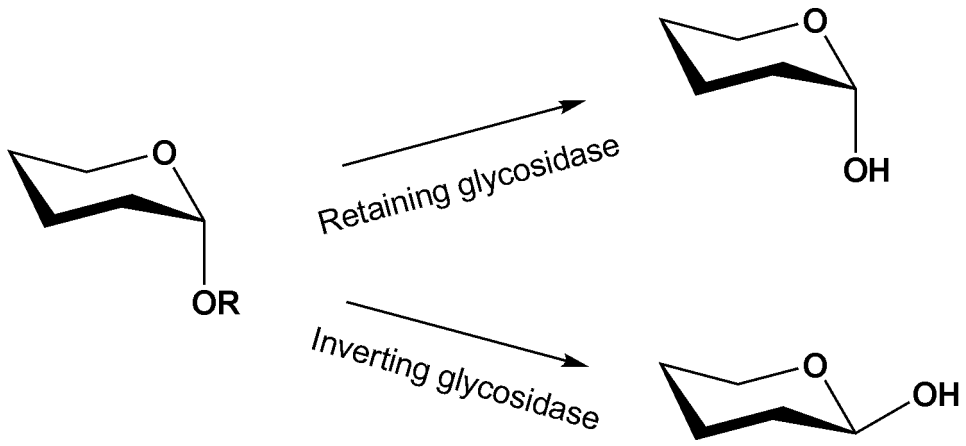



**Figure 1.1** Stereochemical outcome of a hydrolysis reaction catalyzed by an α-glycosidase.

### 1.3 General Catalytic Mechanism of Inverting Glycosidases

The catalytic mechanism of an inverting glycosidase is believed to be a single step substitution reaction at the anomeric carbon[13] which involves displacement of the aglycone with the incoming water molecule. Similar to the acid-catalyzed hydrolysis of glycosides[17], this reaction is believed to proceed via an oxacarbenium ion-like transition state (Figure 1.2). Two conserved carboxylic acid residues in the active site are necessary for the reaction to proceed. One residue acts as general acid and the other as a general base, thereby protonating the aglycone and assisting its departure, and deprotonating the incominmg water molecule, respectively. Since the subject of this research is an α-retaining enzyme, inverting enzymes and their mechanism will not be discussed further.

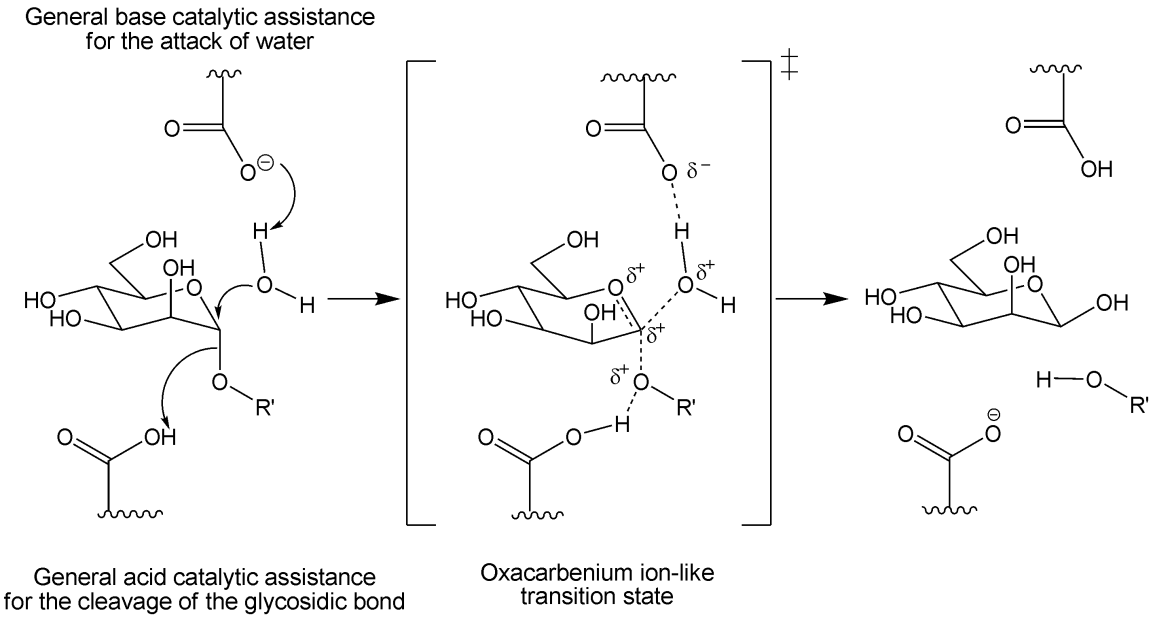



**Figure 1.2.** Proposed single step mechanism for an inverting α-glycosidase.

## 1.4 General Mechanism of Retaining Glycosidases

The original catalytic mechanism proposed by Koshland [13] for retaining glycosidases was a two-step process (also known as the double displacement mechanism) involving two displacement reactions at the anomeric center (Figure 1.3). In the first step, referred to as the glycosylation step, a catalytic nucleophile (generally an aspartate or glutamate residue) attacks the anomeric carbon of the glycone from the β-face. The aglycone departure is assisted by a general acid-base catalytic residue (usually another aspartate or glutamate residue) which protonates the aglycone. This nucelophilic substitution results in the formation of a covalently linked glycosyl-enzyme intermediate with an inversion of the anomeric configuration. The second step, referred to as the deglycosylation step, is a nucelophilic attack by a water molecule at the anomeric center

of this glycosyl-enzyme intermediate. In this step, the conjugate base of the general acid-base residue deprotonates the water molecule, facilitating the attack. This reaction also results in the inversion of the anomeric configuration. The overall reaction thus proceeds with a net retention of the anomeric configuration.

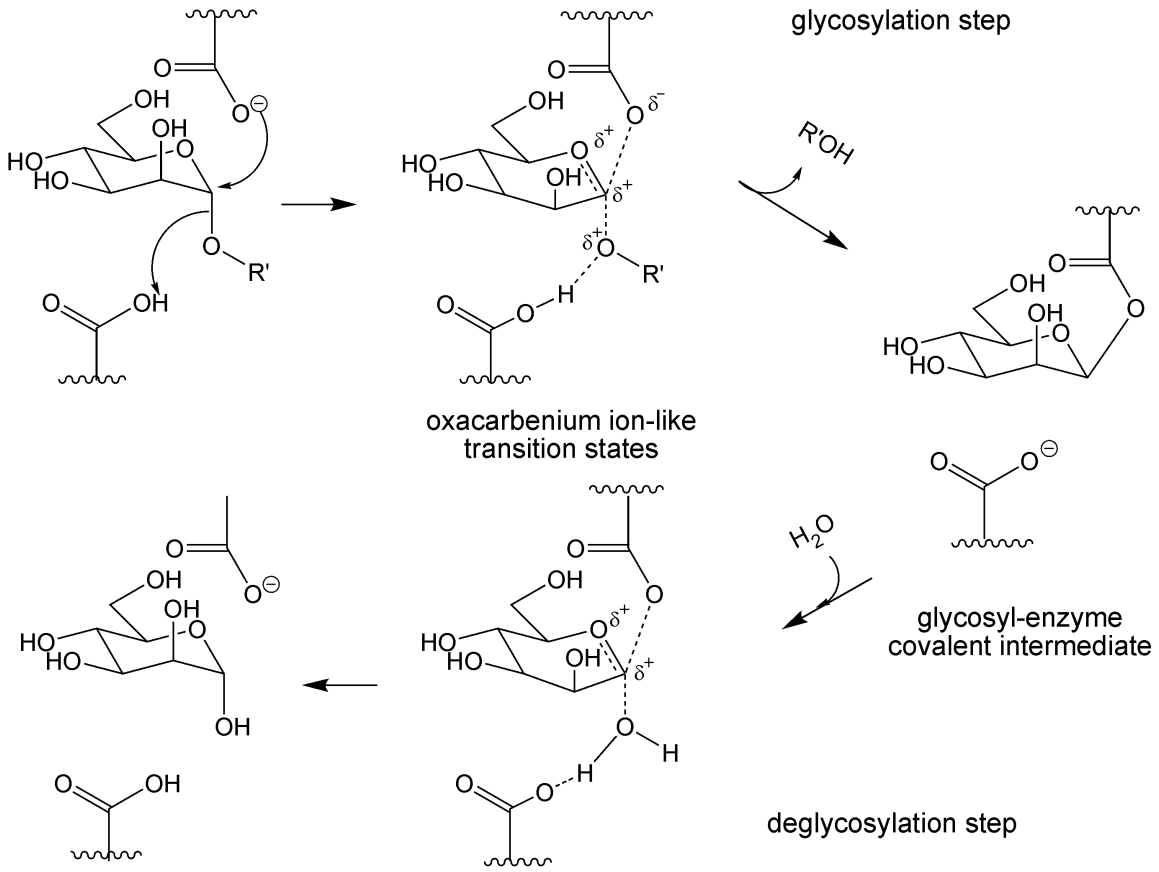



**Figure 1.3.** Proposed double displacement mechanism for a retaining α-glycosidase.

While many essential features of the general mechanism of action of retaining glycosidases are known, [12, 16, 18, 19] several methods have contributed to the refinement of mechanistic details in recent years. These methods include kinetic isotope effects, crystal structure analysis of wild type and mutated glycosidases and of their complexes with

ligands, and the structure and kinetics of the transition state mimics. Many of these mechanistic studies have been carried out on β-glycosidases, largely due to the availability of good, easy to work with model enzymes and relative ease of synthesizing the substrates required for these glycosidases as compared to their α-counterparts. Despite our increasing understanding of these mechanisms some questions remain unanswered, an include the following:

1. Which conformational change is induced in the substrate by an individual glycosidase *en route* to the transition state (distortion towards chair, half-chair, boat, skew-boat or envelope)?
2. Which conformation is adopted at the transition state and the covalent intermediates? Is this conformation the same in all glycosidases?
3. Some glycosidases have a metal ion in the active site. Does this metal ion take part in the mechanism or it is necessary only for structural stability of the ligand-protein complexes?

Several studies in recent years have attempted to address these questions. The following sections present a general review of literature on mechanistic studies of glycosidases.

### 1.5 Distortion of Substrates upon Binding with Glycosidases

Almost all biological processes involve recognition of a ligand by a specific receptor. While the nature of interaction largely depends on the shape and size of the ligand as well as the binding pocket, such an interaction usually results in conformational changes in the ligand and/or receptor that result in what is known as the Michaelis

complex. Primary drivers of such induced conformational changes are different interactions of functional groups in the ligand with protein residues. While these conformational changes arrange the substrate in a conformation which otherwise is not accessible, they also reduce the activation energy for the reaction by preorganizing the substrate in a conformation that lies closer to the transition state on the potential energy surface.

It is known that glycosyl hydrolases distort the substrate to arrive at the Michaelis complex structure. First proposed for the hen egg white lysozyme,[20] such substrate distortion has also been unambiguously observed in high-resolution X-ray structures of Michaelis complexes for two structurally unrelated endoglucanases (from glycosyl hydrolase families 5 and 7[21, 22] and hexosaminidase (family 20).[23]

Ducros et al. recently reported the X-ray structure of the Michaelis complex for a family-26 $\beta$-Mannase, which is a retaining endo-$\beta$-(1-4)-mannose. In these studies the mechanism based inactivator of mannasaes 2,4-dinitrophenyl 2-deoxy-2-fluoro-$\beta$-mannotrioside was used to probe the reaction coordinate of the reaction. Surprisingly, the Michaelis complex displayed $^{1}S_{5}$ conformation of the substrate,[24] which is different than enzymes that work on gluco-configured substrates. This result suggested a different conformational itinerary for the mannosidases as compared to glucosidases.

### 1.6 The Oxacarbenium Ion-like Transition State

The transition state for both the glycosylation and deglycosylation steps of retaining glycosidase catalysis is believed to closely resemble the oxacarbenium ion. In the case of the oxacarbenium ion, the positive charge developed on the sugar is

delocalized between the anomeric carbon and the *endo*-oxygen (Figure 1.4). This introduces a partial double bond character in the bond between the anomeric carbon and the ring oxygen, thus forcing atoms C-5, O-5, C-1 and C-2 to be in one plane. Belief in this kind of transition state has been supported by α-deuterium kinetic isotope effects (α-DKIE) measured with several β-glycosidases.[25-29] For example, in the case of Agrobacterium β-glucosidase, in both the glycosylation and deglycosylation steps the ratio of the rates of hydrolysis of the normal substrate to that of the substrate deuterated at the anomeric carbon ($k_H/k_D$) was found to be positive. [29] This indicated a transition state with increased $sp^2$ character at the anomeric center when compared to the ground state. Moreover, the same study suggested that the transition state for the glycosylation step is 'dissociative', meaning that there is significant cleavage of the bond between the anomeric carbon and anomeric oxygen.

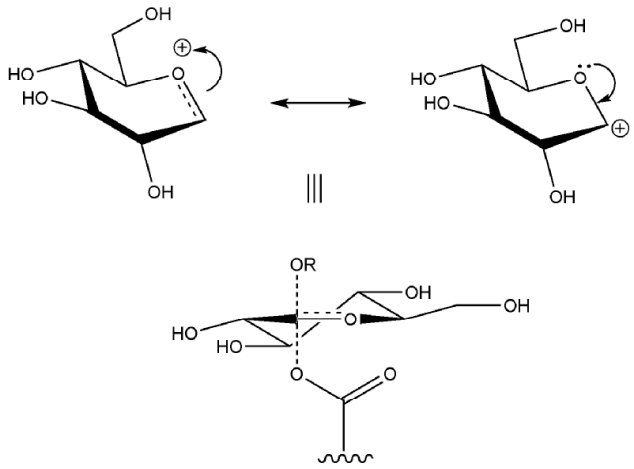



**Figure 1.4.** Proposed structure of the oxa-carbenium ion-like transition state for the hydrolysis reaction catalyzed by α-retaining glycosidases, where C-5, O-5, C-1, and C-2 atoms are in one plane.

In addition to this, fluorinated sugars have also been used as substrates to validate the hypothesis of delocalization of positive charge. Placement of fluorine at the 2- or 5-

position has been shown to reduce the rate of hydrolysis significantly for β-retaining enzymes.[30, 31] These modifications destabilize the positive charge development on the anomeric carbon and the *endo*-oxygen (by inductive effects), thereby making it less susceptible to the attack from the nucleophile. It was also proposed that these modifications might affect the conformational properties of the glycone, possibly resulting in a loss of potential interactions of the glycone with the residues in the active site.[32]

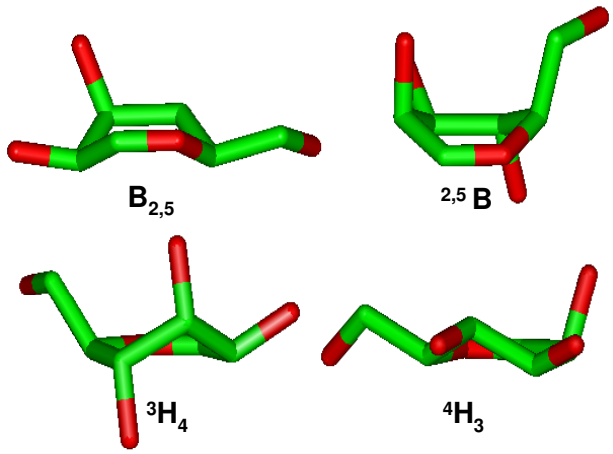



**Figure 1.5.** Possible conformations of the transition state for manno-pyranose ring, which display atoms C-5, O-5, C-1, and C-2 in one plane.

The suggested nature of the transition state leaves only a few conformational itineraries that the pyranosyl ring can follow *en route* to the transition state structure. Ananlysis of the pseudorotation wheel for the pyranose ring suggests that there are only four possible conformers which support the planarity of the C-5, O-5, C-1 and C-2 atoms: $^{4}H_{3}$, $^{3}H_{4}$, $^{2,5}B$ and $B_{2,5}$ (Figure 1.5). Indeed, crystallographic studies of family-5 and -7 retaining cellulases have suggested a $^{4}H_{3}$ half chair conformation at the transition state.[21-23] Furthermore, a $B_{2,5}$ boat conformation was proposed as a transition state conformation for $\beta$-mannosides.

Linus Pauling suggested in the 1940s that the enzymes catalyze the reactions by binding tighter in the transition state conformation, leading to the suggestion that the transition state analogues should act as high-affinity inhibitors of enzymes. In the case of glycosidases, several inhibitors have been reported as transition state analogues. For example, D-manno-1,5-lactone[33-35] and D-manno-1,5-lactam[36] are found to adopt a $B_{2,5}$ conformation, which is the suggested transition state conformation, and both are strong inhibitors of $\beta$-mannosidases. Thus, the knowledge of transition state structure would facilitate efforts aimed at the design of potent inhibitors that mimic transition state conformation.

### 1.7 The Glycosyl-Enzyme Covalent Intermediate

Similar to the transition state, the nature and conformation of the glycosyl-enzyme covalent intermediate (CINT) is also a subject of extensive research. The earliest explanation of the nature of CINT was based on the crystal structure of hen egg-white lysozyme[20], which suggested that the CINT must exist as an ion-pair. A more recent study also suggests that nucelophilic catalysis is unlikely without major conformational changes within the active site of the lysozyme.[37]

α-Deuterium kinetic isotope effect (αDKIE) has also been used to probe the nature of CINT. The use of αDKIE is based on the expectation that if the intermediate formed is a covalent intermediate, the anomeric carbon should be $sp^3$ hybridized. Since the transition state on either side of CINT is oxa-carbenium ion-like with more $sp^2$ character than the ground state, the anomeric carbon must be covalently linked with the catalytic nucelophilic residue. Indeed, in studies in which αDKIE was measured for the

deglycosylation step, kH/kD was found to be greater than unity[15], a finding consistent with a covalent intermediate.

Efforts to isolate covalent intermediates have also been successful using substrate mimics in the presence of wild type as well as mutant enzymes. One of the strategies employed is to use 2- and 5-fluorosugars as mechanism-based deactivator.[30, 31, 38] As discussed above, these substrate mimics reduce the rate of hydrolysis of CINT. If the difference in the rates of glycosylation and deglycosylation reaction is significant, the covalent intermediate accumulates and can be isolated.

## 1.8 Glycosidases in the N-glycan Biosynthesis Pathway and Their Biological Roles

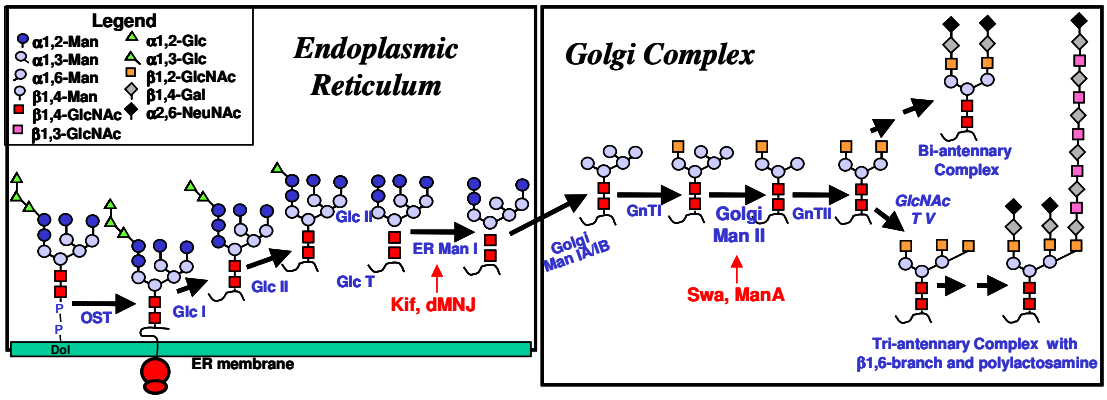



**Figure 1.6.** Biosynthesis of glycoproteins that possess *N*-linked glycans. As the nascent glycoprotein enters the endoplasmic reticulum (ER), a preformed oligosaccharide known as the dolicol-phosphate precursor (DPP) is attached co-translationally to some Asn residues that are part of the consensus sequence Asn-Xaa-Ser/Thr. The biosynthesis of this precursor, its attachment to Asn residues and the subsequent steps of its processing in the ER and the Golgi, are performed by a series of glycosidases and glycosyltransferases.

The biosynthesis of N-linked glycans is a complex process that occurs in secretory pathway of eukaryotic cells and requires the participation of glycoscyl

transferases as well as glycsoidases. The enzymes involved in the maturation of mammalian N-linked glycans are found in the ER and Golgi complex, where they act upon newly synthesized glycoproteins to generate an array of complex structures from a common oligosaccharide precursor.[39] The synthesis and maturation of N-linked oligosaccharides occurs in four stages in the membranes of the ER and Golgi complex.

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 through the action of the oligosaccharide transferase. [40] The second stage involves trimming of the transferred oligosaccharide resulting in the removal of all three glucose residues and six of the nine original mannose residues to generate the Man3GlcNAc2 core structure present on complex type oligosaccharides. The third stage of maturation involves synthesis of GlcNAc branches on the Man3GlcNAc2 core by the action of a collection of GlcNAc transferases. The final stage involves elaboration of these branches with unique capping structures, which are characteristic of complex type oligosaccharides.[1, 41] This final modification of the N-glycans results in extremely diverse oligosaccharide structures which are implicated in the adhesion of cells, either by direct interaction with lectins associated with the extracellular matrix or the surfaces of other cells, or indirectly by influencing the protein-protein interaction of cell adhesion molecules at points of focal contact.

While it is clear from the pathway depicted in Figure 1.6 that glycosidases are essential to the production of substrates required for Golgi glycosyl transferases, glycosidases in the ER are also involved in quality control. As shown in Figure 6, *N*-linked oligosaccharide chains are co-translationally added to proteins in the ER lumen as

$Glc_3Man_9GlcNAc_2$ precursors. The outer and two inner glucoses are trimmed by ER α-glucosidases I (Glc I) and II (Glc II), respectively. Incorrectly folded glycoproteins are recognized by the enzyme UDP-Glucose:glycoprotein lucosyltransferase (GT) which re-attaches a single glucose to the *N*-linked oligosaccharide chain.[5] Monoglucosylated glycans are then recognized by the lectin-like ER chaperones calnexin (CNX) and calreticulin (CRT), which promote correct folding by inhibiting aggregation and preventing premature oxidation and oligomerization. The reattached glucose is then removed by Glc II. Depending on the folded state of the glycoprotein, it either continues down the maturation pathway in the Golgi apparatus or enters a reglucosylation/deglucosylation cycle until it is fully folded and can exit the ER. Inhibitors of these enzymes are known to cause delays in the folding and intracellular transport of glycoproteins and are thus compounds of interests in various therapies.[5] Similar to Glc I and Glc II, recent studies have shown that class I α-mannosidase located in the ER plays an important role in the degradation of misfolded glycoproteins.[5]

### 1.9 Association of the β1,6-Branch with Malignancy

The complexity of N-glycan structures is largely based on the cell-specific expression of a collection of glycosyl transferases, which specify the extension of oligosaccharide structures onto the trimmed $Man_3GlcNAc_2$ core, structure derived from ER and Golgi α-mannosidase action.[1] Linkage specific GlcNAc tranferases establish the branching pattern of the $Man_3GlcNAc_2$ core resulting in common bi-, tri- or tetra-antennary complex type structures that are further extended with either simple or highly complex branched or linear terminal capping structures. Among the most extended and

complex termini are the polylactosamine-containing (repeating Galβ1,4GlcNAcβ1,3-) structures that are preferentially added to a single β1,6GlcNAc branch from the tri-mannosyl core. These structures are high-affinity ligands for extracellular lectins such as galectin 1 and 3, and act as the scaffold for elaboration of branched termini, such as $Le^x$, $Le^y$ and sailyl-$Le^x$ antigen sequences.

One of the rate limiting factors in polylactosamine expression on N-glycans is the synthesis of the β1,6GlcNAc branch on the $Man_3GlcNAc_4$ core. The reaction is catalyzed by the Golgi glycosyltransferase GlcNAc transferase V (Gn TV).[42-44] In mammalian tissues, β1,6GlcNAc branched structures on the plasma membrane are largely restricted to cells that are capable of invasion, including endothelial cells and activated lymphocytes.[45, 46] In transformed cells and human cancer cells, including carcinoma of the breast, colon and skin, a significant increase has been found in β1,6GlcNAc branching.[47-50] Since the extent of β1,6-branching and polylactosamine structures correlates with disease progression, these structures are considered as markers of malignancy in many tissues.[51, 52] The role of β1,6-branching in metastasis and tumor progression was further addressed by the generation of Gn TV-deficient mice, which indicated that β1,6GlcNAc branching catalyzed by Gn TV influences the migration of tumor cells and leucocytes, presumably by affecting their adhesion to the extracellular matrix, and that induction of β1,6-branching contributed to the altered adhesion characteristics of metastatic cells.[53]

## 1.10 Glycoprotein Processing Inhibitors as Anti-tumor Agents

While evidence for the correlation between β1,6-branching, metastasis, tumor progression and Gn TV expression levels is compelling, no natural or synthetic inhibitors of Gn TV have been identified that are cell-permeable and effective in blocking oligosaccharide processing *in vivo*. Since selective and high inhibition of Gn TV has not yet been achieved, inhibitors, that act earlier in the N-glycan biosynthesis pathway and interfere with the formation of the precursor for β1,6-branching, have been tested as anti-tumor agents.

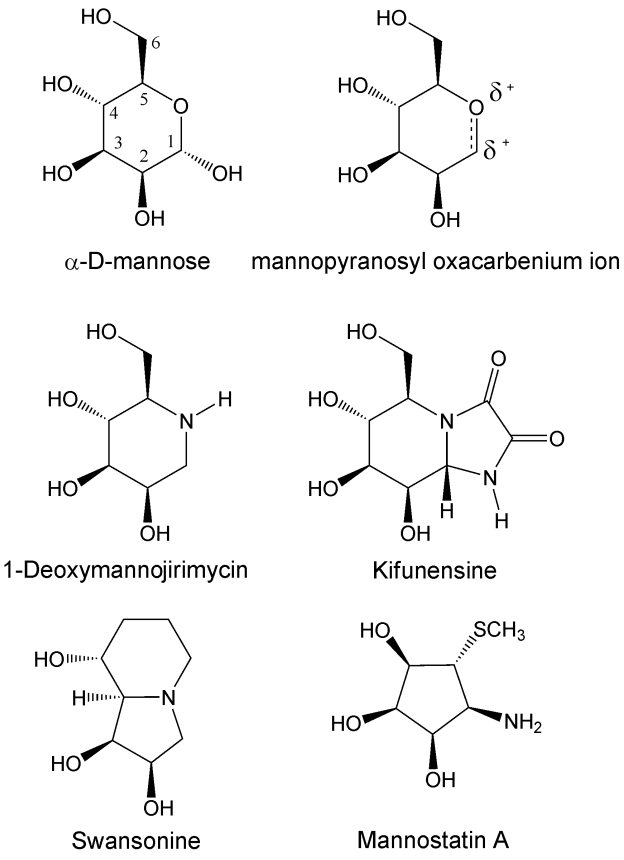



**Figure 1.7.** Glycoprotein processing inhibitors.

Among the known inhibitors, glucosidase inhibitors block the N-glycan processing at an early stage, resulting in the generation of glucosylated high-mannose type oligosaccharides. In addition, these inhibitors tend to be less effective due to multiple routes of early oligosaccharide processing in mammalian cells. The α-mannosidase inhibitors can be divided into two groups based on their selective action toward the two classes of α-mannosidases in mammalian cells.[39, 54] Aza-sugar mannopyranose mimics such as kifunensine and deoxymannojirimycin inhibit class I mannosidases, resulting in extended high mannose structures.[55, 56] In contrast, inhibition of class II mannosidases by presumed manno-furanoses transition state mimics such as swainsonine and mannostatins, results in the formation of hybrid-type structures.[55]

Glucosidase and class I α-mannosidase inhibitors have some potential as inhibitors of β1,6-branch formation because they can block N-glycan processing. However, accumulation of extended high-mannose structures by these compounds can also affect the function and intra cellular transport of some glycoproteins.[54, 57-60] On the other hand, the class II inhibitors do not appear to influence the function or transport of glycoproteins and exhibit low toxicity in tissue cultures and animals, including mice and humans.[61]

Among the known class II α-mannosidase inhibitors, swainsonine, which was first isolated from the Australian plant *Swiansona canescens* and subsequently from additional plants in America[55, 62], is the most studied. Since swainsonine influences the maturation of glycoproteins, yielding hybrid-type rather than complex-type structures, it could be anticipated that the compound might influence cell adhesion characteristics and thus the metastatic behavior of transformed cells.[63] The low toxicity of the compound

was revealed when no changes in cell growth rates were found in transformed cells in culture. Furthermore, when transformed cells were introduced into animal models, swainsonine was found to possess strong immunostimulatory activity which resulted in a significant anti-cancer effect.[63, 64] Swainsonine was shown to stimulate natural killer cells and lymphokine activated killer cells, and significantly blocked tumors in the lungs, liver, and spleen.[64, 65] The compound was also shown to stimulate the anti-tumor activity of macrophages in a time- and dose-dependant manner.[66] Furthermore, in addition to stimulating bone marrow proliferation, swainsonine also shown to enhance bone marrow cellularity, engraftment efficiency, and colony forming units in mice.[63] Two phase I clinical trials have demonstrated that swainsonine has a low toxicity in humans and is well tolerated.[63, 67]

Swainsonine and all other known GMII inhibitors have an additional serious side effect that precludes their use a therapeutic in that they also inhibit the broad specificity lysosomal α-mannosidase, resulting in the accumulation of mannose-containing oligosaccharides in tissues, serum and urine as a phenocopy of the hereditary lysosomal storage disease, α-mannosidosis.[68, 69] This disease results in various clinical features such as coarse features, skeletal abnormalities, dilated cerebral ventricles, and severe mental retardation.[70, 71] Thus, a preferable approach for therapy to inhibit tumor growth and metastasis is to develop alternate lead compounds that specifically inhibit GMII without affecting lysosomal α-mannosidase. It should be mentioned that despite the similarities between the two mannosidases in the sequence, response to inhibitors, and catalytic mechanism as retaining enzymes, GMII and lysosomal α-mannosidase have significant

differences in the aglycone specificity that can be utilized for selective inhibitor design.[72-74]

## 1.11 The X-ray Structure of Swainsonine:GMII Complex

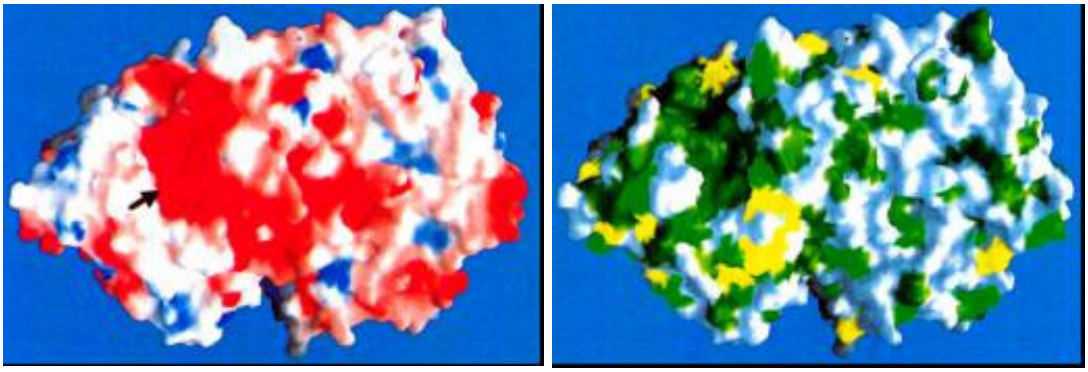



**Figure 1.8.** A. Molecular surface representation of the planer surface of dGMII (adopted from reference 75). The arrow indicates location of the active site region. B. Molecular surface of dGMII, colored for homology with the sequence of human GMII (dark green for identical, light green for conserved, yellow for similar and white for different residues).

Although the crystal structure of human GMII (hGMII) is not yet available, X-ray structures of GMII obtained from *Drosophila melanogaster* (dGMII) in absence of ligand and in complex with swainsonine and 1-deoxymannojirimycin (PDB codes 1HTY, 1HWW and 1HXK) have been described in detail by van den Elsen et. al.[75]. This enzyme has high sequence identity when compared with human counterpart (41% identity and 61% similarity).[75] dGMII has also been shown to display comparable kinetic properties and inhibitor sensitivity as well as identical substrate specificity to hGMII.

dGMII is an oval-shaped molecule with two distinct faces. The N-terminal face of dGMII is convex, whereas the opposing face has a planar surface. The molecular surface of dGMII reveals an expanded pocket in the N-terminal $\alpha/\beta$-domain, formed primarily by

acidic residues (Figure 1.8A). These same residues form the core of a large, surface-exposed patch of highly conserved amino acids, as compared to hGMII sequence (Figure 1.8B). The active site of the enzymes is located in a small cavity inside this region.

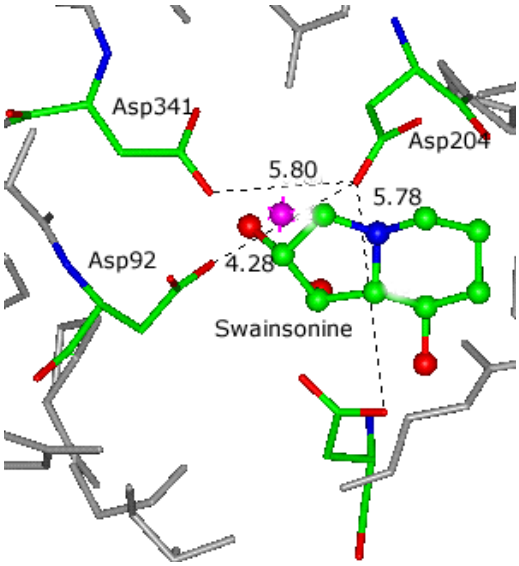



**Figure 1.9.** Active site region of dGMII with a bound swainsonine molecule. Distances between catalytic residues are represented by dotted lines.

From the structures of dGMII-inhibitor complexes, four acidic amino acid residues, specifically Asp92, Asp204, Asp341 and Asp472, were identified as candidates for catalytic side chains based on their proximity in the active site (Figure 1.9). Studies of the catalytic mechanism of Jack-bean α-mannosidase have identified an aspartate residue as a catalytic nucleophile.[76] Comparison of the highly conserved sequence region surrounding this aspartate in Jack-bean α-mannosidase with the same sequence region in dGMII suggests that Asp204 is the catalytic nucleophile that attacks the glycosidic linkage.[75] An identical finding was confirmed for another α-mannosidase belonging to family 38, the bovine lysosomal α-mannosidase.[77]

As dGMII is a retaining α-mannosidase, the distance between the catalytic nucleophile and the base should be about 5.5 Å. Based on this expected distance, Asp341 and Asp 472 are likely candidates as the catalytic base. Geometrically Asp341 is well positioned to be the catalytic base, a notion that is also supported by the fact that the D341N mutant is catalytically inactive.[75]

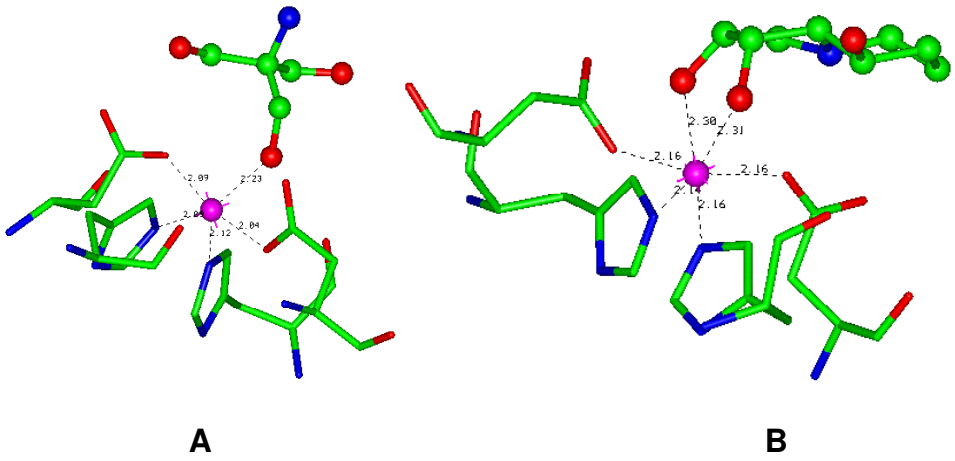



**Figure 1.10.** Binding site of dGMII showing Tris (A) and swainsonine (B) bound to the enzyme and coordination of zinc ion. Zinc ion coordinates with $T_5$ geometry in case of Tris molecule whereas, coordinates with $T_6$ geometry in the case of swainsonine.

The crystal structure of dGMII complexed with 2-amino-2-hydorxy methyl propane-1,3-diol (Tris) also revealed the presence of the enzyme-bound $Zn^{+2}$ ion[75] in the binding site. It was also observed that $Zn^{+2}$ ion binds to the enzyme in a $T_5$ square-based pyramidal geometry (Figure 1.10A) by coordinating to OD1 oxygens of Asp92 and Asp204, the NE2 nitrogens of His90 and His471 and the hydroxyl oxygen O2 of the bound Tris molecule. It was observed that, for dGMII-inhibitors complexes the coordination geometry of $Zn^{+2}$ ion changes to a less favored[78] hexavalent ($T_6$) state (Figure 1.10B). It was proposed that in the actual glycosidase reaction this change in

geometry might energetically facilitate the deglycosylation step by reverting to more stable $T_5$ geometry. Thus, zinc may contribute directly to the glycosidase mechanism.

### 1.12 Covalently Linked Glycosyl-Enzyme Intermediates of dGMII

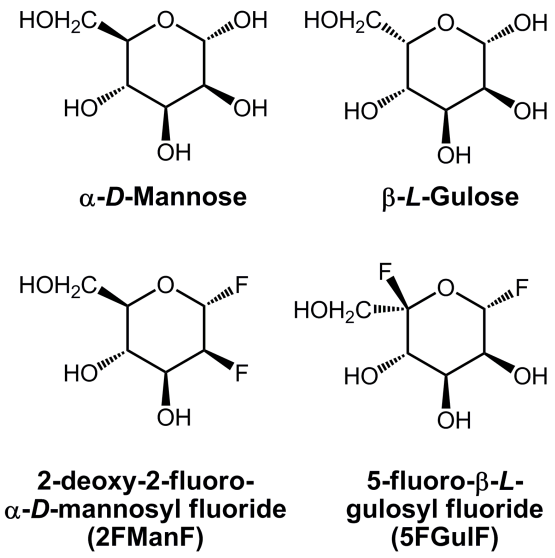



**Figure 1.11.** Substrates used to trap the glycosyl-enzyme covalent intermediate for dGMII.

In addition to the efforts to crystallize dGMII in presence of inhibitors, high resolution X-ray structures of glycosyl-enzyme covalent intermediates using fluorinated sugars have also determined (PDB Ids: 1QWN, 1QWU and 1QX1).[38] The authors employed both the wild type and D341A mutant enzyme and used mechanism-based deactivators 2-hydroxy-2-fluoro-α-*D*-mannosyl fluoride (2FManF) and 5-fluoro-β-*L*-gulosyl fluoride (5FGulF) for this purpose (Figure 1.11). Inspection of all three X-ray structures revealed a $^1S_5$ skew boat conformation of the sugar ring. In addition to moving the C-2 hydroxyl into a pseudo-equatorial orientation, the observed $^1S_5$ skew boat conformation also avoids a steric clash between this group and Asp204OD1.

Furthermore, in this conformation the leaving group was found to be anti-periplanar to the lone pair of the ring oxygen, a requirement for the departure of the leaving group according to Deslongchamps' anti-periplanar lone-pair hypothesis. [79] The electron density maps for all three structures showed the presence of a covalent linkage between the anomeric carbon and nucleophilic carboxylate oxygen of Asp204. Comparison of covalent intermediates with bound conformations of swainsonine and kifunensine revealed that covalent intermediates mimic best swainsonine. This is consistent with the notion that swainsonine mimics the transition state, as the reactive intermediate should itself mimic the transition state according to the Hammons postulate.

### 1.13 Role of Zinc in the Active Site

The active site of GMII is characterized by the presence of a $Zn^{2+}$ ion coordinated by four protein side-chains, specifically His90, Asp92, Asp204, and His471. Zinc atoms are very common in metalloproteins and play a number of roles. There are number of intrinsic characteristics of zinc that are relevant to its role in proteins. With full d-orbital, it is quite stable electronically and not prone to redox changes. It can adapt to anywhere from 2 to 8 coordinates, though 4, 5 and 6 are the most common. Furthermore, it can adapt to its ligation state according to its environment, for example during the course of a reaction.

In the case of GMII, Asp204, which is one of the side-chains coordinating with zinc, is the nucelophilic side-chain in the catalytic mechanism. This coordination thus helps to maintain the deprotonated state of the catalytic nucleophile. Recently Kuntz et al. proposed that in addition to this, zinc is important for substrate binding and subsequent

conformational changes in the substrate. This hypothesis is based on the co-crystal structure of noeuromycin complexed with dGMII [80] and other crystallographic studies.[38, 75, 81, 82]

### 1.14 Features of GMII Inhibitors

Most natural as well as synthetic potent inhibitors of GMII reported thus far display certain features that are common and essential for high inhibitory activity. These features include a *cis*-diol moiety, basic amine nitrogen and other functionalities, typically hydroxyl groups, which presumably mimic hydroxyl groups in the natural substrate. The *cis*-diol moiety is believed to mimic the C-2 and C-3 hydroxyl groups in the natural GMII substrate and coordinates with zinc in the active site. Indeed, the X-ray structure of swainsonine complexed with dGMII showed coordination of the vicinal hydroxyl groups with the $Zn^{+2}$ ion in the active site.[75] Similar coordination has also been reported in the case of other ligands. [38, 75, 80-82] Furthermore, the synthesis and biological evaluation of a small number of mannostatin analogues has revealed that the basicity of the primary amine and neighboring *cis*-diol is essential for inhibitory activity.[83-88] The possible role of methylthio and methyl sulfenyl functions in mannostatin A and B, respectively is intriguing. These features are unique to these natural mannosidase inhibitors. It has been proposed that these functionalities lie near the 4-hydroxyl group of mannosyl oxacarbenium ion but it is also conceivable that these moieties mimic the hydroxyl methyl group in the natural substrate.[89] If these sulfur containing functionalities play a role in the recognition and binding of the mannostatins by GMII, the extent to which these structural elements are important remains unclear.

## 1.15 Objective of the Dissertation

Although the mechanism of retaining glycosidases is common across a variety of different glycosidase families, the detailed atomic characteristics that drive the mechanism are diverse. An understanding of the details of these characteristics at an atomic level is necessary if one is to harness the therapeutic potential of glycosidases. While structural and mechanistic studies on dGMII:ligands complexes have provided a great deal of information, several aspects of the inhibition and catalytic mechanism are not yet clear. We have employed a computational approach in conjunction with experiments to address these issues. The specific aims of the project are:

1. *To enhance our understanding of the mechanism of inhibition of GMII by mannostatin A.* This project involved prediction of the binding mode of mannostatin A and its analogues in the active site of dGMII and subsequent validation of these predictions using X-ray crystallography. We have employed molecular docking in conjunction with molecular dynamics (MD) simulations for this purpose (chapters 2 and 3).
2. *To determine the role of the thiomethyl moiety in mannostatin A in binding.* We have used MD simulations, along with X-ray crystallography and biological evaluation, of some mannostatin A analogues in order to study the dynamics of the ligands and ligand-protein complexes, and to determine the role of the thiomethyl functionality (chapter 3).

3. *To determine the catalytic mechanism of hydrolysis reaction catalyzed by GMII.* We have applied a quantum mechanical (QM) methodology to determine the structures of all reactive species along the reaction coordinate (chapter 4).

The results obtained in this study will be instrumental in achieving the ultimate goal of designing and synthesizing more potent and /or selective inhibitors of GMII.

## 1.16 Overview of the Computational Methods Used in this Dissertation

Interactions between ligands and receptors are fundamental to all biological processes. A theoretical understanding of these interactions is of great significance in understanding the mechanism by which these biomolecules interact. The knowledge of a three-dimensional structure of the receptor, either in the absence or in the presence of a ligand, provides a great deal of information about the interactions critical for ligand binding. When the structure of a receptor is known, structure-based computational strategies can be employed to understand more about the inhibition and catalytic mechanism of the target protein. In the case of GMII, a plethora of structural data is available, which makes it possible to employ computational techniques such as molecular docking and MD simulations in order to understand more about the inhibition and the catalytic mechanism of GMII. The following sections present a general review of computational methods that have been used in this research.

### 1.16.1 Molecular Docking

The docking process involves sampling the coordinate space of the binding site and scoring each possible ligand pose. The pose with the best score is then taken as the

predicted binding mode of that compound. There are a large number of docking programs available and they differ in the sampling algorithms used, handling of the ligand and protein flexibility, and scoring functions they employ. Several commonly used docking methods are listed in Table 1 along with the type of sampling and scoring algorithms used. Some of these programs are commercial software packages and expensive to use. We have used AUTODOCK[90] for molecular docking, which has been successfully used in studies of several ligand-protein complexes.

The majority of the docking tools, including AUTODOCK, currently assume that the protein target is fixed in its crystal structure conformation.[91] This is usually an inaccurate approximation, but a necessary one because of the increased complexity and consequently the computational cost that is required to accurately sample the flexibility of the binding site. In the case of GMII, X-ray structures with several ligands have revealed that there is only a minor change in the biding site upon ligand binding when compared with the X-ray structure of the native enzyme (PDB ID 1HTY[75]). Thus, in order to predict the binding modes of mannostatin A and its analogues, we have ignored the flexibility of the GMII binding site.

Another important aspect of molecular docking is the conformational flexibility of the ligand. The conformation of a compound bound to a macromolecule may be different from its conformation in the unbound form. It is therefore crucial to consider the conformational flexibility of the compounds during the docking process. This can be achieved by pre-computing a set that contains several conformers of the compound to be docked. Several methods are available to generate large conformational databases rapidly and each conformation can be rigidly docked into the targeted biding site. Examples of

methods that work with conformational ensembles are Dock, Fred and slide. Alternatively, since the size of the conformational space for a ligand is directly proportional to the number of rotatable bonds present in the ligand, a variety of efficient algorithms have been employed to explore the conformational flexibility as the calculation proceeds. For example, the genetic algorithm has been successfully employed in GOLD, whereas the Lamarkian genetic algorithm has been used in AUTODOCK.

**Table 1.** Common programs used for molecular docking.

| Method | Ligand flexibility sampling | Scoring function |
|---|---|---|
| Dock | Incremental build | Force field or constant score |
| FlexX | Incremental build | Empirical score |
| Slide | Conformational ensembles | Empirical score |
| Fred (Openeye Software) | Conformational ensembles | Gaussian score or empirical scores |
| Gold | Genetic algorithm | Empirical score |
| Glide (Schrodinger) | Exhaustive search | Empirical score |
| Autodock | Genetic algorithm | Force field |
| LigandFit (Accelrys) | Monte Carlo | Empirical score |
| ICM | Pseudo-Brownian sampling and local minimization | Mixed force field and empirical score |
| QXP | Monte Carlo | Force field |

One limitation of AUTODOCK is that it does not take into account the flexibility of the rings in the cyclic ligand and only rotatable bonds are rotated in order to sample the conformational space for the ligand. However, five-membered rings, such as that in mannostatin A, are inherently flexible, making it necessary to consider the flexibility of the five-membered ring while docking. While MD simulations can be used to sample populated conformations in the solution, it is possible that the solution phase conformer is

different than the one which is biologically recognized. Therefore, we have employed quantum mechanical calculations to derive a set of different ring conformers of mannostatin A and its analogues. These pre-computed conformers were then docked individually into the active site of dGMII.

The molecular docking technique can also be used to estimate the free energy of binding for a particular ligand-protein complex. Currently, most structure-based docking methods employ an empirically derived free energy function that reproduces observed binding constants. Most of these approaches follow an expanded "master equation" to model the free energy of binding, adding entropic terms to the molecular mechanics equations:

$$\Delta G = \Delta G_{vdW} + \Delta G_{h-bond} + \Delta G_{electrostatics} + \Delta G_{conformational} + \Delta G_{torsional} + \Delta G_{solvation} \qquad \text{Equation 1.1}$$

AUTODOCK implements a similar approach using the thermodynamic cycle of Wesson and Eisenberg.[92] The scoring function used in AUTODOCK includes the following five terms:

$$\begin{aligned} \Delta G &= \Delta G_{vdW} \sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) \\ &+ \Delta G_{h-bond} \sum_{i,j} E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) \\ &+ \Delta G_{electrostatics} \sum_{i,j} \frac{q_i q_j}{\varepsilon\left(r_{ij}\right) r_{ij}} \\ &+ \Delta G_{torsional} \; N_{tor} \\ &+ \Delta G_{solvation} \sum_{i,j}\left(S_i V_j + S_j V_i\right) e^{\left(-r_{ij}^2/2\sigma^2\right)} \end{aligned} \qquad \text{Equation 1.2}$$

The five $\Delta G$ terms on the right hand side of equation 1.2 are coefficients empirically determined using linear regression analysis of a set of protein ligand complexes with known binding constants. The summations are performed over all pairs of ligand atoms, *i*, and protein atoms, *j*, in addition to all pairs of atoms in the ligand that are separated by three or more bonds.

The '*in vacuo'* contributions include three interaction energy terms: a Lennard-Jones 12–6 dispersion / repulsion term; a directional 12–10 hydrogen bonding term, where $E(t)$ is a directional weight based on the angle, *t*, between the probe and the target atom; and a screened coulombic electrostatic potential. Each of these terms, including their parameterization, have been described.[93] A measure of the unfavorable entropy of ligand binding due to the restriction of conformational degrees of freedom is added to the '*in vacuo'* function. This term is proportional to the number of $sp^3$ bonds in the ligand, $N_{tor}$.

For the calculation of the desolvation term, AUTODOCK employs the pair-wise, volume-based method described by Stouten et al.[94] since it is consistent with the pre-calculated affinity grid formulation used by AUTODOCK. In this method, for each atom in the ligand, fragmental volumes of surrounding protein atoms are weighted by an exponential function and then summed, evaluating the percentage of volume around the ligand atom that is occupied by protein atoms. This percentage is weighted by the atomic solvation parameter of the ligand atom to give the desolvation energy.

### 1.16.2 MD Simulations of Ligand-Protein Complexes

A fundamental appreciation for how biological molecules interact requires knowledge of their three-dimensional structure and dynamics. While experimental techniques such as X-ray crystallography provide a great deal of information, these techniques are unable to explore the dynamic nature of the interactions that are important for complex formation. MD simulation is a powerful method for exploring the dynamic nature of the interactions between biological molecules and conformational changes in the ligand and the macromolecule following ligand binding. Almost 30 years have passed since the report of the MD simulation of a protein, BPT1.[95] Since then this, computational technique has become an established tool in the study of biomolecules and is complementary to experimental techniques.

The biomolecular simulations find three major areas of application today.[96] First, MD simulations are used for exploring the conformational space of a ligand. This is an advantageous approach as it is possible to sample conformations that are populated in solution. Furthermore, the conformations generated in this manner can be docked individually into the binding site of the target protein and scored according to the predicted binding affinity. The second major application is the study of the interactions of the ligand with protein residues in a ligand-protein complex. In addition to capturing conformational changes in ligand and/or in the protein, such a simulation also reveals the dynamic nature of interactions that are critical for ligand binding. Knowledge of the X-ray structure of the target protein in the presence or absence of the ligand is beneficial for setting up an MD simulation of the ligand-protein complex. Finally, MD simulations can

also be used to estimate free energy changes for chemical processes such as ligand binding.

In this dissertation, we have employed the AMBER[97] suit of programs to perform MD simulations of mannostatin A and its benzylated analogue complexed with dGMII. We have also performed MD simulations of these ligands in an unbound state (chapter 3). The simulations have provided valuable insights into the dynamic nature of interaction that are important for the binding of mannostatin A and the conformational properties of mannostatin A in water. Furthermore, this study has also revealed the importance of the thiomethyl moiety for the binding of mannostatin A by GMII.

Mathematically, the MD simulation method is based on Newton's second law of motion stated in equation 1.3.

$$F = m \cdot a \qquad \text{Equation 1.3}$$

Where, $F$ = the force exerted on the particle,

$m$ = the mass of the particle, and

$a$ = its acceleration

The MD simulation is a deterministic method whereby the knowledge of the total forces on a particular atom at time $t$ can be used to determine the positions and velocities of that particle at any time ($t + \delta t$). The result of a simulation is a trajectory that specifies how the positions and velocities of the atoms in the system vary with time. A general solution to this problem was proposed by Verlet in 1967.[98] ABMER implements a variation of this solution referred to as the leap-frog Verlet algorithm, (ref) which uses a

series of equations to describe atomic position ($x$), velocity ($v$) and acceleration with respect to time ($t$). Since acceleration is the rate of change of the velocity and velocity is the rate of change of displacement, the relationship between these variables can be summarized in the following equation.

$$a = \frac{dv}{dt} = \frac{d^2x}{dt^2} \qquad \text{Equation 1.4}$$

Integrating equation 1.4 we get,

$$v = a \cdot t + v_0 \qquad \text{Equation 1.5}$$

and,

$$x = v \cdot t + x_0 \qquad \text{Equation 1.6}$$

Combining equation 1.6 with the expression for velocity (1.5), the following equation is obtained which gives the value of $x$ at time $t$ as a function of acceleration, $a$, initial position, $x_0$, and initial velocity, $v_0$.

$$x = a \cdot t^2 + v_0 \cdot t + x_0 \qquad \text{Equation 1.7}$$

Furthermore, the force can also be expressed as the first derivative of the potential energy (E) with respect to the position, $x$. Therefore,

$$F = -\frac{dE}{dx_i} \qquad \text{Equation 1.8}$$

Combining equations 1.4 and 1.8 we get,

$$m \cdot a = -\frac{dE}{dx} \qquad \text{Equation 1.9}$$

Thus, the acceleration of each atom is given as,

$$a = -\frac{1}{m}\frac{dE}{dx} \qquad \text{Equation 1.10}$$

The calculated value of acceleration is then used in equation 1.7 to determine the new position of the atom. Thus, initial positions and velocities of the atoms in the system determine their positions and velocities at all other times. Initial positions can be obtained from experimental techniques such as NMR or X-ray crystallography.

In the MD simulation, the potential energy of the system is defined as a sum of different energetic contributors as depicted in equation 1.11 below.

$$E_{total} = E_{bonds} + E_{angles} + E_{torsions} + E_{vdW} + E_{electrostatic} \qquad \text{Equation 1.11}$$

More completely, this equation, also known as the force field equation, can be written as:

$$E_{total} = \sum_{bonds} \frac{k_i}{2}\left(D_i - D_{ref}\right)^2 + \sum_{angles} \frac{k_i}{2}\left(\theta_i - \theta_{ref}\right)^2 + \sum_{torsions} \frac{V_i}{2}\left(1+\cos\left(n\omega_i - \gamma_i\right)\right) + \sum_{vdW} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{electrostatic}\left(\frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right)$$

Equation 1.12

In equation 1.12, the first two terms are simple harmonic and contribute energy penalties for any deviations the reference values. The third terms represents a description of the energy barrier for the torsion potential (V), multiplicity or number of minima during a 360° rotation (n), phase factor (γ), and torsion angle (ω). The van der Waal term is often described by a Lennard-Jones potential where the well depth is $\varepsilon_{ij}$, collision diameter is $\sigma_{ij}$, and inter-atomic distance is $r_{ij}$. The electrostatic term is a coulombic term that includes the charge on the interacting atoms ($q_i$, $q_j$) and the inter-atomic distance.

In addition to the above, it is essential to include solvent effects in the simulation of biomolecular systems. In most simulations, water is the solvent because almost all interactions, which involve biomolecules, take place in the presence of water. There are two methods by which solvent effects can be incorporated into the simulation. In the first method, the solvent is explicitly represented in the system of interest, whereas in the second method, also known as the continuum solvation method, the solvent is represented by a mathematical expression which describes the average effect of bulk solvent on the solute. This is done by approximating the dielectric constant in the solute and solvent

regions. For the purpose of this dissertation, we have used explicit solvation method to perform MD simulations on ligand-dGMII complexes.

The nature of the AMBER force field requires the use of water molecules parameterized with explicit partial charges. There are several empirical water models currently in use, which differ in their treatment of partial charges, overall geometry, and the ability to properly represent bulk properties. We have used the TIP3P water model[99] in our simulations, as it is an established model over last several years. In this model, the partial charges on the hydrogen atoms are balanced by the negative charge on the oxygen atom.

### 1.16.3 Modeling the Chemistry of the Active Site of GMII

Elucidation of the chemistry of a matalloprotein active site at an atomic level of detail has been an enduring challenge to bioorganic and bioinorganic chemistry. It has proven to be extremely difficult by experimental means alone to discriminate between alternative reaction mechanisms and reactive species, or to determine precisely the relative energies of these species.[100] In principle, accurate computational methods are capable of complementing experimental methods to answer these questions at an atomic level of detail.

The use of traditional *ab initio* quantum chemical methods in modeling the catalytic mechanisms of metalloproteins is limited in terms of the size of the system. The selection of model thus becomes critical as inclusion of large number of atoms can lead to intractable CPU times. On the other hand, a smaller model may not be sufficient to produce reasonable results as it lacks explicit effects of the protein residues on the active

site chemistry. In addition, the structurally and electronically diverse nature of transition metals makes it difficult to use computationally less intensive approaches such as semiempirical and force field based methods, as such systems have proven difficult to describe with a set of universal parameters.

Over the past few years, effective methods have emerged to address these problems. Two alternative approaches exist to make modeling of the metalloprotein active site computationally tractable. First, one can construct a purely QM model by truncating the active site core and capping the truncated regions with hydrogen atoms. This traditional approach has been used for many years with increasingly large models. The widespread use of the Density Functional Theory (DFT), which employs gradient-corrected and hybrid functionals, has demonstrated that transition metal systems can be treated in a reasonable, yet cost-effective manner. Advances in DFT-based algorithms, along with a dramatic reduction in the cost and performance of computer hardware, allow quantum chemical treatment of a system containing about 150-200 atoms. (ref)

Another approach is to retain the entire protein, but treat only a specified region in the active site at the QM level, and model the remaining protein via molecular mechanics. This approach, known as the combined quantum mechanical and molecular mechanical (QM/MM) approach, affords a strategy that takes into consideration the effects of protein environments while still keeping the CPU time reasonably low. The most difficult task in this approach is the partitioning of the system into QM and MM regions. It is not obvious how much of the system needs to be included in the QM region [101]. Variations in results are likely when the size and composition of the QM region is altered, because the description of the interactions described by the QM/MM potential function is different as

compared to the same interactions described by the purely QM semiempirical Hamiltonian approximation.[101] Generally, the two subsystems can be separated by cutting the covalent bonds within the neighboring residues of amino acids included in the QM region. For example, a covalent bond, which links atom X and atom Y, can be cut in such a way that X is in the QM region and Y is in the classical region. The problem occurs at this frontier because the electron of X involved in the covalent bond is not paired with any other electron as in molecular mechanics electrons are not explicitly represented (Figure 12). Thus, this electron needs special treatment.

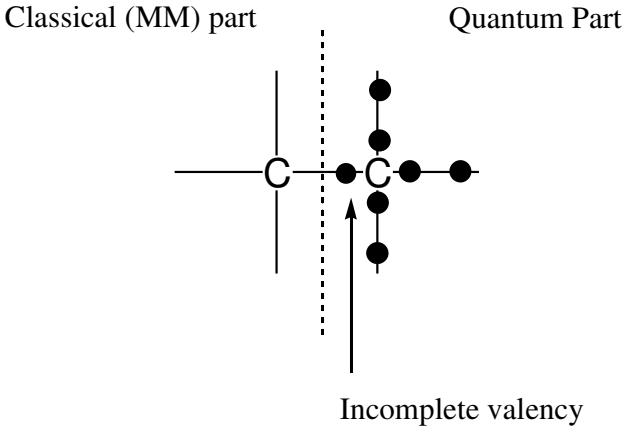



**Figure 1.12.** Example of the problem of cutting covalent bonds at the QM/MM interface: A covalent C–C bond at the frontier between a MM part and a QM part.

Several solutions have been proposed for this problem and can be divided into two broad categories[101]: those that add a pseudoatom to fill the valencies of the quantum frontier atom, (e.g. the link atom method[102] and the connection atom method[103]), and those that deal specifically with the frontier bond orbitals by trying to compute directly its main characteristics from known parameters (e.g. the local self consistent method[104] and the generalized hybrid orbital method.[105]) Although link atoms are an approximate

approach to QM/MM partitioning of a covalently bonded system, they have been found to yield good results for a number of simulations of enzymatic reactions.[106-110]

In order to characterize the catalytic mechanism of GMII, we have employed the DFT based quantum mechanical method. The general strategy used for this task included four major steps: 1. docking the substrate into the active site of GMII; 2. truncation of the active site to create a model of reasonable size; 3. perform a relaxed geometry scan along a pre-defined reaction coordinate; and 4. optimization of the approximate geometries of the intermediates predicted in the scan. We used the JAGUAR program [111] for all the optimizations because it can be run in parallel and is faster than other programs such as Gaussian.[112] The QM/MM implementation used in Jaguar is a program referred to as QSITE. The current version of QSITE does not allow the performing of a relaxed geometry scan. Hence, we have performed density functional calculations on models of differing size (chapter 4).

The results presented in this dissertation are significant not only for GMII, but also for other glycosidases belonging to glycosyl hydrolase family 38, in particular the lysosomal α-mannosidase. Since the binding site of the lysosomal enzyme is very similar to that of GMII, computational insights into the binding of GMII by mannostatin A and its analogues enhance our understanding of the inhibition mechanism of the lysosomal enzyme by mannostatin A. Thus, the research presented in this dissertation is a step forward in the search for potent and, more importantly, selective GMII inhibitors that can be used in cancer therapy.

# CHAPTER 2

# INHIBITION OF GOLGI α-MANNOSIDASE II WITH MANNOSTATIN A ANALOGS: SYNTHESIS, BIOLOGICAL EVALUATION AND STRUCTURE ACTIVITY RELATIONSHIP STUDIES[a]

[a] Li, B.; Kawatkar, S. P.; George, S.; Strachan, H.; Woods, R. J.; Siriwardena, A.; Moremen, K. W.; Boons, G. J., *Chembiochem* **2004,** 5, (9), 1220-1227

## 2.1 Abstract

Mannostatin and aminocyclopentitetrol analogs, which have varying substitutions at the amino function, were synthesized. These compounds were tested as inhibitors of human Golgi- and lysosomal α-mannosidases. Modification of the amine of mannostatin had only marginal effects, whereas similar modifications of aminocyclopentitetrol led to significantly improved inhibitors. *Ab initio* calculations and molecular docking studies were employed to rationalize the results. It was found that Mannostatin and aminocyclopentitretrol could bind to Golgi α-mannosidase II in a similar mode as that of the known inhibitor swainsonine. However, due to the flexibility of the five membered rings of these compounds, additional low energy binding modes could be adopted. These binding modes may be relevant for the improved activities of the benzyl-substituted compounds. The thiomethyl moiety of mannostatin was predicted to make favorable hydrophobic interactions with Arg228 and Tyr727, possibly accounting for its greater inhibitory activity.

## 2.2 Introduction

Cells that have undergone oncogenic transformation often display abnormal cell surface oligosaccharides and these changes in glycosylation are important determinants of the stage, direction and fate of tumor progression.[113] Inhibition of the mannose trimming enzyme human Golgi α-mannosidase II (hGMII), which acts late in the *N*-glycan processing pathway, provides one route to blocking the oncogene-induced changes in cell surface oligosaccharide structures.[63]

HGMII selectively cleaves α1-3 and α1-6 mannosyl residues present in its natural substrate $GlcNAcMan_5GlcNAc_2$.[74] It is a retaining glycosylhydrolase, which employs a two-stage mechanism involving two carboxylic acids positioned within the active site to act in concert: one as a catalytic nucleophile and the other as a general acid/base catalyst.[12, 14, 19, 114, 115] Protonation of the exocyclic glycosyl oxygen of the substrate leads to bond-breaking and simultaneous attack of the catalytic nucleophile to form a glycosyl enzyme intermediate. Subsequent hydrolysis of the covalent intermediate by a nucleophilic water molecule gives a α-mannose product with overall retention of configuration. Studies on retaining mannosidases with 5-fluoro pseudo-substrates and deuterium labeled substrates have shown that the transition states on either side of the covalent intermediate have marked oxocarbenium ion character.[12, 14, 19, 114, 115] Potent inhibitors of glycosidases are thought to mimic oxocarbenium ion-like transition states and for example, the inhibitory activity of the natural product swainsonine has long been attributed to its 5-membered ring resembling a flattened six-membered ring, which had been forced to attain an oxocarbenium-like structure. In fact, the crystal structure of Swainsonine complexed with Golgi α-mannosidase II from *Drosophila* melanogaster shows the inhibitor to be tilted in such a way as to bring the equivalent of its anomeric carbon close to the presumed catalytic nucleophile.[75]

Swainsonine has been much investigated as a consequence of its potent inhibitory properties. Clinical studies have shown[61, 62, 64, 67] it to exhibit potent anti-tumor and anti-metastatic activity. An unfortunate side effect resulting from this compound is blockage in oligosaccharide catabolism arising from inhibition of a related catabolic α-mannosidase found in lysosomes.[63, 67] It is clear that to develop a drug appropriate for

anti-metastatic therapy, an alternate lead compound is required, which is amenable to easy modification in a combinatorial manner to give compounds that specifically inhibit GMII without affecting the function of lysosomal α-mannosidase (LAM).

Mannostatins A and B, isolated from the soil microorganism *Streptoverticillus*, are the some of the most potent inhibitors of class II α-mannosidases reported [116] thus far, and provide interesting compounds for combinatorial modification to give more selective derivatives (Figure 2.1). They were the first non-azasugar type inhibitors to be discovered that possess an aminocyclopentitol structure. The inhibitors are of the reversible, competitive type and do not show the slow-binding phenomenon exhibited by Swainsonine and its analogs. The synthesis and biological evaluation of a small number of mannostatin analogs has revealed that the basicity of the primary amine and also the neighboring *cis*-diol are essential for inhibitory activity.[83-88, 117] The thiomethyl moiety could, for example, be replaced by a hydroxymethyl group.[117, 118] Most mannostatin derivatives have been tested as inhibitors for jack bean and almond mannosidases and structure-activity relationships for the more relevant Golgi and lysosomal enzymes are scarce.

Here, we report the synthesis of a range of mannostatin analogs, which have varying aromatic substitutions at the amino function. In parallel, a range of aminocyclopentitetrols, which are structurally simpler than mannostatin and have as an additional feature that they are non-chiral, were modified in a similar fashion. It was hoped that the aromatic substitutions of **1a** and **2a** would be able to make favorable interactions with aromatic residues in the binding site of the mannosidase enzymes. These two families of compounds have been tested for their ability to inhibit hGMII and

human LAM (hLAM) and the activities of the two series of compounds were compared. Computational studies have been performed to rationalize the data.

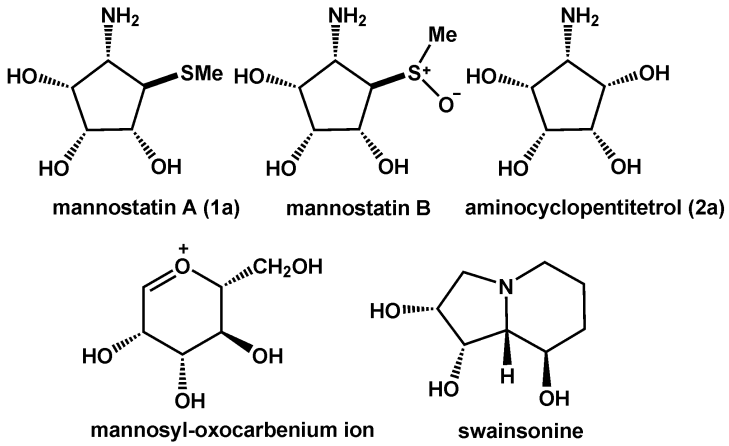



**Figure 2.1.** Aminocyclopentitol inhibitors of Golgi a-mannosidase II.

## 2.3 Synthesis and biological evaluation of mannosidase inhibitors

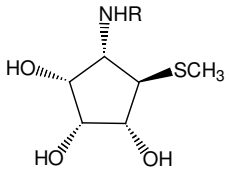
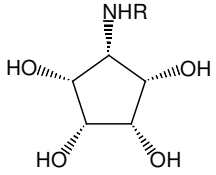
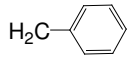
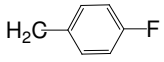
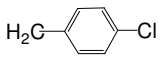
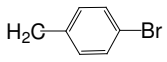
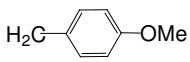
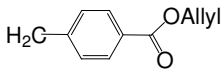
Optically pure mannostatin A (**1a**) and the *meso*-aminocyclopentitetrol **2a** were prepared by a modified literature procedure.[119] A key step of this approach involved an Aldol condensation of nitromethane with a dialdehyde derived from myo-inositol. The physical and spectroscopic data of the two compounds were in agreement with previously reported data. The amino functionalities of **1a** and **2a** were modified by reductive amination using a range of aromatic aldehydes. Conventional procedures involving $Na(CN)BH_3$ in combination with solvents such as Trimethyl Ortho Formate (TMOF)[120] or TMOF/methanol or DCM/methanol in the presence or absence of benzotriazole (BtH)[121] resulted in mixtures of products. Alkylation of the amines with benzyl bromide

in the presence of CsOH[122] gave, apart from the required product, a substantial amount of di-substituted amine. Fortunately, the addition of a methanolic solution of the hydrochloride salts of **1a** and **2a**, pretreated with methanolic sodium hydroxide (0.2 eq), to a slurry of an aromatic aldehyde (2.5 eq), $Na(CN)BH_3$ (1.0 eq), and molecular sieves 3Å followed by stirring for 18 hrs, gave the secondary amines **1b-g** and **2b-g** in reasonable to good yields.[84] The compounds were purified by Iatro bead column chromatography and in each case satisfactory compositional and spectroscopic data were obtained.

Recently, we described recombinant forms of human Golgi α-mannosidase II[123] and human lysosomal α-mannosidase.[73] For these enzymes, the rate of hydrolysis of different concentrations of 4-methylumbelliferone alone and in the presence of different concentrations of inhibitor was measured fluorometrically and $K_i$ values were determined from Dixon plots. As can be seen in Table 2.1, mannostatin A (**1a**) is a markedly better inhibitor than the corresponding aminocyclopentitetrol **2a**. Benzylation of the amino functionality of Mannostatin A gave compounds (**1b-g**) which were slightly less potent inhibitors of the Golgi mannosidase II than the lead compound and *N*-benzyl, *N*-(*p*-fluorobenzyl), *N*-(*p*-chlorobenzyl) and *N*-(*p*-bromobenzyl) and *N*-(*p*-methoxybenzyl) substitution caused similar reductions in inhibitory activity (3-5 fold). Modification of the amino functionality with allyl 4-hydroxymethylenebenzaldehyde (to form **1g**) had a larger impact with a 15-fold loss of activity being observed. Surprisingly, the chemical modifications had no significant effect on HLM indicating that they do not lead to either favorable or unfavorable interactions with the binding site.

The inhibition data for aminocyclopentitetrols (**2a-g**) displayed a different profile than for the mannostatin analogs and for each enzyme, the introduction of a substituted benzyl moiety resulted in an improvement of inhibitory activity. In the case of the HGMII, modification of the amino group of 2a with substituted benzyl moieties led to small improvements in inhibitory activity (5-10 fold) whereas large increases in activity were measured (10-40 fold) with the lysosomal enzyme. Compound **2g**, having a *N*-(allyl 4-hydroxymethylenebenzoic acid) substitution proved to be the best inhibitor tested, displaying a 40-fold more favorable $K_i$ value than parent **2a**. Surprisingly, this substituent caused the largest loss in inhibitory activity for Mannostatin A.

**Table 2.1**. Inhibition of human Golgi Mannosidase II (hGM II) and human Lysosomal Mannosidase (hLAM) by compounds **1a**-**g** and **2a**-**g** ($K_i$ in μM).

| | R |  (**1a-g**) | |  (**2a-g**) | |
|---|---|---|---|---|---|
| | | hGMII $K_i$ | hLAM $K_i$ | hGMII $K_i$ | hLAM $K_i$ |
| a) | H | 0.21 | 0.09 | 50 | 6.6 |
| b) |  | 0.88 | 0.11 | 10 | 0.45 |
| c) |  | 0.53 | 0.17 | 6.0 | 0.33 |
| d) |  | 0.91 | 0.10 | 8.1 | 0.67 |
| e) |  | 0.51 | 0.05 | 7.6 | 0.33 |
| f) |  | 0.52 | 0.10 | 6.6 | 0.48 |
| g) |  | 3.22 | 0.14 | 4.4 | 0.16 |

Despite that the introduction of a substituted benzyl moieties resulted in a less favorable ratios of $K_i$(hGMII) / $K_i$(hLAM), the compounds described here provide important leads for the development of selective inhibitors of hGMII. This Golgi enzyme is able to recognize two potential aglycones during substrate cleavage: $GlcNAcMan_5GlcNAc_2$ to $GlcNAcMan_3GlcNAc_2$. Since the corresponding alternate substrate, $Man_5GlcNAc_2$, is >1000-fold less effective as a substrate, the Golgi enzyme must have an obligatory binding site for recognition of one branch of the aglycone for high affinity substrate interaction. There is no data to support that the broad specific lysosomal α-mannosidase is able to recognize an extended aglycone to any significant degree. Thus, by combinatorial extension of the benzyl derivatives with other chemical functionalities, the promise exits that favorable interactions may be established with the extended binding pocket for the aglycone of hGMII thereby accentuating the selectivity of this enzyme.

### 2.4 Computational studies

The mode of inhibition by an azasugar, such as swainsonine, has been rationalized by their resemblance to the mannosyl oxycarbenium ion; a putative intermediate in the hydrolysis of mannosidases. [124] There has, however, been some debate as to whether this model can be extended to the mode of inhibition of **1a**. It has been contended that in their lowest energy conformations, aminocyclopentitols, such as mannostatin, do not superimpose well onto the hypothetical oxycarbenium ion intermediate.[84, 125] An alternative mode of inhibition has been proposed[84] that is based on the resemblance of 1a to β-mannopyranosylamine.

Recently, crystal structures have been reported for dGMII obtained from *Drosophila melanogaster* (dGMII) in the absence and presence of the inhibitor swainsonine.[75] This enzyme has 41% sequence identity and 61% similarity with hGMII and most importantly, amino acids in catalytic domain are preserved. Furthermore, it has been shown that the two enzymes display similar kinetic and inhibitory properties.[126] Thus, it has been proposed that dGMII is a good model system for the analogous human enzyme.

Molecular docking of **1a** and aminocyclopentitretrol **2a** into the crystal structure of dGMII may provide an opportunity to study the mode of inhibition of these compounds. Furthermore, some of the inhibitory data summarized in Table 2.1 may be rationalized by comparing docking modes of the two compounds with their benzylated counterparts (**1b** and **2b**).

Five-membered ring systems, such as in compounds **1a** and **2a** are inherently flexible due to their ability to assume several twist and envelope conformations, which can interconvert with relative ease *via* pseudo-rotational itineraries.[127, 128] Therefore it is necessary to consider all the possible low energy envelope conformers of **1a** and **2a** for the docking studies.

Conformational properties of isolated five-membered rings such as furanoses have been studied by geometry optimizations of the ten possible envelope conformations, using *ab initio* molecular orbital calculations.[129-131] Thus, the conformational properties of **1a** and **2a** were studied by optimizing the ten possible envelope conformers by constraining a specific endocyclic torsion angle to 0° and allowing all other parameters to be optimized at the B3LYP/6-31G* level. The hydroxyl

and amino groups of the resulting structures were placed in each of the three staggered orientations, which were re-optimized without any restraints at the same level of theory. The effects of solvation were approximated by single point calculations of gas phase and solution phase energies using the Poisson-Boltzmann treatment[132, 133] in the Jaguar program [111] (Table 2.2). It was observed that for both compounds, many hydroxyl rotamers converged to a single structure. In the case of **1a**, thirty initial structures led to twenty unique structures, which represented only six envelope conformations ($E_1$, $^1E$, $E_2$, $E_3$, $^3E$ and $^4E$). In the case of **2a**, fifteen unique structures were obtained, which resembled nine envelope conformations ($E_1$, $^1E$, $E_2$, $^2E$, $E_3$, $^3E$, $E_4$, $^4E$ and $E_5$). The fact that **2a** can adopt a larger number of envelope conformations indicates that its ring structure is more flexible than that of **1a**. The difference in ring flexibility is likely due to the differences in ring substitutions, which are known to determine pseudorotational itineraries of five membered rings.[128]

Each of the low energy conformers of **1a** and **2a** was subsequently docked into the binding site of dGMII. The six envelope conformers of **1a** could be complexed in three different binding modes (Figure 2.2), with conformer $^1E$ exhibiting the most favorable binding energy ($\Delta G$ = –9.1 kcal $mol^{-1}$). In this binding mode the five membered ring of **1a** stacks against the aromatic side chain of Trp95, a type of interaction seen in many carbohydrate-protein complexes.[134] The C-4 and C-5 hydroxyl oxygens of 1a coordinate with the zinc ion in the active site resulting in $T_6$-octahedral coordination geometry. Furthermore, the nitrogen of the exocyclic amino group of **1a** forms hydrogen bonds with the carboxylate oxygen of the putative general acid base residue Asp341 and with the hydroxyl oxygen of Tyr269. In addition, the thiomethyl moiety of **1a** makes

favorable hydrophobic interactions with Arg228 and Tyr727. In the crystal structure of dGMII with swainsonine[75], the $Zn^{+2}$ ion has $T_6$-octahedral coordination geometry complexing the hydroxyl oxygens O4 and O5. However, the bridgehead nitrogen of Swainsonine forms a hydrogen bond with the catalytic nucleophile Asp204. Thus, the two complexes may differ in the interactions of their amines with acidic residues in the binding site.

**Table 2.2**. Relative gas phase and solution phase energies (in kcal $mol^{-1}$) for envelope conformers of Mannostain A (**1a**) and aminocyclopentitetrol (**2a**). Energies for only those conformers are given which were docked in the active site of dGMII.

| Low energy conformers of **1a** | Relative energy (kcal $mol^{-1}$) | | Low energy conformers of **2a** | Relative energy (kcal $mol^{-1}$) | |
|---|---|---|---|---|---|
| | Gas phase | Solution phase | | Gas phase | Solution phase |
| $E_1$ | 0.0 | 0.0 | $E_1$ | 0.0 | 1.8 |
| $^1E$ | 2.8 | 1.5 | $^1E$ | 1.2 | 3.1 |
| $E_2$ | 8.6 | 4.6 | $E_2$ | 6.4 | 3.4 |
| $E_3$ | 5.2 | 1.7 | $^2E$ | 1.6 | 3.5 |
| $^3E$ | 8.9 | 2.8 | $E_3$[a] | 6.4 | 3.4 |
| $^4E$ | 5.4 | 0.5 | $^3E$[a] | 1.6 | 3.5 |
| | | | $E_4$[a] | 0.0 | 1.8 |
| | | | $^4E$[a] | 1.2 | 3.1 |
| | | | $E_5$ | 0.3 | 0.0 |

[a]Conformers $E_3$, $^3E$, $E_4$ and $^4E$ are mirror images of conformers $E_2$, $^2E$, $E_3$ and $^3E$ respectively with same relative energies. These conformers were also docked in the binding site of dGMII.

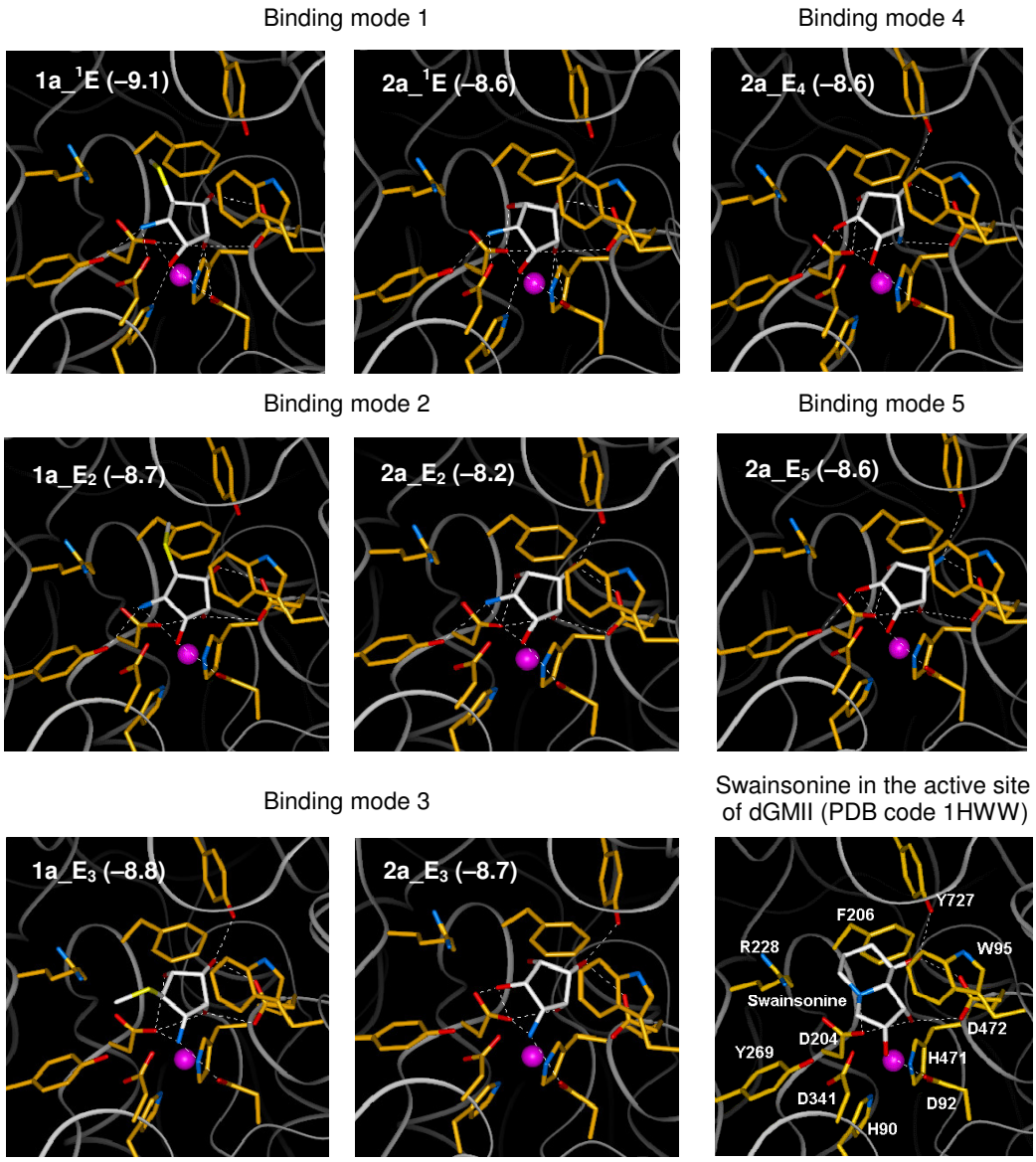



**Figure 2.2.** Different binding modes observed for conformers of **1a** and **2a**. White dashed lines indicate h-bond interactions and free energies of binding are given in the parentheses (in kcal $mol^{-1}$).

Conformers $E_2$ and $^3E$ of **1a**, which were predicted to bind with only slightly smaller computed binding energies than $^3E$ (–8.7 kcal $mol^{-1}$ and –8.8 kcal $mol^{-1}$, respectively), displayed a binding mode very similar to that of swainsonine. In this case, the exocyclic amine of **1a** formed a hydrogen bond with the catalytic nucleophile Asp204. Furthermore, The C-4 and C5-hydroxyl oxygens of **1a** coordinated with the zinc

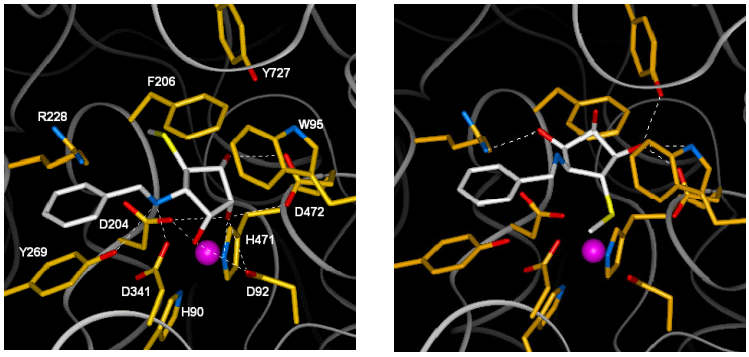
ion in $T_6$-octahedral coordination geometry. The somewhat lower computed binding energies may be due to loss of hydrogen bonding interactions of the C5-hydroxyl with amino hydrogens of His90 and His471, but may also reflect limitations in the computational method.

Conformers $E_1$, $E_3$ and $^4E$ displayed yet another complexation mode of **1a**. In this case, the amine of **1a** coordinated with zinc ion leading to $T_5$ coordination geometry. Pentavalent coordination is energetically more favorable than the hexavalent ($T_6$) coordination of zinc[78] and therefore may be an important contributor to the stabilization of these complexes. Further stabilization came from hydrogen bonds between the nitrogen of the amine with carboxylic oxygen of the catalytic nucleophile Asp204. The interactions between the thiomethyl group and Trp727 and Arg228 were lost in this binding mode, which may account for the slightly lower binding energies.

It has been suggested that the coordination of O4 and O5 of swainsonine with $Zn^{+2}$ mimics a similar chelatation of the C-2 and C-3 hydroxyls of mannosyl oxycarbenium ion. This coordination is an important determinant of the α-mannoside specificity of the enzyme. The binding of swainsonine is stabilized by an important hydrogen bond between N4 and the catalytic nucleophilic residue Asp204. In the case of the mannosyl oxycarbenium ion, its flattened ring structure may places the anomeric center in close proximity to nucleophile Asp204 for favorable interactions. Furthermore, Swainsonine, and probably the oxycarbenium ion, can make van der Waals stacking interactions with Trp95 and Phe206. The conformers $E_2$ and $^3E$ of **1a** displayed a very similar binding mode, in which the exocyclic amino group formed a hydrogen bond with Asp204 and the C-4 and C-5 hydroxyls coordinated with the $Zn^{+2}$ ion. However, due to

the flexibility of its five membered ring, mannostatin can bind in other low energy modes. For example, the small change in ring conformation of $^{1}E$ places the exocyclic amine in a position to form hydrogen bonds with Asp341 and Tyr269. Other ring conformers allowed the amine of **1a** to coordinate with the $Zn^{+2}$ ion in a favorable $T_5$ coordination geometry.

In addition to the three binding modes observed for **1a**, aminocyclopentitretrol **2a** could complex in two additional ways (Figure 2.2). All the five binding modes displayed very similar computed binding energies. Conformers $^{3}E$ and $E_4$ of **2a** complexed in a unique mode, in which the amine and C2- hydroxyl coordinate with the zinc ion resulting in $T_6$ coordination geometry. The ring stacked against Trp95 and the amine formed hydrogen bonds with the catalytic nucleophile Asp204, Asp472 and Asp92. In the case of $E_5$, the amine formed hydrogen bonds with hydroxyls of Tyr727 and Asp472. This binding mode is further stabilized by coordination of C-2 and C-3 hydroxyls with the Zn ion. All the hydroxyls formed hydrogen bonds with Asp204 in addition to hydrogen bond interaction with Tyr269, Asp92 and Asp472.




1b_$^{2}E$ ($\Delta G_{binding}$ = –10.1) 1b_$^{4}E$ ($\Delta G_{binding}$ = –8.9)

**Figure 2.3.** Binding modes observed for conformers of **1b**.

Mannostatin A is a significantly more potent inhibitor of GMII than aminocyclopentitretrol **2a** (Table 2.1). Overall, the computational studies show that **2a** is more flexible allowing additional binding modes. Furthermore, due to inversion of configuration at C-2 and replacement of the thiomethyl moiety with a hydroxyl, the hydrophobic interactions of thiomethyl moiety of **1a** cannot be made by **2a**. Although that the C-2 hydroxyl of **2a** can make hydrogen bonds with Asp204 and Asp341, it is unlikely that these interactions can compensate for the loss of hydrophobic interactions.[135, 136]

Modification of the amine of mannostatin A with benzylic moieties did not improve the inhibitory potential of the resulting compounds (Table 2.1). In order to rationalize these observations, docking studies were also performed with compound **1b**. Thus, benzyl moieties were attached to the low energy conformations of **1a**. The resulting derivatives were docked in the binding site of dGMII to give three different binding modes. In case of conformer $^{1}E$, typical interactions were observed, such as a hydrogen bond between the amine and general acid-base residue Asp341 and coordination of the C-4 and C-5 hydroxyls with the zinc ion. Furthermore, the phenyl ring of **1b** made van der Waal's interactions with the aromatic moiety of Tyr269. The orientation of both the phenyl rings was observed as similar to the preferred off-centered parallel-displaced arrangement with average $R_{cen}$ = 4.4 Å, $\gamma$ = 19.9° and $\theta$ = 41.5° as reported by McGaughey et al.[137] The latter $\pi$-$\pi$ stacking required, however, a tilting of the ring structure of **1b** disrupting a hydrogen bond between the C-5 hydroxyl and Nε2 nitrogen of His90, which is observed in the complex with the parent compound. The phenyl ring in conformers $E_2$, $E_3$ and $^{4}E$ of **1b** made similar stacking interactions with Tyr269. For all

three conformers, the amine formed a hydrogen bond with the catalytic nucleophile Asp204 and hydroxyl oxygen of Tyr269. As in $^{1}E$ the stacking interactions required a tilting of the five membered ring of **1b**, leading to loss of hydrogen bonds. Thus, the computational studies indicate that the favorable interactions made between the phenyl ring of 1b and Tyr269 may be offset by loss of hydrogen bonds due to a tilting of the five membered ring of the inhibitor.

Conformers $E_1$ and $^{3}E$ of **1b** display a very different binding mode than similar conformer of parent compound **1a**. Thus, in these cases the orientation of the inhibitor is flipped in the catalytic site allowing the thiomethyl to make hydrophobic interactions with Trp95. Nevertheless, stacking interactions of the phenyl ring of **1b** with Tyr269, and hydrogen bonds with Arg228 and Tyr727, further stabilizes these complexes.

Docking of the conformers of **2b** in the binding site of dGMII resulted in two binding modes, which mainly differ in the interactions of the amine with either Asp204 ($E_1$, $E_2$, $E_3$, $^{3}E$ and $E_5$) or Asp341 ($^{1}E$, $^{2}E$, $E_4$ and $^{4}E$). A stacking interaction of the phenyl ring of 2b with Tyr269 (average $R_{cen}$ = 4.3 Å, $\gamma = 13.0°$ and $\theta = 41.5°$) plays an important role for positioning the inhibitor in the binding site of the enzyme. In the case of conformers $E_3$, $^{3}E$, $E_4$, $^{4}E$ and $E_5$, the ring structure of **2b** was tilted to allow the stacking interactions. As a result, the hydrogen bond patterns observed in **2a** were altered. However, conformers $E_1$, $^{1}E$, $E_2$ and $^{2}E$ did not require a tilt of the five membered ring and therefore no loss of hydrogen bonds was observed. In this case, the interactions made by the phenyl moiety might be responsible for the improved affinity of compounds **2b-g** for GMII.

Although, the molecular docking analysis predicted that each of the inhibitors had a significant affinity for the enzyme, the method could not distinguish between stronger and weaker binding. The benzylated analogs would be expected to pay a proportionally larger entropic penalty upon binding than that associated with the simpler, more rigid inhibitors. Autodock does not explicitly consider conformational entropy and so would be expected to overestimate the strength of binding of compounds **1b** and **2b**. While a quantitative ranking of affinities was not achieved, much information could be gained from inspection of the predicted binding modes.

## 2.5 Conclusions

A range of mannostatin analogs and structurally simpler aminocyclopentitetrols has been synthesized, which have varying benzylic moieties at their amino function. It was observed that mannostatin A (**1a**) is a markedly better inhibitor than the corresponding aminocyclopentitetrol **2a**. Furthermore, the substitutions of the aminocyclopentitetrols led to significant improvements of inhibitory activity for both hGMII and hLAM whereas similar substituents of mannostatin A had only a marginal effect. Computational studies have been performed to provide a rational for these observations. First, the conformational properties of mannostatin and aminocyclopentitetrol were studied by geometry optimizations of the ten possible envelope conformations, using *ab initio* molecular orbital calculations. Subsequently, the low energy conformers of each compound were docked in the binding site of dGMII and the results compared with interactions of swainsonine observed in a crystal structure with the same enzyme. It has been shown that mannostatin A and aminocyclopentitretrol could

bind to GMII in a similar mode as swainsonine. Probably, the latter compounds mimic the mannosyl oxycarbenium ion, which is a putative intermediate in the hydrolysis of mannosidases. Thus, it appears that mannostatin and aminocyclopentitetrol can inhibit GMII in a similar fashion. However, due to the flexibility of the five membered rings of **1a** and **2a**, additional low energy binding modes could be adopted. It is conceivable that the ring structure of **1a** and **2a** is flexible in the binding site of GMII allowing transitions between different binding modes. The thiomethyl moiety of mannostatin could make favorable hydrophobic interactions with Arg228 and Tyr727. These interactions are not present in complexes with aminocyclopentitretrol (**2a**) providing a rationale for the lower inhibitory potential of this compound. It was also observed that the five membered ring of **2a** is significantly more flexible than that of 1a allowing additional binding modes. Attachment of benzyl moieties to mannostatin let to stacking interactions with the aromatic moiety of Tyr269. These π-π interactions required, however, a tilting of the ring structure of 1b resulting in a disruption of the hydrogen-bonding network observed. Due to the greater flexibility of aminocyclopentitretrol (**2a**), attachment of the benzyl moiety led to conformers that could make the stacking interactions without disrupting hydrogen bonds. This observation may provide a rational for the improved activities of benzyl-substituted aminocyclopentitretrol.

It is to be expected that further combinatorial extension of the benzyl derivatives will lead to compounds that can interact with the extended binding pocket for the aglycone of HGMII thereby accentuating the selectivity of this enzyme. In this respect, the different binding modes computed for both compounds may provide an important opportunity for the design of targeted libraries.

## 2.6 Experimental Section

### 2.6.1 General procedure for reductive amination

A solution of NaOH in MeOH (0.1M, 0.2eq) was added under Argon at room temperature to a solution of compound 1a or 2a in $CH_3OH$ to give a solution having pH 7. To this solution was added a methanolic solution of benzaldehyde (2.5M, 1.0eq) and 3Å MS. A methanolic solution of $NaCNBH_3$ (1.75M, 0.6eq) was added, and after 20hrs, the reaction was diluted with $CH_3OH$ and filtered through Celite. The solution was acidified to pH=1 with 1N HCl and concentrated *in vacuo*. The residue was purified by flash chromatography (Iatrobeads; eluent: $CH_3CN$:HOAc:$H_2O$, 10:0.5:1 v/v to give the expected product.

### 2.6.2 Inhibition studies

Assays of the hLAM were assembled at 4°C in a reaction volume (50μL) containing 4-methylumbelliferyl α-D-mannopyranoside (SIGMA)(3 mM), sodium acetate (100 mM), pH 4.5, and purified recombinant hLAM expressed and purified from HEK293 cells (methods to be published separately). Assays for the hGMII were prepared in a similar reaction volume containing 4-methylumbelliferyl α-D-mannopyranoside (SIGMA)(2.5 mM), sodium acetate (83.3 mM) at pH 5.6, ZnCl (83.3 μM) and purified recombinant HGMII expressed and purified from HEK293 cells (methods to be published separately). Reactions were incubated for one hour at 37°C and stopped with the addition of sodium carbonate (150 μL) to a final concentration of 150 mM. Fluorescence was quantitated on a Spectramax Gemini XS fluorescence reader. All fluorescence values are compared to a 4-methylumbelliferone standard curve. Inhibitors were prepared as stock

solutions in water (1-10 mM depending on the amount of compound) and $IC_{50}$ data were obtained in triplicate over a concentration range of 10 nM – 1 mM and plotted as percent inhibition versus inhibitor concentration. Ki determinations were performed under similar conditions as $IC_{50}$ determinations, with the exception that 4-methylumbelliferyl α-D-mannopyranoside concentrations were also varied from 160 μM - 3 mM. Dixon plots were used to transform the kinetic data into Ki values.

### 2.6.3 Computational studies

All the geometry optimizations were performed with Gaussian 94 program [138] using density functional theory (B3LYP[139-141]) and 6-31G*[142] basis set. The low energy unique conformers of **1a** and **2a** were determined by optimizing ten possible envelope conformers ($^1E$, $E_1$, $^2E$, $E_2$, $^3E$, $E_3$, $^4E$, $E_4$, $^5E$ and $E_5$, the abbreviated nomenclature used here is similar to the one suggested for furanose ring system [143]), by constraining a specific torsion angle to 0° and allowing all other parameters to optimize at BLYP/6-31G* level of theory. The hydroxyl and amino groups of the resulting structures were placed in three different orientations to give different rotamers which were reoptimized without any restraints at B3LYP/6-31G* level. The solvation effects on the energies of the resulting conformers were approximated using the Poisson-Boltzmann treatment[132, 133] provided in the Jaguar program.[111] The unique envelope conformers obtained were docked in the binding site of dGMII.

*Preparation of ligand and receptor molecules for docking* – The crystal structure of drosophila GMII complexed with Swainsonine (PDBid: 1HWW) was used as a model for macromolecule in docking experiments. The protein target, dGMII and the ligands were prepared for docking using AutoDock Version 3.0.5.[90, 144] Charges were assigned using

the Kollman algorithm.[145] Atomic solvation parameters and fragmental volumes were determined using Addsol programs available in AutoDock. AutoTors was used to define torsional angles in the ligand. Polar hydrogen charges of Gasteiger type [146] were assigned and the non-polar hydrogens were merged with the carbons. The macromolecule was kept rigid in all the docking simulations.

*Docking simulations* – Grid maps for docking simulations were generated with 40 grid points (with 0.375 Å spacing) in x, y and z direction, by Autogrid program. The center of the grid was positioned at the $Zn^{+2}$ ion (Zn1102). Lennard-Jones parameters 12-10 and 12-6 (supplied with the program package) were used for modeling H-bonds and van der Waals interactions, respectively. The distance dependant dielectric permittivity of Mehler and Solmajer[147] were used for the calculations of the electrostatic grid maps. The genetic algorithm (GA) and Lamarckian genetic algorithm with the pseudo-Solis and Wets modification (LGA/pSW) methods were used with the default parameters. For all the simulations, the populations in the genetic algorithm were 50 and each simulation comprised of $2.5\times10^5$ energy evaluations. Each docking experiment consisted of a series of 100 simulations.

The $\pi$-$\pi$ stacking interactions between the phenyl ring in 1b and 2b conformers and the phenyl ring in Tyr269 were evaluated by determining the average distance between the centroids of the two phenyl rings ($R_{cen}$), the angle between the planes of the two rings ($\gamma$) and the normal-centroid angle ($\theta$) as discussed in reference 44.

## 2.7 Acknowledgements

The authors are grateful for financial support for this work by NCI of NIH (5UO1CA91295).

# CHAPTER 3

# THE STRUCTURAL BASIS OF THE INHIBITION OF GOLGI α-MANNOSIDASE II BY MANNOSTATIN A AND THE ROLE OF THE THIOMETHYL MOIETY IN LIGAND-PROTEIN INTERACTIONS[a]

[a] Kawatkar, S. P.; Kuntz, D. A.; Woods, R. J.; Rose, D. R.; Boons, G. J., *J. Am. Chem. Soc* .2006, 128, 8310-8319.

Reprinted here with permission of the publisher.

## 3.1 Abstract

The X-ray crystal structures of mannose trimming enzyme *Drosophila* Golgi α–mannosidase II (dGMII) complexed with the inhibitors mannostatin A (**1**) and an *N*-benzyl analog (**2**) have been determined. Molecular dynamics simulations and NMR studies have shown that the five-membered ring of mannostatin A is rather flexible occupying pseudo-rotational itineraries between $^{2}T_{3}$ and $^{5}E$, and $^{2}T_{3}$ and $^{4}E$. In the bound state, mannostatin A adopts a $^{2}T_{1}$ twist envelope conformation, which is not significantly populated in solution. Possible conformations of the mannosyl oxacarbenium ion and an enzyme-linked intermediate have been compared to the conformation of mannostatin A in the co-crystal structure with dGMII. It has been found that mannostatin A best mimics the covalent linked mannosyl intermediate, which adopts a $^{1}S_{5}$ skew boat conformation. The thiomethyl group, which is critical for high affinity, superimposes with the C-6 hydroxyl of the covalent linked intermediate. This functionality is able to make a number of additional polar and non-polar interactions increasing the affinity for dGMII. Furthermore, the X-ray structures show that the environment surrounding the thiomethyl group of **1** is remarkably similar to the arrangements around the methionine residues in the protein. Collectively, our studies contradict the long held view that potent inhibitors of glycosidases mimic an oxacarbenium ion like transition state.

## 3.2 Introduction

Glycosidases are enzymes that catalyze the cleavage of glycosidic bonds and play critical roles in a number of biological processes. Inhibitors of these enzymes have garnered much attention and are currently used for the treatment of diabetes, viral

infections and Gauchers disease.[6-10] Furthermore, inhibition of the mannose trimming enzyme human Golgi α–mannosidase II (GMII; mannosyl-oligosaccharide 1,3-1,6-α–mannosidase II; E.C. 3.2.1.114), which acts late in the *N*-glycan processing pathway, provides a route to blocking the oncogene-induced changes in cell surface oligosaccharide structures.[61, 63, 113, 148] GMII selectively cleaves α(1-3) and α(1-6) mannosyl residues present in its natural substrate $GlcNAcMan_5GlcNAc_2$.[149] It is a retaining Family 38 glycosylhydrolase, which employs a two-stage mechanism involving two carboxylic acids positioned within the active site which act in concert: one as a catalytic nucleophile and the other as a general acid/base catalyst.[12, 14, 19, 114, 115] Protonation of the exocyclic glycosyl oxygen of a substrate molecule leads to bond-breaking and simultaneous attack of the catalytic nucleophile to form a glycosyl enzyme intermediate. Subsequent hydrolysis of the covalent intermediate by a nucleophilic water molecule gives an α-mannose product with overall retention of configuration. Studies with 5-fluoro pseudo-substrates and deuterium labeled substrates have shown that the transition states on either side of the covalent intermediate have marked oxacarbenium ion character.[12, 14, 19, 38, 114, 115] Furthermore, X-ray crystal structures of the wild type and a mutant Drosophila melanogaster GMII (dGMII) in which the acid/base catalyst has been removed, with fluorinated sugar analogs have revealed that the glycosyl enzyme intermediate adopts a distorted $^1S_5$ skew boat conformation.[38] In this conformation, the leaving group is placed anti-periplanar to the lone pair of the ring oxygen, a requirement for the departure of the leaving group according to Deslongchamp's anti-periplanar lone-pair hypothesis.[79] Furthermore, steric clashes between the *syn*-hydrogens at C-3 and C-5 and the attacking water are minimal in the $^1S_5$ skew boat conformation.

Potent inhibitors of glycosidases are believed to mimic oxacarbenium ion-like transition states.[12, 14, 19, 114, 115] The inhibitory activity of the natural product swainsonine has long been attributed to the resemblance of its 5-membered ring to a flattened six-membered ring mannosyl oxacarbenium ion. In fact, the crystal structure of swainsonine complexed with dGMII shows the inhibitor to be tilted in such a way as to bring the equivalent of its anomeric carbon close to the presumed catalytic nucleophile.[75]

Mannostatins A and B, which were isolated from the soil microorganism *Streptoverticillus*, are some of the most potent inhibitors of class II α-mannosidases reported thus far (Figure 3.1).[116] They were the first non-azasugar type inhibitors to be discovered that possess an aminocyclopentitol structure. The inhibitors are of the reversible, competitive type and do not show the slow-binding phenomenon exhibited by swainsonine and its analogs. Mannostatin A effectively blocked the processing of influenza viral hemagglutin in cultured MDCK cells and caused the accumulation of hybrid type protein linked oligosaccharides, which is in agreement with blocking Golgi mannosidase II.[148] The synthesis and biological evaluation of a small number of mannostatin analogs has revealed that the basicity of the primary amine and the neighboring *cis*-diol are essential for inhibitory activity. [83-88, 117]

While interest in mannostatin A continues to grow, the mode of inhibition of this compound is not well understood. In this respect, its carbocyclic structure represents a significant departure from common alkaloid-based glycosidase inhibitors, and it is unclear whether mannostatin resembles either D-mannose or the mannosyl oxacarbenium ion.

We have now determined the X-ray crystal structure of dGMII in complex with

mannostatin A (**1**) and an N-benzyl analog (**2**). We have analyzed in detail key molecular interactions, in particular, the interaction of the thiomethyl ether with the protein to determine the structural basis for inhibition. Furthermore, molecular dynamics simulations of both the unbound mannostatin A and its complex with the enzyme have been performed to obtain information about the solution phase conformational properties and dynamic features of the enzyme-inhibitor interactions. Finally, possible conformations of the mannosyl oxacarbenium ion and the covalent-linked saccharide intermediate have been compared to the conformation of mannostatin A observed in the co-crystal structure.

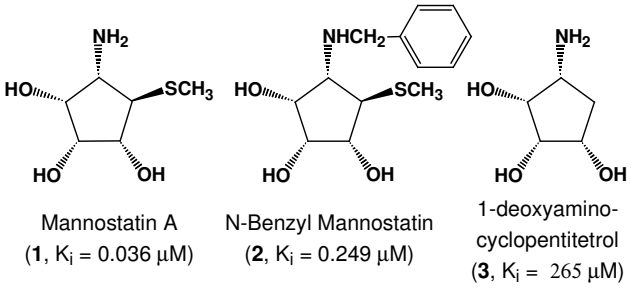



**Figure 3.1**. Cyclopentitol inhibitors of Golgi α-Mannosidase II. The inhibition constants in the parentheses indicate that the compound **3**, which lacks the thiomethyl moiety, is much less potent compared with compounds **1** and **2**.

### 3.3 Results and Discussion

Optically pure mannostatin A (**1**), *N*-benzyl mannostatin A (**2**) and 1-deoxyaminocyclopentitetrol (**3**) were prepared by literature procedures.[84, 87, 119] The rate of hydrolysis of different concentrations of 4-methylumbelliferyl α-D-mannopyranoside alone and in the presence of different concentrations of inhibitor was measured fluorometrically and $K_i$ values were determined from Dixon plots. As can be seen in Figure 3.1, mannostatin A is a more potent inhibitor then its *N*-benzylated counterpart.

On the other hand, 1-deoxyaminocyclopentitetrol **3**, which lacks the thiomethyl ether, is a poor inhibitor indicating that the thiomethyl moiety of **1** makes important interactions with the binding site of the enzyme.

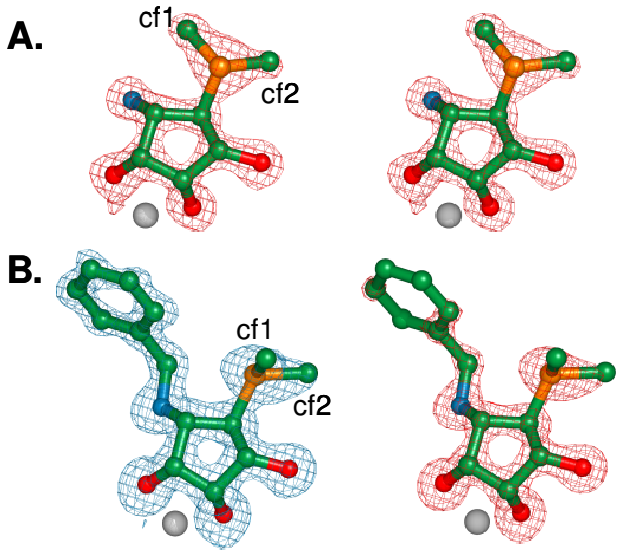
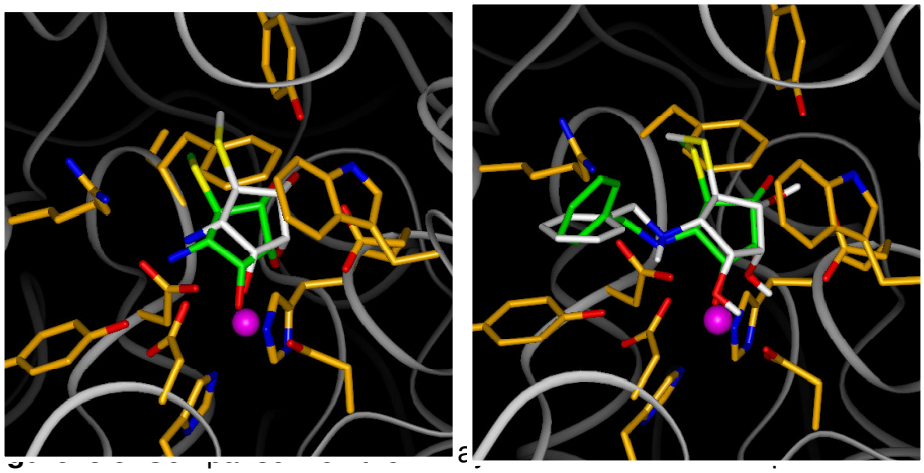



Figure 3.2. Stereoviews of electron density of bound inhibitors: Simulated annealing omit maps (Fo-Fc). A. Mannostatin A 1, contoured at 5 sigma B. N-benzyl mannostatin 2 contoured at 3 sigma (blue) or 5 sigma (red). This figure was created with Pymol. cf1 and cf2 refer to the two possible conformations of C7 which fit into the observed density.

In order to rationalize the inhibition data, X-ray crystal structures of compounds **1** and **2** complexed with dGMII were solved at 1.3 Å. Data collection and refinement statistics are reported in supplemental data Table 1S. The quality of the density around the bound inhibitors is shown in Figure 3.2. The X-ray crystal structure of dGMII complexed with compound **1** did not show any significant changes in the backbone structure compared to the free enzyme (RMSD = 0.129 Å *vs.* PDB 1HTY). The five-membered ring of the inhibitor adopts a $^{2}T_{1}$ twist envelope conformation (phase angle, $\phi$ = 201°), which is stacked against the aromatic ring of Trp95, a type of interaction seen in

many carbohydrate-protein complexes.[150-152] Important close contacts between the inhibitors and dGMII are detailed in Table 1.

Gratifyingly a good agreement was obtained between the X-ray structures and the predicted docking mode from the docking studies (Figure 3.3).[153] There are a number of contacts, which have been observed in previous dGMII crystal structures.[38, 75, 80, 82] The 3,4 *cis*-diol of **1** is complexed with a $Zn^{2+}$ ion in the active site of dGMII resulting in $T_6$ coordination geometry, while the 2- hydroxyl forms hydrogen bonds with Asp472 and Tyr727. Furthermore, the amine of **1** forms hydrogen bonds with catalytic acid residues Asp204 and Asp341, and Tyr269. The binding of the inhibitors within the context of selected active site residues is shown in Figure 3.4.



ocking mode for mannostatin A and its N-benzylated analogue (green: X-ray structure, white: predicted binding mode). A good agrrement was obtained for the binding mode of the five membered ring. However, in the case of benzylated analogue, the phenyl ring assumes different orientation than the one predicted by molecular docking experiments.

Data from SAR experiments (referred to above) has pointed to the importance of the amine and cis-diols in the inhibitory activity of mannostain A and the crystal structure beautifully illustrates how these groups interact with the protein. Similar modes of

interaction, with almost identical distances, were seen in the crystal structure of dGMII complexed with swainsonine (Table 3.1) although only a single interaction with the amine group is observed in that case.

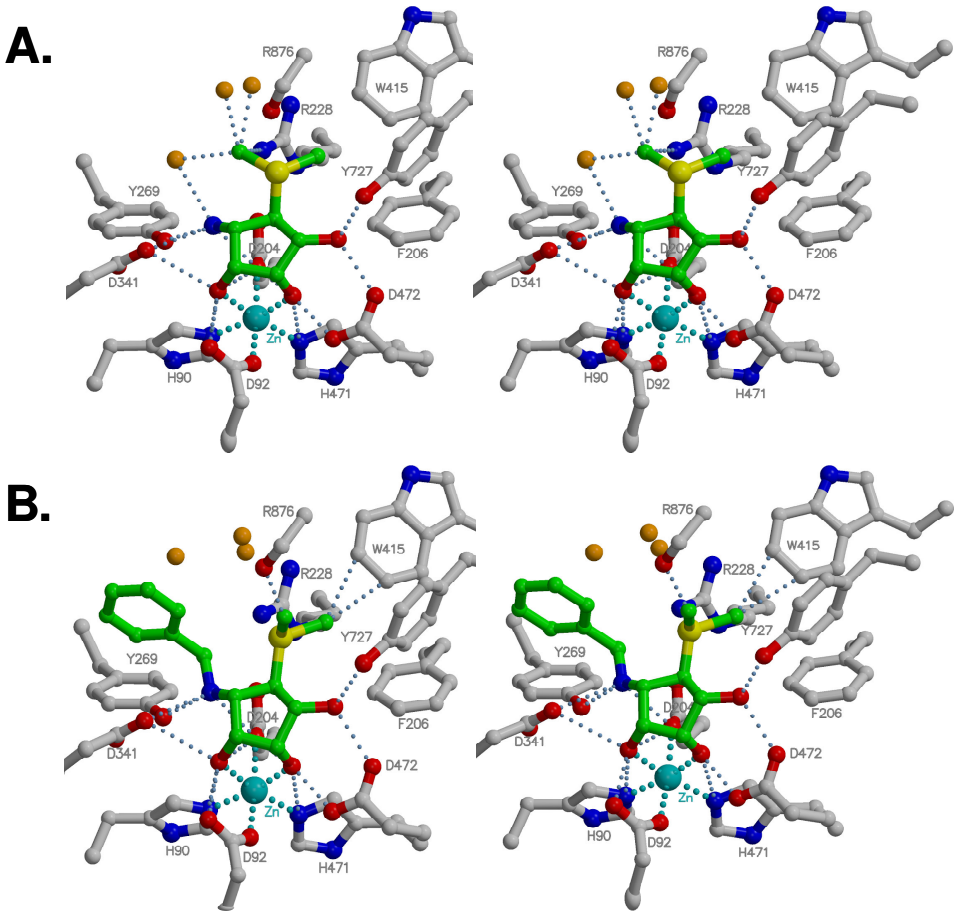



**Figure 3.4.** Stereoview of the interaction of Mannostatin A (**1**, A) and N-Benzyl Mannostatin (**2**, B) with residues in the active site of dGMII. Interactions closer than 3.2 Å are indicated. The interactions with zinc are indicated in cyan. Water molecules appear as orange balls. Distances are presented in Table 1.

The thiomethyl moiety of **1** and **2** is a feature that is not observed in any other glycosidase inhibitors and the data summarized in Figure 3.1 indicate that this structural feature is critical for potent inhibition. It must be noted that the thiomethyl moiety is structurally similar to the side chain of a methionine residue. It has been

proposed that the sulfur atom and ε–$CH_3$ group of methionine residues are involved in several different interactions important for protein stability.[154-158] For example, aryl-sulfur interactions have been proposed favorable because of the observed proximity of the methionine residue to the aromatic side chains in protein X-ray structures and small molecules X-ray structures. In general, these interactions are either hydrophobic or electrostatic, of the types S–$CH_3$---Ar or S---H–Ar, respectively. In addition, the sulfur atom, possessing an empty *d*-orbital is also known to be involved in S---X (X = O and N) interactions.[158] Such an interaction is stabilized by the orbital interaction between the lone pair on O or N and the anti-bonding orbital of the sulfur atom. The strong dipole of the methionine residue also makes it an ideal hydrogen bond acceptor. However, an analysis of the NH---S and OH---S hydrogen bonds in proteins suggests that the sulfur atom in methionine has only a weak character of a hydrogen bond acceptor.[156, 159]

In the context of dGMII at least one third of the methionines occur in a hydrophilic or surface environment, and a number bear a striking resemblance to the environment of the thiomethyl group of the inhibitors. In Figure 3.5, three different methionine residues whose methyl carbon occupies environments similar to **1** are presented, along with their respective distances from water (or a backbone carbonyl in the case of Met769). In two cases (Met264 and Met769) alternative conformations of the thiomethyl group are seen, similar to the situation observed with the inhibitors, while Met224 shows a single strong conformation that places the methyl group in close proximity to waters. In Met264 both confomers are in a hydrophilic environment. In Met769, one orientation places the methyl group in close proximity to 2 waters and the backbone oxygen of Pro684, while the other orientation results in the terminal methyl

group finding itself in a more hydrophobic environment, similar to what is observed for **1.** It should be noted that all the waters that are in close proximity to methionine have other strong hydrogen bond partners and the methionine residues do not appear to be their principal form of interaction with the protein.

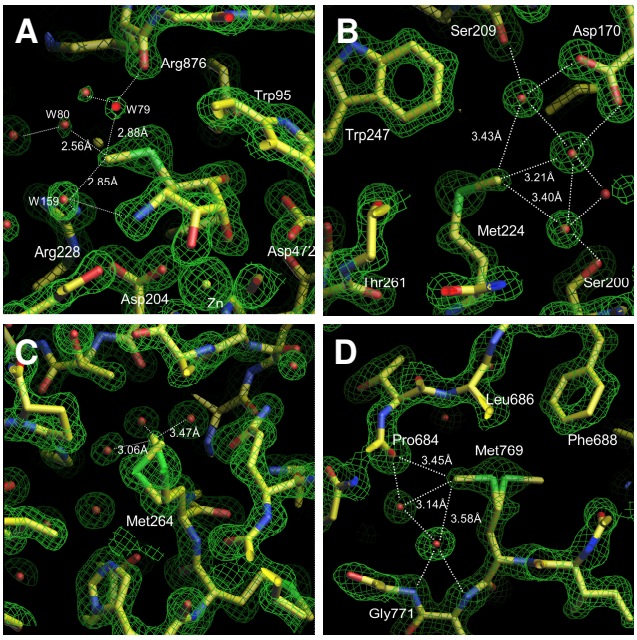



**Figure 3.5.** Comparison of the density of Mannostatin A cf1 with methionine residues of dGMII found in similar environments. Possible water interactions are shown with dotted lines. Distances between the thiomethyl group and proximal waters are indicated. The electron density (2Fo-Fc) is contoured at 1.5 -2 sigma. The figures are generated in O. **A.** Mannostatin A cf1 **B.** Methionine 224 **C.** Methionine264, showing both conformers **D.** Methionine769, showing both conformers.

The sulfur atom of the thiomethyl moiety of **1** seems to form a nonbonded S---O type of interaction with the main chain carbonyl oxygen of Arg876. Statistical analysis of protein structures performed by Iwaoka et al[158] suggests that most S---O interactions are between the sulfur atom in the methionine residue and backbone carbonyl oxygens or side chain carboxylate oxygens. An analysis of the spatial placement of sulfur with

respect to the oxygen revealed that the interaction between the π-orbital on the carbonyl oxygen and the antibonding σ* orbital of S-C bond is largely responsible for this observation. *Ab initio* calculations with small molecules indicate that these interactions could result in stabilization of up to 2.5 kcal $mol^{-1}$ when these atoms are separated by 3.3 Å.[158] It is, however, important to note that these calculations have been performed in vacuo and therefore the binding energies may be different in an aqueous environment due to solvation effects. In the **1**-dGMII complex the sulfur atom is located at 3.49 Å from the Arg876O, suggesting that the interaction observed between the sulfur atom in the thiomethyl moiety and the carbonyl oxygen of Arg876 must be reasonably strong.

**Table 3.1.** Distances between different inhibitor atoms and protein atoms as observed in the complexes of swainsonine (PDB id 1HWW), **1** and **2** with dGMII.

| | Swainsonine | | | 1 | 2 |
|---|---|---|---|---|---|
| **Protein Atom** | **Inhibitor Atom** | **Distance (Å)** | **Inhibitor Atom** | **Distance (Å)** | **Distance (Å)** |
| H90 NE2 | O-2 | 2.97 | O-4 | 3.07 | 3.11 |
| D92 OD1 | O-1 | 3.04 | O-3 | 2.92 | 2.99 |
| D92 OD2 | O-2 | 2.54 | O-4 | 2.55 | 2.51 |
| | O-1 | 2.83 | O-4 | 2.82 | 2.91 |
| D204 OD2 | O-2 | 2.97 | O-3 | 2.92 | 2.95 |
| | N-4 | 2.75 | N | 2.77 | 2.9 |
| R228 NH2 | | | C-7 (cf1) | 3.11 | 5.08 |
| | | | S | 4.01 | 3.54 |
| Y269 OH | N-4 | 4.17 | N | 2.78 | 2.96 |
| D341 OD2 | | | N | 2.99 | 3 |
| W415 CH2 | | | C-7 (cf2) | 3.36 | 3.17 |
| W415 CZ3 | | | C-7 (cf2) | 3.57 | 3.17 |
| H471 NE2 | O-1 | 3.12 | O-3 | 3.1 | 3.02 |
| D472 OD1 | O-8 | 2.56 | O-2 | 2.59 | 2.62 |
| D472 OD2 | O-1 | 2.6 | O-3 | 2.53 | 2.57 |
| Y727 OH | O-8 | 2.64 | O-2 | 2.72 | 2.68 |
| R876 O | ----- | ----- | C-7 (cf1) | 4 | 2.5 |
| | | | S | 3.49 | 4.14 |
| WATERS | | | N | 3.03 | |
| | | | C-7 (cf1) | 2.56, 2.85, 2.88 | |

The most striking structural feature of these inhibitor complexes is the fact that the C-7 carbon (i.e. the methyl group in the thiomethyl functionality) can occupy two possible positions (Figures 3.2 and 3.4). Both these positions are clearly visible in the electron density maps (Figure 3.2) and, in the case of **1**, have B-factors of 5.2 $Å^2$ and 6.8 $Å^2$ (conformer 1 (cf1) and conformer 2 (cf2), respectively) suggesting that cf1 conformer might be slightly favored.

The methyl group obviously forms favorable interactions in each of these positions, but the exact nature of the interactions is somewhat ambiguous and arguments can be made for both hydrophobic and hydrophilic bonds. In the case of the cf1 conformation of **1**, the methyl carbon is 4.3 Å from the CZ carbon in Arg228. The interaction between the methyl and Arg228 could be a C–H---cation type interaction where the C–H acts as donor. A statistical analysis of 1154 protein structures for C–H---π interactions revealed that such interactions are possible in proteins and are mostly intrahelical interactions[160] which could be classified as weak hydrogen bonds and which play an important role in the secondary structure stabilization. It was also found that the terminal methyl in methionine residue is one of the prominent donors for such interactions. The thiomethyl group of **1** can act as a donor to form such interaction with Arg228. There have been several theoretical studies, which estimate stabilization energies in other types of C–H---π interactions. However, to the best of our knowledge, energetic evaluation of the interaction between an aliphatic C–H and guanidinium moiety of an arginine has not been reported. Nevertheless, the relatively close proximity of the

guanidinium moiety as observed in the **1**-dGMII complex, indicate that there is strong possibility of formation of C–H---π type of interaction between the two functionalities.

It is tempting to speculate that the polarizable methyl group may hydrogen bond with neighboring water molecules. Three water molecules (W79, W80 and W159) are present at less than 3.2 Å from the methyl carbon in the cf1 conformation, while a fourth (W80) is 3.3 Å away. W159 also makes a strong hydrogen bond with the amino group of **1** and two other waters. Waters 78-80 show relatively weak electron density while W159 is quite well defined. Iwaoka *et al*[158] showed that a single C–H---O hydrogen bond from the methyl group to a backbone oxygen was weakly stabilizing and one could surmise that 3 such weak hydrogen bonds to water would have an additive effect.

On the other hand, there are arguments for hydrophobic interactions being the predominant binding force.[161, 162] In this respect, it has been shown that methionines display mainly hydrophobic character,[161] at least in the context of artificial hairpin peptides where their measurements were made. In fact, it could be that the water molecules are actually stabilizing the cf2 conformations rather than the cf1 conformation by binding to sulfur. In the cf2 conformation, sulfur can form a hydrogen bond with these water molecules via the lone pair electrons. Moreover, W79 is hydrogen bonded to Arg876O (2.7 Å) and W80 is hydrogen bonded to W79 and W81 suggesting that there may not be any interaction between the methyl in **1** cf1 and the water-sulfur cf2 interaction is what is visible in the electron density.

The X-ray crystal structure shows that in the cf2 conformation the thiomethyl ether of **1** makes two arene-sulfur interactions with Phe206 and Tyr727. In protein structures, methionine is often found in close proximity of aromatic side chain residues

indicating that arene-sulfur interactions are important for protein structure stabilization.[154-157] In general, arene-sulfur interactions are either hydrophobic of the S–$CH_3$---Ar type or electrostatic of the S---H–Ar type. *Ab initio* calculations of a dimethyl sulfide-benzene complex has predicted a stabilization energy of –1.6 kcal $mol^{-1}$ when methyl hydrogens point towards the centroid of an aromatic ring and when the distance between the sulfur atom and the centroid of the benzene ring ($R_{cen}$) is 5.8 Å.[163] In an alternative model, where both methyl groups are at the edges of the aromatic ring and the sulfur is 4.9 Å from the centroid of the aromatic ring, a stabilization energy of –2.5 kcal $mol^{-1}$ was calculated.[163] Furthermore, mutational experiments with β-hairpin model peptides[161] have indicated that arene-sulfur interactions are primarily hydrophobic in nature contributing up to –1 kcal $mol^{-1}$ of stabilization energy.

Interestingly, the X-ray crystal structure shows that the thiomethyl functionality of **1** cf2 complexes with Phe206 of dGMII through a hydrophobic $SCH_3$---Ar type interaction. In this case, the sulfur atom is placed slightly off-centered from the aromatic ring ($R_{cen}$ = 5.8 Å) similar to a mode predicted by *ab initio* calculations. The Tyr727 also makes a hydrophobic interaction with the thiomethyl group but in this case the methyl group is placed at the edge of the aromatic ring ($R_{cen}$=5.3 Å) similar to the second mode of interaction predicted by *ab initio* calculations.[164] Contact analysis shows that the thiomethyl group also makes van der Waals contacts with Trp415 and Arg876.

In the case of the **2**-dGMII complex, the methyl carbon in the thiomethyl functionality (C-7) also occupies two positions, the cf2 position is similar between **1** and **2** but the cf1 has moved. (Figure 3.4B). In the cf1 conformation, the methyl carbon is close to the backbone carbonyl oxygen of Arg876 and the sulfur atom seems to interact

with NH2 of Arg228, which is at 3.5 Å from the sulfur atom, via the lone pairs of the sulfur atom. This interaction has been studied in protein structures before and the structural preferences observed for this interaction suggest a contribution from the $n(S)\rightarrow\sigma^*(NH)$ orbital interaction to the stability.[158] Theoretical studies have shown that such interaction can contribute up to 2 kcal $mol^{-1}$ to the stability with 3.6 Å separation between the two atoms.[158] In the alternate cf2 conformation, this interaction may be weakened by the geometric position of the methyl group. However, this loss in the interaction may easily be compensated by the interaction between the sulfur atom and the Arg876O, which is also observed in **1**-dGMII complex and has been explained earlier. In addition to this interaction, the thiomethyl moiety is also involved in sulfur-π type interactions with Phe206 and Tyr727 in the same way as in the case of **1**. Additionally in the case of **2** cf2 the C-7 is perfectly centered at a distance of 3.17 Å with CH2 and CZ3 of Trp415. The close contacts between the thiomethyl group and water molecules observed in **1**-dGMII complex are not seen in the **2**-dGMII complex; there is no equivalent to W159 whose position has been taken up by the phenyl ring and the W79 equivalent (W361) has moved to 4 Å from the sulfur, and the W80 equivalent (W360) is now 3.8 Å away. There is now an intramolecular interaction between the phenyl ring and sulfur (see below), which displaces the water interactions.

It is important to note that additional modified derivatives of mannostatin (**1**) need to be synthesized and biologically evaluated to determine whether the polar or non-polar interactions of the thiomethyl moiety are responsible for its enhanced activity.

Docking studies[153] had indicated that the aromatic substituent of **2** would be able to make a parallel, off-centered π-π stacking interaction with phenyl ring of Tyr269 and

therefore, it was expected that **2** would be a better inhibitor than **1**. The unexpected finding that **2** is a poorer inhibitor than **1** can, however, be rationalized by careful examination of the crystal structure of dGMII with **2** (Figure 3.4). In both **1** and **2**, the five-membered rings are identically positioned in the binding site of dGMII (Figure 3.4B). The distance between the centroid of the aromatic ring of compound **2** and Tyr269 is 5.7 Å at an angle of 118° between the planes made by the aromatic rings, indicating that no edge to face binding interactions are made which require a distance of 5.0 Å and angle of 90°. The unexpected orientation of the aromatic ring of **2** can be rationalized by an intramolecular S---H–Ar type hydrogen bond between the thiomethyl moiety and the phenyl ring with P = 0.3 Å and L = 5.0 Å (where P is the vertical and L is the horizontal distance between the sulfur atoms and the centroid of the aromatic ring). *Ab initio* calculations have indicated that this special arrangement of phenyl ring and the sulfur atom can contribute up to –1.5 kcal $mol^{-1}$ of bonding energy.[165] These calculations do, however, not take into account entropic and solvation effects.

### 3.4 MD Simulations of Free Ligands

Cyclopentane rings, such as in compounds **1-3**, are inherently flexible due to their ability to assume several twist and envelope conformations (Figure 5), which can interconvert with relative ease *via* pseudo-rotational itineraries.[127, 128] Since ring substitutions may affect these pseudo-rotational itineraries,[128] compounds **1**-**3** may have different conformational properties. These differences may, in part, account for differences in the inhibition constants.

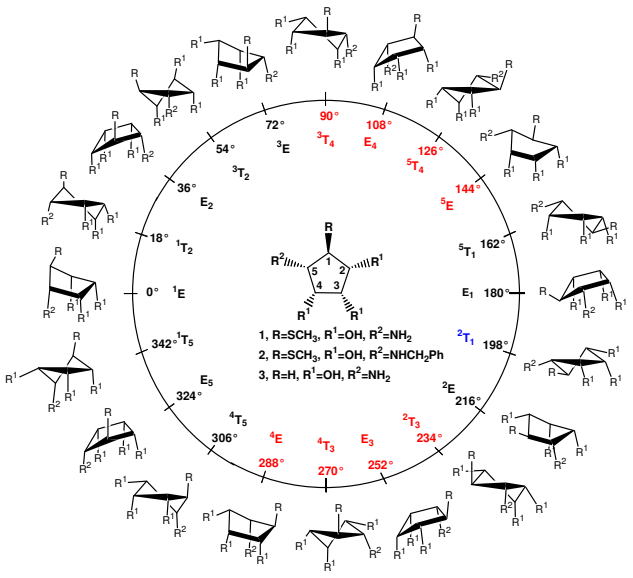



**Figure 3.6.** Pseudo-rotational itinerary for compounds **1**-**3**. Different envelope (E) and twist (T) conformers shown in the figure are uniquely defined by the value of phase angle (ϕ) of puckering. The phase angles for compounds **1-3** were computed using the carnal module in the AMBER suite of programs by employing Cremer and Pople's parameters. The bound conformation of **1** (blue) lies outside the preferred solution phase conformational families (red) as sampled in MD simulation of free ligand.

The conformational properties of **1** were probed by a 5-ns MD simulation of just the ligands in explicit water starting from the conformation of **1** observed in the X-ray crystal structure with dGMII. The most populated conformational family (80% occupancy) occupies pseudo-rotational itineraries between $^{3}T_{4}$ and $^{5}E$ ($\phi$ = 90-144°) whereas a second and minor conformational family (20% occupancy) represents conformers between $^{2}T_{3}$ and $^{4}E$ ($\phi$ = 234-288°) (Figures 3.6 and 3.7A). The $^{2}T_{1}$ conformer observed in the crystal structure was not populated and, thus, it appears that **1** is complexed in a high-energy conformation. The bound conformation of **1** is probably stabilized by a coordination of 3- and 4-oxygens with $Zn^{+2}$ ion and a hydrogen bond between the amine nitrogen and Asp204, which arrange atoms C-3, C-4 and C-5 in single plane resulting in an almost perfect $^{2}T_{1}$ ($\phi$ = 198°) conformation. This conformation also

allows interactions of the thiomethyl moiety with the aromatic residues of Phe206 and Tyr727.

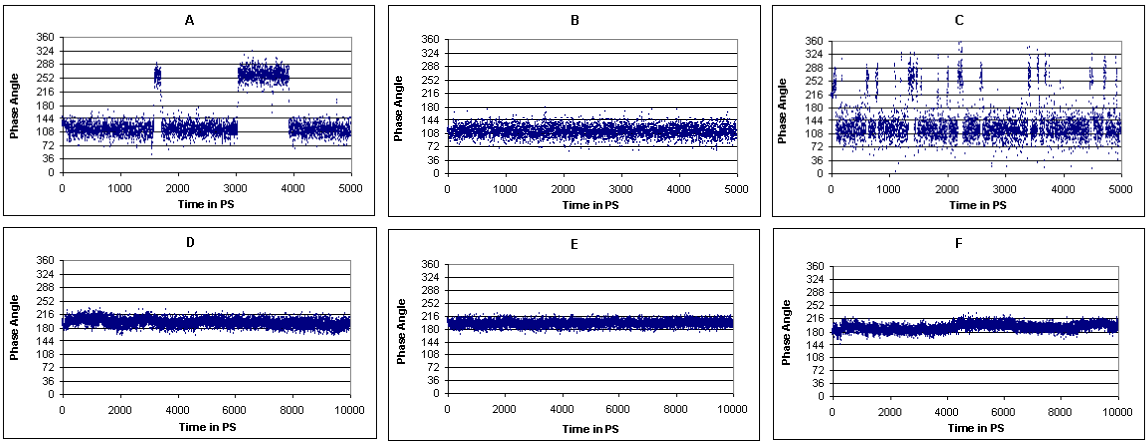



**Figure 3.7.** Plots of phase angles (ϕ) as a function of time in the MD simulation trajectories. A: **1** in water, B: **2** in water, C: **3** in water, D: **1** bound to dGMII, E: **2** bound to dGMII and F: **2** bound to dGMII with the phenyl ring in **2** making parallel off-center type of interaction with the phenyl ring in Tyr269. MD trajectories for ligands in a box of TIP3P water molecules (top row) indicate that compound **1** populates two clusters (A): the first consists of conformers $^{3}T_{4}$, $E_{4}$, $^{5}T_{4}$ and $^{5}E$ (ϕ = 90-144°) and the second consists of conformers $^{2}T_{3}$, $E_{3}$, $^{4}T_{3}$ and $^{4}E$ (ϕ = 234-288°). Compound **2** populates only one conformational cluster (B) with conformers $^{2}T_{3}$, $E_{3}$, $^{4}T_{3}$ and $^{4}E$. In the case of **3**, the interconversion barrier is reduced resulting in frequent conversion between different conformers (C). It can be noted that the bound conformation ($^{2}T_{1}$, ϕ = 201°) of the five-membered ring in **1** and **2** is different than the conformers populated in the 5-ns MD simulations and remains in the same conformational family comprising of $E_{1}$, $^{2}T_{1}$ and $^{2}E$ conformers (ϕ = 180-216°) during the 10 ns trajectory (bottom row, D and E). Although, the alternate orientation of the phenyl ring in **2** induces some flexibility in the five-membered ring, during the entire simulation the five-membered ring is restricted to the same conformational family (F), indicating that this arrangement of phenyl ring is also possible.

The MD simulation of **2** (Figure 3.7B) showed only the cluster of conformers between $^{3}T_{4}$ and $^{5}E$ (ϕ = 90-144°) indicating that the benzyl substituent increases the barrier for interconversion between conformational states. Interestingly, the intramolecular S---H–Ar interaction between the thiomethyl moiety and the phenyl ring

was not stable and after 200 ps the distance between the phenyl and sulfur was already 7 Å. Although the removal of the thiomethyl group of compound **3** decreased the interconversion barrier as evident from the frequent interconversions between the two clusters of conformers (Figure 3.7C), it populated the same conformational clusters as **1**.

In order to validate the computational results, proton-proton J-coupling constants for compounds **1-3** were calculated over the entire trajectory using a Karplus type equation[166] and the resulting values compared with experimental data from NMR solution experiments (Table 3.2). Gratifyingly, a good agreement was obtained between the computed and experimental *J*-values. The best agreements were obtained for compound **3,** which is aided by a more effective statistical sampling of the two conformational states in **3**.

**Table 3.2.** Computed and experimental homonuclear *J*-coupling constants (in Hz) for compounds **1-3**. The coupling constants were calculated using a karplus type equation proposed by Haasnoot et.al. (reference 48) over the entire 5-ns trajectory from MD simulations of ligands in a box of TIP3P water molecules. The values in the parentheses indicate the experimental values.

| | **1** | **2** | | **3** |
|---|---|---|---|---|
| | | | $J_{Hb\text{-}H1}$ | 1.5 (1.5) |
| $J_{H1\text{-}H2}$ | 6.2 (7.5) | 5.0 (6.3) | $J_{Ha\text{-}H1}$ | 7.9 (7.6) |
| $J_{H2\text{-}H3}$ | 6.3 (4.8) | 6.2 (5.1) | $J_{H1\text{-}H2}$ | 7.5 (6.2) |
| $J_{H3\text{-}H4}$ | 4.8 (4.1) | 4.8 (4.3) | $J_{H2\text{-}H3}$ | 5.9 (5.6) |
| $J_{H4\text{-}H5}$ | 5.1 (6.3) | 5.0 (6.1) | $J_{H3\text{-}H4}$ | 5.2 (5.6) |
| $J_{H5\text{-}H1}$ | 7.0 (7.0) | 8.0 (7.9) | $J_{H4\text{-}Ha}$ | 7.6 (7.6) |
| | | | $J_{H4\text{-}Hb}$ | 6.0 (5.7) |

### 3.5 MD Simulations of Protein-bound Ligands

While the X-ray crystal structures of **1** and **2** complexed with dGMII provided a great deal of insight into the binding modes of the inhibitors, MD simulations offer

additional information about the dynamic properties of the complexes. These simulations also allow an unambiguous analysis of hydrogen bonding patterns and occupancies and make it possible to determine the dynamic nature of H-π type, sulfur-π type hydrogen bonds and π-π type stacking interactions.

MD simulations of the complexes of **1** and **2** with dGMII only provided stable complexes when started with inhibitors in which the amino groups were protonated. This observation supports the notion that protonation of an amine of a glycosidase inhibitor is crucial for binding.[167] Furthermore, this behavior is in agreement with previously published SAR studies indicating the necessity of the amino group for inhibitory activity of the compounds.[83-88, 117] During the entire simulations no significant changes were observed in the protein backbone compared to the X-ray structure suggesting that MD simulations did not induce any conformational changes in the backbone. In addition to this, during the entire 10-ns simulation the cyclopentane rings of compounds **1** and **2** exhibit little flexibility as only one a small region of the pseudorotation wheel was occupied ($E_1$, $^2T_1$, $^2E$: $\phi$ = 180-216°, Figure 3.7D and 3.7E). This conformational family is different than the conformational states sampled in the simulations of ligands in explicit water, indicating that the cyclopentane rings do not revert to their low energy conformation.

Analysis of hydrogen bonding patterns and occupancies provides information about which interactions are the most important ones for complexation. Typically, the strongest hydrogen bonds have the highest occupancies, the smallest standard deviations and the shortest heavy atom separations. For compound **1**, very strong hydrogen bonds were observed between the 2-hydroxyl oxygen and Asp472 OD1-OD2, 3-hydroxyl

oxygen and Asp472 OD2 and 4-hydroxyl oxygen and Asp341 OD1 (Table 3.3). Moderately strong hydrogen bonds were assigned between 4-hydroxyl oxygen and Tyr269 OH, the amine nitrogen and Asp204 OD1-OD2, and Tyr269 OH and Asp341 OD1-OD2. The 2-hydroxyl oxygen also accepted a proton from Tyr727 to form a moderate hydrogen bond. In addition, the methyl group in the thiomethyl ether remained in cf2 conformation during the entire 10-ns period. The interactions between the thiomethyl moiety and the aromatic rings of Phe206 and Tyr727 were observed during the entire simulation further supporting the notion that these are important for ligand binding.

**Table 3.3.** Hydrogen bonding interactions between **1** and **2,** and dGMII over the 10-ns trajectories, computed using the carnal module in AMBER 7.0. The first value is the average distance between the specified atoms in Å, the values in parentheses indicate the percent occupancies.

| Donor atom | Acceptor atom | 1 | 2 | 2[a] |
|---|---|---|---|---|
| R228NH2 | S | 3.8 (5) | 3.9 (0.1) | 3.8 (32) |
| Y727OH | O-2 | 3.0 (40) | 3.1 (74) | 3.2 (51) |
| O-2 | D472OD1 | 3.2 (100) | 3.0 (100) | 3.0 (98) |
| O-2 | D472OD2 | 2.6 (100) | 2.9 (100) | 2.9 (100) |
| O-3 | D472OD1 | 3.9 (0.6) | 3.8 (29) | 3.9 (14) |
| O-3 | D472OD2 | 2.5 (100) | 2.5 (100) | 2.5 (100) |
| O-4 | Y269OH | 3.7 (54) | 3.8 (0.2) | 3.8 (1.5) |
| O-4 | D341OD1 | 2.5 (100) | 3.6 (90) | 3.6 (93) |
| O-4 | D341OD2 | 3.8 (7) | 2.6 (100) | 2.6 (100) |
| N | D204OD1 | 3.6 (33) | 3.6 (97) | 3.6 (83) |
| N | D204OD2 | 3.2 (49) | 3.4 (97) | 3.4 (81) |
| N | Y269OH | 2.9 (52) | 2.9 (91) | 3.0 (96) |
| N | D341OD1 | 3.0 (43) | 3.1 (100) | 3.4 (99) |
| N | D341OD2 | 3.5 (43) | 3.4 (99) | 3.3 (99) |

The properties of the hydrogen bonds of the complex of **2** with dGMII were very

similar compared to those of **1**. Furthermore, the intramolecular hydrogen bond between the thiomethyl moiety and phenyl ring was maintained throughout the simulation. In order to investigate in more detail the importance of the intramolecular S---H–Ar interaction of **2** in the complex with dGMII, a MD simulation was performed whereby the phenyl ring of **2** was stacked against Tyr269 (Figure 3.7F). Interestingly, the π-π stacking interaction was found to be stable (data not shown) with average $R_1$ = 3.3 Å and $R_2$ = 2.1 Å, where $R_1$ is the vertical and $R_2$ is the horizontal distance between the centroids of the phenyl rings. Theoretical study on benzene dimer indicates that this geometrical arrangement of the aromatic rings results in the stabilization energy of up to –1.5 kcal $mol^{-1}$.[168] Furthermore, the conformation of the cyclopentane ring and hydrogen bonding and other interactions were unchanged. This result is in agreement with our previous modeling studies,[153] even though this interaction is not observed in the crystal structure. This discrepancy between the crystal structure and the modeling study may be a result of the fact that the docking program used does not explicitly consider conformational entropy.

### 3.6 Overlay studies

The mode of inhibition of GMII by azasugars such as swainsonine has been rationalized by their resemblance to the mannosyl oxacarbenium ion, a putative intermediate in the hydrolysis of glycosides. However, extension of this model to the aminocyclopentitols, such as mannostatin, has been a subject of debate.[84, 89, 125, 169] Previous attempts to predict the precise binding mode of mannostatin A and its analogs resulted in multiple binding models that differed in the conformation of the five-

membered ring of these inhibitors.[153] The X-ray structures of dGMII complexed with inhibitors **1** and **2** offers a unique opportunity to compare the three dimensional arrangement of the functional groups of **1** and **2** relative to that of the mannosyl oxacarbenium ion.

Previously, it has been reported that the mannosyl oxacarbenium ion can adopt at least two conformations in which C-4 is either above (flap-up) or below (flap-down) the plane of the C-2, C-1, O-5 and C-5 atoms.[89, 169] Optimized geometries of both the conformers were obtained by quantum mechanical calculations at the DFT/B3LYP/6-31G* level of theory. It was found that the flap-up conformer was more stable than flap-down conformer by 4.4 kcal $mol^{-1}$, which is in agreement with previous calculations.[169] Subsequently, the two conformers of the oxacarbenium ion were superimposed on the conformation of mannostatin A observed in the X-ray crystal structure. The zinc coordinating C-3, C-4 *cis*-diol of **1** mimics the C-3, C-2 *cis*-hydroxyl of the oxacarbenium ion and therefore the O-3, C-3, C-4 and O-4 of **1** were superimposed onto O-3, C-3, C-2 and O-2 of the oxacarbenium ion conformers (Figure 7).

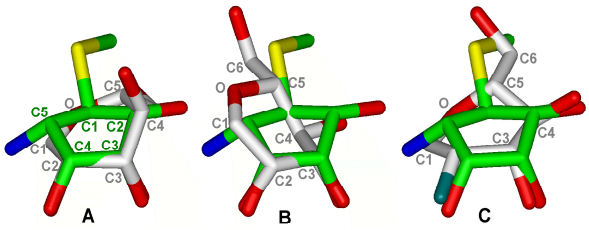



**Figure 3.7.** Superimposition of Mannostatin A (**1**) in bound conformation onto the flap-up (A) and flap-down (B) oxacarbenium ion, and covalent intermediate (C, PDB ID 1QX1). Mannostatin A mimics best the covalent intermediate.

In the case of the flap-up conformer, the cyclopentane ring of compound **1** seems to mimic the sugar ring with C-2, C-3 and C-4 carbon atoms lying very close to C-4, C-3

and C-2 carbon atoms in the sugar ring and the C-1 carbon is positioned very close to the ring oxygen. The zinc-coordinating 3- and 4-oxygen atoms in **1** display a nearly perfect overlap with corresponding hydroxyl oxygens with RMS deviation of 0.16 Å. The C-5 carbon is located above the anomeric carbon of the oxacarbenium ion placing the amino nitrogen in the region of the exo-cyclic anomeric oxygen of the substrate. The protonated form of this amine at physiological pH raises the possibility that in this arrangement the amine nitrogen mimics the positive charge on the exo-cyclic anomeric oxygen, which develops during the formation of the oxacarbenium ion. In the case of the inhibitor, swainsonine, the amine nitrogen is positioned very close to the ring oxygen suggesting that it probably mimics the positive charge on the ring oxygen in the transition state. The C-2 hydroxyl and thiomethyl moiety of **1** do not seem to mimic any functionality of the oxacarbenium ion. However, these moieties are involved in multiple interactions with protein residues enhancing the affinity of **1** for dGMII.

Attempts to superimpose **1** with the flap-down conformer were less successful. The C-3 and C-4 carbon atoms lie near the C-3 and C-2 atoms in the oxacarbenium ion, respectively, while all the other carbons are at a considerable distance. Moreover, the 3- and 4-hydroxyl groups display poorer overlap with the corresponding hydroxyl groups in the oxacarbenium ion with greater RMS deviation of 0.33 Å, compared to 0.16 Å in the flap-up conformer. The amine group now points away from the β-face of the oxacarbenium ion. It has been proposed that β-mannanase catalyzed hydrolysis proceeds through a transition state in which the oxacarbenium ion adopts a $B_{2,5}$ conformation.[24] Therefore, the geometries of the mannosyl oxacarbenium ion in the $B_{2,5}$ and $^{2,5}B$ conformation were optimized and the resulting structures superimposed on the bound

conformation of 1. In each case, only the C-3, C-4 diol of **1** coordinating with Zn overlayed well with the C-2, C-3 diol of the mannosyl oxacarbenium ions (results not shown) demonstrating that **1** does not mimic a TS in the boat conformation.

Interestingly, the overlay of **1** with a covalently trapped fluorinated sugar analog,[38] which mimics the covalently linked reaction intermediate and is found in a distorted $^{1}S_{5}$ skew boat conformation, shows a good correspondence with all three hydroxyls of **1**. Furthermore, in this overlay the sulfur is in a close proximity to the C-6 hydroxyl of the covalently linked sugar analog (Figure 7C). The thiomethyl group of **1** makes, however, additional polar and non-polar interactions with protein residues compared to the equivalent C-6 hydroxyl, thereby enhancing the affinity of compound **1** for dGMII. The amine is in close proximity to C-1 of the covalently linked intermediate and therefore is ideally positioned to interact with the catalytic nucleophile (Asp204).

Thus, the overlay studies indicate that mannostatin mimics best the enzyme-linked mannosyl intermediate, which has been proposed to adopt a $^{1}S_{5}$ skew boat conformation. This conformation has been implicated in the reaction mechanism of α- as well as β-retaining glycosidases. For example, the three-dimensional structure of a β-mannanase from *Pseudomonas cellulosa* (endo β-retaining mannosidase) was determined at various points along its reaction pathway.[24] Interestingly, the conformation of the Michaelis complex of the β-mannanase was identical to that of the fluorinated sugar analog linked to dGMII.

The five-membered ring of swainsonine is believed to resemble the flattened six-member ring of the mannosyl oxacarbenium ion. However, superimpositions of swainsonine onto the two mannosyl oxacarbenium ion conformers and the enzyme

trapped intermediate showed that the best overlay is obtained with the latter intermediate (overlays shown in the supplementary material). It is generally believed that potent glycosidase inhibitors mimic an oxacarbenium-ion like transition state. The results of our studies highlight that they may also mimic reaction intermediates such as the Michaelis complex or enzyme-linked intermediate.

In his superimposition study,[89] Winkler reported that Mannostatin A mimics the flap-up conformer with the functional groups present in mannostatin A mimicking different hydroxyl groups in the oxacarbenium ion. Winkler used the X-ray structure of Mannostatin A tetraacetate for these superimposition studies.[170] It must be noted that conformation of the five-membered ring in Mannostatin B tetraacetate is $^{3}T_{4}$ ($\phi = 93.5°$), which is different than the bound conformation of the cyclopentane ring ($^{2}T_{1}$) in both the X-ray structures reported here. Moreover, 2- and 3-hydroxyl groups in mannostatin A were superimposed onto the zinc-coordinating 3- and 2-hydroxyl groups in the oxacarbenium ion, respectively. However, from the X-ray structure of **1**-dGMII complex, it is revealed that 3- and 4-hydroxyl groups in **1** coordinate with the $Zn^{2+}$ ion and thus mimic the zinc-coordinating cis-diol in the actual substrate. The important role of the zinc ion in the catalytic mechanism of dGMII has recently been discussed.[80]

### 3.7 Conclusions

Glycosidases are enzymes that play crucial roles in the biosynthesis of glycoproteins. Inhibitors of these enzymes have garnered much attention as lead compounds for drug discovery for diseases such as viral- and bacterial infections, diabetes, Gauchers disease and cancer. In many cases, the mode of inhibition by these

compounds is not well understood, complicating efforts to design and synthesize more potent- and/or selective inhibitors. As a result, only a very small number of glycosidase inhibitors have been successfully developed as therapeutics.

The studies reported here indicate that mannostatin A mimics the covalently linked mannosyl intermediate, which has been shown to adopt a $^{1}S_{5}$ skew boat conformation. In particular, the zinc-coordinating 3- and 4-hydroxyls of **1** display a good overlap with the corresponding *cis*-diol of the mannosyl residue. Furthermore, the amine nitrogen of **1** is positioned close to the C-1 carbon of enzyme-linked mannosyl intermediate and therefore can interact with the carboxylic acid of the catalytic nucleophile Asp204. The thiomethyl group of mannostatin A, which is required for high affinity binding, showed a good overlay with the C-6 hydroxyl of the covalently linked intermediate. This functionality is, however, able to make a number of additional interactions increasing the affinity for dGMII. The thiomethyl moiety of **1** is structurally similar to the side chain of methionine residues. Interestingly, the environment surrounding the thiomethyl group of **1** is remarkably similar to the arrangements around the methionine residues in the protein structure. While this observation supports the importance of interactions made by the thiomethyl moiety in methionine with respect to protein folding and secondary structure stability, it also supports the role of thiomethyl moiety in mannostatin A binding. Another important finding is that the methyl carbon of the thiomethyl moiety of **1** in the complex with GMII adopts two different conformations. In both conformations, the sulfur interacts with the backbone oxygen of Arg876 through the $\pi$-orbital on the carbonyl oxygen and the antibonding $\sigma^*$ orbital of the S-C bond. In cf1 conformation, the polarizable methyl group is in a hydrophilic environment

interacting with three water molecules and the π-system of the Arg228 side chain. The side chains of methionine residues 224 and 264 are in similar hydrophilic environments. On the other hand, the methyl group of the cf2 conformation is in a hydrophobic environment where it forms C–H---π type interactions with the phenyl rings of Phe206 and Tyr727. In the protein structure, similar interactions have been observed for the Met769.

The MD simulations and NMR studies of the uncomplexed inhibitor showed that the five-membered ring can adopt two different conformational families. In the bound state, however, it is restricted to a single conformation, which is different from the conformations observed for the free ligand. This finding is surprising because it is generally believed that potent inhibitors are complexed in their ground state conformations. Probably, the flexibility of five membered rings renders the energy barrier for the conformational change relatively low.

Compound **2** was designed in such a way that its benzyl moiety can interact with aromatic residues of the binding site of the enzyme thereby increasing its affinity. However, it was found that this compound has a slightly reduced affinity indicating that the benzyl group cannot make favorable interactions with the enzyme. The unexpected finding could be rationalized by careful examination of the crystal structure of dGMII with **2**. Thus, the five-membered ring of compounds **1** and **2** were identically positioned in the binding site of dGMII. The distance between the centroid of the aromatic ring of compound **2** and Tyr269 was such that no edge to face binding interactions could be made. The unexpected orientation of the aromatic ring of **2** could be rationalized by an

intramolecular S---H–Ar type hydrogen bond between the thiomethyl moiety and the phenyl ring.

### 3.8 Experimental Procedures

Compounds **1-3** were prepared by reported procedures.[84, 87, 119] Inhibition constants were determined as detailed in Li *et al*.[153]

#### 3.8.1 Crystallization, Data Collection and Structure Refinement

Measurement of inhibition, crystallization, data collection and structural refinement were carried out essentially as outlined by Kuntz *et al*[82] with the exceptions noted below. Inhibitors were dissolved at 100 mM in water (**1**) or methanol (**2**) Crystals of dGM2 were grown up overnight, washed with phosphate buffered reservoir solution and soaked with 10 mM **1** for at least 3-6 hrs. **2** was co-crystallized with dGMII in Tris-buffered reservoir solution without phosphate washing.[81] Crystal quality was assessed on our home source and data was collected on Beamline F1 at the Cornell High Energy Synchrotron Source. Approximately 400 frames with 0.5 degree oscillation/frame were collected for each dataset. Each of the structures was refined independently. The last stage involved addition of alternate conformations and low quality waters followed by energy minimization and individual B-factor refinement. The quality of the final model was assessed using a number of structure validation programs including MolProbity, WhatIf, and hetze.

#### 3.8.2 Molecular Dynamics Simulations

The sander[171] module of AMBER 7.0[97] in conjunction with PARM99 parameter set for proteins, was utilized in molecular dynamics simulations. A modified GLYCAM[172]

parameter set for glycosides and glycoproteins was used for all three ligands. The X-ray crystal structures were protonated using the leap module in AMBER and were solvated by a 25 Å droplet of TIP3P[99] waters around the $Zn^{+2}$ ion. Initially the solvent positions were optimized by energy minimization followed by energy minimization of whole system. This was followed by MD simulations at 300 K for 10-ns. In each simulation the system was heated gradually to 300 K in 100-ps and then held at 300 K for the entire simulation. For all simulations, protein residues within a sphere of 10 Å (with $Zn^{+2}$ ion as the center of the sphere), all waters and ligands were allowed to move. All other atoms were frozen to their positions in the crystal structures.

Simulations of free ligands were performed for 5-ns, under periodic boundary conditions at constant pressure. These simulations followed similar protocols for energy minimizations as used for the simulations of protein-bound ligands. The initial conformations of ligands in these simulations were kept same as observed in the crystal structure. In order to determine hydrogen bond occupancies, hydrogen bonding interactions were assumed to be present if the participating heavy atoms were 4 Å apart and the angle between the heavy atoms and the donating hydrogen was 60°. All MD simulations used an integration time step of 2 fs and scaling of 1-4 electrostatic and van der Waal interactions by standard values of 1/1.2 and 1/2.0, restraint of all hydrogen-containing bonds through the SHAKE algorithm[171] and a cutoff of 8-Å for all non-bonded interactions. The MD trajectories were analyzed using the carnal module in AMBER 7.0. The phase angles of puckering were computed using Cremer and Pople's parameters.[173]

### 3.8.3 Geometry optimization of flap-up and flap-down conformers of mannosyl oxacarbenium ion and superimpositon

All the geometry optimizations were performed with Gaussian 98 program[112] using density functional theory (B3LYP[139-141]) and 6-31G* basis set.[142] The optimized conformers of flap-up and flap-down oxocarbenium ions were determined by constraining the dihedral angle made by C-2, C-1, O-5 and C-5 to 0° and allowing all other parameters to optimize at BLYP/6-31G* level of theory. The hydroxyl groups of the resulting structures were placed in three different orientations representing three rotamers and the conformers were reoptimized without any restraints at B3LYP/6-31G* level. For flap-up conformer two- and for flap-down conformer three unique conformers were identified with different orientations of hydroxy oxygen atoms. For each conformer, the C-2, C-1, O-5 and C-5 were in one plane with a distance between C1 and endocyclic oxygen of 1.27 Å, indicating oxacarbenium-like character. The lowest energy conformation of flap-up and flap-down conformers was used for superimposition with inhibitors **1** and **2**. The covalent intermediate geometry was obtained from PDB 1QX1. The superimpositions were performed using Insight II 2000.

## 3.9 Acknowledgement

The authors wish to thanks the staff of CHESS for data collection time and support. The authors are also grateful for financial support for this work by NCI of NIH (5UO1CA91295) and the Canadian Institutes of Health Research.

## 3.10 Supporting Information

Protein coordinates and structure factors are available from the Protein Data Bank. The PDB IDs are 2F7O (Mannostatin A) and 2F7P (benzylated Mannostatin A).

# CHAPTER 4

# CATALYTIC MECHANISM OF GOLGI α-MANNOSIDASE II: A QUANTUM MECHANICAL STUDY

## 4.1 Abstract

A high-level quantum mechanical (QM) investigation of the hydrolysis reaction catalyzed by Golgi a-mannosidase II (GMII), which acts late in the *N*-glycan biosynthesis pathway, has been conducted. In all, four models have been studies, which differ in the number of atoms and the overall charge. Best results were obtained for the model, which contained explicit water molecules. These calculations indicate that in the transition state of the glycosylation reaction, the mannopyranose ring adopts a conformation that lies between an ideal $B_{2,5}$ boat conformation and $^{1}S_{5}$ skew boat conformation. It has also been shown that the transition state conformation of the sugar ring resembles oxacarbenium ion. In the case of deglycosylation step, the sugar ring adopts a conformation between $B_{2,5}$ and $^{O}S_{2}$ skew boat conformation. The glycosyl-enzyme covalent intermediate displays a conformation close to $B_{2,5}$ boat conformation, and is in excellent agreement with the X-ray strutcutre of trapped covalent intermediate. Finally, the bound conformations of swainsonine and mannostatin A have been compared. These superimposition studies support long held view that swainsonine and mannostatin A are transition state mimics.

## 4.2 Introduction

Glycosidases are enzymes that catalyze the cleavage of glycosidic bonds and play critical roles in a number of biological processes. Inhibitors of these enzymes have garnered much attention and are currently used for the treatment of diabetes, viral infections and Gauchers disease.[6-10] Furthermore, inhibition of the mannose trimming enzyme human Golgi α–mannosidase II (GMII; mannosyl-oligosaccharide 1,3-1,6-

α–mannosidase II; E.C. 3.2.1.114), which acts late in the *N*-glycan processing pathway, provides a route to blocking the oncogene-induced changes in cell surface oligosaccharide structures.[61, 63, 113, 148] GMII selectively cleaves α(1-3) and α(1-6) mannosyl residues present in its natural substrate $GlcNAcMan_5GlcNAc_2$.[149] It is a retaining Family 38 glycosylhydrolase, which employs a two-stage mechanism (Figure 4.1) involving two carboxylic acids positioned within the active site which act in concert: one as a catalytic nucleophile and the other as a general acid/base catalyst.[12, 14, 19, 114, 115] Protonation of the exocyclic glycosyl oxygen of a substrate molecule leads to bond-breaking and simultaneous attack of the catalytic nucleophile to form a glycosyl enzyme intermediate. Subsequent hydrolysis of the covalent intermediate by a nucleophilic water molecule gives an α-mannose product with overall retention of configuration. Studies with 5-fluoro pseudo-substrates and deuterium labeled substrates have shown that the transition states on either side of the covalent intermediate have marked oxacarbenium ion character.[12, 14, 19, 38, 114, 115] Furthermore, X-ray crystal structures of the wild type and a mutant Drosophila melanogaster GMII (dGMII) in which the acid/base catalyst has been removed, with fluorinated sugar analogs have revealed that the glycosyl enzyme intermediate adopts a distorted $^1S_5$ skew boat conformation.[38] In this conformation, the leaving group is placed anti-periplanar to the lone pair of the ring oxygen, a requirement for the departure of the leaving group according to Deslongchamp's anti-periplanar lone-pair hypothesis.[79] Furthermore, steric clashes between the *syn*-hydrogens at C-3 and C-5 and the attacking water are minimal in the $^1S_5$ skew boat conformation.

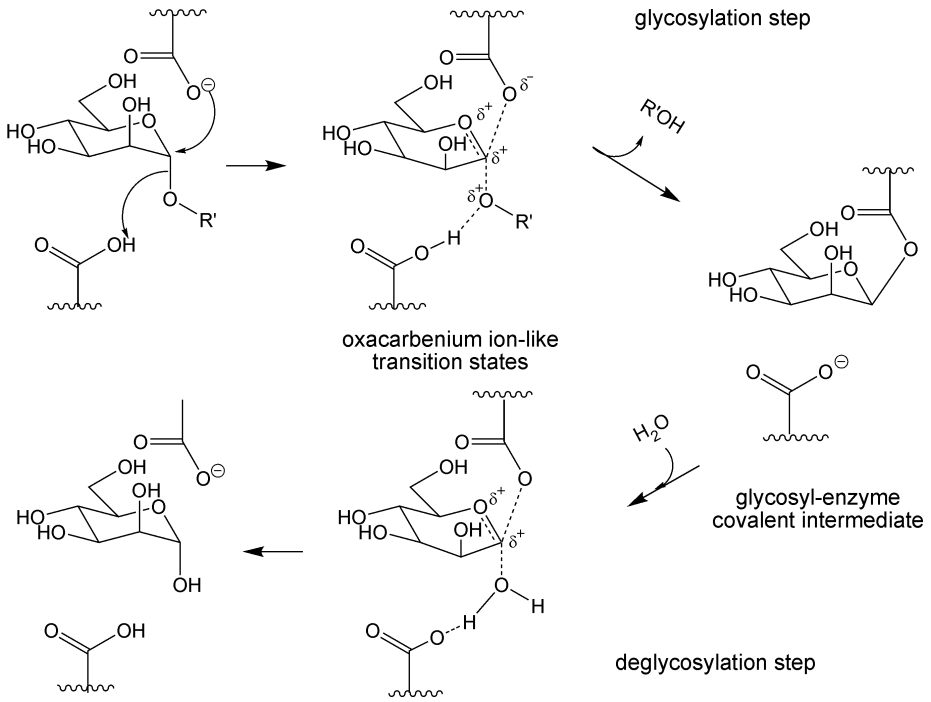



**Figure 4.1.** General double displacement mechanism proposed for retaining mannosidases.

Potent inhibitors of glycosidases are believed to mimic oxacarbenium ion-like transition states.[12, 14, 19, 114, 115] The inhibitory activity of the natural product swainsonine has long been attributed to the resemblance of its 5-membered ring to a flattened six-membered ring mannosyl oxacarbenium ion. In fact, the crystal structure of swainsonine complexed with dGMII shows the inhibitor to be tilted in such a way as to bring the equivalent of its anomeric carbon close to the presumed catalytic nucleophile. [75] While similarities between swainsonine and mannostatin to the trapped covalent intermediates have recently been shown,[174] it is till unclear which conformation is adopted by the mannopyranose ring at the transition state, and whether potent inhibitors of GMII mimic the transition state.

In principle, computational methods are capable of determining enzymatic catalytic mechanisms to an appropriate level of detail when used in complement with experiments. However, formidable problems have had to be overcome in order to elucidate the mechanism using computational methods. The central role of transition metal atom in the active site chemistry demand use of *ab initio* quantum mechanical methods is the desired level of accuracy is to be achieved. While significant progress has been made in semiempirical methods and force field methods, these methods are system-specific and often require parameterization of the system under consideration.[100] The diverse nature of transition metals (structurally and electronically) makes it difficult to describe any system containing a transition metal atom with a set of universal parameters.[175, 176] Furthermore, the number of atoms that can be treated by *ab initio* methods, although has increased substantially in recent years, is still limited. The selection of model thus becomes critical as inclusion of large number of atoms can lead to intractable CPU times. On the other hand, smaller model may not be sufficient to produce reasonable results as it lacks explicit effects of the protein residues on the active site chemistry.

Over past few years, effective computational methods have emerged to address these problems. The widespread use of density functional theory (DFT)[177, 178] deploying gradient-corrected and hybrid functionals has demonstrated that transition metal systems can be treated in a reasonably accurate, yet cost-effective fashion (for review see reference [100] and references there in). Advances in DFT-based algorithms, along with dramatic reductions in the cost/performance of computational hardware, allow on the order of 150 – 200 atoms to be modeled at a quantum chemical level. Progress has also

been made specifically in the treatment of transition metals, including areas such as construction of an improved initial guess.[179] While the accuracy of DFT methods is best established for organic systems, a substantial number of comparisons have been made between theory and experiment (see reference [100] and references there in). Furthermore, it has been shown that Becke's three-parameter hybrid exchange functional[139] gives notable improvements in metal-ligand bond energies. Similar trend was observed in benchmark calculations using small organic molecules.[100]

## 4.3 Computational Procedure and Models Used

In this work, we have investigated the mechanism of the hydrolysis reaction catalyzed by GMII using QM methods. The general strategy employed for this purpose includes four major steps: 1. docking the substrate into the active site of GMII; 2. truncation of the active site to create a model of reasonable size; 3. performing a relaxed geometry scan along a pre-defined reaction coordinate; and 4. optimization of the approximate geometries of the intermediates predicted in the scan. Tvaroska et al. have successfully used this strategy to determine the catalytic mechanism of glycosyl transferases.[180-182]

### 4.3.1 Docking of 1-methyl mannopyranoside into the active site of dGMII

In this study, 1-methyl mannopyranoside was used as the substrate. The substrate was placed in $^4C_1$ conformation and was optimized at B3LYP/6-31G** level using Gaussian98 program.[183] The optimized structure of 1-methyl mannopyranoside was docked in to the active site of dGMII by the procedure explained as follows.

*Preparation of ligand and receptor molecules for docking* – The crystal structure of dGMII (PDB id: 1HTY[75]) was used as a model for macromolecule in docking experiments. The protein target, dGMII and the ligands were prepared for docking using AutoDock 3.0.5.[90, 144] Charges were assigned using the Kollman algorithm.[145] Atomic solvation parameters and fragmental volumes were determined using Addsol programs available in AutoDock. AutoTors was used to define torsional angles in the ligand. Polar hydrogen charges of Gasteiger type [146] were assigned and the non-polar hydrogens were merged with the carbons. The macromolecule was kept rigid in all the docking simulations.

*Docking simulations* – Grid maps for docking simulations were generated with 40 grid points (with 0.375 Å spacing) in x, y and z direction, by Autogrid program. The center of the grid was positioned at the $Zn^{+2}$ ion (Zn1102). Lennard-Jones parameters 12-10 and 12-6 (supplied with the program package) were used for modeling H-bonds and van der Waals interactions, respectively. The distance dependant dielectric permittivity of Mehler and Solmajer[147] were used for the calculations of the electrostatic grid maps. The genetic algorithm (GA) and Lamarckian genetic algorithm with the pseudo-Solis and Wets modification (LGA/pSW) methods were used with the default parameters. For all the simulations, the populations in the genetic algorithm were 50 and each simulation comprised of $2.5\times10^5$ energy evaluations. Each docking experiment consisted of a series of 100 simulations.

The docking experiments resulted in exclusively one docking mode. In this orientation, the mannopyranosyl ring is stacked against the aromatic ring of Trp95. This type of interaction is common for several carbohydrate-protein complexes and is an

important interaction for carbohydrate recognition.[134, 150-152] The 2- and 3-hydroxyl groups are coordinated to $Zn^{+2}$ ion, whereas 4-hydroxyl group forms a hydrogen bond with the carboxylate oxygen atom of Asp472.

### 4.3.2 Selection of the model

In the case of quantum mechanical calculations, the selection of the model used is critical, as the size of the model affects the required time and the accuracy of calculations. However, the selection of residues to be included in the model is not obvious. We have used four models, which differ in the number of atoms and overall charge (Table 4.1) in this study.

**Table 4.1.** Different models used in the computational study of the catalytic mechanism of GMII.

| | **Model 1** | **Model 2** | **Model 3** | **Model 4** |
|---|---|---|---|---|
| Number of atoms | 76 | 130 | 160 | 150 |
| Protein Residues | Asp92, Asp204, Asp341, Asp472, Trp95 | Model 1 + Ile96 and Arg228 | Model 2 + His90 and His471 | Model 1 + Ile96, His90, His471 and 4 water molecules |
| Overall charge on the model | –1 | 0 | 0 | –1 |
| Zn coordination sphere | Asp92, Asp204 and *cis*-diol in the substrate | Same as Model 1 | Asp92, Asp204, His90, His471 and *cis*-diol in the substrate | Same as Model 3 |
| Level of theory | B3LYP/LACVP* | B3LYP/LACVP* | B3LYP/6-31G** (LACVP** for zinc) | B3LYP/6-31G** (LACVP** for zinc) |

### 4.3.3 Performing a relaxed geometry scan

The catalytic mechanism was investigated using relaxed geometry scans, which include finding minimum-energy geometries while holding a particular coordinate fixed to various values. Typically, this coordinate is the distance between the atoms involved in the process of bond formation or bond breaking. In the case of GMII, each step of the double displacement reaction can be represented as a function of different reaction coordinates. For example, the glycosylation step can be represented as a function of two reaction coordinates (Figure 4.2): the distance between the nucleophilic oxygen of Asp204 and the anomeric carbon in the substrate, and the distance between the anomeric oxygen and the proton from the general acid Asp341. The first coordinate mimics the formation of bond between the anomeric carbon and the catalytic nucleophilic oxygen whereas the second coordinate mimics the protonation of the leaving group. During a relaxed geometry scan, these distances will be changed and corresponding potential energy is calculated. The potential energy of the system can then be represented as a function of the coordinates in a three-dimensional plot. Such a two dimensional, relaxed geometry scan would result in the prediction of the alternative reaction pathways. Such geometry scans are useful to locate the approximate geometries of the Michaelis complex, the product of the catalytic reaction and other reactive reactive intermediates observed along the reaction pathway. The exact structures and energies can be obtained by optimizing the approximate geometries without any restraints on the distances that have been used as reaction coordinates.

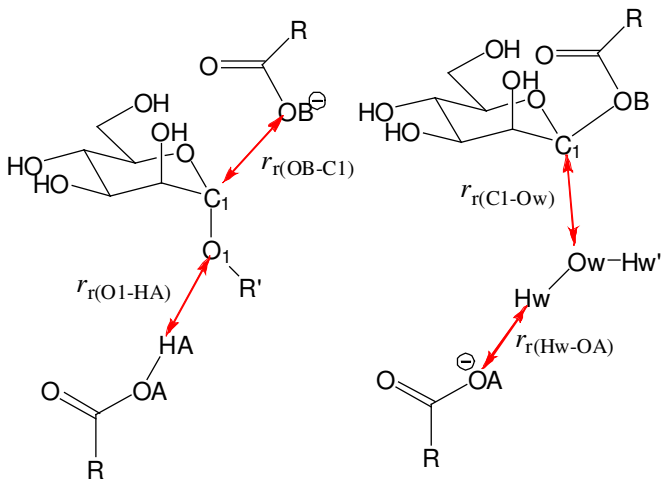



**Figure 4.2.** General representation of the relaxed geometry scans used in the investigation of the catalytic mechanism of GMII. Each step can be represented with the help of two reaction coordinates.

## 4.4 Model 1

### 4.4.1 Description of the Model and Computational Procedure

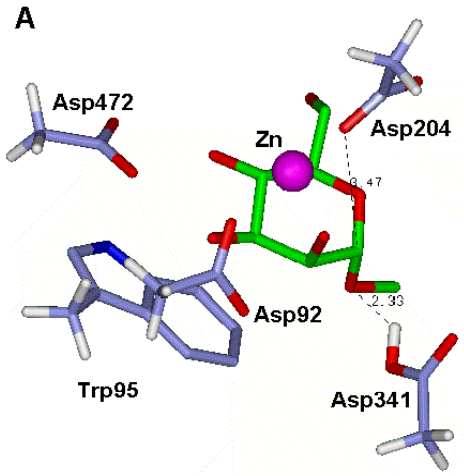



**Figure 4.3.** Model 1 used for the quantum mechanical study of the catalytic mechanism of GMII. The model consists of 76 atoms and the over all charge on the model is –1. The model includes the substrate, 1-methyl mannopyranoside; $Zn^{+2}$ ion; four aspartic acid residues (Asp92, Asp204, Asp341, and Asp472); and Trp95.

In this investigation, the structural model used to analyze computationally the catalytic mechanism of GMII consists of minimum essential residues, or their fragments that are involved in the reaction and are in direct contact with the substrate. The reaction site model, Model 1, consists of 76 atoms and contains: the substrate, 1-methyl mannopyranoside; $Zn^{+2}$ ion; four aspartic acid residues (Asp92, Asp204, Asp341, and Asp472); and Trp95 (Figure 4.3). The overall charge on the model is –1. All protein residues are truncated at the Cβ carbon atoms and are frozen at their positions in the X-ray structure of dGMII (PDB ID 1HTY). All Cβ atoms were represented by methyl groups. Furthermore, all atoms in Trp95 were frozen at the crystallographic position. Since Asp341 has been identified as the general acid-base catalyst in GMII,[75] this residue is protonated in the model and acts as general acid residue in the glycosylation step of the double displacement mechanism.

In the case of the glycosylation step, two distances were considered as the reaction coordinates: the distance ($r_1$) between the anomeric carbon and the catalytic nucelophilic oxygen Asp204 OD1, and the distance ($r_2$) between the anomeric oxygen and the proton from the general acid residue Asp341. The relaxed geometry scan was performed at DFT/B3LYP/LACVP* level of theory. During the scan, distance $r_1$ was decreased from 2.93 Å to 1.33 Å by 0.2 Å, whereas the distance $r_2$ was decreased from 1.69 Å to 0.89 Å by 0.2 Å. The approximate geometries obtained from this scan were further optimized at DFT/B3LYP/LACVP*. LACVP* uses 6-31G* basis for atoms up to Ar in the periodic table, whereas for zinc it uses effective core potential.

### 4.4.2 Results and Discussion

The glycosylation step of the double displacement reaction mechanism is described by means of a potential energy surface (PES) represented in the form of two-dimensional reaction coordinate contour diagram (Figure 4.4). The x-axis of the plot represents distance $r_2$, which mimics the protonation of the leaving group. The y-axis of the plot represents distance $r_1$, which mimics the formation of the covalent linkage between the anomeric carbon and the catalytic nucleophilic oxygen of Asp204.

The mannopyranose ring adopts a $B_{2,5}$ boat conformation at the Michaelis complex (MC). In this conformation, the 2- and 3-hydroxyl group coordinate with the $Zn^{+2}$ ion in the active site. Furthermore, the anomeric oxygen forms a hydrogen bond with the proton from general acid Asp341 (1.7 Å), suggesting that this interaction may facilitate the protonation of the leaving group and its subsequent departure. The 4-hydroxyl group also forms a strong hydrogen bond with Asp472.

While it was possible to locate the Michaelis complex MC, it was not possible to identify the transition state and the glycosyl-enzyme covalent intermediate using the PES obtained from the geometry scan of Model 1. The PES displayed a plateau at about 45 kcal $mol^{-1}$, which ranged from $r_1$ = 1.93 Å to 1.43 Å and $r_2$ = 1.29 Å to 0.89 Å. Further approach of the catalytic nucleophile towards the anomeric carbon resulted in the higher potential energy (more than 48 kcal $mol^{-1}$). Furthermore, all attempts to optimize the transition state (TS1) and the glycosyl-enzyme covalent intermediate (CINT) failed. These calculations resulted in the Michaelis complex structure.

In order to confirm our results, we performed another PES scan where we changed r1 from 1.93 Å to 1.43 Å by 0.05 Å and r2 from 1.29 Å to 0.89 Å by 0.05 Å. This is the region where the first PES scan displayed a plateau. The analysis of the

refined PES (Figure 4) revealed three intermediates with identical energies (45.6 kcal $mol^{-1}$, 45.9 kcal $mol^{-1}$ and 46.6 kcal $mol^{-1}$). Optimization of all three intermediates resulted in the Michaelis complex.

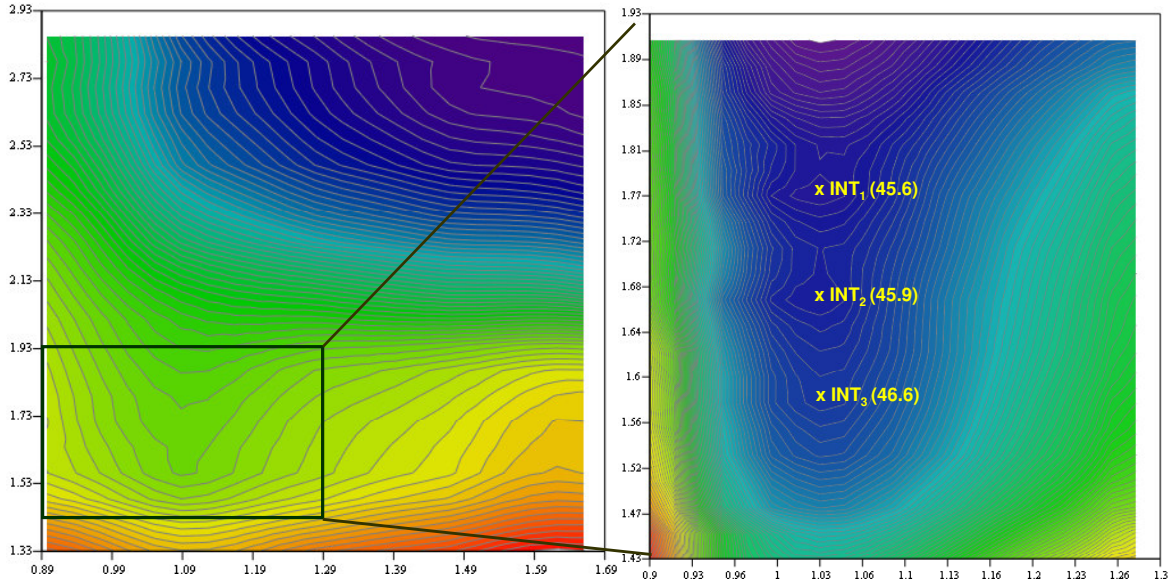



**Figure 4.4.** Potential energy surface obtained from the two-dimensional scan of Model 1 at DFT/B3LYP/LACVP* level (purple – lowest energy region, orange – highest energy region). Initial PES scan resulted in single reaction pathway. However, the approximate locations of the transition state and the glycosyl-enzyme covalent intermediate could not be identified. A refined scan resulted in three intermediates of identical energies. These intermediates converged to the structure of the Michaelis complex following the optimization at the same level.

The fact that the identification of approximate location of TS1 and CINT was not possible on the PES obtained from Model 1 suggests that Model 1 probably is an insufficient model to determine the catalytic mechanism of GMII. Furthermore, the potential energy (~ 45 kcal $mol^{-1}$) in the region, where the transition state is expected, is high for an enzymatic catalytic reaction. This may be the result of the lack of flexibility in the model. The model consists of protein residues that are truncated at the Cβ carbon

atoms and frozen at their crystallographic positions. In reality, the backbone of the protein is relatively stable than the side chain atoms. Thus, truncating the protein residues at Cβ carbon atoms and freezing them at their crystallographic positions may not reproduce the protein environment required for the reaction.

## 4.5 Model 2

### 4.5.1 Description of the Model and Computational Procedure

In order to increase the flexibility of protein residues, we increased the size of the model. In Model 2 (Figure 4.5), all protein residues were truncated at the Cα carbon atoms. The Cα carbon atoms of all residues were represented by a methyl groups and were frozen at the crystallographic positions. In addition to this, we also included Arg228 and Ile96 in the model. Arg228 is located close to 6-hydroxyl group, suggesting a possible hydrogen bond between the two moieties. Such an interaction may also affect the orientation as well and the conformation of the mannopyranosyl ring in the binding site. Furthermore, all atoms in Trp95 in Model 1 were fixed to their crystallographic positions. This was done to avoid movement of Trp95 from its original position that may result from the conformational changes in the substrate. However, a careful analysis of the X-ray structure of dGMII revealed that Ile96 is positioned just above the aromatic ring of Trp95. This arrangement probably locks Trp95 in such a way that it remains stacked against the mannopyranosyl ring in the substrate during the course of the reaction.

The reaction coordinates chosen for the PES scan were same as that in the PES scan of Model 1. All optimizations were performed at DFT/B3LYP/LACVP* level of

theory. During the PES scan for the gycosylation step, $r_1$ was changed from 3.28 Å to 1.48 Å by 0.2 Å. The second reaction coordinate $r_2$ was decreased from 2.15 Å to 0.85 Å by 0.2 Å.

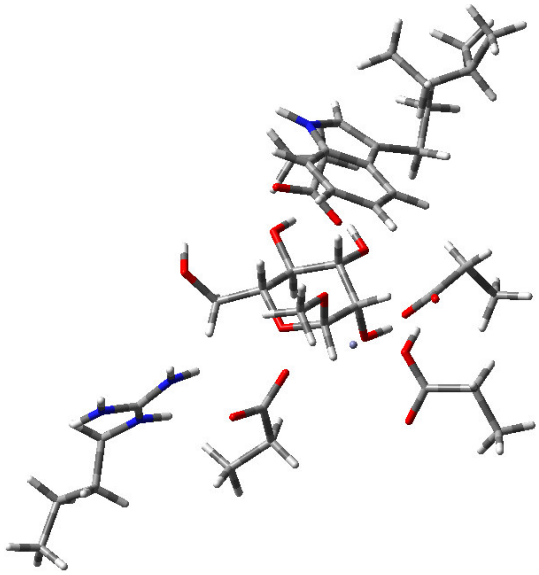


**Figure 4.5.** Model 2 used for the quantum mechanical study of the catalytic mechanism of GMII. The model consists of 130 atoms and the over all charge on the model is 0. The model includes the substrate, 1-methyl mannopyranoside; $Zn^{+2}$ ion; four aspartic acid residues (Asp92, Asp204, Asp341, and Asp472); Trp95; Ile96 and Arg228.

### 4.5.2 Results and Discussions

Unlike Model 1, the PES scan of Model 2 displayed approximate location of MC, CINT and two transition states located very close to each other (Figure 4.6). In the case of Model 2, the sugar ring adopts $B_{2,5}$ conformation in the Michaelis complex MC. Similar conformation was also observed in the case of Model 1. All interactions made by different functionalities in the substrate are similar to those observed in Model 1. The inclusion of Ile96 proved to be important, as this residue seem to facilitate the stacking

interaction of the substrate with Trp95, by avoiding the movement of Trp95 away from the sugar ring. Arg228, however, does not form a hydrogen bond with 6-hydroxyl group.

Unlike Model 1, it was possible to identify two transition states for the glycosylation step. Both transition states, when optimized further, converged to a single structure (TS1), suggesting that there is only one reaction pathway. However, the calculated reaction barriers are very high (61.5 kcal $mol^{-1}$) for either structure. The transition state conformation is slightly twisted $B_{2,5}$ boat conformation and lies between an ideal $B_{2,5}$ and an ideal $^{1}S_5$ skew boat conformation. In this conformation the distance $r_1$ is 1.97 Å. Furthermore, the glycosidic bond is completely cleaved in the transition state suggesting the dissociative nature of the transition state as proposed by experimental studies.[29]

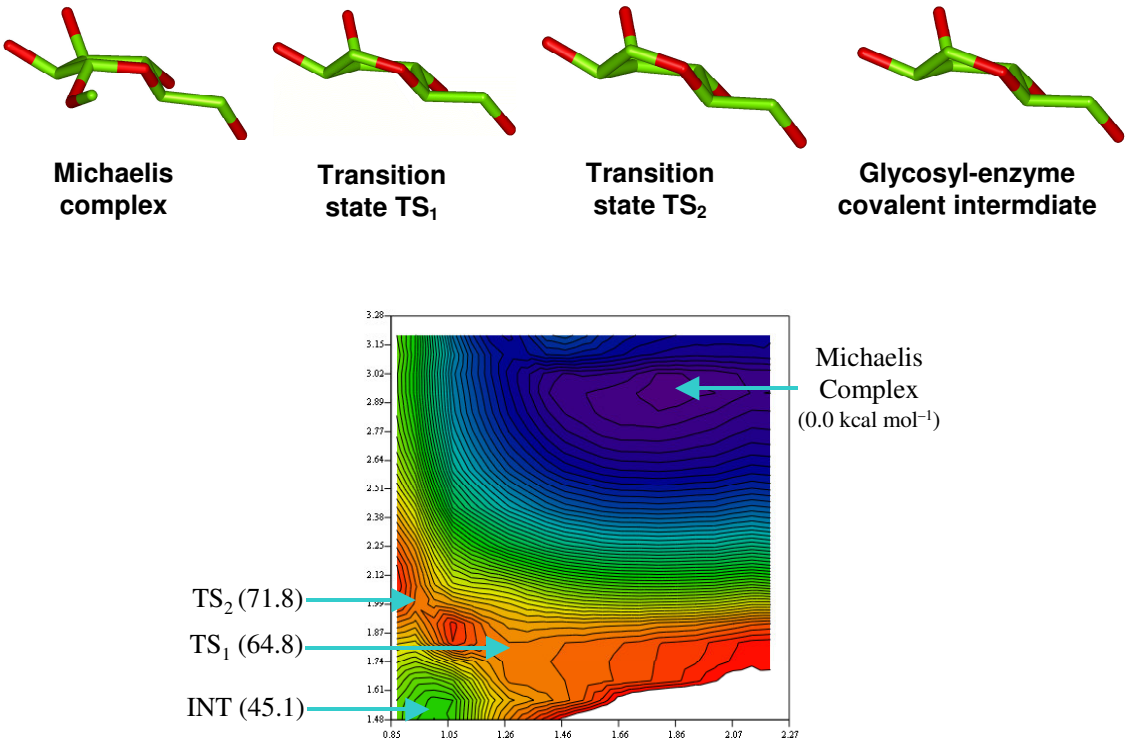



**Figure 4.6.** Potential energy surface obtained from the two-dimensional scan of Model 2 at DFT/B3LYP/LACVP* level (purple – lowest energy region, white – highest energy

region). The Y-axis of the plot represents the distance between the anomeric carbon and the catalytic nuceluophile, whereas the x-axis represents the distance between the proton from general acid catalyst and the anomeric oxygen. The approximate locations of the Michaelis complex, glycosyl-enzyme covalent intermediate and two transition states that are close to each other could be identified. However, computed reaction barriers were very high. Optimized structures of the mannopyranose ring in the Michaelis complex, transition states and the covalent intermediates are shown at the top. Optimization of two transition state structure converged into a single structure suggesting that only one reaction pathway is possible. This pathway is a function of the distance between the anomeric carbon and the catalytic nucleophile.

Interestingly, in the optimized structure of $TS_1$, the carboxylate oxygen of Asp472 abstracts the proton from 3-hydroxyl group following a strong ionic linkage formation between the $Zn^{+2}$ ion and the 2-hydroxyl oxygen. This may be the result of an incomplete coordination sphere for the $Zn^{+2}$ ion. In the X-ray structure of dGMII complexed with swainsonine, the $Zn^{+2}$ ion is coordinated to Asp92, Asp204, His90 and His471. In addition to this, the *cis*-diol moiety in swainsonine coordinates with the $Zn^{+2}$ ion resulting in T6 geometry of coordination for the $Zn^{+2}$ ion. In the case of Model 2, the histidine residues, which coordinate with the $Zn^{+2}$ ion, were not included. This may result in larger partial charge on zinc and stronger attraction between the $Zn^{+2}$ ion and the functionalities coordinating with zinc. Such an interaction would result in tighter binding of the substrate, which may prevent the substrate from leaving the binding site following the hydrolysis reaction. The optimized structure of CINT, which is 41.4 kcal $mol^{-1}$ higher in energy than the MC, displayed similar proton abstraction.

The PES scan for Model 2 revealed that increase in the size affected the energy profile when compared with Model 1. While it was possible to identify the transition state for the glycosylation step and the glycosyl-enzyme covalent intermediate on the two-dimensional PES for Model 2, the reaction barrier and the reaction energy were too high for the reaction to be feasible under physiological conditions. Furthermore, abstraction of

proton from 3-hydroxyl oxygen by Asp472 suggests that complete coordination sphere of the $Zn^{+2}$ ion must be included in the calculations.

## 4.6 Model 3

### 4.6.1 Description of the Model and Computational Procedure

This model consists of His90 and His471 in addition to all residues in Model 2 (Figure 4.7). Introduction of these two histidine residues completes the coordination sphere for the $Zn^{+2}$ ion. Furthermore, we performed one-dimensional relaxed geometry scan where we chose the distance between the anomeric carbon and the nucelophilic oxygen as the reaction coordinate. This decision was based on the fact that the PES for both models previously studied did not display multiple reaction pathways for the glycosylation reaction. The observed reaction pathway in either case could be represented as the function of the distance between the anomeric carbon and the catalytic nucleophilic oxygen of Asp204. Thus, to avoid unnecessarily long computer time we have performed one-dimensional scan along this distance.

The geometry scan was performed at DFT/B3LYP/6-31G**. The $Zn^{+2}$ ion was represented by LACVP**, which includes the effective core potential.[184] During the scan, the distance between the anomeric carbon and the catalytic nucleophile is decreased from 3.91 Å to 1.41 Å by 0.2 Å. The approximate geometries of MC, TS1 and CINT were further optimized at the same level.

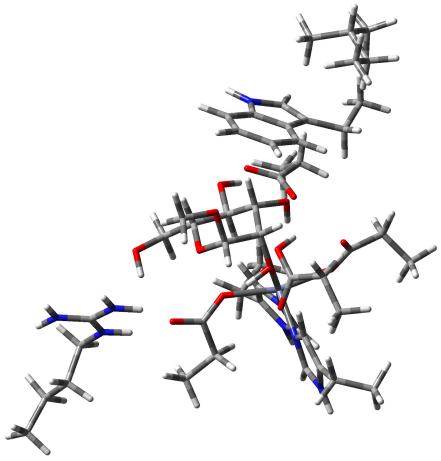


**Figure 4.7.** Model 3 used for the quantum mechanical study of the catalytic mechanism of GMII. The model consists of 160 atoms and the over all charge on the model is 0. The model includes the substrate, 1-methyl mannopyranoside; $Zn^{+2}$ ion; four aspartic acid residues (Asp92, Asp204, Asp341, and Asp472); Trp95; Ile96, Arg228, His90 and His471.

### 4.6.2 Results and Discussio

The energy profile obtained from the one-dimensional scan for Model 3 is depicted in Figure 4.8. It was possible to identify the MC, TS1 and CINT from the one-dimensional energy profile. As the geometry scan progressed, the conformation of the mannopyranosyl ring changes from $^4C_1$ to $B_{2,5}$ at the lowest energy structure. This structure was further optimized as the Michaelis complex which resulted in slightly twisted $B_{2,5}$ boat conformation of the mannopyranose ring in the substrate. In this conformation, the 2- and 3- hydroxyl oxygen atoms are coordinated with the $Zn^{+2}$ ion. Furthermore, interactions observed in previous models were also retained in the optimized structure of the Michaelis complex MC.

The calculated barrier for the reaction is 46.5 kcal $mol^{-1}$ at DFT/B3LYP/6-31G** level. In the transition state TS1, the mannopyranosyl ring adopts a conformation close to

$^1S_5$ skew boat conformation. Interestingly, the proton abstraction observed in Model 1 and Model 2 studies was not observed in Model 3. The proton of 3-hydroxyl group forms a hydrogen bond with Asp472. Furthermore, the dihedral angle between C-5, O-5, C-1 and C-2 atoms is 2.2°, suggesting a strong double bond character in the bond between O-5 and C-1 atoms. Such an arrangement supports the maximum delocalization of the positive charge developed at the anomeric carbon during the formation of the transition state.

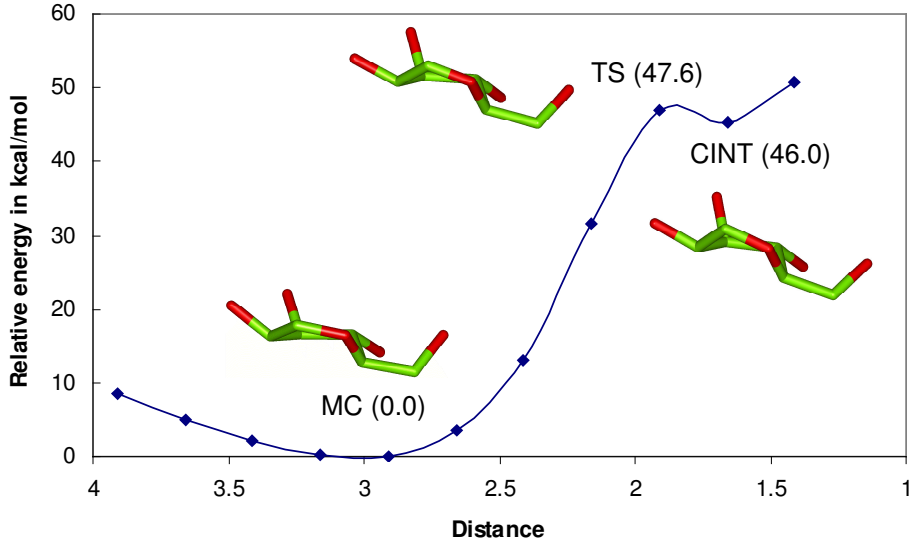



**Figure 4.8.** Potential energy profile obtained from the one-dimensional scan of Model 3 at DFT/B3LYP/6-31G** level (zinc is represented by LACVP** basis set) The approximate locations of the Michaelis complex, glycosyl-enzyme covalent intermediate and transition state could be identified. These approximate geometries were further optimized at same level. The geometries of the mannopyranose rings in the optimized structures have been shown.

The optimized structure of CINT displayed a skew boat conformation for the sugar ring, which lies between $B_{2,5}$ and $^1S_5$ conformations. X-ray structures of the wild type and a mutant dGMII in which the general acid/base catalyst has been removed, with

fluorinated sugar analogs have revealed that the glycosyl covalent intermediate adopts a $^1S_5$ skew boat conformation.[38] The optimized structure of glycosyl-enzyme covalent intermediate CINT provides an excellent opportunity to compare the bound conformation of the fluorinated sugar with that of the computed CINT conformation. Superimposition of two structures led to comparable geometrical parameters (Figure 4.9). However, the covalent bond between the catalytic nucleophile and the anomeric carbon is longer than that observed in the X-ray structure.

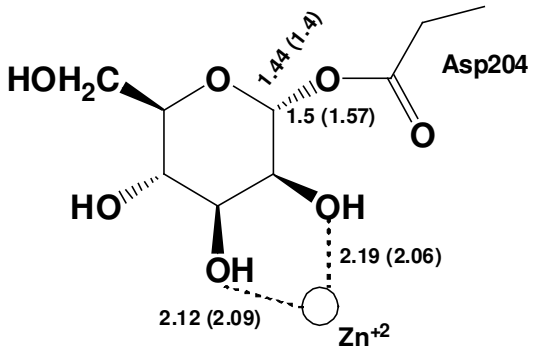



**Figure 4.9.** Comparison of the optimized structure of CINT with the X-ray structure of trapped covalent intermediate.

We have computed the Wiberg bond indices[185] for MC, TS1 and CINT (Figure 4.10), which shows the change in the bond order of O-5–C-1 bond and O-1–C-1 bond as the glycosylation progresses. An increased bond order for O-5–C-1 bond at TS1 suggests increased double bond character in this particular bond. The bond order of this bond decreases from 1.2 in TS1 to 0.99 in CINT indicating formation of an $sp^3$ hybridized carbon atom following formation of the glycosyl-enzyme covalent intermediate CINT. The natural population analysis (NPA) using NBO program[186] also suggested

delocalization of the positive charge between the endocyclic oxygen and the anomeric carbon (Figure 4.11).

While the computed energy barrier for the glycosylation step is still higher, the inclusion of two histidine residues, which coordinate with the $Zn^{+2}$ ion, seems to prevent abstraction of the proton from 3-hydroxyl by Asp472. The higher barriers could be attributed to the lack of solvent effects, as Model 3 does not include water molecules. It is possible that water molecules present in the binding site stabilize the charge developed in the transition state, thus, reducing the activation barrier for the glycosylation step. For this reason we have investigated the catalytic mechanism of GMII using a model that included four water molecules in the binding site.

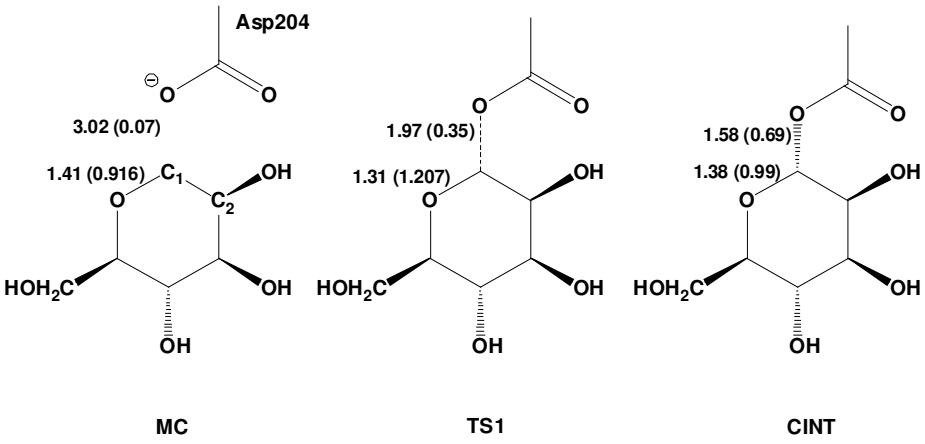



**Figure 4.10.** Bond order analysis for MC, TS1 and CINT obtained from QM study of Model 3. The bond distances are given for the O-5–C-1 bond and the bond between the anomeric carbon and the catalytic nucelophilic oxygen. Wiberg bond indices are given in the parenthesis. The increased sp2 character in the O-5–C-1 bond in TS1 is clear from the reduced length of the bond and increased bond order.

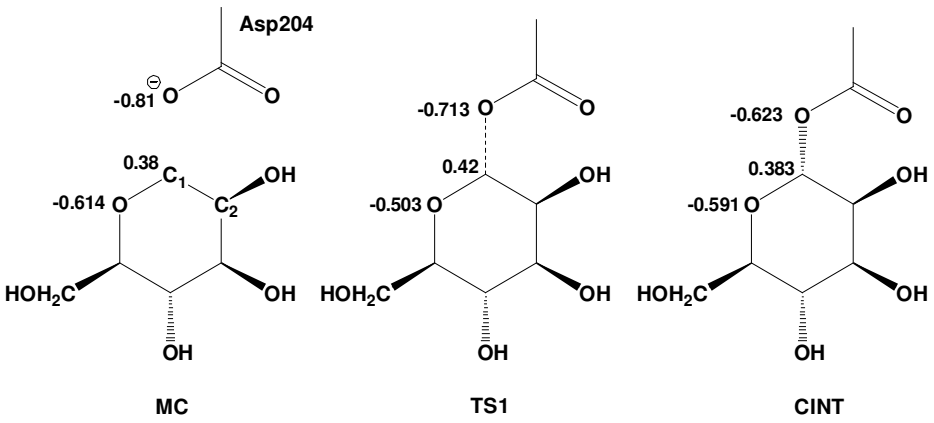



**Figure 4.11.** Natural population analysis for MC, TS1 and CINT obtained from QM study of Model 3. Increased charge on the ring oxygen indicates delocalization of the positive charge developed during the formation of the transition state.

## 4.7 Model 4

### 4.7.1 Description of the Model and Computational Procedure

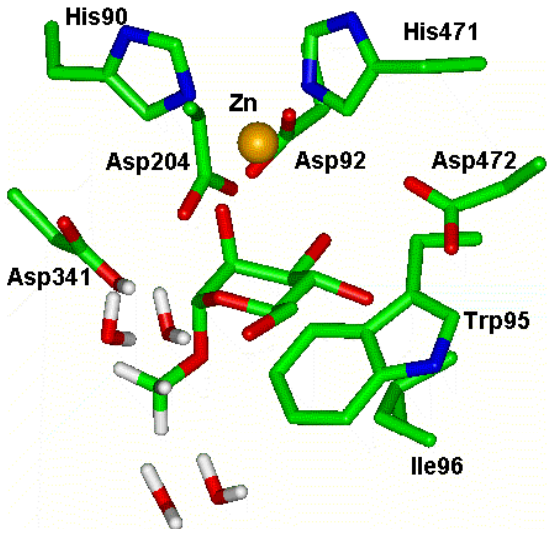



**Figure 4.12.** Model 4 used in the quantum mechanical (DFT/B3LYP/6-311++G**// DFT/B3LYP/6-31G**) analysis of the mechanism of hydrolysis catalyzed by GMII. Protein residues have been truncated at Cα carbon, which are represented by methyl groups. The substrate is mimicked by 1-methyl mannopyranoside.

The fourth model used for investigating the catalytic mechanism of GMII consists of 150 atoms (Table 1, Figure 4.12). The overall charge on the model is –1. This model contains four water molecules, which are conserved in the X-ray structures of swainsonine and mannostatin A complexed with dGMII. These water molecules were included to compute the effect of water on relative energies of all reactive species during the cource of the hydrolysis reaction. The wate molecules obtained from the X-ray structures do not contain hydrogen atoms. Hence, in the case of model 4, GWH utility in AMBER suit[97] of programs was sused to add hydrogen atoms to the water molecules. This program adds hydrogen atoms depending on the electronic environment of water molecules. The computational protocol used for this model is the same as that used in the QM study of Model 3. In the case of the glycosylation step, the distance ($r_1$) between the anomeric carbon and the catalytic nucelophilic oxygen Asp204 OD1, which mimics the formation of a covalent bond between these two atoms, is considered as the reaction coordinate. During the scan, distance $r_1$ was decreased from 3.15 Å to 1.35 Å by 0.2 Å. For the deglycosylation step, the distance ($r_2$) between the anomeric carbon and the oxygen of the closest water molecule, which mimics the attack of water on CINT, is chosen as the reaction coordinate. All optimizations were performed at DFT/B3LYP/6-31G** level. Similar to Model 3, zinc was represented by LACVP** basis set which include effective core potential. Optimized structure of the glycosyl-enzyme covalent intermediate CINT is taken as the starting geometry for the relaxed geometry scan for the deglycosylation step.

### 4.7.2 Results and Discussion

The calculated energy profiles for the glycosylation and the deglycosylation steps are presented in Figure 15. For the glycosylation step, the calculated barrier is 15.8 kcal $mol^{-1}$ at B3LYP/6-311++G** level. The mannopyranose ring adopts an almost perfect $^{1}S_{5}$ skew boat conformation at the Michaelis complex MC where atoms C-2, C-3, C-4 and O-5 are in one plane with the dihedral angle of –1.9°. In this case, the anomeric carbon is above while the C-5 atom is below the plane made by these four atoms. In the Michaelis complex MC, the mannopyranosyl ring is stacked against the Trp95 aromatic ring, a feature common in several carbohydrate-protein complexes.[150-152] The $Zn^{+2}$ ion in the active site coordinates with Nε2 nitrogen atoms of His90 and His471, and the carboxylate oxygens of Asp204 and Asp92. Furthermore, 2- and 3- hydroxyl groups coordinate with the $Zn^{+2}$ ion present in the active site, resulting in $T_6$ type of coordination geometry for $Zn^{+2}$ ion. Such coordination geometry has also been reported in many ligand-dGMII complexes where the ligands possessed a *cis*-diol moiety that coordinates with the zinc.[38, 75, 80, 81] The orientation of the mannopyranosyl ring is further fixed by a strong hydrogen bond between 4-hydroxyl group and Asp472OD1 (1.8 Å between the hydroxyl hydrogen and carboxylate oxygen). In this conformation, the anomeric oxygen forms a very strong hydrogen bond (1.61 Å) with the proton from general acid Asp341, suggesting that this interaction may facilitate protonation of the leaving group and its subsequent departure.

The conformation of the transition state TS1 lies between an ideal $^{1}S_{5}$ skew boat conformation and an ideal $B_{2,5}$ boat conformation with the dihedral angle of 9.6° between atoms C-2, C-3, C-4 and O-5, suggesting a minor conformational change in the sugar ring form its conformation in the Michaelis complex. The bond between C-1 and O-1 atoms is

completely cleaved with these atoms separated by 2.38 Å. Furthermore, atoms C-5, O-5, C-1 and C-2 are co-planar with a dihedral angle of –0.5°, a condition which supports the maximum proposed $sp^2$ character on the anomeric carbon and the ring oxygen during the formation of the transition state. The 2- and 3-hydroxyl groups remain coordinated to Zn+2 ion at 2.18 Å and 2.14 Å, respectively, whereas the 4-hydroxyl group remained hydrogen bonded to carboxylate oxygen of Asp472.

The conformation of the mannopyranosyl ring further changes at the glycosyl-enzyme covalent intermediate (CITN) where it is closer to an ideal $B_{2,5}$ boat conformation. The relative energy of CINT is 4.8 kcal $mol^{-1}$ at B3LYP/6-311++G** level when compared with MC. Interestingly, the nucleophilic oxygen Asp204OD1 moves away from the $Zn^{+2}$ ion after the formation of covalent linkage between the anomeric carbon and Asp204OD1 (Table 2). This movement is consistent with the X-ray structures of trapped covalent intermediates for dGMII and may be a result of decreased negative charge on the nucleophilic oxygen following the covalent bond formation.

For the deglycosylation step, the calculated barrier is relatively high (30.0 kcal $mol^{-1}$) at DFT/B3LYP/6-311++G** level. The transition state TS2 conformation lies between an ideal $B_{2,5}$ boat conformation and $^{O}S_2$ skew boat conformation. Interestingly, the dihedral angle between C-5, O-5, C-1, and C-2 is –24.5° as compared with –0.5° in TS1, indicating that these atoms are not co-planar as expected for an oxacarbenium ion-like transition state. Interactions made by 2-, 3- and 4-hydroxyl groups with protein residues were similar to those observed in TS1 and CINT.

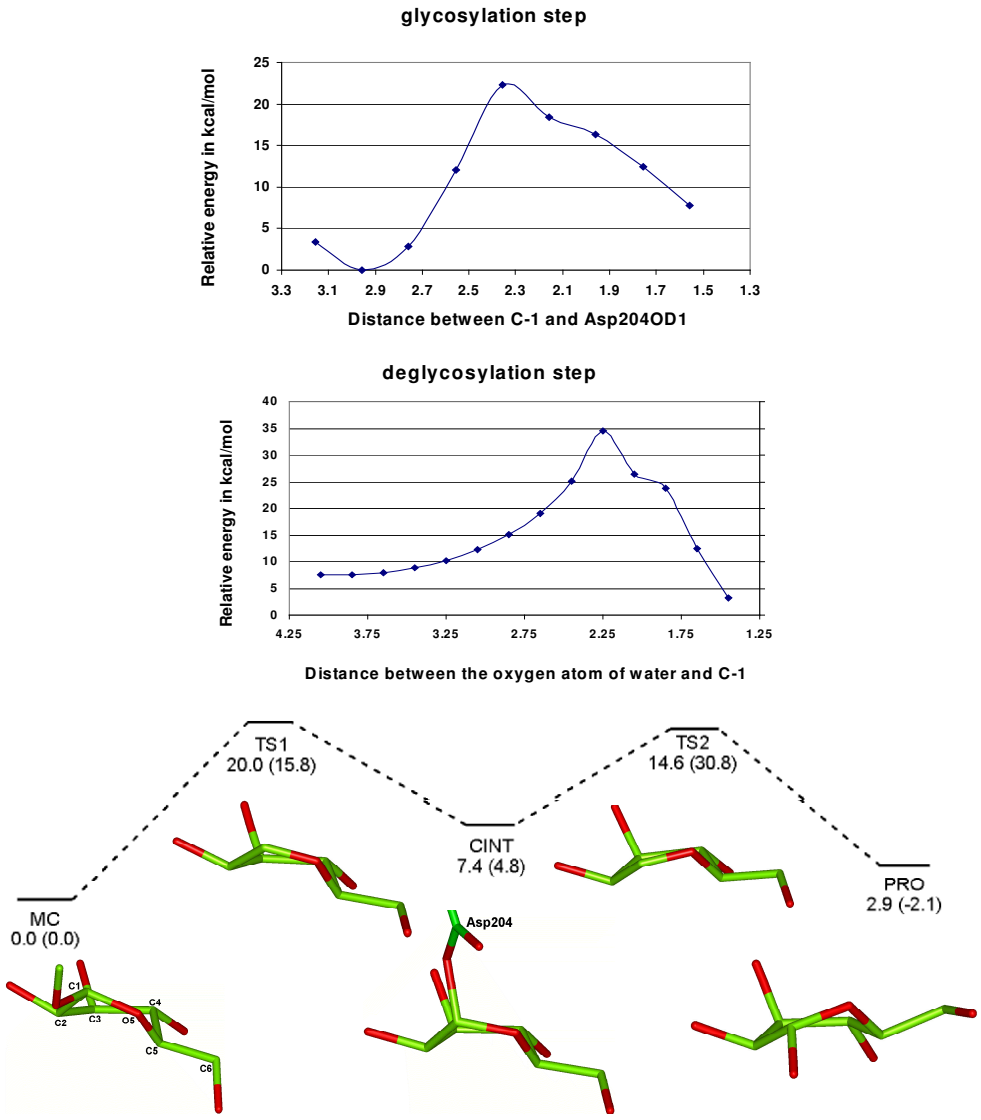



**Figure 4.13.** Energy profile and schematic representation of relative energies (in kcal $mol^{-1}$, DFT/B3LYP/6-31G** // DFT/B3LYP/6-311++G**) for the glycosylation and deglycosylation steps involved in the hydrolysis reaction catalyzed by GMII. Values in parenthesis are the single point energies at DFT/B3LYP/6-311++G**. The conformation of the mannopyranosyl ring in the Michaelis complex is $^{1}S_{5}$. At the transition state the conformation lies between an ideal $^{1}S_{5}$ and $B_{2,5}$ boat conformation, with atoms C-5, O-5, C-1 and C-2 in one plane as expected for an oxacarbenium ion-like transition state. In the case of CINT the conformation was found to be slightly distorted $B_{2,5}$ boat and was comparable to that in the trapped covalent intermediate (PDB id 1QWN). For the deglycosylation step, the transtion state conformation lies between $B_{2,5}$ and $^{O}S_{2}$ skew boat conformation while in the hydrolyzed product it was found to be very close to an ideal $^{O}S_{2}$

The hydrolyzed product displayed a conformation very close to an ideal $^{O}S_2$ conformation with relative energy of – 2.1 kcal mol$^{-1}$ when compared with MC. In this structure, unlike CINT, the nucleophilic carboxylate oxygen Asp204OD1 again moves close to the $Zn^{+2}$ ion and forms a $T_6$ type of geometry. Furthermore, the 4-hydroxyl group remained hydrogen bonded to carboxylate oxygen of Asp472.

**Table 4.2.** Computed (DFT/B3LYP/6-31G**) bond lengths for different geometries observed in the glycosylation step.

| | Distances in Å | | | | | |
|---|---|---|---|---|---|---|
| | $r_1$ | $r_{C1-O1}$ | $r_{O2-Zn}$ | $r_{O3-Zn}$ | $r_{204OD1-Zn}$ | C1-O5 |
| Starting geometry | 3.70 | 1.41 | 2.18 | 3.63 | 2.13 | 1.41 |
| MC | 3.11 | 1.42 | 2.19 | 2.12 | 2.07 | 1.43 |
| TS1 | 2.42 | 2.38 | 2.18 | 2.14 | 2.14 | 1.28 |
| CINT | 1.49 | 3.02 | 2.12 | 2.11 | 2.66 | 1.39 |
| TS2 | 2.16[a] | 3.02 | 2.21 | 2.15 | 2.09 | 1.29 |
| PRO | 1.42[a] | 3.57 | 2.2 | 2.11 | 2.08 | 1.41 |

[a] Distance is between C-1 and the oxygen of attacking water molecule.

Analysis of important distances (Table 4.2) suggests that significant conformational changes occur during early stages of glycosylation step where the sugar donor undergoes distortion from the initial $^{4}C_1$ conformation, resulting in $^{1}S_5$ conformation in MC. It is known that glycosidases distort the substrate to yield the Michaelis complex structure. First proposed for the hen egg white lysozyme,[20] such substrate distortion has also been unambiguously observed in high-resolution X-ray structures of Michaelis complexes for two structurally unrelated endoglucanases (from glycosyl hydrolase families 5 and 7),[21, 22] and hexosaminidase (family 20).[23]

Furthermore, X-ray structure of Michaelis complex for Mannase 26A, which is a retaining endo-$\beta$-(1-4)- mannase, displayed $^{1}S_5$ conformation of the substrate.[24] During the initial optimization, similar distortion was observed in the mannopyranoside ring resulting in a $B_{2,5}$ conformation. This change in the conformation seems to be mainly controlled by the interaction between the C-3 hydroxyl oxygen and the $Zn^{+2}$ ion. The distance between these atoms decreases from 3.63 Å in $^{4}C_1$ conformation to 2.12 Å in the MC ($^{1}S_5$ conformation). Furthermore, in this conformation, C-2 hydroxyl group moves in pseudo-equatorial orientation as against axial orientation in $^{4}C_1$ conformation, thereby reducing the activation energy for the formation of the covalent glycosyl-enzyme intermediates. The distance between O-2 and $Zn^{+2}$ ion, however, does not change (2.18 Å in $^{4}C_1$ to 2.19 Å in $^{1}S_5$ conformation) during this transformation. Thus, it appears that the coordination of C-2 and C-3 hydroxyl oxygen atoms with the $Zn^{+2}$ ion controls the conformational itinerary in early stages of the hydrolysis reaction.

The role of $Zn^{+2}$ ion in GMII has been recently proposed by Kuntz et al. based on the co-crystal structure of noeuromycin bound to dGMII[80] and previous crystallographic studies.[38, 75, 80, 81] In addition to maintaining the deprotonated state of the catalytic nucleophile Asp204, it was proposed that $Zn^{+2}$ ion drives the conformational itinerary of the substrate with the help of the interaction between C-2 and C-3 hydroxyl groups, and $Zn^{+2}$ ion. Our calculations support this hypothesis. Furthermore, in $^{4}C_1$ conformation, O-2 is coordinated with $Zn^{+2}$ ion resulting in $T_5$ geometry (square pyramidal) of zinc coordination sphere. This type of coordination geometry is one of the most common geometries observed in zinc binding proteins and was also proposed to be favorable compared to the $T_6$ octahedral geometry.[78] Substrate distortion driven by O-2 and O-3

interaction with $Zn^{+2}$ ion, thus, helps achieving otherwise unfavorable $T_6$ coordination geometry. Moreover, natural population analysis (Table 4) reveals that the charge on zinc remains unchanged during the progress of reaction suggesting that the role of zinc is mainly structural and coordination of substrate with $Zn^{+2}$ ion is essential for driving the conformational itinerary towards the transition state.

**Table 4.3.** Computed (DFT/B3LYP/6-31G**) dihedral angles for the mannopyranosyl ring along the reaction coordinate.

| | Dihedral angle in ° | | | | | |
|---|---|---|---|---|---|---|
| | C5-O5-C1-C2 | O5-C1-C2-C3 | C1-C2-C3-C4 | C2-C3-C4-C5 | C3-C4-C5-O5 | C4-C5-O5-C1 |
| Starting geometry | -59.6 | 55.4 | -55.0 | 53.4 | -54.5 | 59.3 |
| MC | 11.4 | -50.8 | 26.5 | 31.2 | -70.9 | 49.5 |
| TS1 | -0.5 | -44.6 | 28.4 | 23.8 | -65.1 | 55.5 |
| CINT | 2.5 | -53.0 | 41.8 | 14.5 | -63.5 | 57.0 |
| TS2 | -24.5 | -27.1 | 32.5 | 9.5 | -57.8 | 70.4 |
| PRO | -38.3 | -15.9 | 36.2 | -6.4 | -41.1 | 68.9 |

It is thought that the transition state structures on either side of the covalent intermediate of retaining glycosidases have significant oxacarbenium ion character.[25-29] It has also been proposed that the transition state for the glycosylation step is 'dissociative' in nature suggesting a significant cleavage of the scissile bond.[29] The computed structure of the transition state is consistent with these experimental observations. In the optimized structure of TS1, atoms C-5, O-5, C-1 and C-2 are planar (dihedral angle of –0.5°) and the bond between C-1 and O-5 shortens to 1.28 Å from 1.43 Å in the Michaelis complex as a result of delocalization of the positive charge developed during the formation of the transition state. Indeed, NPA of TS1 reveals increased charge on O-5, thus indicating the delocalization of charge between C-1 and O-5. In order to get a quantitative measure of

the extent of bond-formation or bond-breaking process in the glycosylation step, we performed bond order analysis on the optimized geometries of MC, TS1 and CINT. Wiberg bond indices[185] have been calculated (Table 4.5) for the four bonds involved the transformation using NBO program[186] as implemented in Jaguar.[111] An increased bond order of C-1–O-5 bond (from 0.88 in MC to 1.38 in TS1) confirms an induced double bond ($sp^2$) character. In the case of CINT, the bond order for this bond decreases to 0.97 indicating the anomeric carbon becomes $sp^3$ hybridized, following the formation of covalent linkage between C-1 and Asp204OD1.The scissile glycosidic bond is completely cleaved in TS1 suggesting the dissociative nature of the transition state.

**Table 4.4.** Charge distribution along the reaction coordinate calculated using natural population analysis at DFT/B3LYP/6-31G**.

| | C-1 | O-5 | O-1 | Zn |
|---|---|---|---|---|
| MC | 0.383 | -0.634 | -0.624 | 1.64 |
| TS1 | 0.469 | -0.445 | -0.803 | 1.646 |
| CINT | 0.38 | -0.591 | -0.827 | 1.645 |
| TS2 | 0.46 | -0.47 | -0.824 | 1.643 |
| PRO | 0.38 | -0.61 | -.803 | 1.643 |

**Table 4.5.** Bond order analysis performed with NBO program using Wiberg bond indices.

| Bond | MC | TS1 | CINT | TS2 | PRO |
|---|---|---|---|---|---|
| C1–O1 | 0.88 | 0.13 | 0.01 | 0.01 | 0.0 |
| C1–O5 | 0.89 | 1.33 | 0.97 | 1.27 | 0.91 |
| C1–OD1 | 0.01 | 0.13 | 0.77 | 0.07 | 0.0 |
| C1–$O_{Wat}$ | 0.0 | 0.0 | 0.0 | 0.25 | 0.90 |

Analysis of the pseudorotational wheel for pyranose ring suggests that there are only four possible conformers that can satisfy the condition of the planarity of C-5, O-5, C-1 and C-2. They are $^{3}H_{4}$, $^{4}H_{3}$, $^{2,5}B$ and $^{1}S_{5}$. Indeed, the transition states for the glycosylation steps of family-5 and 7 retaining celluloses[21, 22], and family-11 xylanases[23] have been proposed to adopt $^{4}H_{3}$ and $^{2,5}B$ conformations, respectively. Furthermore, $B_{2,5}$ transition state has also been suggested for $\beta$-mannase 26A.[24] In the case of class I α-mannosidase, the transition state is proposed to adopt $^{3}H_{4}$ conformation.[187] However, our calculations show that in the case of glycosylation step, the substrate can adopt $^{1}S_{5}$ skew boat conformation while still maintaining the planarity among C-5, O-5, C-1 and C-2 atoms. This conformation seems to be a result of the hydrogen bond formation between the anomeric oxygen and the proton from the general acid Asp341. This interaction also facilitates the protonation of the leaving group by general acid Asp341.

While $^{1}S_{5}$ skew boat conformation in MC facilitates the protonation of the leaving group by the formation of a hydrogen bond between glycosidic hydrogen and the proton from Asp341, the conformation of the mannopyranose ring in CINT seems to assist with the deglycosylation step. In this conformation the leaving group is placed anti-periplanar to the lone pair of the ring oxygen, a requirement for the departure of the leaving group according to Deslongchamp's anti-periplanar lone-pair hypothesis.[79] Furthermore, steric clashes between the *syn*-hydrogens at C-3 and C-5, and the attacking water molecule are minimal in the $^{1}S_{5}$ skew boat conformation. This facilitates the attack of water leading to deglycosylation of the glycosyl-enzyme covalent intermediate.

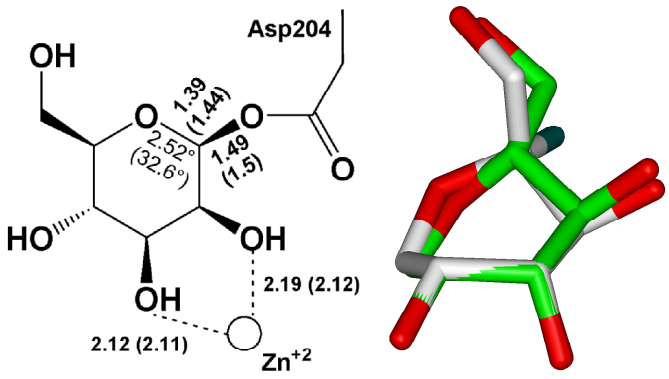



**Figure 4.14.** Comparison of the geometrical parameters and superimposition of glycosyl-enzyme covalent intermediate obtained from PDB 1QWN (white) and optimized structure of glycosyl covalent intermediate (green). While a good correlation can be seen in different functionalities on the sugar rings, the conformations of the ring are slightly different. The values in parenthesis are for the X-ray structure.

X-ray structures of the wild type and a mutant dGMII in which the general acid/base catalyst has been removed, with fluorinated sugar analogs have revealed that the glycosyl covalent intermediate adopts a $^{1}S_{5}$ skew boat conformation.[38] With optimized structure of glycosyl-enzyme covalent intermediate in our hand, we compared the bound conformation of the fluorinated sugar with that of the computed CINT conformation. Superimposition of two structures led to staggering similarities (Figure 4.14), with many of the geometrical parameters comparable.

The transition state for the deglycosylation was also expected to be similar to the oxacarbenium ion. Interestingly, the dihedral angle between atoms C-5, O-5, C-1, and C-2 was found to be –24.5° as compared with –0.5° in TS1, indicating the transition state for the deglycosylation step may have a reduced oxacarbenium ion character in the C-1–O-5 bond. However, the bond order analysis revealed an increased bond order of C-1–O-5 bond (from 0.97 in CINT to 1.27 in TS2) suggesting that a double bond character is induced in that bond during the formation of TS2. Furthermore, the bond length of C-1–

O-5 also reduces from 1.39 Å in CINT to 1.28 Å in TS2, which is an indicative of charge delocalization between the two atoms. This is also supported by an increased charge on the ring oxygen (from –0.59 in CINT to –0.47 in TS2, Table 4.4). This charge again decreases to –0.61 in the hydrolyzed product. The bond between C-1 and O-5 atoms also lengthens to 1.41 with decreased bond order of 0.91.

Ducros et al.[24] recently solved X-ray structures of Michaelis complex and glycosyl-enzyme covalent intermediate for β-mannase, a family 26 β-mannosidase, and proposed conformational itinerary for the glycosylation step as ${}^{1}S_{5} \rightarrow B_{2,5} \rightarrow {}^{O}S_{2}$. In this case, the enzyme transforms a β-mannosyl species into an α-mannoside. Similar transformation is achieved by GMII in the deglycosylation step suggesting that the conformational changes during the deglycosylation step must be similar to those in the glycosylation step in β-mannase. Indeed, computed itinerary for the deglycosylation step is consistent with this similarity proposed by Numao et al.[38]

In the glycosylation step, as the distance between C-1 and the nucleophilic carboxylate oxygens decreases, the mannopyranosyl ring moves closer to Asp204 (Figure 4.15). The extent of this movement is highest for O-5, which moves 1.5 Å from its position in MC. The C-5 carbon atom also moves 1.2 Å from its position in MC. In contrast, C-2 and C-3 move only 0.66 Å and 0.5 Å, respectively, from their positions in MC. This restricted motion is a result of coordination these two atoms with $Zn^{+2}$ ion. This forces the conformation of the sugar ring towards a $B_{2,5}$ boat conformation from an ideal ${}^{1}S_{5}$ skew boat conformation in MC. In the case of deglycosylation step, however, maximum movement is seen in C-1, which moves 1.12 Å from its positioning CINT. This movement brings C-1, C-3, C-4, and C-5 atoms close to be planar (dihedral angle is

–12.9°) where the atom O-5 lies above where as C-2 lies below this plane resulting in a conformation close to an ideal $^{O}S_{2}$ skew boat conformation.

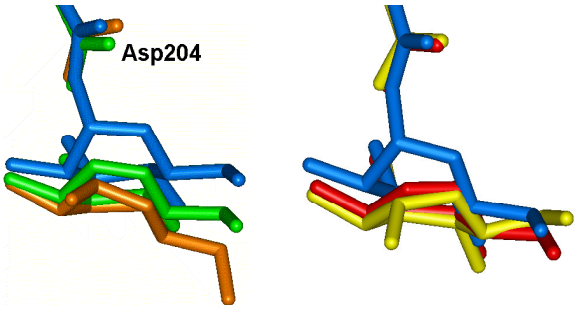



**Figure 4.15.** Superimposition of the active site in MC (orange), TS1 (green), CINT (blue), TS2 (red) and PRO (yellow) reveals the movement of the mannopyranosyl ring in the active site of dGMII.

Potent inhibitors of glycosidases are thought to mimic oxocarbenium ion-like transition states. For example, the inhibitory activity of the natural product swainsonine has long been attributed to its 5-membered ring resembling a flattened six-membered ring, which had been forced to attain an oxocarbenium-like structure. In fact, the crystal structure of swainsonine complexed with dGMII shows the inhibitor to be tilted in such a way as to bring the equivalent of its anomeric carbon close to the presumed catalytic nucleophile.[75] On the other hand, recently solved co-crystal structure of mannostatin A complexed with dGMII revealed that in the bound conformation mannostatin A mimics best the covalent intermediate.[174] Studies presented here provide an excellent opportunity to compare the bound conformations of swainsonine and mannostatin A with the transition state for the glycosylation step, TS1.

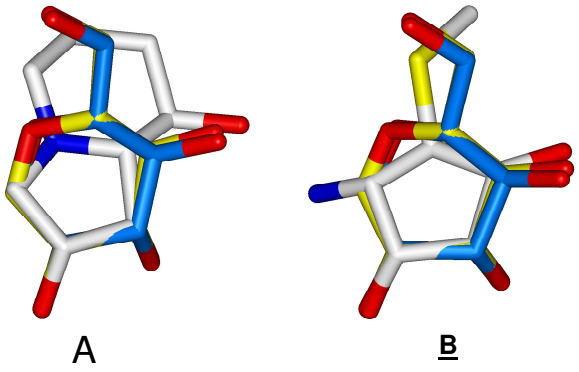



**Figure 4.16.** Superimposition of optimized structures of transition state (TS1, yellow) and glycosyl-enzyme covalent intermediate (CINT, blue) onto bound conformations of swainsonine (A) and mannostatin A (B). Clearly either inhibitor seems to mimic both TS1 and CINT. Interestingly, the ring nitrogen in swainsonine lies near the endocyclic oxygen of the sugar rings, where as the amine nitrogen of mannostatin A lies near the anomeric carbon.

In the case of swainsonine, superimposition of the *cis*-diol onto 2- and 3-hydroxyl groups in TS1 resulted in extremely good overlay (Figure 4.16, RMSD = 0.06) between these hydroxyl functionalities. The 6-hydroxyl group seems to mimic 4-hydroxyl moiety in the mannopyranose ring with the hydroxyl oxygen placed at 1.5 Å from the 6-hydroxyl oxygen in swainsonine. It is believed that the ring oxygen in swaisonine mimics the positive charge developed on the ring oxygen in the actual substrate. Indeed, as can be seen in Figure 4.16, the ring nitrogen in swainsonine lies very close to the ring oxygen of the substrate (1.0 Å) suggesting that this nitrogen, in the protonated state, may mimic the positive charge on the ring oxygen developed during the formation of the transition state. In the case of mannostatin A, unlike swainsonine, the amine nitrogen lies close to the anomeric carbon (1.1 Å), thus possibly mimicking the positive charge build up on the anomeric carbon. The 3- and 4-hydroxyl groups in mannostatin display excellent overlay (RMSD = 0.07) with corresponding *cis*-diol moiety in the TS1. The 2- hydroxyl group in

mannostatin lies in the same region as that of 4-hydroxyl group in TS1 at 1.3 Å from the 4-hydroxyl oxygen of TS1.

The superimposition study performed here supports the belief that swainsonine and mannostain are transition state mimics. It is also clear from Figure 4.16 that that these inhibitors also mimic the glycosyl-enzyme covalent intermediate. This is a result of similarity between the conformations of the mannopyranose rings in TS1 and CINT. Both these conformations lie between an ideal $^{1}S_{5}$ skew boat conformation and $B_{2,5}$ boat conformation, suggesting that the structural differences between them are minor.

While the calculations described in this study involve a reasonably large region of the active site, it is possible that inclusion of more residues in the model may alter the results (geometries and energies) to some extent. For this reason we are currently investigating the effect of protein environments on the results reported here by using mixed quantum mechanical / molecular mechanical (QM/MM) calculations where a small portionof the active site is treated at QM level whereas, the rest of the protein is treated at MM level. Furthermore, we are also performing QM calculations on a bigger model in order to compre results from QM/MM calculations with QM calculations. All these calculations use optimized geometries of MC, TS1, CINT, TS2 and PRO as starting geometries.

## 4.8 Conclusions

Glycosidases are enzymes that play crucial roles in the biosynthesis of glycoproteins. Inhibitors of these enzymes have garnered much attention as lead compounds for drug discovery for diseases such as viral- and bacterial infections,

diabetes, Gauchers disease and cancer. However, efforts to design more potent and/or more selective inhibitors often are limited due to a dearth of knowledge pertaining to the catalytic mechanisms and the mechanism of inhibition by natural inhibitors of these enzymes. Computational methods are, in principle, capable of elucidating the details of an enzymatic catalytic reaction. When combined with experimental observation, such computational insights are significant in understanding the conformational changes that occur in the substrate *en route* to the transition state.

Golgi α-mannosidase II, an α-retaining mannosidase, which acts late in the *N*-glycan biosynthesis pathway, has been validated as a target for blocking tumor growth and metastasis. The enzyme catalyzes a hydrolysis reaction by a double displacement mechanism. Although general aspects of this mechanism has been accepted and refined over the years using experimental techniques, the nature of the transition state and the conformational itinerary of the mannopyranosyl ring in the glycone has been a subject of debate. Furthermore, the knowledge of the transition state structure would also facilitate the efforts of designing potent transition state analogues. In this work, a high-level quantum mechanical study of the hydrolysis reaction catalyzed by GMII has been conducted with the main focus on the conformational analysis of the substrate as the reaction progresses.

In all, four models, which differ in the size and the overall charge, have been investigated. In the first three models, which lack explicit water molecules, the reaction barriers for the glycosylation step were very high (46.5 – 61.5 kcal $mol^{-1}$). For this reason the deglycosylation step was not performed for these models. Although these models did not yield more realistic energy profiles, the results obtained were beneficial for the

refinement of the model. From the QM study of Model 1 – Model 3, it was concluded that all the residues that coordinate with the metal ion present in the active site must be included in the quantum mechanical calculations. Model 3, which includes complete coordination sphere for the $Zn^{+2}$ ion, resulted in more stable complex. Unlike Model 1 and Model 2, this model did not display proton abstraction from 2-hydroxyl group by Asp472.

The best results were obtained for Model 4, where the computed activation barrier for the glycosylation step is 15.8 kcal $mol^{-1}$ at DFT/B3LYP/6-311++G** level. This model includes four water molecules in addition to all the residues in Model 3. In the optimized Michaelis complex, the mannopyranosyl ring in the substrate adopts a perfect $^{1}S_{5}$ skew boat conformation, where C-2, C-3, C-4 and O-5 atoms are in one plane. This conformation seems to facilitate the departure of the leaving group by forming a hydrogen bond between the anomeric oxygen and the proton from general acid catalyst Asp341.

It has been previously proposed that in the transition state the mannopyranose ring adopts one of $^{3}H_{4}$, $^{4}H_{3}$, $B_{2,5}$, and $^{2,5}B$ conformation. Our calculations suggested a slightly twisted $B_{2,5}$ conformation at the transition state that lies between $B_{2,5}$ boat conformation and $^{1}S_{5}$ skew boat conformation. In this conformation, the atoms C-5, O-5, C-1 and C-2 are in one plane suggesting that the transition state resembles mannopyranosyl oxacarbenium ion. Bond order analysis and natural bond order analysis revealed an increased $sp^{2}$ character was observed in the C-1–O-5 bond along with delocalization of positive charge between C-1 and O-5 atoms. At the transitin state the C-

1–O-1 bond is completely cleaved, suggesting the dissociative nature of the transition state for the glycosylation step in accord with the experimental observations.

The optimized structure of CINT revealed that the sugar ring adopts a skew boat conformation that is very close to ideal $^{1}S_{5}$ conformation. This conformation seems to facilitate the deglycosylation step, as in this conformation the leaving group (the catalytic nucleophile) is anti-periplanar to the lone pair in the ring oxygen. Furthermore, comparison of the optimized CINT structure with the trapped covalent intermediate resulted in staggering similarities. The catalytic nucleophilic oxygen moves away from the $Zn^{+2}$ ion. Similar movement of Asp204 was also observed in the X-ray structures of trapped covalent intermediates.

Unlike the glycosylation step, that transition state for the deglycosylation step (TS2) revealed that atoms C-5, O-5, C-1 and C-2 are not co-planar. However, increased $sp^2$ character was observed for the C-1–O-5 bond along with the delocalization of the charge across this bond. In this structure the conformation of the sugar ring is slightly skewed boat conformation that lies between $B_{2,5}$ and $^{O}S_{2}$ skew boat conformation. This conformation of the mannopyranosyl ring changes in the hydrolyzed product to a conformation very close to an ideal $^{O}S_{2}$ skew boat conformation.

The calculations presented here suggest that the $Zn^{+2}$ ion in the active site is essential for substrate binding to the active site, and the coordination 2- and 3- hydroxyl group with the $Zn^{+2}$ ion is the main driving force for the conformational changes in the substrate in the early stages of the hydrolysis reaction. Furthermore, the natural population analysis revealed that the charge on zinc remains unchanged over the course of the reaction. The role of zinc is, thus, structural and not electronic in nature.

It is thought that potent inhibitors of mannosidases mimic the transition state structure of the substrate. We have compared the bound conformation of mannostatin A and swainsonine to the mannopyranosyl ring in the optimized TS1. These superimposition studies revealed that mannostatin A and swainsonine in bound conformation mimic the transition state for the glycosylation step, which has been shown to adopt a conformation that lies between $B_{2,5}$ and $^{1}S_{5}$ skew boat conformation. In particular, the zinc-coordinating 3- and 4-hydroxyls of **1** display a good overlap with the corresponding *cis*-diol of the mannosyl residue. Furthermore, the amine nitrogen of mannostatin A is positioned close to the C-1 carbon of TS1, whereas in the case of swainsonine the amine nitrogen is close to the ring oxygen suggesting that it may mimic the charge build up at the ring oxygen during the formation of the transition state. Collectively these studies indicate that mannostatin A and swainsonine are transition state mimics.

## 4.9 Supporting information

**Table S1.** Optimized geometries of Michaelis complex, transition state and glycosyl-enzyme covalent intermediate for Model 2.

| Michaelis complex (Model 2)<br>$E_{absolute}$ = -2867.34354482303 hartree | | | | Transition state TS1 (Model 2)<br>$E_{absolute}$ = -2867.26110158471 hartree | | | |
|---|---|---|---|---|---|---|---|
| Zn1 | 0.0578635532 | 0.0138279425 | 0.0230321071 | Zn1 | 0.0082149349 | 0.0110415875 | 0.0130147353 |
| C2 | 0.0515042443 | 0.0357764100 | 4.8540410579 | C2 | 0.0101624129 | 0.0148942770 | 4.8440828103 |
| H3 | 0.9044713233 | 0.0027593623 | 4.2083016256 | H3 | 0.8622706602 | -0.0082246727 | 4.1967841512 |
| H4 | 0.1550299828 | -0.7229469291 | 5.6016687275 | H4 | 0.1270339280 | -0.7448653363 | 5.5886847714 |
| H5 | -0.0594791675 | 0.9894437132 | 5.3259185260 | H5 | -0.1634825230 | 0.9572731698 | 5.3197253633 |
| C6 | -1.1703060982 | -0.2304879372 | 3.9444866265 | C6 | -1.1762713906 | -0.3135665265 | 3.8990828625 |
| H7 | -1.8980923411 | -0.9077801681 | 4.3976220426 | H7 | -1.8364538951 | -1.0897528255 | 4.2924839313 |
| H8 | -1.6847674920 | 0.7115157180 | 3.7158660941 | H8 | -1.7879124803 | 0.5820727814 | 3.7277469319 |
| C9 | -0.7390536709 | -0.8004602287 | 2.6056386725 | C9 | -0.7074399322 | -0.7517561077 | 2.5177190347 |
| O10 | 0.1666326211 | -0.1029839191 | 1.9755564348 | O10 | 0.1905449600 | 0.0296573589 | 1.9545924963 |
| O11 | -1.2497481406 | -1.8663089477 | 2.2016156946 | O11 | -1.1931859813 | -1.7674534950 | 1.9962058854 |
| C12 | -6.1590717869 | 0.4847225607 | 4.2401550135 | C12 | -6.2177445584 | 0.0489657134 | 4.2429181649 |
| H13 | -5.7357866497 | 0.6598781338 | 5.2075844842 | H13 | -5.8051864462 | 0.2485877211 | 5.2102428501 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| H14 | -6.3701562938 | 1.4227459648 | 3.7687850721 | H14 | -6.4921449264 | 0.9724376232 | 3.7754371988 |
| H15 | -7.0647116910 | -0.0750599946 | 4.3436978044 | H15 | -7.0836299822 | -0.5706139763 | 4.3459947935 |
| C16 | -5.1693669436 | -0.3348493393 | 3.3831336783 | C16 | -5.1988164552 | -0.6992483416 | 3.3599270329 |
| H17 | -4.8562007580 | -1.2199776228 | 3.9529787801 | H17 | -4.8618674683 | -1.6040698586 | 3.8855968841 |
| H18 | -4.2539829659 | 0.2497095947 | 3.2160739991 | H18 | -4.2972104748 | -0.0839913520 | 3.2275156229 |
| C19 | -5.7309757977 | -0.7899355739 | 2.0590364681 | C19 | -5.7433309344 | -1.0942400871 | 2.0108607459 |
| C20 | -5.3798536566 | -2.0008937476 | 1.3506880100 | C20 | -5.1923854610 | -2.0943025827 | 1.1225105662 |
| C21 | -6.1706764709 | -2.0469292639 | 0.1679443735 | C21 | -6.0160917432 | -2.1415613981 | -0.0355634644 |
| C22 | -4.4790597463 | -3.0568665354 | 1.6104687318 | C22 | -4.0796558884 | -2.9540128050 | 1.2014134559 |
| C23 | -6.6843399637 | -0.1649897640 | 1.2875472512 | C23 | -6.8510448998 | -0.5917496496 | 1.3719191228 |
| N24 | -6.9757069212 | -0.9250823865 | 0.1670389048 | N24 | -7.0265747722 | -1.2202453727 | 0.1495589164 |
| C25 | -6.0897904844 | -3.1058195901 | -0.7431988258 | C25 | -5.7521331305 | -3.0086116822 | -1.1027811203 |
| C26 | -4.4055815977 | -4.1188012780 | 0.7138000468 | C26 | -3.8079977257 | -3.8191596435 | 0.1460228674 |
| C27 | -5.2026411380 | -4.1432664772 | -0.4513963471 | C27 | -4.6411282701 | -3.8402355610 | -0.9950995563 |
| H28 | -3.8331739493 | -3.0324577464 | 2.4839952499 | H28 | -3.4177339177 | -2.9210167071 | 2.0610288634 |
| H29 | -7.2109145458 | 0.7677035755 | 1.4537744720 | H29 | -7.5544718060 | 0.1637185625 | 1.6942110082 |
| H30 | -6.6842868947 | -3.1073296662 | -1.6530746795 | H30 | -6.3906518362 | -3.0290539770 | -1.9834013764 |
| H31 | -3.7182196270 | -4.9455946587 | 0.9069947537 | H31 | -2.9290586591 | -4.4631587529 | 0.1702667949 |
| H32 | -5.1161719338 | -4.9841297790 | -1.1361304705 | H32 | -4.3962697386 | -4.5203904132 | -1.8074886037 |
| C33 | -8.5110614298 | -2.5490541919 | 4.7426118150 | C33 | -8.3602039755 | -3.1374106916 | 4.7378284478 |
| H34 | -9.0127345136 | -2.5943731014 | 5.6865785848 | H34 | -8.8558056439 | -3.2197557804 | 5.6824977426 |
| H35 | -7.5796307843 | -2.0345826804 | 4.8584973891 | H35 | -7.4650931675 | -2.5621300083 | 4.8540291867 |
| H36 | -9.1221336340 | -2.0260876087 | 4.0381467566 | H36 | -9.0063550740 | -2.6539348853 | 4.0364626988 |
| C37 | -8.2379298454 | -3.9768232527 | 4.2559731076 | C37 | -7.9849533246 | -4.5449545447 | 4.2604036344 |
| H38 | -7.2967524413 | -4.3250420606 | 4.6281292682 | H38 | -7.0256470523 | -4.8261116830 | 4.6428351456 |
| C39 | -9.3033156544 | -4.9541516738 | 4.7877468842 | C39 | -8.9981533378 | -5.5928757586 | 4.7531163639 |
| H40 | -9.1467701206 | -5.9687404579 | 4.3999072200 | H40 | -8.7587847705 | -6.5937775214 | 4.3718217800 |
| H41 | -9.2728603421 | -5.0075077846 | 5.8832066751 | H41 | -9.0035002112 | -5.6458618158 | 5.8495574538 |
| H42 | -10.3162164388 | -4.6432897652 | 4.5048487256 | H42 | -10.0196753923 | -5.3507523398 | 4.4337496582 |
| C43 | -8.1160967369 | -4.0617662253 | 2.7145557627 | C43 | -7.7970381761 | -4.6055367273 | 2.7201195474 |
| H44 | -7.7289836582 | -5.0558116013 | 2.4509930449 | H44 | -7.3515358484 | -5.5774904923 | 2.4650793682 |
| H45 | -7.3539888323 | -3.3475486976 | 2.3892627797 | H45 | -7.0546585222 | -3.8509566964 | 2.4305129934 |
| C46 | -9.4025885202 | -3.8136600462 | 1.9175796198 | C46 | -9.0591699808 | -4.4110232580 | 1.8693089503 |
| H47 | -9.1890189128 | -3.8376467742 | 0.8434366552 | H47 | -8.8013415540 | -4.4297300525 | 0.8039949951 |
| H48 | -10.1666423969 | -4.5736047284 | 2.1188156128 | H48 | -9.7994591895 | -5.2013277959 | 2.0434994814 |
| H49 | -9.8434871695 | -2.8337345653 | 2.1374442105 | H49 | -9.5484926116 | -3.4490720229 | 2.0663126543 |
| C50 | 3.0249653416 | 1.7132866061 | -4.1120876266 | C50 | 2.8469914125 | 1.9206907923 | -4.1204704807 |
| H51 | 3.1392551191 | 0.9981735865 | -4.8996310828 | H51 | 3.0073678426 | 1.2177614611 | -4.9109403759 |
| H52 | 3.9809541854 | 2.1234529050 | -3.8624269596 | H52 | 3.7738612482 | 2.3929997392 | -3.8708268641 |
| H53 | 2.3740653354 | 2.4984533120 | -4.4358522037 | H53 | 2.1443850616 | 2.6617253465 | -4.4401462779 |
| C54 | 2.5612260910 | 1.0740070790 | -2.8144435388 | C54 | 2.2867801057 | 1.1592872818 | -2.9063153713 |
| H55 | 3.3110580241 | 0.3493867401 | -2.4635414078 | H55 | 3.0862472022 | 0.5841662900 | -2.4169549111 |
| H56 | 2.4808910334 | 1.8260435850 | -2.0196995206 | H56 | 1.8876147952 | 1.8427267581 | -2.1476031119 |
| C57 | 1.2491600771 | 0.3357141324 | -2.9106421837 | C57 | 1.2142605729 | 0.1413041542 | -3.2886871757 |
| O58 | 0.6970902547 | 0.0096895486 | -1.8204069906 | O58 | 0.4259895143 | -0.2463708619 | -2.3189351471 |
| O59 | 0.7932219316 | 0.0472440878 | -4.0455202333 | O59 | 1.1494819119 | -0.3099399469 | -4.4407947540 |
| C60 | 3.6967082360 | 2.4859406039 | -10.3483765617 | C60 | 3.4532220027 | 2.7598398697 | -10.3548779403 |
| H61 | 4.2019819899 | 1.5759721557 | -10.0965704817 | H61 | 4.0186682968 | 1.8847322396 | -10.1074728640 |
| H62 | 4.3783409152 | 3.3063983533 | -10.2685420103 | H62 | 4.0786282947 | 3.6237396633 | -10.2730795025 |
| H63 | 3.3295823285 | 2.4234159245 | -11.3512989434 | H63 | 3.0891081242 | 2.6766315441 | -11.3573955267 |
| C64 | 2.5558498307 | 2.7225908718 | -9.3232385172 | C64 | 2.2763097709 | 2.8387444857 | -9.3549453918 |
| H65 | 2.9553306778 | 3.3139551039 | -8.4881029786 | H65 | 2.5872949772 | 3.3837138304 | -8.4533982701 |
| H66 | 1.7666180939 | 3.3393987473 | -9.7735449048 | H66 | 1.4386820942 | 3.4062868372 | -9.7823409142 |
| C67 | 1.9426160029 | 1.4441918476 | -8.7235435200 | C67 | 1.8158062100 | 1.4363336886 | -8.9261329772 |
| H68 | 2.7357243386 | 0.8419891490 | -8.2622674537 | H68 | 2.6563304101 | 0.9231313570 | -8.4412572768 |
| H69 | 1.4998676009 | 0.8183681078 | -9.5126484748 | H69 | 1.5703950515 | 0.8400110378 | -9.8189867823 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| C70 | 0.8924115102 | 1.7804164596 | -7.6476586501 | C70 | 0.6323869630 | 1.4536199597 | -7.9424629930 |
| H71 | -0.0082325883 | 2.2078297547 | -8.1037348139 | H71 | -0.3206573264 | 1.6273664966 | -8.4571301587 |
| H72 | 1.2998538166 | 2.5511000921 | -6.9843380439 | H72 | 0.7576580660 | 2.2793808203 | -7.2354232244 |
| N73 | 0.5035282983 | 0.6609521003 | -6.7791287582 | N73 | 0.5240553236 | 0.2444746473 | -7.1202756188 |
| H74 | 0.9270765057 | 0.5954595159 | -5.8423078241 | H74 | 0.9003043047 | 0.2618539658 | -6.1658397861 |
| C75 | -0.6256435220 | -0.0517853358 | -6.8963699367 | C75 | -0.0511529267 | -0.9043741719 | -7.4743711474 |
| N76 | -1.3072536461 | -0.0533827716 | -8.0726567168 | N76 | -0.4966874134 | -1.0939201819 | -8.7360511419 |
| H77 | -2.1259719305 | -0.6338130647 | -8.1546032942 | H77 | -0.9488764159 | -1.9598362972 | -8.9896462275 |
| H78 | -0.7943114119 | 0.1128244359 | -8.9280114714 | H78 | -0.1435269649 | -0.5256348916 | -9.4905246958 |
| N79 | -1.0806580539 | -0.7254030421 | -5.8491814005 | N79 | -0.1998712815 | -1.8636308991 | -6.5573247211 |
| H80 | -1.8679121445 | -1.3596471217 | -5.8921158275 | H80 | -0.6158709189 | -2.7570659856 | -6.7782752527 |
| H81 | -0.6423798689 | -0.5725500690 | -4.9234554454 | H81 | -0.1097899658 | -1.6227827657 | -5.5687140621 |
| C82 | 1.5252589176 | -7.0426056574 | 1.3724852507 | C82 | 1.5397449806 | -7.0373117327 | 1.3336901532 |
| H83 | 2.1874919555 | -6.2987212487 | 0.9819506643 | H83 | 2.1947829041 | -6.2862246582 | 0.9448103420 |
| H84 | 1.7556638567 | -7.9903030139 | 0.9326386853 | H84 | 1.7764564668 | -7.9802462573 | 0.8870095490 |
| H85 | 1.6425682362 | -7.1011518346 | 2.4347254410 | H85 | 1.6593845842 | -7.0987668951 | 2.3955075768 |
| C86 | 0.0754812577 | -6.6668581700 | 0.9886851801 | C86 | 0.1146251836 | -6.5689033345 | 0.9642779884 |
| H87 | -0.6371578471 | -7.1447333554 | 1.6760148273 | H87 | -0.6175222036 | -7.3821296548 | 1.0759134531 |
| H88 | -0.0928894070 | -5.5855596842 | 1.0664903305 | H88 | -0.2177476618 | -5.7777515942 | 1.6485124092 |
| C89 | -0.2864826367 | -7.1970052438 | -0.3974661398 | C89 | -0.0073482027 | -6.0164145880 | -0.4809582761 |
| O90 | 0.0107937383 | -8.3092190053 | -0.7680916932 | O90 | 0.8625369121 | -6.3249850592 | -1.3176797400 |
| O91 | -1.0256686024 | -6.4070714394 | -1.2042047119 | O91 | -1.0366031366 | -5.2727364733 | -0.6932660823 |
| H92 | -1.0877963805 | -5.4802573470 | -0.9135081295 | H92 | -1.1052030098 | -4.6983200942 | -2.1488778458 |
| C93 | -4.5367970128 | 4.9965859389 | 1.2891867980 | C93 | -4.9068589401 | 4.6691934939 | 1.3049330017 |
| H94 | -3.8236155478 | 5.0119159924 | 0.4910269145 | H94 | -4.1978460073 | 4.7349405085 | 0.5056221121 |
| H95 | -4.1849652050 | 5.6067667799 | 2.0947614555 | H95 | -4.5949955129 | 5.2987536601 | 2.1120608228 |
| H96 | -5.4737577991 | 5.3755702393 | 0.9362143691 | H96 | -5.8677665486 | 4.9858986160 | 0.9549784876 |
| C97 | -4.6445782610 | 3.5313705937 | 1.8064172814 | C97 | -4.9665731028 | 3.1819781660 | 1.7706485339 |
| H98 | -4.0970338017 | 3.4101455226 | 2.7446052167 | H98 | -4.4707634630 | 3.0472885750 | 2.7346327320 |
| H99 | -5.6910772232 | 3.2650299755 | 1.9956936824 | H99 | -6.0058530125 | 2.8535076189 | 1.8779971590 |
| C100 | -4.0903447488 | 2.5454961879 | 0.7869746974 | C100 | -4.2958472029 | 2.2796283468 | 0.7484718191 |
| O101 | -4.5642933110 | 2.5001478106 | -0.3540235954 | O101 | -4.7190028855 | 2.1673173144 | -0.4020064878 |
| O102 | -3.0852762908 | 1.8295388949 | 1.2220191795 | O102 | -3.2042172438 | 1.6976288708 | 1.2079474611 |
| O103 | -1.415898024 | -3.8305230219 | -2.4450376475 | O103 | -1.1984970662 | -4.2706336617 | -3.083426224 |
| C104 | -0.891806216 | -4.4239241005 | -3.6238024350 | C104 | -0.6925596670 | -5.210180666 | -4.0074433297 |
| H105 | -0.840951015 | -5.4929014519 | -3.4361164559 | H105 | 0.0578599868 | -5.8539714979 | -3.5309796032 |
| H106 | 0.1145237219 | -4.0379028959 | -3.8402767333 | H106 | -0.2146770540 | -4.680234827 | -4.8538124819 |
| H107 | -1.5403218792 | -4.228956913 | -4.4867207893 | H107 | -1.4894477194 | -5.846654542 | -4.4335583383 |
| C108 | -1.4138026716 | -2.418748552 | -2.4688705474 | C108 | -0.6895522683 | -1.653451230 | -2.7775753040 |
| H109 | -0.410007019 | -2.0517788240 | -2.7039880543 | H109 | 0.1167292360 | -2.3560803351 | -2.9355802271 |
| C110 | -1.8271551298 | -1.976269922 | -1.0562981588 | C110 | -1.4595768084 | -1.752283898 | -1.4962670105 |
| H111 | -2.494703727 | -2.7591494075 | -0.6877219393 | H111 | -2.069849043 | -2.6571731816 | -1.5808702894 |
| O112 | -0.630919280 | -2.0091285046 | -0.2666356505 | O112 | -0.476047552 | -1.9338825935 | -0.4818066375 |
| H113 | -0.893799718 | -2.1433254270 | 0.7048960916 | H113 | -0.877510421 | -2.2131245368 | 0.3966780235 |
| C114 | -2.616277855 | -0.6152470488 | -0.9102711882 | C114 | -2.3942722056 | -0.507515265 | -1.2563490480 |
| H115 | -3.4178986579 | -0.831137953 | -0.1927525265 | H115 | -3.2629932683 | -0.880101414 | -0.6975637340 |
| O116 | -1.8486328307 | 0.4588587669 | -0.4077956096 | O116 | -1.7668071910 | 0.531726830 | -0.5404328181 |
| H117 | -2.5935203334 | 1.2212865206 | 0.4491002088 | H117 | -2.6929232625 | 1.1939190283 | 0.4642069619 |
| C118 | -3.325433900 | -0.2348009214 | -2.2279781923 | C118 | -2.8990870525 | 0.0536663345 | -2.6043651220 |
| H119 | -2.6484627631 | 0.4336655323 | -2.7828196517 | H119 | -2.2320876527 | 0.8955869344 | -2.8537506926 |
| O120 | -4.5820741708 | 0.3918659579 | -2.0170849399 | O120 | -4.2474153541 | 0.4792207062 | -2.5551999504 |
| H121 | -4.4675641044 | 1.2157363025 | -1.4716517807 | H121 | -4.3692137425 | 1.0903573380 | -1.7862210895 |
| C122 | -3.547359297 | -1.4738782538 | -3.1021657604 | C122 | -2.7817249278 | -0.968155313 | -3.7342739064 |
| H123 | -4.076196637 | -2.2544529234 | -2.5447985682 | H123 | -3.3443467835 | -1.880279430 | -3.5118081332 |
| C124 | -4.3403338408 | -1.186544900 | -4.3800342781 | C124 | -3.2038990229 | -0.412462889 | -5.0975391899 |
| H125 | -4.3259601644 | -2.086269626 | -5.0095973979 | H125 | -2.9551849953 | -1.142581985 | -5.8772349553 |

| | |
|---|---|
| H126 -3.8342457361 -0.372481427 -4.9333609971 | H126 -2.6135225318 0.5022832685 -5.2967176110 |
| O127 -5.6894679510 -0.882925369 -4.1166145266 | O127 -4.588985728 -0.1860295124 -5.1607995109 |
| H128 -5.6499431996 -0.219761675 -3.3960996089 | H128 -4.8260852730 0.2690878796 -4.3264338398 |
| O129 -2.2492503696 -1.9331982759 -3.508379529 | O129 -1.366808954 -1.3498588194 -3.8686722525 |
| H130 -7.474913507 -0.5813695026 -0.6404117239 | H130 -7.7290276913 -0.978227771 -0.5308673032 |
| Transition state TS2 (Model 2) | Covalent intermediate (Model 2) |
| $E_{absolute}$ = -2867.26125014924 hartree | $E_{absolute}$ = -2867.29332013373 hartree |
| Zn1 -0.0297687033 0.0002933136 0.0506679399 | Zn1 -0.0000534791 -0.0008410127 -0.0024686421 |
| C2 -0.0343311928 -0.0143523629 4.8817135897 | C2 0.0027932190 0.0010553370 4.8286001410 |
| H3 0.8189078675 -0.0041016465 4.2355738727 | H3 0.8550752229 -0.0022125499 4.1811256510 |
| H4 0.1099071097 -0.7720637075 5.6235993840 | H4 0.1106498416 -0.7603527777 5.5728785389 |
| H5 -0.2451134553 0.9187277149 5.3606695231 | H5 -0.1427956772 0.9479755955 5.3046382437 |
| C6 -1.2061912088 -0.3846972441 3.9335660908 | C6 -1.1587789098 -0.3051684833 3.8617636217 |
| H7 -1.8379742967 -1.1855913193 4.3241136556 | H7 -1.5777617238 -1.3041869581 4.0216436170 |
| H8 -1.8503952496 0.4880409647 3.7632465991 | H8 -1.9982427159 0.3899998455 3.9909557103 |
| C9 -0.7211267273 -0.8016451342 2.5509055192 | C9 -0.7987655819 -0.2511027773 2.3728479501 |
| O10 0.1509892667 0.0123364369 1.9925906913 | O10 0.3226841544 0.2954461587 1.9814550492 |
| O11 -1.1727282655 -1.8301054730 2.0237376216 | O11 -1.6065227760 -0.7346608624 1.5598506941 |
| C12 -6.2580565318 -0.2201241621 4.2715612036 | C12 -6.2214721170 0.2197469117 4.2285338505 |
| H13 -5.8548459184 -0.0083336430 5.2402387381 | H13 -5.8030270059 0.4066863250 5.1958639561 |
| H14 -6.5675414336 0.6937716714 3.8071821721 | H14 -6.4684855389 1.1511220204 3.7614587387 |
| H15 -7.0993307894 -0.8732909233 4.3710205215 | H15 -7.1053060048 -0.3739639856 4.3315211295 |
| C16 -5.2133835700 -0.9273044160 3.3839380018 | C16 -5.2375656155 -0.5583954930 3.3281018372 |
| H17 -4.8522074876 -1.8283641066 3.8998147353 | H17 -4.9102967875 -1.4687057446 3.8510345534 |
| H18 -4.3298973504 -0.2847220153 3.2624127946 | H18 -4.3202431684 0.0245093063 3.1628844573 |
| C19 -5.7414829638 -1.3199854505 2.0271387675 | C19 -5.8369072017 -0.9490859275 1.9998925135 |
| C20 -5.1642496017 -2.2951242647 1.1272735955 | C20 -5.3291653088 -1.9391287625 1.0753866275 |
| C21 -5.9816405439 -2.3428026294 -0.0355933856 | C21 -6.2347795637 -2.0089707137 -0.0197526488 |
| C22 -4.0345886324 -3.1337927194 1.1986450402 | C22 -4.1972093503 -2.7732840587 1.0728388959 |
| C23 -6.8573172485 -0.8330605233 1.3898664680 | C23 -6.9900098579 -0.4638189246 1.4330956175 |
| N24 -7.0109138364 -1.4443671849 0.1563777956 | N24 -7.2414957900 -1.1010299607 0.2289725946 |
| C25 -5.6961472520 -3.1886423189 -1.1141721712 | C25 -6.0522963442 -2.8940252423 -1.0875512250 |
| C26 -3.7433362913 -3.9800237854 0.1330888265 | C26 -4.0128586378 -3.6580649876 0.0165727396 |
| C27 -4.5707173267 -4.0011669574 -1.0123608398 | C27 -4.9350795069 -3.7218056869 -1.0492525861 |
| H28 -3.3761889513 -3.0994011660 2.0611546127 | H28 -3.4562920158 -2.6969362468 1.8617492598 |
| H29 -7.5791268537 -0.0985984616 1.7197802044 | H29 -7.6780257657 0.2860646355 1.7979224285 |
| H30 -6.3293693711 -3.2073982116 -1.9985528530 | H30 -6.7557154803 -2.9326740058 -1.9160573559 |
| H31 -2.8557823386 -4.6116953873 0.1538160761 | H31 -3.1342248161 -4.2982459952 0.0022150448 |
| H32 -4.3097633493 -4.6655603505 -1.8328365619 | H32 -4.7718670020 -4.4312075740 -1.8580601784 |
| C33 -8.2756486295 -3.4892648874 4.7512529609 | C33 -8.4572556936 -2.9019915948 4.7225820614 |
| H34 -8.7689166086 -3.5944464896 5.6948764744 | H34 -8.9549245121 -2.9699975036 5.6673056968 |
| H35 -7.4037372388 -2.8800700167 4.8709441109 | H35 -7.5454852050 -2.3535116862 4.8388498596 |
| H36 -8.9391864903 -3.0285830762 4.0507789875 | H36 -9.0889225658 -2.3993208255 4.0215163929 |
| C37 -7.8448652371 -4.8794665303 4.2699087364 | C37 -8.1257584715 -4.3190825804 4.2410325490 |
| H38 -6.8760884980 -5.1243300832 4.6534256089 | H38 -7.1742707962 -4.6297734306 4.6201290022 |
| C39 -8.8179454287 -5.9677188522 4.7555497289 | C39 -9.1674456427 -5.3376479020 4.7378996987 |
| H40 -8.5388363766 -6.9570474543 4.3711074485 | H40 -8.9617212106 -6.3445135618 4.3520121755 |
| H41 -8.8244418283 -6.0252167944 5.8517645348 | H41 -9.1672350779 -5.3940437432 5.8340917569 |
| H42 -9.8471204467 -5.7640347604 4.4339407831 | H42 -10.1832899892 -5.0638262448 4.4262871605 |
| C43 -7.6500603113 -4.9258403445 2.7298913084 | C43 -7.9499810529 -4.3856597220 2.6998859414 |
| H44 -7.1664039969 -5.8785979994 2.4718810822 | H44 -7.5362308474 -5.3714862777 2.4427796831 |
| H45 -6.9365032604 -4.1417525515 2.4460795819 | H45 -7.1882329612 -3.6535277576 2.4039971095 |
| C46 -8.9159473300 -4.7760397048 1.8757232242 | C46 -9.2123222368 -4.1546394518 1.8585078661 |
| H47 -8.6539797807 -4.7785413173 0.8112706124 | H47 -8.9631772263 -4.1816069804 0.7913058317 |
| H48 -9.6253317061 -5.5956325429 2.0430202933 | H48 -9.9748982962 -4.9219422902 2.0394415064 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| H49 | -9.4431683456 | -3.8350727853 | 2.0763422094 | H49 | -9.6702377668 | -3.1780997572 | 2.0582504336 |
| C50 | 2.7380456897 | 2.0347399690 | -4.0713876517 | C50 | 2.8933477501 | 1.8255418310 | -4.1357079282 |
| H51 | 2.9266732944 | 1.3416210283 | -4.8642802421 | H51 | 3.0327124920 | 1.1184826086 | -4.9264782007 |
| H52 | 3.6455202990 | 2.5417634715 | -3.8185984699 | H52 | 3.8338362270 | 2.2701024073 | -3.8860427761 |
| H53 | 2.0075782183 | 2.7491217809 | -4.3892813604 | H53 | 2.2129384122 | 2.5871800878 | -4.4549739756 |
| C54 | 2.2038249357 | 1.2453102764 | -2.8630009457 | C54 | 2.2627658673 | 0.9783552641 | -3.0176225008 |
| H55 | 3.0233898311 | 0.6995313358 | -2.3732850326 | H55 | 3.0446038980 | 0.4549447195 | -2.4512499367 |
| H56 | 1.7756746560 | 1.9087543282 | -2.1022239093 | H56 | 1.6658094932 | 1.5654990567 | -2.3150719867 |
| C57 | 1.1731830156 | 0.1873103113 | -3.2538053629 | C57 | 1.3991405461 | -0.1110379419 | -3.6432076935 |
| O58 | 0.4021962641 | -0.2406742194 | -2.2867879705 | O58 | 0.1907689590 | -0.1599666379 | -3.1379209967 |
| O59 | 1.1274512892 | -0.2580637886 | -4.4087871912 | O59 | 1.8216487893 | -0.8401214693 | -4.5490822454 |
| C60 | 3.3193942434 | 2.9207831427 | -10.3016753200 | C60 | 3.5231439506 | 2.6488211949 | -10.3698951957 |
| H61 | 3.9181138221 | 2.0673759261 | -10.0567607707 | H61 | 4.0624718987 | 1.7572581391 | -10.1229279315 |
| H62 | 3.9106281455 | 3.8080234753 | -10.2156873378 | H62 | 4.1738681952 | 3.4937926964 | -10.2878708876 |
| H63 | 2.9601109702 | 2.8273361299 | -11.3050333012 | H63 | 3.1565628794 | 2.5768238505 | -11.3723814818 |
| C64 | 2.1387724316 | 2.9527224434 | -9.3037696845 | C64 | 2.3622974378 | 2.7571647788 | -9.3500448085 |
| H65 | 2.4282602217 | 3.5053529354 | -8.3996997458 | H65 | 2.6928806157 | 3.3380637462 | -8.4780729463 |
| H66 | 1.2820983658 | 3.4915382823 | -9.7307039647 | H66 | 1.5126684249 | 3.3058040115 | -9.7781446121 |
| C67 | 1.7275802207 | 1.5336108845 | -8.8803439000 | C67 | 1.9183916103 | 1.3689112106 | -8.8592257342 |
| H68 | 2.5849465265 | 1.0483822279 | -8.3961605346 | H68 | 2.7931260443 | 0.8531414888 | -8.4397970172 |
| H69 | 1.5033939661 | 0.9314102791 | -9.7747660877 | H69 | 1.5692156619 | 0.7502223429 | -9.6967359123 |
| C70 | 0.5435291146 | 1.5079706861 | -7.8977466187 | C70 | 0.8353042982 | 1.4225057640 | -7.7721309342 |
| H71 | -0.4142831559 | 1.6543114606 | -8.4121925883 | H71 | -0.1375704133 | 1.6912077181 | -8.1950278552 |
| H72 | 0.6416293391 | 2.3339785624 | -7.1866994970 | H72 | 1.0749087082 | 2.1999151065 | -7.0379822363 |
| N73 | 0.4736456927 | 0.2922954488 | -7.0814656224 | N73 | 0.7279663861 | 0.1653627941 | -7.0225868910 |
| H74 | 0.8552570838 | 0.3158117929 | -6.1294560352 | H74 | 1.4268216869 | -0.0045578071 | -6.2990406789 |
| C75 | -0.0631702727 | -0.8733542456 | -7.4406584761 | C75 | -0.1752974324 | -0.8058257568 | -7.1381051170 |
| N76 | -0.5138807419 | -1.0682334209 | -8.6997344868 | N76 | -1.2534141916 | -0.6708382545 | -7.9419768129 |
| H77 | -0.9310634540 | -1.9501261463 | -8.9582876278 | H77 | -1.9517130526 | -1.4015210137 | -7.9134368405 |
| H78 | -0.1927275402 | -0.4781478115 | -9.4517421031 | H78 | -1.6147593455 | 0.2538530434 | -8.1274029157 |
| N79 | -0.1655013224 | -1.8464246137 | -6.5313447191 | N79 | 0.0314066251 | -1.9792980633 | -6.5080155691 |
| H80 | -0.5710704021 | -2.7445764114 | -6.7530542291 | H80 | -0.8026125827 | -2.4074209797 | -6.1148716198 |
| H81 | -0.0797526104 | -1.6087728282 | -5.5414013484 | H81 | 0.7831113830 | -1.9515153548 | -5.8132334719 |
| C82 | 1.7543696059 | -6.9929796192 | 1.3468302681 | C82 | 1.6932747495 | -7.0126916527 | 1.3149526887 |
| H83 | 2.3822783376 | -6.2171684991 | 0.9617566221 | H83 | 2.3308044169 | -6.2465924307 | 0.9262601520 |
| H84 | 2.0262728527 | -7.9254840243 | 0.8982931277 | H84 | 1.9508554492 | -7.9492411344 | 0.8663898060 |
| H85 | 1.8747571720 | -7.0541408663 | 2.4085801203 | H85 | 1.8145118334 | -7.0718222583 | 2.3767209092 |
| C86 | 0.3146097187 | -6.5711231707 | 0.9762628570 | C86 | 0.2806417211 | -6.5031262043 | 0.8848153890 |
| H87 | -0.3881623737 | -7.4119887500 | 1.0703016343 | H87 | 0.0380051295 | -6.8466375837 | -0.1253907363 |
| H88 | -0.0502125540 | -5.8035461513 | 1.6707003631 | H88 | -0.5094522840 | -6.8715765698 | 1.5508981130 |
| C89 | 0.1806469712 | -6.0001312381 | -0.4611206317 | C89 | 0.2446623515 | -4.9671681142 | 0.9157765459 |
| O90 | 1.0670010447 | -6.2647199222 | -1.2961063642 | O90 | 0.6392072128 | -4.3084977861 | 1.8524206529 |
| O91 | -0.8723476104 | -5.2901950002 | -0.6693302535 | O91 | -0.2413548357 | -4.4117725684 | -0.2254213076 |
| H92 | -0.9818393366 | -4.7379716495 | -2.1385313732 | H92 | -0.1545628435 | -4.5985603568 | -2.1676147106 |
| C93 | -5.1239458599 | 4.4588348735 | 1.3530310951 | C93 | -4.7748407516 | 4.8002886514 | 1.2921112458 |
| H94 | -4.4169741265 | 4.5551692925 | 0.5550144569 | H94 | -4.0643220361 | 4.8453277921 | 0.4927023781 |
| H95 | -4.8378602339 | 5.0969363362 | 2.1629950804 | H95 | -4.4443423749 | 5.4200220885 | 2.0994286615 |
| H96 | -6.0959798710 | 4.7392853724 | 1.0028723777 | H96 | -5.7260067942 | 5.1454433283 | 0.9424469253 |
| C97 | -5.1232486744 | 2.9683040271 | 1.8114839788 | C97 | -4.7936604603 | 3.2885425945 | 1.7015577936 |
| H98 | -4.6229552869 | 2.8491242327 | 2.7752018165 | H98 | -4.2310024493 | 3.1329115566 | 2.6256285961 |
| H99 | -6.1481134023 | 2.5966238241 | 1.9165543383 | H99 | -5.8157991653 | 2.9294193996 | 1.8541214733 |
| C100 | -4.4152956701 | 2.0993485412 | 0.7852086867 | C100 | -4.1590360125 | 2.4477229828 | 0.5930446207 |
| O101 | -4.8252886956 | 1.9859253941 | -0.3698892479 | O101 | -4.7689123728 | 2.1890066318 | -0.4444424764 |
| O102 | -3.3080831748 | 1.5489299547 | 1.2466807879 | O102 | -2.9086833536 | 2.1102171331 | 0.8410855203 |
| O103 | -1.097742978 | -4.3173170796 | -3.0714741846 | O103 | -0.306968996 | -4.6361318577 | -3.1310329142 |
| C104 | -0.472500046 | -5.1886754382 | -3.9902297963 | C104 | -1.515950540 | -5.3453971408 | -3.3343510653 |

| | |
|---|---|
| H105 0.3048960033 -5.7813419909 -3.4913907309 | H105 -1.475045840 -6.3709816722 -2.9355045241 |
| H106 -0.001130685 -4.6009893768 -4.8017230178 | H106 -1.6786601108 -5.414154583 -4.4157854033 |
| H107 -1.1943510528 -5.876560158 -4.4671961399 | H107 -2.385823427 -4.8394268223 -2.8879135732 |
| C108 -0.6520543599 -1.682693772 -2.7454432183 | C108 -0.650993168 -1.3852863691 -3.2421137071 |
| H109 0.1802295221 -2.3534222524 -2.9073179014 | H109 0.014243405 -2.2148105818 -3.4770884339 |
| C110 -1.416955674 -1.8172108416 -1.4637402118 | C110 -1.2293540559 -1.508735329 -1.8368179648 |
| H111 -1.9885884337 -2.745893856 -1.5493915621 | H111 -1.8900687377 -2.389545206 -1.8486827915 |
| O112 -0.427377042 -1.9606106658 -0.4489166101 | O112 -0.1600753709 -1.713401894 -0.9573472266 |
| H113 -0.819780359 -2.2618004925 0.4264446854 | H113 -0.196425577 -3.4027828016 -0.2132006002 |
| C114 -2.403291776 -0.6133899097 -1.2213028318 | C114 -2.1646648637 -0.258821446 -1.5229768041 |
| H115 -3.253133802 -1.0231249613 -0.6596113811 | H115 -2.9211965896 -0.630344058 -0.8172220146 |
| O116 -1.820106022 0.4534130350 -0.5075426225 | O116 -1.4814879474 0.8527996291 -0.9726132741 |
| H117 -2.7757563394 1.0717615497 0.5008496037 | H117 -2.4365648215 1.5950443357 0.0754764437 |
| C118 -2.9360864065 -0.075521679 -2.5675923349 | C118 -2.9047607668 0.2011380215 -2.8042255915 |
| H119 -2.306941255 0.7940200536 -2.8203688366 | H119 -2.3403999125 1.0649970855 -3.1930697881 |
| O120 -4.3012971833 0.2923410669 -2.5134941457 | O120 -4.2634867390 0.5476076310 -2.5754233871 |
| H121 -4.4435635206 0.9057146103 -1.7501910026 | H121 -4.3434875619 1.1572930350 -1.8003222568 |
| C122 -2.7765836899 -1.093563416 -3.6956515160 | C122 -2.9299377007 -0.871921998 -3.8954676386 |
| H123 -3.293030031 -2.0317853567 -3.4687575262 | H123 -3.467121329 -1.7563985010 -3.5300797750 |
| C124 -3.230969237 -0.5615775612 -5.0579860779 | C124 -3.5986810067 -0.390054501 -5.1886765907 |
| H125 -2.9539520799 -1.2818904608 -5.837278511 | H125 -3.458374194 -1.1591373497 -5.9644565641 |
| H126 -2.6836772141 0.3783707538 -5.2630602109 | H126 -3.0670072529 0.5211185959 -5.5288213174 |
| O127 -4.6251011399 -0.39809976 -5.1141738843 | O127 -4.979696924 -0.1849147333 -5.0498830053 |
| H128 -4.8777180781 0.0488222643 -4.2799411057 | H128 -5.086881071 0.2579486712 -4.1795550264 |
| O129 -1.3447524197 -1.405081667 -3.8351872547 | O129 -1.579066035 -1.2483904594 -4.2726124943 |
| H130 -7.723703508 -1.2181305566 -0.5186612501 | H130 -7.9769621879 -0.854805318 -0.4141555180 |

Table S2. Optimized geometries of Michaelis complex, transition state and glycosyl-enzyme covalent intermediate for Model 3.

| Michaelis complex (Model 3)<br>Eabsolute = -3477.27757863120 hartrees | | | |
|---|---|---|---|
| Zn1 | 0.0059518077 | -0.0045420373 | 0.0027571077 |
| C2 | 0.0111847474 | -0.0005226064 | 6.4443614577 |
| H3 | 1.1006414128 | 0.0002560020 | 6.4329569189 |
| H4 | -0.4158167618 | -0.2136587975 | 7.4244355120 |
| H5 | -0.3450975592 | 0.9752556681 | 6.1140739757 |
| C6 | -0.4241263087 | -1.1456555268 | 5.5019067411 |
| H7 | -1.5087263738 | -1.1877057923 | 5.3827867968 |
| H8 | -0.1283484899 | -2.1011905642 | 5.9600277421 |
| C9 | 0.1632911203 | -1.0212630134 | 4.1383915781 |
| N10 | 1.5235855233 | -0.8732251740 | 3.8960609190 |
| C11 | -0.4245101351 | -0.9463358657 | 2.9050157300 |
| H12 | -1.4728104356 | -1.0219180024 | 2.6560396677 |
| C13 | 1.7013522745 | -0.7045020600 | 2.5610446007 |
| H14 | 2.6633343700 | -0.5517948173 | 2.0947827996 |
| N15 | 0.5399838578 | -0.7487981042 | 1.9392063186 |
| C16 | 0.3177869213 | 4.7236668275 | 1.9601519125 |
| H17 | 1.2586291168 | 4.2500761978 | 1.6781825778 |
| H18 | -0.0404457240 | 4.3692763268 | 2.9260006039 |
| H19 | 0.3855501668 | 5.8106805161 | 1.9277714231 |
| C20 | -0.7218301460 | 4.2912241137 | 0.9006764148 |
| H21 | -0.3762621627 | 4.5780292740 | -0.1017992036 |

| | | | |
|---|---|---|---|
| H22 | -1.6845146458 | 4.7810664074 | 1.0715229561 |
| C23 | -0.9675922383 | 2.7836221809 | 0.8573641860 |
| O24 | 0.0502798651 | 2.0155460409 | 0.7984087392 |
| O25 | -2.1708390872 | 2.3994626233 | 0.8692770090 |
| C26 | -3.1788977475 | 6.4465099340 | -3.1018897308 |
| H27 | -2.5610201130 | 6.2126631625 | -3.9690305192 |
| H28 | -2.6994808666 | 7.1821662546 | -2.4554337733 |
| H29 | -4.1789539980 | 6.7641787922 | -3.3974984202 |
| C30 | -3.3027055574 | 5.1343417696 | -2.2761315195 |
| H31 | -3.8992682800 | 5.3247024621 | -1.3753814985 |
| H32 | -2.3051741539 | 4.8435977319 | -1.9212399475 |
| C33 | -3.9093476133 | 3.9799899061 | -3.0335353327 |
| C34 | -3.5316696077 | 3.5090963148 | -4.2651013600 |
| H35 | -2.7182174846 | 3.8180573334 | -4.9054569373 |
| C36 | -4.9876602584 | 3.1184544754 | -2.5989697901 |
| N37 | -4.3450597679 | 2.4578720684 | -4.6699993305 |
| H38 | -3.9700576836 | 1.7225588226 | -5.2568490462 |
| C39 | -5.2334546539 | 2.1858836267 | -3.6440899076 |
| C40 | -5.7337189385 | 3.0143594470 | -1.4110979515 |
| H41 | -5.5309036431 | 3.6775548116 | -0.5748406931 |
| C42 | -6.2366955371 | 1.2154243141 | -3.5535096603 |
| H43 | -6.4369385796 | 0.5372743679 | -4.3796298927 |
| C44 | -6.7110797625 | 2.0303021814 | -1.3057568049 |
| H45 | -7.2815698272 | 1.9351217024 | -0.3858669156 |
| C46 | -6.9705256768 | 1.1473376356 | -2.3726470900 |
| H47 | -7.7489928228 | 0.3960593105 | -2.2687686823 |
| C48 | -6.8929534527 | 7.2871486064 | -3.0144600258 |
| H49 | -6.9893295413 | 7.3188500796 | -4.1000119886 |
| H50 | -5.9471415442 | 6.8508512046 | -2.6936765934 |
| H51 | -7.0854612426 | 8.2540483196 | -2.5500730915 |
| C52 | -8.0590803717 | 6.3836808332 | -2.5184547212 |
| H53 | -8.9920642287 | 6.9457805447 | -2.6885287514 |
| C54 | -7.9426667057 | 6.1471853690 | -1.0003185315 |
| H55 | -7.8123783140 | 7.1217207913 | -0.5077454275 |
| H56 | -7.0242415054 | 5.5797976258 | -0.7991071675 |
| C57 | -8.1522827446 | 5.0873610041 | -3.3351610359 |
| H58 | -8.9979475468 | 4.4644681118 | -3.0254131296 |
| H59 | -8.2874529715 | 5.3143271769 | -4.3998180097 |
| H60 | -7.2446121187 | 4.4856408005 | -3.2358163472 |
| C61 | -9.1332479178 | 5.4240375026 | -0.3613194563 |
| H62 | -9.0224612432 | 5.3700095400 | 0.7275451569 |
| H63 | -10.0755127182 | 5.9453521031 | -0.5718137813 |
| H64 | -9.2298280249 | 4.3983914315 | -0.7313609038 |
| C65 | 2.3867571408 | -4.3585160874 | -0.5260575894 |
| H66 | 1.6232966953 | -5.1358193588 | -0.5647423043 |
| H67 | 3.1143034003 | -4.4877503947 | 0.2758274691 |
| H68 | 2.8291625577 | -4.1219961621 | -1.4938940098 |
| C69 | 1.5563805372 | -3.1548175474 | -0.0527408363 |
| H70 | 2.1070746078 | -2.2142116156 | -0.1106075863 |
| H71 | 1.2885250813 | -3.2929952323 | 1.0025523558 |
| C72 | 0.2434001893 | -3.0337387407 | -0.8130417572 |
| O73 | -0.2458745016 | -1.8824422971 | -0.9868215364 |
| O74 | -0.2940326360 | -4.1094983262 | -1.2166766686 |
| C75 | 3.1094856551 | -10.2310290119 | -2.8857613782 |
| H76 | 2.1917141675 | -10.7908691004 | -2.7071217635 |
| H77 | 3.6509228639 | -9.9712029481 | -1.9763170096 |

| | | | |
|---|---|---|---|
| H78 | 3.7390516877 | -10.6544531689 | -3.6682706997 |
| C79 | 2.6577370788 | -8.9051562791 | -3.5959487035 |
| H80 | 2.2681760894 | -9.1934429452 | -4.5835668645 |
| H81 | 3.5490507120 | -8.2957615363 | -3.8000190684 |
| C82 | 1.5969787977 | -8.0053686508 | -2.9248423724 |
| H83 | 2.0108506932 | -7.5164402757 | -2.0333741899 |
| H84 | 0.7330751303 | -8.5962895472 | -2.5957280142 |
| C85 | 1.1263699148 | -6.9408382217 | -3.9389954733 |
| H86 | 0.6284155541 | -7.4619436929 | -4.7652971895 |
| H87 | 2.0127630099 | -6.4374459742 | -4.3576286027 |
| N88 | 0.1892791362 | -5.9531760657 | -3.4008383095 |
| H89 | 0.2969891503 | -5.5996989733 | -2.4458528642 |
| C90 | -0.6461357028 | -5.2070120078 | -4.1532147367 |
| N91 | -0.7383455075 | -5.4203234749 | -5.4967801037 |
| H92 | -1.3128075903 | -4.7642990183 | -6.0094690056 |
| H93 | 0.0539812492 | -5.7966359317 | -5.9914879290 |
| N94 | -1.4338019820 | -4.3091264564 | -3.5910360830 |
| H95 | -1.9084793337 | -3.6150499879 | -4.1823972653 |
| H96 | -1.2706087432 | -4.0936391400 | -2.5834024768 |
| C97 | -4.4807708493 | -0.4881041594 | 5.9085464467 |
| H98 | -5.0037248519 | -1.2325944964 | 6.5095373451 |
| H99 | -4.4171765531 | 0.4473206300 | 6.4644341660 |
| H100 | -3.4764656966 | -0.8619606393 | 5.7071721663 |
| C101 | -5.1698630785 | -0.2409125582 | 4.5533388109 |
| H102 | -5.1546508874 | 0.8232480506 | 4.2883535766 |
| H103 | -6.2306906102 | -0.5166029158 | 4.5607532601 |
| C104 | -4.5089345852 | -0.9629897121 | 3.3849042167 |
| O105 | -3.5111249945 | -1.6592977827 | 3.4752977789 |
| O106 | -5.1487814821 | -0.7158524524 | 2.2408947356 |
| H107 | -4.6725270734 | -1.1119515257 | 1.4681181567 |
| C108 | 4.5734719845 | 1.6214117603 | -4.4381554421 |
| H109 | 5.3044530773 | 1.3920287258 | -5.2141835309 |
| H110 | 4.9493502120 | 2.3644397112 | -3.7354495989 |
| H111 | 3.6407839957 | 1.9988073005 | -4.8581960937 |
| C112 | 4.3551591248 | 0.3247910982 | -3.6338992172 |
| H113 | 5.3373268889 | -0.0707844231 | -3.3337038782 |
| H114 | 3.8923532122 | -0.4475509760 | -4.2572337074 |
| C115 | 3.4927726232 | 0.5252816525 | -2.4331062778 |
| N116 | 3.6022752324 | 1.6124853093 | -1.5744226057 |
| C117 | 2.4618686959 | -0.2126858828 | -1.9243762352 |
| H118 | 2.0211360674 | -1.1172565329 | -2.3118172917 |
| C119 | 2.6443241183 | 1.5050861409 | -0.6205649161 |
| H120 | 2.4282069890 | 2.2471418332 | 0.1329766751 |
| N121 | 1.9538368209 | 0.3955538530 | -0.7977390650 |
| C122 | 1.7649226173 | 3.9603630163 | -5.5961447851 |
| H123 | 1.3919921398 | 3.2332069172 | -6.3178483560 |
| H124 | 2.6109660207 | 3.5405638317 | -5.0515525795 |
| H125 | 2.0731508887 | 4.8682079530 | -6.1145304368 |
| C126 | 0.6189233605 | 4.2594739315 | -4.6097773657 |
| H127 | -0.1686643816 | 4.8150232729 | -5.1371886174 |
| H128 | 0.9582541698 | 4.8841414799 | -3.7780491669 |
| C129 | -0.0284119631 | 2.9814814592 | -4.0421398187 |
| O130 | -0.4463016072 | 2.1325134627 | -4.8765469001 |
| O131 | -0.1015754376 | 2.8999308500 | -2.7750006878 |
| O132 | -4.4460305179 | -1.9019945224 | -0.1516434078 |
| C133 | -4.7849628338 | -3.2723253443 | 0.0556802272 |

| | | | |
|---|---|---|---|
| H134 | -5.6698730126 | -3.2842774206 | 0.6962817485 |
| H135 | -5.0124284208 | -3.7634351963 | -0.8957893374 |
| H136 | -3.9676253497 | -3.8109920548 | 0.5548512244 |
| C137 | -3.2631542310 | -1.7259626350 | -0.9398760143 |
| H138 | -2.4531912876 | -2.3128048700 | -0.5037763366 |
| C139 | -2.9295153564 | -0.2191335036 | -0.8617421991 |
| H140 | -3.8927605979 | 0.2957142443 | -0.7589603840 |
| O141 | -2.1831192669 | -0.0332433975 | 0.3324660304 |
| H142 | -2.2567985590 | 0.9717699633 | 0.6122981082 |
| C143 | -2.2094452635 | 0.4186651733 | -2.1147192325 |
| H144 | -2.7320657731 | 1.3570700668 | -2.3355728193 |
| O145 | -0.8555344268 | 0.7080900054 | -1.8320341141 |
| H146 | -0.5746679467 | 1.6266288202 | -2.2384321433 |
| C147 | -2.3110702479 | -0.4904776819 | -3.3419528437 |
| H148 | -1.4937547575 | -1.2183072849 | -3.2449192458 |
| O149 | -2.2304982249 | 0.1830158866 | -4.5971445868 |
| H150 | -1.4915753102 | 0.8682480800 | -4.5965273499 |
| C151 | -3.6206460100 | -1.2890410079 | -3.2929722348 |
| H152 | -4.4785822676 | -0.6345557965 | -3.0892496269 |
| C153 | -3.8967451778 | -2.0357573616 | -4.6039059968 |
| H154 | -4.5065826819 | -1.4022254251 | -5.2594494598 |
| H155 | -4.4569011050 | -2.9504352286 | -4.3887459717 |
| O156 | -2.6995596465 | -2.4216996407 | -5.3004148265 |
| H157 | -2.2272890455 | -1.5698754278 | -5.4233992277 |
| O158 | -3.4763644397 | -2.2502385758 | -2.2390747779 |
| H159 | 2.2563470108 | -0.8772475693 | 4.5892063904 |
| H160 | 4.2246781546 | 2.3983001930 | -1.6832550001 |
| Transition state (Model 3) | | | |
| $E_{absolute}$ = -3477.04012756984 hartrees | | | |
| Zn1 | 0.0053084763 | 0.0113637134 | -0.0165702041 |
| C2 | 0.0026237566 | 0.0074766186 | 6.3834280516 |
| H3 | 1.0920919523 | 0.0095570803 | 6.3733674449 |
| H4 | -0.4261845554 | -0.2056218865 | 7.3627211344 |
| H5 | -0.3537685749 | 0.9834727283 | 6.0539037829 |
| C6 | -0.4452535620 | -1.1333739086 | 5.4396928976 |
| H7 | -1.5304958288 | -1.1656654864 | 5.3131895752 |
| H8 | -0.1570377347 | -2.0918660806 | 5.8988662882 |
| C9 | 0.1430077921 | -1.0209830775 | 4.0750973686 |
| N10 | 1.5017315387 | -0.8561334255 | 3.8295736064 |
| C11 | -0.4508300921 | -0.9931023647 | 2.8414951813 |
| H12 | -1.5099307229 | -1.1025034410 | 2.6224802659 |
| C13 | 1.6806288278 | -0.7206146270 | 2.4938534214 |
| H14 | 2.6406667155 | -0.5573549338 | 2.0261792805 |
| N15 | 0.5180030459 | -0.8056520883 | 1.8699243762 |
| C16 | 0.3091385910 | 4.7375171680 | 1.9470373451 |
| H17 | 1.2508870521 | 4.2653875293 | 1.6656439127 |
| H18 | -0.0498633295 | 4.3815113725 | 2.9120059718 |
| H19 | 0.3756554583 | 5.8246493219 | 1.9160809354 |
| C20 | -0.7291111137 | 4.2908234322 | 0.8863301054 |
| H21 | -0.4070894216 | 4.6227661117 | -0.1113201231 |
| H22 | -1.7056757994 | 4.7444612006 | 1.0803185868 |
| C23 | -0.9351424414 | 2.7751494875 | 0.7795643659 |
| O24 | 0.1028400278 | 2.0291526081 | 0.7304842229 |
| O25 | -2.1260619047 | 2.3597749634 | 0.7103185248 |
| C26 | -3.1818428462 | 6.4859569877 | -3.1101617546 |

| | | | |
|---|---|---|---|
| H27 | -2.5623617219 | 6.2579400727 | -3.9777109743 |
| H28 | -2.7042904023 | 7.2183660339 | -2.4586583189 |
| H29 | -4.1818227543 | 6.8041776934 | -3.4054349834 |
| C30 | -3.3624806040 | 5.1778201930 | -2.2920959526 |
| H31 | -3.8418584402 | 5.4196216351 | -1.3327255377 |
| H32 | -2.3768687599 | 4.7633076409 | -2.0353139998 |
| C33 | -4.1823813531 | 4.1191283243 | -2.9877607790 |
| C34 | -4.4658375489 | 4.0315573678 | -4.3282903417 |
| H35 | -4.1756769796 | 4.6841100852 | -5.1400932629 |
| C36 | -4.8259821896 | 2.9755992062 | -2.3730925299 |
| N37 | -5.2499826881 | 2.9175033253 | -4.5849425529 |
| H38 | -5.5339229100 | 2.6101008516 | -5.5012992484 |
| C39 | -5.4773521053 | 2.2425044469 | -3.4048913591 |
| C40 | -4.9120051750 | 2.5024043123 | -1.0481615612 |
| H41 | -4.3948324084 | 3.0181476412 | -0.2446757552 |
| C42 | -6.1943332412 | 1.0665466523 | -3.1468351250 |
| H43 | -6.6955441569 | 0.5270067214 | -3.9485930397 |
| C44 | -5.6187979110 | 1.3315703537 | -0.7857974195 |
| H45 | -5.6671690375 | 0.9360798954 | 0.2271744606 |
| C46 | -6.2501932852 | 0.6177864715 | -1.8290793275 |
| H47 | -6.7789355910 | -0.3011116640 | -1.5909830856 |
| C48 | -6.8970253507 | 7.3216819150 | -3.0235013639 |
| H49 | -6.9917782882 | 7.3596501936 | -4.1089951154 |
| H50 | -5.9511869635 | 6.8846232574 | -2.7038341889 |
| H51 | -7.0913934139 | 8.2856071896 | -2.5537368998 |
| C52 | -8.0343250507 | 6.3989216362 | -2.4990617418 |
| H53 | -8.9894037375 | 6.9145978465 | -2.6932234169 |
| C54 | -7.9044890172 | 6.2198008783 | -0.9729235147 |
| H55 | -7.7580204851 | 7.2098115164 | -0.5155163312 |
| H56 | -6.9891977298 | 5.6471923754 | -0.7610205476 |
| C57 | -8.0645119562 | 5.0591975697 | -3.2465358390 |
| H58 | -8.9058475197 | 4.4319194039 | -2.9301177569 |
| H59 | -8.1626533404 | 5.2177999617 | -4.3285633099 |
| H60 | -7.1430304414 | 4.4927637131 | -3.0745872111 |
| C61 | -9.0982445760 | 5.5368443255 | -0.2957279386 |
| H62 | -8.9588458027 | 5.4900244779 | 0.7907937697 |
| H63 | -10.0298879586 | 6.0854493864 | -0.4880578418 |
| H64 | -9.2365785985 | 4.5093868774 | -0.6499197006 |
| C65 | 2.3940136404 | -4.3066707321 | -0.6088230144 |
| H66 | 1.6315343487 | -5.0846299286 | -0.6532477197 |
| H67 | 3.1204853274 | -4.4397615059 | 0.1934054140 |
| H68 | 2.8376181975 | -4.0639457074 | -1.5745720077 |
| C69 | 1.4708601175 | -3.1215065575 | -0.2122744632 |
| H70 | 1.9476663147 | -2.1483251778 | -0.3273868674 |
| H71 | 1.2058840834 | -3.1970279369 | 0.8500434091 |
| C72 | 0.1220027602 | -3.1368806502 | -0.9514454223 |
| O73 | -0.5113204491 | -2.0067470925 | -1.0169758635 |
| O74 | -0.3280896984 | -4.2061425718 | -1.4138364341 |
| C75 | 3.1273143273 | -10.1643518106 | -3.0018937632 |
| H76 | 2.2099356614 | -10.7263128420 | -2.8279577830 |
| H77 | 3.6670521735 | -9.9092408054 | -2.0901074216 |
| H78 | 3.7585776674 | -10.5824257980 | -3.7859109666 |
| C79 | 2.7298098650 | -8.8460288688 | -3.7351199785 |
| H80 | 2.3538723579 | -9.1368961542 | -4.7279035440 |
| H81 | 3.6480834399 | -8.2717835186 | -3.9269417914 |
| C82 | 1.6920882474 | -7.9008564010 | -3.1123309448 |

| | | | |
|---|---|---|---|
| H83 | 2.0797178859 | -7.4368053754 | -2.1953028915 |
| H84 | 0.7755053901 | -8.4411647532 | -2.8427397318 |
| C85 | 1.3640946982 | -6.8235829763 | -4.1572235900 |
| H86 | 0.9309346491 | -7.3180651394 | -5.0349473021 |
| H87 | 2.3017661233 | -6.3431342677 | -4.4799255471 |
| N88 | 0.4148700371 | -5.8148803116 | -3.6987895475 |
| H89 | 0.3804457917 | -5.5416890023 | -2.7135080177 |
| C90 | -0.3328436327 | -5.0618309036 | -4.5178277672 |
| N91 | -0.2315456535 | -5.2032540481 | -5.8671619244 |
| H92 | -0.7843713679 | -4.5755191131 | -6.4367348768 |
| H93 | 0.6350787624 | -5.5234178564 | -6.2715444078 |
| N94 | -1.2225923678 | -4.2151505975 | -4.0279599049 |
| H95 | -1.6273779439 | -3.4830987272 | -4.6310109648 |
| H96 | -1.2835727950 | -4.1116912228 | -3.0092156491 |
| C97 | -4.4873427572 | -0.4987645382 | 5.8482465566 |
| H98 | -5.0063818510 | -1.2485102686 | 6.4460898258 |
| H99 | -4.4277835291 | 0.4343430638 | 6.4084588668 |
| H100 | -3.4814226313 | -0.8669383220 | 5.6444773976 |
| C101 | -5.1606192880 | -0.2608719031 | 4.4818988668 |
| H102 | -5.1016158753 | 0.8043732213 | 4.2166109416 |
| H103 | -6.2312460370 | -0.4941454049 | 4.5000494304 |
| C104 | -4.5277687591 | -1.0166630674 | 3.2788277024 |
| O105 | -3.3902647493 | -1.5261334846 | 3.4110518233 |
| O106 | -5.2520717333 | -1.0024749958 | 2.2303801698 |
| H107 | -4.8851310437 | -1.7635498368 | 0.9148400977 |
| C108 | 4.5520186405 | 1.7004875061 | -4.4950590180 |
| H109 | 5.2814670828 | 1.4774105523 | -5.2743607385 |
| H110 | 4.9277009186 | 2.4408636111 | -3.7894553533 |
| H111 | 3.6170379969 | 2.0773585338 | -4.9104476314 |
| C112 | 4.3510530979 | 0.3821867685 | -3.7195457965 |
| H113 | 5.3410193686 | -0.0105449231 | -3.4397835588 |
| H114 | 3.8948271625 | -0.3759766835 | -4.3661338541 |
| C115 | 3.4979121244 | 0.5211864715 | -2.5039339333 |
| N116 | 3.6350365485 | 1.5380101394 | -1.5679850063 |
| C117 | 2.4624765136 | -0.2371586713 | -2.0355916122 |
| H118 | 2.0128875483 | -1.1064080931 | -2.4901910917 |
| C119 | 2.6907592116 | 1.3718207142 | -0.6092874979 |
| H120 | 2.4950168933 | 2.0651774831 | 0.1951035481 |
| N121 | 1.9775354232 | 0.2883253226 | -0.8578914900 |
| C122 | 1.7140839117 | 4.0038724725 | -5.6526969122 |
| H123 | 1.3511983033 | 3.2726892072 | -6.3754515343 |
| H124 | 2.5651222273 | 3.5945550762 | -5.1079084236 |
| H125 | 2.0108383136 | 4.9162105712 | -6.1698881709 |
| C126 | 0.5550028010 | 4.2827813867 | -4.6759755272 |
| H127 | -0.2459659706 | 4.8091706011 | -5.2136234730 |
| H128 | 0.8704191968 | 4.9222072257 | -3.8456168344 |
| C129 | -0.0556166254 | 2.9898235300 | -4.1085908594 |
| O130 | -0.4144499390 | 2.1103463269 | -4.9367308739 |
| O131 | -0.1483961311 | 2.9201693178 | -2.8395174969 |
| O132 | -4.8467490944 | -2.2485728220 | 0.0116155024 |
| C133 | -5.7147505967 | -3.3526195600 | 0.0986325490 |
| H134 | -5.4313561086 | -4.0546007422 | 0.9007727044 |
| H135 | -6.7596353669 | -3.0506177377 | 0.2798984326 |
| H136 | -5.6834963972 | -3.9007972500 | -0.8538277494 |
| C137 | -2.4764419584 | -1.9672614988 | -1.1792125671 |
| H138 | -2.5698102033 | -2.6426173590 | -0.3416552127 |

| | | | |
|---|---|---|---|
| C139 | -2.6358296185 | -0.4845445216 | -0.9787927047 |
| H140 | -3.7240094038 | -0.3414668913 | -0.9418457533 |
| O141 | -2.0694419001 | -0.1232845857 | 0.2527063422 |
| H142 | -2.2437428149 | 0.8887758599 | 0.4725480373 |
| C143 | -2.0515200764 | 0.3421589721 | -2.1796193062 |
| H144 | -2.6575405669 | 1.2487224096 | -2.2888205804 |
| O145 | -0.7096742645 | 0.6753216628 | -1.9084981905 |
| H146 | -0.4846285689 | 1.6291476842 | -2.3103940369 |
| C147 | -2.1315057736 | -0.4872186606 | -3.4672601058 |
| H148 | -1.2079468703 | -1.0866956831 | -3.4947713171 |
| O149 | -2.2829009772 | 0.2439537671 | -4.6649366284 |
| H150 | -1.5630723288 | 0.9559445307 | -4.7132795601 |
| C151 | -3.2953152539 | -1.4740111042 | -3.3470356034 |
| H152 | -4.2009815007 | -0.9963755494 | -2.9600297071 |
| C153 | -3.6481758813 | -2.1740138978 | -4.6569397579 |
| H154 | -4.4357925615 | -1.5887141566 | -5.1446950429 |
| H155 | -4.0361473128 | -3.1787648647 | -4.4637529208 |
| O156 | -2.5306195486 | -2.3014094352 | -5.5467758180 |
| H157 | -2.2031185473 | -1.3711671636 | -5.6313016813 |
| O158 | -2.8848261777 | -2.4631382858 | -2.3263317326 |
| H159 | 2.2311337377 | -0.8177396706 | 4.5264864762 |
| H160 | 4.2700107303 | 2.3208409972 | -1.6271516739 |
| Covalent intermediate (Model 3) | | | |
| $E_{absolute}$ = -3477.20503780738 hartrees | | | |
| Zn1 | -0.0266398513 | -0.0045068011 | 0.0053509863 |
| C2 | 0.0212312592 | 0.0203359865 | 6.4467324273 |
| H3 | 1.1105887514 | 0.0207457068 | 6.4281155764 |
| H4 | -0.3965407250 | -0.1950061606 | 7.4302957431 |
| H5 | -0.3305341494 | 0.9974652972 | 6.1155986863 |
| C6 | -0.4414812009 | -1.1055401728 | 5.4885209667 |
| H7 | -1.5268637403 | -1.1228897723 | 5.3604385865 |
| H8 | -0.1646789159 | -2.0737044412 | 5.9324592051 |
| C9 | 0.1365283941 | -0.9941228126 | 4.1160080368 |
| N10 | 1.4909134169 | -0.8184215740 | 3.8558889971 |
| C11 | -0.4659543879 | -1.0044136385 | 2.8845921732 |
| H12 | -1.5263745304 | -1.1326680855 | 2.6800462692 |
| C13 | 1.6604852164 | -0.7156244749 | 2.5160796260 |
| H14 | 2.6151483261 | -0.5475829959 | 2.0398315651 |
| N15 | 0.4942952058 | -0.8307121381 | 1.9006331271 |
| C16 | 0.2994841790 | 4.7299154920 | 1.9453203499 |
| H17 | 1.2383052082 | 4.2551284198 | 1.6586652430 |
| H18 | -0.0524478894 | 4.3787612203 | 2.9146610714 |
| H19 | 0.3673395138 | 5.8167980133 | 1.9089728037 |
| C20 | -0.7455757344 | 4.2649235172 | 0.8982359060 |
| H21 | -0.4438854836 | 4.6012343722 | -0.1037439436 |
| H22 | -1.7280207008 | 4.6996701070 | 1.1022675523 |
| C23 | -0.9297941588 | 2.7451188545 | 0.7843823516 |
| O24 | 0.1157002571 | 2.0085944288 | 0.7593375572 |
| O25 | -2.1131454033 | 2.3150459241 | 0.6792397712 |
| C26 | -3.2301343683 | 6.4374233256 | -3.0990299128 |
| H27 | -2.6180724363 | 6.2005873151 | -3.9694775834 |
| H28 | -2.7462461187 | 7.1750231632 | -2.4581440450 |
| H29 | -4.2320359080 | 6.7544346214 | -3.3890430534 |
| C30 | -3.4036658380 | 5.1365337035 | -2.2666224423 |
| H31 | -3.8900299779 | 5.3846075316 | -1.3130947504 |

| | | | |
|---|---|---|---|
| H32 | -2.4162440988 | 4.7354501599 | -1.9977009250 |
| C33 | -4.2063803584 | 4.0556836152 | -2.9488074434 |
| C34 | -4.4855274801 | 3.9442870266 | -4.2883349961 |
| H35 | -4.2052600576 | 4.5912320641 | -5.1075595084 |
| C36 | -4.8277411863 | 2.9069204440 | -2.3188386210 |
| N37 | -5.2494556733 | 2.8129188277 | -4.5307227664 |
| H38 | -5.5103024498 | 2.4775525348 | -5.4429163617 |
| C39 | -5.4643714376 | 2.1491249645 | -3.3413743107 |
| C40 | -4.8995652639 | 2.4438705540 | -0.9890918197 |
| H41 | -4.3880140808 | 2.9747257074 | -0.1926265054 |
| C42 | -6.1577416244 | 0.9623574207 | -3.0696764179 |
| H43 | -6.6464765537 | 0.4031188507 | -3.8647596280 |
| C44 | -5.5809416866 | 1.2615530273 | -0.7137161676 |
| H45 | -5.6118128872 | 0.8713857293 | 0.3013446375 |
| C46 | -6.2032215316 | 0.5275384880 | -1.7472830621 |
| H47 | -6.7188082951 | -0.3959779142 | -1.5002340992 |
| C48 | -6.9432977016 | 7.2794484281 | -2.9897640407 |
| H49 | -7.0468419988 | 7.3076678035 | -4.0747519052 |
| H50 | -5.9955061427 | 6.8439081582 | -2.6738313810 |
| H51 | -7.1324623503 | 8.2479023017 | -2.5272455599 |
| C52 | -8.0668842251 | 6.3513285665 | -2.4476816942 |
| H53 | -9.0297027269 | 6.8555020794 | -2.6298482374 |
| C54 | -7.9132452469 | 6.1775649506 | -0.9232879198 |
| H55 | -7.7744992813 | 7.1698289839 | -0.4702180141 |
| H56 | -6.9865721136 | 5.6200099899 | -0.7243483542 |
| C57 | -8.0908460726 | 5.0090245800 | -3.1909136119 |
| H58 | -8.9169136728 | 4.3709100638 | -2.8602218939 |
| H59 | -8.2086315424 | 5.1628463973 | -4.2705149148 |
| H60 | -7.1588641540 | 4.4568602206 | -3.0328792353 |
| C61 | -9.0849966382 | 5.4765910903 | -0.2278097127 |
| H62 | -8.9308390933 | 5.4356879546 | 0.8557759740 |
| H63 | -10.0279904862 | 6.0068579801 | -0.4091493048 |
| H64 | -9.2093772383 | 4.4464646994 | -0.5762925765 |
| C65 | 2.3493184806 | -4.3608671861 | -0.5251587427 |
| H66 | 1.5853953606 | -5.1380557302 | -0.5562744843 |
| H67 | 3.0821243103 | -4.4877367807 | 0.2723024940 |
| H68 | 2.7853698512 | -4.1276091706 | -1.4966634913 |
| C69 | 1.3988920526 | -3.1763051286 | -0.1633003577 |
| H70 | 1.8465170424 | -2.1951422488 | -0.3144956131 |
| H71 | 1.1346271778 | -3.2138928211 | 0.9007951709 |
| C72 | 0.0463008790 | -3.2556364404 | -0.8899573239 |
| O73 | -0.6095926385 | -2.1051642411 | -0.9762493548 |
| O74 | -0.4022039516 | -4.3184918197 | -1.3264086027 |
| C75 | 3.0546996329 | -10.2411936745 | -2.8706106329 |
| H76 | 2.1379695440 | -10.8001684404 | -2.6840879254 |
| H77 | 3.6022244573 | -9.9786003929 | -1.9656165912 |
| H78 | 3.6789456592 | -10.6673383365 | -3.6559003104 |
| C79 | 2.6772489264 | -8.9335709262 | -3.6260442631 |
| H80 | 2.2887666744 | -9.2338194591 | -4.6102802965 |
| H81 | 3.6080384757 | -8.3869995267 | -3.8356547553 |
| C82 | 1.6709506553 | -7.9484655312 | -3.0212470050 |
| H83 | 2.0482201476 | -7.5188963544 | -2.0840524868 |
| H84 | 0.7186155923 | -8.4437715918 | -2.7952239262 |
| C85 | 1.4439624824 | -6.8433364214 | -4.0603516427 |
| H86 | 1.0733107808 | -7.3105475243 | -4.9807134332 |
| H87 | 2.4125400256 | -6.3752621560 | -4.2981722369 |

| | | | |
|---|---|---|---|
| N88 | 0.4793131745 | -5.8311732820 | -3.6508625753 |
| H89 | 0.3453623817 | -5.6122613528 | -2.6636624704 |
| C90 | -0.1627226026 | -5.0275692712 | -4.5105942322 |
| N91 | 0.0861055132 | -5.1158995138 | -5.8428460781 |
| H92 | -0.4052901563 | -4.4784089859 | -6.4527638626 |
| H93 | 0.9649441283 | -5.4790615663 | -6.1726341010 |
| N94 | -1.0761276241 | -4.1745917275 | -4.0887023457 |
| H95 | -1.4513110000 | -3.4447591849 | -4.7198091784 |
| H96 | -1.3049955016 | -4.1184642622 | -3.0966827338 |
| C97 | -4.4743164699 | -0.4675661984 | 5.9422516400 |
| H98 | -4.9934956183 | -1.2099500112 | 6.5490927193 |
| H99 | -4.4067691957 | 0.4696273004 | 6.4946832053 |
| H100 | -3.4714760957 | -0.8423847315 | 5.7354392584 |
| C101 | -5.1429961292 | -0.2548293100 | 4.5704015368 |
| H102 | -5.0998417001 | 0.8057250145 | 4.2874937419 |
| H103 | -6.2085915690 | -0.5078883532 | 4.5784962143 |
| C104 | -4.4805682432 | -1.0212376101 | 3.3901755205 |
| O105 | -3.3396388711 | -1.5153046185 | 3.5571353912 |
| O106 | -5.1790642729 | -1.0365757199 | 2.3260014952 |
| H107 | -4.7865078638 | -1.8622247704 | 1.0775348674 |
| C108 | 4.5607265153 | 1.6026027476 | -4.4484828531 |
| H109 | 5.2917874857 | 1.3689565375 | -5.2231627003 |
| H110 | 4.9383470973 | 2.3459581121 | -3.7470586240 |
| H111 | 3.6297968274 | 1.9822028077 | -4.8704332869 |
| C112 | 4.3570796443 | 0.2922919515 | -3.6613880053 |
| H113 | 5.3456700746 | -0.0956320117 | -3.3726065434 |
| H114 | 3.9049017844 | -0.4738091242 | -4.3003204238 |
| C115 | 3.4988900640 | 0.4419134187 | -2.4513741329 |
| N116 | 3.6237774831 | 1.4705317175 | -1.5275983499 |
| C117 | 2.4638199525 | -0.3158030325 | -1.9824181955 |
| H118 | 2.0243017764 | -1.1904143187 | -2.4355979460 |
| C119 | 2.6734310120 | 1.3112447474 | -0.5734570521 |
| H120 | 2.4650447849 | 2.0135175968 | 0.2201869431 |
| N121 | 1.9670843337 | 0.2210968263 | -0.8145887646 |
| C122 | 1.7590725816 | 3.9513154250 | -5.6034144696 |
| H123 | 1.3811655158 | 3.2244221652 | -6.3227903460 |
| H124 | 2.6047608966 | 3.5289321777 | -5.0602707203 |
| H125 | 2.0697790863 | 4.8568906799 | -6.1242846179 |
| C126 | 0.6027389764 | 4.2345990099 | -4.6250083521 |
| H127 | -0.1897418695 | 4.7803176071 | -5.1539304384 |
| H128 | 0.9232252598 | 4.8517506457 | -3.7808726449 |
| C129 | -0.0255497103 | 2.9380670730 | -4.0864261264 |
| O130 | -0.4030683772 | 2.0810314898 | -4.9276868865 |
| O131 | -0.1096874198 | 2.8404874153 | -2.8170827321 |
| O132 | -4.7183279167 | -2.3963631460 | 0.2071950740 |
| C133 | -5.7533552771 | -3.3450900121 | 0.2171246818 |
| H134 | -6.7511015593 | -2.8866052100 | 0.3214162881 |
| H135 | -5.7410270948 | -3.8980833317 | -0.7332880762 |
| H136 | -5.6522968405 | -4.0851517450 | 1.0297789252 |
| C137 | -2.2584767373 | -2.0679723194 | -1.1643928938 |
| H138 | -2.5827301937 | -2.6978963966 | -0.3397258702 |
| C139 | -2.5783374720 | -0.5878507084 | -0.9904887748 |
| H140 | -3.6717981175 | -0.5477716430 | -0.9747455960 |
| O141 | -2.0760839316 | -0.1593600340 | 0.2543372118 |
| H142 | -2.2508459322 | 0.8612395173 | 0.4354792295 |
| C143 | -2.0182568541 | 0.2730657368 | -2.1702016906 |

| | | | |
|---|---|---|---|
| H144 | -2.6523866038 | 1.1616998370 | -2.2729770334 |
| O145 | -0.6851783565 | 0.6410900463 | -1.8849135375 |
| H146 | -0.4560844453 | 1.6033321903 | -2.3070729888 |
| C147 | -2.0573288176 | -0.5369019204 | -3.4707655333 |
| H148 | -1.1014188663 | -1.0842548188 | -3.5102469589 |
| O149 | -2.2435802124 | 0.2091322839 | -4.6558687580 |
| H150 | -1.5355083227 | 0.9286214508 | -4.7091229160 |
| C151 | -3.1524208505 | -1.5971541402 | -3.3464536543 |
| H152 | -4.0820079140 | -1.1749102933 | -2.9515887189 |
| C153 | -3.4889707863 | -2.2762938570 | -4.6734561505 |
| H154 | -4.3162606938 | -1.7240020624 | -5.1323600331 |
| H155 | -3.8176083710 | -3.3064111377 | -4.5053805164 |
| O156 | -2.3793795805 | -2.3098321842 | -5.5858921330 |
| H157 | -2.1087883396 | -1.3619552099 | -5.6405795937 |
| O158 | -2.6636221651 | -2.5766406914 | -2.3656133515 |
| H159 | 2.2242470801 | -0.7531267034 | 4.5454522220 |
| H160 | 4.2563920284 | 2.2541237276 | -1.5891635640 |

Table S4. Optimized geometries of Michaelis complex, transition state and glycosyl-enzyme covalent intermediate for Model 4.

| Michaelis complex<br>Eabsolute = -3419.84999610339 hartrees | | | |
|---|---|---|---|
| C1 | -0.0136617122 | 0.0263772276 | -0.0033454692 |
| H2 | -0.0125235050 | 0.0227314633 | 1.0677337105 |
| H3 | 0.9988983893 | 0.0261400407 | -0.3501462705 |
| H4 | -0.5043604874 | 0.9272321443 | -0.3110995909 |
| C5 | -1.7013591156 | -0.9660647165 | 6.1536711551 |
| H6 | -2.6917547740 | -0.5670715974 | 6.2117271044 |
| H7 | -1.1006910237 | -0.5412631660 | 6.9295846287 |
| H8 | -1.2764056563 | -0.7230233743 | 5.2009135610 |
| C9 | -2.0296403239 | -5.8873697069 | 10.0020610005 |
| H10 | -3.0901207061 | -6.0551085610 | 10.2000334071 |
| H11 | -1.7261518055 | -4.9329228284 | 10.4320598201 |
| H12 | -1.4441825125 | -6.6896444679 | 10.4494600031 |
| C13 | 0.9966510692 | -8.2122340362 | 10.6489452512 |
| H14 | 0.4641864046 | -9.0475321446 | 11.0518016080 |
| H15 | 1.5504428052 | -7.7408457812 | 11.4349221829 |
| H16 | 0.2878818523 | -7.4919042196 | 10.2418124523 |
| C17 | -7.6871217939 | -3.5790958934 | -0.8818299485 |
| H18 | -8.2407048277 | -4.4619565493 | -0.6466216196 |
| H19 | -8.3029677839 | -2.7167645236 | -0.7297931287 |
| H20 | -7.3727903577 | -3.6146479259 | -1.9046114062 |
| C21 | -9.5343180138 | -3.0399103556 | 6.3420834539 |
| H22 | -10.5540153811 | -3.2495387930 | 6.6650806270 |
| H23 | -9.2334046070 | -2.0553036248 | 6.7014527388 |
| H24 | -8.8612797368 | -3.7947612036 | 6.7484487489 |
| C25 | 2.5852073048 | -3.5534830187 | -0.1210956140 |
| H26 | 2.1559329432 | -2.6788412648 | 0.3213081365 |
| H27 | 3.6465019163 | -3.4351727837 | -0.1831488062 |
| H28 | 2.1825265573 | -3.6884139614 | -1.1029283107 |
| C29 | -7.4982127054 | -5.1515975422 | 8.8496374832 |

| | | | |
|---|---|---|---|
| H30 | -8.0216150837 | -6.0414999550 | 8.5730984790 |
| H31 | -7.7165390869 | -4.9273431544 | 9.8730296015 |
| H32 | -7.8413258933 | -4.3214725574 | 8.2330895785 |
| Zn33 | -4.2110508636 | -4.0770526849 | 2.8671415791 |
| C34 | -0.7542662529 | -1.1925669214 | -0.5934344878 |
| C35 | -1.9921576339 | -1.5360242666 | 0.1615110981 |
| N36 | -3.0269859532 | -0.6511692065 | 0.4356787814 |
| C37 | -3.9245876134 | -1.2891556278 | 1.2374818386 |
| N38 | -3.5241941234 | -2.5176460432 | 1.4850767710 |
| C39 | -2.3301293528 | -2.6857578630 | 0.8222941882 |
| C40 | -1.7160127103 | -2.5012116945 | 6.3126165722 |
| C41 | -2.1158907715 | -3.2585779759 | 5.0435806429 |
| O42 | -3.1460501163 | -2.8715959026 | 4.4100318602 |
| O43 | -1.3710743172 | -4.2286285995 | 4.7134805882 |
| C44 | -6.4601028158 | -3.5446940308 | 0.0436029872 |
| C45 | -5.8576341142 | -4.9440545164 | 0.2823912237 |
| O46 | -5.1808867453 | -5.1210526774 | 1.3589751241 |
| O47 | -6.0711773306 | -5.8201501959 | -0.5764306202 |
| C48 | -9.4522973133 | -3.0150743571 | 4.8055540292 |
| C49 | -8.0399683594 | -2.8081589169 | 4.3717975550 |
| N50 | -7.2162703012 | -1.8267169085 | 4.9141366650 |
| C51 | -5.9591136511 | -2.0114935863 | 4.4357268378 |
| N52 | -5.9414008411 | -3.0276998571 | 3.5989057956 |
| C53 | -7.2201037205 | -3.5356951960 | 3.5554577686 |
| H54 | -0.9718840693 | -1.0009632592 | -1.6544653625 |
| H55 | -0.1244586448 | -2.0859106294 | -0.5705694659 |
| H56 | -3.1058779607 | 0.2965562969 | 0.1007704216 |
| H57 | -4.8331663170 | -0.8345053146 | 1.6031233984 |
| H58 | -1.7919508232 | -3.6200042448 | 0.8573107174 |
| H59 | -0.7276999552 | -2.8653041765 | 6.6069543212 |
| H60 | -2.4187118427 | -2.7960749842 | 7.1023534907 |
| H61 | -5.6665473196 | -2.9140178746 | -0.3774619701 |
| H62 | -6.7091795736 | -3.1022821353 | 1.0117311877 |
| H63 | -9.8005827186 | -3.9609709124 | 4.3791670988 |
| H64 | -10.1156046397 | -2.2306666593 | 4.4116336013 |
| H65 | -7.4788344241 | -1.1522553362 | 5.6157323979 |
| H66 | -5.0847267718 | -1.4589861984 | 4.7445880115 |
| H67 | -7.4470888815 | -4.4104579655 | 2.9659110427 |
| C68 | -1.7573122859 | -5.8797649425 | 8.4779225628 |
| C69 | -1.9466898224 | -7.2166186088 | 7.8041521912 |
| C70 | -3.1021368309 | -7.9555369474 | 7.7643510968 |
| N71 | -2.9268787215 | -9.1293734815 | 7.0498595773 |
| C72 | -1.6391908772 | -9.1365413016 | 6.5492048194 |
| C73 | -1.0081165930 | -10.0732678403 | 5.7246678742 |
| C74 | 0.3151985002 | -9.8302168167 | 5.3621982564 |
| C75 | 0.9777550361 | -8.6590725140 | 5.7956970625 |
| C76 | 0.3390574290 | -7.7216732056 | 6.6060970569 |
| C77 | -0.9853327719 | -7.9563721738 | 7.0139830664 |
| H78 | -2.4102182771 | -5.1426485284 | 7.9941373530 |
| H79 | -0.7298672494 | -5.5379972426 | 8.3024441490 |
| H80 | -4.0813638396 | -7.7319307488 | 8.1605631415 |
| H81 | -3.7114142114 | -9.5084698096 | 6.5244069577 |
| H82 | -1.5446773369 | -10.9377398720 | 5.3424356824 |
| H83 | 0.8275196145 | -10.5378117598 | 4.7132893441 |
| H84 | 2.0118801292 | -8.4938544990 | 5.4997040032 |
| H85 | 0.8546290786 | -6.8129006262 | 6.9064211161 |

| | | | |
|---|---|---|---|
| C86 | 2.0029566762 | -8.7353071805 | 9.5834108476 |
| C87 | 3.4276815158 | -8.8107139659 | 10.1598999691 |
| C88 | 1.5573967392 | -10.0818868605 | 8.9678627075 |
| C89 | 1.6104082421 | -11.2956809970 | 9.9064558778 |
| H90 | 2.0300564916 | -8.0173674921 | 8.7558946153 |
| H91 | 3.4739445999 | -9.4510639349 | 11.0492528488 |
| H92 | 4.1367071192 | -9.2119905202 | 9.4252568228 |
| H93 | 3.7831992982 | -7.8164744008 | 10.4570165848 |
| H94 | 0.5384742741 | -9.9668810073 | 8.5823304382 |
| H95 | 2.1827226297 | -10.2892385222 | 8.0906109490 |
| H96 | 1.0040800862 | -11.1518280943 | 10.8088392774 |
| H97 | 1.2250225061 | -12.1863111365 | 9.3983487093 |
| H98 | 2.6327524893 | -11.5217461377 | 10.2308942680 |
| C99 | 2.2273327992 | -4.7933552265 | 0.7061980855 |
| C100 | 0.7691951419 | -5.1787689714 | 0.5309302734 |
| O101 | 0.0101858120 | -4.6252086604 | -0.2469148736 |
| O102 | 0.4087098877 | -6.2058584057 | 1.3028912064 |
| H103 | 2.8280837424 | -5.6613075229 | 0.4034407845 |
| H104 | 2.4212858144 | -4.6570454173 | 1.7765802992 |
| H105 | -0.5528202166 | -6.4765491553 | 1.1385014591 |
| C106 | -5.9873103383 | -5.3481261582 | 8.6623824995 |
| C107 | -5.6480384624 | -5.8130891958 | 7.2390736449 |
| O108 | -6.1316769957 | -6.9142792340 | 6.8730623606 |
| O109 | -4.9008059573 | -5.0387529455 | 6.5576986610 |
| H110 | -5.4442883390 | -4.4233971535 | 8.8805622211 |
| H111 | -5.6367471210 | -6.1179959374 | 9.3627703748 |
| O112 | -5.1927342613 | -8.4868649400 | -0.5992232723 |
| H113 | -5.5302027056 | -7.5655286861 | -0.5366251542 |
| H114 | -4.9220431910 | -8.7133352027 | 0.3014156068 |
| O115 | 0.9275882715 | -8.6744739125 | 2.5451053342 |
| H116 | 0.7468182144 | -8.5348669927 | 3.4891390499 |
| H117 | 0.9636534382 | -7.7939373183 | 2.1344125247 |
| O118 | -1.1019509344 | -10.3797413798 | 1.5316566960 |
| H119 | -0.4754969625 | -9.7421265841 | 1.9299071806 |
| H120 | -1.8206303400 | -9.8091520789 | 1.2110848328 |
| O121 | 1.6060325564 | -11.4655335926 | 2.4768464484 |
| H122 | 1.6802023907 | -10.5012781056 | 2.3937533920 |
| H123 | 0.7263858635 | -11.6320587699 | 2.1085783984 |
| O124 | -1.8784358571 | -7.3353536064 | 0.8203551100 |
| C125 | -2.1702642907 | -7.2908974608 | -0.5988016051 |
| C126 | -3.0204825750 | -7.3363282102 | 1.6720359334 |
| C127 | -2.6436697331 | -6.6535607124 | 2.9873054203 |
| O128 | -2.3878808823 | -5.2886783767 | 2.7033031963 |
| C129 | -3.7305908955 | -6.8335138877 | 4.1042947888 |
| O130 | -4.5737057678 | -5.6989574617 | 4.1792177464 |
| C131 | -4.5556535474 | -8.1217747514 | 3.8926131533 |
| O132 | -5.0654965887 | -8.6511531968 | 5.1001653437 |
| C133 | -3.6312203903 | -9.1484418526 | 3.2230265433 |
| O134 | -3.4065686037 | -8.7019895095 | 1.8583202876 |
| C135 | -4.1288930498 | -10.5935328402 | 3.1648197533 |
| O136 | -3.0773332348 | -11.5435227627 | 3.2259326917 |
| H137 | -2.2096894962 | -6.2489029217 | -0.9310987452 |
| H138 | -1.3453926660 | -7.8042238114 | -1.1008549268 |
| H139 | -3.1185871739 | -7.7963046783 | -0.8095139731 |
| H140 | -3.8392104964 | -6.7968885934 | 1.1895415898 |
| H141 | -1.7164907828 | -7.1232594137 | 3.3446793410 |

| | | | |
|---|---|---|---|
| H142 | -1.8902556673 | -4.8693736135 | 3.5274486024 |
| H143 | -3.1960127412 | -6.9366808977 | 5.0569640182 |
| H144 | -4.7191083592 | -5.4382320499 | 5.1846229718 |
| H145 | -5.3726955162 | -7.8829804086 | 3.1926478011 |
| H146 | -5.4977835530 | -7.9312146727 | 5.6351049351 |
| H147 | -2.6755828172 | -9.1734792718 | 3.7626220453 |
| H148 | -4.7449465619 | -10.7120918275 | 2.2575058842 |
| H149 | -4.7728392328 | -10.7719152158 | 4.0303943174 |
| H150 | -2.3562237739 | -11.2502918922 | 2.6312948604 |

Transition state (Model 3)

$E_{absolute}$ = -3419.81809396550 hartrees

| | | | |
|---|---|---|---|
| C1 | 0.0000000000 | 0.0000000000 | 0.0000000000 |
| H2 | 0.0000000000 | 0.0000000000 | 1.0710859869 |
| H3 | 1.0129296111 | 0.0000000000 | -0.3457200542 |
| H4 | -0.4916254253 | 0.8991154856 | -0.3113411458 |
| C5 | -1.6928754250 | -0.9738480410 | 6.1585643246 |
| H6 | -2.6838888477 | -0.5760470258 | 6.2142088115 |
| H7 | -1.0936291346 | -0.5455670576 | 6.9336640290 |
| H8 | -1.2675549188 | -0.7328815882 | 5.2054435526 |
| C9 | -2.0183759825 | -5.8824833176 | 10.0233373224 |
| H10 | -3.0788311531 | -6.0510339487 | 10.2207545244 |
| H11 | -1.7166807176 | -4.9261540690 | 10.4504057544 |
| H12 | -1.4322739066 | -6.6824091800 | 10.4740861440 |
| C13 | 1.0104728023 | -8.2008869941 | 10.6813436119 |
| H14 | 0.4787478708 | -9.0355551009 | 11.0864763566 |
| H15 | 1.5627667626 | -7.7260506125 | 11.4662981992 |
| H16 | 0.3011311245 | -7.4829414222 | 10.2710089926 |
| C17 | -7.6674810144 | -3.6191649314 | -0.8743904834 |
| H18 | -8.2200840743 | -4.5019934261 | -0.6367694040 |
| H19 | -8.2846915677 | -2.7571831987 | -0.7259466486 |
| H20 | -7.3520050853 | -3.6577564408 | -1.8967092904 |
| C21 | -9.5231724148 | -3.0579882329 | 6.3456702867 |
| H22 | -10.5429222323 | -3.2679419410 | 6.6682903803 |
| H23 | -9.2240177499 | -2.0717448137 | 6.7020071795 |
| H24 | -8.8495181197 | -3.8105105633 | 6.7553201558 |
| C25 | 2.6039858388 | -3.5766003605 | -0.1027976154 |
| H26 | 2.1730187455 | -2.7010597951 | 0.3361691664 |
| H27 | 3.6651800935 | -3.4570174576 | -0.1641233668 |
| H28 | 2.2025438301 | -3.7154343322 | -1.0845934340 |
| C29 | -7.4868128088 | -5.1582846124 | 8.8625671009 |
| H30 | -8.0086770197 | -6.0498535399 | 8.5885000175 |
| H31 | -7.7065482510 | -4.9308558490 | 9.8849566354 |
| H32 | -7.8304223816 | -4.3307422380 | 8.2428320415 |
| Zn33 | -4.2439469913 | -3.7761458922 | 2.9073781739 |
| C34 | -0.7182042977 | -1.2201945210 | -0.6098676516 |
| C35 | -2.0043430797 | -1.5063663142 | 0.0785190217 |
| N36 | -3.0290372950 | -0.5882388851 | 0.2727782025 |
| C37 | -3.9615478722 | -1.1626531745 | 1.0790571528 |
| N38 | -3.5915326909 | -2.3854471853 | 1.4022088401 |
| C39 | -2.3809924346 | -2.6149786034 | 0.7834100092 |
| C40 | -1.7172472932 | -2.5065869919 | 6.3104074568 |
| C41 | -2.1694820056 | -3.2273750537 | 5.0415734466 |
| O42 | -3.1599326269 | -2.7330991804 | 4.4088131668 |
| O43 | -1.5190502049 | -4.2574389739 | 4.7071022490 |
| C44 | -6.4506554635 | -3.5441932144 | 0.0659476316 |

| | | | |
|---|---|---|---|
| C45 | -5.7846875815 | -4.9063627036 | 0.3363138824 |
| O46 | -5.2313252078 | -5.0527267031 | 1.5004319022 |
| O47 | -5.8066516419 | -5.7724084928 | -0.5505404860 |
| C48 | -9.4837338754 | -3.0282052912 | 4.8074908253 |
| C49 | -8.1143075289 | -2.6937883309 | 4.3267396386 |
| N50 | -7.4124054205 | -1.5672659178 | 4.7434465875 |
| C51 | -6.1450106568 | -1.6531361504 | 4.2620739389 |
| N52 | -6.0093295948 | -2.7479330687 | 3.5449608730 |
| C53 | -7.2179815003 | -3.4051645462 | 3.5815045289 |
| H54 | -0.8629174985 | -1.0463602009 | -1.6871625559 |
| H55 | -0.1160357809 | -2.1303712487 | -0.5187359413 |
| H56 | -3.0726941932 | 0.3448179346 | -0.1093081620 |
| H57 | -4.8734357831 | -0.6740706004 | 1.3902370370 |
| H58 | -1.8472438433 | -3.5516354477 | 0.8701009159 |
| H59 | -0.7263694611 | -2.8907914183 | 6.5714708960 |
| H60 | -2.4018056563 | -2.8021432393 | 7.1166535001 |
| H61 | -5.6773491001 | -2.8906144110 | -0.3576003705 |
| H62 | -6.7321393927 | -3.0949477007 | 1.0214432177 |
| H63 | -9.7628278190 | -4.0018729884 | 4.3912239115 |
| H64 | -10.2262726909 | -2.3046481722 | 4.4376636518 |
| H65 | -7.7583190345 | -0.8456219494 | 5.3584112530 |
| H66 | -5.3567625283 | -0.9485160109 | 4.4816850866 |
| H67 | -7.3359482102 | -4.3618922347 | 3.0962472146 |
| C68 | -1.7413249609 | -5.8171782059 | 8.5011925438 |
| C69 | -1.9258106170 | -7.0986645644 | 7.7283301994 |
| C70 | -3.0667014527 | -7.8522759956 | 7.6258318795 |
| N71 | -2.8764346249 | -8.9412444756 | 6.7865733071 |
| C72 | -1.5943254067 | -8.8661467551 | 6.2730100798 |
| C73 | -0.9391824149 | -9.7064755284 | 5.3687147025 |
| C74 | 0.3840169056 | -9.4133008379 | 5.0517737343 |
| C75 | 1.0299822275 | -8.2895211822 | 5.6104126149 |
| C76 | 0.3689525611 | -7.4453838538 | 6.4983810429 |
| C77 | -0.9589171038 | -7.7320090454 | 6.8576430306 |
| H78 | -2.3865558388 | -5.0467468450 | 8.0592883160 |
| H79 | -0.7114292925 | -5.4675840932 | 8.3527979497 |
| H80 | -4.0416372266 | -7.7001289784 | 8.0640785367 |
| H81 | -3.6566131924 | -9.3144338314 | 6.2520529027 |
| H82 | -1.4416070827 | -10.5622611793 | 4.9251997551 |
| H83 | 0.9175354332 | -10.0642268293 | 4.3650515826 |
| H84 | 2.0660307140 | -8.0855860313 | 5.3462930915 |
| H85 | 0.8731638416 | -6.5711272633 | 6.9031913892 |
| C86 | 1.9964566802 | -8.7216764902 | 9.6013475874 |
| C87 | 3.3761487192 | -9.0208306140 | 10.2124455242 |
| C88 | 1.4283060799 | -9.9359514024 | 8.8298317866 |
| C89 | 1.3301736602 | -11.2475345944 | 9.6227838669 |
| H90 | 2.1385807270 | -7.9258557689 | 8.8583240197 |
| H91 | 3.3092693231 | -9.7474220656 | 11.0326140040 |
| H92 | 4.0660590778 | -9.4292499712 | 9.4623947950 |
| H93 | 3.8296782771 | -8.1086992132 | 10.6216222293 |
| H94 | 0.4366155909 | -9.6711008308 | 8.4461345442 |
| H95 | 2.0514796037 | -10.1086220211 | 7.9426942637 |
| H96 | 0.7168650136 | -11.1441347704 | 10.5271017934 |
| H97 | 0.8677562657 | -12.0281996719 | 9.0072133460 |
| H98 | 2.3153966467 | -11.6166529425 | 9.9339304809 |
| C99 | 2.2408283272 | -4.8264280638 | 0.7120133001 |
| C100 | 0.7847856595 | -5.2819240590 | 0.4952167151 |

| | | | |
|---|---|---|---|
| O101 | -0.0115916264 | -4.5392859598 | -0.1068104945 |
| O102 | 0.5098540191 | -6.4502655859 | 0.9609605831 |
| H103 | 2.8915130999 | -5.6693027200 | 0.4451474597 |
| H104 | 2.3916703561 | -4.6687398211 | 1.7889879963 |
| H105 | -0.8469513464 | -7.1279910216 | 0.6176032645 |
| C106 | -5.9757561341 | -5.3627435418 | 8.6948385947 |
| C107 | -5.6200533427 | -5.7714006576 | 7.2610333929 |
| O108 | -6.1162787530 | -6.8479805942 | 6.8360867731 |
| O109 | -4.8534721125 | -4.9748329806 | 6.6315860897 |
| H110 | -5.4284143035 | -4.4515562298 | 8.9564711894 |
| H111 | -5.6467156384 | -6.1614753392 | 9.3743314880 |
| O112 | -5.0825265986 | -8.5131520610 | -1.3904134361 |
| H113 | -5.3618510074 | -7.6458515619 | -1.0369372562 |
| H114 | -4.9692387710 | -9.0734535519 | -0.6038446921 |
| O115 | 1.8164491436 | -8.3023223612 | 2.3325599004 |
| H116 | 1.4079020990 | -8.1085134538 | 3.1922635133 |
| H117 | 1.4416301136 | -7.5850208183 | 1.7456805917 |
| O118 | -1.9177476996 | -10.2003579900 | 1.4776226273 |
| H119 | -1.0062377147 | -10.5103158180 | 1.6957167399 |
| H120 | -1.7825559216 | -9.3505049843 | 0.9989524302 |
| O121 | 0.7697846284 | -10.7467914936 | 1.8657361402 |
| H122 | 1.1502811116 | -9.8302460158 | 1.9936495632 |
| H123 | 1.0794616348 | -11.0156449406 | 0.9880161646 |
| O124 | -1.6710986234 | -7.6686266398 | 0.3505498701 |
| C125 | -1.8092573971 | -7.5502744457 | -1.0693691194 |
| C126 | -3.6334662831 | -6.8650970997 | 1.4312525698 |
| C127 | -2.9879116157 | -6.4215465917 | 2.7074969510 |
| O128 | -2.5469380375 | -5.0997388202 | 2.5790456424 |
| C129 | -3.9500833382 | -6.6262920161 | 3.9404761674 |
| O130 | -4.6570338797 | -5.4342022661 | 4.1888883619 |
| C131 | -4.9288666503 | -7.8051511347 | 3.7199255545 |
| O132 | -5.2019239006 | -8.5296926222 | 4.9015519937 |
| C133 | -4.3497581552 | -8.7533213187 | 2.6705301361 |
| O134 | -4.2815183723 | -7.9659412238 | 1.4163967805 |
| C135 | -5.1487497642 | -9.9986253666 | 2.3158376535 |
| O136 | -4.6231479235 | -10.6515990453 | 1.1791826068 |
| H137 | -1.6611728882 | -6.5074327595 | -1.3759436889 |
| H138 | -1.0618107588 | -8.1737096618 | -1.5825223419 |
| H139 | -2.8122552438 | -7.8909142104 | -1.3468228702 |
| H140 | -3.6079370429 | -6.3199220685 | 0.5045700659 |
| H141 | -2.1314259891 | -7.0979528205 | 2.8486938930 |
| H142 | -2.0189186725 | -4.8164298231 | 3.4423016231 |
| H143 | -3.3122964178 | -6.8609570170 | 4.8006632111 |
| H144 | -4.7251113247 | -5.2667927921 | 5.2223822637 |
| H145 | -5.8512643202 | -7.3729737227 | 3.3032474044 |
| H146 | -5.5699391145 | -7.8812841607 | 5.5692755468 |
| H147 | -3.3202763384 | -9.0507568758 | 2.8938777741 |
| H148 | -6.1866931154 | -9.7280429908 | 2.0874098830 |
| H149 | -5.1608179118 | -10.6401641025 | 3.2107833096 |
| H150 | -3.6379635463 | -10.6847951813 | 1.2657726007 |
| Covalent intermediate (Model 3)<br>$E_{absolute}$ = -3419.83818022281 hartrees | | | |
| C1 | 0.0060422221 | -0.0047110369 | 0.0052853278 |
| H2 | 0.0064762705 | -0.0058848107 | 1.0763705836 |
| H3 | 1.0188290395 | -0.0066306726 | -0.3408474952 |

| | | | |
|---|---|---|---|
| H4 | -0.4836674226 | 0.8958583661 | -0.3048706625 |
| C5 | -1.6865432541 | -0.9814635716 | 6.1634693657 |
| H6 | -2.6766286399 | -0.5814760248 | 6.2199534500 |
| H7 | -1.0860124206 | -0.5553931477 | 6.9387931570 |
| H8 | -1.2614052624 | -0.7397872705 | 5.2104468981 |
| C9 | -2.0216174988 | -5.8935799284 | 10.0229977524 |
| H10 | -3.0823723948 | -6.0599400137 | 10.2206626421 |
| H11 | -1.7175794698 | -4.9384063250 | 10.4509900121 |
| H12 | -1.4371498428 | -6.6953271780 | 10.4726314207 |
| C13 | 1.0022280030 | -8.2195703053 | 10.6772301217 |
| H14 | 0.4687742855 | -9.0534731603 | 11.0816654711 |
| H15 | 1.5559162869 | -7.7468488878 | 11.4624789087 |
| H16 | 0.2943513862 | -7.4995677229 | 10.2679711124 |
| C17 | -7.6699868687 | -3.6055074285 | -0.8699412095 |
| H18 | -8.2244958494 | -4.4873395747 | -0.6330616479 |
| H19 | -8.2851792544 | -2.7422905508 | -0.7203019867 |
| H20 | -7.3550136457 | -3.6436943408 | -1.8924302011 |
| C21 | -9.5214738902 | -3.0480340243 | 6.3514856979 |
| H22 | -10.5415667781 | -3.2560266283 | 6.6742914464 |
| H23 | -9.2199372218 | -2.0628630782 | 6.7087802839 |
| H24 | -8.8493633303 | -3.8025315431 | 6.7600366388 |
| C25 | 2.6018620345 | -3.5870962415 | -0.1024901679 |
| H26 | 2.1730611054 | -2.7110615817 | 0.3376107190 |
| H27 | 3.6633000213 | -3.4698545226 | -0.1641176295 |
| H28 | 2.1997081200 | -3.7239429123 | -1.0842736498 |
| C29 | -7.4888666682 | -5.1557028584 | 8.8652507758 |
| H30 | -8.0128640861 | -6.0457844000 | 8.5904204122 |
| H31 | -7.7076710423 | -4.9288966251 | 9.8879782300 |
| H32 | -7.8308482478 | -4.3267042344 | 8.2465623693 |
| Zn33 | -4.2247574777 | -3.5094145815 | 3.0563445600 |
| C34 | -0.7264585664 | -1.2236331755 | -0.5919321286 |
| C35 | -2.0329969194 | -1.4687773644 | 0.0784237702 |
| N36 | -3.0410299199 | -0.5258521399 | 0.2361424164 |
| C37 | -4.0047152752 | -1.0624201408 | 1.0282653774 |
| N38 | -3.6728152448 | -2.2934067269 | 1.3784578998 |
| C39 | -2.4516589286 | -2.5605613610 | 0.7878233518 |
| C40 | -1.6856974325 | -2.5058116625 | 6.3378104946 |
| C41 | -2.2159100570 | -3.2100480355 | 5.1028959253 |
| O42 | -3.1635553255 | -2.5943993708 | 4.4918282192 |
| O43 | -1.6984538819 | -4.2989930803 | 4.7565404572 |
| C44 | -6.4424970362 | -3.5940239116 | 0.0532960775 |
| C45 | -5.6585743277 | -4.9015508241 | -0.0059181420 |
| O46 | -5.0944777035 | -5.2033991253 | 1.1910908993 |
| O47 | -5.5459857541 | -5.5817825134 | -1.0028415646 |
| C48 | -9.5329403018 | -3.0336041019 | 4.8100596871 |
| C49 | -8.2202504619 | -2.6165929030 | 4.2477745880 |
| N50 | -7.6921565292 | -1.3393955483 | 4.3824777216 |
| C51 | -6.4228059047 | -1.3510179100 | 3.8980369478 |
| N52 | -6.1184860800 | -2.5533724177 | 3.4514965312 |
| C53 | -7.2252811617 | -3.3478581895 | 3.6653330395 |
| H54 | -0.8537394273 | -1.0716620784 | -1.6738799417 |
| H55 | -0.1502097057 | -2.1468828102 | -0.4667208978 |
| H56 | -3.0543741187 | 0.4006733898 | -0.1614316601 |
| H57 | -4.9102339217 | -0.5456069272 | 1.3096404891 |
| H58 | -1.9378863682 | -3.5064162418 | 0.8952986774 |
| H59 | -0.6804123618 | -2.8864427693 | 6.5409845404 |

| | | | |
|---|---|---|---|
| H60 | -2.3179090429 | -2.7992695972 | 7.1856027926 |
| H61 | -5.7418411796 | -2.7989458886 | -0.2246972830 |
| H62 | -6.7314507221 | -3.3996567335 | 1.0868168360 |
| H63 | -9.7683262158 | -4.0294730516 | 4.4218813617 |
| H64 | -10.3339405109 | -2.3667344058 | 4.4596382015 |
| H65 | -8.1523365530 | -0.5482756086 | 4.8067300605 |
| H66 | -5.7641818155 | -0.4953688850 | 3.9131579712 |
| H67 | -7.1940094894 | -4.4032445552 | 3.4439342390 |
| C68 | -1.7404418334 | -5.7897668172 | 8.5009082113 |
| C69 | -1.9462191121 | -7.0119025637 | 7.6393177749 |
| C70 | -3.1179994317 | -7.6924975921 | 7.4175265778 |
| N71 | -2.9427917291 | -8.6980592726 | 6.4808107906 |
| C72 | -1.6350730093 | -8.6529379455 | 6.0374879244 |
| C73 | -0.9807608813 | -9.4437351642 | 5.0887545557 |
| C74 | 0.3698490358 | -9.1987396698 | 4.8626617735 |
| C75 | 1.0431343486 | -8.1623844701 | 5.5441775755 |
| C76 | 0.3835213533 | -7.3651151060 | 6.4745564224 |
| C77 | -0.9721690870 | -7.6127809668 | 6.7532666943 |
| H78 | -2.3591083488 | -4.9762502012 | 8.0994644002 |
| H79 | -0.7039460118 | -5.4557330827 | 8.3708218344 |
| H80 | -4.1011818298 | -7.5413428497 | 7.8353253161 |
| H81 | -3.7199179917 | -9.0343642859 | 5.9161338317 |
| H82 | -1.5057023725 | -10.2160491060 | 4.5341916065 |
| H83 | 0.9009235768 | -9.8215865176 | 4.1496181518 |
| H84 | 2.0971702693 | -7.9883236923 | 5.3382687533 |
| H85 | 0.9112837201 | -6.5579874829 | 6.9771592755 |
| C86 | 1.9753401275 | -8.7415336425 | 9.5894794700 |
| C87 | 3.3288752698 | -9.1358634025 | 10.2041918820 |
| C88 | 1.3547536787 | -9.8899403484 | 8.7594412870 |
| C89 | 1.1785286094 | -11.2271363809 | 9.4932339233 |
| H90 | 2.1668963639 | -7.9219944077 | 8.8845356539 |
| H91 | 3.2137509829 | -9.8922482630 | 10.9900833140 |
| H92 | 4.0080731730 | -9.5458080962 | 9.4468217111 |
| H93 | 3.8213999836 | -8.2675229274 | 10.6585420549 |
| H94 | 0.3843048766 | -9.5563963495 | 8.3769536106 |
| H95 | 1.9777247421 | -10.0557328053 | 7.8716654737 |
| H96 | 0.5584040830 | -11.1319180528 | 10.3925803404 |
| H97 | 0.6861844400 | -11.9552510717 | 8.8396726519 |
| H98 | 2.1379014677 | -11.6598135720 | 9.8000901437 |
| C99 | 2.1899963682 | -4.8397883071 | 0.6886342222 |
| C100 | 0.7236708151 | -5.2298420480 | 0.3947171664 |
| O101 | -0.0617145712 | -4.3526566908 | -0.0132927459 |
| O102 | 0.4302846840 | -6.4648219287 | 0.5816564498 |
| H103 | 2.8248314886 | -5.6970183756 | 0.4386459759 |
| H104 | 2.2958672946 | -4.6900652943 | 1.7719404742 |
| H105 | -0.9646613132 | -7.1420089943 | 0.1196486443 |
| C106 | -5.9910091095 | -5.3787481052 | 8.6262283340 |
| C107 | -5.7803418445 | -5.7475514451 | 7.1534389442 |
| O108 | -6.2208770294 | -6.8589268481 | 6.7692575348 |
| O109 | -5.2071079836 | -4.8592200330 | 6.4404994010 |
| H110 | -5.4124242362 | -4.4819856927 | 8.8658321953 |
| H111 | -5.6365145177 | -6.2030669909 | 9.2570825422 |
| O112 | -5.1659039131 | -8.9285929361 | -1.5399318713 |
| H113 | -5.2218345309 | -8.0238348737 | -1.2027892344 |
| H114 | -5.0631141364 | -9.4552075986 | -0.7243658474 |
| O115 | 1.6871853163 | -8.2605224525 | 2.0201439761 |

| | | | |
|---|---|---|---|
| H116 | 1.3622942056 | -8.0074614943 | 2.8956178056 |
| H117 | 1.2843027556 | -7.5678249596 | 1.4189723478 |
| O118 | -2.2569813176 | -9.9330758258 | 1.3726421977 |
| H119 | -1.3680289330 | -10.2947848662 | 1.5748903130 |
| H120 | -2.0811776318 | -9.1352505487 | 0.8210289725 |
| O121 | 0.4179912373 | -10.6287467425 | 1.7613920825 |
| H122 | 0.8720002367 | -9.7424667126 | 1.8046102669 |
| H123 | 0.7193678508 | -11.0155537928 | 0.9307275262 |
| O124 | -1.7828945924 | -7.6732877589 | -0.1132214978 |
| C125 | -1.8398649164 | -7.8913323060 | -1.5147505524 |
| C126 | -4.1574016704 | -6.3612064566 | 1.2267376908 |
| C127 | -3.3759927255 | -6.2256525131 | 2.5300526382 |
| O128 | -2.7548320747 | -4.9525142991 | 2.5584555029 |
| C129 | -4.2925006273 | -6.4298617598 | 3.7624207224 |
| O130 | -4.9902725890 | -5.2151658951 | 4.0259432943 |
| C131 | -5.2824381444 | -7.6090167321 | 3.5497538038 |
| O132 | -5.4015690944 | -8.4312357822 | 4.6984540593 |
| C133 | -4.7979816989 | -8.4189219159 | 2.3459951925 |
| O134 | -4.8648636002 | -7.5551984612 | 1.1775200187 |
| C135 | -5.5651620284 | -9.6858313336 | 2.0008951461 |
| O136 | -4.9336782461 | -10.4168916064 | 0.9594461605 |
| H137 | -1.7386536243 | -6.9475943549 | -2.0699480531 |
| H138 | -1.0458848528 | -8.5742743543 | -1.8605578203 |
| H139 | -2.8145187033 | -8.3345232620 | -1.7438823972 |
| H140 | -3.5122145461 | -6.2767703471 | 0.3535149429 |
| H141 | -2.6044607055 | -7.0012548516 | 2.5109960177 |
| H142 | -2.2136106111 | -4.8259363566 | 3.4257262503 |
| H143 | -3.6525729793 | -6.6694185087 | 4.6202114245 |
| H144 | -5.1075917520 | -5.1017130114 | 5.0861134666 |
| H145 | -6.2571907211 | -7.1680085429 | 3.2877249897 |
| H146 | -5.7541145588 | -7.8560968523 | 5.4264329407 |
| H147 | -3.7595566067 | -8.7316610382 | 2.5019746738 |
| H148 | -6.5775480661 | -9.4366477640 | 1.6591693459 |
| H149 | -5.6530563025 | -10.2809553186 | 2.9217106197 |
| H150 | -3.9588584313 | -10.4124402298 | 1.1195461738 |

Transition state TS2 (Model 4)

$E_{absolute}$ = -3419.79485197980 hartrees

| | | | |
|---|---|---|---|
| C1 | -0.0002514101 | -0.0020821259 | -0.0067253212 |
| H2 | 0.0007162989 | -0.0020755382 | 1.0643602286 |
| H3 | 1.0123629569 | -0.0043238575 | -0.3533603922 |
| H4 | -0.4901687647 | 0.8981162487 | -0.3176294811 |
| C5 | -1.6897110153 | -0.9721469062 | 6.1533738939 |
| H6 | -2.6797917914 | -0.5721553650 | 6.2099104698 |
| H7 | -1.0888192011 | -0.5451871441 | 6.9279283846 |
| H8 | -1.2622516922 | -0.7369675422 | 5.1997645587 |
| C9 | -2.0225719297 | -5.8800266677 | 10.0184794768 |
| H10 | -3.0832183789 | -6.0462309773 | 10.2168560711 |
| H11 | -1.7183772113 | -4.9243641786 | 10.4452673519 |
| H12 | -1.4378329272 | -6.6812436807 | 10.4687051157 |
| C13 | 1.0017366904 | -8.2051174731 | 10.6737679335 |
| H14 | 0.4685338448 | -9.0386053862 | 11.0793877263 |
| H15 | 1.5557881454 | -7.7314985471 | 11.4582193406 |
| H16 | 0.2936136916 | -7.4856078154 | 10.2640684939 |
| C17 | -7.6765026322 | -3.6042905069 | -0.8741588000 |
| H18 | -8.2308413903 | -4.4858935635 | -0.6360313928 |

| | | | |
|---|---|---|---|
| H19 | -8.2916714400 | -2.7409452970 | -0.7251643935 |
| H20 | -7.3620365679 | -3.6435857123 | -1.8967618953 |
| C21 | -9.5244248588 | -3.0389681275 | 6.3475710007 |
| H22 | -10.5443445051 | -3.2466646468 | 6.6711139586 |
| H23 | -9.2227684913 | -2.0533863882 | 6.7036294899 |
| H24 | -8.8520662365 | -3.7929756038 | 6.7566181753 |
| C25 | 2.5957262683 | -3.5844318164 | -0.1118464213 |
| H26 | 2.1670928296 | -2.7079378274 | 0.3275024325 |
| H27 | 3.6571264851 | -3.4671958830 | -0.1741318942 |
| H28 | 2.1930913901 | -3.7223838917 | -1.0932780120 |
| C29 | -7.4904409517 | -5.1437465314 | 8.8626440121 |
| H30 | -8.0145225855 | -6.0341610808 | 8.5890557362 |
| H31 | -7.7087492070 | -4.9158262506 | 9.8852298121 |
| H32 | -7.8327797270 | -4.3154502151 | 8.2432129512 |
| Zn33 | -4.2200136469 | -4.0108936070 | 2.8662118964 |
| C34 | -0.7215890178 | -1.2307884824 | -0.6002984541 |
| C35 | -1.9613818732 | -1.5763645841 | 0.1468591214 |
| N36 | -2.9656054414 | -0.6725694619 | 0.4697539808 |
| C37 | -3.8690562132 | -1.3096361677 | 1.2611080234 |
| N38 | -3.5008643418 | -2.5600380876 | 1.4537884310 |
| C39 | -2.3214314025 | -2.7432982230 | 0.7651089183 |
| C40 | -1.7067402363 | -2.5130774009 | 6.2995276137 |
| C41 | -2.2020435458 | -3.3144705131 | 5.0801636990 |
| O42 | -3.1676037394 | -2.8466920733 | 4.3964415950 |
| O43 | -1.6125578303 | -4.4112201575 | 4.8567067651 |
| C44 | -6.4484695068 | -3.5506367229 | 0.0565016601 |
| C45 | -5.7704393733 | -4.9243815471 | 0.2504332778 |
| O46 | -5.1629751831 | -5.1335179908 | 1.3702522277 |
| O47 | -5.8617126004 | -5.7421227408 | -0.6796171467 |
| C48 | -9.4583747658 | -3.0121862156 | 4.8099772472 |
| C49 | -8.0618911933 | -2.7616197031 | 4.3508306879 |
| N50 | -7.2748496119 | -1.7210135310 | 4.8332428904 |
| C51 | -6.0180997579 | -1.8733252183 | 4.3393430003 |
| N52 | -5.9675226572 | -2.9263509380 | 3.5515134342 |
| C53 | -7.2231075351 | -3.4902140331 | 3.5554013137 |
| H54 | -0.9341491530 | -1.0417905743 | -1.6630828138 |
| H55 | -0.0925328376 | -2.1258080005 | -0.5668845385 |
| H56 | -3.0142166199 | 0.2906654216 | 0.1764858575 |
| H57 | -4.7562698119 | -0.8430865141 | 1.6621414341 |
| H58 | -1.7966802549 | -3.6883524366 | 0.7336810685 |
| H59 | -0.7033791761 | -2.8838250968 | 6.5293480907 |
| H60 | -2.3471057280 | -2.8054976406 | 7.1425876297 |
| H61 | -5.6869368047 | -2.8720263507 | -0.3484550102 |
| H62 | -6.7203981016 | -3.1456934670 | 1.0337344915 |
| H63 | -9.7855566105 | -3.9673264085 | 4.3873182314 |
| H64 | -10.1505491336 | -2.2478691403 | 4.4264863543 |
| H65 | -7.5591482518 | -1.0203856096 | 5.5000174287 |
| H66 | -5.1690057139 | -1.2627124245 | 4.6069361375 |
| H67 | -7.4187417195 | -4.4013480701 | 3.0117080961 |
| C68 | -1.7439049404 | -5.8364470651 | 8.4984104784 |
| C69 | -1.9534412868 | -7.1303668713 | 7.7546598323 |
| C70 | -3.0844065489 | -7.9053385416 | 7.7274030074 |
| N71 | -2.9261125214 | -8.9953936103 | 6.8814989414 |
| C72 | -1.6787645262 | -8.9004713238 | 6.2926832641 |
| C73 | -1.0664511035 | -9.7232779445 | 5.3418261501 |
| C74 | 0.2263220709 | -9.4006411891 | 4.9400780997 |

| | | | |
|---|---|---|---|
| C75 | 0.8901708472 | -8.2742345696 | 5.4700921981 |
| C76 | 0.2717631668 | -7.4457494234 | 6.4015128263 |
| C77 | -1.0283516411 | -7.7549335878 | 6.8356482267 |
| H78 | -2.3708738905 | -5.0625825856 | 8.0378004825 |
| H79 | -0.7083342875 | -5.5099005087 | 8.3434604900 |
| H80 | -4.0341846583 | -7.7640000283 | 8.2208442677 |
| H81 | -3.7330374158 | -9.3805250502 | 6.3956695858 |
| H82 | -1.5783793903 | -10.5894054720 | 4.9308530248 |
| H83 | 0.7163472287 | -10.0169322800 | 4.1938946499 |
| H84 | 1.9044703551 | -8.0554172601 | 5.1434889585 |
| H85 | 0.7836549333 | -6.5643611911 | 6.7788060914 |
| C86 | 1.9831700769 | -8.7120180124 | 9.5839506121 |
| C87 | 3.3606976637 | -9.0353818782 | 10.1870699161 |
| C88 | 1.4040256961 | -9.9014730990 | 8.7832731084 |
| C89 | 1.2865312008 | -11.2274263321 | 9.5481506435 |
| H90 | 2.1311022938 | -7.8995674712 | 8.8603278474 |
| H91 | 3.2884864476 | -9.7778391273 | 10.9911116023 |
| H92 | 4.0468364775 | -9.4326281162 | 9.4291181406 |
| H93 | 3.8214296292 | -8.1369001029 | 10.6152589738 |
| H94 | 0.4180200211 | -9.6159491917 | 8.4018041267 |
| H95 | 2.0282546454 | -10.0615064880 | 7.8950580585 |
| H96 | 0.6700001914 | -11.1347376354 | 10.4501980655 |
| H97 | 0.8186815074 | -11.9899364428 | 8.9160375601 |
| H98 | 2.2645910164 | -11.6140857001 | 9.8572513461 |
| C99 | 2.1709312194 | -4.8363997966 | 0.6650677798 |
| C100 | 0.7572996314 | -5.2724329902 | 0.2337284869 |
| O101 | 0.0233387345 | -4.4626050329 | -0.3645552400 |
| O102 | 0.4377067260 | -6.4889150317 | 0.4954573971 |
| H103 | 2.8503338052 | -5.6766743588 | 0.4826225049 |
| H104 | 2.1744177040 | -4.6774575766 | 1.7515730916 |
| H105 | -0.9340846201 | -6.8158384162 | -0.2828832846 |
| C106 | -5.9792818858 | -5.3479126842 | 8.7227727532 |
| C107 | -5.6571227094 | -5.8191628532 | 7.3106863122 |
| O108 | -6.1196413015 | -6.9455687956 | 6.9854079300 |
| O109 | -4.9900237175 | -5.0151982134 | 6.5895000467 |
| H110 | -5.4362600327 | -4.4235168987 | 8.9403016195 |
| H111 | -5.6541346341 | -6.1165265008 | 9.4353946536 |
| O112 | -4.9163440745 | -8.2871388290 | -1.3384031554 |
| H113 | -5.2466198978 | -7.3978749236 | -1.0911187605 |
| H114 | -5.0445584883 | -8.8069958373 | -0.5365835535 |
| O115 | 1.6427059387 | -7.9791376689 | 2.2431812419 |
| H116 | 1.2546067000 | -7.6726552047 | 3.0747586995 |
| H117 | 1.2522454440 | -7.3611071757 | 1.5485407024 |
| O118 | -2.3612774928 | -8.9225750687 | 1.0994566565 |
| H119 | -1.4840260074 | -9.2331187165 | 1.4478361386 |
| H120 | -2.1556151412 | -8.4479226719 | 0.2417825143 |
| O121 | 0.0311271232 | -10.0196322740 | 1.7147941070 |
| H122 | 0.6918678740 | -9.2928234799 | 1.9109100527 |
| H123 | 0.2766783810 | -10.3481388535 | 0.8408875958 |
| O124 | -1.7848928887 | -7.1795607883 | -0.6746113900 |
| C125 | -1.8754848197 | -6.9384533808 | -2.0663482287 |
| C126 | -3.6673608016 | -7.3052345359 | 1.6936062915 |
| C127 | -3.0974928343 | -6.7964030839 | 2.9995545718 |
| O128 | -2.5860534299 | -5.4989242103 | 2.7983078088 |
| C129 | -4.1764211149 | -6.8218371660 | 4.1476574523 |
| O130 | -4.8395698564 | -5.5811300030 | 4.1956931998 |

| | | | |
|---|---|---|---|
| C131 | -5.2081288647 | -7.9874208927 | 4.0242710872 |
| O132 | -5.3781597581 | -8.7180023121 | 5.2181474656 |
| C133 | -4.7564563241 | -8.9453452821 | 2.9352086760 |
| O134 | -4.6448902137 | -8.1427502026 | 1.7113122068 |
| C135 | -5.6321959246 | -10.1168086799 | 2.5523292000 |
| O136 | -5.0055005973 | -10.9106298655 | 1.5531391538 |
| H137 | -1.7908069601 | -5.8655786427 | -2.2796475703 |
| H138 | -1.0852412001 | -7.4636525711 | -2.6284638364 |
| H139 | -2.8505354052 | -7.3096586349 | -2.3898581814 |
| H140 | -3.4668871327 | -6.8160857128 | 0.7462962685 |
| H141 | -2.2942020064 | -7.4852841189 | 3.2903185598 |
| H142 | -2.1138664655 | -5.1405508208 | 3.6616228674 |
| H143 | -3.6212526334 | -6.9695202688 | 5.0806540384 |
| H144 | -4.9249364851 | -5.3231818288 | 5.2107135382 |
| H145 | -6.1531596232 | -7.5241919479 | 3.7042276038 |
| H146 | -5.6953671691 | -8.0517377968 | 5.8982386734 |
| H147 | -3.7525383834 | -9.3299240231 | 3.1513781528 |
| H148 | -6.6056934902 | -9.7460337099 | 2.1970527394 |
| H149 | -5.8065765291 | -10.7206721660 | 3.4505444617 |
| H150 | -4.2266209129 | -10.4351421006 | 1.2191113025 |

Hydrolyzed producte PRO (Model 4)

$E_{absolute}$ = -3419.84544562564 hartrees

| | | | |
|---|---|---|---|
| C1 | 0.0019826568 | 0.0037881084 | -0.0014307813 |
| H2 | 0.0021713608 | 0.0033109788 | 1.0696550896 |
| H3 | 1.0148513619 | 0.0038372192 | -0.3473292359 |
| H4 | -0.4896043841 | 0.9030931334 | -0.3122847447 |
| C5 | -1.6899082774 | -0.9726276090 | 6.1569975669 |
| H6 | -2.6808706193 | -0.5747487285 | 6.2129939416 |
| H7 | -1.0904809623 | -0.5447540134 | 6.9321822866 |
| H8 | -1.2609695791 | -0.7388447758 | 5.2037091791 |
| C9 | -2.0152367256 | -5.8829501176 | 10.0196411297 |
| H10 | -3.0756745590 | -6.0514787665 | 10.2171701959 |
| H11 | -1.7133670089 | -4.9268424649 | 10.4470822845 |
| H12 | -1.4291382006 | -6.6831374276 | 10.4699302776 |
| C13 | 1.0134874674 | -8.2019606048 | 10.6760807463 |
| H14 | 0.4817473643 | -9.0367539687 | 11.0809354048 |
| H15 | 1.5659690121 | -7.7275311669 | 11.4611493822 |
| H16 | 0.3041479625 | -7.4837588035 | 10.2661910060 |
| C17 | -7.6660270988 | -3.6141928677 | -0.8760844624 |
| H18 | -8.2186798163 | -4.4970699381 | -0.6387595319 |
| H19 | -8.2831222336 | -2.7522134912 | -0.7271479185 |
| H20 | -7.3507350356 | -3.6523614991 | -1.8984758649 |
| C21 | -9.5203889081 | -3.0560416249 | 6.3445523022 |
| H22 | -10.5401036331 | -3.2660335589 | 6.6672584250 |
| H23 | -9.2210693746 | -2.0699880552 | 6.7012759363 |
| H24 | -8.8467403086 | -3.8088161146 | 6.7537479916 |
| C25 | 2.6055797227 | -3.5730357428 | -0.1062805638 |
| H26 | 2.1747806919 | -2.6976462363 | 0.3331522066 |
| H27 | 3.6667755429 | -3.4535353916 | -0.1677399962 |
| H28 | 2.2039504033 | -3.7113907111 | -1.0880673941 |
| C29 | -7.4838030085 | -5.1576703016 | 8.8601536057 |
| H30 | -8.0058077960 | -6.0490629990 | 8.5857811685 |
| H31 | -7.7033346591 | -4.9306745963 | 9.8826831511 |
| H32 | -7.8274361924 | -4.3298162220 | 8.2408480889 |
| Zn33 | -4.1470727252 | -4.1729129194 | 2.8788661333 |

| | | | |
|---|---|---|---|
| C34 | -0.7247468395 | -1.2274086116 | -0.5843568856 |
| C35 | -1.9332156099 | -1.5993810625 | 0.2005124999 |
| N36 | -2.9953964087 | -0.7471337400 | 0.4698909327 |
| C37 | -3.8497040576 | -1.3941461625 | 1.3109023694 |
| N38 | -3.3928189664 | -2.5955605695 | 1.5885013993 |
| C39 | -2.2078945988 | -2.7393550799 | 0.9059849777 |
| C40 | -1.6928931463 | -2.4978327186 | 6.3814942816 |
| C41 | -2.1550200106 | -3.3758546083 | 5.2110456524 |
| O42 | -3.1040283579 | -2.9648714680 | 4.4723658495 |
| O43 | -1.5405189404 | -4.4730932366 | 5.0859981607 |
| C44 | -6.4285532339 | -3.5680297709 | 0.0316858087 |
| C45 | -5.7229178207 | -4.9356957722 | 0.1737943815 |
| O46 | -4.9348320743 | -5.0771315355 | 1.1742365982 |
| O47 | -5.9586473510 | -5.8001999905 | -0.6927628458 |
| C48 | -9.4197549739 | -3.0406702175 | 4.8082356640 |
| C49 | -8.0003286434 | -2.8540892617 | 4.3811256425 |
| N50 | -7.1586504811 | -1.8959365931 | 4.9357943192 |
| C51 | -5.9091653736 | -2.0863278758 | 4.4391832627 |
| N52 | -5.9107862078 | -3.0809308309 | 3.5760578423 |
| C53 | -7.1976137285 | -3.5719176882 | 3.5389859753 |
| H54 | -0.9715663948 | -1.0370179513 | -1.6390310498 |
| H55 | -0.0762813006 | -2.1076969276 | -0.5784097686 |
| H56 | -3.1216226962 | 0.1827368412 | 0.1018335564 |
| H57 | -4.7715170312 | -0.9697989992 | 1.6793013924 |
| H58 | -1.6299175934 | -3.6501449996 | 0.9479348329 |
| H59 | -0.6890684655 | -2.8424022202 | 6.6474807642 |
| H60 | -2.3366268580 | -2.7428849424 | 7.2374285585 |
| H61 | -5.6860469524 | -2.8580441333 | -0.3578026344 |
| H62 | -6.6857985886 | -3.2049286684 | 1.0300311784 |
| H63 | -9.7763874106 | -3.9845841015 | 4.3839054332 |
| H64 | -10.0697511648 | -2.2511754740 | 4.4021373405 |
| H65 | -7.4004332388 | -1.2444522198 | 5.6657063141 |
| H66 | -5.0235042270 | -1.5559915100 | 4.7555215504 |
| H67 | -7.4468702134 | -4.4316612142 | 2.9365326274 |
| C68 | -1.7083769115 | -5.8118401360 | 8.5021083754 |
| C69 | -1.7848583619 | -7.1066970839 | 7.7347269158 |
| C70 | -2.8705767912 | -7.9365524584 | 7.6064440787 |
| N71 | -2.5840506756 | -9.0119473659 | 6.7805175936 |
| C72 | -1.2932286342 | -8.8621015425 | 6.3115211518 |
| C73 | -0.5491095833 | -9.6688456792 | 5.4462054067 |
| C74 | 0.7602511031 | -9.2823948950 | 5.1672723901 |
| C75 | 1.3053956296 | -8.1014652218 | 5.7146045762 |
| C76 | 0.5585960937 | -7.2981179383 | 6.5704833492 |
| C77 | -0.7535408753 | -7.6790500402 | 6.9005550082 |
| H78 | -2.3769573587 | -5.0788101082 | 8.0334082301 |
| H79 | -0.6987560465 | -5.4018061055 | 8.3780430728 |
| H80 | -3.8718678539 | -7.8333679959 | 8.0002285464 |
| H81 | -3.3375094964 | -9.4484482481 | 6.2487364399 |
| H82 | -0.9742183128 | -10.5565592286 | 4.9885218305 |
| H83 | 1.3708124107 | -9.9049354806 | 4.5189292135 |
| H84 | 2.3270493004 | -7.8214937705 | 5.4668100421 |
| H85 | 0.9816229378 | -6.3806472348 | 6.9699741532 |
| C86 | 2.0128463570 | -8.7198852252 | 9.6075204038 |
| C87 | 3.3918593170 | -8.9925160172 | 10.2326139182 |
| C88 | 1.4701918639 | -9.9494724200 | 8.8442971577 |
| C89 | 1.3748144923 | -11.2510987045 | 9.6529104082 |

| | | | |
|---|---|---|---|
| H90 | 2.1483294674 | -7.9311641058 | 8.8575069334 |
| H91 | 3.3299074919 | -9.7118042965 | 11.0583327963 |
| H92 | 4.0941145552 | -9.3968486774 | 9.4931345716 |
| H93 | 3.8281833303 | -8.0710859742 | 10.6370549572 |
| H94 | 0.4822116371 | -9.7002832634 | 8.4440660214 |
| H95 | 2.1068368029 | -10.1252863440 | 7.9682291079 |
| H96 | 0.7488410343 | -11.1419775816 | 10.5465805928 |
| H97 | 0.9286645658 | -12.0447283014 | 9.0437649584 |
| H98 | 2.3582195766 | -11.6060323676 | 9.9826109953 |
| C99 | 2.2040832182 | -4.8542785879 | 0.6620476665 |
| C100 | 0.9046930987 | -5.4035801414 | 0.0807792829 |
| O101 | 0.0434954299 | -4.6815835528 | -0.4021723995 |
| O102 | 0.8268132164 | -6.7300941447 | 0.0889066038 |
| H103 | 2.9702212226 | -5.6314267830 | 0.5845709480 |
| H104 | 2.0532614153 | -4.6661939637 | 1.7325033942 |
| H105 | -0.0194080847 | -7.0600954502 | -0.4271407687 |
| C106 | -5.9776141258 | -5.3966549169 | 8.6836253653 |
| C107 | -5.7197532600 | -5.9363230199 | 7.2741301475 |
| O108 | -5.9123666618 | -7.1690170866 | 7.0979037111 |
| O109 | -5.3895755579 | -5.0672351187 | 6.4093222546 |
| H110 | -5.4158386819 | -4.4698323218 | 8.8380471907 |
| H111 | -5.6435132541 | -6.1395844144 | 9.4168994768 |
| O112 | -5.1316889675 | -8.5070479504 | -0.8866987224 |
| H113 | -5.5117167729 | -7.6013540597 | -0.8098794724 |
| H114 | -4.5820015304 | -8.5448161240 | -0.0884069375 |
| O115 | 1.1889697985 | -8.7580999452 | 2.1135541086 |
| H116 | 0.9874048768 | -8.3476817761 | 2.9684909033 |
| H117 | 1.0657963565 | -8.0528520464 | 1.4507564481 |
| O118 | -2.1023630515 | -8.1856908907 | 1.3062165384 |
| H119 | -1.8955190951 | -9.0138296528 | 1.8023173195 |
| H120 | -1.6325986884 | -8.0305072348 | -0.2520892842 |
| O121 | -0.9657635622 | -10.4409537937 | 2.2363883281 |
| H122 | -0.1185085724 | -9.9393090037 | 2.1391332117 |
| H123 | -1.0315847572 | -10.9975878354 | 1.4501913606 |
| O124 | -1.1598429233 | -7.7620992031 | -1.0834834796 |
| C125 | -2.0670138584 | -6.9698800993 | -1.8709273205 |
| C126 | -3.2103256153 | -7.4650733881 | 1.8222339199 |
| C127 | -2.9106649422 | -6.9015253668 | 3.2218012933 |
| O128 | -2.4437119623 | -5.5682299471 | 3.0191138672 |
| C129 | -4.1050018878 | -6.9343601120 | 4.2086047389 |
| O130 | -4.8853098399 | -5.7531137840 | 4.0711598591 |
| C131 | -5.0302430523 | -8.1829267838 | 4.0712628360 |
| O132 | -5.0778803754 | -8.9737533976 | 5.2529811968 |
| C133 | -4.6823692404 | -9.0793706339 | 2.8781843741 |
| O134 | -4.3871704584 | -8.2322452830 | 1.7646519212 |
| C135 | -5.8808644512 | -9.9322408045 | 2.4379049036 |
| O136 | -5.6700480536 | -10.6309338574 | 1.2292018096 |
| H137 | -2.0759481196 | -5.9292560204 | -1.5287624177 |
| H138 | -1.7183523055 | -6.9990488057 | -2.9081594451 |
| H139 | -3.0767331994 | -7.3897593247 | -1.8158665628 |
| H140 | -3.3858824134 | -6.6330781313 | 1.1430357087 |
| H141 | -2.1067315072 | -7.5031893779 | 3.6664123159 |
| H142 | -2.0234305189 | -5.1969983543 | 3.8908281154 |
| H143 | -3.6809482785 | -6.9441639156 | 5.2180132718 |
| H144 | -5.1567160983 | -5.4707795230 | 5.0348238006 |
| H145 | -6.0234269266 | -7.7573262743 | 3.8686645616 |

| | | | |
|---|---|---|---|
| H146 | -5.4455982186 | -8.3661001930 | 5.9462018023 |
| H147 | -3.8252329279 | -9.7318509605 | 3.1079225814 |
| H148 | -6.7568850941 | -9.2666251795 | 2.3692035398 |
| H149 | -6.0842884795 | -10.6681547881 | 3.2231572384 |
| H150 | -5.6946827799 | -9.9816673076 | 0.5013407023 |

# CHAPTER 5

# CONCLUSIONS

Glycosidases are enzymes that play crucial roles in the biosynthesis of glycoproteins. Inhibitors of these enzymes have garnered much attention as lead compounds for drug discovery for diseases such as viral and bacterial infections, diabetes, Gauchers disease, and cancer. In many cases, the mode of inhibition by these compounds is not well understood, complicating efforts to design and synthesize more potent and/or selective inhibitors. As a result, only a very small number of glycosidase inhibitors have been successfully developed as therapeutics. In the case of Golgi a-mannosidase II, which is a glycosyl hydrolase family 38 enzyme that acts late in the N-glycan biosynthesis pathway and cleaves selectively α1,3- and α1,6 glycosidic linkage between two mannose units, several aspects of the inhibition and the catalytic mechanism are not apparent. In this respect, potent inhibitors of GMII have potential as biochemical tools with which one can delineate further the catalytic and the inhibition mechanism, with a possibility of discovering or designing more potent and selective GMII inhibitors.

Mannostatins A and B, isolated from the soil microorganism *Streptoverticillus*, are some of the most potent inhibitors of class II α-mannosidases reported thus far, and are interesting lead compounds. Despite their structural simplicity, the mannostatins are some of the most potent inhibitors of GMII. A straightforward correlation of their structure with their inhibitory properties remains nonetheless elusive, at least using classical models. In particular it is not clear whether mannostatin A mimics the oxacarbenium ion-like transition state of the hydrolysis reaction catalyzed by GMII. In addition to this, the possible role of thiomethyl moiety in mannostatin A is intriguing, all the more so in that this feature is unique to mannostain A. If the thiomethyl function plays a role in the binding of mannostatin A by GMII, the extent to which this structural

feature is important remains unclear. Finally, from the viewpoint of the catalytic mechanism of GMII, it is not known what conformational itinerary is followed by the substrate *en route* to the hydrolyzed product. An understanding of this itinerary would be significant in designing, synthesizing and evaluating novel inhibitors of GMII. Moreover, with the knowledge of the transition state structure it would be possible to determine to what extent the potent inhibitors of GMII mimic the transition state. In this dissertation, we have employed a computational approach in conjunction with experiments to address these unanswered questions.

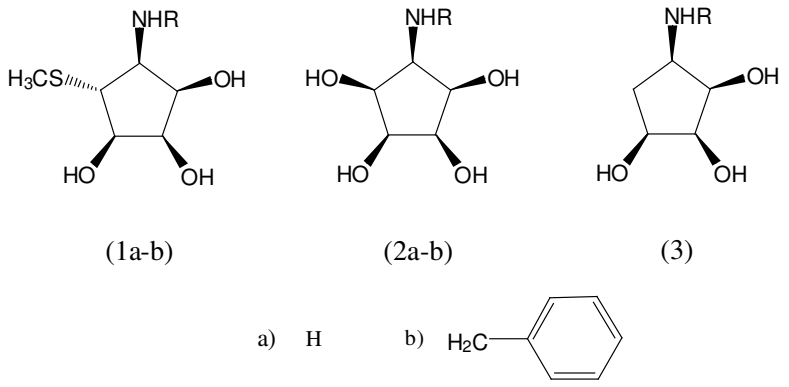



**Figure 5.1.** GMII inhibitors studied inthis disseration.

In order to understand the mechanism of inhibition of a protein target by a specific ligand, the knowledge of the orientation of the ligand in the binding site essential. The first project dealt with prediction of the binding mode of mannostatin in dGMII binding site to rationalize of biological data for a set of mannostatin analogues. A range of Mannostatin analogs and structurally simpler aminocyclopentitetrols has been synthesized, which have varying benzylic moieties at their amino function. It was

observed that mannostatin A is a markedly better inhibitor than the corresponding aminocyclopentitetrol. Furthermore, the substitutions of the aminocyclopentitetrols led to significant improvements of inhibitory activity for both hGMII and hLM whereas similar substituents of Mannostatin A had only a marginal effect. Computational studies have been performed to provide a rational for these observations.

First, the conformational properties of Mannostatin and its analogue aminocyclopentitetrol were studied by geometry optimizations of the ten possible envelope conformations, using *ab initio* molecular orbital calculations. Subsequently, the low energy conformers of each compound were docked in the binding site of dGMII and the results compared with interactions of swainsonine observed in a crystal structure with the same enzyme. It has been shown that Mannostatin and aminocyclopentitretrol could bind to GMII in a similar mode as swainsonine. Probably, the latter compounds mimic the mannosyl oxycarbenium ion, which is a putative intermediate in the hydrolysis of mannosidases. Thus, it appears that Mannostatin and aminocyclopentitretrol can inhibit GMII in a similar fashion. However, due to the flexibility of the five membered rings of **1a** and **2a**, additional low energy binding modes could be adopted. It is conceivable that the ring structure of **1a** and **2a** is flexible in the binding site of GMII allowing transitions between different binding modes. The thiomethyl moiety of Mannostatin could make favorable hydrophobic interactions with Arg228 and Tyr727. These interactions are not present in complexes with aminocyclopentitretrol (**2a**) providing a rationale for the lower inhibitory potential of this compound. It was also observed that the five membered ring of **2a** is significantly more flexible than that of **1a** allowing additional binding modes. Attachment of benzyl moieties to Mannostatin let to stacking interactions with the

aromatic moiety of Tyr269. These π-π interactions required, however, a tilting of the ring structure of **1b** resulting in a disruption of the hydrogen-bonding network observed. Due to the greater flexibility of aminocyclopentitretrol (**2a**), attachment of the benzyl moiety led to conformers that could make the stacking interactions without disrupting hydrogen bonds. This observation may provide a rational for the improved activities of benzyl-substituted aminocyclopentitretrol.

The predicted binding modes of mannostatin A and its benzylated analogue were validated using X-ray crystallography. An excellent agreement was observed between the predictions and the bound conformations of cyclopentane rings in either inhibitor. Furthermore, from the biological data it was concluded that the thiomethyl group of mannostatin A is required for high affinity binding, as the compound which lacks this moiety was found to be 7000 times less potent than mannostatin A. Analysis of the X-ray structure of dGMII:mannostatin A complex suggested that this functionality is able to make a number of additional interactions increasing the affinity for dGMII. The thiomethyl moiety of **1a** is structurally similar to the side chain of methionine residues. Interestingly, the environment surrounding the thiomethyl group of **1a** is remarkably similar to the arrangements around methionine residues in the protein structure. Another important finding is that the methyl carbon of the thiomethyl moiety of **1a** in the complex with GMII adopts two different conformations. In both conformations, the sulfur interacts with the backbone oxygen of Arg876 through the π-orbital on the carbonyl oxygen and the antibonding σ* orbital of the S-C bond. In cf1 conformation, the polarizable methyl group is in a hydrophilic environment interacting with three water molecules and the π-system of the Arg228 side chain. Whereas, the methyl group of the cf2 conformation is in

a hydrophobic environment where it forms C–H---π type interactions with the phenyl rings of Phe206 and Tyr727.

The bound conformation was further compared with the proposed transition state conformers. The studies reported here indicate that Mannostatin A mimics the covalently linked mannosyl intermediate, which has been shown to adopt a $^{1}S_{5}$ skew boat conformation. In particular, the zinc-coordinating 3- and 4-hydroxyls of **1a** display a good overlap with the corresponding *cis*-diol of the mannosyl residue. Furthermore, the amine nitrogen of **1** is positioned close to the C-1 carbon of enzyme-linked mannosyl intermediate and therefore can interact with the carboxylic acid of the catalytic nucleophile Asp204. The thiomethyl functionality showed a good overlay with the C-6 hydroxyl of the covalently linked intermediate.

The MD simulations and NMR studies of the uncomplexed inhibitor showed that the five-membered ring could adopt two different conformational families. In the bound state, however, it is restricted to a single conformation, which is different from the conformations observed for the free ligand. This finding is surprising because it is generally believed that potent inhibitors are complexed in their ground state conformations. Probably, the flexibility of five membered rings renders the energy barrier for the conformational change relatively low.

Compound **1b** was designed in such a way that its benzyl moiety can interact with aromatic residues of the binding site of the enzyme thereby increasing its affinity. However, it was found that this compound has a slightly reduced affinity indicating that the benzyl group cannot make favorable interactions with the enzyme. The unexpected finding could be rationalized by careful examination of the crystal structure of dGMII

with **1b**. Thus, the five-membered ring of compounds **1a** and **1b** were identically positioned in the binding site of dGMII. The distance between the centroid of the aromatic ring of compound **1b** and Tyr269 was such that no edge to face binding interactions could be made. The unexpected orientation of the aromatic ring of **1b** could be rationalized by an intramolecular S---H–Ar type hydrogen bond between the thiomethyl moiety and the phenyl ring.

In addition to the mechanism of inhibition of GMII by mannostatin A, several aspects of the catalytic mechanism of the hydrolysis reaction catalyzed by GMII are not clear. In particular, it is not known which conformation is adopted by the mannopyransoyl ring at the transition state. Furthermore, the exact role of $Zn^{+2}$ ion in the catalytic mechanism is intriguing. In order to address these questions, we have studied the hydrolysis reaction catalyzed by GMII using quantum mechanical methods. We have used four models, which differ in the size and the overall charge, to investigate the mechanism. In the first three models, which lack explicit water molecules, the reaction barriers for the glycosylation step were very high (46.5 – 61.5 kcal $mol^{-1}$). For this reason the deglycosylation step was not performed for these models. Although these models did not result in more realistic energy profiles, the results obtained were beneficial for the refinement of the model. From the QM study of Model 1 – Model 3, it was concluded that all the residues that coordinate with the metal ion present in the active site must be included in the quantum mechanical calculations. Model 3, which includes complete coordination sphere for the $Zn^{+2}$ ion, resulted in more stable complex. Unlike Model 1 and Model 2, this model did not display proton abstraction from 2-hydroxyl group by Asp472.

The best results were obtained for Model 4, where the computed activation barrier for the glycosylation step is 15.8 kcal $mol^{-1}$ at DFT/B3LYP/6-311++G** level. This model includes four water molecules in addition to all the residues in Model 3. In the optimized Michaelis complex, the mannopyranosyl ring in the substrate adopts a perfect $^{1}S_{5}$ skew boat conformation, where C-2, C-3, C-4 and O-5 atoms are in one plane. This conformation seems to facilitate the departure of the leaving group by forming a hydrogen bond between the anomeric oxygen and the proton from general acid catalyst Asp341.

The mannopyranosyl ring in the transition state TS1 adopts a conformation between $S_{2,5}$ and $^{1}S_{5}$. Calculation of Wiberg bond indices and natural bond order analysis supported the resemblance of the transition state with the mannopyranosyl oxacarbenium ion. An increased sp2 character was observed in the C-1–O-5 bond along with delocalization of positive charge between C-1 and O-5 atoms. At the transition state the C-1–O-1 bond is completely cleaved, suggesting a dissociative nature of the transition state for the glycosylation step in accord with the experimental observations.

The optimized structure of CINT revealed that the sugar ring adopts a skew boat conformation that is very close to an ideal $^{1}S_{5}$ conformation. The comparison of the optimized CINT structure with the trapped covalent intermediate resulted in staggering similarities. The catalytic nucleophilic oxygen moves away from the $Zn^{+2}$ ion. Similar movement of Asp204 was also observed in the X-ray structures of trapped covalent intermediates.

Unlike the glycosylation step, that transition state for the deglycosylation step (TS2) revealed that these atoms are not co-planar. However, increased $sp^{2}$ character was

observed for the C-1–O-5 bond along with the delocalization of the charge across this bond. In this structure the conformation of the sugar ring is slightly skewed boat conformation that lies between $B_{2,5}$ and $^{O}S_2$ skew boat conformation. This conformation of the mannopyranosyl ring changes in the hydrolyzed product to a conformation very close to an ideal $^{O}S_2$ skew boat conformation.

The calculation presented here suggests that the $Zn^{+2}$ ion in the active site is essential for substrate binding to the active site and the coordination 2- and 3- hydroxyl group with the $Zn^{+2}$ ion in the main driving force for the conformational changes in the substrate. Furthermore, the natural population analysis revealed that the charge on zinc remains unchanged over the course of the reaction. The role of zinc is, thus, structural and not electronic in natural.

Finally, we compared the bound conformation of mannostatin A and swainsonine to the mannopyranosyl ring in the optimized TS1. These superimposition studies revealed that mannostatin A and swainsonine in bound conformation mimic the transition state for the glycosylation step, which has been shown to adopt a conformation that lies between $B_{2,5}$ and $^{1}S_5$ skew boat conformation. In particular, the zinc-coordinating 3- and 4-hydroxyls of mannostatin A display a good overlap with the corresponding *cis*-diol of the mannosyl residue. Furthermore, the amine nitrogen of mannostatin A is positioned close to the C-1 carbon of TS1, whereas in the case of swainsonine the amine nitrogen is close to the ring oxygen suggesting that it may mimic the charge build up at the ring oxygen during the formation of the transition state. Collectively these studies indicate that mannostatin A and swainsonine are transition state mimics.

The studies presented in this dissertation are significant in the search of potent and /or selective inhibitors of GMII. With the knowledge of the mechanism of inhibition of GMII by mannostatin A, it is possible to design mannostatin analogues that are more potent or more selective inhibitors of GMII. Furthermore, it is also possible to design molecules, which mimic the transition state conformation. Such molecules would have higher binding affinities against GMII.

# REFERENCES

(1) Varki, A., *Glycobiology* **1993,** 3, 97-130.

(2) Witczak, Z. J., *Curr. Med. Chem.* **1995,** 1, 392.

(3) Sears, P.; Wong, C. H., *Proc. Natl. Acad. Sci. U. S. A.* **1996,** 93, 12086-12093.

(4) Williams, S. J.; Davies, G. J., *Trends Biotechnol.* **2001,** 19, 356-362.

(5) Dwek, R. A.; Butters, T. D.; Platt, F. M.; Zitzmann, N., *Nat. Rev. Drug Discov.* **2002,** 1, 65-75.

(6) Asano, N.; Nash, R. J.; Molyneux, R. J.; Fleet, G. W. J., *Tetrahedron Asymmetry* **2000,** 11, 1645-1680.

(7) Berecibar, A.; Grandjean, C.; Siriwardena, A., *Chem. Rev.* **1999,** 99, 779-844.

(8) Lillelund, V. H.; Jensen, H. H.; Liang, X. F.; Bols, M., *Chem. Rev.* **2002,** 102, 515-553.

(9) vonItzstein, M.; Thomson, R. J., *Curr. Med. Chem.* **1997,** 4, 185-210.

(10) Yuasa, H.; Saotome, C.; Kanie, O., *Trends Glycosci. Glycotechnol.* **2002,** 14, 231-251.

(11) Wolfenden, R.; Lu, X. D.; Young, G., *J. Am. Chem. Soc.* **1998,** 120, 6814-6815.

(12) Zechel, D. L.; Withers, S. G., *Acc. Chem. Res.* **2000,** 33, 11-18.

(13) Koshland, D. E., *Biol. Rev.* **1953,** 28, 416-436.

(14) Zechel, D. L.; Withers, S. G., *Curr. Opin. Chem. Biol.* **2001,** 5, 643-649.

(15) Sinnott, M. L., *Chem. Rev.* **1990, 90**, 1172-1202.

(16) Rye, C. S.; Withers, S. G., *Curr. Opin. Chem. Biol.* **2000,** 4, 573-580.

(17) Capon, B., *Chem. Rev.* **1969,** 69, 407-498.

(18) Withers, S. G., *Carbohydrate Polymers* **2001,** 44, 325-337.

(19) Vasella, A.; Davies, G. J.; Bohm, M., *Curr. Opin. Chem. Biol.* **2002,** 6, 619-629.

(20) Blake, C. C. F.; Johnson, L. N.; Mair, G. A.; North, A. C. T.; Phillips, D. C.; Sarma, V. R., *Proc. Royal Soc. Lon. B* **1967,** 167, 378-&.

(21) Davies, G. J.; Mackenzie, L.; Varrot, A.; Dauter, M.; Brzozowski, A. M.; Schulein, M.; Withers, S. G., *Biochemistry* **1998,** 37, 11707-11713.

(22) Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schulein, M.; Davies, G. J., *Biochemistry* **1996,** 35, 15280-15287.

(23) Tews, I.; Perrakis, A.; Oppenheim, A.; Dauter, Z.; Wilson, K. S.; Vorgias, C. E., *Nat. Struct. Biol.* **1996,** 3, 638-648.

(24) Ducros, V. M. A.; Zechel, D. L.; Murshudov, G. N.; Gilbert, H. J.; Szabo, L.; Stoll, D.; Withers, S. G.; Davies, G. J., *Angew. Chem., Int. Ed.* **2002,** 41, 2824-2827.

(25) Umezurike, G. M., *Biochem. J.* **1988,** 254, 73-76.

(26) Tull, D.; Withers, S. G., *Biochemistry* **1994,** 33, 6363-6370.

(27) Sinnott, M. L., *FEBS Lett.* **1978,** 94, 1-9.

(28) Li, B. F. L.; Holdup, D.; Morton, C. A. J.; Sinnott, M. L., *Biochem. J.* **1989,** 260, 109-114.

(29) Kempton, J. B.; Withers, S. G., *Biochemistry* **1992,** 31, 9961-9969.

(30) Withers, S. G.; Rupitz, K.; Street, I. P., *J. Biol. Chem.* **1988,** 263, 7929-7932.

(31) McCarter, J. D.; Withers, S. G., *J. Am. Chem. Soc.* **1996,** 118, 241-242.

(32) Namchuk, M. N.; McCarter, J. D.; Becalski, A.; Andrews, T.; Withers, S. G., *J. Am. Chem. Soc.* **2000,** 122, 1270-1277.

(33) Levvy, G. A.; Conchie, J.; Hay, A. J., *Biochem. J.* **1964,** 91, 378-384.

(34) Levvy, G. A.; Snaith, S. M., *Advances in Enzymology and Related Areas of Molecular Biology* **1972,** 36, 151-181.

(35) Walaszek, Z.; Horton, D.; Ekiel, I., *Carbohydr. Res.* **1982,** 106, 193-201.

(36) Nishimura, Y.; Adachi, H.; Satoh, T.; Shitara, E.; Nakamura, H.; Kojima, F.; Takeuchi, T., *J. Org. Chem.* **2000,** 65, 4871-4882.

(37) Strynadka, N. C. J.; James, M. N. G., *J. Mol. Biol.* **1991,** 220, 401-424.

(38) Numao, S.; Kuntz, D. A.; Withers, S. G.; Rose, D. R., *J. Biol. Chem.* **2003,** 278, 48074-48083.

(39) Kornfeld, R.; Kornfeld, S., *Annu. Rev. Biochem.* **1985,** 54, 631-664.

(40) Knauer, R.; Lehle, L., *Biochim. Biophys. Acta* **1999,** 1426, 259-273.

(41) Varki, A., *Trends Cell Biol.* **1998,** 8, 34-40.

(42) Buckhaults, P.; Chen, L.; Fregien, N.; Pierce, M., *J. Biol. Chem.* **1997,** 272, 19575-19581.

(43) Perng, G. S.; Shoreibah, M.; Margitich, I.; Pierce, M.; Fregien, N., *Glycobiology* **1994,** 4, 867-871.

(44) Pierce, M.; Buckhaults, P.; Chen, L.; Fregien, N., *Glycoconjugate J.* **1997,** 14, 623-630.

(45) Fernandes, B.; Sagman, U.; Auger, M.; Demetrio, M.; Dennis, J. W., *Cancer Res.* **1991,** 51, 718-723.

(46) Lemaire, S.; Derappe, C.; Michalski, J. C.; Aubery, M.; Neel, D., *J. Biol. Chem.* **1994,** 269, 8069-8074.

(47) Easton, E. W.; Blokland, I.; Geldof, A. A.; Rao, B. R.; Vandeneijnden, D. H., *FEBS Lett.* **1992,** 308, 46-49.

(48) Do, K. Y.; Fregien, N.; Pierce, M.; Cummings, R. D., *J. Biol. Chem.* **1994,** 269, 23456-23464.

(49) Dennis, J. W.; Kosh, K.; Bryce, D. M.; Breitman, M. L., *Oncogene* **1989,** 4, 853-860.

(50) Dennis, J. W.; Laferte, S.; Yagel, S.; Breitman, M. L., *Cancer Cells-a Monthly Review* **1989,** 1, 87-92.

(51) Itzkowitz, S. H.; Yuan, M.; Fukushi, Y.; Palekar, A.; Phelps, P. C.; Shamsuddin, A. M.; Trump, B. F.; Hakomori, S.; Kim, Y. S., *Cancer Res.* **1986,** 46, 2627-2632.

(52) Kim, Y. S.; Yuan, M.; Itzkowitz, S. H.; Sun, Q.; Kaizu, T.; Palekar, A.; Trump, B. F.; Hakomori, S., *Cancer Res.* **1986,** 46, 5985-5992.

(53) Granovsky, M.; Fata, J.; Pawling, J.; Muller, W. J.; Khokha, R.; Dennis, J. W., *Nat. Med.* **2000,** 6, 306-312.

(54) Duronio, V.; Jacobs, S.; Romero, P. A.; Herscovics, A., *J. Biol. Chem.* **1988,** 263, 5436-5445.

(55) Daniel, P. F.; Winchester, B.; Warren, C. D., *Glycobiology* **1994,** 4, 551-566.

(56) Moremen, K. W.; Trimble, R. B.; Herscovics, A., *Glycobiology* **1994,** 4, 113-125.

(57) Knight, J.; Beug, H.; Marshall, J.; Hayman, M. J., *Oncogene* **1988,** 2, 317-326.

(58) Keegan, A. D.; Conrad, D. H., *J. Immunol.* **1987,** 139, 1199-1205.

(59) Fischer, T.; Thoma, B.; Scheurich, P.; Pfizenmaier, K., *J. Biol. Chem.* **1990,** 265, 1710-1717.

(60) Akiyama, S. K.; Yamada, S. S.; Yamada, K. M., *J. Biol. Chem.* **1989,** 264, 18011-18018.

(61) Dennis, J. W.; Koch, K.; Yousefi, S.; Vanderelst, I., *Cancer Res.* **1990,** 50, 1867-1872.

(62) Bowlin, T. L.; Sunkara, P. S., *Biochem. Biophys. Res. Commun.* **1988,** 151, 859-864.

(63) Goss, P. E.; Baker, M. A.; Carver, J. P.; Dennis, J. W., *Clin. Cancer Res.* **1995,** 1, 935-944.

(64) Olden, K.; Breton, P.; Grzegorzewski, K.; Yasuda, Y.; Gause, B. L.; Oredipe, O. A.; Newton, S. A.; White, S. L., *Pharm. Ther.* **1991,** 50, 285-290.

(65) Humphries, M. J.; Olden, K., *Pharmacol. Ther.* **1989,** 44, 85-105.

(66) Das, P. C.; Roberts, J. D.; White, S. L.; Olden, K., *Oncology Res.* **1995,** 7, 425-433.

(67) Goss, P. E.; Reid, C. L.; Bailey, D.; Dennis, J. W., *Clin. Cancer Res.* **1997,** 3, 1077-1086.

(68) Novikoff, P. M.; Touster, O.; Novikoff, A. B.; Tulsiani, D. P., *J. Cell Biol.* **1985,** 101, 339-349.

(69) Tulsiani, D. R. P.; Touster, O., *Archives of Biochemistry and Biophysics* **1983,** 224, 594-600.

(70) Berg, T.; Riise, H. M. F.; Hansen, G. N.; Malm, D.; Tranebjaerg, L.; Tollersrud, O. K.; Nilssen, O., *American Journal of Human Genetics* **1999,** 64, 77-88.

(71) Nilssen, O.; Berg, T.; Riise, H. M. F.; Ramachandran, U.; Evjen, G.; Hansen, G. M.; Malm, D.; Tranebjaerg, L.; Tollersrud, O. K., *Human Molecular Genetics* **1997,** 6, 717-726.

(72) Aldaher, S.; Degasperi, R.; Daniel, P.; Hirani, S.; Warren, C.; Winchester, B., *Biochem. J.* **1992,** 286, 47-53.

(73) Liao, Y. F.; Lal, A.; Moremen, K. W., *J. Biol. Chem.* **1996,** 271, 28348-28358.

(74) Tulsiani, D. R. P.; Hubbard, S. C.; Robbins, P. W.; Touster, O., *J. Biol. Chem.* **1982,** 257, 3660-3668.

(75) van den Elsen, J. M. H.; Kuntz, D. A.; Rose, D. R., *EMBO J.* **2001,** 20, 3008-3017.

(76) Howard, S.; He, S. M.; Withers, S. G., *J. Biol. Chem.* **1998,** 273, 2067-2072.

(77) Numao, S.; He, S. M.; Evjen, G.; Howard, S.; Tollersrud, O. K.; Withers, S. G., *FEBS Lett.* **2000,** 484, 175-178.

(78) Alberts, I. L.; Nadassy, K.; Wodak, S. J., *Protein Sci.* **1998,** 7, 1700-1716.

(79) Deslongchamps, P., *Stereoelectronic effects in organic chemistry*. Pergamon Press: Oxford, 1983; Vol. 1.

(80) Kuntz, D. A.; Liu, H.; Bols, M.; Rose, D. R., *Biocatalysis Biotransform.* **2006,** 24, 1-7.

(81) Shah, N.; Kuntz, D. A.; Rose, D. R., *Biochemistry* **2003,** 42, 13812-13816.

(82) Kuntz, D. A.; Ghavami, A.; Johnston, B. D.; Pinto, B. M.; Rose, D. R., *Tetrahedron Asymmetry* **2005,** 16, 25-32.

(83) Boss, O.; Leroy, E.; Blaser, A.; Reymond, J. L., *Org. Lett.* **2000,** 2, 151-154.

(84) King, S. B.; Ganem, B., *J. Am. Chem. Soc.* **1994,** 116, 562-570.

(85) Kleban, M.; Hilgers, P.; Greul, J. N.; Kugler, R. D.; Li, J.; Picasso, S.; Vogel, P.; Jager, V., *Chembiochem* **2001,** 2, 365-368.

(86) Ogawa, S.; Morikawa, T., *Bioorg. Med. Chem. Lett.* **1999,** 9, 1499-1504.

(87) Ogawa, S.; Morikawa, T., *Bioorg. Med. Chem. Lett.* **2000,** 10, 1047-1050.

(88) Popowycz, F.; Gerber-Lemaire, S.; Demange, R.; Rodriguez-Garcia, E.; Asenjo, A. T. C.; Robina, I.; Vogel, P., *Bioorg. Med. Chem. Lett.* **2001,** 11, 2489-2493.

(89) Winkler, D. A., *J. Med. Chem.* **1996,** 39, 4332-4334.

(90) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J., *J. Comput. Chem.* **1998,** 19, 1639-1662.

(91) Lyne, P. D., *Drug Discov. Today* **2002,** 7, 1047-1055.

(92) Wesson, L.; Eisenberg, D., *Protein Sci.* **1992,** 1, 227-235.

(93) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J., *J. Comput.-Aided Mol. Design* **1996,** 10, 293-304.

(94) Stouten, P. F. W.; Frommel, C.; Nakamura, H.; Sander, C., *Molecular Simulation* **1993,** 10, 97-&.

(95) Mccammon, J. A.; Gelin, B. R.; Karplus, M., *Nature* **1977,** 267, 585-590.

(96) Hansson, T.; Oostenbrink, C.; van Gunsteren, W. F., *Curr. Opin. Struct. Biol.* **2002,** 12, 190-196.

(97) Case, D. A., et al. *Amber 7.0*, University of California: San Francisco, CA, 2002.

(98) Verlet, L., *Physical Review* **1967,** 159, 98-103.

(99) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *J. Phys. Chem.* **1983,** 79, 926-935.

(100) Friesner, R. A.; Baik, M. H.; Gherman, B. F.; Guallar, V.; Wirstam, M.; Murphy, R. B.; Lippard, S. J., *Coord. Chem. Rev.* **2003,** 238, 267-290.

(101) Monard G., P.-R. X., Gonzalez-Lafont A. and Lluch J., *Int. J. Quantum Chem.* **2003,** 93, 229-244.

(102) Field, M. J.; Bash, P. A.; Karplus, M., *J. Comput. Chem.* **1990,** 11, 700-733.

(103) Antes, I.; Thiel, W., *J. Phys. Chem. A* **1999,** 103, 9290-9295.

(104) Thery, V.; Rinaldi, D.; Rivail, J. L.; Maigret, B.; Ferenczy, G. G., *J. Comput. Chem.* **1994,** 15, 269-282.

(105) Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J., *J. Phys. Chem. A* **1998,** 102, 4714-4721.

(106) Cunningham, M. A.; Ho, L. L.; Nguyen, D. T.; Gillilan, R. E.; Bash, P. A., *Biochemistry* **1997,** 36, 4800-4816.

(107) Mulholland, A. J.; Richards, W. G., *Proteins* **1997,** 27, 9-25.

(108) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M., *J. Phys. Chem. A* **2000,** 104, 1720-1735.

(109) Ridder, L.; Mulholland, A. J.; Rietjens, I. M. C. M.; Vervoort, J., *J. Am. Chem. Soc.* **2000,** 122, 8728-8738.

(110) Ridder, L.; Mulholland, A. J.; Vervoort, J.; Rietjens, I. M. C. M., *J. Am. Chem. Soc.* **1998,** 120, 7641-7642.

(111) Schrodinger, I. *Jaguar 4.2*, Portland, OR, 2001.

(112) Frisch, M. J., et al. *Gaussian 98*, Revision A.11.3 ed; Gaussian, Inc.: Pittsburgh, PA, 2002.

(113) Dennis, J. W.; Granovsky, M.; Warren, C. E., *Biochim. Biophys. Acta* **1999,** 1473, 21-34.

(114) Heightman, T. D.; Vasella, A. T., *Angew. Chem., Int. Ed.* **1999,** 38, 750-770.

(115) McCarter, J. D.; Withers, S. G., *Curr. Opin. Struct. Biol.* **1994,** 4, 885-892.

(116) Aoyagi, T.; Yamamoto, T.; Kojiri, K.; Morishima, H.; Nagai, M.; Hamada, M.; Takeuchi, T.; Umezawa, H., *J. Antibiot.* **1989,** 42, 883-889.

(117) Ogawa, S.; Washida, K., *Eur. J. Org. Chem.* **1998**, 1929-1934.

(118) Ogawa, S.; Morikawa, T., *Eur. J. Org. Chem.* **2000**, 1759-1765.

(119) Uchida, C.; Kimura, H.; Ogawa, S., *Bioorg. Med. Chem.* **1997,** 5, 921-939.

(120) Szardenings, A. K.; Burkoth, T. S.; Look, G. C.; Campbell, D. A., *J. Org. Chem.* **1996,** 61, 6720-6722.

(121) Schwarz, M. K.; Tumelty, D.; Gallop, M. A., *J.Org.Chem.* **1999,** 64, 2219-2231.

(122) Salvatore, R. N.; Nagle, A. S.; Jung, K. W., *J. Org. Chem.* **2002,** 67, 674-683.

(123) Misago, M.; Liao, Y. F.; Kudo, S.; Eto, S.; Mattei, M. G.; Moremen, K. W.; Fukuda, M. N., *Proc. Natl.Acad.Sci. USA* **1995,** 92, 11766-11770.

(124) Lalegerie, P.; Legler, G.; Yon, J. M., *Biochimie* **1982,** 64, 977-1000.

(125) Knapp, S.; Dhar, T. G. M., *J. Org. Chem.* **1991,** 56, 4096-4097.

(126) Rabouille, C.; Kuntz, D. A.; Lockyer, A.; Watson, R.; Signorelli, T.; Rose, D. R.; van den Heuvel, M.; Roberts, D. B., *J. Cell Sci.* **1999,** 112, 3319-3330.

(127) Kilpatrick, J. E.; Pitzer, K. S.; Spitzer, R., *J. Am. Chem. Soc.* **1947,** 69, 2483-2488.

(128) Pitzer, K. S.; Donath, W. E., *J. Am. Chem. Soc.* **1959,** 81, 3213-3218.

(129) Chung-Phillips, A.; Chen, Y. Y., *J. Phys. Chem. A* **1999,** 103, 953-964.

(130) Cloran, F.; Carmichael, I.; Serianni, A. S., *J. Phys. Chem. A* **1999,** 103, 3783-3795.

(131) Gordon, M. T.; Lowary, T. L.; Hadad, C. M., *J. Org. Chem.* **2000,** 65, 4954-4963.

(132) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B., *J. Phys. Chem.* **1996,** 100, 11775-11788.

(133) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A.; Honig, B., *J. Am. Chem. Soc.* **1994,** 116, 11875-11882.

(134) Boraston, A. B.; Chiu, P.; Warren, R. A. J.; Kilburn, D. G., *Biochemistry* **2000,** 39, 11129-11136.

(135) Lemieux, R. U., *Acc. Chem. Res.* **1996,** 29, 373-380.

(136) Lemieux, R. U.; Delbaere, L. T. J.; Beierbeck, H.; Spohr, U., *Ciba Foundation Symposia* **1991,** 158, 231-248.

(137) McGaughey, G. B.; Gagne, M.; Rappe, A. K., *J. Biol. Chem.* **1998,** 273, 15458-15463.

(138) Frisch, M. J. T., G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revision E.2; Gaussian, Inc.: Pittsburgh, 1995.

(139) Becke, A. D., *J. Chem. Phys.* **1993,** 98, 5648-5652.

(140) Becke, A. D., *Phys. Rev. A* **1988,** 38, 3098-3100.

(141) Lee, C. T.; Yang, W. T.; Parr, R. G., *Phys. Rev. B* **1988,** 37, 785-789.

(142) Hariharan, P. C.; Pople, J. A., *Theor. Chim. Acta* **1973,** 28, 213-222.

(143) Sundaralingam, M., *J. Am. Chem. Soc.* **1971,** 93, 6644-&.

(144) Goodsell, D. S.; Morris, G. M.; Olson, A. J., *J. Mol. Recognit.* **1996,** 9, 1-5.

(145) Chiche, L.; Gregoret, L. M.; Cohen, F. E.; Kollman, P. A., *Proc. Natl. Acad. Sci. U. S. A.* **1990,** 87, 3240-3243.

(146) Gasteiger, J.; Marsili, M., *Tetrahedron* **1980,** 36, 3219-3228.

(147) Mehler, E. L.; Solmajer, T., *Protein Eng.* **1991,** 4, 903-910.

(148) Elbein, A. D., *FASEB J.* **1991,** 5, 3055-3063.

(149) Tulsiani, D. R. P.; Harris, T. M.; Touster, O., *J. Biol. Chem.* **1982,** 257, 7936-7939.

(150) Boraston, A. B.; Nurizzo, D.; Notenboom, V.; Ducros, V.; Rose, D. R.; Kilburn, D. G.; Davies, G. J., *J. Mol. Biol.* **2002,** 319, 1143-1156.

(151) Chavez, M. I., et al., *Chem. Eur. J.* **2005,** 11, 7060-7074.

(152) Fernandez, M. D.; Canada, F. J.; Jimenez-Barbero, J.; Cuevas, G., *J. Am. Chem. Soc.* **2005,** 127, 7379-7386.

(153) Li, B.; Kawatkar, S. P.; George, S.; Strachan, H.; Woods, R. J.; Siriwardena, A.; Moremen, K. W.; Boons, G. J., *Chembiochem* **2004,** 5, 1220-1227.

(154) Reid, K. S. C.; Lindley, P. F.; Thornton, J. M., *FEBS Lett.* **1985,** 190, 209-213.

(155) Pal, D.; Chakrabarti, P., *J. Biomol. Struct. Dynamics* **1998,** 15, 1059-1072.

(156) Pal, D.; Chakrabarti, P., *J. Biomol. Struct. Dynamics* **2001,** 19, 115-128.

(157) Samanta, U.; Pal, D.; Chakrabarti, P., *Proteins* **2000,** 38, 288-300.

(158) Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S., *Bull. Chem. Soc. Jpn.* **2002,** 75, 1611-1625.

(159) Ippolito, J. A.; Alexander, R. S.; Christianson, D. W., *J. Mol. Biol.* **1990,** 215, 457-471.

(160) Brandl, M.; Weiss, M. S.; Jabs, A.; Suhnel, J.; Hilgenfeld, R., *J. Mol. Biol.* **2001,** 307, 357-377.

(161) Tatko, C. D.; Waters, M. L., *Protein Sci.* **2004,** 13, 2515-2522.

(162) Alber, F.; Kuonen, O.; Scapozza, L.; Folkers, G.; Carloni, P., *Proteins* **1998,** 31, 453-459.

(163) Pranata, J., *Bioorganic Chem.* **1997,** 25, 213-219.

(164) Cheney, B. V.; Schulz, M. W.; Cheney, J., *Biochim. Biophys. Acta* **1989,** 996, 116-124.

(165) Duan, G. L.; Smith, V. H.; Weaver, D. F., *Mol. Phys.* **2001,** 99, 1689-1699.

(166) Haasnoot, C. A. G.; Deleeuw, F.; Altona, C., *Tetrahedron* **1980,** 36, 2783-2792.

(167) Varrot, A.; Tarling, C. A.; Macdonald, J. M.; Stick, R. V.; Zechel, D. L.; Withers, S. G.; Davies, G. J., *J. Am. Chem. Soc.* **2003,** 125, 7496-7497.

(168) Sinnokrot, M. O.; Sherrill, C. D., *J. Phys. Chem. A* **2004,** 108, 10200-10207.

(169) Winkler, D. A.; Holan, G., *J. Med. Chem.* **1989,** 32, 2084-2089.

(170) Morishima, H.; Kojiri, K.; Yamamoto, T.; Aoyagi, T.; Nakamura, H.; Iitaka, Y., *J. Antibiot.* **1989,** 42, 1008-1011.

(171) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., *J. Comput. Phys.* **1977,** 23, 327-341.

(172) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraserreid, B., *J. Phys. Chem.* **1995,** 99, 3832-3846.

(173) Cremer, D.; Pople, J. A., *J. Am. Chem. Soc.* **1975,** 97, 1354-1358.

(174) Kawatkar, S. P.; Kuntz, D. A.; Woods, R. J.; Rose, D. R.; Boons, G. J., *J. Am. Chem. Soc.* **2006**.

(175) Comba, P.; Zimmer, M., *J. Chem. Edu.* **1996,** 73, 108-110.

(176) Cundari, T. R., *J. Chem. Soc. Dalton Transactions* **1998**, 2771-2776.

(177) Parr, R. G.; Yang, W., *Density Functional Theory of Atoms and Molecules*. Oxford University Press: New York, 1989.

(178) Ziegler, T., *Chem. Rev.* **1991,** 91, 651-667.

(179) Vacek, G.; Perry, J. K.; Langlois, J. M., *Chem. Phys. Lett.* **1999,** 310, 189-194.

(180) Andre, I.; Tvaroska, I.; Carver, J. P., *Carbohydr. Res.* **2003,** 338, 865-877.

(181) Tvaroska, I.; Andre, I.; Carver, J. P., *Glycobiology* **2003,** 13, 559-566.

(182) Tvaroska, I.; Andre, I.; Carver, J. P., *J. Am. Chem. Soc.* **2000,** 122, 8762-8776.

(183) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, Jr., R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A.G.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E.S. and Pople, J. A. *Gaussian 98*, Revision A9; Gaussian, Inc.: Pittsburgh, PA, 1998.

(184) Hay, P. J.; Wadt, W. R., *J. Chem. Phys.* **1985,** 82, 299-310.

(185) Wiberg, K. B., *Tetrahedron* **1968,** 24, 1083-&.

(186) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Weinhold, F. *NBO 5.0*, Theoretical Chemistry Institute, University of Wisconsin: Madison, 2001.

(187) Karaveg, K.; Siriwardena, A.; Tempel, W.; Liu, Z. J.; Glushka, J.; Wang, B. C.; Moremen, K. W., *J. Biol. Chem.* **2005,** 280, 16197-16207.