

New inhibitors of Golgi α -mannosidase II



Doctoral thesis

submitted in fulfilment of the regulations for attaining the degree of
"Doktor der Naturwissenschaften (Dr. rer. nat.)"
in the subject of Medicinal Chemistry

at the Faculty of Chemistry, Pharmacy and Geoscience
at the Johannes Gutenberg University of Mainz

Lina Irsheid
born in Saarbrücken

Mainz 2018

Submitted at the Faculty of Chemistry,
Pharmacy and Geoscience:

Dean: [REDACTED]

Name of 1st reviewer: [REDACTED]

Name of 2nd reviewer: [REDACTED]

Date of doctoral examination:

D77 (Dissertation Universität Mainz)

Declaration of Authorship

I, Lina Irsheid, declare that this thesis titled "New inhibitors of Golgi- α -mannosidase II" and the work presented in it are my own. I confirm that

- This work was done wholly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

- [REDACTED]

Signed:

Date:

"La vie n'est pas facile pour aucun d'entre nous. Mais que dire de qui? Nous devons persévérer et surtout avoir confiance en nous-mêmes. Nous devons croire que nous sommes doués pour quelque chose et que nous devons l'atteindre."

Marie Curie

Pour Marc

List of Abbreviation

[x]	Concentration [x]
°C	Degree Celcius
λ_{emm}	Emmissions wavelength
λ_{ex}	Excitation wavelength
μL	Microliter
4-MU-Man	4-Methylumbelliferyl-α-D-mannopyranoside
4-MU-Glc	4-Methylumbelliferyl-α-D-glucopyranoside
Arg/ R	Arginine/R
Asp/ D	Aspartate
dGMII	Golgi-α-Mannosidase of <i>Drosophila melanogaster</i>
E*I	High affinity enzyme-inhibitor complex after conformational change
EI	Enzyme-Inhibitor-complex
ES	Enzyme-Substrate-complex
ESI	Enzyms-Substrate-Inhibitor-complex
et al.	<i>et alii/aliae/alia</i>
F	Fluorescence unit
Glu/E	Glutamate
hGMII	Human Golgi-α-Mannosidase
His/H	Histidine
hLM	Human lysosomale Mannosidase
IC₅₀	Half maximal inhibitory concentration
JBM	Jack Bean α-Mannosidase
JGU	Johannes-Gutenber Universität
kDA	Kilodalton
Lys/K	Lysine
K_i	Dissociation constant of the, enzyme-inhibitor complex (EI)
K_i'	Dissociation constant of the enzyme-substrate-inhibitor complex (ESI)
K_i[*], K_i^{app}	Dissociation constant of the high-affiniy, isomerized enzyme-inhibitor complex (E*I)

K_M	Michalis-Menten-constant
k_{obs}	Pseudo-first order rate constant
JMU	Julius-Maximilians-Universität Würzburg
mg	Milligramm
min	Minute
mL	Milliliter
mM	Millimolar
MES	2-(<i>N</i> -Morpholino)ethan-sulfonsäure
MOE	Molecular Operating Environment
nM	nanomolar
nm	nanometer
off	Offset
Phe/F	Phenylalanine
pNP-Glc	4-Nitrophenyl-β- <i>D</i> -glucopyranoside
Pro/P	Proline
Thr/T	Threonine
Trp/W	Tryptophane
TS	Transition state
Tyr/Y	Tyrosine
U	Enzyme activity
v	velocity
ZnSO₄	Zinc sulfate

Abstract

Faculty of Chemistry, Pharmaceutical Sciences and Geoscience Institute of
Pharmacy and Biochemistry

New inhibitors of Golgi α -mannosidase II

by Lina Irsheid

Golgi α -mannosidase II (GMII) plays a crucial role in the *N*-glycosylation pathway. In various tumor cell lines, the distribution of the *N*-linked sugars on the cell surface is modified and correlates with the progression of tumor metastasis. Thus, GMII is a molecular target for anticancer agents and its inhibition has shown to induce tumor repression. GMII, a member of the family 38 glycoside hydrolases, cleaves two mannose units (α -(1,3) and α -(1,6)) of the intermediate $\text{GlcNAcMan}_5(\text{GlcNAc})_2$. The active site of the enzyme mainly consists of two aspartate residues and a zinc cation. GMII acts as a retaining glycosidase and cleaves the sugars in a two-step S_N2 mechanism resulting in a covalent glycosyl-enzyme complex. The mechanism preserves the configuration of the anomeric C-atom. Several natural product-based or synthetic inhibitors have been investigated so far. However, the clinical use of the known potent inhibitor swainsonine is restricted due to the side effects resulting from inhibition of the closely related lysosomal α -mannosidases. Three virtual screening approaches with a library of purchasable compounds to identify new chemotypes inhibiting the GMII were followed. The docking studies were carried out against the known active site and a potential allosteric binding site identified by molecular modeling. The most promising candidates were purchased and subjected to enzyme inhibition assays. The aim of the present work was to find new lead compounds for the inhibition of GMII. Two hits were identified by the enzyme assays, namely compound **18** ($IC_{50} \approx 200 \mu\text{M}$) consisting of a pyridinyl-2-acetonitrile and a nicotinoyl-piperidine moiety and compound **32** ($K_i = 63 \mu\text{M}$) showing a spiroketal-indolin-2-one structure. Both compounds will represent the first building blocks for the design of new inhibitors of GMII.

Zusammenfassung

Fachbereich 09 - Chemie, Pharmazie und Geowissenschaften Institut für Pharmazie
und Biochemie

Neue Inhibitoren der Golgi α -mannosidase II

Von Lina Irsheid

Die Golgi α -mannosidase II (GMII) ist eines der Schlüsselenzyme im *N*-Glykosylierungsprozess. In verschiedenen Tumorzelllinien ist die Verteilung der Glykosylierungsmuster modifiziert und korreliert mit dem Tumorwachstum. Die GMII stellt daher ein attraktives Target für neue Zytostatika dar und die Hemmung dieses Enzyms führte bereits in klinischen Studien bei Patienten mit fortgeschrittenem Tumorstadium zur Reduktion des Tumorwachstums und der Metastasen. GMII, ein Mitglied der Familie 38 der Glykosylhydrolasen, spaltet zwei Mannoseeinheiten (α -(1,3) und α -(1,6)) des Zwischenprodukts $\text{GlcNAcMan}_5(\text{GlcNAc})_2$ ab. Das aktive Zentrum des Enzyms besteht hauptsächlich aus zwei Aspartat-Resten und einem Zink-Ion. GMII ist eine retendierende Glycosidase und spaltet die Zucker in einem zweistufigen S_N2 -Mechanismus, was zu einem kovalenten Glykosyl-Enzym-Komplex führt. Dieser Mechanismus bewahrt die Konfiguration des anomeren C-Atoms. Mehrere Inhibitoren auf Basis natürlicher oder synthetischer Produkte wurden untersucht. Die klinische Verwendung des bekannten Inhibitors Swainsonin ist jedoch aufgrund der Nebenwirkungen eingeschränkt, die aus der Hemmung der eng verwandten lysosomalen α -Mannosidasen resultieren. Es wurden drei virtuelle Screening-Ansätze mit einer Bibliothek von kommerziell erhältlichen Verbindungen verfolgt, um neue Chemotypen zu identifizieren, welche die GMII inhibieren. Das Docking wurde an dem bekannten aktiven Zentrum und an einer potentiellen allosterischen Bindungsstelle durchgeführt, die durch *molecular modeling* identifiziert wurde. Die Treffer des virtuellen Screenings wurden in Enzym Assays experimentell untersucht. Ziel dieser Arbeit war es, neue Leitstrukturen für die Hemmung der GMII zu finden. Zwei Treffer konnten identifiziert werden, wobei Verbindung **18** ($IC_{50} \approx 200 \mu\text{M}$) eine Pyridinyl-2-acetonitril- sowie eine Nicotinoyl-piperidin-Struktur und Verbindung **32** ($K_i = 63 \mu\text{M}$) eine Spiroketal-indolin-2-on-Struktur aufweist. Diese stellen die ersten Bausteine für das Design neuer GMII-Inhibitoren dar.

Acknowledgements

I would like to thank the many people who contributed to the realisation of this work .

First and foremost, I would like to express my gratitude to [REDACTED] for the opportunity to work on such an interesting project and for the trust she has always placed in me concerning research and teaching. Her unwavering support made it possible for me to juggle motherhood and the work.

Many thanks to [REDACTED] for his open ear for my inquiries. He always willingly and knowledgeably answered all my many questions and remains a person of good-heartedness and patience. I have learnt a great deal from him, reaching well beyond his scientific expertise.

I am much obliged to [REDACTED] for his continuous support during the course of my PhD. He has been of tremendous help for the achievement of this work whereby I have gained a lot from his linux skills and shared his scientific curiosity. Not only has Thomas been an esteemed colleague, he has become a true friend.

My special thanks go to [REDACTED]. Her advice on enzyme assays has been invaluable. I will miss our delightful conversations over coffee. I am also so grateful to [REDACTED] for being a good interlocutor.

Thanks always to the former and current members of the [REDACTED]
[REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]. They have always been ambitious in their scientific research, willing to help and share information and fortunately inclined to have fun. Thus they have provided a friendly and cooperative atmosphere at work.

I tip my hat to all the diploma-, bachelor-, and pharmacy students for their help and support. They have greatly contributed to many of the results presented in this thesis.

I am sincerely appreciative for the support of the working group of [REDACTED]
[REDACTED] for supplying us with the enzyme dGMII.

Last but not least, I am thankful to [REDACTED] for his unconditional support and for having my back whilst realising my PhD. My perpetual thanks to [REDACTED]. He lifted my spirits across all these years while commuting between Mainz and Bonn. Likewise, I am deeply indebted to

[REDACTED], who have always encouraged me and backed my determination.

Contents

Declaration of Authorship	i
List of Abbreviation	iii
Abstract	v
Zusammenfassung	vi
Acknowledgements	vii
1 Introduction	1
1.1 Metastasis process in detail: selectins, altered glycosylation patterns and overexpressed glycosidases	1
1.2 Golgi- α -Mannosidase II	4
1.2.1 Structure and characteristics	5
1.3 β -glucosidase	10
1.4 Jack Bean Mannosidase	13
1.5 Glycosidase inhibitors - Current State	13
1.6 Virtual screening and molecular docking	18
1.6.1 DOCK	20
1.6.2 FlexX	23
1.7 Enzyme assays and enzyme kinetics	25
1.7.1 Enzyme kinetic	26
1.7.2 Enzyme Inhibition	28
1.7.3 The Z-Factor and the Z'-Factor	35
2 Objectives	39
3 Material and Methods	41
3.1 Homology modeling	41
3.2 RMSD value	42
3.3 Molecular graphics	43
3.4 Virtual ligand screening of the catalytic site of the homology modell of hGMII	44
3.4.1 Receptor preparation	44
3.4.2 Docking of swainsonine for testing the receptor setup	44
3.4.3 Database preparation	44
3.4.4 Pharmacophore search	45
3.4.5 Molecular docking	46
3.4.6 Docking analysis and molecular graphics	46

3.5	Virtual ligand screening of the potential allosteric site of dGMII.....	47
3.5.1	Identification of a potential allosteric binding site of Golgi α -mannosidase II.....	47
3.5.2	Comparison of binding site properties – active site vs. potential allosteric site	47
3.5.3	Receptor preparation and sphere set generation for virtual screening.....	48
3.5.4	Small molecule preparation.....	48
3.5.5	Pharmacophore search.....	49
3.5.6	Molecular docking	50
3.5.7	Docking analysis and molecular graphics.....	50
3.6	Virtual ligand screening of the active site of dGMII.....	52
3.6.1	Receptor preparation and sphere set generation	52
3.6.2	Redocking of swainsonine	52
3.6.3	Small molecule preparation.....	52
3.6.4	Pharmacophore search.....	53
3.6.5	Molecular docking	53
3.6.6	Docking analysis and molecular graphics.....	53
3.7	Enzyme Assay	54
3.7.1	Photometric assay	54
3.7.2	Fluorometric assay.....	55
4	Results	73
4.1	Homology modeling	73
4.1.1	Homology Model of the hGMII.....	73
4.1.2	Homology model of JBM	75
4.1.3	Homology model of the sweet almond β -glucosidase.....	77
4.2	Virtual ligand screening of the catalytic site of hGMII	79
4.2.1	Docking of swainsonine for testing the receptor setup.....	79
4.2.2	Filtering the database.....	80
4.2.3	The pharmacophore model	80
4.2.4	Screening of a small compound library	82
4.3	Virtual ligand screening of the potential allosteric site of dGMII.....	85
4.3.1	Identification of a potential allosteric binding site of dGMII	85
4.3.2	Comparison of binding site properties- active site versus potential allosteric site	88
4.3.3	Pharmacophore hypothesis.....	89
4.3.4	Molecular docking	91
4.4	Virtual ligand screening for ligands of the catalytic site of dGMII	93
4.4.1	Redocking of swainsonine	93
4.4.2	Filtering the database.....	94

4.4.3	The pharmacophore model	94
4.4.4	Screening of a large compound library	95
4.5	Enzyme assays	97
4.5.1	Photometric Assay	97
4.5.2	Fluorometric Assay	99
5	Discussion	135
5.1	Homology modeling	135
5.1.1	Homology model of the hGMII.....	135
5.1.2	Homology model of JBM	135
5.1.3	Homology model of the sweet almond β -glucosidase.....	136
5.2	Enzyme Assay	137
5.2.1	Enzyme assays with <i>drosophila melanogaster</i> Golgi α -mannosidase II.....	137
5.2.2	Enzyme assays with JBM	142
5.2.3	Enzyme assays with the β -glucosidase from sweet almonds	146
5.3	Discussion of the virtual screening results.....	147
5.3.1	Virtual ligand screening of the catalytic site of the homology model of hGMII.....	147
5.3.2	Virtual ligand screening of the potential allosteric site of dGMII.....	152
5.3.3	Virtual ligand screening of the catalytic site of dGMII	154
6	Conclusion and future perspectives	157
7	Bibliography	161
8	Curriculum Vitae.....	171

1 Introduction

Studies with late-stage cancer patients showed that a specific inhibition of Golgi- α -Mannosidase II (GMII) by swainsonine resulted in a reduction of tumour growth and metastasis [1], [2]. The reasons for the reduction of tumour growth and metastasis and the major role of GMII will be discussed hereafter.

1.1 Metastasis process in detail: selectins, altered glycosylation patterns and overexpressed glycosidases

Metastases are responsible for 90% of all cancer deaths. The term metastasis refers to the spread of one or more tumour cells from the primary tumour to distant organ sites of the body. The invasion-metastasis cascade refers to a sequence of biological processes which lead to the formation of metastases (see Figure 1). Metastasis starts with the local invasion of one or more cells of the primary tumour. These tumour cells penetrate the vascular or lymphatic channel where they either grow at the place of penetration or detach and are transported via the circulatory system to a distant organ site. Circulating tumour cells undergo a variety of stress such as lymphocytes, monocytes and natural killer cells or blood turbulence. Surviving tumour cells arrest at a distant organ site either by adhering to endothelial cells or by attaching at the sites of endothelial cell retraction leading to a higher vascular permeability. The next step is the process of leaving a vessel or wall and penetrating the surrounding tissue, which is called extravasation. Extravasated cancer cells must survive in the parenchym environment of the distant organ tissue to multiply and eventually to form a secondary (or metastatic) tumour [3]–[5]. However, it is widely known that most injected cancer cells are not able to form metastatic tumours [6], [7]. For example, experimental data from Sakurai *et al.* [8] have shown that only 1.5 % of human lung cancer cells injected into the circulation survived after 24h post-injection. This study suggested that very few surviving tumour cells are needed to cause metastasis.

The arrest of cancer cells is mediated by the interaction between the carbohydrate ligands on the outer surface of cancer cells and the E-selectins on the endothelium [9]. On the surface of many cancer cells, the carbohydrate structures of sialyl Lewis A (sLe^a) and sialyl Lewis X (sLe^x) were found [9]–[13]. The expression of E-selectin is limited to activated endothelium [14].

Figure 1 displays a simplified representation of the steps of metastasis:

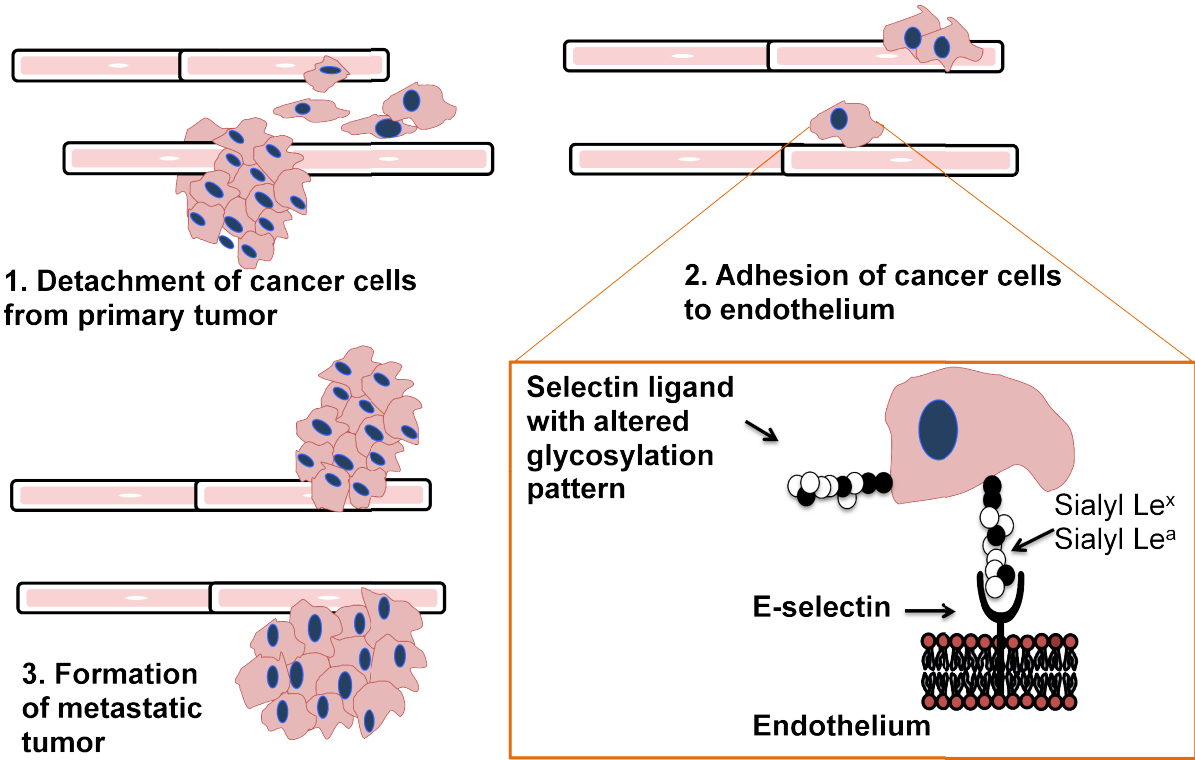


Figure 1: Simplified representation of the different steps involved in cancer metastasis after Kannagi et al. [15] and Jubeli et al.[16]

Furthermore, other studies showed that not only E-selectins but also P- and L- selectins are involved in cancer metastasis. P-selectin is stored in α -granules of endothelial cells or platelets. L-selectin is expressed on leukocytes. The selectin ligands on cancer cells may interact with leukocytes and platelets facilitating the attachment to the endothelium and thus the formation of metastases (see Figure 2) [15], [17].

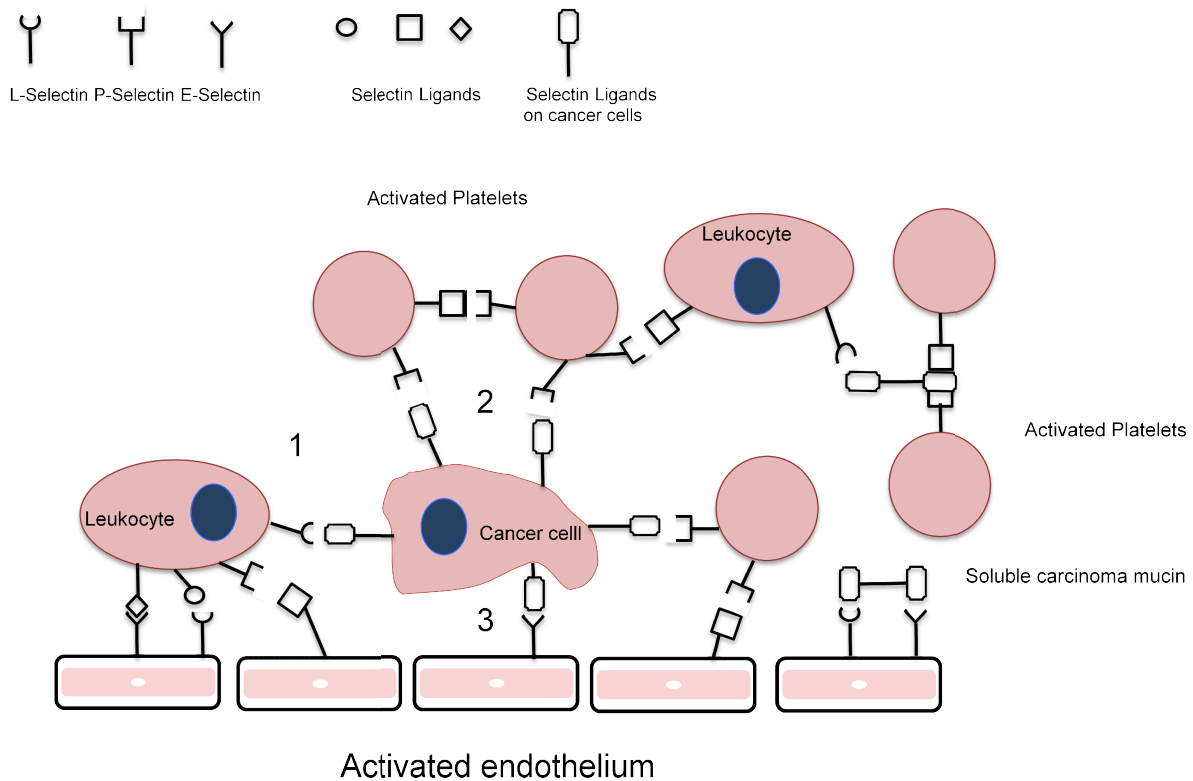


Figure 2: Possible tumor-selectin-cell interactions during the metastasis process after Läubli et al. [17]

Selectins preferably recognize SLe^x and SLe^a tetrasaccharide motifs that are synthesized by glycosyltransferases such as α -1,3-fucosyltransferases IV or VII, α -2,3-sialyltransferases, β -1,4 galactosyltransferases and *N*-acetylglucosaminyltransferases [14], [18]. In many cases of cancer, the altered glycosylation patterns are the consequence of a malfunction of the involved glycosyltransferases [19]–[22]. For example, the fucosyltransferase IV and VII, which are responsible for the synthesis of SLe^x , are overexpressed on epithelial cancer cells. It has been shown that an increased expression of the fucosyltransferase IV and VII gene is observed in lung cancers and that a correlation between SLe^x and fucosyltransferase VII expression is remaining. Furthermore, a poor prognosis of patients with overexpressed fucosyltransferase IV and VII genes is reported [23]. It has been proven difficult to develop effective glycosyltransferase inhibitors due to the limited amount of available 3D-structures for these enzymes, low substrate affinity of the enzymes for their natural substrates and complex transition states [24]–[26]. Thus it was attempted to intervene earlier in the processing of the glycosylation chains.

The asparagine-linked (*N*-linked) glycosylation is a co-translational process in the endoplasmic reticulum (ER). Initially, the same high-mannose-type tetradecasaccharide consisting of three glucose (Glc), nine mannose (Man) and two *N*-acetylglucosamine (GlcNAc) residues is attached to each protein. Then, this oligosaccharide chain is truncated

by the glucosidases α -glucosidase I and II as well as α -mannosidase I. The protein arrives at the Golgi apparatus, where glycosyltransferases and glycosidases continue the processing. GlcNAc-transferase I transmits a GlcNAc residue. The GMII continues its processing and trims the α 1,3-Man and α 1,6-Man in order to form GlcNAcMan₃ (GlcNAc)₂-protein [27]–[29]. This enzyme plays an essential role in the *N*-glycosylation pathway and its inhibition has shown to suppress the formation of altered glycosylation patterns [28], [30]–[32].

1.2 Golgi- α -Mannosidase II

The Golgi α -mannosidase II (GMII) is a key enzyme of the *N*-glycosylation pathway and displays a potential target for the development of novel inhibitors for the treatment of metastatic cancers. It shortens the branched mannose intermediate GlcNAcMan₅GlcNAc₂ by cleaving the two mannose residues M4 and M5, which are linked by two different glycosidic bonds (α -1,3 and α -1,6, respectively) to the core structure GlcNAcMan₃GlcNAc₂ (see Figure 3) [1],[2],[32].

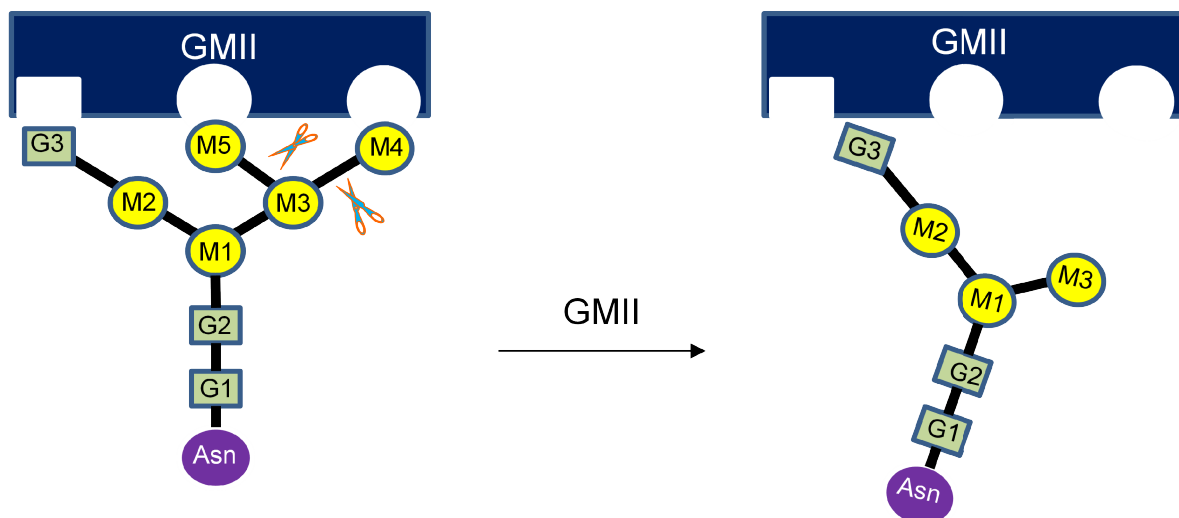


Figure 3: Schematic representation of the reactions catalyzed by the GMII. Yellow circles: mannose, gray squares: N-acetylglucosamine. Figure following Shah et al. [3]

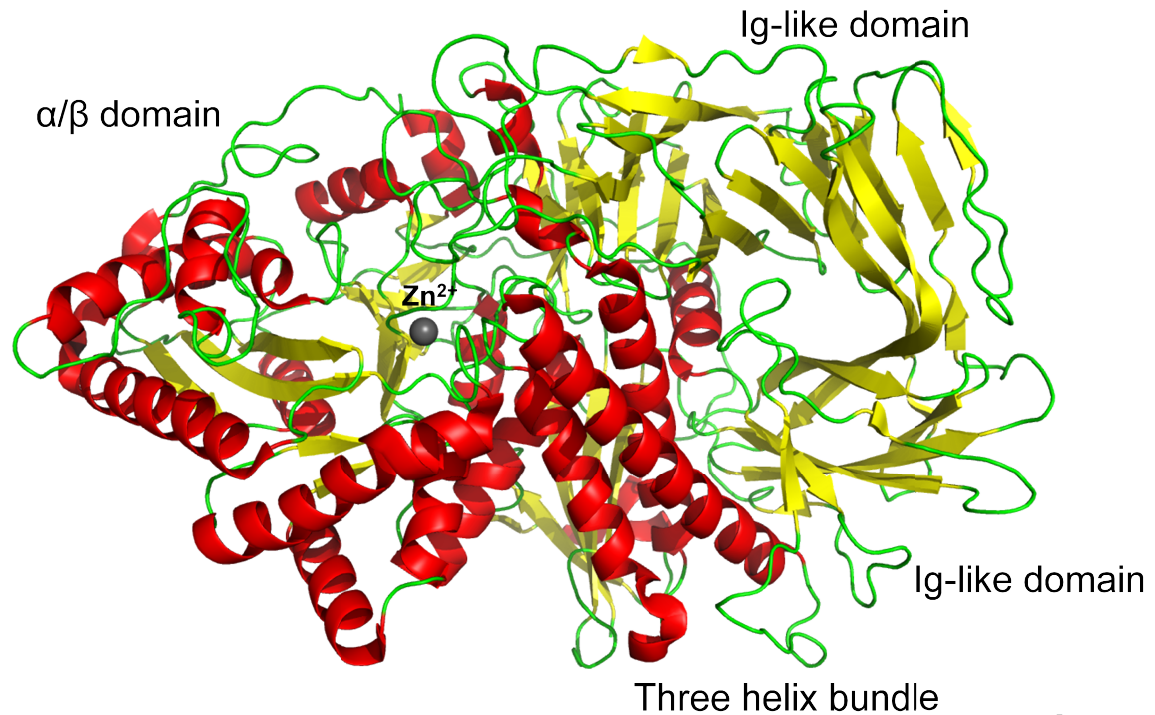
Since it is difficult to obtain human GMII in sufficient quantity and purity, GMII of *Drosophila melanogaster* (dGMII) is used as a model system. The sequence of dGMII possesses 41% identity and 61% similarity compared to human GMII (hGMII). It has been shown that hGMII and dGMII possess comparable kinetic properties, inhibitor sensitivity as well as the same substrate specificity [30],[33].

Jack Bean α mannosidase (JBM), which is a commercially available glycosidase, serves as an additional model enzyme. The complete primary sequence or structural details have not been characterized yet [34]. Via fragment ion spectra of peptides derived from proteolytic digests, 959 amino acids of 981 were determined. LC/MS studies revealed a high sequence similarity in the catalytic site with human lysosomal- α mannosidase (hLM) and hGMII [34], [35].

1.2.1 Structure and characteristics

Human golgi α -mannosidase (hGMII) belongs to the family 38 of the glycoside hydrolases (EC 3.2.1.114) [36], [37]. This enzyme is a type II transmembrane protein, approximately 125 kDa in size and located in the Golgi apparatus. GMII consists of a small N-terminal α/β domain, a three α -helix bundle and an all- β C-terminal domain (see Figure 4). All these components are forming a single compact entity which is linked by five disulfide bonds and stabilized by a zinc-binding site. As a type II transmembrane protein, the N-terminus is located on the cytoplasmic site [38],[39]. The catalytic domain is located on the luminal site at the C-terminus, where it is coupled via a link to an integral membrane anchor [38]. The molecule presents two different faces: the convex N-terminal face and the planar surface on the opposite site of the molecule. The last N- terminal residue Cys31 is placed at the convex face and implying that this surface faces the inner side of the Golgi membrane whereas the planar surface containing the catalytic site points toward the Golgi lumen. The C-terminal β -domain can additionally be subdivided into two immunoglobulin-like units [30]. An extended pocket within the α/β domain is located on the planar site of GMII, which is mainly formed by acidic amino acid residues. These residues form the core of a large and surface exposed area consisting of highly conserved amino acids. The catalytic site occupies a small cavity at the side of this highly conserved and negatively charged region. Of particular importance is a central zinc ion which provides stability and plays an important role in substrate recognition [30].

Contrary to other metalloenzymes, in which the zinc ion is tetra-coordinated, the zinc atom in the unliganded GMII is penta-coordinated (four times by amino acid side chains and once by a solvent molecule). Upon binding of the substrate or an inhibitor, a sixfold coordination takes place whereby the solvent molecule is displaced and the fifth and sixth binding site are occupied by ligand atoms [28].



5

Figure 4: Molecular structure of the Golgi α -mannosidase II of *Drosophila melanogaster* (dGMII) (obtained from X-ray structure data, pdb: 1HTY from van den Elsen et al. [30])

Mechanism of mannose cleavage

When entering the Golgi apparatus, the polysaccharide chain consists of $GlcNAc_2Man_5$, characterized by the three terminal mannose units M2, M4 and M5 (see Figure 3). The enzyme GlcNAc-transferase I transfers a GlcNAc residue (G3) onto the terminal mannose moiety M2. The presence of the G3 unit is a prerequisite for the activity of GMII which cleaves the two remaining terminal mannose units M4 and M5, which are linked to M3 via an $\alpha 1,3$ and an $\alpha 1,6$ -glycosidic bond [28]. The substrate binding site comprises three distinct sites:

1. the catalytic site
2. the holding site
3. the anchor site

In the following 2D-diagramm (see Figure 5) generated with LigPlot+, the interactions between the substrate and dGMII are depicted.

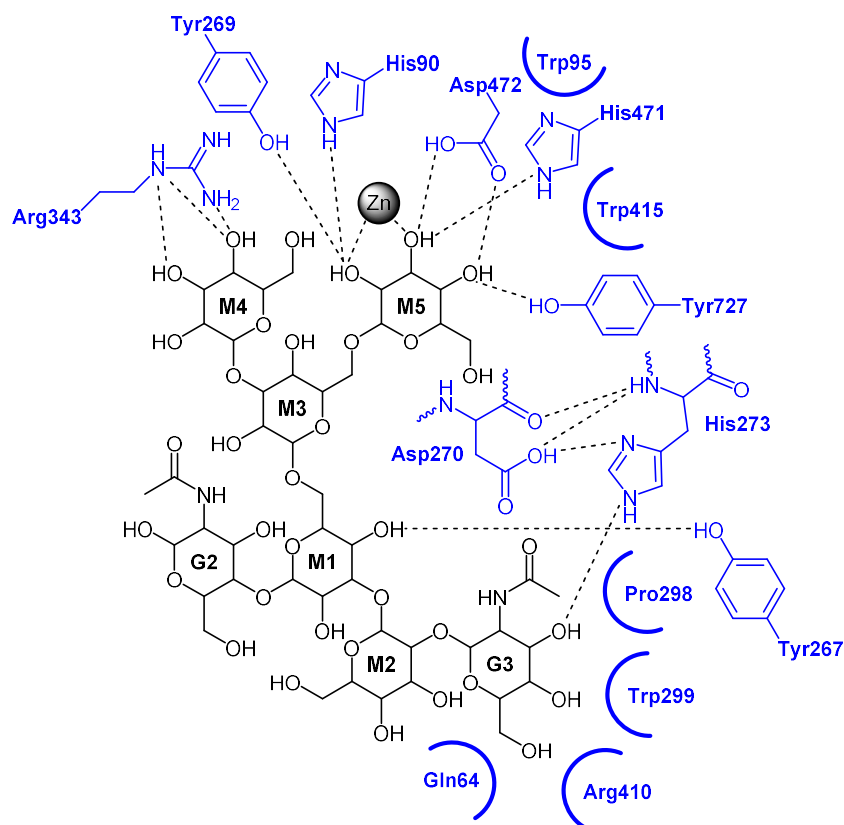


Figure 5: Overview of the interactions between the substrate and dGMII as determined by LigPlot+ (obtained from X-ray structure data, pdb:3CZN, published by Shah N. et al. [32])

1. Catalytic site

The α -1,6-linked mannose (M5) is bound to the catalytic site via hydrogen bonds and π - π interactions. The hydroxyl groups at positions 2 and 4 form hydrogen bonds to two tyrosine residues (Tyr269 and Tyr727). The hydroxyl groups at position 2 and 3 interact with His90 and His471, respectively. The hydroxyl group at position 3 forms a hydrogen bond with Asp472. In addition, π - π interactions occur between the hydrophobic surface of the saccharide ring and a tryptophan as well as a tyrosine residue (Trp95 and Tyr727). The free electrons of the oxygens at positions 2 and 3 coordinate to the zinc atom [32].

2. Holding site

The α -1,3-linked mannose (M4) is bound to the holding site and interacts with Arg343. A distance of 9 Å separates M4 from the side chain of Asp204 acting as a nucleophile, which is indicating that M5 undergoes the catalytic process in first place. Due to the spacious sidechain of Arg343, this site is closely limited in length. No inhibitor binding in this adjacent site has been described yet [32].

3. Anchor site

The *N*-acetylglucosamine (G3) is tightly bound in the so called anchor site and is 13 and 14 Å away from M5 and M4, respectively. G3 forms π - π interactions with a tyrosine residue (Tyr267), its acetyl group is embedded in a hydrophobic region composed of tryptophan (Trp299) and proline (Pro298), and it forms a hydrogen bond with a histidine residue (His273). The presence of this anchor site is crucial considering the high flexibility of the substrate. G3 is responsible for binding and orienting the substrate for the catalytic reaction and thus assuring the stability of this oligosaccharide complex by binding in the anchor site [32]. G3 also ensures the rotation of M4 from the holding site to the catalytic site by keeping the substrate in the active site [28].

The active site of *d*GMII crystallized with its natural branched substrate is depicted in following figure (see Figure 6):

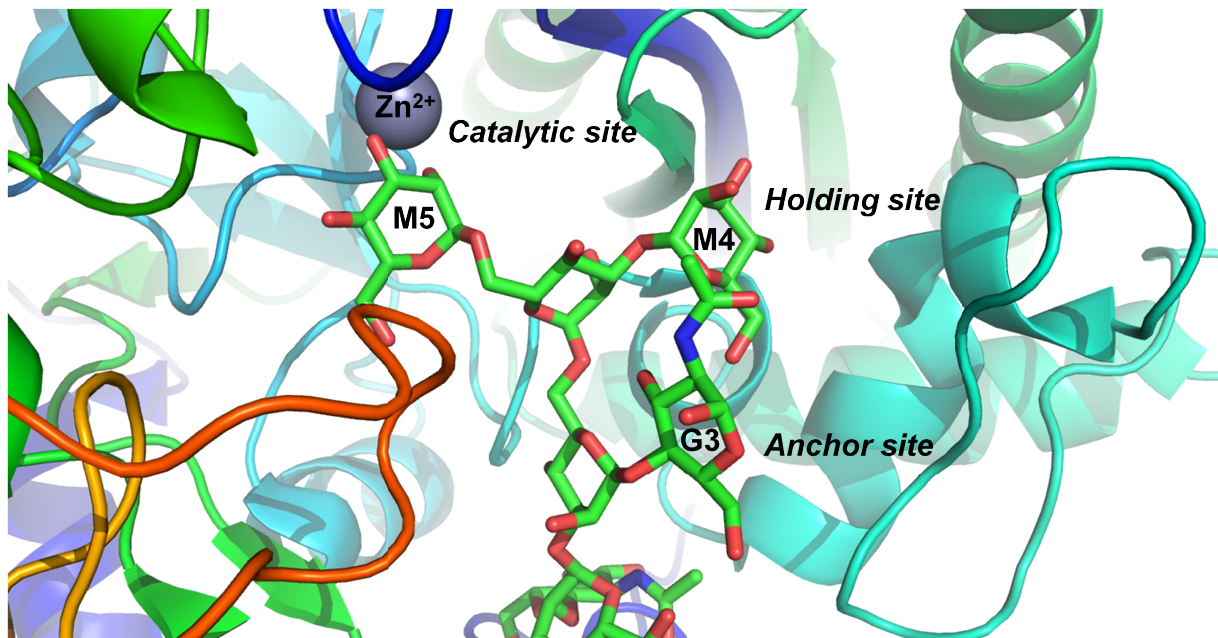


Figure 6: The active site constitutes of three sugar binding sites: the catalytic, the holding and the anchor site. GMII substrate with green carbon atoms bound to *d*GMII is shown. The zinc ion is depicted as gray sphere. (obtained from X-ray structure data, pdb: 3CZN published from Shah N. et al. [32])

In the following figure, the different steps of the hydrolysis cascade of GMII are depicted (see Figure 7).

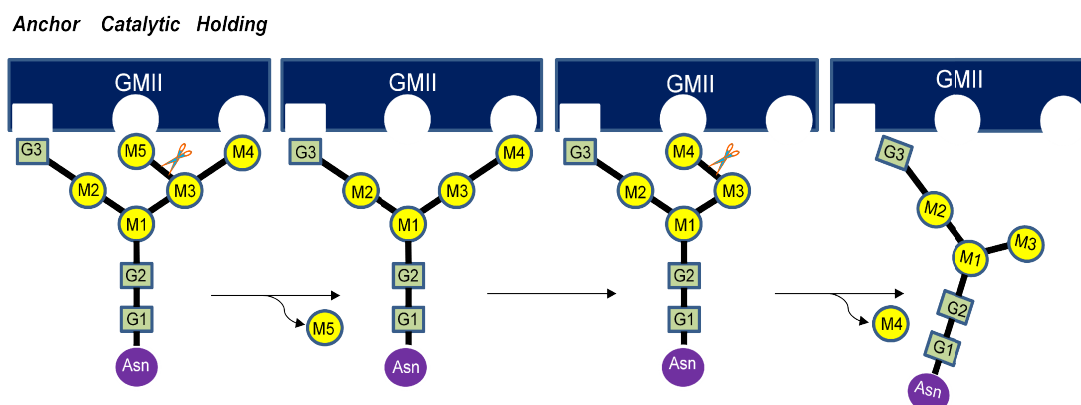


Figure 7: Schematic representation of the GMII catalyzed hydrolysis cascade reproduced from Shah N. et al. [32]

Hydrolysis mechanism

The enzyme belongs to the family of retaining glycoside hydrolases and catalyzes the reaction by a double displacement mechanism resulting in a covalent glycosyl-enzyme complex passing through a transition state with oxacarbenium ion character [40],[41],[31]. The configuration of the anomeric C-atom is preserved during the hydrolysis of the glycoside [40], [42]. It has been proposed that GMII binds the substrate in a distorted conformation and has different forms between ${}^0S_2 \rightarrow B_{2,5} [TS] \rightarrow {}^1S_5$ along the reaction pathway of the glycosylation step [43], [31]. GMII adopts a $B_{2,5}$ type of transition state (TS) [31], [42]. The conformation of the substrate in the covalent intermediate is better defined as wild-type GMII has been co-crystallized with covalent inhibitors (fluorinated mannose analogues) and suggested distorted 1S_5 skew boat conformations [31]. This conformation provides some advantages. In the deglycosylation step, steric hindrance is minimized between the two syn hydrogens at C-3 and C-5 as well as for the nucleophile water. The skew boat conformation forces the hydroxyl groups into a pseudo-equatorial position, thus favouring the formation of a planar transition state by decreasing the energy barrier [31]. Furthermore, the leaving group is positioned anti-periplanar to the non-bonding ion pair of the oxygen atom facilitating the departure of the leaving group [31].

Within the catalytic site, deprotonated Asp204 acts as the nucleophile, while the other conserved Asp341 acts as a general acid/base catalyst (see Figure 8) [44]. The nucleophile attacks the anomeric carbon atom in a S_Ni type reaction. Meanwhile the second aspartate protonates the exocyclic glycosyl oxygen and activates the OR group as a leaving group. This leads to the formation of a covalent complex between the enzyme and the mannose moiety. In a subsequent step, a nucleophilic water molecule is deprotonated by Asp341 and attacks the anomeric C-atom resulting in the release of the mannose residue with net

retention of the initial stereochemistry [45]. The cleaved M5 residue is released and M4 moves from the holding to the catalytic site and is trimmed in the same way as described for M5 (see Figure 7).

Figure 8 shows the hydrolysis mechanism of dGMII involving Asp204 and Asp341:

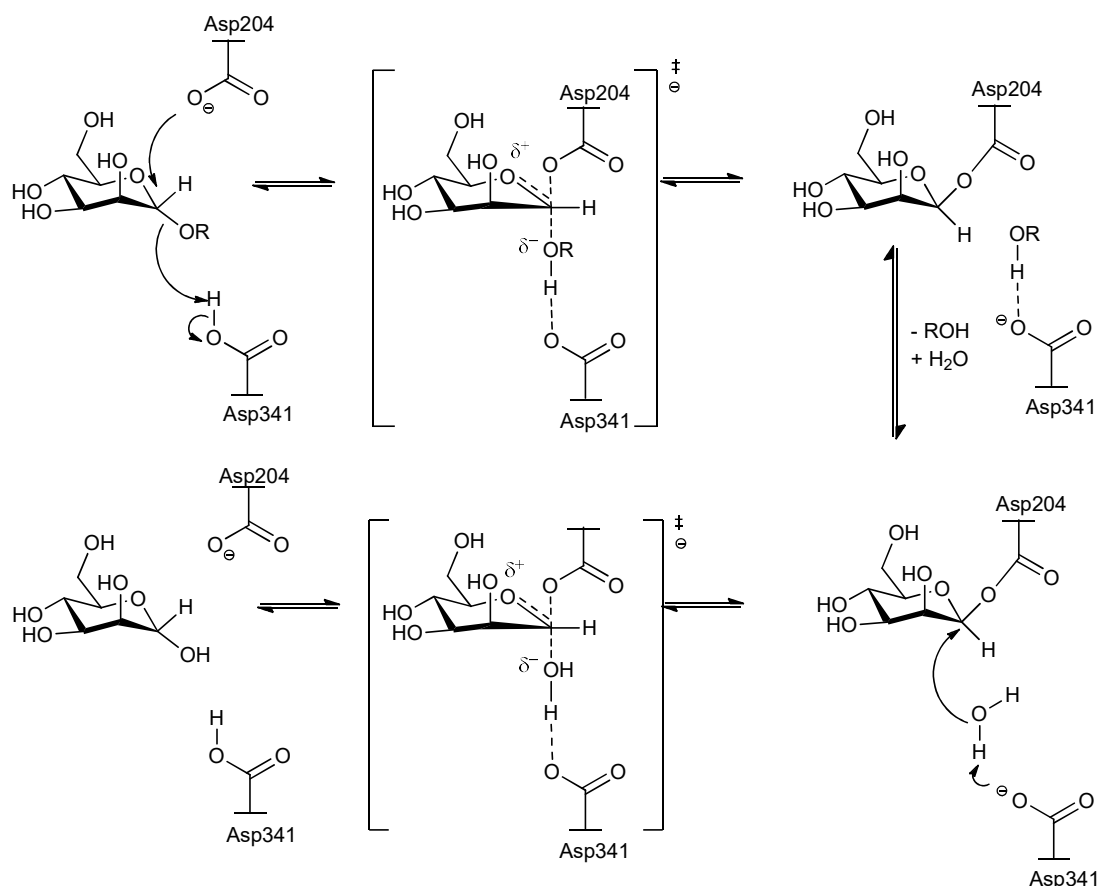


Figure 8: Schematic representation of the hydrolysis mechanism of dGMII involving Asp204 and Asp341 according to Reilly et al. [40] and Koshland D.E. [41]

1.3 β -glucosidase

β -glucosidases (EC 3.2.1.21) belong to the family of glycosyl hydrolases 1 and 3. They split β -glycosidic linkages in disaccharides or glucose-substituted molecules [46]. These enzymes are used in agricultural, biotechnological, industrial and medical applications and plays key roles in biological processes [46]. For example alterations or mutations of β -glucosidase are linked to diseases such as Parkinson's disease, Lewy body dementia and lysosomal storage disease (Gaucher's disease) [47], [48]. The β -glucosidases, belonging to the glycosyl hydrolase family 1 are β -retaining glucosidases and use a glutamate residue as the catalytic nucleophile [21], [49], whereas β -glucosidases belonging to the glycosyl hydrolase family 3,

use aspartate as catalytic nucleophile instead [46]. Known inhibitors are for example the irreversible inhibitor conduritol- β -epoxide [50], [51] and isofagomine [52], [53] (see Figure 9).

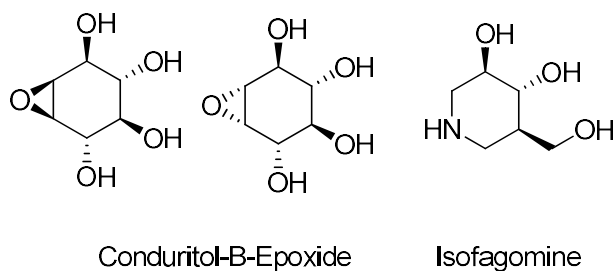


Figure 9: Structures of the Conduritol-B-epoxide and Isofagomine

The major difference between α - and β -glycosidases is the position of the catalytic nucleophile and the catalytic proton donor. The protonation of β -glycosidases happens either anti or syn to the C1-O bond [54]–[57] (see Figure 10).

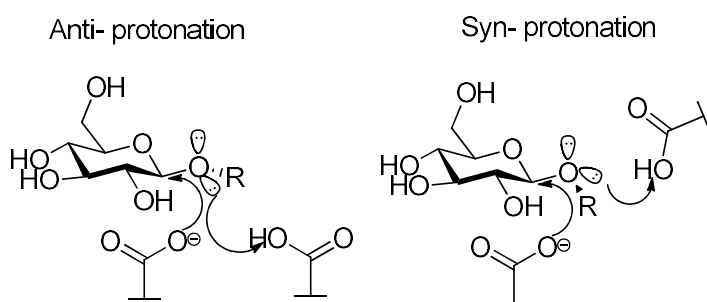


Figure 10: β -Glycosidases protonate their substrate anti or syn to the C1-O bond after Heightman et Vasella [55]

The difference between β - and α -glucosidases in regard to the oxacarbenium ion character in the transition state is illustrated in Figure 11 by comparing the structures of known β - and α -glucosidase inhibitors with the respective transition states. The β -glucosidase inhibitor isofagomine imitates the carbocation transition state, while the α -glucosidase inhibitor deoxynojirimycin is an analogue of the oxacarbenium state. The carbocation transition state seems to be more essential for equatorial hydrolysis (β -glucosidase), while the oxacarbenium transition state is preferred for axial hydrolysis [54].

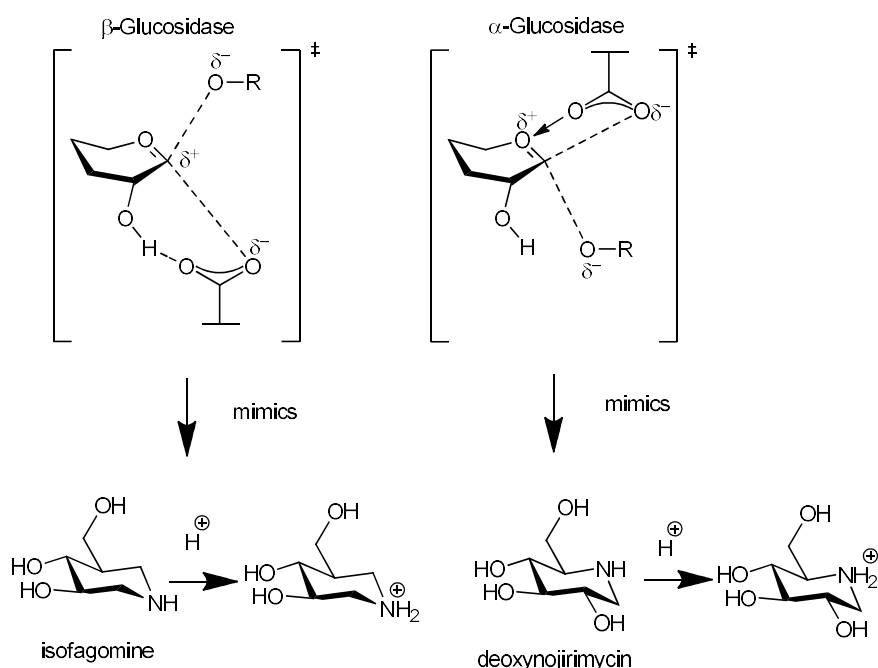


Figure 11: Transition states of α - and β -glucosidase, comparison of mechanism of action of isofagomine and deoxynojirimycin adapted from Borges de Melo et al. [54]

The rationale for enzyme inhibition by mimicking the transition state is the assumption that the transition state is probably responsible for the highest degree of enzymatic stabilization throughout hydrolysis [57]. Conduritol-B-epoxide is a competitive irreversible β -glucosidase inhibitor [58], [51]. The proposed mechanism of action of Conduritol-B-epoxide is shown in Figure 12. The inactivation of the β -glucosidase by Conduritol-B-epoxide presumably takes place as follows: the oxygen of the epoxide is protonated by an acidic group AH and the anomeric C atom undergoes a nucleophilic attack from the carboxylate ion resulting in the formation of an ester.

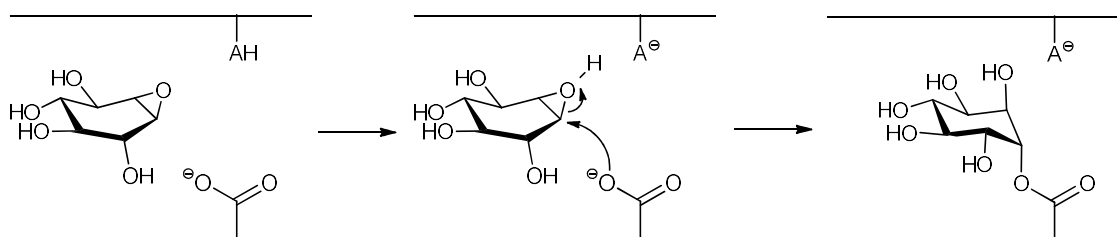


Figure 12: Postulated mechanism of conduritol b epoxides according to Borges de Melo et al. [54]

The β -glucosidase of sweet almonds, which has been used in our studies, has been subject of numerous kinetic studies and inhibition assays (for reversible and irreversible inhibitors) due to its commercial availability resulting from the feasibility to isolate it [59]. The β -glucosidase from sweet almonds is classified in the family 1 of glycosyl hydrolases using a

glutamate as active site residue for the nucleophilic attack with net retention of anomeric configuration [59], [60].

1.4 Jack Bean Mannosidase

The third model enzyme in this work is the commercially available Jack Bean mannosidase from *Canavalia ensiformis* (JBM). Neither the entire primary sequence, nor its 3D structure have yet been characterized [34]. In 2013, Kumar *et al.* [35] succeeded to elucidate 98 % (959 amino acids) of the protein sequence of JBM using "bottom up" proteomic methods. The enzyme is a zinc metalloprotease and consists of two subunits. The larger subunit with a molecular mass of ~ 66,000 Da is glycosylated, whereas the smaller unit with a molecular mass of ~44,000 Da has no glycosylation [61], [62]. JBM cleaves α 1–2-, α 1–3- and α 1–6-linked mannose residues from high mannose oligosaccharides but also has a broad aglycone specificity [34]. JBM is a retaining enzyme and follows the double-displacement mechanism [63]. Furthermore, LC /MS studies from Withers *et al.* [34] have shown that the catalytic site of the JBM has a high sequence similarity with a number of class II Golgi and lysosomal α -mannosidases. An aspartate residue has been determined as the catalytic nucleophile in JBM and therefore it can be classified within family 38 of glycosyl hydrolases [34]. JBM is inhibited by swainsonine and mannostatin, like other mannosidase II enzymes [64]–[66]. This enzyme is more readily available than the mammalian α -mannosidases and is therefore widely used for sugar analysis [63].

1.5 Glycosidase inhibitors - Current State

Glycosidases play a leading role in protein modification and recognition processes. These enzymes have been identified as being related to numerous diseases, which are the result of a lack or a dysfunction of the concerned glycosidase. In the last years, the efforts for the design of glycosidase inhibitors have increased. Currently, only a few glycosidase inhibitors are used as treatment options in the EU and the US (Figure 13). Among these approved glycosidase inhibitors are the antidiabetic drugs miglitol (Glyset®) and acarbose (Glucobay®) used for treating type II diabetes, the antiviral drug oseltamivir (Tamiflu®) against influenza virus or miglustat (Zavesca®) for the treatment of Morbus Gaucher, a lysosomal storage disease [67], [68], [69].

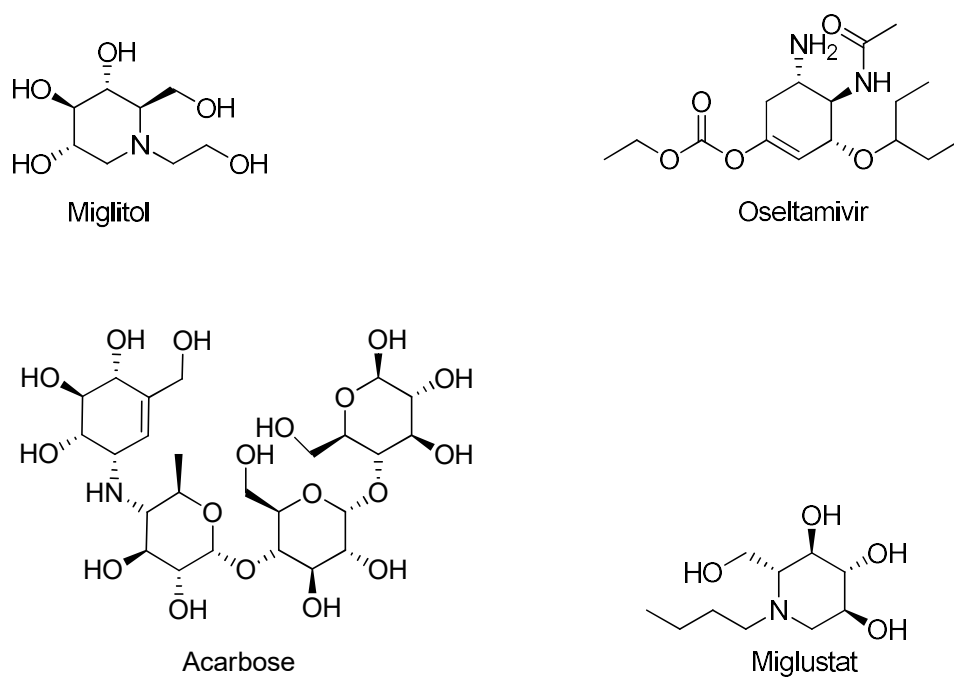
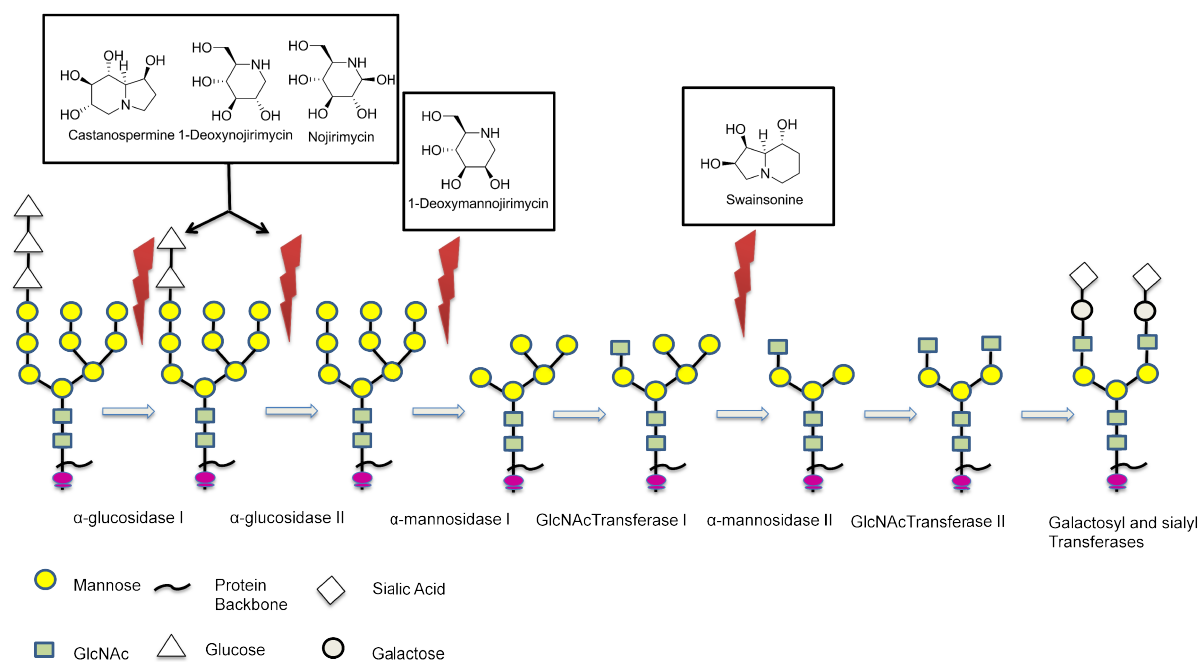


Figure 13: Structures of approved glycosidase inhibitors

Increasing attention has been dedicated to the design of iminosugars because they are protonated at physiological pH and mimic the postulated transition state of glycoside hydrolysis. The biosynthesis of glycoproteins can be specifically inhibited by various iminosugars as shown in Figure 14. It is assumed that suitable inhibitors must be capable to mimic the oxacarbenium-like transition state [44], [55], [70].



8

Figure 14: Inhibitors against the biosynthetic pathway of N-linked oligosaccharides according to Kajimoto and Node [71]

Nojirimycin and 1-deoxynojirimycin (DNJ) (see Figure 14) are inhibitors of α -glucosidases. They are D-glucose analogues and mimic the transition state of the glucosidase-catalyzed reaction [71]. Nojirimycin was one of the first iminosugars ever isolated (1966 from *Streptomyces* sp) [72]. Deoxynojirimycin was first synthesized from Nojirimycin, before it was isolated from mulberry leaves in 1970 [73]. The leaves of the mulberry tree have long been used as a traditional medicine in China due to their antihyperglycemic effect [74]. Miglitol (see Figure 13), which is an oral anti-diabetic drug, is a derivative of deoxynojirimycin [71]. Castanospermine is a 1,6,7,8-tetrahydroxyoctahydroindolizine alkaloid from seeds of *Castanospermum australe* [75]. Castanospermine is a potent inhibitor of lysosomal α - and β -glucosidase without influencing other glycosidases [106],[107]. Swainsonine (Figure 15) is an indolizidine alkaloid and was obtained from *Swainsona canescens*, a toxic legume found in Australia [78]. Swainsonine is an inhibitor of Golgi α -mannosidase II, lysosomal α -mannosidase, and jack bean α -mannosidase [64], [79]–[81]. Elbein *et al.* [65], [82] were the first to show that swainsonine suppressed the formation of glycosylation patterns of the asparagine-linked class of glycoproteins by inhibiting α -1,2-mannosidases. In a report of Tulsiani *et al.* [80], it is described that swainsonine inhibits the α -1,3- and α -1,6-mannosidase activity. Swainsonine exhibits anticancer activities by reducing tumour growth and metastasis [83]. Further investigations revealed that swainsonine has an influence on various cancer types such as lymph cancer [1], colorectal cancer [84], Ehrlich ascites carcinoma [85],

human hepatoma [86] and leukemia [87]. All patients experienced a good therapeutic effect [84].

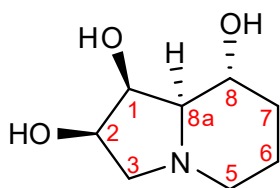


Figure 15: Structure of swainsonine

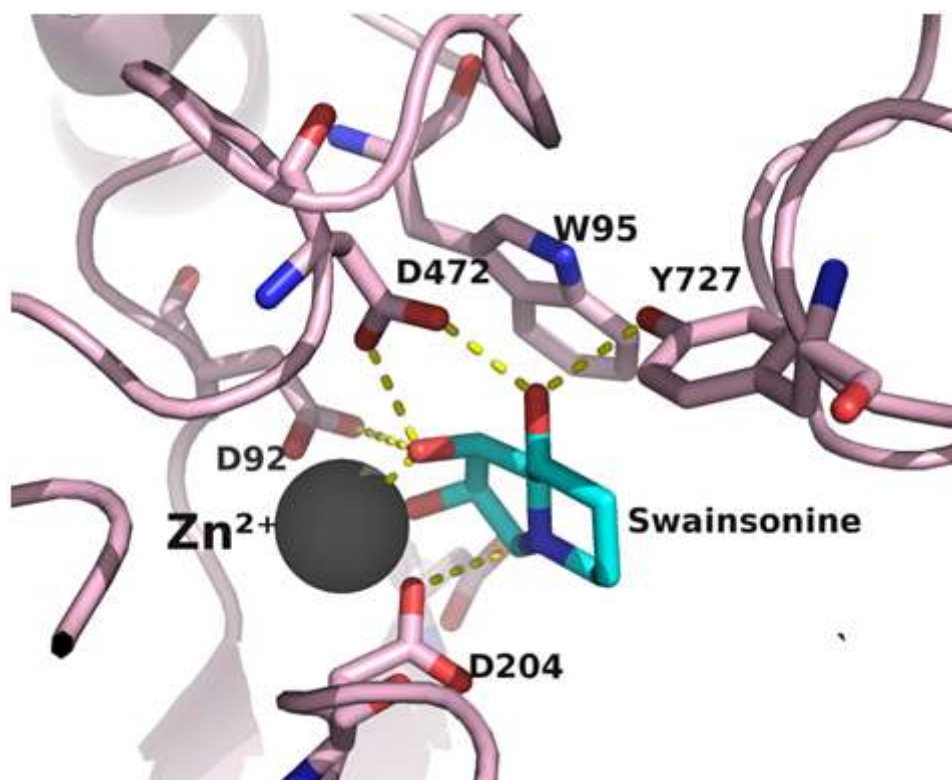


Figure 16: Binding mode of swainsonine to the active site of dGMII (obtained from crystal structure data, pdb: 1HWW)

Crystal structure analysis revealed that when bound to dGMII the five-membered ring of swainsonine occupies a similar position as the hexose ring of the substrate intermediate, i.e. the two hydroxyl groups are positioned in the same manner as the 2- and 3-hydroxyl groups of the covalent intermediate. This enables interactions with D92 and the zinc ion (see Figure 16). The ring nitrogen occupies the same position as the ring oxygen and the anomeric carbon of the covalent sugar intermediate. Furthermore, the positively charged nitrogen interacts with the negatively charged nucleophile D204 via ionic interactions. The hydroxyl groups of the six-membered ring of swainsonine interact via hydrogen bonding with D472 and Y727 which corresponds to the interactions occurring between the 4-hydroxyl group in the substrate intermediate and these amino acids. The remaining parts of swainsonine fit into the hydrophobic pocket of the catalytic site [31].

Swainsonine represents one of the most potent mannosidase II inhibitors (IC_{50} (JBM) = 40 nM; IC_{50} (dGM II) = 40 nM; IC_{50} (hGM II) = 20-50 nM) [29]. Due to its inhibitory effects swainsonine modifies the glycosylation patterns on the cell surface and consequently alters the expression of specific membrane glycoproteins [88]. Moreover, Sun *et al.* [89] propose an additional anticancer effect based on the study with human gastric carcinoma SGC-7901 cells. Swainsonine is able to induce tumour cell apoptosis by inhibiting the expression of mtp53 protein and increasing the expression of wtp53 protein [90],[91]. Furthermore, swainsonine seems to be responsible for a Ca^{2+} -overload, which leads to the activation of the endonuclease to cause apoptosis of the human gastric carcinoma SGC-7901 cells [89]. The presumed anticancer mechanism is depicted in the following figure (see Figure 17):

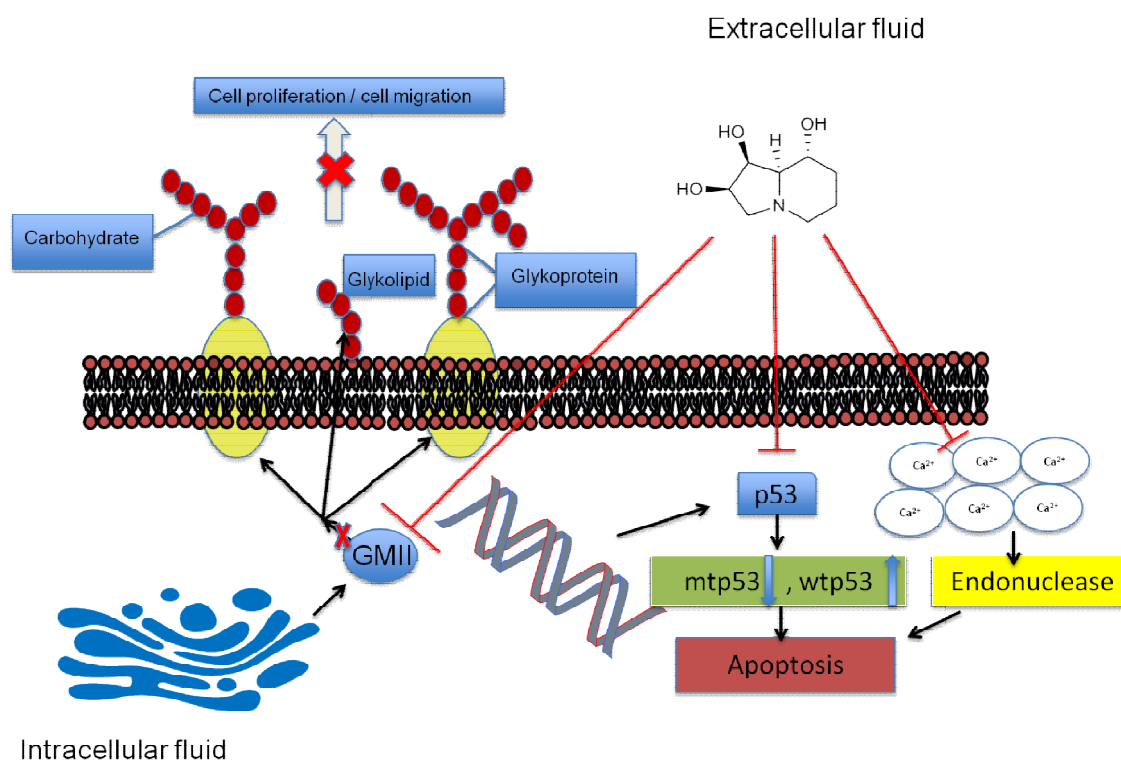


Figure 17: The postulated anticancer mechanisms of swainsonine according to Lu *et al.* [90]

The side effects of swainsonine, including the storage disease mannosidosis resulting from the concomitant inhibition of the lysosomal mannosidase, or the poor cellular bioavailability, have led to the need to search for new α -mannosidase inhibitors with an improved selectivity profile [92],[29].

1.6 Virtual screening and molecular docking

After the identification of a suitable target, a virtual screening approach is commonly used in computer aided drug design to discover new chemical structures which bind to a target of interest. Thus, the goal of a virtual screening is to find active compounds from a set of a large library of a commercially or virtually available compounds. In addition to high throughput screenings (HTS), virtual screenings represent a powerful tool for the identification of lead structures and consequently for the optimization of hits to candidate drugs [93], [94]. Virtual screening approaches can be subdivided into ligand-based and structure-based methodologies (see Figure 18). In the ligand-based method, the structure of a binding ligand or a whole set of binding ligands is known. Based on this information, it is possible to identify other compounds from databases which have similar structures or properties as the known ligand [94]. This identification can be carried out as follows: similarity and substructure searching [95], pharmacophore [96] or 3D shape matching [97]. In the structure-based method, the three-dimensional structure of the target macromolecule must be available. X-ray crystallography or NMR proved to be helpful tools for the determination of these 3D-structures [98]. The Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, <http://rcsb.org>) is a source for freely accessible 3D-structures [99]. Generally, it is much easier to perform docking, if a 3D-structure with a co-crystallized ligand is available. On the one hand, a redocking of the ligand can be carried out in order to test the suitability of the chosen docking setup. On the other hand, structural features that are important for binding can be identified and included in the calculations.

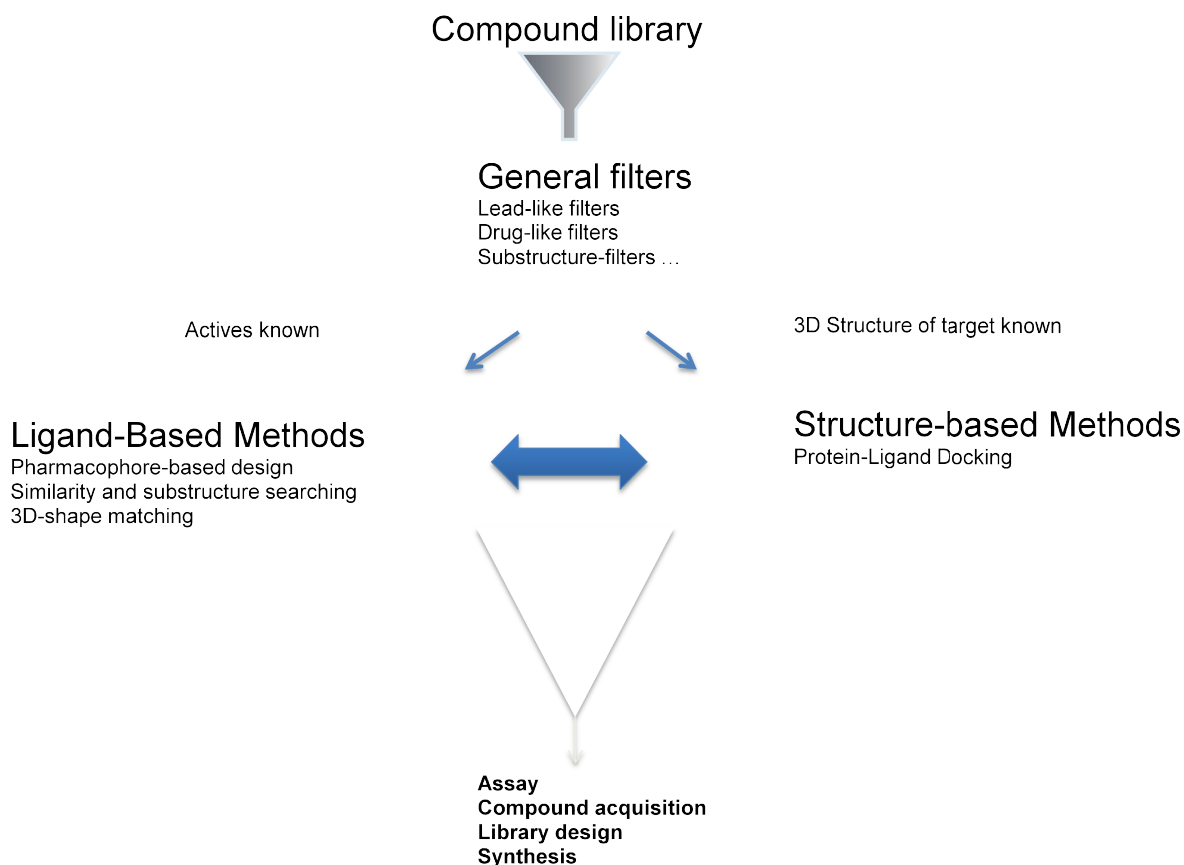


Figure 18 : Overview of common virtual screening approaches adapted from Leach et Gillet [93]

A successful docking run encompasses the correct prediction of the ligand's binding mode (posing) and the generation of a resulting list which is rank-ordered by the expected binding affinities (scoring) [93], [94]. For performing the docking, different docking programs using different algorithms are available. DOCK [100], AutoDock [101], FlexX [102] and Glide [103] are among the most widely used docking engines.

In this work, Dock 3.6 and FlexX were used, which will thus be briefly described here. Structure-based virtual screening was performed. Each of the following process stages will be discussed in this work: target and database preparation, the docking procedure, the analysis of the docking results, the choice of the compounds for experimental testing and the results of the enzyme assays.

1.6.1 DOCK

As depicted in Figure 19, the molecular docking with DOCK comprises the following components: characterization of the receptor site, calculation of the grids, docking and scoring.

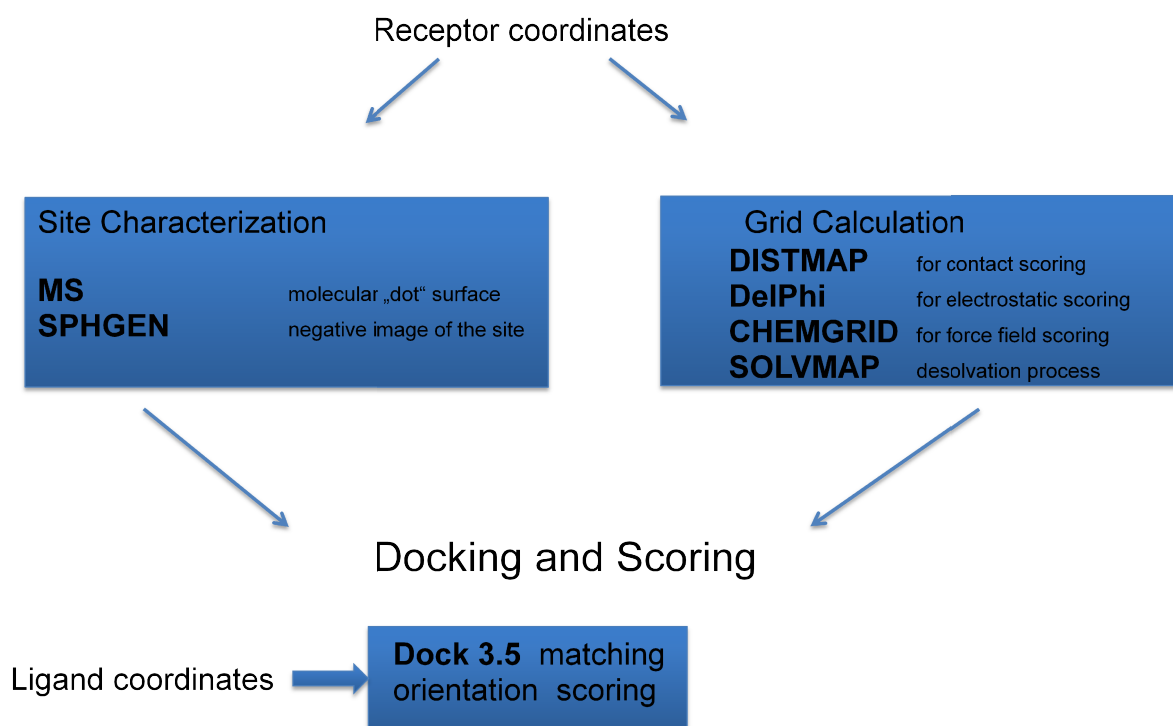


Figure 19: Dock procedure adapted from Meng C. et al. [104]

Site Characterization

A molecular surface is generated using the Connolly MS algorithm [105], [106]. This algorithm calculates the part of the surface, which is accessible to a solvent molecule. The solvent molecule, modeled as a sphere, rolls over the surface of the molecule and creates a three-dimensional contour of the surface [106]. The program SPHGEN creates a "negative" of the binding pocket using the molecular surface calculated with MS. For this purpose, the program fills the binding pocket with spheres of different sizes. Each created sphere touches the surface at two points (see Figure 20).

Grids calculation

The steric and energetic information of each part of the receptor is stored in terms of so-called grid points.

Dock generates the following grids for scoring:

- **DISTMAP** is necessary for determining the contact score. Each grid point is assigned a precalculated “score” based on the distance to receptor atoms. A penalty score happens when the distance between a grid point and a receptor atom is too small. The score for a certain grid point is the number of receptor atoms within the cutoff distance in Ångstroms [107], [108].
- **DelPhi** is used for calculation of the electrostatic potential (ESP) map of the receptor and is based on the Poisson-Boltzmann equation [104].
- **SOLVMAP** generates the grid, which is used for calculating the desolvation cost that has to be paid when a ligand binds to a receptor.
- **Chemgrid** generates grid files required for force field scoring. For that purpose, van der Waals and electrostatic forces are calculated based on the following equation:

$$E = \sum_{i=1}^{lig} \cdot \sum_{j=1}^{rez} \cdot \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + 332.0 \frac{q_i q_j}{D r_{ij}} \right]$$

Equation 1

where A_{ij} and B_{ij} are van der Waals repulsion and attraction parameters, q_i and q_j are the charges of the ligand i and the receptor atoms j , D the dielectric function and r the distance between the atoms i and j . The factor 332.0 converts the energy into kcal/mol.

Scoring

Dock uses a forcefield-based scoring function. In this force field-based scoring function, E_{elec} represents the electrostatic energies, E_{vdw} the energies based on van der Waals interactions, and ΔG_{solv} the desolvation energies, respectively (Equation 2) [109].

$$\Delta G = \sum E_{elec} + \sum E_{vdw} + \sum \Delta G_{solv}$$

Equation 2

Docking

Docking poses are created by matching the ligand atoms with the centers of the spheres so that the distances between the atoms correspond to the distances between the centers of the spheres, within a certain tolerance limit (see Figure 20). A set of atom-sphere pairs is then used to calculate an orientation of the ligand within the binding site by calculation of the translation vector and rotation matrix. After refinement of the orientation by excluding steric clashes, the generated pose is scored using a grid-based approach that predicts interactions between the docked molecules and the receptor at each lattice point [93].

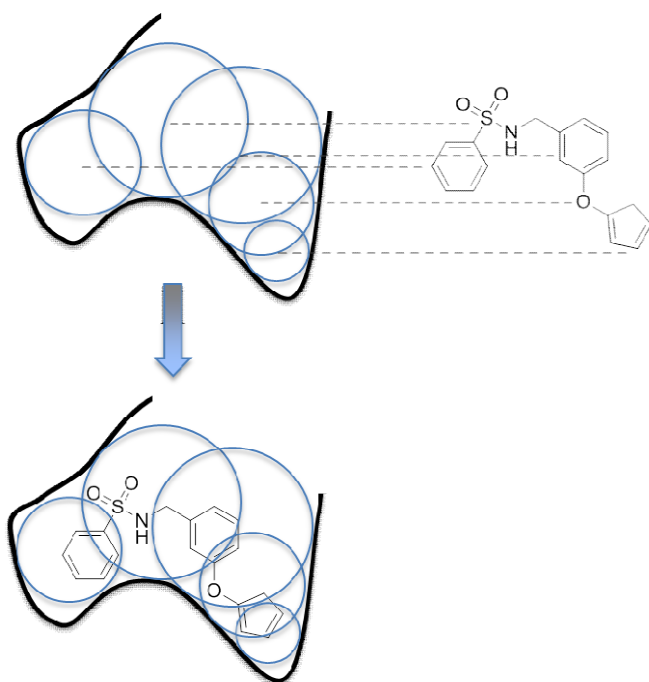


Figure 20: The overlapping receptor spheres create a "negative image" of the binding pocket. Ligand atoms are aligned with the sphere centers in order to evaluate if the ligand fits geometrically to the binding site. Figure adapted from Leach et al. [93]

1.6.2 FlexX

The algorithm of FlexX (BioSolveIT) is based on the initial placement of a "core fragment" and the ensuing flexible adaptation of the remaining groups. Thus, this algorithm is based on an incremental construction strategy and can be subdivided into three stages:

- **Base selection:** selection of the base fragment
- **Base placement:** placement of the base fragment into the site
- **Complex construction:** attachment of other groups to the nascent structure

The base fragments are often rigid parts of the molecule such as ring systems, and not too small in size. In order to predict the correct binding mode, the placement and the choice of the base fragment is of particular importance. If the size of the core fragment and the number of possible interactions with the receptor increase, the probability to postulate the correct binding mode simultaneously increases. However, the larger the size of the base fragment the more conformations are possible and thus the computing time increases. To summarize, the number of possible conformations should be kept low, whereas the number of interactions with the receptor should be maximized. Once the base fragment is chosen, the remaining parts of the ligand are cut at acyclic or rotatable bonds and positioned in various conformations, then completed and the previously separated bonds are rebuilt [93], [102].

The scoring function (Equation 3) used by FlexX is a variant of the simple empirical scoring function of Boehm [110] and can be described as follows:

$$\Delta G_{bind} = \Delta G_{match}F_{match} + \Delta G_{lipo}F_{lipo} + \Delta G_{ambig}F_{ambig} + \Delta G_{clash}F_{clash} + \Delta G_{rot}F_{rot}$$

Equation 3

where ΔGi are coefficients of the functions F_i describing the protein and ligand coordinates. ΔG_{match} is the sum of the individual energy contributions for receptor-ligand interactions, such as hydrogen bonds, metal contacts, and aromatic interactions multiplied by a penalty function. This penalty function takes into account the deviations for the distance of H---O/N hydrogen-bonds from the ideal distance of 1.9 Å and angle deviations of the hydrogen bond $\angle_{N/O-H---O/N}$ from an ideal angle of 180°. This penalty function accepts deviations up to 0.2 Å and 30 ° from the ideal geometry [110], [111]. F_{lipo} and F_{ambig} represent the hydrophobic contacts of the atom pairs of receptor and ligand, whereby F_{lipo} considers the completely nonpolar contact pairs only and F_{ambig} only the pairs with one polar and one nonpolar atom [111]. F_{clash} describes the protein-ligand overlap and F_{rot} the number of rotatable bonds in the ligand.

HYDE Score

The HYDE [112] scoring function is another scoring function applied in the present study and is also implemented in LeadIT (BioSolveIT). The HYDE score (Equation 3) is calculated separately from the intrinsic score of the FlexX program (Equation 3). HYDE (**H**ydrogen bonding and **D**esolvation) estimates the free binding energy that is released upon desolvation of the polar and nonpolar groups of a ligand and by the formation of hydrogen bonds during protein-ligand binding [112]. $\Delta G_{binding}$ is calculated as follows:

$$\Delta G_{binding} = \sum_i \Delta G_{dehydration}^i + \Delta G_{H-bond}^i$$

Equation 4

where $\Delta G_{dehydration}^i$ represents the dehydration free energy contributions per atom type and ΔG_{H-bond}^i the contribution of the atoms i to the formation of hydrogen bonds in the protein-ligand interface. For these calculations, HYDE uses logP-derived parameters [112]. HYDE scoring has the following advantages: it is target-independent and allows an estimation of the reliability of the predicted pose and whether a compound binds with high or low affinity to the target. Furthermore, the individual contributions of the atoms of the ligand are calculated, which gives indications of further potential modification sites of the molecule [112]. In Figure 21, the concept of HYDE is described.

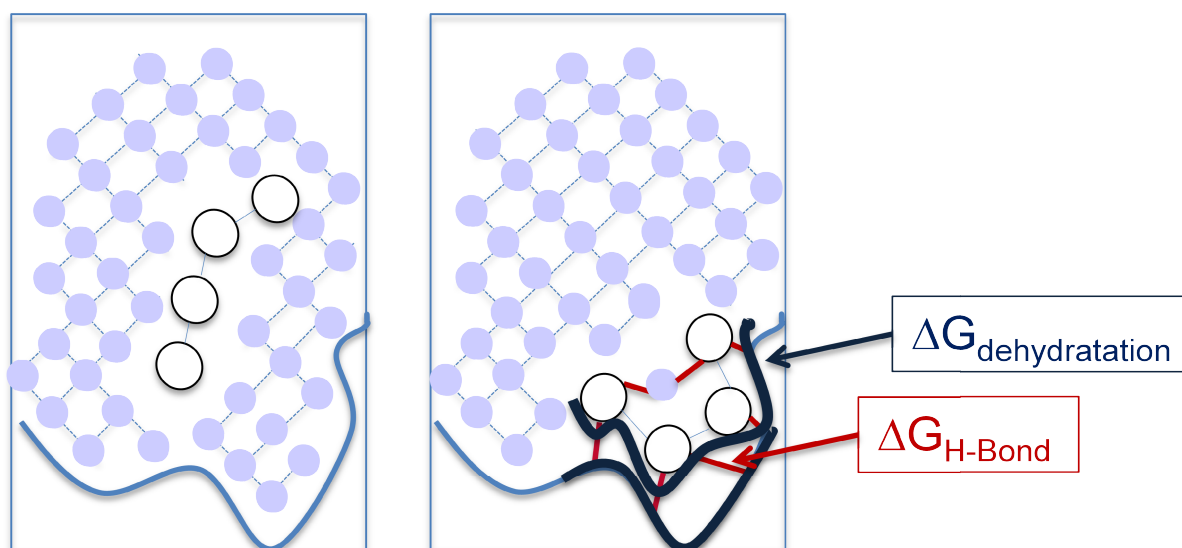


Figure 21: The concept of HYDE according to Reulecke et al. [112]

1.7 Enzyme assays and enzyme kinetics

With the following enzyme assays, both the compounds purchased based on the virtual screening results and individually synthesized potential inhibitors were tested for their inhibitory potency and the type of inhibition. In this chapter, a photometric assay established by [REDACTED] (diploma student, working group [REDACTED], JGU, Mainz) for the α -mannosidase from *Canavalia ensiformis* (Jack Bean) (JBM) is described. Furthermore, fluorometric enzyme assays with dGMII and sweet almond β -glucosidase are described.

Photometric assay

For the photometric assay, the substrate is p-nitrophenyl- α -D-mannopyranoside (pNP-Man). The cleavage of this substrate by the enzyme leads to the release of p-nitrophenol (pNP) (see Figure 22). The reaction is stopped with sodium carbonate and the absorbance A ([OD]) of the deprotonated nitrophenol can be measured at 405 nm.

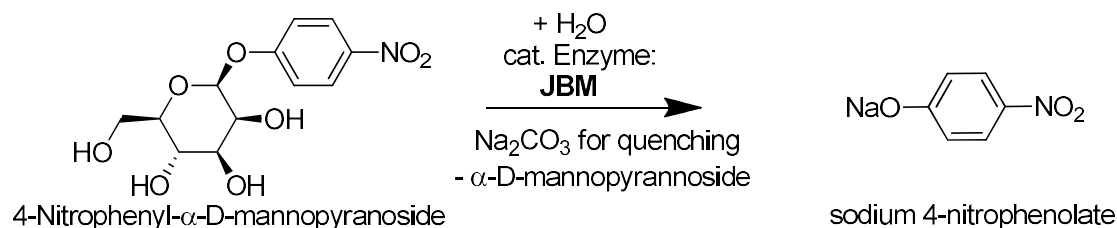


Figure 22: Substrate conversion of pNP-Man by JBM

Fluorometric assay

For the fluorometric assay with dGMII and JBM, 4-methyl-umbelliferyl- α -D-mannopyranoside (4-MU-Man) serves as substrate (see Figure 23). The cleavage of this substrate by the enzyme leads to the release of 4-methylumbelliferone (355 nm excitation and 485 nm emission).

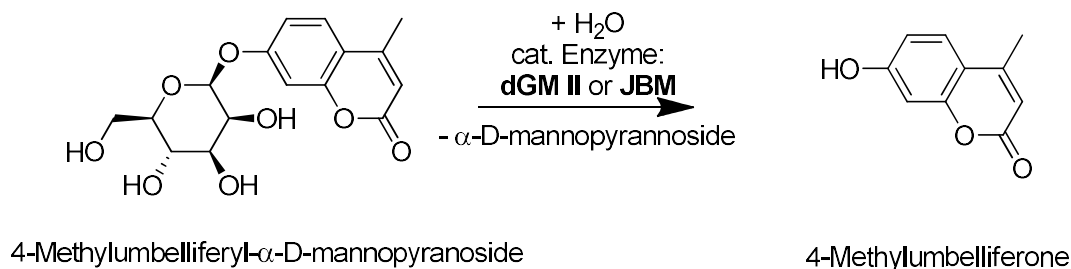


Figure 23: Substrate conversion of 4-Mu-MAN by dGMII and JBM

For the fluorometric assay with sweet almond β -glucosidase, 4-methylumbelliferyl- β -D-glucopyranoside (4-MU-Glc) serves as substrate. The cleavage of this substrate by the enzyme also leads to the release of 4-methylumbelliferone (355 nm excitation and 485 nm emission) (see Figure 24).

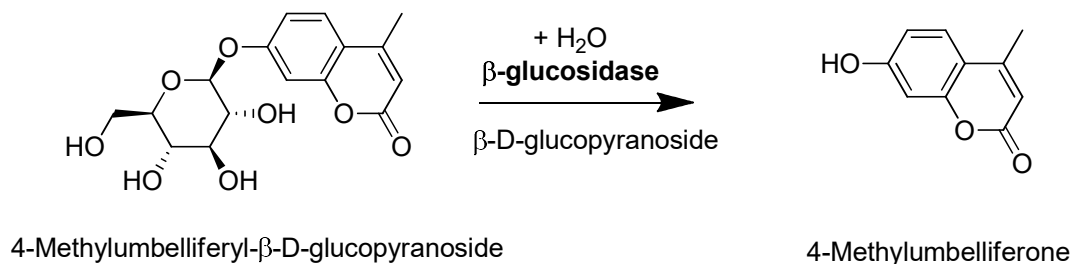


Figure 24: Substrate conversion of 4-MU-Glc by sweet almond β -glucosidase

The α -mannosidase assays were performed in MES (2- (N-morpholino) ethanesulfonic acid) buffer (40 mM, pH = 5.75) and for the β -glucosidase assays in citrate buffer (20 mM, pH = 5.0). First, the assay systems were established with known inhibitors: the highly potent inhibitor swainsonine (purchased from Sigma Aldrich) for JBM and dGMII and isofogamine synthesized by [REDACTED] (PhD student, working group [REDACTED], Johannes Gutenberg University, Mainz) for the β -glucosidase. IC₅₀-value determinations and classical kinetic experiments were carried out.

1.7.1 Enzyme kinetic

The following paragraphs contain theoretical informations about enzyme kinetics. The books of Robert Copeland ("Evaluation of Enzyme Inhibitors in Drug Discovery A Guide for Medicinal Chemists and Pharmacist") [113] and ("Enzyme: A Practical Introduction to Structure, Mechanism, and Data Analysis") [114] were used as primary literature source throughout this section.

The enzyme kinetics describes the reaction rates of chemical reactions catalyzed by the enzyme (see Figure 25). Based on studies of the enzyme kinetics, the nature of the inhibitor can be assessed. Hence, it can be evaluated whether the inhibitor is a tight binder, a competitive inhibitor, a non-competitive inhibitor, etc. Brown (1902) described the enzyme-catalyzed reaction by the following scheme (see Figure 25):

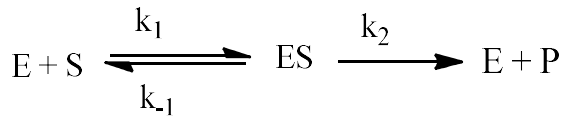


Figure 25: Enzyme catalyzed reaction- E: enzyme; S: Substrate; ES: Enzyme-substrate complex, P: Product; k_n : rate constants

Michaelis and Menten (1913) defined the rate of an enzymatic reaction, which led to the following equation:

$$v = \frac{v_{\max} [S]}{K_M + [S]}$$

Equation 5: v_{\max} : maximum velocity at complete saturation of the enzyme, $[S]$ substrate concentration, K_M : Michaelis Menten constant

The prerequisite for the validity of the equation is the assumption of a steady-state equilibrium, which means that the rate of formation of the ES-complex equals the rate of decay to E + S ($k_1 = k_{-1}$). The reverse reaction $E+P \rightarrow ES$ at the beginning of the reaction is negligible, since the amount of product P is close to 0. Thus, the model of Michaelis and Menten presumes that $k_2 \ll k_{-1}$. K_M , referred to as the Michaelis constant, is the substrate concentration at which the reaction rate v_{\max} is half-maximal. K_M is thus a measure of the affinity of the enzyme to the substrate. The lower the K_M -value the higher is the affinity.

To graph the Michaelis-Menten equation, the velocity is plotted against the substrate concentration (see Figure 26):

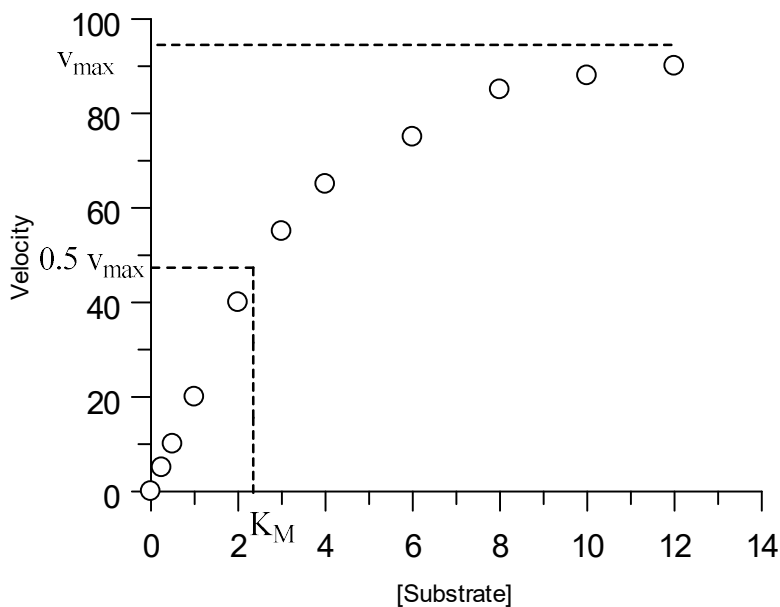


Figure 26: Michaelis-Menten saturation curve

Fitting the initial velocity as a function of substrate concentration results in a hyperbolic curve. The kinetic constants K_M and v_{\max} can be determined as illustrated (see Figure 26).

Lineweaver-Burk plot

The Michaelis-Menten equation can be linearized by plotting the reciprocal values of velocity ($1/v$) against the reciprocal substrate concentration ($1/[S]$) resulting in the so-called Lineweaver-Burk plot. K_M and v_{max} can then be determined from the slope and the y-intercept of the linear fit, respectively (see Figure 27).

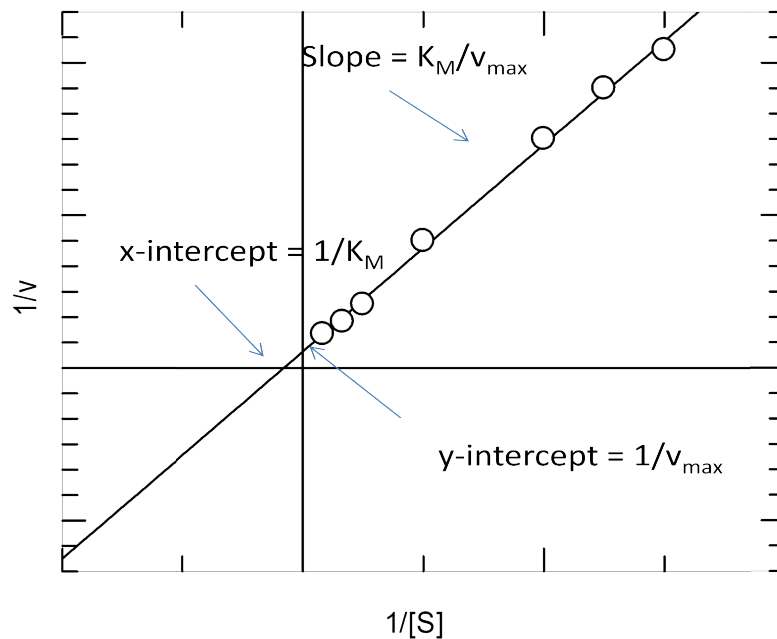


Figure 27: Lineweaver-Burk plot

1.7.2 Enzyme Inhibition

Generally, an inhibitor is able to bind through reversible and/or irreversible interactions to the target. In the case of irreversible inhibition, a covalent bond is formed between the target and the inhibitor. In contrast, most inhibitors with a reversible binding mode, interact with the target through noncovalent forces like electrostatic forces, hydrogen bonding, hydrophobic or van der Waals forces. Reversible inhibition in turn can be subdivided into three different inhibition types: competitive, uncompetitive and non-competitive inhibition.

1.7.2.1 Competitive inhibition

In case of a competitive inhibition, the substrate and the inhibitor compete for a common binding site. An enzyme-inhibitor complex is formed in addition to the formation of the enzyme-substrate complex (see Figure 28). At high substrate concentrations, the substrate

displaces the inhibitor from the common binding site. In the presence of an inhibitor, more substrate is needed to reach the half maximal velocity. As a result, v_{\max} remains unchanged, while K_M apparently increases. The binding affinity of competitive inhibitors is indicated by the inhibition constant K_i . The velocity of binding of a competitive inhibitor is described by the following equation (Equation 6):

$$v = \frac{v_{\max} [S]}{K_M \left(1 + \frac{[I]}{K_i} \right) + [S]}$$

Equation 6: v_{\max} : maximum velocity at complete saturation of the enzyme, $[S]$: substrate concentration, $[I]$: inhibitor concentration, K_M : Michaelis Menten constant, K_i : inhibition constant

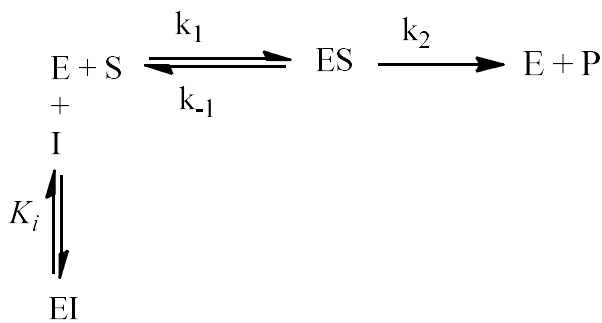


Figure 28: Schematic representation of the competitive inhibition, adapted from Ramsay R. et Tipton K. [115]

In the case of a competitive inhibition mechanism, the intersecting lines from the double reciprocal plots converge at the y-axis as depicted in following figure (see Figure 29):

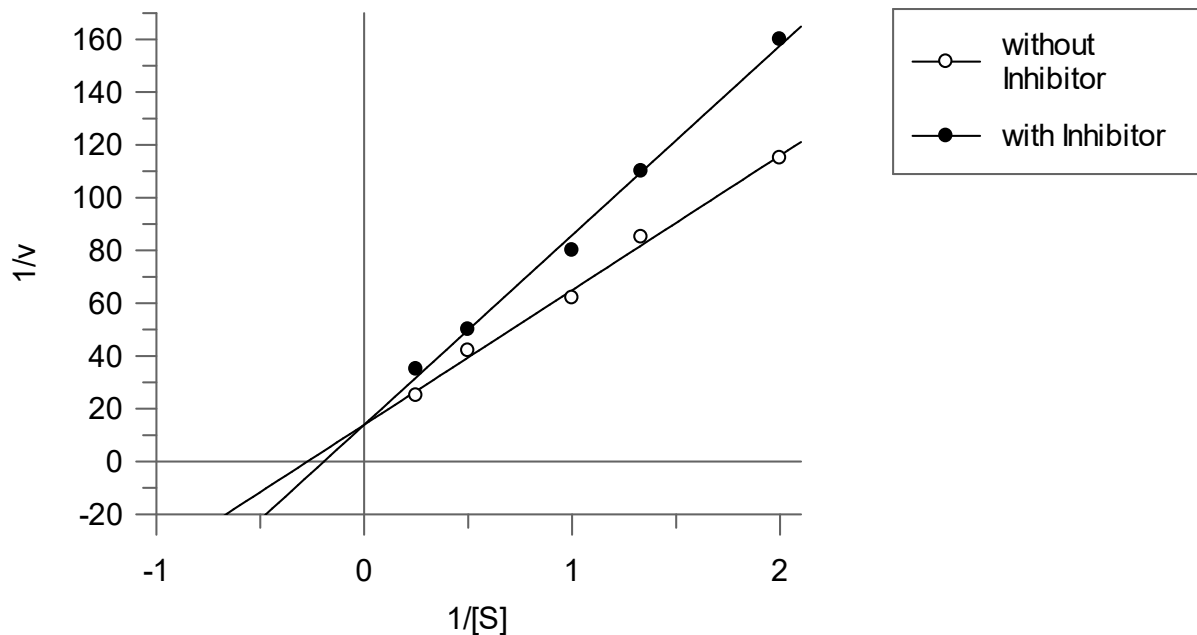


Figure 29: Lineweaver -Burk plot of competitive inhibition

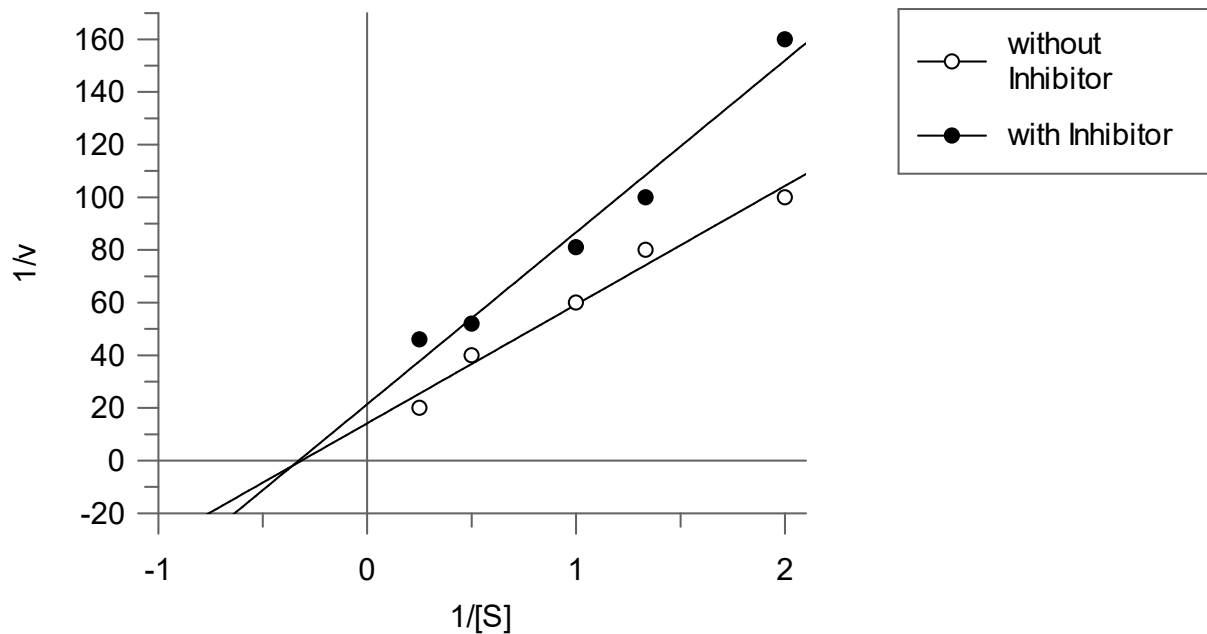


Figure 33: Lineweaver -Burk plot of non-competitive inhibition

If the inhibitor preferentially binds to the free enzyme ($K_i > K'_i$), the value of K_M increases with increasing inhibitor concentration. In this case, the intersecting lines of the double reciprocal plot (Lineweaver-Burk) converge to the left of the y-axis and below the x-axis.

If on the other hand the inhibitor has a higher affinity to the ES-complex ($K'_i < K_i$), the value of K_M decreases with increasing inhibitor concentration. The intersecting lines of the double reciprocal plot (Lineweaver-Burk) merge to the left of the y-axis and above the x-axis.

1.7.2.4 Slow binder

Slow binding inhibitors are time-dependent inhibitors. These inhibitors bind slowly to the enzyme and cause a change in the turnover rate. The turnover-time curves are not linear over the time period. Two different slopes, v_i (initial slope) and v_s (steady-state slope), can be determined (see Figure 34). At the beginning of the reaction, there is a linear relationship between the product formation and the time. This initial velocity turns then into a lower velocity (steady-state velocity). Due to the slow onset of these inhibitors, a difference in the inhibition should be observed before and after preincubation of the enzyme with the inhibitor.

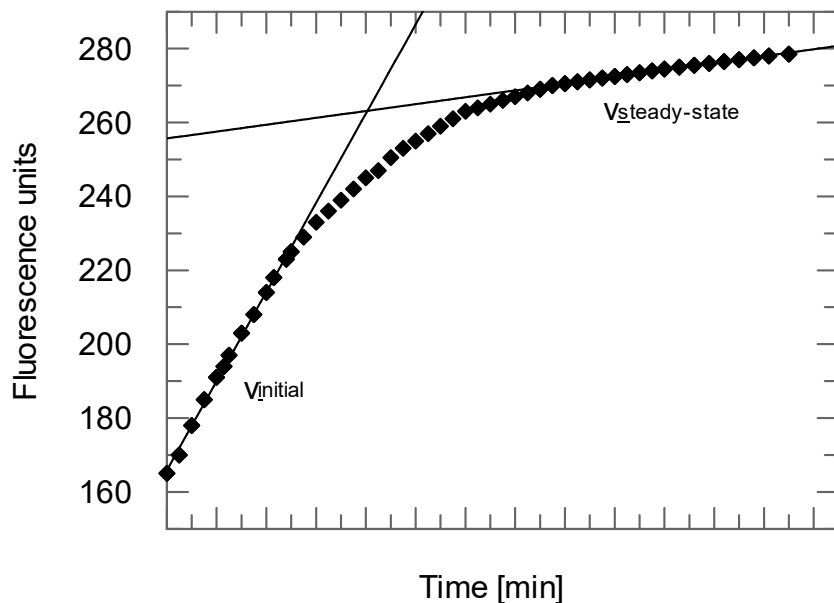
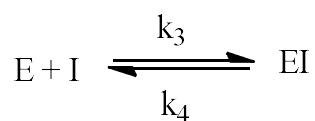


Figure 34: Progress curve of a slow binding inhibitor.

There are four different modes of time-dependent inhibition. In this chapter, two of them are described in more detail.

In a simple single-step mechanism (Figure 35, Scheme A) the association and dissociation rate constants k_3 and k_4 are very small compared to the ones for simple reversible inhibitors. This means that the inhibition slowly progresses. The K_i -value is calculated according to Equation 9.

(A)



(B)

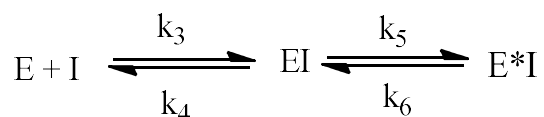


Figure 35: (A) simple reversible slow binding (B) enzyme isomerization

$$K_i = \frac{k_4}{k_3} = \frac{[E][I]}{[EI]}$$

Equation 9: K_i : inhibition constant, $[E]$: enzyme concentration, $[I]$: inhibitor concentration, $[EI]$: enzyme-inhibitor complex concentration, k_n : rate constant

In the above figure (Figure 35, Scheme B) a two-step binding mechanism is displayed, where the enzyme and the inhibitor form an EI-complex first and then the high-affinity, isomerized enzyme-inhibitor complex E^*I is built. K_i is the dissociation constant of the EI complex, which is formed first. The dissociation constant of the E^*I -complex is K_i^* (Equation 10).

$$K_i^* = \frac{K_i k_6}{k_5 + k_6} = \frac{[E][I]}{[EI] + [E^*I]}$$

Equation 10: K_i : inhibition constant, $[E]$: enzyme concentration, $[I]$: inhibitor concentration, $[EI]$: enzyme-inhibitor complex concentration, $[E^*I]$: isomerized enzyme-inhibitor complex concentration, k_n : rate constants

1.7.2.5 IC_{50} determination

The IC_{50} -value represents the concentration of an inhibitor at which the enzyme reaction is half-maximal inhibited. This value is often used to describe the inhibitory effect. For reversible inhibitors, the IC_{50} -value is dependent from the type of inhibition. In the following table, the equations used to determine the IC_{50} -values are listed according to the type of inhibition:

Table 1: Comparison of IC_{50} values depending on the type of inhibition [115]

Inhibitor type	IC_{50}
Competitive	$\left(\frac{[S]}{K_M} + 1\right) \cdot K_i$
	Cheng-Prusoff equation [116]
Uncompetitive	$\left(\frac{K_M}{[S]} + 1\right) \cdot K_i$
Non-competitive	K_i

In the case of a non competitive inhibitor, the IC_{50} is equal to the K_i and not dependent from the substrate and enzyme concentration.

For the IC_{50} -value determination, the fluorescence was measured over a period of 15 minutes in the present study, and the turnover rate of the enzymatic reaction was plotted

against the various inhibitor concentrations. However, a distinction is made between the initial slopes (v_i) and the slopes in the steady-state (v_s) state. A non-linear fit of the measured data was performed according to Equation 15 (see section 3.7.2.4). An IC_{50} value is obtained when plotting the v_i -values against the inhibitor concentration. Using the Cheng-Prusoff equation (Equation 20, see section 3.7.2.4) the K_i value for the different inhibitor types can be calculated. For the non-linear regression of the v_s -values against the inhibitor concentration, an IC_{50} can be obtained and converted into the K_i^* value.

1.7.3 The Z-Factor and the Z'-Factor

As part of the virtual screening performed a large number of substances are tested. By the determination of the Z-factor it is possible to evaluate the reliability of the assay. The Z-factor is a dimensionless, simple statistical parameter reported by Zhang *et al.* [117] to assess the quality of high-throughput screening (HTS) assays. The Z-factor is a powerful tool to compare different assays. The aim of a high-throughput screening is to find lead compounds from large molecule libraries that are for example enzyme inhibitors. In order to identify promising inhibitors, it is important that the assays are well designed for the respective target. The lower the variations in the measurements are, the higher is the reliability of the test system and thus, an identified hit can be considered as true binder. For the validation of a HTS assay, unknown samples and control samples are tested. Negative controls (usually the background) shall correspond to the minimum signal and positive controls to the maximum signal. The resulting data of the positive and negative control can be used to calculate their means and the standard deviations. The dynamic range of the assay signal is defined as the difference between the mean of positive controls and the mean of the negative controls. Because a great number of substances in the tested library display very low or no pharmacological activity, the means and the standard deviations of most of the sample data should be in the same range as the positive control for inhibition assays. The means of the signals are denoted as μ and standard deviations as σ , whereby $\mu(s)$ and $\mu(c)$ are defined as mean values of the sample and control data, respectively, and $\sigma(s)$ and $\sigma(c)$ as the corresponding standard deviations. The variation band (see Figure 36) for a signal is defined as the mean $\pm 3\sigma$, which corresponds to a confidence limit of 99.73 %. The separation band between the sample data and the corresponding control data is defined as

$$| \mu(s) - \mu(c) | - (3\sigma(s) + 3\sigma(c))$$

where $|\mu(s) - \mu(c)|$ represents the dynamic range and is the absolute value of the difference of the mean of the sample data and the control data. The separation band is the signal window, where hits can be identified.

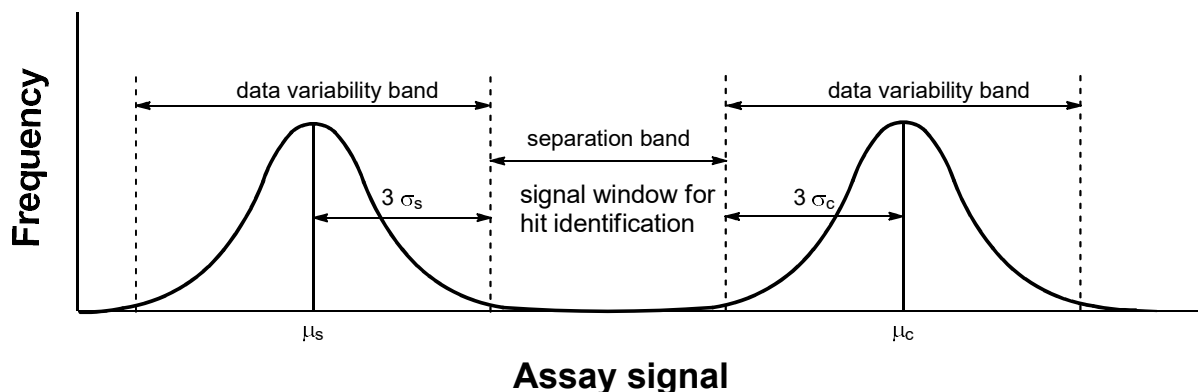


Figure 36: Representation of the variation bands and the separation band in a HTS assay based on Zhang et al. [117]

The statistical measure is Z and is defined as follows (Equation 11):

$$Z = \frac{|\mu_s - \mu_c| - (3\sigma_s + 3\sigma_c)}{|\mu_s - \mu_c|} = 1 - \frac{(3\sigma_s + 3\sigma_c)}{|\mu_s - \mu_c|}$$

Equation 11: $\mu(s)$: mean of the sample data values, $\mu(c)$ mean of the control data values, $\sigma(s)$ standard deviation of the sample data values and $\sigma(c)$ standard deviation of the control data values

The Z-factor can be any value less than or equal to one ($-\infty < Z \leq 1$). Only the range of $-1 < Z \leq 1$ is of interest for any test system since $Z = -1$ is the lower detection limit of the assay system. The higher the Z-factor is the higher is the quality of the assay. This factor gives a statement about the quality of the assay. In the following table, the assay quality is compared depending on their Z-factors.

Table 2: Interpretation of the quality of assays by the value of the Z-factor

Z-factor value	Structure of Assay	Assay quality
1	SD = 0 , or dynamic range $\rightarrow \infty$	ideal assay
$1 > Z \geq 0,5$	Separation band is large	excellent assay
$0,5 > Z > 0$	Separation band is small	double assay
0	No separation band, the sample signal variation and control signal variation bands touch	"yes/no type assay"
< 0	No separation band, the sample signal variation and control signal variation bands overlap	screening is impossible

Figure 37 illustrates two types of assays, **A** and **B**. Figure 37 **A** shows an assay with a Z factor of 0.5 and Figure 37 **B** displays an assay with a Z factor of 0.1. It is evident that the assay of Figure 37 **A** is more suitable for screening due to the fact that the data varies less.

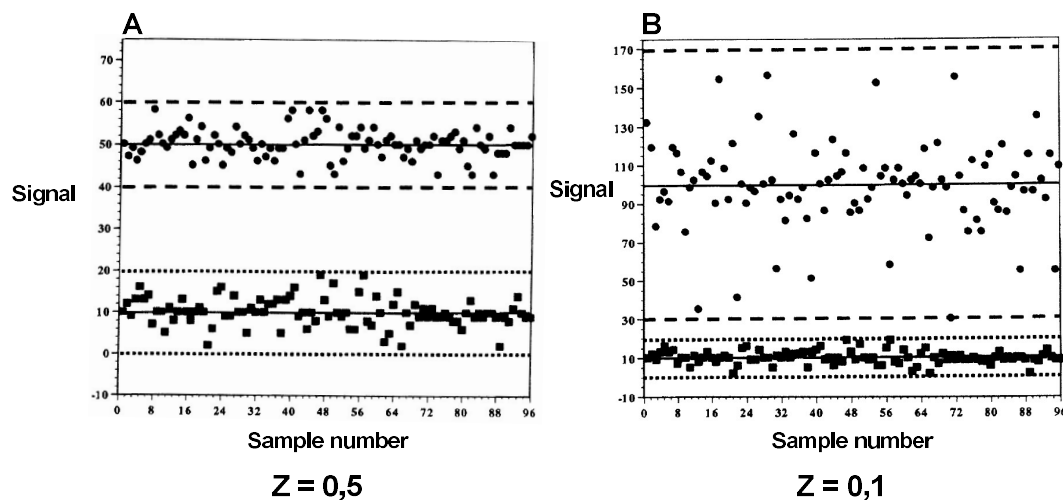


Figure 37: General example of an assay performed in two different formats (A and B). The solid lines represent the mean of the sample signal and the means of the negative control data, the dashed lines are each three σ away from the solid lines according to Zhang et al. [117]

The Z-factor is of particular importance to determine the quality of any HTS assay. To assess the quality of an assay itself, a Z'-factor can be calculated. The Z'-factor is a characteristic parameter used to determine the overall quality of an assay and to identify the interference of test samples. The Z'-factor is calculated from the control data (positive control (c_+) and negative control (c_-)) and is defined as follows (Equation 12):

$$Z' = 1 - \frac{(3\sigma_{c_+} + 3\sigma_{c_-})}{|\mu_{c_+} - \mu_{c_-}|}$$

Equation 12: μ_{c_+} : mean of the positive control data values, μ_{c_-} : mean of the negative control data values, σ_{c_+} standard deviation of the positive control data values and σ_{c_-} : standard deviation of the negative control data values

In this work, the Z'-factor of the fluorescence assay was determined by testing positive controls containing the reaction buffer, dGMII and the substrate 4-methylumbelliferyl- α -D-mannopyranoside and negative controls additionally containing the known inhibitor swainsonine (50 nM per well). The Z'-factor is calculated according to the above equation from the determined means and standard deviations. The data variability and the reproducibility of the assay should at best be determined on a plate itself (intra-plate), from plate-to-plate (inter-plate) and from day to day [118].

2 Objectives

The main focus of this work was the discovery of new inhibitors against GMII. GMII plays a crucial role in the mammalian *N*-glycosylation pathway. A hallmark for cancer cells are altered glycoproteins (especially *N*-glycans) on the cell surface [119]–[121]. The greatest successes in suppressing abnormal glycosylation patterns are observed by inhibition of GMII [122], [123]. Over the past two decades, only few novel ligands addressing the catalytic site have been identified. Few progress has been made in the last years in the development of irreversible [124] and reversible inhibitors. Known irreversible inhibitors are cyclophellitol (1991) [125] or 5-fluoro-(1996) [126] and 2-deoxy-2-fluoro-glycosides (1987) [124], [127]. In a recent study from Overkleeft *et al.* [128], a new class of irreversible glycosidase inhibitors, cyclophellitol cyclosulfates, was introduced. However, since the discovery of the known reversible inhibitors of GMII, swainsonine (1982) [80] and mannosatin (1989) [129], [130] no significant advances have been made with regard to this inhibitor class. In order to choose an expedient approach to initiate the study described in this work, methods from the rational drug design were used. Rational drug design provides an important basis for the development and optimization of new drugs. Virtual screening studies were performed with the overall goal to find new scaffolds for active-site directed inhibitors as well as possible allosteric binding inhibitors. For the efficient screening of potential inhibitors identified by the virtual screenings, a fluorometric assay was established. For this purpose, the activity of dGMII was assayed using 4-MU-Man as substrate. IC₅₀ values and the mode of inhibition of virtual screening hits were also determined. In addition, fluorometric assays for JBM and β-glucosidase from sweet almonds were also established in order to investigate the selectivity of the potential inhibitors.

3 Material and Methods

3.1 Homology modeling

In the last years, considerable progress has been made in X-ray crystallography and high-field NMR spectroscopy in solving protein structures. However, the 3D-structures of many therapeutically relevant targets are still lacking. Since these 3D-structures are often very important in the early stage of drug discovery, homology modeling is an useful tool to obtain an approximation of a 3D-structure [131]. In this work, homology models were built within the software MOE (Molecular Operating Environment). First the target protein amino acid sequence was searched (fasta (.fst) sequence). In the next step, a template protein with a high sequence identity to the target protein is required. For this purpose, the **Search PDB** application in MOE was used to search suitable template protein chains within the PDB database. Using the **Align** application in MOE, the sequence of the proposed templates and the sequence of the target were aligned and metrics like the similarity, the identity and the RMSD value were reported as a result of the alignment. Once the template with a high sequence identity was chosen, the homology model was built using **MOE-Homology** [132], [133]. The used force field for the generation of the homology models was Amber12EHT [134], [135]. Afterward, the homology model was assessed using the **Protein Geometry** application. A Ramachandran plot was generated in order to check whether the homology model contains outlier in terms of backbone dihedral angles, which can be then corrected via an energy minimization in order to obtain more typical phi and psi angles [132], [133]. In Figure 38 below, the major steps of the homology modeling process performed in this work are described:

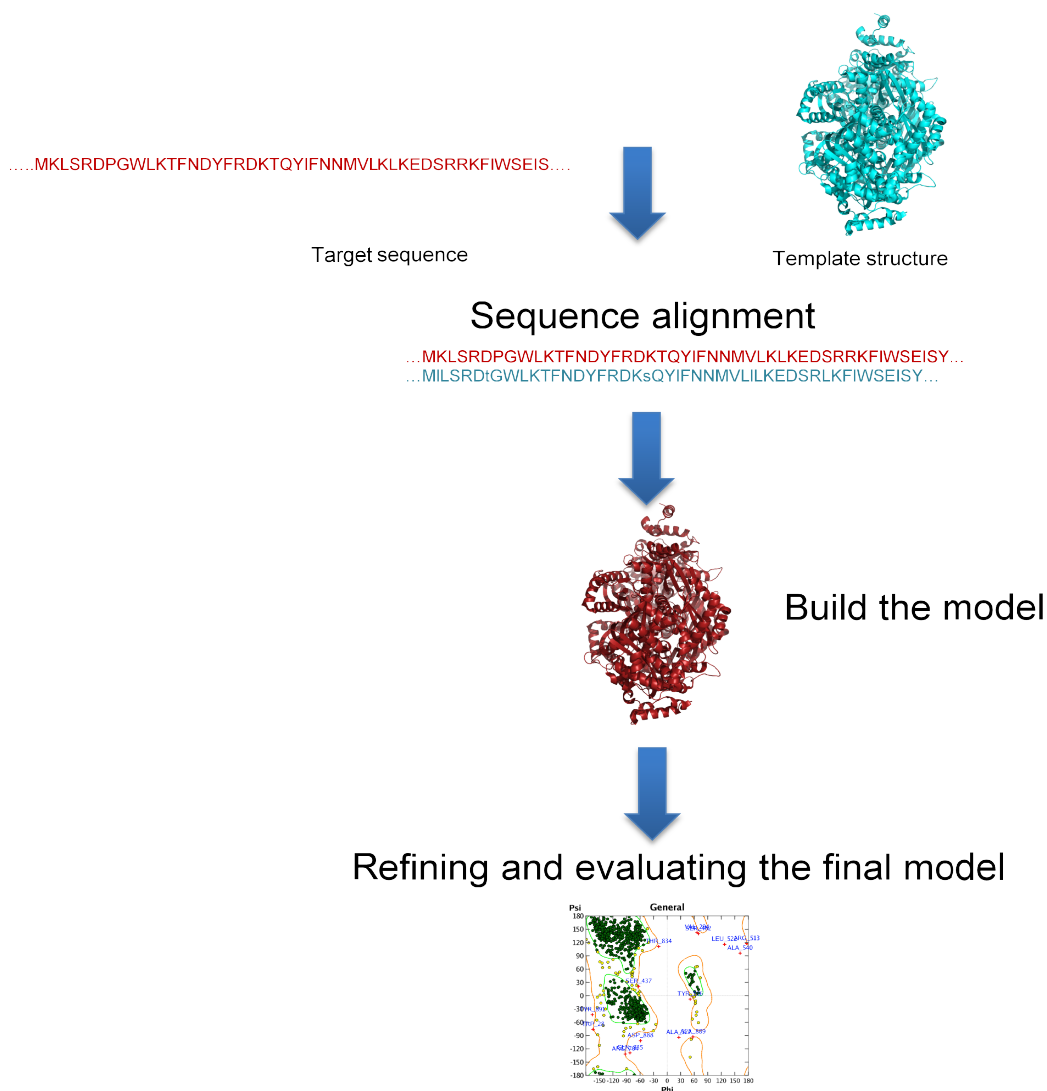


Figure 38: Homology modeling consists of different steps including the identification of the template, the alignment of the sequences, building a model and evaluating the final model

3.2 RMSD value

The RMSD (root mean square deviation) is an important value, which gives informations about the similarity between structures. A RMSD value, which is less than 3 Å, indicates a good similarity between structures [136]. It is possible to validate a receptor model by redocking the reference ligand and calculating the RMSD value.

The RMSD value is given by following equation:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^N d_i^2}{N}}$$

Equation 13: d_i is the distance (in Å) between the i th atom of the two conformations and N the number of such distances [137].

3.3 Molecular graphics

The figures showing the protein structures were made using PyMOL (<http://pymol.org>) [138]. The script "colorbyrmsd.py" (<https://pymolwiki.org/index.php/ColorByRMSD>) was used to compare the similarity of two aligned protein structures due to evaluation of the RMSD. Hereby, good alignments are indicated by blue colour, higher deviations by orange/yellow/red colour and residues that are not aligned are coloured in white.

3.4 Virtual ligand screening of the catalytic site of the homology modell of hGMII

3.4.1 Receptor preparation

The input receptor molecule for the virtual screening approach was the homology model of hGMII which was built as described in section 3.1. Before the filtered database was docked using LeadIT [139], the optimal docking parameters were first determined by docking of swainsonine into the catalytic site of hGMII. The following parameters have been set in the LeadIT [139] program:

- Residues in the active site were deprotonated via the Protoss module of LeadIT, thereby considering the physiological pH and neighbouring group effects.
- Asp289 presumably acting as nucleophile in the catalytic reaction was deprotonated and Asp426 acting as acid/base catalyst was protonated
- The zinc ion was octahedrally coordinated

3.4.2 Docking of swainsonine for testing the receptor setup

First a docking of swainsonine was performed in order to test the suitability of the chosen receptor setup.

3.4.3 Database preparation

A virtual database of about 10 0000 commercially available compounds from the suppliers Sigma Aldrich (<https://www.sigmaaldrich.com/germany.html>) and Timtec (<http://www.timtec.net/>) was used as starting point. The compounds were initially stored in sd-format and loaded in MOE v 2012.10 (Chemical computing group, Montreal, QC, Canada) in SMILES[140] format. Physicochemical properties of the database compounds such as the molecular weight, number of hydrogen bond acceptors or donors, logP values and rotational bonds were calculated. Subsequently, the database wash application within MOE was applied since it was a database from an external source. The MOE wash application was performed to correct the protonation states (deprotonate strong acids and protonate strong bases) and disconnect Group I metals into simple salt in order to keep only the largest

molecular fragments. The adjusted library was then filtered for desired drug-like properties (see Table 3).

Table 3: Compound selection criteria

Selection criterion	Definition
SlogP	> - 1
MW [g/mol]	≤ 500
TPSA [Å]	< 140
Rotatable bonds	≤ 7
Fused ring systems	< 5
Heavy atoms	< 30
Hydrogen-bond donors	< 6
Hydrogen-bond acceptors	< 11

3.4.4 Pharmacophore search

The database was filtered for desired properties. A pharmacophore step was then applied to introduce a second filter step. The pharmacophore model was based on the binding mode of swainsonine to dGMII. Both, the pharmacophore model and the pharmacophore search were carried out using MOEv.2012.10 (Chemical Computing Group, Montreal, QC, Canada). In preparation of the pharmacophore search, all the compounds passing the first filter step were converted into a three-dimensional format (sd-file). Low energy conformers of the compounds were calculated using OMEGA2 (Openeye, Santa Fe, NM), which were then applied to the pharmacophore search. Only the best conformation of a database compound was kept for storage in the output file so that each compound passing the pharmacophore search only occurred once. Finally, these molecules were used for docking.

3.4.5 Molecular docking

LeadIT v.2.1.6. was used to dock the remaining database molecules. The following parameters were chosen:

- All amino acid residues within 8 Å of the zinc ion were considered as part of the binding site
- Docking strategy: Hybrid approach (enthalpy and entropy)
- The maximum allowed overlap volume has been set to 2.9 Å³
- The intra-ligand clash-factor was set to 0.6

Only the best scoring pose was taken into account. For each pose, a HYDE [112] assessment was applied.

3.4.6 Docking analysis and molecular graphics

The docking poses were stored in the FlexX solutions database alongside with the scores in kJ/mol. As described in section 1.6.2, the scoring function (see Equation 3, [110]) contains the following terms: match, lipophilicity, ambiguity, clash and rotatable bonds. These terms are listed in the database. Postulated binding modes in 2D were visualized in LeadIT with the PoseViewer. The 3D diagrams of the docking poses as well as figures of the pharmacophore search were made using PyMOL (<http://pymol.org>) [138].

3.5 Virtual ligand screening of the potential allosteric site of dGMII

3.5.1 Identification of a potential allosteric binding site of Golgi α -mannosidase II

The potential allosteric binding pocket of the GMII protein was identified by using the Site Finder application of MOE v2014.10 (Chemical Computing Group, Montreal, QC, Canada) [141], which uses a methodology based on so-called alpha spheres that are derived from the receptor atoms and describe the shape of the protein. Each alpha sphere was classified as either “hydrophilic” or “hydrophobic”. Collections of neighbouring spheres were clustered to produce a collection of sites, which were ranked according to the number of hydrophobic contacts with the receptor. In the rank-ordered list for each potential site the number of receptor atoms (size), the number of hydrophobic contacts, the number of side chain contacts and the PLB (Propensity for Ligand Binding) score [142] were reported along with the residues containing the respective contacting atoms. During the site identification with the Site Finder application, the radii of a hypothetical hydrophilic hydrogen bonding atom (such as N or O) and of a hypothetical hydrophobic atom (such as C) were 1.4 Å and 1.8 Å, respectively. Each hydrophilic alpha sphere that has no neighbouring hydrophobic alpha sphere within 3 Å was discarded. Clusters of alpha spheres were merged, if at least two pairs of spheres were within 2.5 Å. Potential sites were only kept, if they consisted of at least three alpha spheres.

3.5.2 Comparison of binding site properties – active site vs. potential allosteric site

DogSite Scorer was used to compare the characteristics and the druggability of the active and the potential allosteric site. The DoGSite Scorer is a web server (<http://dogsite.zbh.uni-hamburg.de>) employed for identification and characterization of potential new binding pockets as well as the investigation of subpockets. General properties of the detected pockets such as the volume, the surface, the shape and the so-called druggability score are calculated. A SVM (support vector machine) model based on a discriminant analysis separates druggable from less druggable binding pockets. The druggability score is estimated for each pocket and subpocket resulting in values between zero and one, whereupon pockets with a score approaching one are assumed to be the more druggable ones [143], [144].

3.5.3 Receptor preparation and sphere set generation for virtual screening

The crystal structure of the *Drosophila melanogaster* GMII in complex with the inhibitor swainsonine (PDB code 1HWW [30]) was chosen as the receptor for the docking study at the potential allosteric site. First, polar hydrogen atoms were added to the protein and their positions were minimized using the AMBER force field [145] as implemented in MOE. Crystallographic water molecules and ions were removed in order to obtain a final receptor setup consisting of protein atoms only. Amber partial charges [146] were assigned to the protein atoms. Due to the lack of a reference ligand, the proposed binding site was manually filled with dummy atoms. Subsequently, spheres were placed into the proposed binding site around the dummy atoms to lower the effective dielectric constant in this region and to serve as matching points for ligand atoms during docking. The position of the sphere points in the final receptor setup used for docking was based on a simple cubic grid wrapping the dummy ligand atoms. The distance between neighbouring grid points was 2.2 Å at a maximum distance of 4 Å to any dummy atom. No sphere was closer than 2.5 Å to any receptor atom. Binding site properties such as the electrostatic and van der Waals potential as well as the solvent-excluded volume were stored as grids. These grids were calculated as described earlier [147],[148].

3.5.4 Small molecule preparation

A virtual in-house library of 5,574,182 commercially available small molecules was used as starting point for the pharmacophore search and docking, respectively. The molecules were stored in SMILES format [140] alongside with their calculated physicochemical properties (e.g. molecular weight, number of H-bond donors/acceptors, logP) [149]. The molecules were filtered first and then converted in a dockable format similar as described previously (section 3.4.3). The library compounds were filtered for desired lead-like properties (see Table 4). Additionally, all compounds containing unwanted groups were removed. For the remaining molecules, the most relevant protomers, tautomers and stereoisomers were calculated using in-house python scripts based on the OEToolkit (Openeye, Santa Fe, NM). Afterwards, three-dimensional structures and low energy conformations of all stereoisomers of compounds passing both the physicochemical property filter step and the pharmacophore search were generated with OMEGA2 [150] (version 2.4.6) (Openeye, Santa Fe, NM). Partial charges and desolvation energies for the transfer of the small molecules from high dielectric

medium (bulk solvent) to low dielectric medium within the binding site were calculated using AMSOL (<http://comp.chem.umn.edu/amsol/>) [151]. Finally, the multiple conformers of each molecule were aligned to their ring systems and stored in a hierarchical (flexibase) format [152].

Table 4: Compound selection criteria

Selection criterion	Definition
SlogP	< 3.5 & > - 3.5
MW [g/mol]	<= 350 & > 250
TPSA [Å]	< 140
Rotatable bonds	< 7
Fused ring systems	< 4
Total charge	< 2 & > -2
Heavy atoms	< 30
Hydrogen-bond donors	< 6
Hydrogen-bond acceptors	< 11
Absence of unwanted functionalities	No unwanted groups*

*mutagenic groups such as nitro groups, reactive groups such as thiols or michael acceptors, groups interfering with typical HTS assays such as quinones, groups with unfavourable pharmacokinetic properties such as sulfates [149]

3.5.5 Pharmacophore search

The pharmacophore model was developed on the basis of the spatial arrangement of the amino acids within the proposed allosteric binding pocket with special considerations to potential electrostatic or ionic interactions that could be formed. Both the pharmacophore model development and the pharmacophore search were carried out using MOE v.2014.10. In order to find suitable annotation points for the pharmacophore query, the standard “Unified” annotation scheme was applied. In total, a set of 6 pharmacophore features was created. For the pharmacophore search, the stereoisomers of molecules passing the first filter step for desirable properties were converted into a three-dimensional format as

described above (sd-file) (see section 3.5.4). Only those molecules with at least one low-energy conformation meeting the criteria of the pharmacophore search were kept for docking.

3.5.6 Molecular docking

DOCK 3.6 was used to place the ligands into the potential allosteric binding site [151]–[153]. The following parameters were chosen for sampling of the ligand orientations: ligand and receptor bin sizes were 0.5 Å, whereas the overlap bin size was 0.4 Å and 0.5 Å for receptor and ligand bins, respectively. The distance tolerance for matching ligand atoms to the receptor spheres used for matching were 1.2 Å. Each docking pose was scored for electrostatic and van der Waals complementarity and corrected for partial desolvation [152]. Only the best-scoring representation (correct protomer, tautomer and stereoisomer) for each docked molecule was stored in the output database.

3.5.7 Docking analysis and molecular graphics

The docking poses were stored in a MOE database alongside with their corresponding scores in kJ/mol. Subsequently, predicted binding modes were visualized separately for closer inspection to manually select the most promising candidate compounds. For that purpose, ranking of the database entries was first performed according to the total docking score and in a second step according to the predicted ligand efficiency (LE) by using the following equation (see Equation 14):

$$LE = -1 * \left(\frac{Score}{HAC} \right)$$

Equation 14: "Score" represents the total docking score obtained and HAC is the heavy atom count

The diagrams for the representation of predicted binding modes and the pharmacophore search were prepared using PyMOL (<http://pymol.org>) [138]. The two-dimensional figures were generated using the Ligand Interaction Diagram tool of MOE.

The figures comparing the properties of the active and the potential allosteric site with regard to charged and neutral regions within the binding sites were prepared by using the the

PyMOL script “resicolor.py” (<https://pymolwiki.org/index.php/Resicolor>). Hereby, acidic and basic residues were coloured red and blue, respectively. All other protein parts were coloured grey.

3.6 Virtual ligand screening of the active site of dGMII

3.6.1 Receptor preparation and sphere set generation

The crystal structure of the *Drosophila melanogaster* GMII in complex with the inhibitor swainsonine (PDB code 1HWW [30]) was chosen as the receptor for the docking study at the potential allosteric site. First, polar hydrogen atoms were added to the protein and their positions were minimized using the AMBER force field [145] as implemented in MOE. Crystallographic water molecules and ions, except for Zn^{2+} , were removed in order to obtain a final receptor setup consisting of protein atoms only. All water molecules have been removed from the binding pocket because swainsonine directly interacts with the protein. Amber partial charges [146] were assigned to the protein atoms. A partial charge of +2 was assigned to Zn^{2+} . Asp204 acting as nucleophile was deprotonated and Asp341 serving as acid/base catalyst was protonated. A sphere set around swainsonine was generated in the catalytic site. The distance between neighbouring grid points was 2.4 Å at a maximum distance of 4 Å to any dummy atom. No sphere was closer than 1.7 Å to any receptor atom.

3.6.2 Redocking of swainsonine

In order to validate the docking preparation, swainsonine was docked into the catalytic site. The RMSD value of the docking pose of swainsonine compared to the crystallographically determined conformation was evaluated. The RMSD value between the crystallised binding mode of swainsonine and the docked pose of swainsonine was calculated by using the python script `ligand_alignment.py` (<http://compbio.cs.toronto.edu/ligalign/downloads.html>) [154].

3.6.3 Small molecule preparation

The protocol for the preparation of the small molecules is the same as described in section 3.5.4

3.6.4 Pharmacophore search

In order to have an additional filter step beyond choosing compounds solely based on their physicochemical properties, a pharmacophore search was carried out. The pharmacophore model was based on the binding mode of swainsonine in the catalytic site. Both the pharmacophore model development and the pharmacophore search were carried out using MOE v.2014.10 (see section 3.4.4).

3.6.5 Molecular docking

DOCK 3.6 was also used to place the ligands into the catalytic binding site [151]–[153]. The sampling parameters were chosen as already described in section 3.5.6.

3.6.6 Docking analysis and molecular graphics

For details regarding the performance of the docking analysis it is referred to section 3.5.7.

3.7 Enzyme Assay

3.7.1 Photometric assay

The commercially available JBM was the first model enzyme available, before the working group of [REDACTED] (JMU, Würzburg) supplied dGMII. The former diploma student Michael Juchum (working group [REDACTED], JGU, Mainz, 2012) [155] established the photometric assay for JBM. In the work of Michael Juchum, the assay system was tested with the commercially available (Sigma Aldrich) inhibitor swainsonine. Measurements were carried out to determine IC_{50} values and classical enzyme kinetics experiments were performed in addition [155]. With this photometric enzyme assay, the purchased hit compounds of the first virtual screening with the hGMII (section 3.4) were tested for their inhibitory potency. The absorbance measurements were carried out on a Tecan Infinite F200 PRO microplate reader (absorbance at 405 nm). The following substances were obtained from Sigma Aldrich: *Canavalia ensiformis* GMII (ammonium sulfate suspension), MES monohydrate (BioUltra, for molecular biology), zinc sulfate solution (2.0 M in H₂O, BioUltra, for molecular biology), albumin (from bovine serum, BioReagent), substrate 4-nitrophenyl- α -D-mannopyranoside, and swainsonine.

3.7.1.1 Enzyme assays with the Jack Bean Mannosidase

3.7.1.1.1 Testing of the virtual screening hits

Reaction buffer:

MES-buffer (40 mM) supplemented with bovine serum albumin 1.1% (0,1% final concentration in MES-buffer) und ZnSO₄-solution 1.1% (0.1% final concentration in MES-buffer), pH = 5,75.

Enzyme:

Jack Bean-Mannosidase, in MES-buffer supplemented with bovine serum albumin 1.7% (1.0% final concentration in MES-buffer) and ZnSO₄-solution 1.7% (1.0% final concentration in MES-buffer), concentration per well: 0.04 U/mL

Substrate:

4-Nitrophenyl- α -D-mannopyranose in reaction buffer, concentration per well: 1 mM

Inhibitor:

Virtual screening hit compound dissolved in buffer: 1 mM final compound concentration per well.

Pipetting per well:

15 μ L inhibitor + 15 μ L substrate + 15 μ L enzyme + 105 μ L buffer

Positive control: 15 μ L buffer + 15 μ L substrate + 15 μ L enzyme + 105 μ L buffer

Negative control: 15 μ L inhibitor + 15 μ L substrate + 15 μ L buffer + 105 μ L buffer

Experimental procedure:

Compound solutions and the buffer were pipetted into the wells in triplicate. This was followed by a 2 minute incubation of the solutions at 37 °C. Using a multichannel pipette, 15 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 15 minutes. Finally, 15 μ L of the substrate solution was added and the reaction was incubated again for 5 minutes. The reaction was then stopped with 15 μ L sodium carbonate solution and the absorption was measured at 405 nm.

3.7.2 Fluorometric assay

The establishment of these assays and the measurements of the compounds were performed in cooperation with students of the Johannes Gutenberg University, namely [REDACTED], in the context of their bachelor thesis and with [REDACTED] in the context of her diploma thesis.

The fluorescence measurements were carried out on a Tecan Reader Infinite F200 PRO microplate reader equipped with a filter system (355 nm excitation and 485 nm emission). The release of 4-methylumbelliferone was measured for 15 min by fluorescence spectroscopy using Greiner 96 Flat Bottom Black Polystyrol plates. The synthesis of the substrates 4-MU-Man and 4-MU-Glc, respectively was accomplished by [REDACTED] (PhD student, working group [REDACTED], JGU, Mainz). Studies based on the use of 4-MU-Man as substrate and the monitoring of the release of fluorescence by cleavage through GMII proteins have been described earlier [156]. The dGMII enzyme was recombinantly expressed and purified by [REDACTED] (working group [REDACTED], JMU, Würzburg). JBM and the β -glucosidase from sweet almonds were purchased from Sigma Aldrich.

3.7.2.1 Enzyme assays with dGMII

The reaction buffer and the enzyme were prepared as follows:

Reaction buffer:

MES-buffer (40 mM) supplemented with bovine serum albumin 1.1% (0.1% final concentration in MES-buffer) und ZnSO₄-solution 1.1% (0.1% final concentration in MES-buffer), pH = 5,75.

Enzyme:

GMII from *Drosophila melanogaster*, in MES-buffer supplemented with bovine serum albumin 1.7 % (1.0% final concentration in MES-buffer) und ZnSO₄-Lösung 1.7 % (1.0 % final concentration in MES-buffer), concentration per well: 0.002 mg/mL.

3.7.2.1.1 K_M value determination

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) diluted in reaction buffer to concentrations of: 7 mM; 6 mM; 5 mM; 4 mM, 3 mM; 2 mM; 1.75 mM; 1.5 mM; 1 mM; 0.75 mM; 0.5 mM per well.

Pipetting per well:

Positive control: 40 μ L substrate + 10 μ L enzyme

Negative control: 40 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 40 μ L of the differently concentrated substrate solutions were pipetted into the wells. This was followed by a 2 minute incubation of the substrate solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution was pipetted into each well (for the negative control 10 μ L buffer was used instead), the reaction volume was then mixed with a multichannel pipette and the plate was incubated for 5 minutes at 37 °C. The determination was made in triplicate. The well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.1.2 IC₅₀ determination of swainsonine

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 0.5 mM; 1 mM, 1.5 mM; 2 mM.

Inhibitor:

Swainsonine in buffer per well: 600 nM, 450 nM, 300 nM, 150 nM, 120 nM, 60 nM, 48 nM, 45 nM, 36 nM, 30 nM, 24 nM, 15 nM, 12 nM, 6 nM

Pipetting per well:

30 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme

Positive control: 30 μ L buffer + 10 μ L substrate + 10 μ L enzyme

Negative control: 30 μ L inhibitor + 10 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 30 μ L of the differently concentrated swainsonine solutions were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2 minute incubation of the solutions at 37°C. Using a multichannel pipette, 10 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes (negative control: 10 μ L buffer). Finally, 10 μ L of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C. The negative control contained 150 nM swainsonine per well to exclude a too high autofluorescence of the inhibitor and was determined in triplicate. The positive control was measured in quadruplicate on the plate in order to obtain the uninhibited enzyme activity.

3.7.2.1.3 IC₅₀ determination of swainsonine as slow binder

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 1 mM

Inhibitor:

Swainsonine in buffer per well: 1000 nM, 750 nM, 600 nM, 500 nM, 250 nM, 100 nM, 50 nM, 25 nM, 10 nM

Pipetting per well:

30 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme

Positive control: 30 μ L buffer + 10 μ L substrate + 10 μ L enzyme

Negative control: 30 μ L inhibitor + 10 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 30 μ L of the differently concentrated swainsonine solutions were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2 minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for **10** minutes, **5** minutes and **0** minutes, respectively (negative control: 10 μ L buffer). Finally, 10 μ L of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.1.4 K_i value of swainsonine

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 10 mM; 9 mM; 8 mM; 7 mM; 6 mM; 5 mM; 4 mM, 3 mM; 2 mM; 1.75 mM; 1.5 mM; 1 mM; 0.75 mM; 0.5 mM

Inhibitor:

Swainsonine in reaction buffer 40 nM per well

Pipetting per well:

10 μ L inhibitor + 30 μ L substrate + 10 μ L enzyme

10 μ L buffer + 30 μ L substrate + 10 μ L enzyme (for the K_M value)

Experimental procedure:

On a 96-well plate, the enzyme and the inhibitor or buffer were pre-pipetted into the wells in triplicate. Both were incubated for 5 min at 37 °C. After the incubation time, 30 µL of the differently concentrated substrate solutions were pipetted into the wells by using a multichannel pipette and the reaction volume was then mixed with the same multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.1.5 Z'-factor determination

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 1 mM.

Inhibitor:

Swainsonine in reaction buffer: 50 nM per well

Pipetting per well:

Positive control: 30 µL reaction buffer + 10 µL enzyme + 10 µL substrate.

Negative control: 27.5 µL reaction buffer + 2.5 µL inhibitor + 10 µL enzyme + 10 µL substrate.

Experimental procedure:

On a 96-well plate **16** positive controls and **16** negative controls were pipetted. All the components, except for the substrate and the enzyme, were first pipetted in the wells. This step was followed by a 2 minute pre-incubation period at 37 °C. After that, 10 µL of the enzyme was added and the plate was incubated for 5 minutes at 37 °C. Subsequently 10 µL of the substrate was added and the reaction volume was then mixed with a multichannel pipette. The well plate was shaken for 10 seconds and the fluorescence measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.1.6 DMSO effect on the activity of dGMII

The effect of DMSO was tested by investigation on the overall reaction rate of the enzymatic reaction. All the purchased compounds and most of the synthesized inhibitors were dissolved in DMSO. In order to evaluate the effect of DMSO on the enzyme activity, different DMSO concentrations were tested in the assay mixture. In a recent study of L. Misuri *et al.* [157], DMSO showed to act as a weak aldose reductase inhibitor. This publication revealed that it is important to keep the DMSO concentration low in order to avoid interferences of the solvent with the enzymatic assay. DMSO has to be considered as an additional factor which may have an influence on inhibition parameters. The DMSO effect on the activity of dGMII was tested as described in the following.

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man, 1 mM per well) in reaction buffer containing different concentrations of DMSO (20%, 10%, 5%, 4%, 3%, 2%, 1% and 0% in relation to the total reaction volume)

Pipetting per well:

Positive control: 40 μ L substrate + 10 μ L enzyme

Experimental procedure:

On a 96-well plate, 40 μ L of the substrate solutions with different DMSO concentration were pipetted into the wells. This was followed by a 2-minute incubation of the substrate solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution were finally pipetted into each well (for the negative control 10 μ L buffer instead of enzyme was used), the reaction volume was then mixed with the multichannel pipette and the plate was incubated for 5 minutes at 37 °C. The well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.1.7 Testing of the virtual screening hits

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer at a concentration of 1 mM per well

Inhibitor:

Compounds **1, 2, 3, 4, 5, 6, 7, 8, 9, 10,11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27** with a well concentration of 300 μ M

Compound **24** with a well concentration of 100 μ M due to the small available amount

Compounds **28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41** with a well concentration of 100 μ M due to the small available amount

Compounds **42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58** with a well concentration of 300 μ M

All the compounds were dissolved in DMSO.

Pipetting per well:

2.5 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme + 27.5 μ L buffer

Positive control: 10 μ L substrate + 10 μ L enzyme + 30 μ L buffer

Negative control: 2.5 μ L inhibitor + 10 μ L substrate + 37.5 μ L buffer

Experimental procedure:

On a 96-well plate, 2.5 μ L of the compounds solutions were pipetted into the wells in sextuplicate (positive control: 30 μ L buffer, negative control: 37.5 μ L buffer). This was followed by a 2-minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μ L of enzyme solution were pipetted into each well and the reaction mixture was then incubated again for five minutes. Finally, 10 μ L of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C. The positive control was measured in duplicate (for each compound) on the plate in order to obtain the uninhibited enzyme activity.

3.7.2.1.8 IC₅₀ determination of compound 18

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 0.1 mM, 0.5 mM, 1.0 mM, 1.5 mM, 2.0 mM, 3.0 mM

Inhibitor:

Compound **18** in at a concentration of DMSO 1000 μM , 800 μM , 600 μM , 400 μM , 200 μM , 100 μM , 50 μM , 10 μM

Pipetting per well:

2.5 μL inhibitor + 10 μL substrate + 10 μL enzyme + 27.5 μL buffer

Positive control: 10 μL substrate + 10 μL enzyme + 30 μL buffer

Negative control: 2.5 μL inhibitor + 10 μL substrate + 37.5 μL buffer

Experimental procedure:

On a 96-well plate, 2.5 μL of the compounds solutions were pipetted into the wells in triplicate (positive control: 30 μL buffer). This was followed by a two-minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μL of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for five minutes. Finally, 10 μL of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C. The positive control was measured in triplicate for each IC_{50} determination in order to obtain the uninhibited enzyme activity.

3.7.2.1.9 IC_{50} determination of compound **32**

████████████████████ (JGU Mainz) did the IC_{50} measurements of compound **32** as part of her diploma thesis.

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 0.1 mM, 1 mM and 2 mM

Inhibitor:

Compound **32** dissolved in DMSO and buffer (5 % v/v DMSO per well): 700 μM , 600 μM , 500 μM , 400 μM , 300 μM , 200 μM , 100 μM and 10 μM

Pipetting per well:

30 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme

Positive control: 30 μ L buffer + 10 μ L substrate + 10 μ L enzyme

Negative control: 30 μ L inhibitor + 10 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 30 μ L of the differently concentrated solutions were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2-minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes. Finally, 10 μ L of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.1.10 K_i determination of compound **32**

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 7 mM, 6 mM, 5mM, 4 mM, 3 mM, 2mM, 1.75 mM, 1.50 mM, 0.75 mM and 0,50 mM

Inhibitor:

Compound **32** dissolved in DMSO and buffer (5 % v/v DMSO per well): 150 μ M

Pipetting per well:

30 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme

Positive control: 30 μ L buffer + 10 μ L substrate + 10 μ L enzyme

Negative control: 30 μ L inhibitor + 10 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 30 μ L of the differently concentrated solutions were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2-minute incubation of the

solutions at 37 °C. Using a multichannel pipette, 10 µL of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes. Finally, 10 µL of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.2 Enzyme assays with the Jack Bean Mannosidase

For the fluorescence assay detecting the enzymatic activity of JBM, the reaction buffer and the enzyme were composed as follows:

Reaction buffer:

MES-buffer (40 mM) supplemented with bovine serum albumin 1.1% (0.1% final concentration in MES-buffer) und ZnSO₄-solution 1.1% (0.1% final concentration in MES-buffer), pH = 5.75.

Enzyme:

Jack Bean-Mannosidase, in MES-buffer supplemented with bovine serum albumin 1.7 % (1.0% final concentration in MES-buffer) und ZnSO₄-solution 1.7 % (1.0 % final concentration in MES-buffer), concentration per well: 0.04 U/mL

3.7.2.2.1 K_M-value determination for JBM

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 5 mM; 4 mM; 3 mM; 2mM ;1.75 mM; 1.5 mM; 1.25 mM; 1 mM; 0.8 mM; 0.6 mM; 0.4 mM; 0.2 mM; 0.1 mM; 0.05 mM; 0.01 mM

Pipetting per well:

Positive control: 40 µL substrate + 10 µL enzyme

Negative control: 40 µL substrate + 10 µL buffer

Experimental procedure:

On a 96-well plate, 40 μL of the differently concentrated substrate solutions were pipetted in triplicate into the wells. This was followed by a 2-minute incubation of the substrate solutions at 37 °C. Using a multichannel pipette, 10 μL of the enzyme solution were finally pipetted into each well (for the negative control 10 μL buffer instead of enzyme), the reaction volume was then mixed with a multichannel pipette and the plate was incubated for another 5 minutes at 37 °C. The determination was carried out in triplicate. The well plate was shaken for 10 seconds and the fluorescence measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.2.2 IC₅₀ determination of swainsonine

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 0.05 mM; 0.1 mM; 0.2 mM; 0.3 mM; 0.7 mM

Inhibitor:

Swainsonine in buffer, concentrations per well: 1000 nM; 750 nM; 500 nM; 250 nM; 200 nM; 100 nM; 80 nM; 75 nM; 60 nM; 50 nM; 40 nM; 25 nM, 20 nM; 10 nM.

Pipetting per well:

30 μL inhibitor + 10 μL substrate + 10 μL enzyme

Positive control: 30 μL buffer + 10 μL substrate + 10 μL enzyme

Negative control: 30 μL inhibitor + 10 μL substrate + 10 μL buffer

Experimental procedure:

On a 96-well plate, 30 μL of the differently concentrated swainsonine solutions were pipetted into the wells in triplicate (positive control: 30 μL buffer). This was followed by a 2-minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μL of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes (negative control: 10 μL buffer). Finally, 10 μL of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a

period of 15 minutes at a temperature of 37 °C. The negative control contained 150 nM swainsonine per well to exclude a too high autofluorescence of the inhibitor and was determined in triplicate. The positive control was measured in quadruplicate on the plate in order to obtain the uninhibited enzyme activity.

3.7.2.2.3 K_i value of swainsonine as slow binder

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 4 mM; 3 mM; 2mM; 1.75 mM; 1.5 mM; 1.25 mM; 1 mM; 0.8 mM; 0.6 mM; 0.4 mM; 0.2 mM; 0.1 mM; 0.05 mM; 0.01 mM.

Inhibitor:

Swainsonine at a concentration of 150 nM in reaction buffer per well

Pipetting per well:

10 μ L inhibitor + 30 μ L substrate + 10 μ L enzyme

10 μ L buffer +30 μ L substrate + 10 μ L enzyme (for the K_M value)

Experimental procedure:

On a 96-well plate, the enzyme and the inhibitor or buffer were pipetted in triplicate into the wells. Enzyme and inhibitor were incubated for 5 minutes at 37 °C. After the incubation time, 30 μ L of the differently concentrated substrate solutions were pipetted into the wells using a multichannel pipette and the reaction volume was then mixed with the same multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.2.4 Testing of compound **18**

Substrate:

4-Methylumbelliferyl- α -D-mannopyranoside (4-MU-Man) in reaction buffer, concentration per well: 0.4 mM

Inhibitor:

Compound **18** dissolved in DMSO (final concentration per well: 300 μ M)

Pipetting per well:

2.5 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme + 27.5 μ L buffer

Positive control: 10 μ L substrate + 10 μ L enzyme + 30 μ L buffer

Negative control: 2.5 μ L inhibitor + 10 μ L substrate + 37.5 μ L buffer

Experimental procedure:

On a 96-well plate, 2.5 μ L of the solution of compound **18** in DMSO were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2-minute incubation of the solutions at 37°C. Using a multichannel pipette, 10 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes (negative control: 10 μ L buffer). Finally, 10 μ L of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C. The negative control contained 300 μ M of compound **18** per well to exclude a too high autofluorescence of the compound and was determined in triplicate. The positive control was measured in triplicate on the plate in order to obtain the uninhibited enzyme activity.

3.7.2.3 Enzyme assays with the β -glucosidase from sweet almonds

██████████ (diploma student, JGU, Mainz) and ██████████ (B.Sc. Chemie, JGU, Mainz) carried out the K_M determination and IC_{50} determination. For the fluorescence-based enzyme assay with β -glucosidase, the reaction buffer and the enzyme were composed as follows:

Reaction buffer:

Citrate buffer, pH = 5

Enzyme:

β -Glucosidase from sweet almonds at a concentration of 40 μ g/ml (0.748 U/ml) in citrate buffer.

3.7.2.3.1 K_M Value determination

Substrate:

4-Methylumbelliferyl- α -D-glucopyranoside (4-MU-Glc) in reaction buffer, concentration per well: 7 mM, 6 mM, 5 mM, 4 mM, 3 mM, 2 mM, 1 mM, 0.75 mM, 0.5 mM, 0.25 mM.

Pipetting per well:

Positive control: 40 μ L substrate + 10 μ L enzyme

Negative control: 40 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 40 μ L of the differently concentrated substrate solutions were pipetted in triplicate into the wells. This was followed by a 2-minute incubation of the substrate solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution were finally pipetted into each well (for the negative control 10 μ L buffer instead of enzyme), the reaction volume was then mixed with a multichannel pipette and the plate was incubated for five minutes at 37 °C. The determination was carried out in triplicate. The well plate was shaken for 10 seconds and the fluorescence measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C.

3.7.2.3.2 IC_{50} determination of isofagomine

Substrate:

4-Methylumbelliferyl- α -D-glucopyranoside (4-MU-Glc) in reaction buffer, concentration per well: 1 mM

Inhibitor:

Isofagomine at the following concentrations in reaction buffer per well: 1000 nM, 900 nM, 800 nM, 700 nM, 600 nM, 500 nM, 400 nM, 300 nM, 100 nM

Pipetting per well:

30 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme

Positive control: 30 μ L buffer + 10 μ L substrate + 10 μ L enzyme

Negative control: 30 μ L inhibitor + 10 μ L substrate + 10 μ L buffer

Experimental procedure:

On a 96-well plate, 30 μ L of the differently concentrated isofagomine solutions were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2-minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes (negative control: 10 μ L buffer). Finally, 10 μ L of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C. The negative control contained 500 nM isofagomine per well to exclude a too high autofluorescence of the inhibitor and was determined in triplicate. The positive control was measured in triplicate on the plate in order to obtain the uninhibited enzyme activity.

3.7.2.3.3 Testing of compound **18**

Substrate:

4-Methylumbelliferyl- α -D-glucopyranoside (4-MU-Glc) in reaction buffer, concentration per well: 1 mM

Inhibitor:

Compound **18** dissolved in DMSO (final concentration per well: 300 μ M)

Pipetting per well:

2.5 μ L inhibitor + 10 μ L substrate + 10 μ L enzyme + 27.5 μ L buffer

Positive control: 10 μ L substrate + 10 μ L enzyme + 30 μ L buffer

Negative control: 2.5 μ L inhibitor + 10 μ L substrate + 37.5 μ L buffer

Experimental procedure:

On a 96-well plate, 2.5 μ L of the solution of compound **18** in DMSO were pipetted into the wells in triplicate (positive control: 30 μ L buffer). This was followed by a 2-minute incubation of the solutions at 37 °C. Using a multichannel pipette, 10 μ L of the enzyme solution were pipetted into each well and the reaction mixture was then incubated again for 5 minutes

(negative control: 10 µL buffer). Finally, 10 µL of the substrate solution was added and the reaction volume was then mixed with a multichannel pipette. After that, the well plate was shaken for 10 seconds and the fluorescence was measured every 30-40 seconds over a period of 15 minutes at a temperature of 37 °C. The negative control contained 300 µM of compound **18** per well to exclude a too high autofluorescence of the compound and was determined in triplicate. The positive control was measured in triplicate on the plate in order to obtain the uninhibited enzyme activity.

3.7.2.4 Formulas and general informations

We considered the compounds from the virtual screening potential allosteric site of dGMII as inactive if less than 40% of the enzyme activity was inhibited at a the highest possible well concentration of 300 µM. In the case of the virtual screening of the catalytic site of dGMII, a less stringent cut off was applied and substances were considered as inactive if less than 20% of the enzyme activity was inhibited at well concentrations of 300 µM (or the highest soluble concentration). For the virtual screening based on the homology model of hGMII, compounds were classified as inactive if less than 50% of the enzyme activity was inhibited at a well concentration of 1000 µM.

The inhibition was calculated with the following equation (see Equation 15):

$$\text{Inhibition (\%)} = \left(1 - \frac{\text{activity (with inhibitor)}}{\text{activity (without inhibitor)}} \right) \cdot 100$$

Equation 15

IC₅₀ values were calculated using following equation (see Equation 16):

$$y = \frac{y_{\max} - y_{\min}}{1 + \left(\frac{[I]}{IC_{50}}\right)^s} + y_{\min}$$

Equation 16: y_{\max} and y_{\min} are defined as 100% and 0% enzyme activity, respectively, s represents the slope of the linear part of the curves (Hill coefficient), and $[I]$ is the inhibitor concentration.

K_i values for a competitive inhibition were calculated using following equation (see Equation 17):

$$v = \frac{v_{\max} \cdot [S]}{K_M \left(1 + \frac{[I]}{K_i}\right) + [S]}$$

Equation 17: v_{\max} : maximum velocity at complete saturation of the enzyme, $[S]$: substrate concentration, $[I]$: inhibitor concentration, K_M : Michaelis Menten constant, K_i : inhibition constant

K_i values for a non-competitive inhibition were calculated using following equation (see Equation 18):

$$v = \frac{1}{1 + \frac{[I]}{K_i}} \cdot \frac{v_{max} \cdot [S]}{K_M + [S]}$$

Equation 18: v_{max} : maximum velocity at complete saturation of the enzyme, $[S]$: substrate concentration, $[I]$: inhibitor concentration, K_M : Michaelis Menten constant, K_i : inhibition constant

The Z'-factor value was calculated using following equation (see Equation 19):

$$Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|}$$

Equation 19: μ_{c+} : mean of the positive control data values, μ_{c-} : mean of the negative control data values, σ_{c+} : standard deviation of the positive control data values and σ_{c-} : standard deviation of the negative control data values

IC_{50} and K_i values are related by the Cheng-Prusoff equation [116] (see Equation 20)

$$IC_{50} = \left(\frac{[S]}{K_M} + 1 \right) \cdot K_i$$

Equation 20: IC_{50} : half maximal inhibitory concentration, $[S]$: substrate concentration, $[I]$: inhibitor concentration, K_M : Michaelis Menten constant, K_i : inhibition constant

In addition, the Lineweaver-Burk plot was utilized to determine the inhibition type. The reciprocal velocity data were plotted against the reciprocal substrate concentrations and a classical linear regression ($y = a + bx$) was used. Frequently, only the values of higher substrate concentrations were used. The program GraFit® (5.0.13, Erithacus Software, Horley, Surrey, UK (2006)) was used to calculate the IC_{50} and K_i -values as well as for the preparation of the respective diagrams.

4 Results

4.1 Homology modeling

4.1.1 Homology Model of the hGMII

The sequence alignment of hGMII (Q16706) [37] with dGMII (pdb:1HWW) [30] using MOE showed 41.0% identity and 57.9% similarity between both amino acid sequences. The superposition of the dGMII crystal structure with the 3D model of hGMII yielded to a RMSD value of 1.16 Å (see Figure 39).

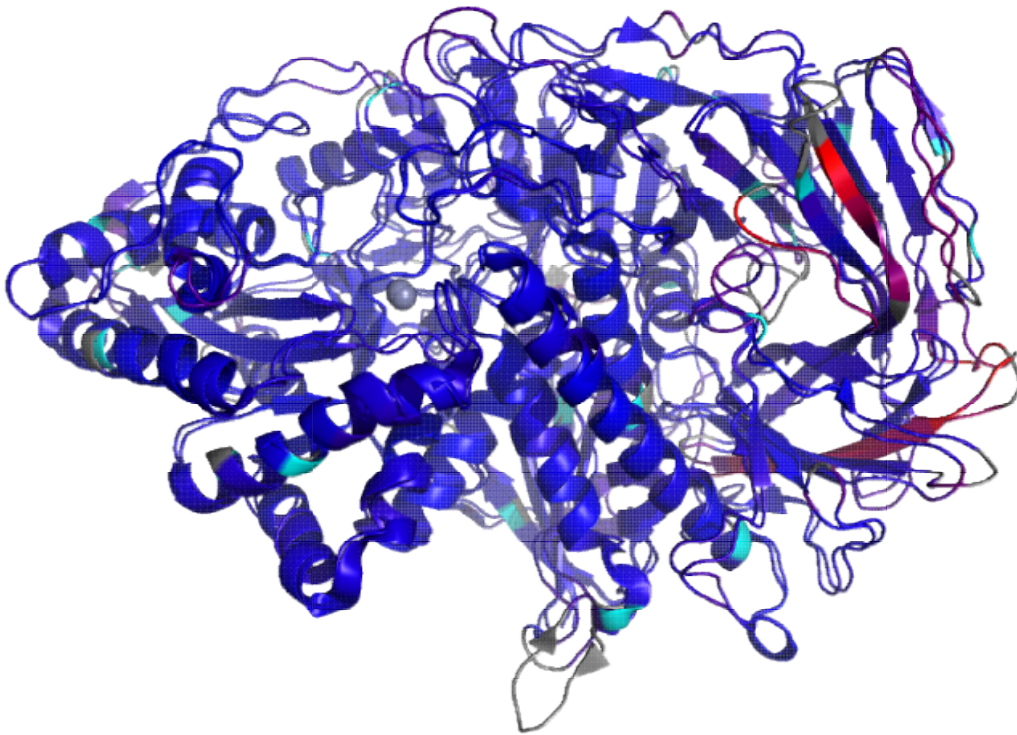


Figure 39: 1HWW and the homology model of hGMII aligned and coloured by RMSD using the pymol script colorbyrmsd.py. Hereby, blue colour indicates good alignment, whereas higher deviations are coloured red. Non-aligned residues are shown in grey colour.

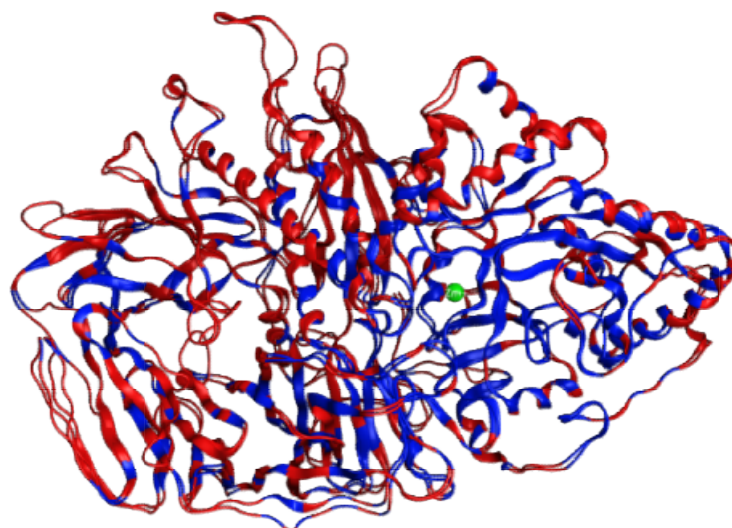


Figure 40: Superimposed structures of dGMII (1HWW) and the homology model of the hGMII. Areas with identical amino acids are coloured blue, areas with similar amino acids are coloured red. Representation of the identity of the homology model compared to the crystal structure of dGMII

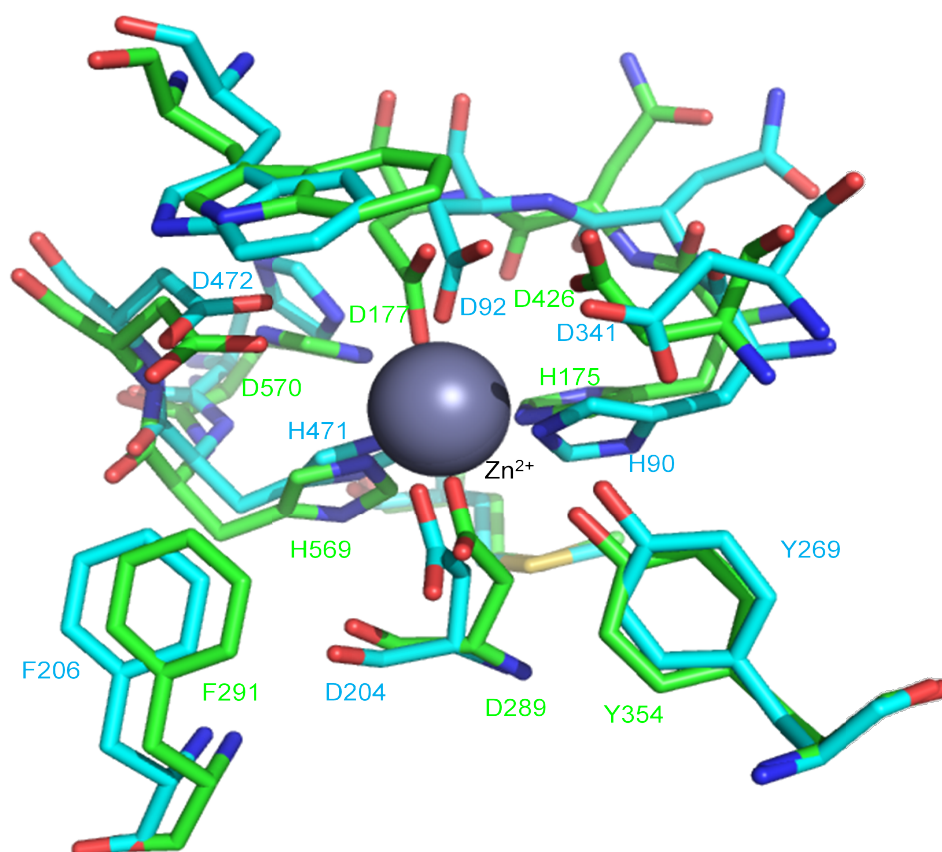


Figure 41: Superimposed structures of the dGMII (1HWW) and the homology model of the hGMII in an area of 6 Å around Zn^{2+} . Carbon atoms of 1HWW are coloured in cyan blue and hGMII in green.

With regard to the catalytic sites the structures and sequences of both dGMII and the homology model of hGMII are even more similar to each other as it is the case when considering the entire protein. In fact, dGMII and hGMII are 100% structurally identical within a radius of 6 Å around the central Zn²⁺-ion according to our homology modelling studies (see Figure 41). An overlay of both structures in this area resulted in a RMSD value of 0.76 Å.

4.1.2 Homology model of JBM

The α-mannosidase is also found in the Jack Bean (JBM), but its structure has not yet been fully elucidated [35]. The sequence alignment of JBM with the lysosomal α-mannosidase (bLAM) [158] (pdb code: 1O7D) using MOE showed 41.5% similarity and 58.7% identity with regard to the amino acid sequences. This is consistent with the sequence alignment of JBM and bLAM by Kumar *et al.* [35], which resulted in an amino acid identity of 41% between both proteins. In the following table (see Table 5), the identity and the similarity of JBM in comparison to bLM and dGMII is listed:

Table 5: Comparison of the amino acid sequence of Jack Bean α-mannosidase with the sequences of dGMII (pdb: 1HWW) and the lysosomal α-mannosidase (pdb: 1O7D) in terms of identity and similarity (in %).

Template	Identity	Similarity
Golgi-Mannosidase <i>Drosophila melanogaster</i> pdb: 1HWW	25.4%	42.9%
Lysosomal-Mannosidase <i>Bos taurus</i> pdb: 1O7D	41.5%	58.7%

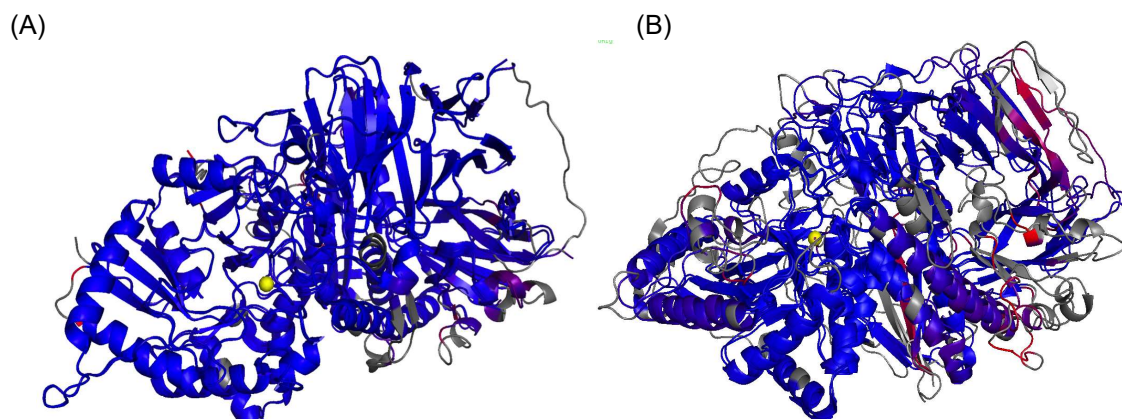


Figure 42: Comparison of the homology model of JBM using either bLAM (pdb: 1O7D) as template (A) or dGMII (pdb: 1HW) as template (B). The applied colouring scheme (coloured by RMSD) is the same as already described above (see Figure 39).

As expected, the homology model of JBM aligns very well with the bLAM crystal structure (see Figure 42 (A)). A superposition of both structures yielded a RMSD value of 0.40 Å. Superimposing the postulated JBM structure with dGMII resulted in a significantly higher RMSD value of 1.44 Å. Furthermore, larger deviations from a good alignment in several parts of the protein are obtained (see Figure 42 (B)) when using dGMII as template. The general quality of the homology model with bLAM as template is indicated by the quality of a Ramachandran plot performed in MOE (see Figure 43).

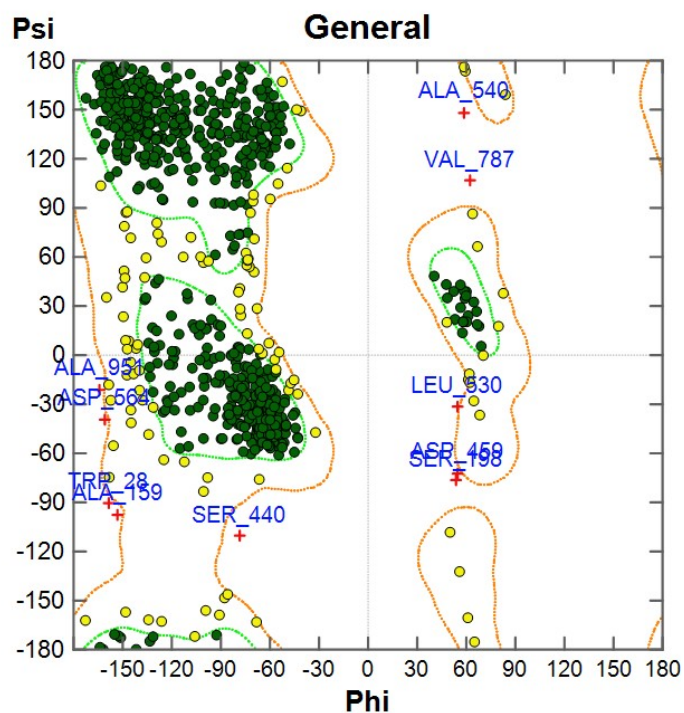


Figure 43: Ramachandran plot for the conformational analysis of amino acids as a qualitative assessment of the homology model. Amino acids with favourable psi and phi angles are indicated by green dots and outliers with unfavourable angles are depicted by red crosses.

Twelve outliers were identified: Ala951, Val787, Asp564, Pro562, Ala540, Leu530, Asp459, Ser440, Pro207, Ser198, Ala159 and Trp28. The Ramachandran plot considers the angles Psi and Phi of the protein backbone. In this context, outliers or amino acids with unfavourable angles are also identified. Only a few outliers were observed in the Ramachandran plot. However, the sequence of JBM consists of 959 amino acids of which eight amino acids are unknown [35]. These unknown amino acids are not located within the active site.

4.1.3 Homology model of the sweet almond β -glucosidase

The tertiary structure of the sweet almond β -glucosidase is not yet known [159]. The primary structure consists of 542 amino acids. The protein has a molecular mass of 61.5 kDa [37]. According to previous studies, the enzyme is naturally occurring as a dimer. However, the activity of the monomer has not yet been elucidated [159]. For the development of the homology model of sweet almond β -glucosidase, cyanogenic β -glucosidase from the white clover (pdb code: 1CBG [160]) was used as template, since it was proposed as such by the “search pdb”-function within MOE. The sequence alignment of both structures showed 57.2% identity and 68.6% similarity between both sequences. A superposition of both the template crystal structure and the sweet almond β -glucosidase resulted in a value of 0.99 Å. As shown in Figure 44 (A), the very large portion of blue color indicates a good alignment of both structures. Higher deviations are not observed. In Figure 44 (B), both catalytic glutamates of the structural model of sweet almond β -glucosidase are represented in sticks.

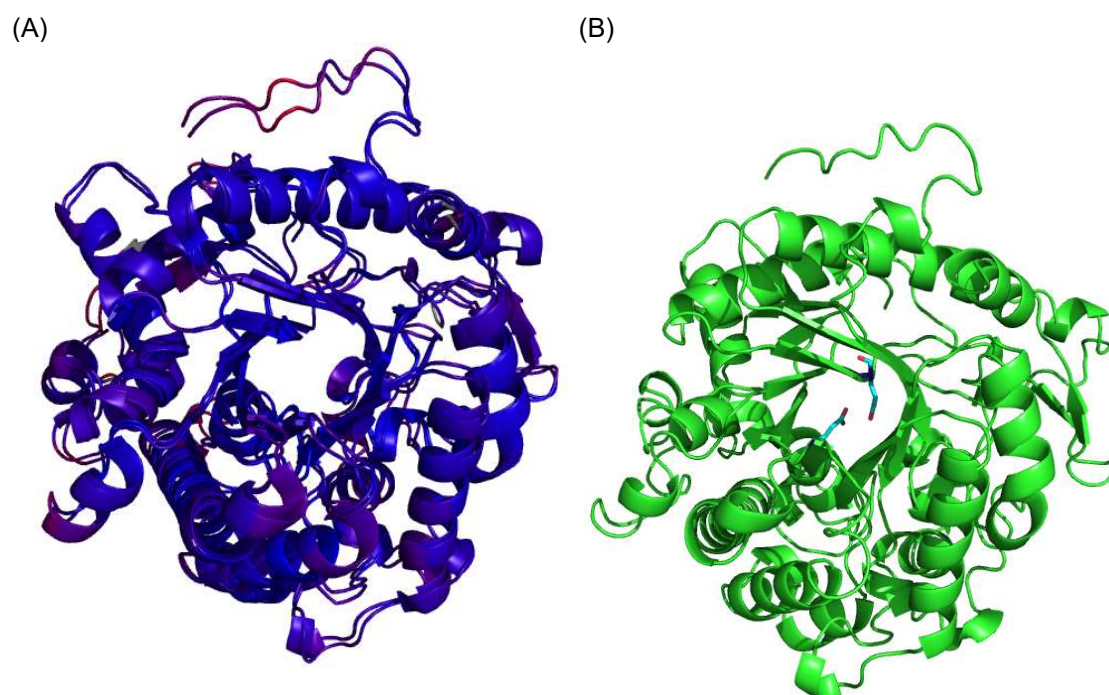


Figure 44: A) Overlay of the homology model of sweet almond β -glucosidase with the cyanogenic β -glucosidase (pdb: 1CBG) as template coloured by RMSD. (B) Structural model of sweet almond β -glucosidase. The two catalytic glutamates are shown as stick representation with carbon atoms coloured in sky blue.

In the Ramachandran plot (see Figure 45), Serin 266 was detected to be a outlier.

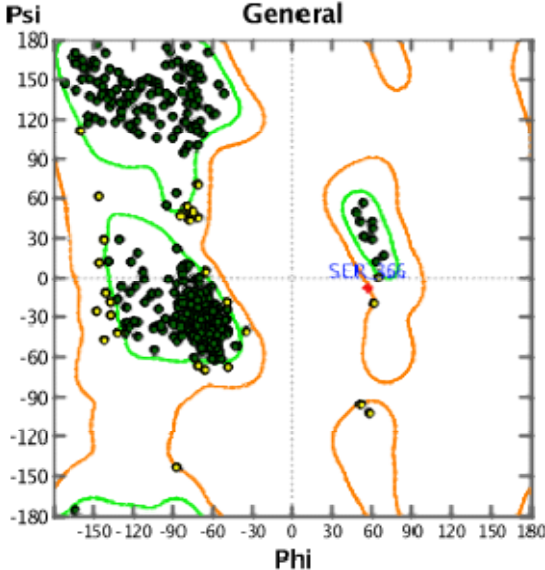


Figure 45: Ramachandran plot for the conformational analysis of amino acids as a qualitative assessment of the homology model. Amino acids with favourable psi and phi angles are indicated by green dots and outliers with unfavourable angles are depicted by red crosses.

4.2 Virtual ligand screening of the catalytic site of hGMII

A small set of compounds was screened to find potential active-site directed inhibitors of hGMII using the software FlexX (LeadIT 2.1.6.). Promising candidates were selected and were tested in enzyme assays in order to confirm the docking predictions.

4.2.1 Docking of swainsonine for testing the receptor setup

Swainsonine is a known inhibitor of hGMII, dGMII and JBM as described in section 1.5. This compound exhibited inhibition in the two-digit nM range in the fluorometric assay (see sections 4.5.2.1.2 and 4.5.2.2.2). When docked into the catalytic site of the homology model of hGMII, swainsonine had a score of -20.1 kJ/mol. The HYDE score was also considerably negative (-44 kJ/mol). Based on this, the calculated ligand efficiency was 0.88 kJ/mol/HA. Furthermore, the Hyde assessment diagram also indicated the inhibitory activity to be in the nanomolar range (see Figure 46):

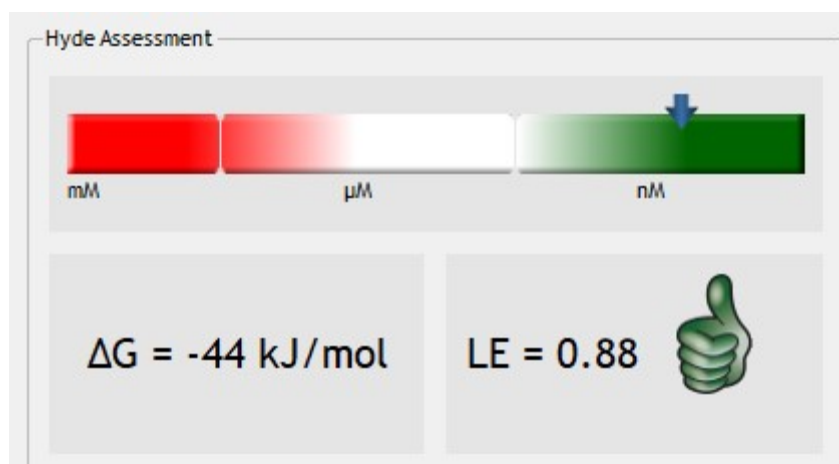


Figure 46: Hyde assessment based on the predicted binding mode of swainsonine bound to the homology model of hGMII. This figure was extracted from LeadIT 2.1.6.

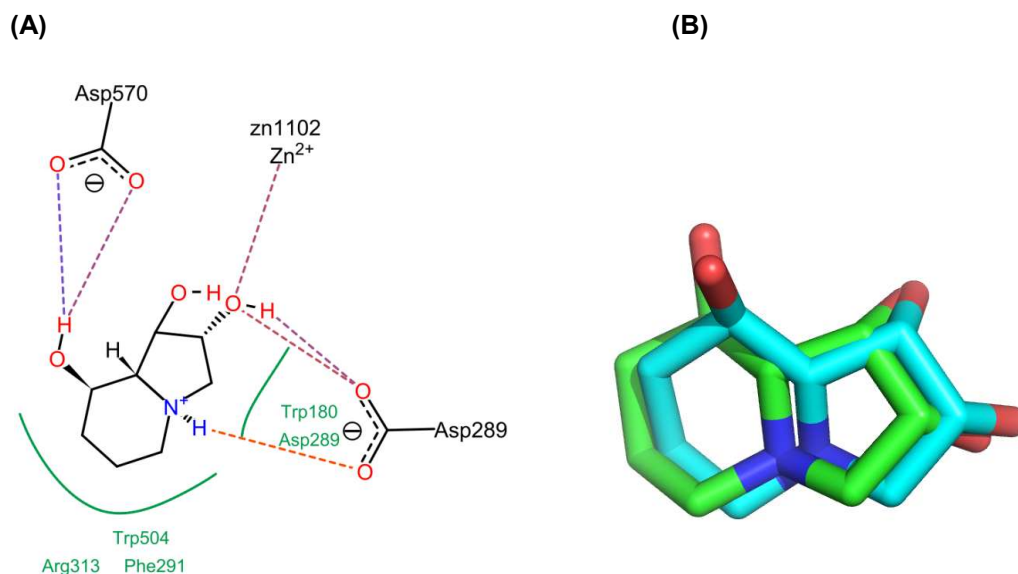


Figure 47: (A) Predicted binding mode of swainsonine within the receptor binding site. The ionic, metal- and hydrogen bond-interactions are displayed by differently coloured dashed lined, whereas hydrophobic interactions are displayed as green curves. (B) Overlay of the predicted binding mode of swainsonine to the homology model(cyan carbon atoms) and crystallographically determined binding mode (pdb: 1HWW; green carbon atoms)

The predicted binding mode of swainsonine bound to the homology model of hGMII closely resembles the crystallographically determined binding mode within the dGMII catalytic site (see Figure 47 (B)). An overlay of both structures yielded a RMSD value of 0.8 Å.

4.2.2 Filtering the database

The small database of about 10,000 compounds was filtered for desired properties as described in section 3.4.3. The compounds were filtered for drug-like properties. After this filter step, 7273 molecules of the database remained for the pharmacophore search. The application of these criteria helped to find compounds that follow the "Rule of five" for orally active drugs [161].

4.2.3 The pharmacophore model

The pharmacophore model was based on the postulated binding mode of swainsonine to hGMII. First, an excluded volume was determined covering all receptor atoms to avoid conformations passing the pharmacophore search which potentially cause steric clashes. Six

other pharmacophore features were also included (see Figure 48). The area covered by the six-membered ring of the indolizidine heterocycle from swainsonine was defined as hydrophobic and/or aromatic region in the size of a sphere with a radius of 1.5 Å. In an area adjacent to this hydrophobic region, a hydrogen bond donor function was determined with a radius of 1 Å. This corresponds to the area covered by the basic nitrogen atom of the heterocycle. The metal-ligand coordination between the hydroxyl groups of the five-membered ring of swainsonine and Zn^{2+} was defined by two spheres with a radius of 1 Å each. To further define the metal-ligand coordination, another sphere corresponding to the acceptor position of the projection was defined around the Zn^{2+} ion with a radius of 1 Å. The sphere corresponding to the hydroxy function of the six-membered ring of swainsonine had a radius of 1 Å and was determined as being either a hydrogen bond donor or a hydrogen bond acceptor function. The two features associated with the Zn^{2+} coordination were defined as being essential. In order to pass the pharmacophore filter, four out of the seven features (incl. the excluded volume) had to be met during the pharmacophore search.

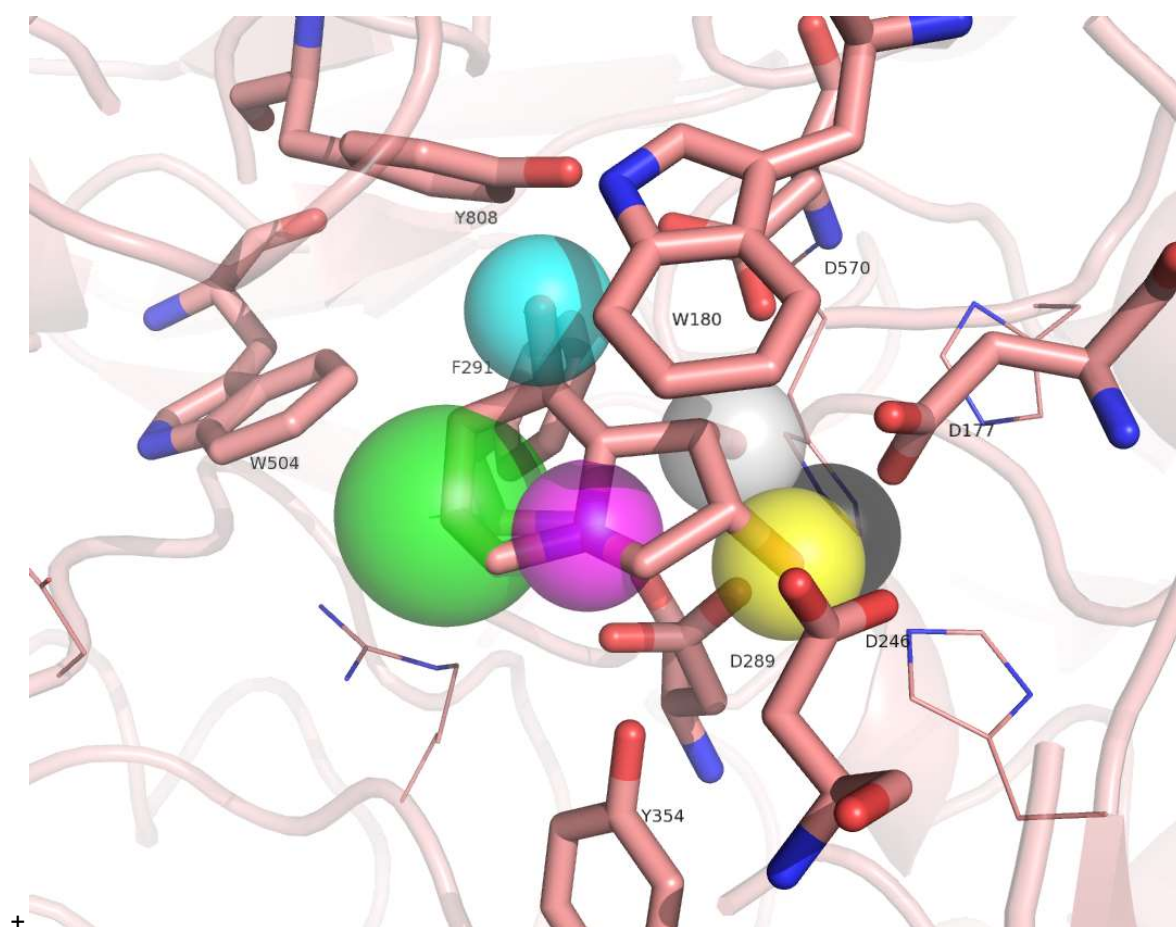


Figure 48: Pharmacophore hypothesis represented by colour-coded spheres. The volumes of the spheres define the positions and sizes of the corresponding electronic and steric features in matching ligand conformations. Binding site residues are displayed with salmon carbon atoms. The green sphere indicates an area capable of mediating hydrophobic interactions with W504 and F291. The magenta sphere represents the placement of a hydrogen bond donor. The region for a hydrogen-bond donor interacting with Y808 and D570 is depicted by a cyan sphere. The desired positions for Zn^{2+} -interactions are shown in grey and yellow spheres, respectively. The black sphere corresponds to the acceptor region of the metal-ligand coordination.

16 621 low energy conformations of compounds passing the first filter step were calculated. A total of 539 molecules passed the second filter step. These molecules were then docked into the catalytic site of hGMII.

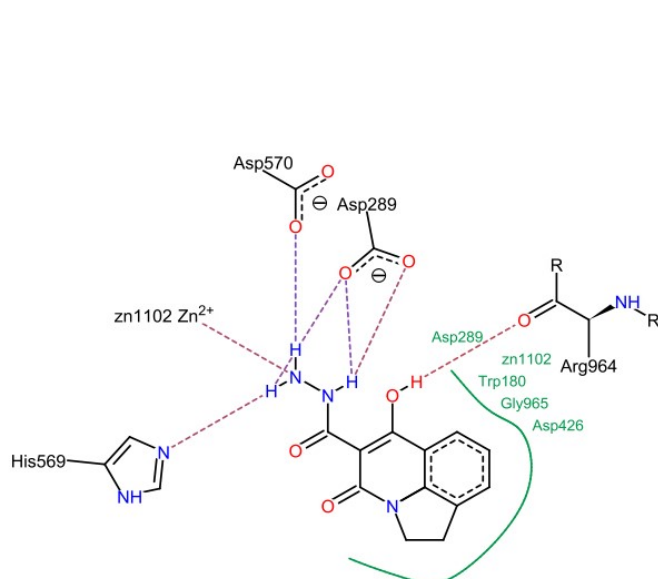
4.2.4 Screening of a small compound library

A total of 319 compounds were successfully docked into the binding pocket of hGMII. The most promising candidates were chosen according to the following criteria:

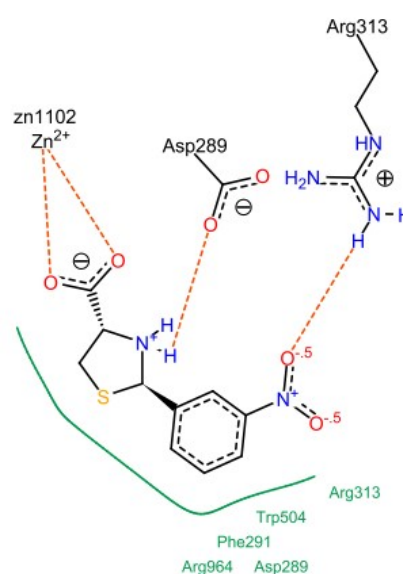
- predicted docking poses with scores ≤ -21 kJ/Mol
- good shape complementarity to the binding pocket
- at least one interaction with the Zn^{2+} -ion
- hydrogen bonding and/or ionic interactions to aspartate residues including Asp289, Asp570, Asp177 and Asp246 predicted
- contribution of hydrophobic interactions involving aromatic residues Trp180, Phe291, Trp504 and Tyr808

Based on these selection criteria, nine compounds were purchased from Sigma Aldrich. It was possible to obtain the correct stereoisomers that had been considered during the docking procedure. On the basis of four selected compounds (Figure 49), the choice of the purchased compounds is exemplified.

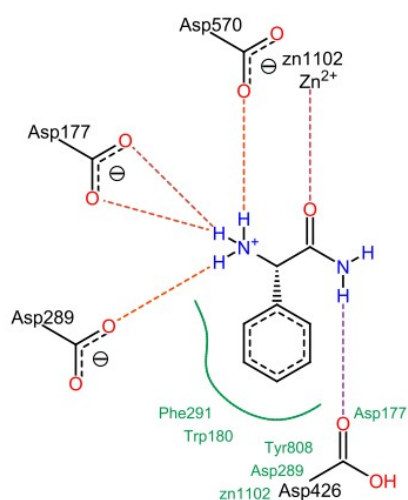
(a) LiSA 24



(b) LiSA 17



(c) LiSA 22



(d) LiSA 1

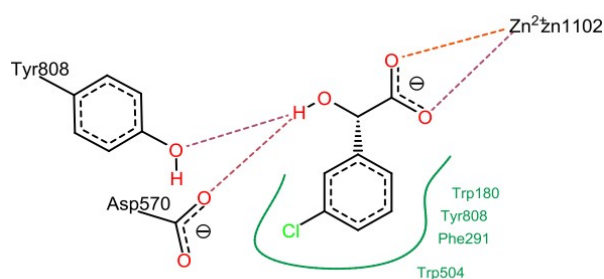


Figure 49: The predicted binding modes of four selected virtual screening hits from the virtual screening against hGMII. Predicted interactions are indicated as differently coloured dashed lines. The hydrophobic interactions are indicated by the green solid lines (a) LiSA 24 (b) LiSA 17 (c) LiSA 22 (d) LiSA 1. The PoseView images were extracted from Leadit (FlexX 2.1.6).

The common basis of these compounds is the predicted interaction with Zn^{2+} mediated via different Zn-binding groups (carboxylate, carbonyl). In addition, all compounds feature hydrophobic moieties, which are predicted to interact with active site residues of hGMII. The carbohydrazide moiety of LiSA 24 potentially forms hydrogen bonds with Asp570, Asp289 and His569. According to the postulated binding mode, the hydroxy group of the heterocycle forms an interaction with the carbonyl backbone of Arg964. Furthermore, π - π -stacking to Trp180 was predicted for LiSA 24. In the predicted binding mode of LiSA 17, an ionic interaction is observed between the nitro group and the guanidino group of Arg313. In addition, the predicted pose of LiSA 17 showed a further ionic interaction between Asp 289 and the positively charged nitrogen atom of the thiazolidine moiety. LiSA 22 is predicted to form several hydrogen bonds and ionic interactions with the receptor atoms of the catalytic site. For example, the protonated amino function of the phenylacetamide moiety might form ionic interactions with carboxylate groups of Asp289, Asp177 and Asp570, respectively. Hydrophobic interactions are postulated between LiSA 22 and Phe291, Trp190 and Tyr808 in addition. Furthermore, this compound displays a good shape complementarity to the catalytic site. LiSA 1 potentially also shows a good shape complementarity to the catalytic site. For LiSA 1 hydrophobic interactions are predicted between the phenyl group and Trp180, Tyr808 and Phe291. Moreover, hydrogen bonds are postulated between the hydroxy group and Tyr808 and Asp570.

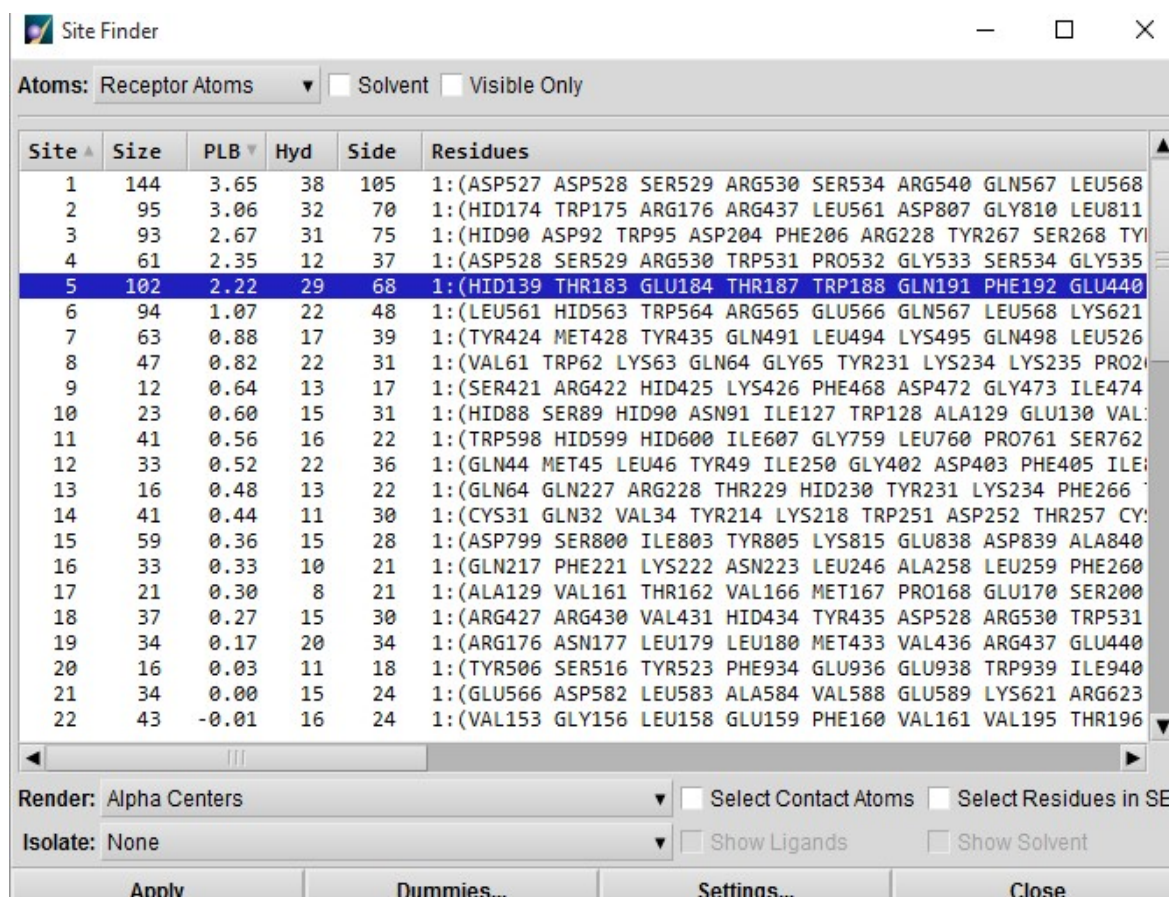
Generally, the presented docking hits are larger in molecular weight than swainsonine and thus, the ligand atoms occupy the binding site almost entirely.

The experimental testing of the virtual screening hits is described in section 3.7.1.1.1. The virtual screening hits are listed together with their scores and inhibition data in section 4.5.1.1.1.

4.3 Virtual ligand screening of the potential allosteric site of dGMII

4.3.1 Identification of a potential allosteric binding site of dGMII

The potential allosteric binding pocket of the GMII protein was identified by using the Site Finder application of MOE v2014.10. A total of 60 cavities were identified in the protein structure (see Figure 50) (PDB Code: 1HWW [162]).



Site	Size	PLB	Hyd	Side	Residues
1	144	3.65	38	105	1:(ASP527 ASP528 SER529 ARG530 SER534 ARG540 GLN567 LEU568
2	95	3.06	32	70	1:(HID174 TRP175 ARG176 ARG437 LEU561 ASP807 GLY810 LEU811
3	93	2.67	31	75	1:(HID90 ASP92 TRP95 ASP204 PHE206 ARG228 TYR267 SER268 TYR
4	61	2.35	12	37	1:(ASP528 SER529 ARG530 TRP531 PRO532 GLY533 SER534 GLY535
5	102	2.22	29	68	1:(HID139 THR183 GLU184 THR187 TRP188 GLN191 PHE192 GLU440
6	94	1.07	22	48	1:(LEU561 HID563 TRP564 ARG565 GLU566 GLN567 LEU568 LYS621
7	63	0.88	17	39	1:(TYR424 MET428 TYR435 GLN491 LEU494 LYS495 GLN498 LEU526
8	47	0.82	22	31	1:(VAL61 TRP62 LYS63 GLN64 GLY65 TYR231 LYS234 LYS235 PRO2
9	12	0.64	13	17	1:(SER421 ARG422 HID425 LYS426 PHE468 ASP472 GLY473 ILE474
10	23	0.60	15	31	1:(HID88 SER89 HID90 ASN91 ILE127 TRP128 ALA129 GLU130 VAL
11	41	0.56	16	22	1:(TRP598 HID599 HID600 ILE607 GLY759 LEU760 PRO761 SER762
12	33	0.52	22	36	1:(GLN44 MET45 LEU46 TYR49 ILE250 GLY402 ASP403 PHE405 ILE
13	16	0.48	13	22	1:(GLN64 GLN227 ARG228 THR229 HID230 TYR231 LYS234 PHE266
14	41	0.44	11	30	1:(CYS31 GLN32 VAL34 TYR214 LYS218 TRP251 ASP252 THR257 CY
15	59	0.36	15	28	1:(ASP799 SER800 ILE803 TYR805 LYS815 GLU838 ASP839 ALA840
16	33	0.33	10	21	1:(GLN217 PHE221 LYS222 ASN223 LEU246 ALA258 LEU259 PHE260
17	21	0.30	8	21	1:(ALA129 VAL161 THR162 VAL166 MET167 PRO168 GLU170 SER200
18	37	0.27	15	30	1:(ARG427 ARG430 VAL431 HID434 TYR435 ASP528 ARG530 TRP531
19	34	0.17	20	34	1:(ARG176 ASN177 LEU179 LEU180 MET433 VAL436 ARG437 GLU440
20	16	0.03	11	18	1:(TYR506 SER516 TYR523 PHE934 GLU936 GLU938 TRP939 ILE940
21	34	0.00	15	24	1:(GLU566 ASP582 LEU583 ALA584 VAL588 GLU589 LYS621 ARG623
22	43	-0.01	16	24	1:(VAL153 GLY156 LEU158 GLU159 PHE160 VAL161 VAL195 THR196

Figure 50: The identified sites are listed with following properties: the size, the PLB (propensity ligand binding), the number of hydrophobic contacts and the number of side chain contacts [163],[164].

After visual inspection of the most promising sites with regard to the size, the amino acid composition, the accessibility and the distance of the proposed site to the active site, one site has been chosen as potential allosteric site. This site consists of 102 receptor atoms, of which 29 are hydrophobic and 68 are side chain atoms. In contrast, the catalytic site is composed of 93 receptor atoms, of which 31 are hydrophobic and 75 are side chain atoms. Approximately 30 Å separate the active site from the potential binding site when measured from the center of each site (see Figure 51). A comparison of the amino acid composition revealed some differences: the known catalytic site is compact, hydrophilic and

predominantly negatively charged in the region where potential ligands are deeply buried (Figure 52 (A)). The potential allosteric site seems to be more open compared to the active site. Furthermore, the amino acid composition is more balanced in terms of the distribution of negatively and positively charged residues (Figure 52 (B)).

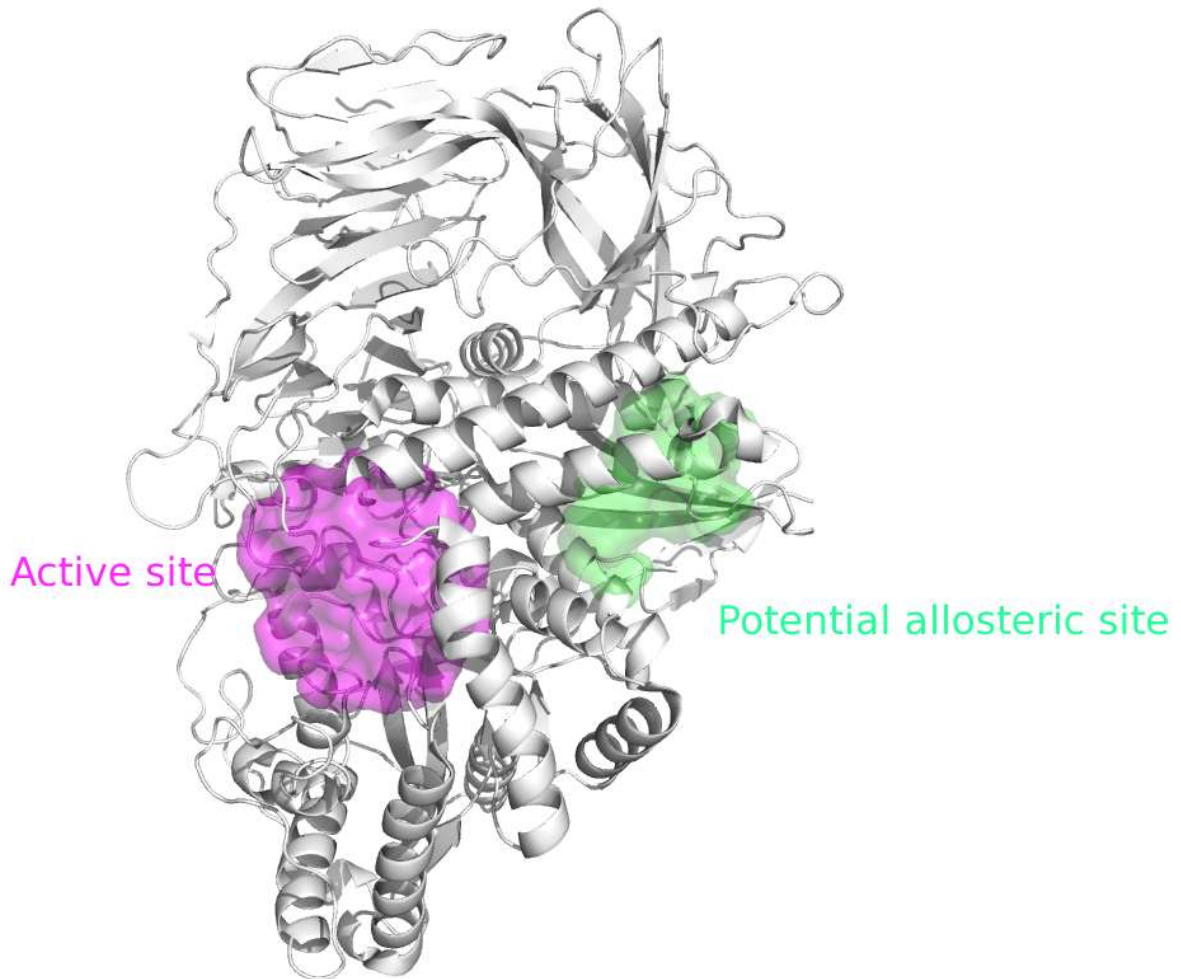
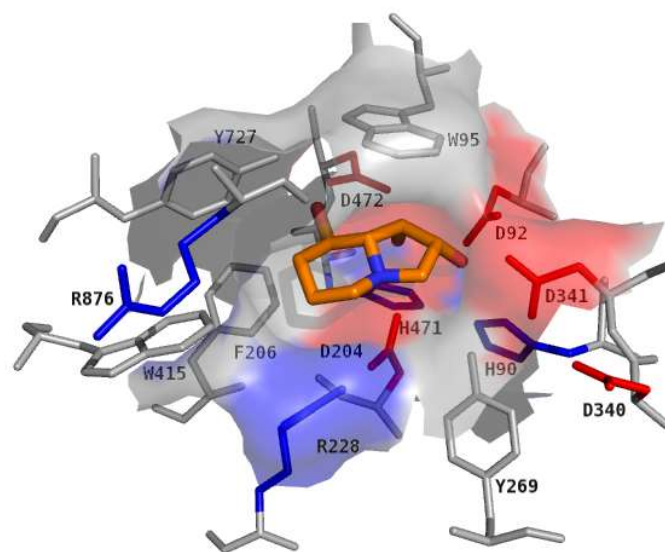


Figure 51: Active (magenta) and potential allosteric (green) binding site within dGMII

(A)



(B)

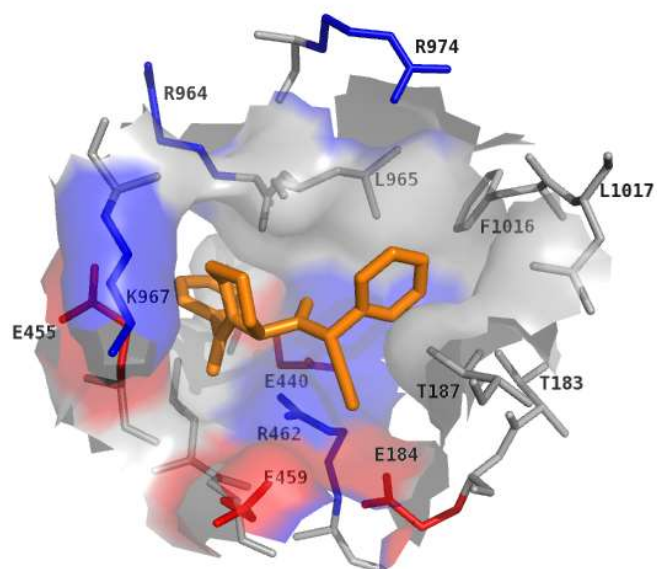


Figure S2: Color-coded surfaces around the active site with the co-crystallized ligand swainsonine depicted with ocre carbon atoms (A) and the identified pocket with virtual screening hit displayed by ocre-coloured sticks (B). The red surface represents regions with acidic residues, blue surface regions indicate basic residues, and the grey surface corresponds to polar and nonpolar residues and the backbone.

The active site has a largely negatively charged surface due to the acidic residues in its pocket (Figure 52 (A)). Amino acids such as D204, D340, D92 or D472 may accordingly interact with positively charged binding partners. The GMII is a metalloenzyme and contains a zinc ion in its active site. Noticeably, as with all zinc sites, the metal ion is enclosed by a shell of hydrophilic groups which is surrounded from a larger shell of hydrophobic groups [165]. On the other hand, the cavity of the potential allosteric site has a positively charged and a negatively charged surface patch. The positively charged patch is formed by the residues R462, R463, R963 or K967 (Figure 52 (B)), whereas amino acid residues such as E459, E184 or E440 build the negatively charged surface. In all other regions of the potential allosteric site neutral amino acids are predominant.

4.3.2 Comparison of binding site properties- active site versus potential allosteric site

The druggability of the potential allosteric binding site was evaluated using the DoGSite Scorer by considering the following properties: pocket volume, surface area, depth, druggability score and the apolar amino acid ratio (see Table 6).

Table 6: Druggability prediction of the active and the potential allosteric site of GMII using DoGSiteScorer (input pdb: 1HWW).

	Volume [\AA^3]	Surface [\AA^2]	Depth [\AA]	Druggability score	Apolar amino acid ratio
Active site	194.65	229.38	8.97	0.28	0.21
Potential allosteric site	879.26	820.14	22.46	0.83	0.46

The DoGSite Scorer predicted a high druggability for the potential allosteric site, since scores > 0.5 indicate a druggable binding site [143]. Furthermore, the properties of both sites differed significantly, reconfirming what has already been partly described above (see Figure 52). The larger pocket volume, greater depth, and a higher apolar amino acid ratio of the potential allosteric site are positive predictors of its druggability [143]. Thus, the potential allosteric site could potentially bind drug-like compounds.

4.3.3 Pharmacophore hypothesis

In the absence of a known ligand for this potential allosteric binding pocket, the pharmacophore model was built based on its amino acid composition. Here, a special focus was set on charged amino acids buried inside the potential binding pocket to facilitate favourable electrostatic interactions between ligands and protein. The pharmacophore model included an excluded volume covering all receptor atoms to prevent a steric clash upon binding. Additionally, six other pharmacophore features were included in the model (see Figure 53). Two pharmacophore features were included to achieve interactions with E459. On the one hand side, a hydrogen-bond donor or a cationic functionality was required close to the glutamate within a radius of 2.4 Å. On the other hand, the counterpart of this interaction, the carboxylic acid function of E459, was determined as a hydrogen-bond acceptor and had a radius of 1.7 Å. Further, a hydrophobic moiety was required close to the side chains of L965 and F1016. The hydrophobic feature is given a tolerance radius of 2.1 Å. Another pharmacophore feature was located between E459 and R462 to either represent a hydrogen-bond donor function interacting with E459 or a hydrogen bond acceptor interacting with R462. An additional feature determined an area of a hydrogen bond acceptor that may interact with the protonated side chain of K967 and the last feature represented the position of an anionic and hydrogen bond-accepting group presumably interacting with R462. The remaining 3 features all had radii of 1.0 Å. In addition to the excluded volume area, the features accounting for the interaction of proposed ligands with E459 via either hydrogen bonding or electrostatic interactions and those representing the interaction network to E459 and/or R462, respectively, were determined as being mandatory during the pharmacophore search. Ligand conformations could only pass the pharmacophore search, if six out of seven features (with excluded volume) were matched.

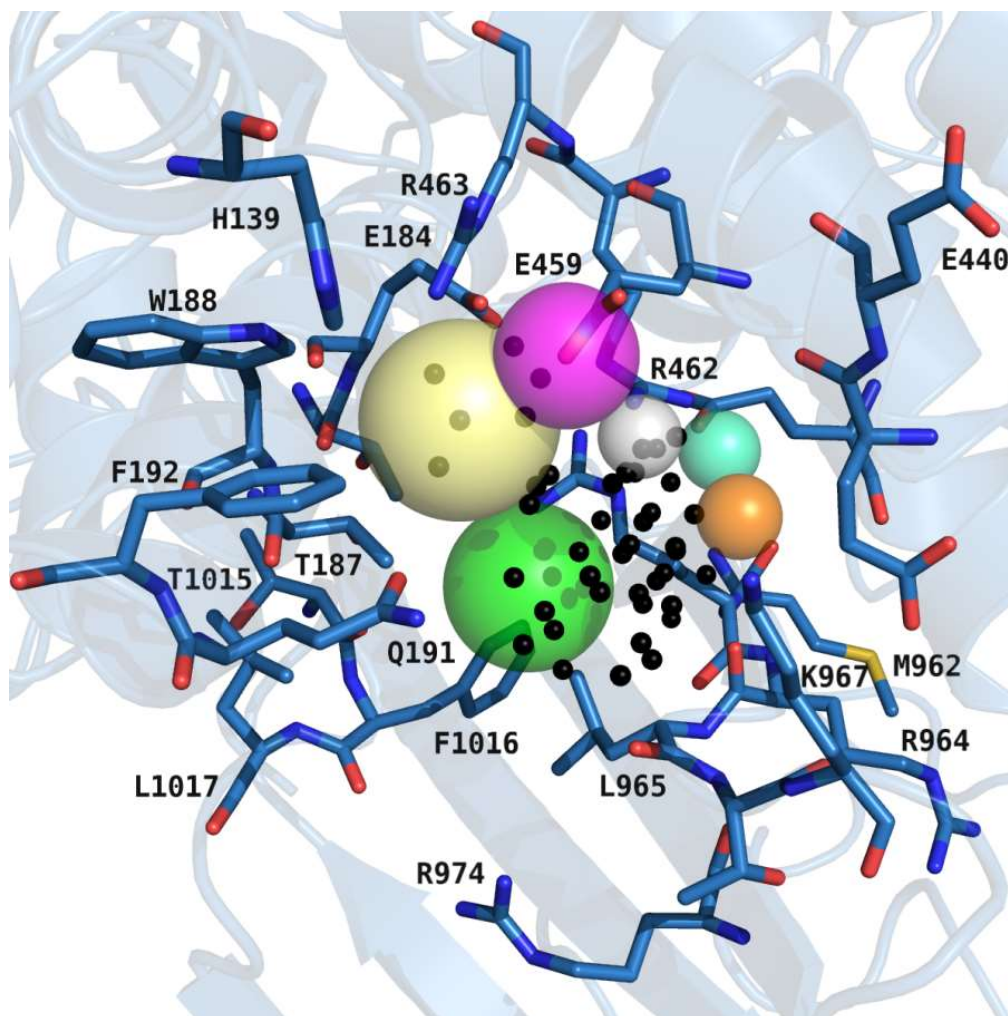


Figure 53: Pharmacophore hypothesis represented by colour-coded spheres. The volumes of the spheres define the positions and sizes of the corresponding electronic and steric features in matching ligand conformations. Binding site residues are displayed with carbon atoms in sky-blue colour. Dummy atoms used to define the binding site region are indicated by black spheres. The pale-yellow and magenta spheres, respectively, indicate the placement of a hydrogen-bond donor or cationic centre to interact with E459. The green sphere represents the area that had to be addressed with hydrophobic substituents. The desired position for either a hydrogen-bond donor or acceptor interacting with E459 and R462, respectively, is indicated by a grey sphere. The anionic centre with hydrogen bond acceptor properties is depicted by a pale-cyan sphere, whereas the orange sphere accounts for a hydrogen bond acceptor possibly interacting with K967.

Several million conformations of the > 600,000 molecules passing the first filter step for molecular properties were calculated. However, only the conformation with the lowest rmsd-value between the matching atom positions and the pharmacophore query points was kept in the results table. Altogether, 41,815 molecules passed this second filter step. Subsequently, these molecules were subjected to docking into the potential allosteric binding site.

4.3.4 Molecular docking

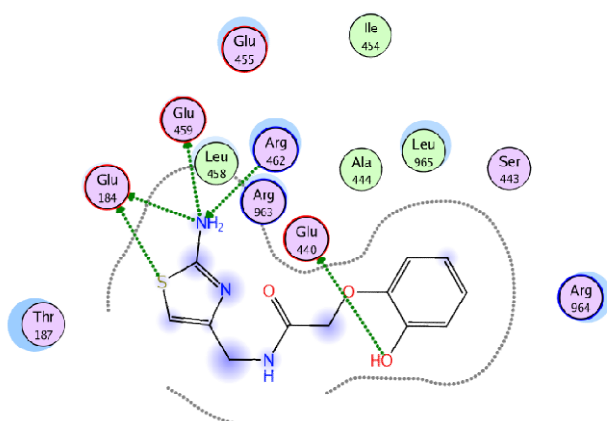
For 36,607 compounds, a docking pose was obtained. Compounds were selected for purchase and the subsequent determination of the binding affinity according to the following criteria:

- total docking score (ΔG) predicted to be ≤ -25 kJ/mol
- ligand efficiency (LE) (as defined by Equation 14) ≥ 1 kJ/mol per heavy atom
- and visual inspection of the predicted pose including an accurate analysis of the postulated protein-ligand interactions by considering the following requirements: hydrophobic parts should not be solvent-exposed and strong polar ligand groups were required to interact with at least one binding site residue via electrostatic and/or ion interactions (particularly with E459 und R462).

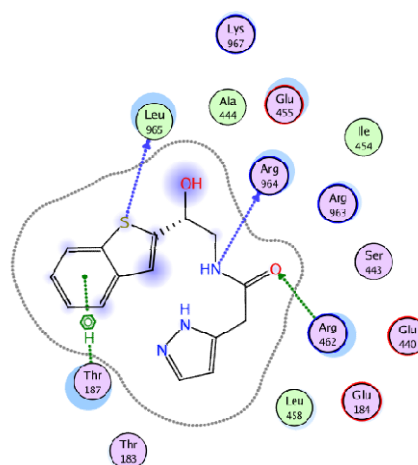
In total, 27 out of the 36,607 compounds were selected for testing. These 27 compounds were purchased from ChemBridge and Enamine and tested using a fluorometric enzyme assay (see section 4.5.2.1.7). For the compounds with stereoisomerism, it was only possible to obtain racemic mixtures, although stereoisomers had been taken into account for the molecular docking.

On the basis of four selected compounds (see Figure 54), the choice of the purchased compounds is exemplified.

(a) Compound 11



(b) Compound 12



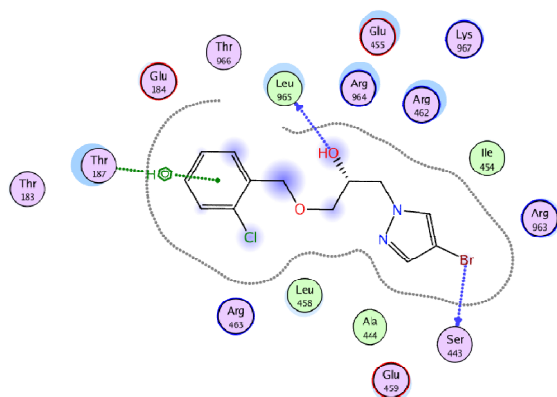
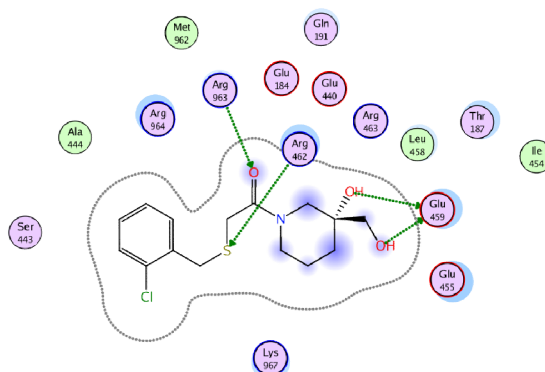
(c) Compound **13**(d) Compound **20**

Figure 54: The predicted binding modes for four selected hits from the virtual screening against dGMII. Predicted interactions are indicated as dashed lines. The hydrophobic interactions are shown in green: (a) Compound **11** (b) Compound **12** (c) Compound **13** (d) Compound **20**. The ligand interaction diagrams are extracted from MOE

Compound **11** is predicted to form several interactions with the amino acid residues of the potential allosteric site. The amino group of the aminothiazole ring potentially forms hydrogen bonds with Glu184, Glu459 and Arg462. Moreover, the phenolic OH-group is postulated to form a hydrogen bond with Glu440. The benzene ring of the benzothiazophene bicycle of compound **12** presumably forms π -H-interactions to Thr187. The nitrogen atom of the amide group might form a hydrogen bond to the backbone of Arg964. The predicted binding mode of compound **13** is also interesting. Here, the secondary hydroxyl group might donate a hydrogen bond to the backbone carbonyl oxygen of Leu965. Furthermore, a π -H-interaction to Thr187 is also predicted for compound **13**. Compound **20** is also forming different interactions: hydrogen bonds with Arg462, Arg963 and Glu459. The carbonyl oxygen of compound **20** possibly acts as a hydrogen bond acceptor by interacting with Arg963. The hydroxyl groups might donate a hydrogen bond to Glu459. The sulfur atom is predicted to act as hydrogen bond acceptor to Arg462.

4.4 Virtual ligand screening for ligands of the catalytic site of dGMII

A large database was screened against the catalytic site of dGMII in order to find new inhibitors. The docking procedure was carried out using the software DOCK. The molecular docking approach was first validated by redocking the known ligand swainsonine into the catalytic site. The most promising compounds were selected by different criteria and were purchased for experimental testing. This approach differs from the virtual screening described in section 3.4 with respect to the following points: the docking software used, the size of the compound database, and the choice of the receptor structure.

4.4.1 Redocking of swainsonine

In order to assess the accuracy of binding mode predictions, swainsonine was docked in the catalytic site of dGMII first. As already described in section 1.5, swainsonine exhibits inhibition in the two-digit nanomolar range. Docking of swainsonine into the catalytic binding site of dGMII resulted in a negative score of -12.62 kJ/Mol. A RMSD-value of 1.099 between the co-crystallized ligand and the docking pose was determined indicating a successful pose prediction. Depending on the ligand size, it is generally assumed that RMSD-values below 2 Å are indicative for a successful docking performance [166], [167].

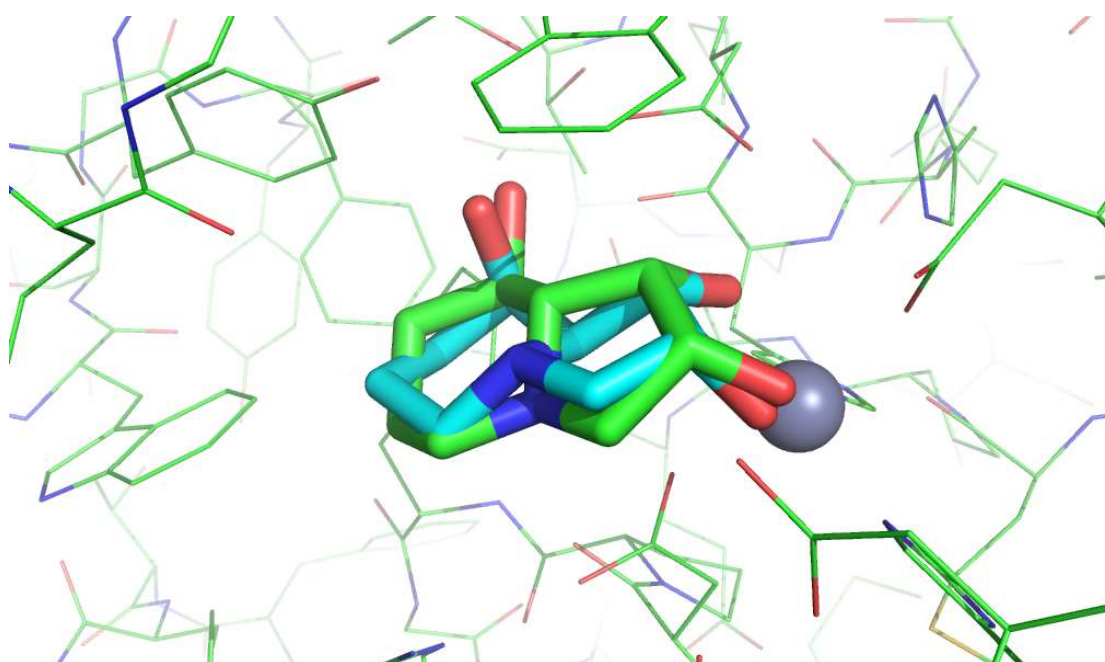


Figure 55: Overlay of the predicted binding mode of swainsonine (cyan-coloured carbon atoms) with its crystallographically determined binding mode (pd : 1HWW; green carbon atoms). Zn²⁺ is displayed as grey sphere.

The predicted binding pose reflects the important interactions between the Zn^{2+} ion with the hydroxyl groups of the five-membered ring of swainsonine.

4.4.2 Filtering the database

The database was filtered as described in section 3.5.4.

4.4.3 The pharmacophore model

The pharmacophore model was built based on the binding mode of swainsonine as it was also done during preparation of the pharmacophore model for hGMII (see Figure 48, section 4.2.3). The six pharmacophore features were selected as described in section 3.4.4.

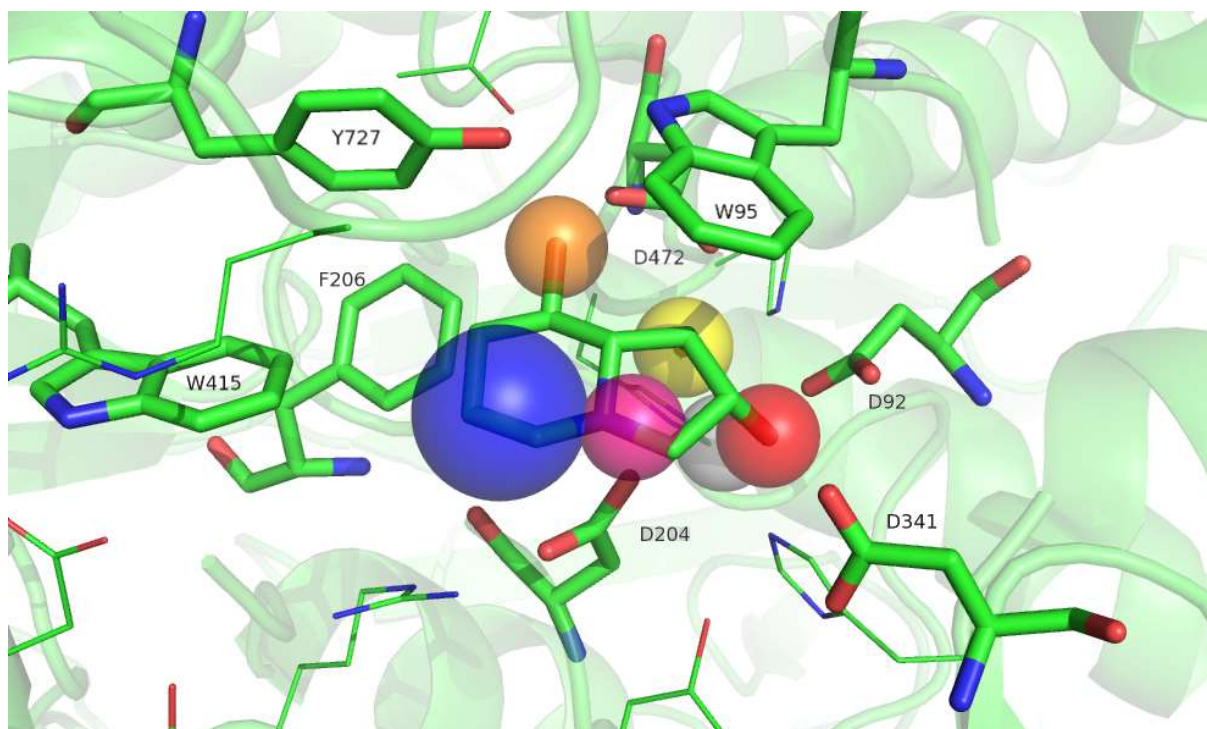


Figure 56: Pharmacophore hypothesis represented by colour-coded spheres. The volumes of the spheres define the positions and sizes of the corresponding electronic and steric features in matching ligand conformations. Binding site residues are displayed with green carbon atoms. The blue sphere indicates an area that had to be addressed with hydrophobic and/or aromatic substituents interacting with W415 and F291. The magenta sphere represents a hydrogen bond donor function. The region for a hydrogen-bond donor interacting with Y727 and D472 is depicted by an orange sphere. The desired positions for interactions with Zn^{2+} interactions are indicated by yellow and red spheres, respectively. The "metal ligator projection" is defined around the Zn^{2+} as a grey sphere.

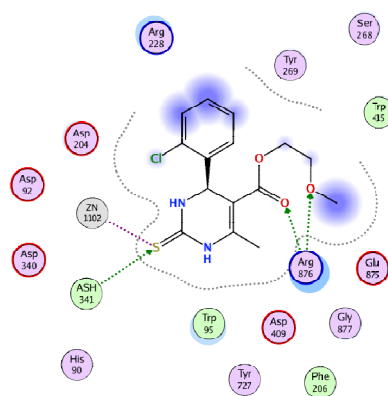
4.4.4 Screening of a large compound library

A total of about 123,300 compounds were docked into the binding pocket of dGMII. The most promising candidates were chosen according to the following criteria:

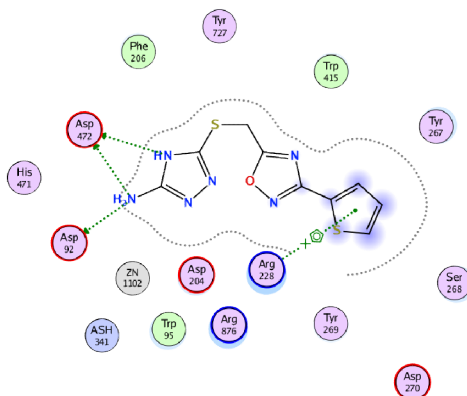
- total docking score (ΔG) predicted to be ≤ -25 kJ/mol
- ligand efficiency (LE, Equation 14) ≥ 1.3 kJ/mol per heavy atom
- good shape complementarity to the binding pocket
- interactions to the central Zn^{2+} -ion predicted
- visual inspection of the predicted pose including an accurate analysis of the reliability of the postulated protein-ligand interaction.
- special emphasis on potential interactions to aspartate residues within the binding pocket (including Asp204, Asp341, Asp92 and Asp472)
- a contribution of hydrophobic interactions involving aromatic residues Trp95, Trp415 and Tyr80

The criteria concerning the docking score and the LE had to be met. Based on these selection criteria, 30 compounds were purchased from Enamine and Chembridge. On the basis of four selected compounds (Figure 57), the choice of the purchased compounds is exemplified.

(A) Compound **33**



(B) Compound **34**



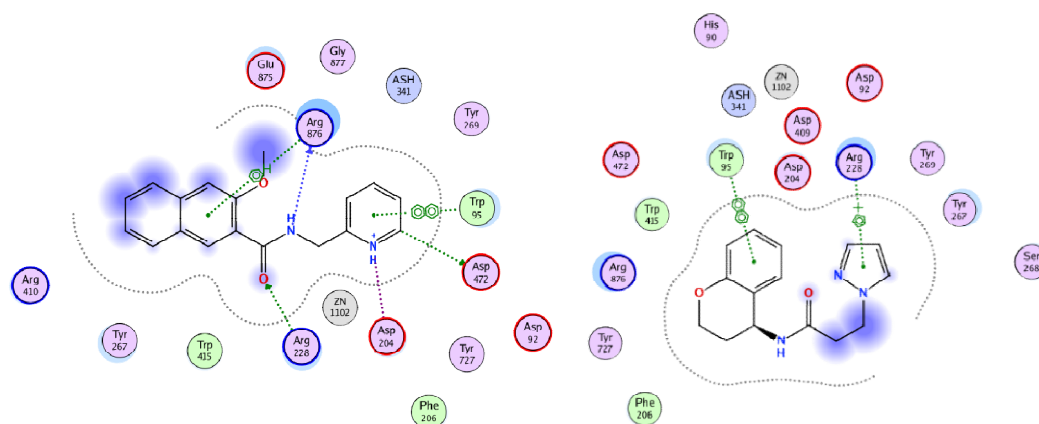
(C) Compound **40**(D) Compound **52**

Figure 57: The predicted binding modes for four selected screening hits from a virtual screening of the dGMII (A) compound **34** (B) compound **33** (C) compound **40** (D) compound **52**

Compound **33** was mainly chosen due to its postulated interaction of the sulfur atom of the tetrahydropyrimidine ring with the Zn^{2+} -ion. The predicted interactions of the ether oxygen and the carbonyl oxygen acting as hydrogen bond acceptors to Arg876 were also a reason for the choice of compound **33**. The amino function of the aminotriazole of compound **34** is predicted to form a hydrogen bond interaction to Asp92 and Asp472. The nitrogen atom on position 4 of the 1,2,4-triazole potentially interacts with Asp472 via an additional hydrogen bond. Furthermore, the thiophene moiety of compound **34** might be involved in an arene-cation interaction with the positively charged side chain of Arg228. The pyridine ring of compound **40** is predicted to form π - π -stacking interactions to Trp95. The 3-methoxy aromatic ring from the naphthylamide moiety of compound **40** possibly forms an arene-H interaction with Arg876. The carbonyl function of compound **40** is predicted to act as hydrogen bond acceptor in an interaction with Arg228. An electrostatic interaction between the protonated nitrogen atom of the pyridine ring and Asp204 is also predicted. The chroman moiety of compound **52** might take part in π - π -stacking interactions to Trp95. On the other hand, an arene-cation interaction is predicted between the pyrazole ring and Arg228. This compound was purchased mainly due to its aromatic interactions. In the choice of the compounds, attention was always paid to the diversity of chemical structures and to the various postulated binding modes.

4.5 Enzyme assays

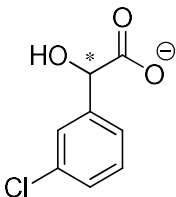
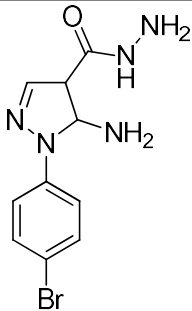
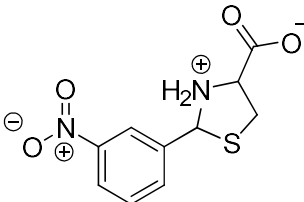
4.5.1 Photometric Assay

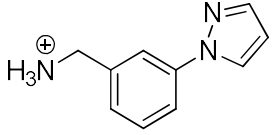
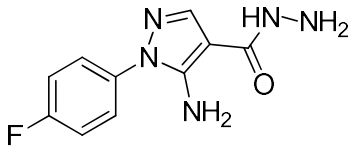
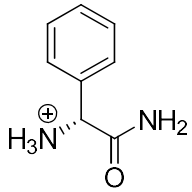
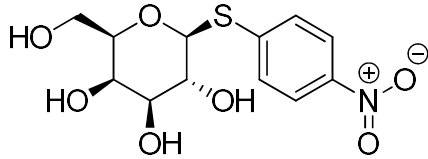
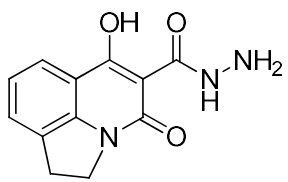
4.5.1.1 Enzyme assays with the Jack Bean Mannosidase

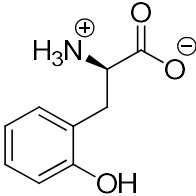
4.5.1.1.1 Experimental testing of the virtual screening hits

In the following table (see Table 7), the virtual screening hits are listed together with their scores and internal compound identifiers. The compounds were dissolved in assay buffer and subsequently tested in the photometric assay of the JBM described in section 3.7.1.1.1. The compounds were all tested at a well concentration of 1 mM.

Table 7: List of score-ranked virtual screening hits with corresponding inhibition data in the enzymatic assay.

Cpd.	Molecular structure	Score [kJ/mol]	Hyde Score [kJ/mol]	Inhibition in enzymatic assay with JBM** [%]
LiSA 1***		-23.5	-13	52 ± 3
LiSA 13		-25.0	-4	No inhibition
LiSA 17		-32.3	-8	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Hyde Score [kJ/mol]	Inhibition in enzymatic assay with JBM** [%]
LiSA 20		-34.4	-6	No Inhibition
LiSA 21		-28.8	-5	No inhibition
LiSA 22		-34.3	-6	No inhibition
LiSA 23		-22.7	-18	No inhibition
LiSA 24		-24.1	-20	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Hyde Score [kJ/mol]	Inhibition in enzymatic assay with JBM** [%]
LiSA 26		-27.4	- 8	No inhibition

**Inhibition (see Equation 15, section 3.7.2.4)

*** The purchased compound was only available as (*R*)-enantiomer (Score: -22.3 kJ/mol)

4.5.2 Fluorometric Assay

4.5.2.1 Enzyme assay with the dGMII

4.5.2.1.1 K_M value determination

The non-linear fit to determine the K_M value of dGMII with 4-MU-Man as substrate is shown in the following graph (see Figure 58):

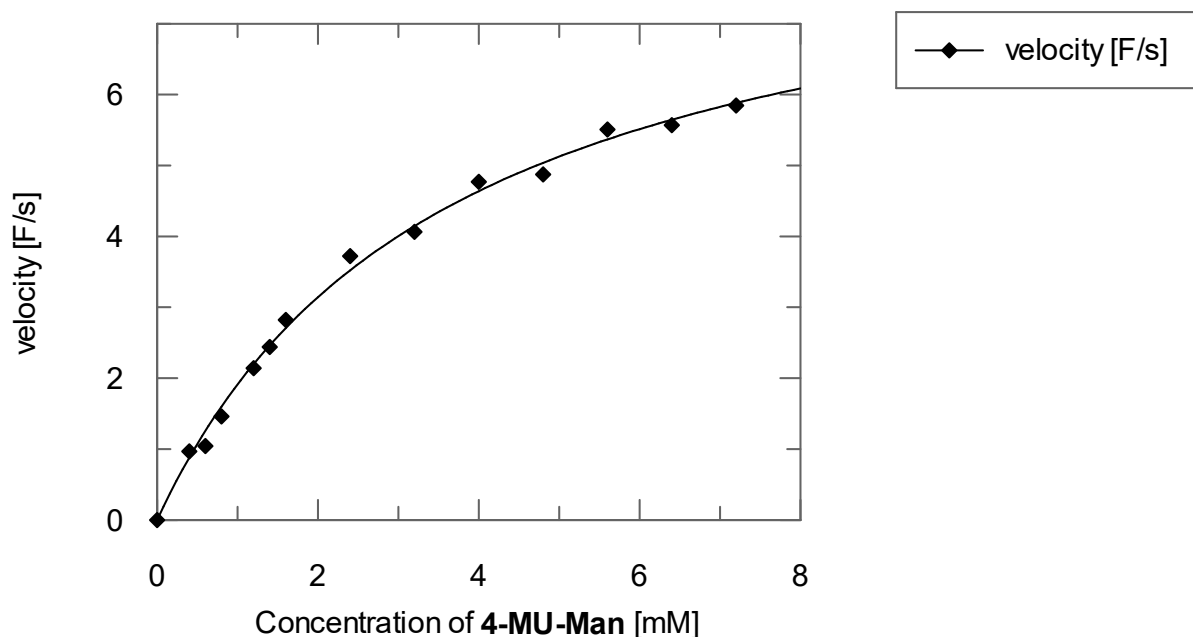


Figure 58: Non-linear fit for the determination of the K_M value of the dGMII

In the following table (Table 8), the determined K_M value and the v_{max} value are listed.

Table 8: K_M and v_{max} values from the non linear regression of the graph from figure 58

K_M	$3.6 \pm 0.3 \text{ mM}$
V_{max}	$8.8 \pm 0.3 \text{ F/s}$

4.5.2.1.2 IC_{50} determination of swainsonine

In order to establish the assay, the known inhibitor swainsonine was used as a reference and inhibition datas such as IC_{50} values and K_i values were determined.

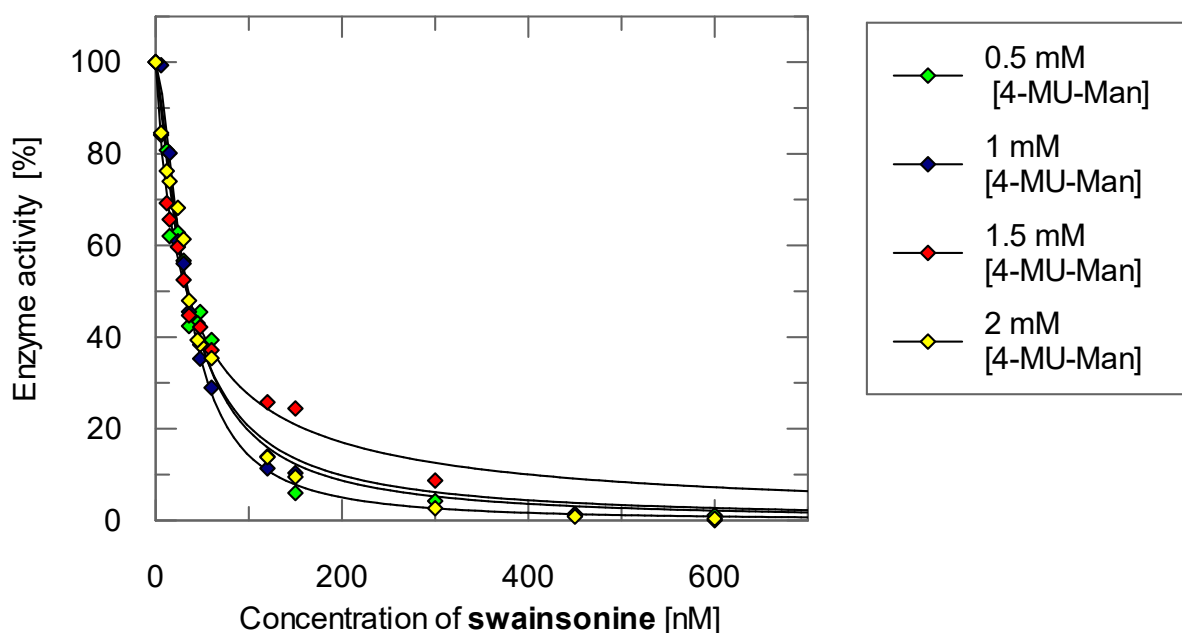


Figure 59: Determination of the IC_{50} value of swainsonine at different substrate concentrations of 4-MU-Man. Non-linear regression was applied to the data by using Equation 20.

The respective IC_{50} values of swainsonine in the presence of different substrate concentrations are listed in Table 9.

Table 9: Effect of different substrate concentration on the IC_{50} values of swainsonine

Concentration of 4-Mu-Man	IC_{50} values
0.5 mM	37 ± 4.3 nM
1.0 mM	34 ± 0.7 nM
1.5 mM	34 ± 1.4 nM
2.0 mM	33 ± 2.8 nM

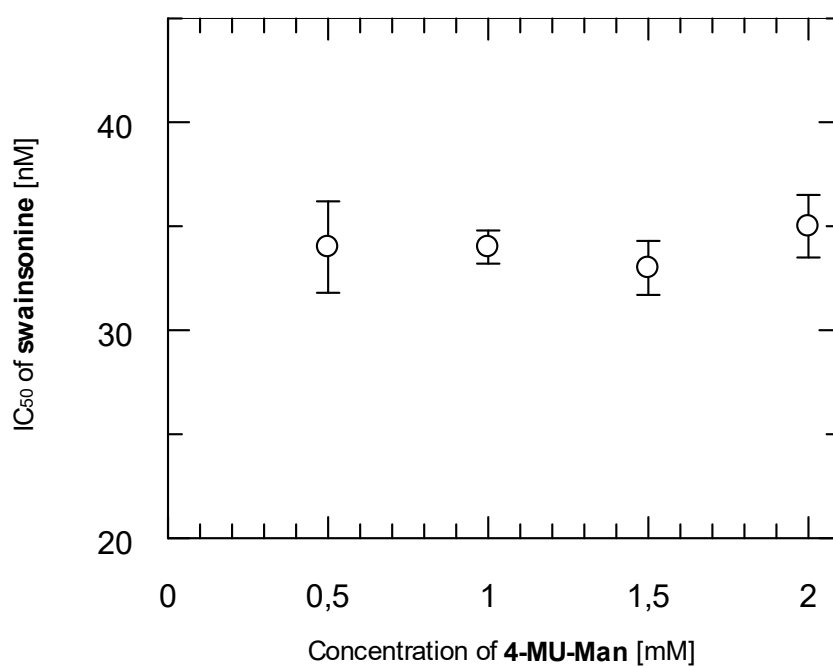


Figure 60: Plot of the determined IC_{50} values for the dGMII inhibitor swainsonine in the presence of different substrate concentrations of 4 MU-Man

4.5.2.1.3 IC₅₀ determination of swainsonine as slow binder

IC₅₀ values of swainsonine were determined at different incubation times (0, 5 and 10 minutes). The IC₅₀ values of swainsonine are listed in Table 10:

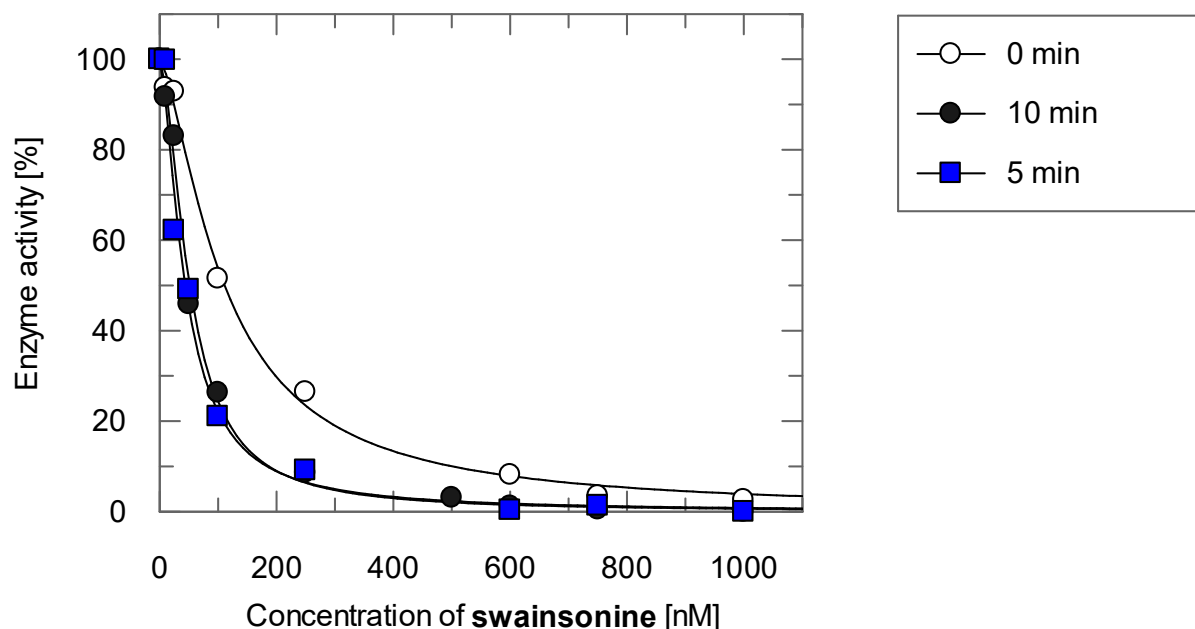


Figure 61: Determination of the IC₅₀ value of swainsonine at a substrate concentration of 1 mM of 4-MU-Man depending on the incubation time (10 min, 5 min and 0 min). Non-linear regression was applied to the data by using Equation 16

Table 10: Effect of different incubation times on the IC₅₀ values of swainsonine

Incubation Time [min]	IC ₅₀ values [nM]
10	51 ± 3
5	44 ± 4
0	111 ± 6

4.5.2.1.4 K_i value determination of swainsonine

In the following figure (see Figure 62), the slopes were plotted against the substrate concentrations. Using the Michaelis-Menten equation, a non-linear regression was applied for both series of measurements. The first series of measurements was carried out in the presence and the other one in the absence of the inhibitor swainsonine. In Table 11, the corresponding K_M values obtained from the Michaelis-Menten plot are reflected.

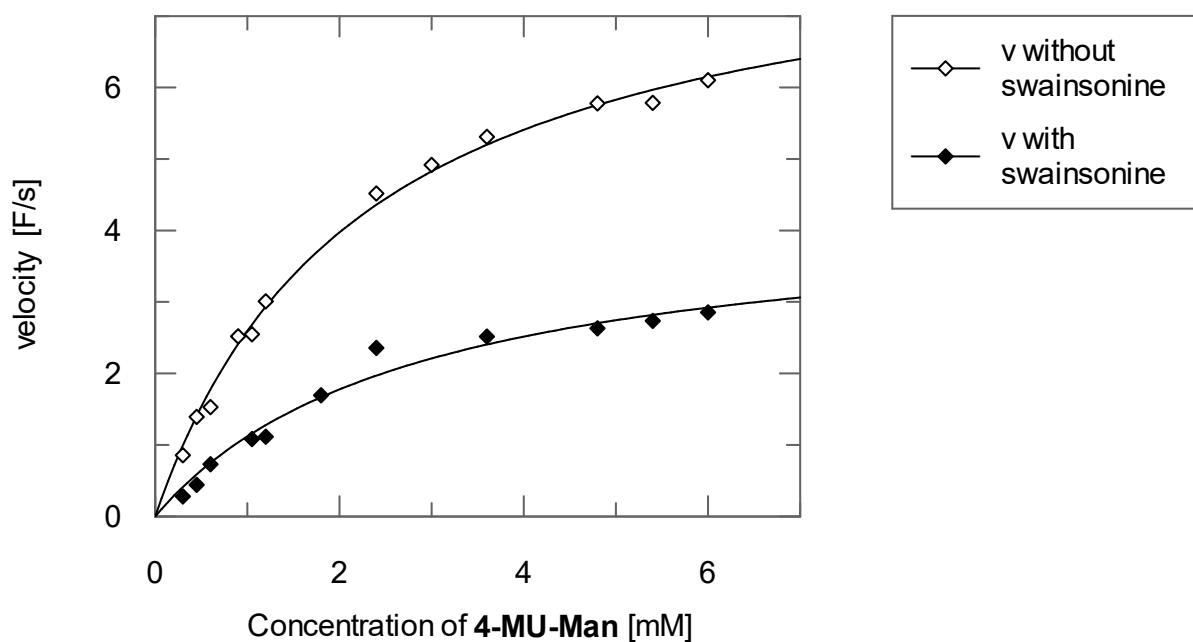


Figure 62: Example of a Michaelis-Menten plot for the determination of K_M and v_{max} values of dGM II with and without swainsonine

Table 11: K_M and v_{max} values of the Michaelis-Menten plot with and without swainsonine

	K_M (K_M' with swainsonine)	v_{max}
Without Swainsonine	2.3 ± 0.2 mM	8.5 ± 0.2 F/s
With Swainsonine	2.9 ± 0.6 mM	4.3 ± 0.4 F/s

Non-linear regression analysis was performed with the data points from Figure 62 in two different ways: once by assuming that swainsonine competitively inhibits dGMII and further by assuming that swainsonine acts in a non-competitive manner.

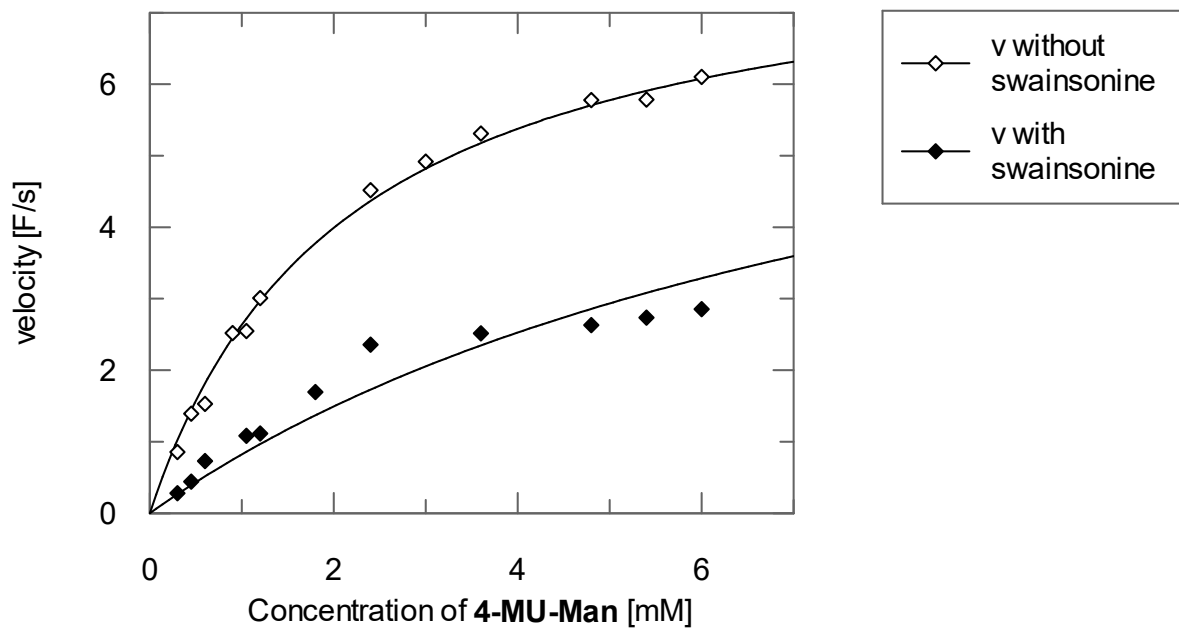


Figure 63: Competitive fit to determine the K_M , v_{max} and K_i values of dGM II with and without swainsonine using Equation 17

Table 12: K_M , v_{max} and K_i values obtained from the competitive fit of Figure 63

K_M	2.1 ± 0.3 mM
v_{max}	8.2 ± 0.4 F/s
K_i	12.0 ± 1.0 nM

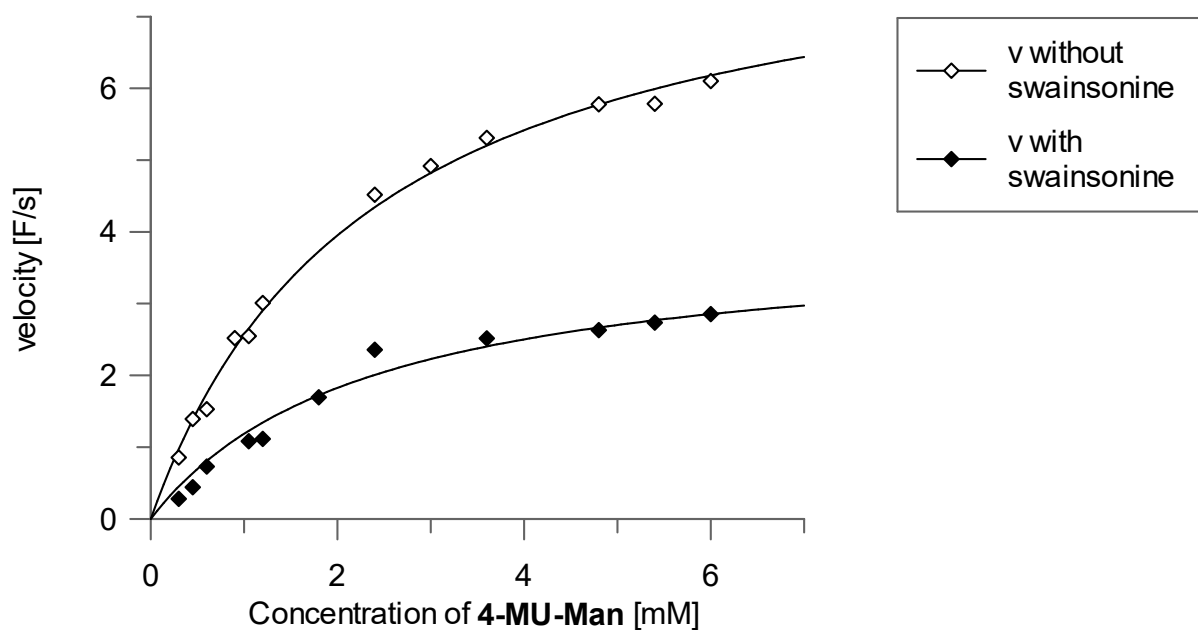


Figure 64: Non-competitive fit to determine the K_M , v_{max} and K_i values of dGM II with and without swainsonine using Equation 18

Table 13: K_M , v_{max} and K_i values obtained from the non-competitive fit of Figure 64

K_M	2.4 ± 0.2 mM
v_{max}	8.6 ± 0.3 F/s
K_i	34.0 ± 2.0 nM

Furthermore, a Lineweaver-Burk plot (see Figure 65) was performed on the data obtained from the data points of Figure 62.

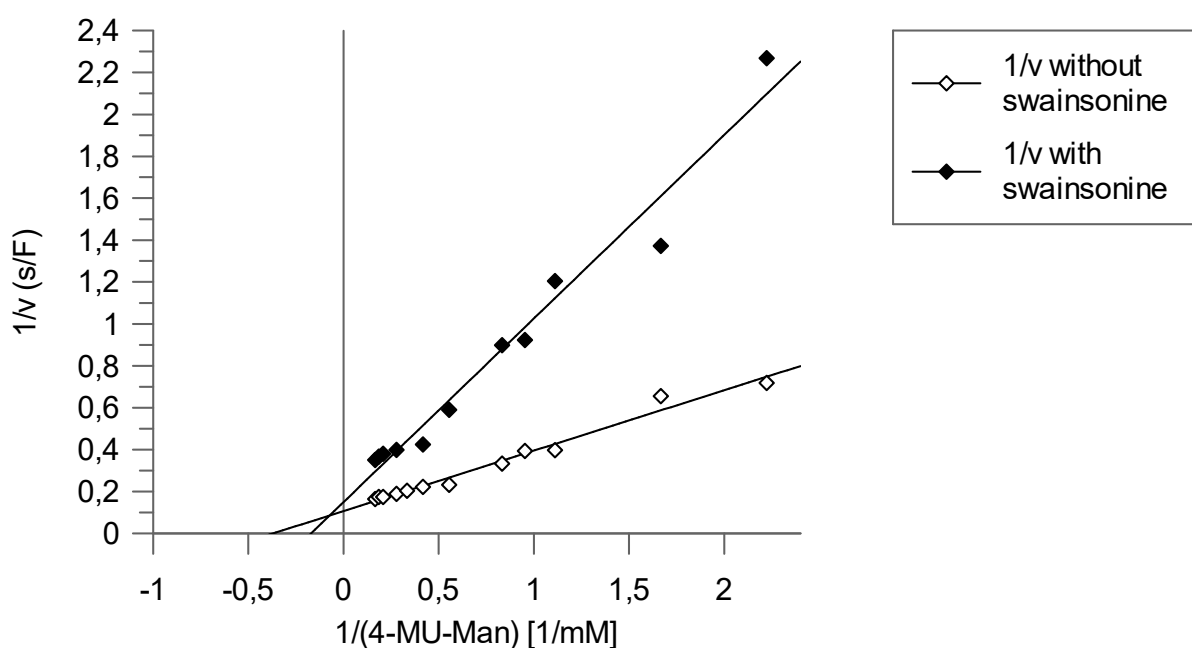


Figure 65: Lineweaver-Burk plot for the determination of the inhibitor type

K_i and K_i^* value of swainsonine as slow binder

As described in the literature by Kang *et al.* [64] and Tulsiani *et al.* [79], swainsonine is a slow-binding inhibitor. The dissociation constants K_i for the $[EI]$ complex and K_i^* for the high-affinity isomerized $[E^*I]$ complex are calculated via the initial rates (v_i) and the steady-state velocities (v_s). The values for v_i and v_s are obtained by IC_{50} value measurements at different substrate concentration over a period of 15-minutes. They were plotted against the inhibitor concentrations in each case. Due to the slow binding character of swainsonine, the inhibition is time-dependent. The turnover-time curves are not linear but asymptotic (see Figure 66).

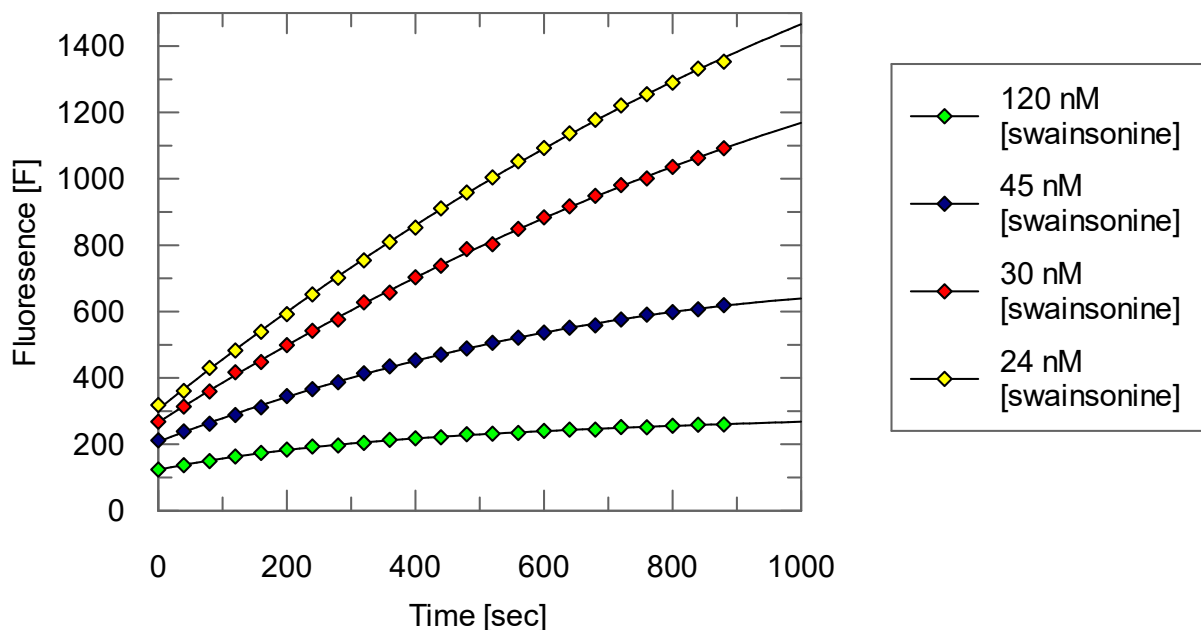


Figure 66: Examples of turnover-time curves for inhibition of dGM II by the inhibitor swainsonine

Table 14: List of K_i values and K_i^* values resulting from the non-linear fits

[4-MU-Man]	K_i	K_i^*
0.5 mM	32 ± 6 nM	15 ± 2 nM
1.0 mM	28 ± 2 nM	16 ± 1 nM
1.5 mM	37 ± 4 nM	14 ± 2 nM
2.0 mM	35 ± 4 nM	15 ± 2 nM

4.5.2.1.5 Effect of the assay additive DMSO on the activity of dGMII

In the following table (see Table 15), the effect of the assay additive DMSO on the activity of the dGMII is described.

Table 15: The effect of DMSO on the activity of dGMII

DMSO concentration (v/v)	0%	1%	2%	3%	4%	5%	10%	20%
Enzyme activity	100%	97%	99%	109%	96%	105%	50%	23%

4.5.2.1.6 Z'- factor determination

The measurements for the determination of the Z'-factor were carried out under the described conditions (see section 3.7.2.1.5). The velocities were plotted against the well numbers. In Figure 67, the result from two measurements are displayed.

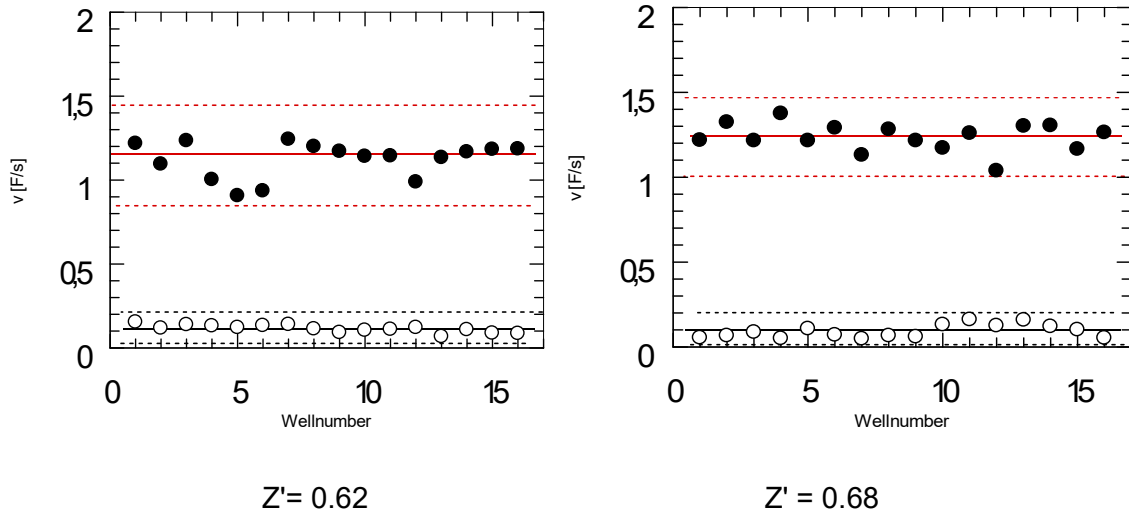


Figure 67: The data corresponding to the positive controls are shown as filled black circles and those for the negative controls as white circles, respectively. The solid lines represent the average of the positive and negative controls, the dashed lines are each three standard deviations away from the solid lines.

For each measurement, the mean μ and the standard deviation σ of the positive and negative control data were determined. Subsequently, the Z 'factor was calculated according to Equation 12. The results are summarized in Table 16.

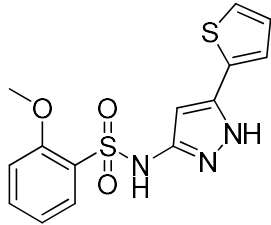
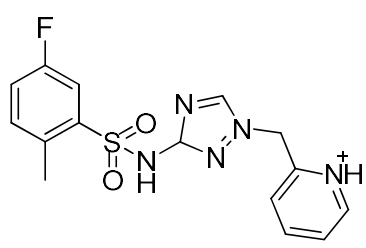
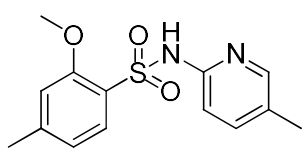
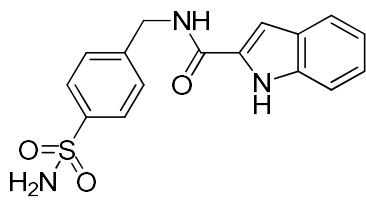
Table 16: Average of the calculated means μ and standard deviations σ for the positive and negative controls and the Z 'factor for two measurements

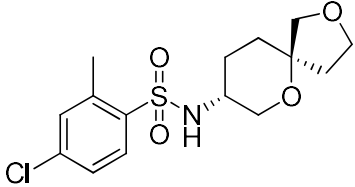
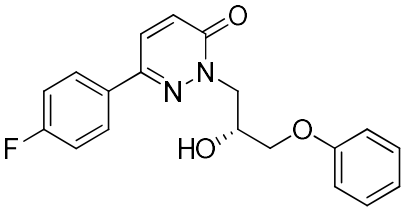
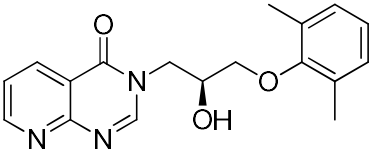
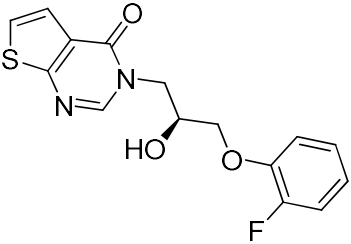
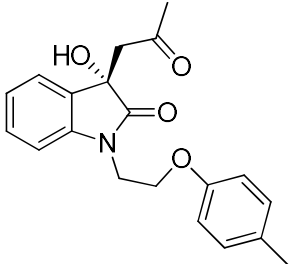
Measurement	μ_{c+} / [F/s]	σ_{c+} / [F/s]	μ_{c-} / [F/s]	σ_{c-} / [F/s]	Z'
1	1.12	0.11	0.11	0.02	0.62
2	1.24	0.09	0.09	0.04	0.68

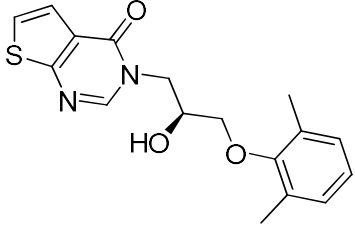
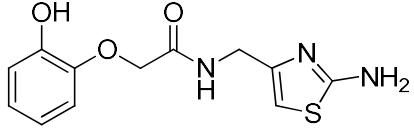
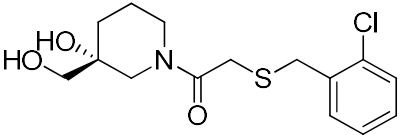
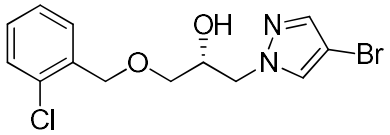
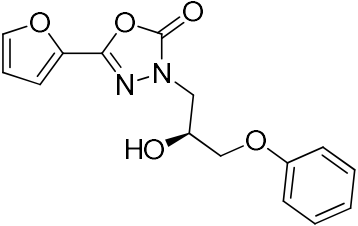
4.5.2.1.7 Testing of the virtual screening hits

In Table 17, the results of the virtual screening of the potential allosteric site described in section 3.5 with the corresponding inhibition data in the enzymatic assay, the predicted score and ligand efficiency are listed.

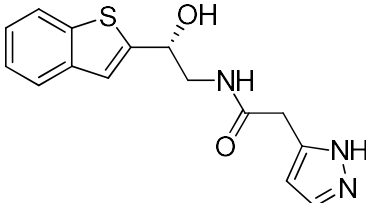
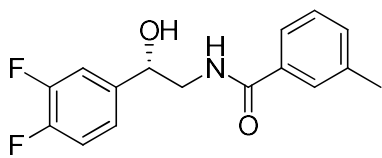
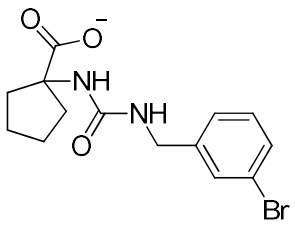
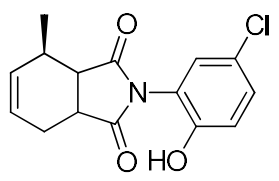
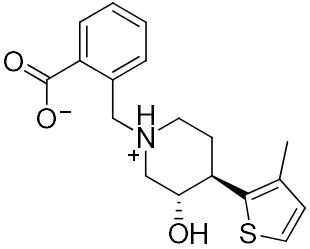
Table 17: List of score-ranked virtual screening hits with corresponding inhibition data in the enzymatic assay.

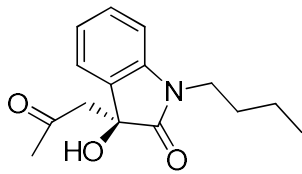
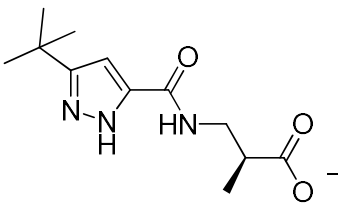
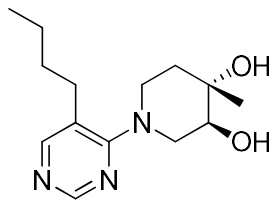
Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
1 ^a		-69.11	3.14	No inhibition
2 ^a		-60.48	2.52	No inhibition
3 ^a		-56.03	2.80	No inhibition
4 ^a		-53.66	2.33	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
5 ^a		-46.75	2.12	No inhibition
6 ^a		-39.96	1.60	No inhibition
7 ^a		-38.81	1.62	No inhibition
8 ^a		-37.87	1.72	No inhibition
9 ^a		-37.63	1.50	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
10 ^a		-37.27	1.62	No inhibition
11 ^b		-35.78	1.88	No inhibition
12 ^b		-35.76	1.70	No inhibition
13 ^a		-35.29	1.86	No inhibition
14 ^a		-34.53	1.57	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
15 ^a		-32.38	1.35	No inhibition
16 ^b		-32.32	1.70	No inhibition
17 ^b		-31.43	1.50	No inhibition
18 ^a		-31.38	1.25	46.3 ± 7.5
19 ^a		-31.37	1.43	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
20 ^a		-31.25	1.49	No inhibition†
21 ^a		-31.24	1.49	No inhibition
22 ^a		-31.22	1.56	No inhibition
23 ^b		-31.08	1.55	No inhibition
24 ^b		-26.77	1.16	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
25 ^a		-26.40	1.40	No inhibition
26 ^a		-25.95	1.44	No inhibition
27 ^b		-25.66	1.50	No inhibition

* Predicted ligand efficiency (see Equation 14)

** Inhibition (see Equation 15 at 300 μM (100 μM in case of cpd. **24**))

† Assay interference due to autofluorescence

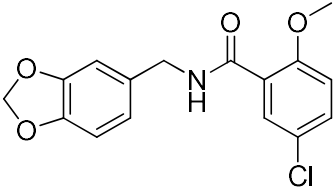
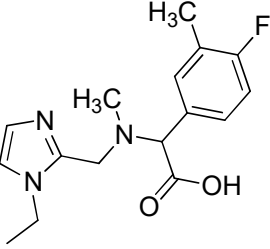
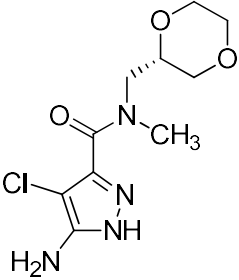
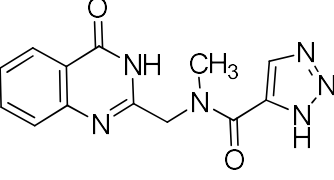
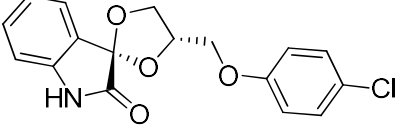
^a Compounds have been purchased from Enamine

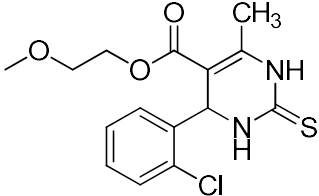
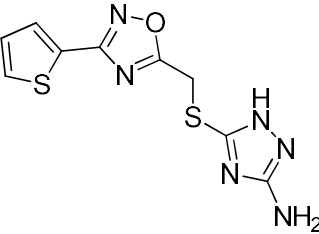
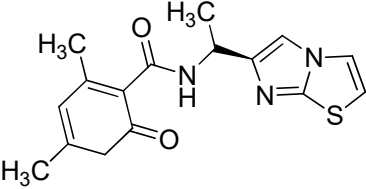
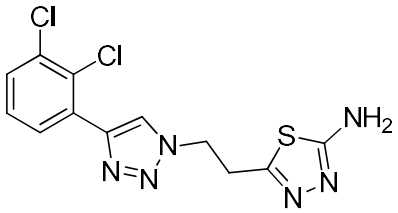
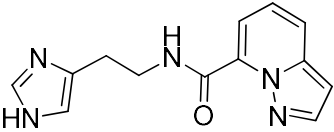
^b Compounds have been purchased from Chembridge

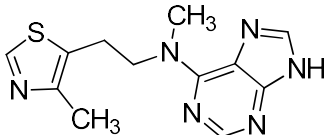
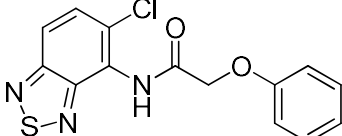
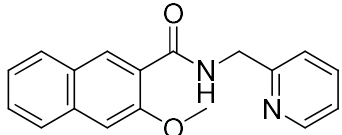
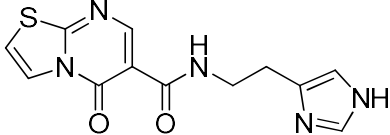
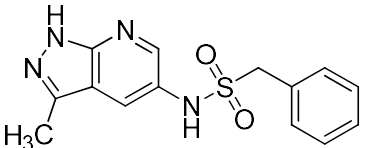
Two out of the 27 compounds exhibited inhibition > 40% at 300 μM, namely the bis-pyridyl derivative **18** (ca. 46% inhibition at 300 μM) and the benzothiophene **20** (ca. 41% inhibition at 300 μM). However, due to interference with the fluorometric assay compound **20** was not further pursued as inhibition of GMII was expected to be caused by an experimental artefact.

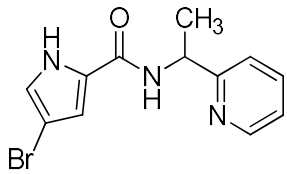
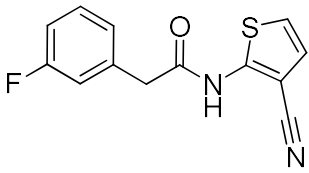
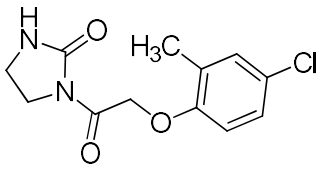
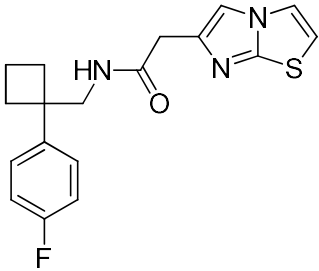
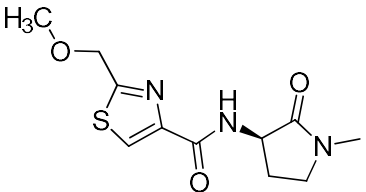
In Table 18, the results of the virtual screening of the catalytic site described in section 3.6 with the corresponding inhibition data of the enzymatic assay, the predicted score and ligand efficiency are listed.

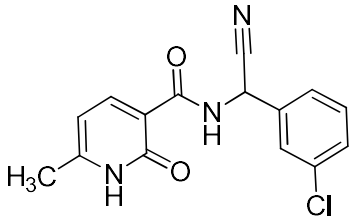
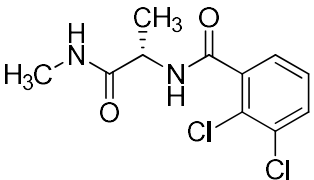
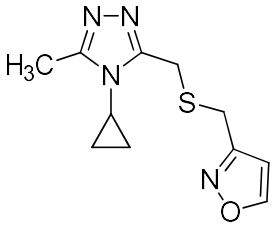
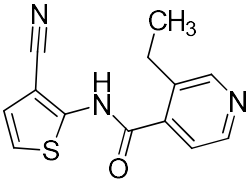
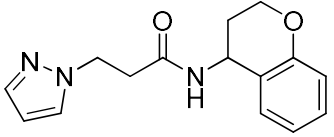
Table 18: List of score-ranked virtual screening hits with corresponding inhibition data in the enzymatic assay.

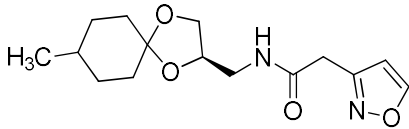
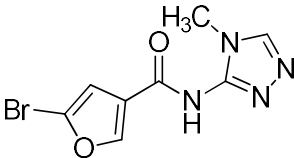
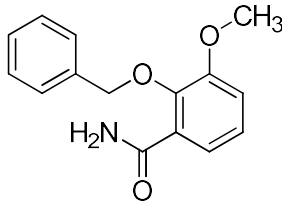
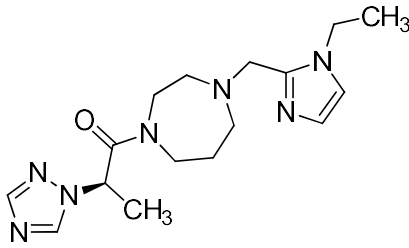
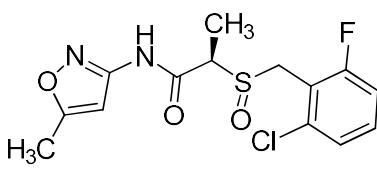
Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
28 ^a		-31.67	1.44	No inhibition
29 ^a		-31.78	1.44	No inhibition
30 ^a		-33.36	1.85	No inhibition
31 ^a		-31.87	1.52	No inhibition
32 ^a		-31.15	1.35	22.4 ± 10.1

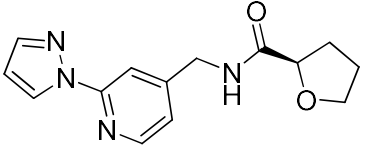
Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
33 ^a		-36.41	1.65	18.8 ± 12.9
34 ^a		-44.16	2.45	No inhibition
35 ^a		-30.84	1.40	No inhibition
36 ^a		-43.10	2.65	No inhibition
37 ^a †		-32.62	1.72	No inhibition †

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
38 ^a		-39.61	2.08	No inhibition
39 ^a		-31.26	1.49	No inhibition
40 ^a		-34.58	1.57	No inhibition
41 ^a		-35.12	1.76	No inhibition
42 ^{bt}		-31.66	1.51	No inhibition †

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
43 ^b		-31.88	1.89	No inhibition
44 ^b		-31.86	1.77	No inhibition
45 ^b		-33.69	1.87	No inhibition
46 ^b		-33.18	1.38	No inhibition
47 ^b		-31.76	1.76	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
48 ^b		-30.89	1.47	19.5 ± 12.6
49 ^b		-31.18	1.83	No inhibition
50 ^b		-31.14	1.83	No inhibition
51 ^b		-31.47	1.75	No inhibition
52 ^b		-30.91	1.54	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
53 ^b		-31.22	1.49	No inhibition
54 ^b		-31.42	2.09	No inhibition
55 ^b		-42.27	2.2	No inhibition
56 ^b		-34.39	1.43	No inhibition
57 ^b		-31.23	1.42	No inhibition

Cpd.	Molecular structure	Score [kJ/mol]	Predicted LE* [kJ×mol ⁻¹ ×HA ⁻¹]	Inhibition** [%]
58 ^b		-32.09	1.60	No inhibition

* Predicted ligand efficiency (see Equation 14)

** Inhibition (see Equation 15) at 300 μM (100 μM in case of compounds purchased from chembridge)

† Assay interference due to autofluorescence

^a Compounds have been purchased from Chembridge

^b Compounds have been purchased from Enamine

Three out of the 30 compounds exhibited very weak inhibition (compound **32** and **33**) at a concentration of 100 μM per well and one compound (compound **48**) at a concentration of 300 μM per well. As part of her diploma thesis, Malena dos Santos Guilherme [168] determined an inhibitory activity of $84.1 \pm 4.0\%$ for compound **32** at a concentration of 500 μM per well. For this reason, compound **32** was further pursued as potential inhibitor.

4.5.2.1.8 IC₅₀ and determination and mode of inhibition of compound **18**

For compound **18**, the IC₅₀ value was determined to be $216.7 \pm 44.3 \mu\text{M}$ (see Figure 68). Assays with varying substrate concentrations indicated that the IC₅₀ value is independent of the substrate concentrations proving a non-competitive mode of inhibition (Table 19 & Figure 69). The IC₅₀ determination of compound **18** was carried out as described in section 3.7.2.1.8.

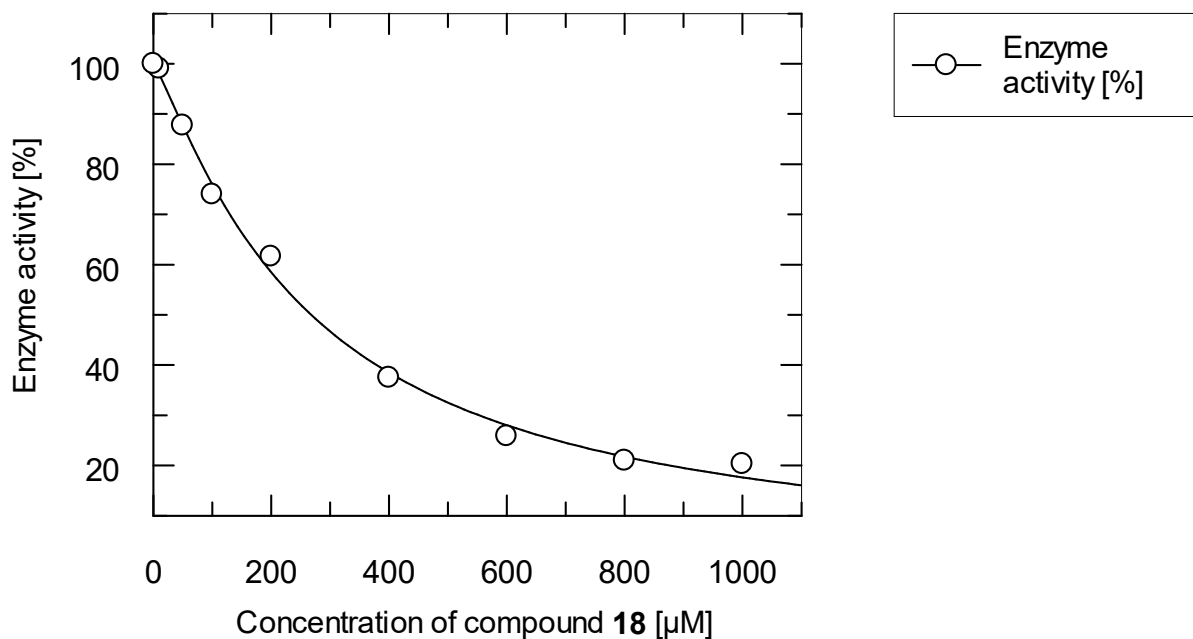


Figure 68: Example of a determination of the IC₅₀ value of compound **18** at a substrate concentration of 4-MU-Man of 1 mM. Non-linear regression was applied to the data using Equation 16

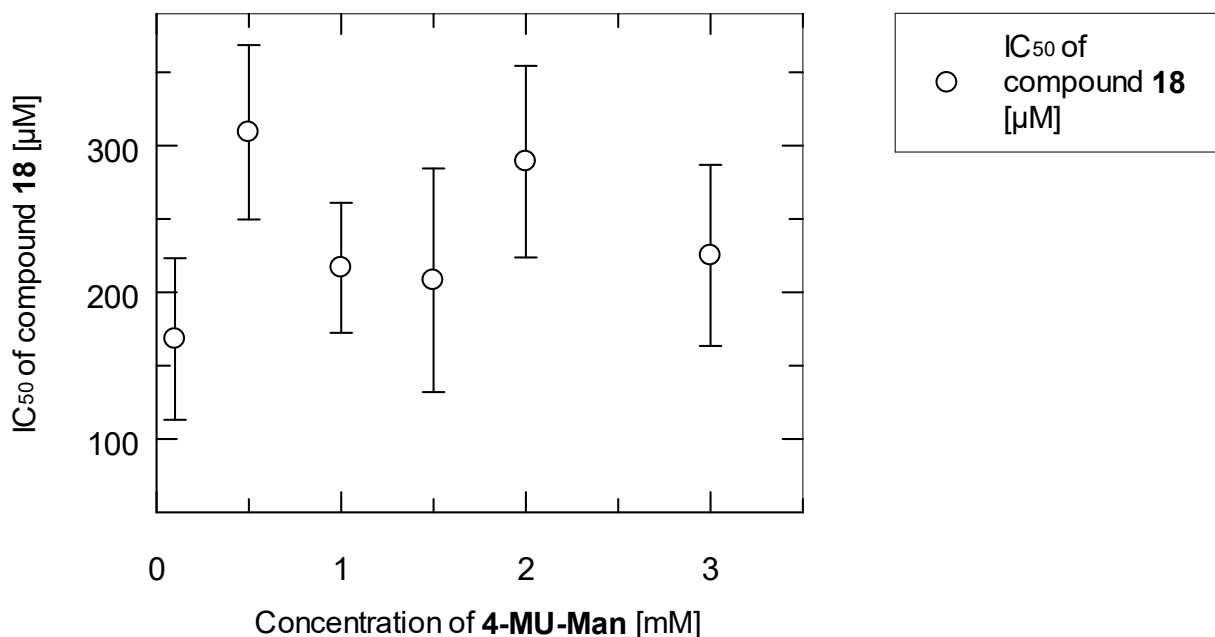


Figure 69: Effect of different substrate concentrations on the IC₅₀ values of compound

Table 19: Effect of different substrate concentrations on the IC₅₀ values of compound 18

[4-MU-Man] Substrate concentration	IC ₅₀
0.1 mM	168.1 ± 55.1 μM
0.5 mM	309.1 ± 59.4 μM
1.0 mM	216.7 ± 44.3 μM
1.5 mM	208.2 ± 76.2 μM
2.0 mM	289.1 ± 65.3 μM
3.0 mM	225.2 ± 61.7 μM

In order to determine the selectivity of compound 18, it was also tested whether it inhibits the β-glucosidase of sweet almonds. At 300 μM only 15% inhibition was observed (see section 3.7.2.3.3). In addition, compound 18 was also tested for inhibitory activity against JBM. At 300 μM, 70 % inhibition was observed (see section 3.7.2.2.4).

4.5.2.1.9 IC₅₀ and determination and mode of inhibition of compound 32

The IC₅₀ determination of compound 32 was carried out as described in section 3.7.2.1.9.

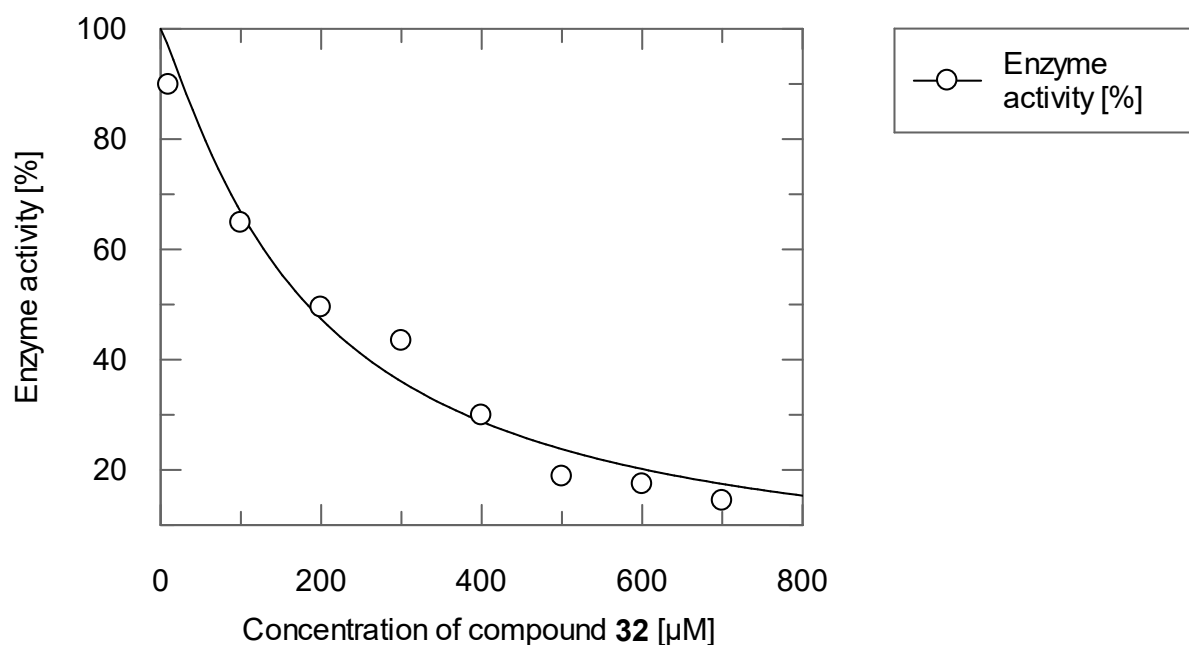


Figure 70: Example of a determination of the IC₅₀ value of compound 32 at a substrate concentration of 4-MU-Man of 1 mM. Non-linear regression analysis was applied to the data using by Equation 16.

In the following table (see Table 20), the effect of different substrate concentrations on the IC_{50} values of compound **32** are displayed (see section 3.7.2.1.9).

Table 20: Effect of different substrate concentrations on the IC_{50} values of compound **32**

[4-MU-Man] Substrate concentration	IC_{50}
0.1 mM	$217.3 \pm 26,9 \mu\text{M}$
1.0 mM	$196.5 \pm 13.9 \mu\text{M}$
2.0 mM	$322.4 \pm 50.8 \mu\text{M}$

4.5.2.1.10 K_i value determination of compound **32**

In the following figure (see Figure 71), the slopes were plotted against the substrate concentrations. Using the Michaelis-Menten equation, a non-linear regression was applied for both series of measurements: the measurements with and without compound **32** being added.

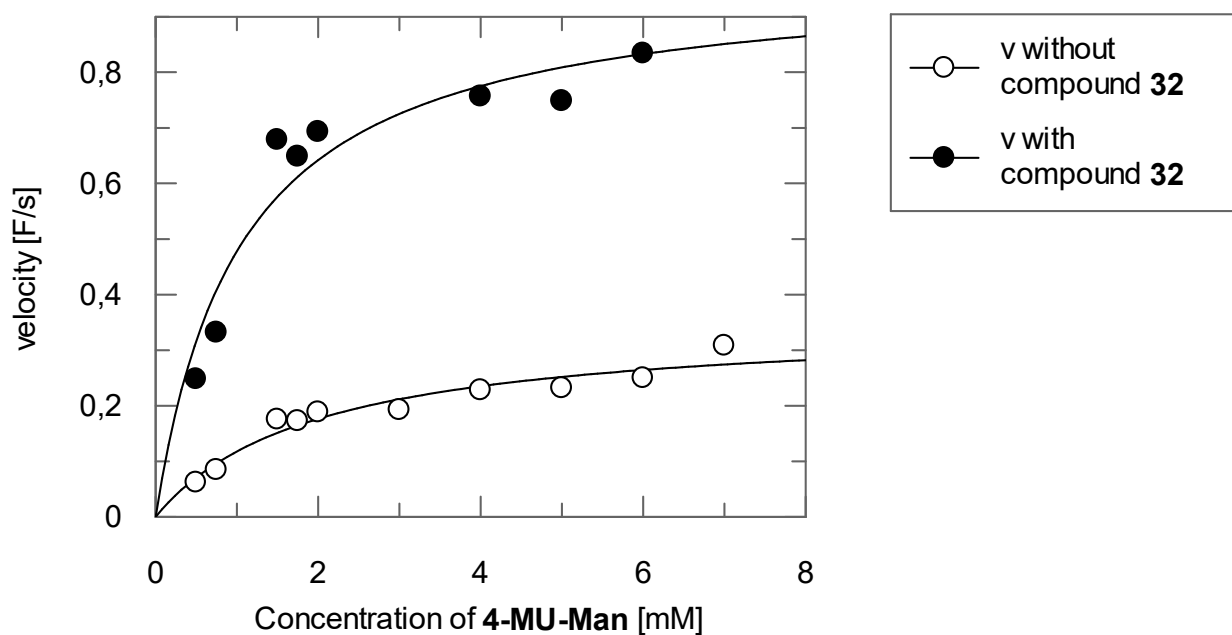


Figure 71: Example of a Michaelis-Menten plot for the determination of K_M and v_{max} values of dGM II with and without compound **32**

Table 21 lists the values obtained from the Michaelis-Menten plot of Figure 71 (see section 3.7.2.1.10).

Table 21: K_M and v_{max} values of the Michaelis-Menten plot of Figure 71 with and without compound 32

	K_M (K_M' with compound 32)	v_{max}
Without compound 32	1.0 ± 0.3 mM	0.98 ± 0.08 F/s
With compound 32	2.0 ± 0.5 mM	0.4 ± 0.03 F/s

Non-linear regression using the data points from Figure 71 was carried out once by assuming compound 32 acting as a competitive inhibitor, and secondly under the assumption that it non-competitively inhibits the enzyme (see Figure 72 and Figure 73).

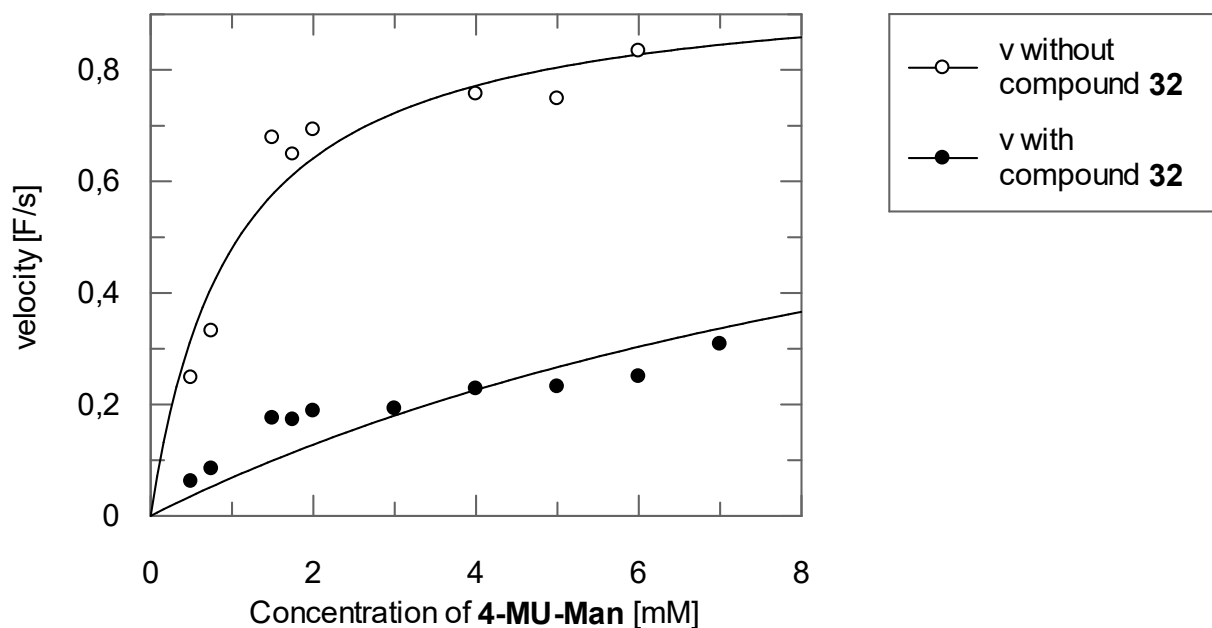


Figure 72: Competitive fit to determine the K_M , v_{max} and K_i values of dGM II with and without addition of compound 32 using Equation 17

Table 22: Kinetic constants obtained from the competitive fit as shown in Figure 72

K_M	1.0 ± 0.3 mM
v_{max}	1.0 ± 0.1 F/s
K_i	12.6 ± 3.0 μ M

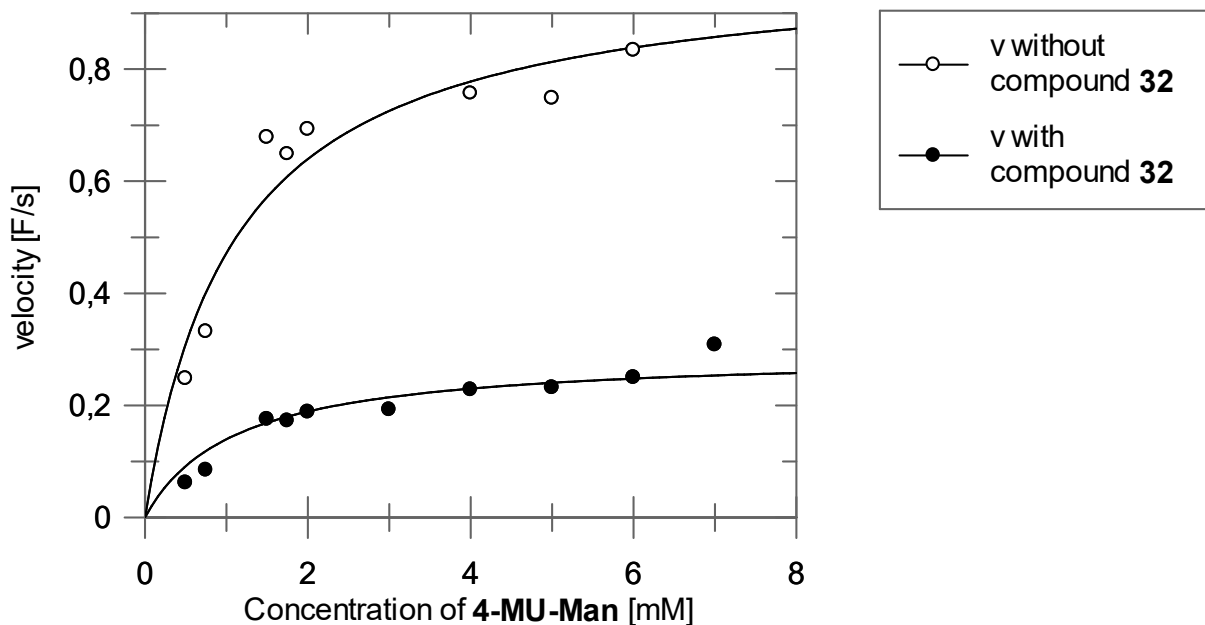


Figure 73: Non-Competitive fit to determine the K_M , v_{max} and K_i values of dGM II with and without compound **32** using Equation 18

In Table 23, the kinetic constants obtained from the non-competitive fit as shown in Figure 73 are listed.

Table 23: Kinetic constants obtained from the non-competitive fit as shown in Figure 73

K_M	1.1 ± 0.2 F/s
v_{max}	1.0 ± 0.06 mM
K_i	62.9 ± 7.0 μ M

A Lineweaver-Burk plot was calculated on the data obtained from the data points displayed in Figure 73.

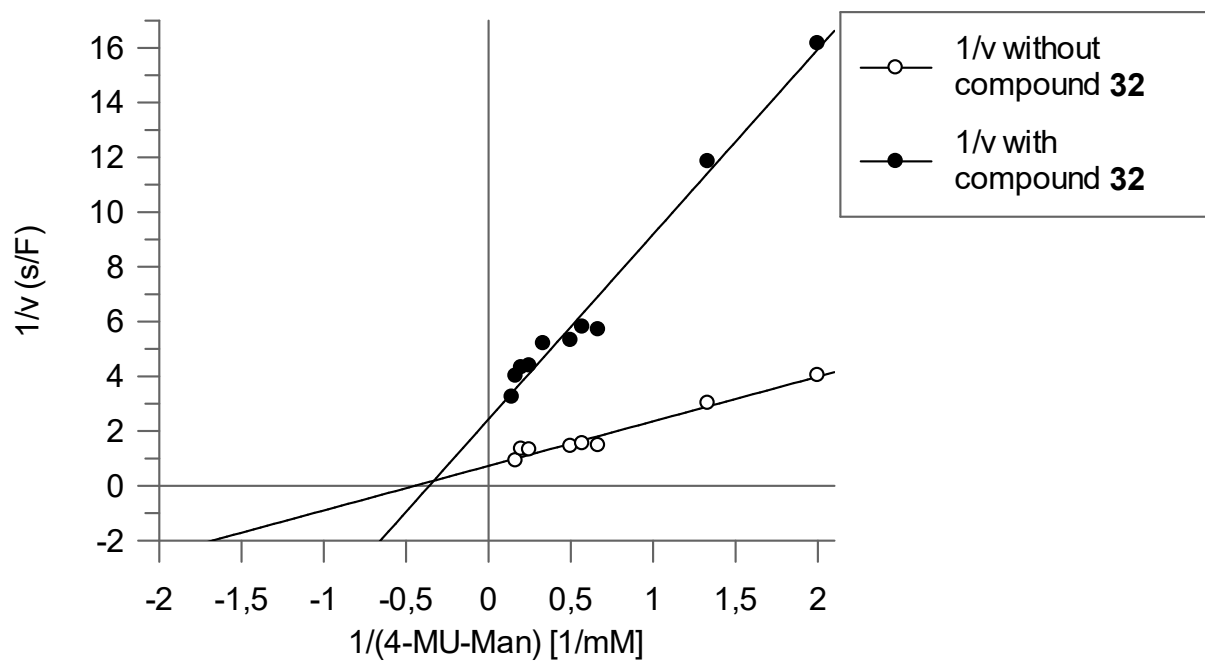


Figure 74: Lineweaver-Burk plot for the determination of the inhibitor type of compound 32

4.5.2.2 Enzyme assay with JBM

4.5.2.2.1 K_M value determination

The non-linear fit to determine the K_M value of dGMII with the substrate 4-MU-Man is shown in the following graph (see Figure 75):

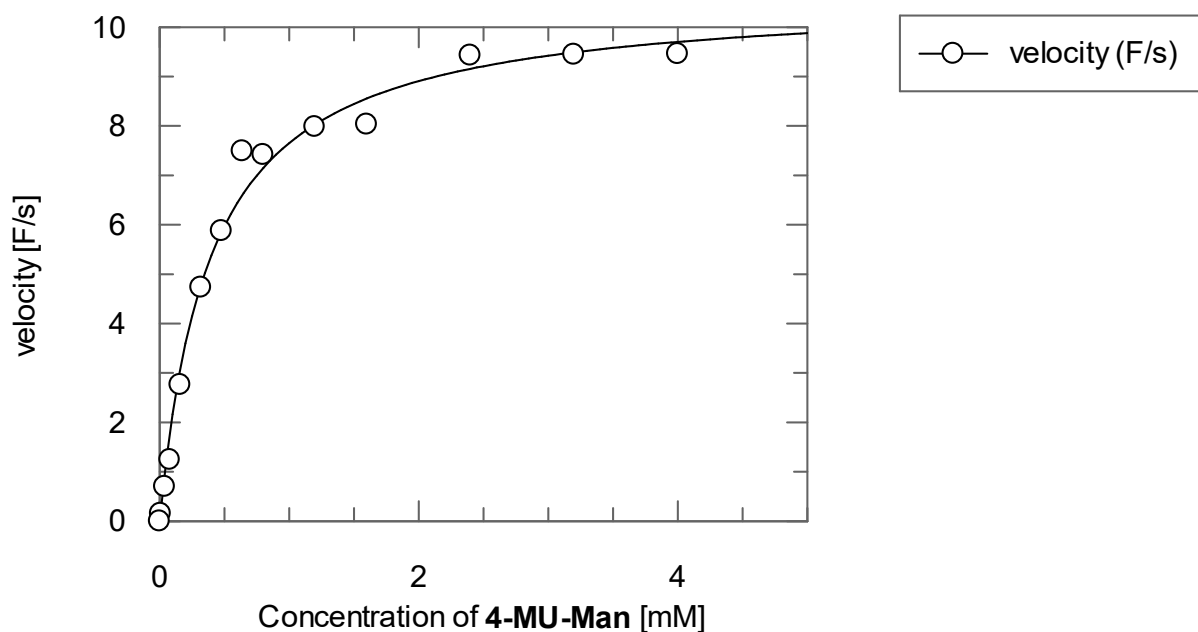


Figure 75: Example of the non-linear fit for the determination of the K_M value of JBM

In Table 24, the kinetic constants obtained from the Michaelis-Menten graph as shown in Figure 75 are listed.

Table 24: K_M and v_{max} values from the non-linear regression from graph from Figure 75

K_M	0.4 ± 0.05 mM
v_{max}	10.6 ± 0.4 F/s

4.5.2.2.2 IC₅₀ determination of swainsonine

In order to establish the assay, the known inhibitor swainsonine was used as a reference and inhibition data such as IC₅₀ values and K_i values were determined.

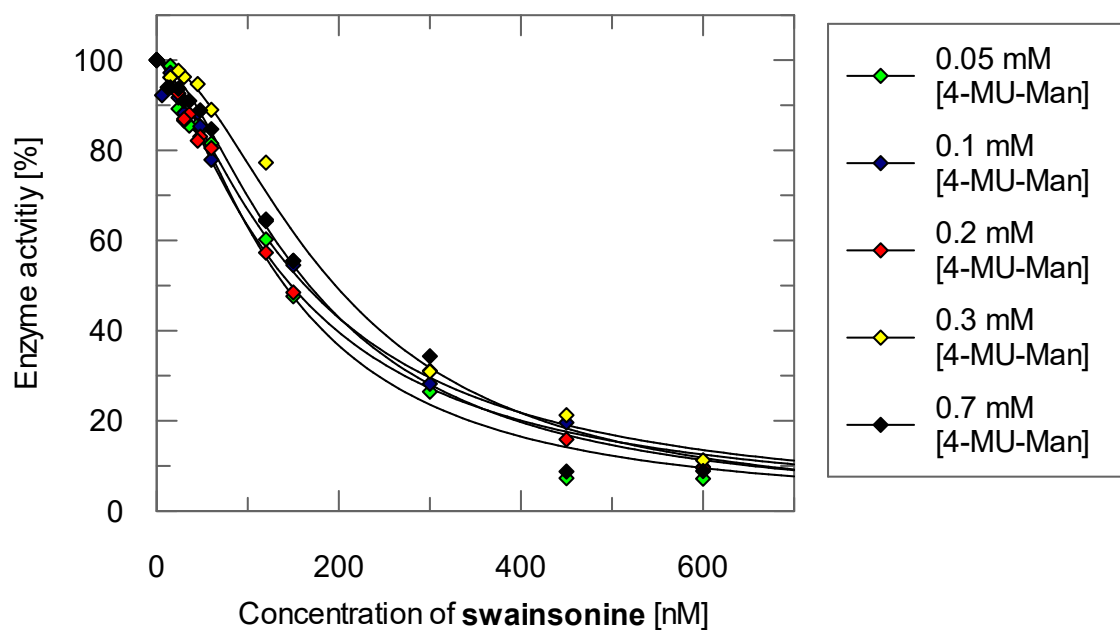


Figure 76: Determination of the IC₅₀ value of swainsonine at different concentrations of the substrate 4-MU-Man

In Table 25, the effect of different substrate concentration on the IC₅₀ value of swainsonine is shown.

Table 25: Effect of different substrate concentration on the IC₅₀ values of swainsonine

Concentration of 4-Mu-Man	IC ₅₀ values
0.05 mM	141 ± 9 nM
0.1 mM	160 ± 4 nM
0.2 mM	160 ± 16 nM
0.3 mM	207 ± 16 nM
0.7 mM	278 ± 15 nM

4.5.2.2.3 K_i value of swainsonine

In Figure 77, the slopes were plotted against the substrate concentrations. Using the Michaelis-Menten equation, a non-linear regression was applied for both series of measurements. The first series of measurements was carried out in the presence and the other one in the absence of the inhibitor swainsonine. In Table 26, the corresponding K_M values obtained from the Michaelis-Menten plot are reflected.

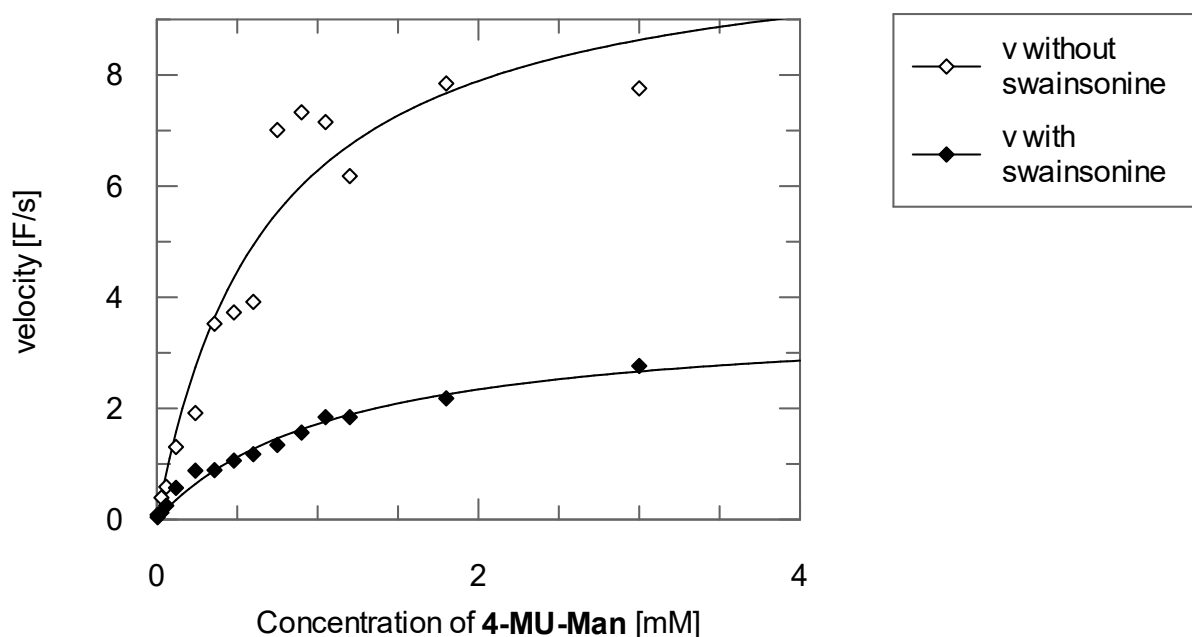


Figure 77: Example of the Michaelis-Menten plot for the determination of K_M and v_{max} values of JBM with and without swainsonine

Table 26: K_M and v_{max} values of the Michaelis-Menten plot with and without swainsonine

	K_M (K_M' with swainsonine)	v_{max}
Without Swainsonine	0.7 ± 0.2 mM	11 ± 1.2 F/s
With Swainsonine	1.1 ± 0.2 mM	3.7 ± 0.3 F/s

Non-linear regression analysis was performed with the data points from Figure 77 in two different ways: once by assuming that swainsonine competitively inhibits dGMII and further by assuming that swainsonine acts in a non-competitive manner.

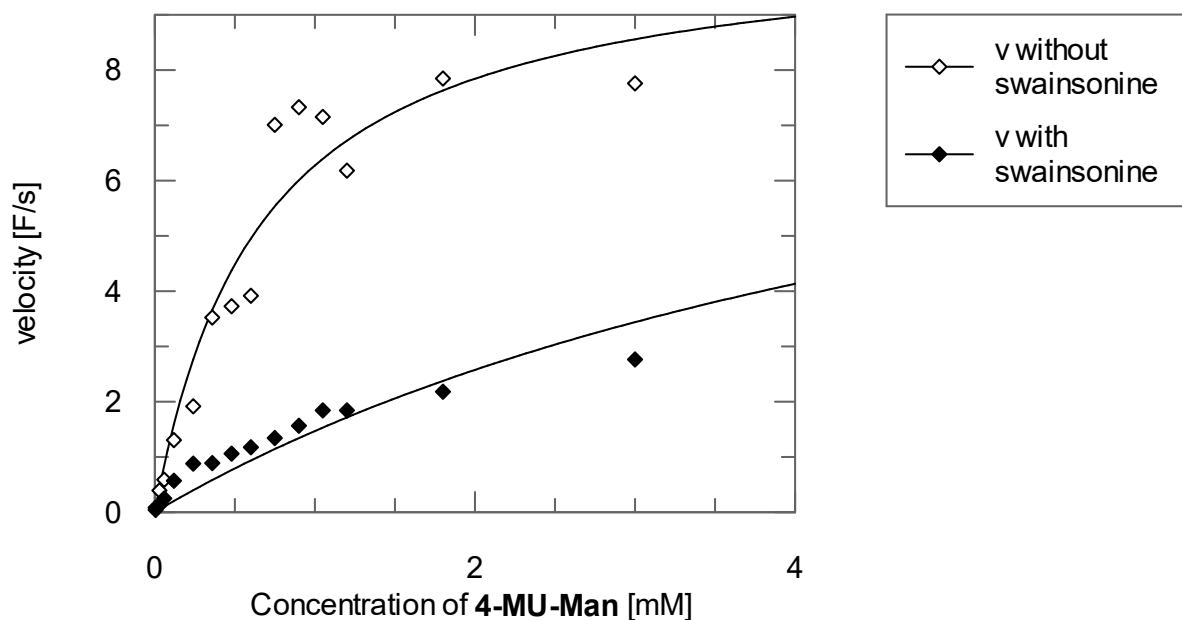


Figure 78: Competitive fit to determine the K_M , v_{max} and K_i With and without addition of swainsonine using Equation 17

In Table 27, kinetic parameters of the competitive fit of the data presented in Figure 78 are displayed.

Table 27: Kinetic parameters obtained from the competitive fit of data presented in Figure 78

K_M	0.67 ± 0.14 mM
v_{max}	10.5 ± 0.9 F/s
K_i	18 ± 4 nM

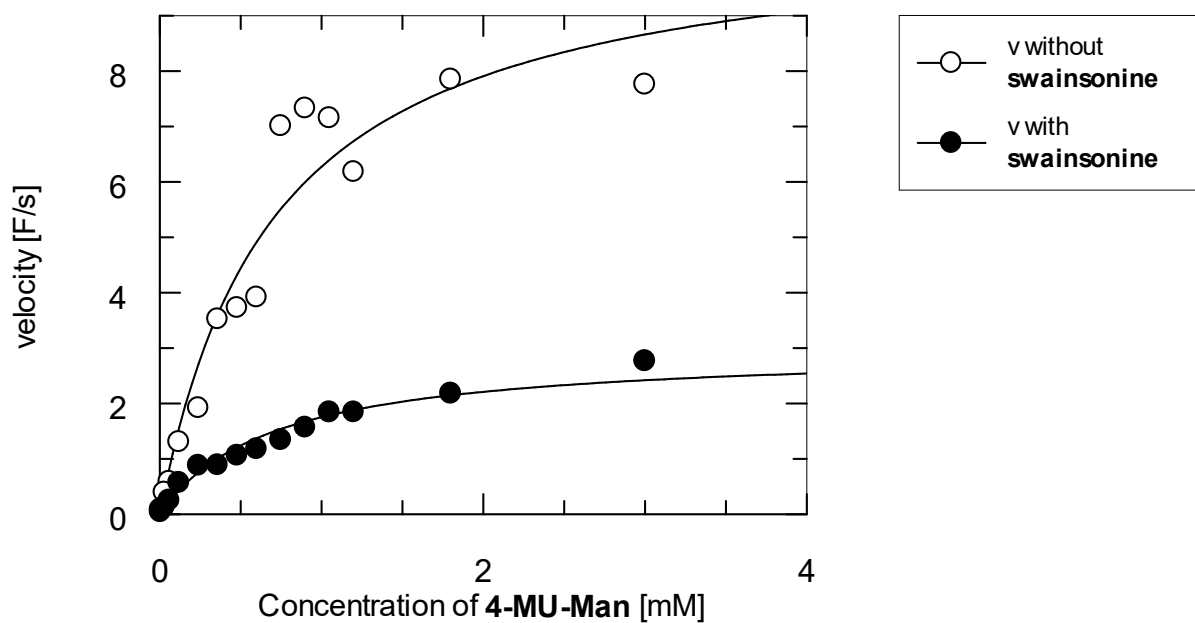


Figure 79: Non-competitive fit to determine the K_M , v_{max} and K_i values with and without addition of swainsonine using Equation 18

In Table 28, kinetic parameters of the non-competitive fit of the data presented in Figure 79. are displayed.

Table 28: Kinetic parameters obtained from the non-competitive fit of data presented in Figure 79

K_M	$0.70 \pm 0,14$ mM
v_{max}	10.7 ± 0.9 F/s
K_i	58 ± 9 nM

A Lineweaver-Burk plot was constructed with the data as obtained from the data points displayed in Figure 77.

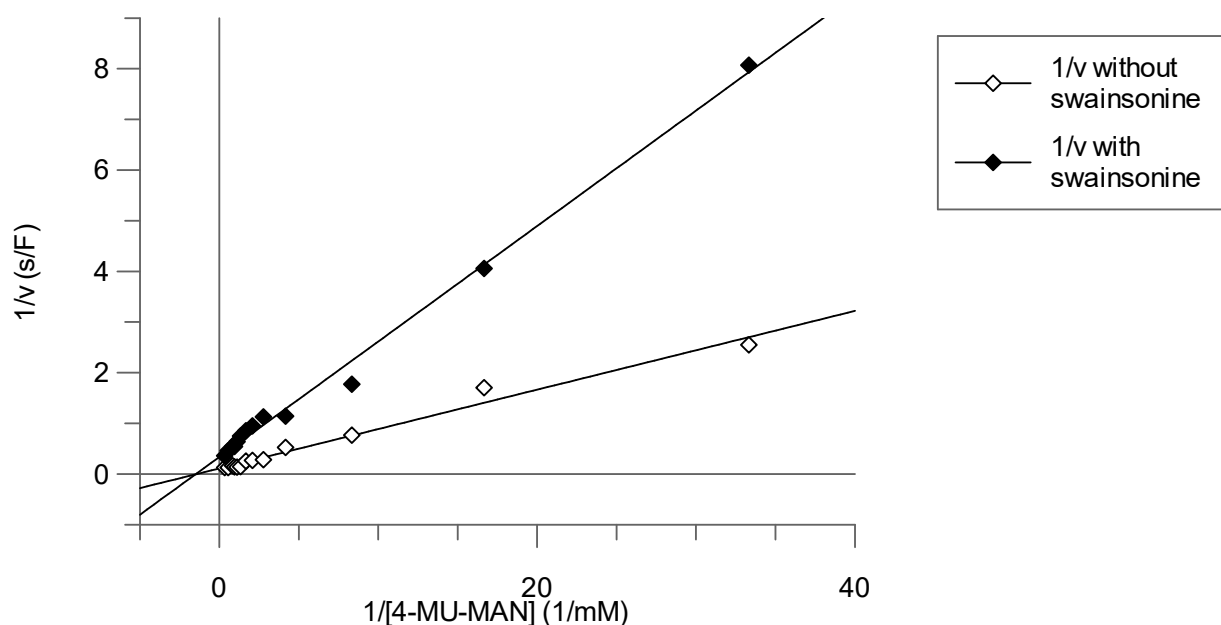


Figure 80: Lineweaver-Burk plot for the determination of the inhibitor type of swainsonine

K_i and K_i^* value of swainsonine as slow binding inhibitor of JBM

For JBM, the dissociation constants K_i and K_i^* of swainsonine were determined by taking the initial and steady-state rates into account, respectively. The v_i and v_s -values are derived from the determination of the IC_{50} value (see section 4.5.2.2.2). In the following table, the values of K_i and K_i^* are listed.

Table 29: List of K_i values and K_i^* values resulting from the nonlinear fits

[4-MU-Man]	K_i	K_i^*
0.05 mM	173 ± 14 nM	112 ± 6 nM
0.1 mM	197 ± 11 nM	128 ± 7 nM
0.2 mM	151 ± 9 nM	115 ± 8 nM
0.3 mM	169 ± 13 nM	147 ± 8 nM
0.7 mM	194 ± 10 nM	135 ± 9 nM

4.5.2.3 Enzyme assays with the β -glucosidase from sweet almonds

4.5.2.3.1 K_M value determination

The non-linear fit to determine the K_M value of the β -glucosidase from sweet almonds with the substrate 4-MU-Glc is shown in the following graph (see Figure 81).

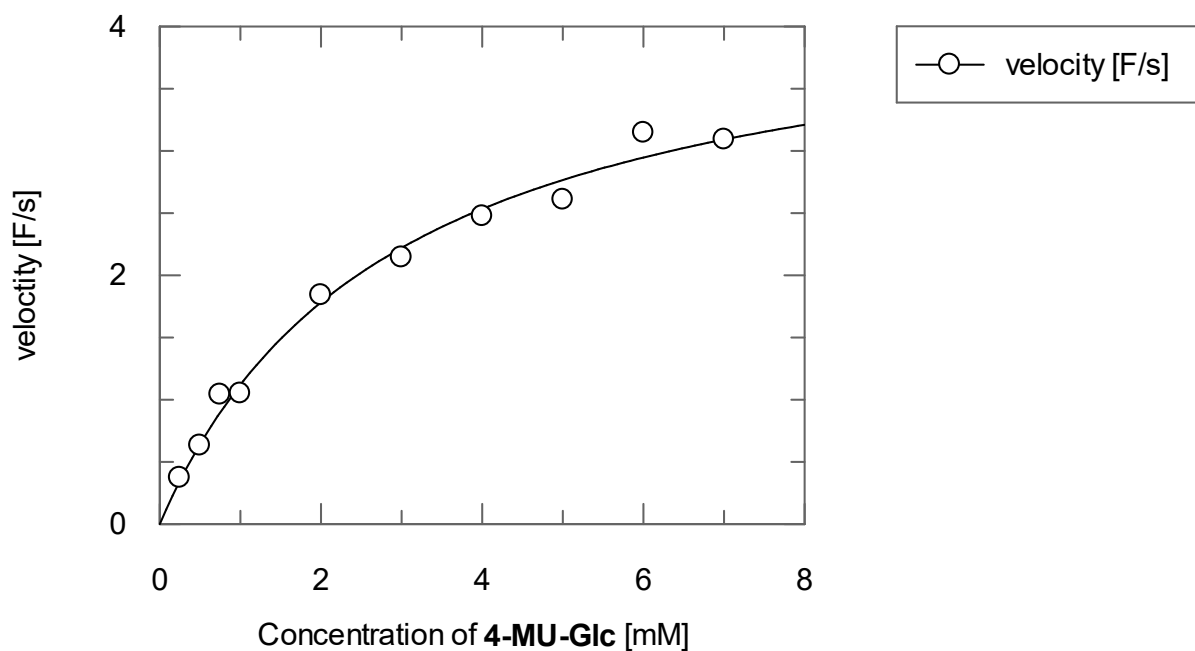


Figure 81: Example of a non-linear fit for the determination of the K_M value of the β -Glucosidase from sweet almonds

Considering all the three measurements, an average K_M value of 2.4 ± 0.5 mM and a v_{\max} value of 2.4 ± 1.7 F/s were determined (see Table 30).

Table 30: K_M and v_{\max} values from the non-linear regression from graph of Figure 81

K_M	2.9 ± 0.4 mM
v_{\max}	4.4 ± 0.2 F/s

4.5.2.3.2 IC₅₀ determination of isofagomine

Isofagomine is a known inhibitor of the β -glucosidase and was used as reference for the establishment of the fluorometric assay with β -glucosidase from sweet almonds with 4-MU-Glc as substrate.

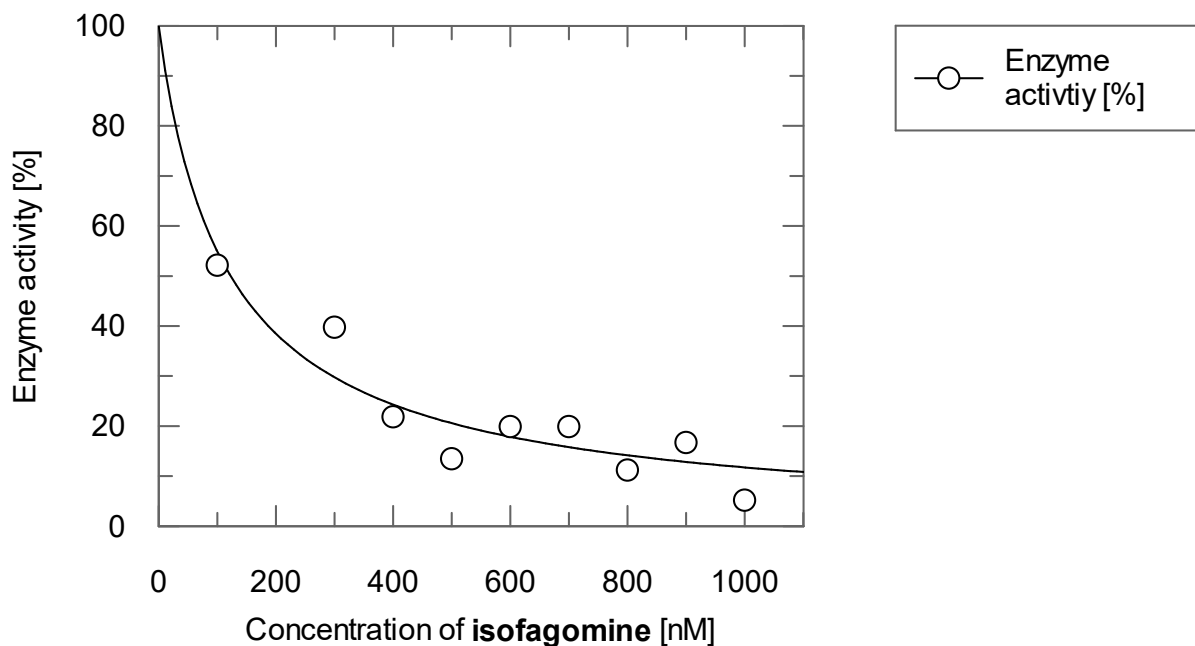


Figure 82: Example of a determination of the IC₅₀ value of isofagomine. Non-linear regression was applied to the data using Equation 16

The mean average IC₅₀ value for both measurements was 174.7 ± 73.4 nM.

4.5.2.3.3 Selectivity of compound **18**

In order to determine the selectivity of compound **18**, it was also tested for inhibition of the β -glucosidase of sweet almonds. At 300 μ M only 15% inhibition was observed.

5 Discussion

5.1 Homology modeling

5.1.1 Homology model of the hGMII

The determined sequence similarity and identity correspond to what has been reported in the earlier literature (41% identity and 61% similarity) [30]. Furthermore, the superposition of both structures yielded a RMSD-value of 1.16 Å, which leads to the assumption that the proposed 3D structure of hGMII closely resembles the structure of its template. The differences in the structures seem to be more prominent only in the periphery of the structure (see Figure 39, section 4.1.1). Figure 40 (see section 4.1.1) displays the identity (A) and the similarity (B) of both structures. The similarity extends over the entire enzyme, whereas regions with identical amino acids are restricted to several parts of the enzyme structures. These areas with the highest degree of identity are located in close proximity to the Zn²⁺-ion and thus in the catalytic binding pocket. In the area of 6 Å around the Zn²⁺-ion, the amino acid composition is highly conserved (see Figure 41, section 4.1.1). The structural similarity of the two enzymes with regard to the binding pocket was already described in the literature by D. Rose [28] and Petersen *et al.* [43]. Consequently, the homology model was found suitable to be used for further docking studies. However, sequence identity and similarity alone are not sufficient to prove functional homology. Rabouille *et al.* [169] showed that dGMII shares similar kinetic properties and inhibitor sensitivity when compared to mouse GMII. Based on this knowledge, the homology modelling results support the assumption that dGMII can be used as a suitable surrogate for mammalian GMII in enzyme assays.

5.1.2 Homology model of JBM

JBM serves as a model for a potential off-target due to its sequence similarity to bLAM (see Table 5, section 4.1.2). However, the fact that JBM represents an off-target model may prove useful for testing potential inhibitors in enzymatic assays. Testing the hit compounds for activity against the off-target helps determining the selectivity profile. Selective inhibitors should discriminate between closely related targets. Thus, the structural similarity to GMII proteins with regard to the active site is of particular importance for the present study. This similarity can be determined upon superposition of the homology model of JBM with dGMII (see Figure 42, section 4.1.2).

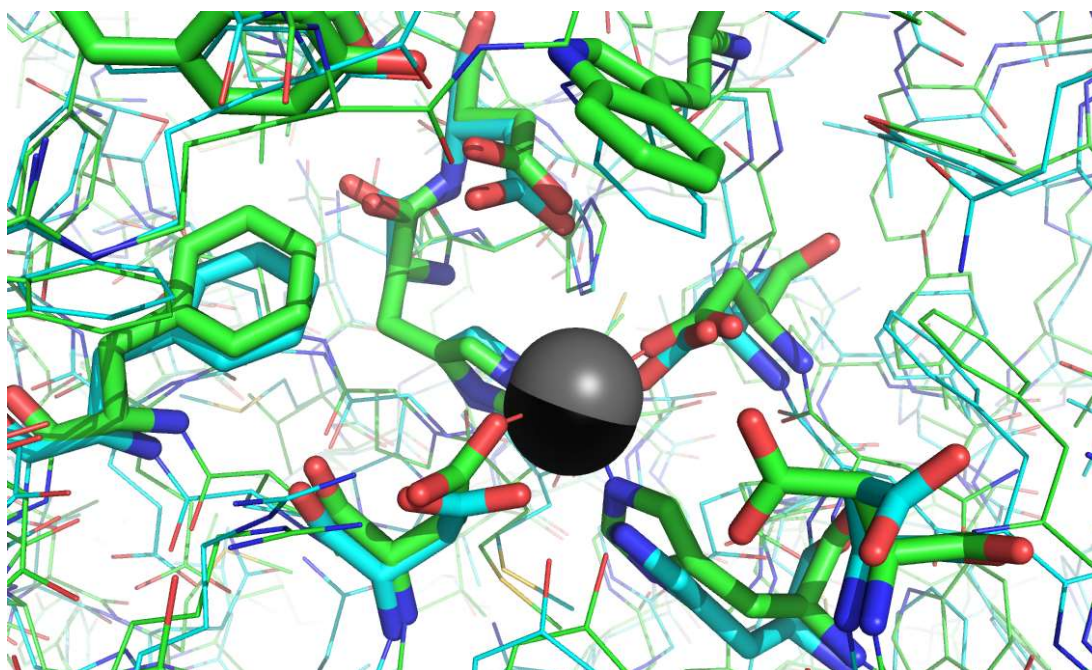


Figure 83: Superposition of JBM (cyan-coloured carbon atoms) and dGMII (green-coloured carbon atoms) (PDB code): 1HWW [30]). Amino Acids near the zinc ion are depicted in stick representation. The zinc ions are shown as spheres (black sphere from the homology model and grey sphere from the crystal structure).

Regions up to 10 Å around the zinc ion are almost identical and differ only marginally with regard to the atom coordinates of the amino acids. JBM is, like other Class II α -mannosidases, inhibited by swainsonine [64], certainly due to the close structural similarity of the catalytic site.

5.1.3 Homology model of the sweet almond β -glucosidase

As already described (see section 1.3), the β -glucosidase from sweet almonds is classified in the family 1 of glycosyl hydrolases and uses glutamate as active site residue for the nucleophilic attack with net retention of the anomeric configuration [59], [60]. In contrast, dGMII belongs to the family 38 of glycosyl hydrolases [36] and uses an aspartate as active site residue. The phi-psi plot for the homology model is presented in Figure 45 (section 4.1.3). Most of the amino acids are assigned to the favourable regions of the plot (see Figure 45, section 4.1.3). Its primary structure consists of 542 amino acids compared to 1108 amino acids in the case of dGMII [37]. The sequence alignment of the cyanogenic β -glucosidase from the white clover (490 amino acids and pdb code: 1CBG [160]) with the sequence of the sweet almond β -glucosidase revealed 57.2% identity and 68.6% similarity. An overlay of both structures resulted in a RMSD value of 0.99 Å. On the other hand, the sequence alignment of dGMII and sweet almond β -glucosidase resulted in a low identity of 12.2% identity and

similarity 29.3%. An overlay of both structures resulted in a RMSD of 24.5 Å indicating a high difference between both 3D-structures.

5.2 Enzyme Assay

5.2.1 Enzyme assays with *drosophila melanogaster* Golgi α -mannosidase II

5.2.1.1 K_M value determination

The determined K_M value for dGMII is 3.6 ± 0.3 mM ($v_{max} = 8.8 \pm 0.3$ F/s). In the literature, the K_M (4.9 ± 2.6 mM) is only documented within a personal communication from H. Strachnan and K. Moremen (Complex Carbohydrate Research Center and Department of Biochemistry and Molecular Biology, University of Georgia) as part of the publication from Coleman *et al.* [156]. However, it should be noticed that the determined K_M value is in the same range as the one reported by H. Strachnan and K. Moremen. The affinity of dGMII for 4-MU-Man is low compared to the natural substrate ($85 \mu\text{M}$) [156], [170]. However, this assay is sensitive and requires very low enzyme concentrations, which enables the fast screening of potential inhibitors. By Coleman *et al.* [156], a fluorimetric assay using resorufin α -D-mannopyranoside (see Figure 84) as substrate was presented (K_M value: $200 \pm 73 \mu\text{M}$ and v_{max} : 11 nmol/min). The K_M value using the substrate resorufin α -D-mannopyranoside ($200 \mu\text{M}$) is significantly lower than that in the presence of 4-Mu-Man. Thus, this substrate has 20-fold higher affinity than 4-MU-Man and 250-fold higher than PNP-Man. Subsequently, a lower substrate concentration would be a cost effective alternative. On the other hand, due to the high affinity of resorufin α -D-mannopyranoside, weak competitive inhibitors would not be detected in the pre-screening ($300 \mu\text{M}$ or $100 \mu\text{M}$ well concentration of each inhibitor) and thus would not be considered for the possible optimization process. This would make the search for new inhibitors significantly more difficult. Since the substrate resorufin α -D-mannopyranoside has a similar affinity as the natural substrate, it might however be an option to use this assay in later stages of the optimization process.

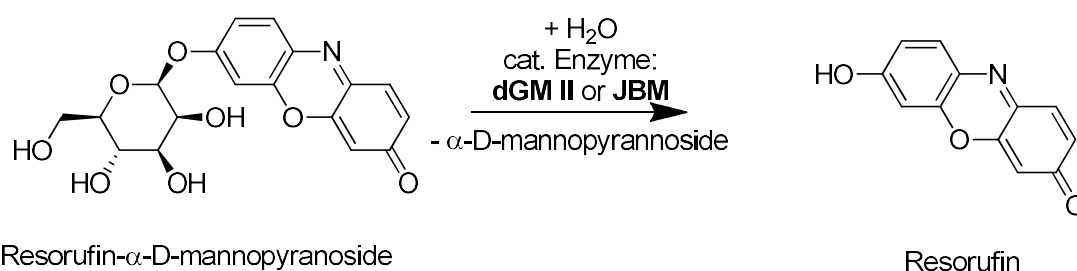


Figure 84: Catalyzed reaction of dGMII or JBM using Resorufin- α -D-mannopyranoside as substrate

The determination of the K_M is crucial for the establishment of an assay. A substrate concentration around or below the K_M is recommendable for the determination of inhibitory activity, especially for competitive inhibitors [171]. For this reason, we fixed a substrate concentration of 1 mM of 4-MU-Man for the inhibition assays.

5.2.1.2 IC_{50} determination of swainsonine

In the following (see Table 31), IC_{50} values of swainsonine with substrates other than 4-MU-Man are listed.

Table 31: IC_{50} values of swainsonine with different substrates from the literature

Substrate	IC_{50} values [nM]
4-nitrophenyl- α -D-mannopyranoside	40 [29] 12-20 [169]
resorufin- α -D-manno-pyranoside	100 (measurement 30 minutes) [156]

Since the cited literature values were provided without indicating the error range, it cannot be concluded whether the specifically determined values are within the error ranges. All measured IC_{50} values are between 31-40 nM. It thus becomes clear that both the IC_{50} values determined in this study and the literature values are in the lower nanomolar range. Furthermore, it can be concluded from the results that swainsonine most probably is a non-competitive inhibitor, due to the fact that the IC_{50} values remain unchanged regardless of the substrate concentrations. Moreover, Kang S.M. *et al.* [64] reported that swainsonine is a slow binder with an optimal pre-incubation time of 5 minutes. In order to demonstrate the slow binding character of swainsonine, IC_{50} values were determined with varying incubation time (0, 5 and 10 minutes, see section 4.5.2.1.3). The slow-binding effect was also demonstrated by plotting the fluorescence against the time over a period of 15 or 30 min. In this case, two different gradients are obtained in the respective graph: the initial velocity (v_i) and the steady-state velocity (v_s). The measurements were always carried out over a period of 15 min, but the IC_{50} value was determined only via the linear portion (first 6 minutes) of the initial slope (v_i). In the case of a non-competitive inhibitor, the IC_{50} is equal to the K_i [115].

5.2.1.3 IC₅₀ determination of swainsonine with different incubation time

Table 10 (see section 4.5.2.1.3) clearly shows that swainsonine is a slow-binding inhibitor. The IC₅₀ values decrease upon previous incubation. The inhibitory effect increases by 54% upon previous incubation (10 min incubation). However, it has been reported by Kang S.M. *et al.* [64] that a preincubation time of 5 min of swainsonine is optimal, since the greatest inhibition was detected at 5 min preincubation.

5.2.1.4 K_i value determination of swainsonine

Based on the measurements in the fluorescence based assay with 4-MU-Man as substrate, a K_M value of 2.3 ± 0.2 mM for the uninhibited enzyme reaction of dGMII with 4-MU-Man was determined. The K_M value of this reaction in the presence of swainsonine is 2.9 ± 0.6 mM (see Table 11). In view of the error range, the K_M value has not changed in presence of the inhibitor. A comparison of the v_{max}-values indicates a decrease in the presence of swainsonine (without swainsonine: v_{max} = 8.5 ± 0.2 F/s, with swainsonine: v_{max} = 4.3 ± 0.4 F/s). Taken together, it can be concluded that under the given assay conditions, swainsonine acts as a non-competitive inhibitor of dGMII. Furthermore, the non-competitive fit matches the data better than the competitive one (see Figure 63 and Figure 64, see section 4.5.2.1.4). Moreover, the Lineweaver-Burk plot also indicates non-competitive behaviour, as judged by the position of the intersect with the x-axis (see Figure 65, section 4.5.2.1.4).

Swainsonine is described in the literature as a competitive inhibitor of JBM [29]. In crystal structures of dGMII (1HWW [162] and 3BLB [172]), swainsonine is bound in the catalytic site. Tulsiani *et al.* [79] reported that swainsonine acts as a competitive inhibitor of JBM at concentrations of ≤ 500 nM and a non-competitive inhibitor at higher concentration (> 2 μM). However, the inhibition of the homogeneous GMII derived from Golgi membranes of rat liver was non-competitive, even at low concentrations of 200 nM. This would be consistent with the assay results with dGMII. The observed non-competitive inhibition mode of swainsonine can be explained by the presence of two domains in the enzyme that the substrate 4-MU-Man can bind to, namely the catalytic site and the holding site (see Figure 85).

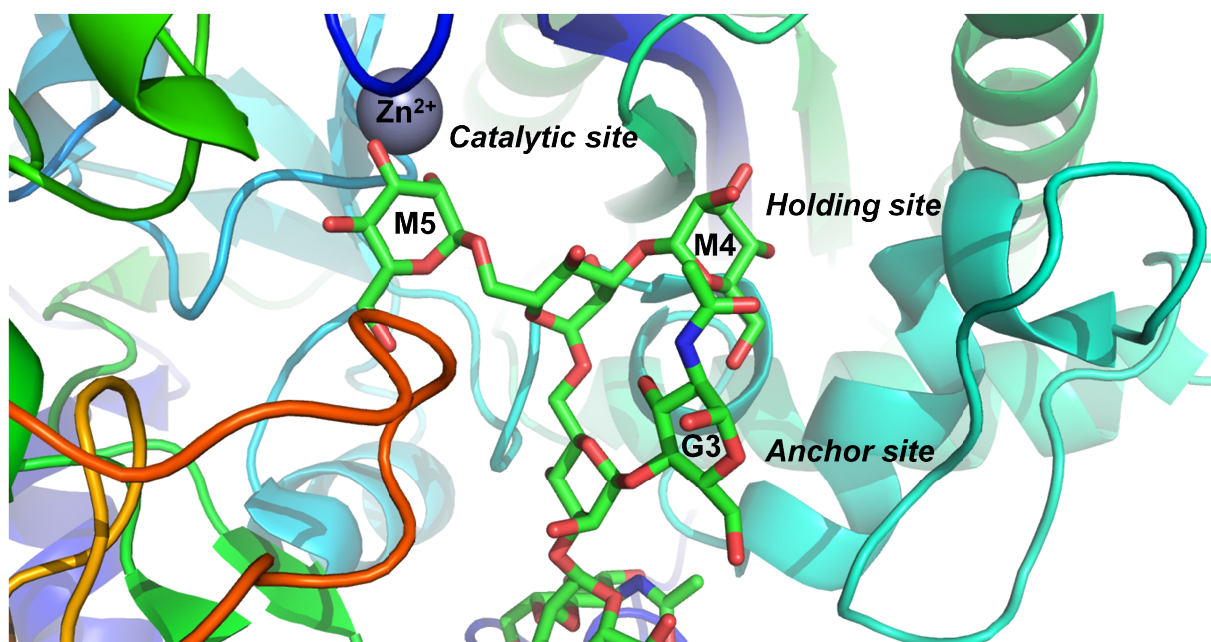


Figure 85: The domains of the active site of the dGMII (pdb code: 1HWW) [162]. The natural substrate is depicted in sticks with green carbon atoms (M (mannose) and G (N-Acetylglucosamine))

Assuming non-competitive behaviour, a K_i value of 34 ± 2 nM was determined. The reported literature values are 20 nM [29], 40 nM [173] and 44.7 nM, respectively [138]. Both the K_i value determined in the present study and the literature value lie within the lower nanomolar range. In summary, it can be concluded that the fluorometric assay can be regarded as suitable and thus also be used for further studies of the enzyme kinetics of the dGMII. As already described in the above paragraphs, the linear fits of the Lineweaver-Burk indicate non-competitive inhibition of the enzyme (see Figure 65, section 4.5.2.1.4). This is the case for a non-competitive inhibitor with $\alpha > 1$ [113]. With the help of α , it can be determined if the inhibitor preferentially binds to the free enzyme or to the enzyme-substrate complex (ES complex) [113].

α is defined as $\frac{K_{i'}}{K_i}$, where $K_{i'}$ is the dissociation constant of the enzyme-substrate-inhibitor complex (ESI complex) and K_i is the dissociation constant of the enzyme-inhibitor complex (EI complex). If $\alpha > 1$, the inhibitor binds preferentially to the ES complex [113]. Thus, swainsonine seems to also favourably bind the ES complex.

K_i and K_i^* value of swainsonine as slow binder

The mean values for K_i and K_i^* determined in this study are 15 ± 2 nM and 33 ± 4 nM, respectively (see Table 5, section 3.7.2.1.4). This means that the isomerized EI complex (E^*I complex) displays a higher affinity between enzyme and inhibitor than the initially formed enzyme-inhibitor complex (EI complex). A comparison with the K_i value determined in section

4.5.2.1.4 (34 ± 2 nM) clearly shows that this value is in the same range as the value calculated when assuming that swainsonine acts as a slow binder.

5.2.1.5 DMSO effect on the activity of dGMII

Significant decreases of enzyme activity were mainly observed in the presence of 10% DMSO, whereas the addition of 5% DMSO caused no perturbation. A too high DMSO concentration has a considerable effect on enzyme activity. It is important to keep the DMSO concentration as low as possible (generally $\leq 5\%$) and equally high in the assay mixture. The inhibition of the compounds was calculated using a positive control containing the same concentration of DMSO (5% (v/v)). The only exceptions were assays performed in the presence of swainsonine and isofagomine, which were completely dissolved in buffer for the assays and were tested using a positive control containing 0% DMSO.

5.2.1.6 Z'- factor determination

The Z' factor values were in the range $1 > Z \geq 0.5$. According to the classification of Zhang *et al.* [120], the assay can be considered as well-suited for its purpose. The separation band is large enough as depicted in Figure 67 (section 4.5.2.1.6). Nevertheless, there is still potential for optimization. The determined Z'-factor is at the lower limit. A possibility to improve the quality of the assay is to increase the volume of the components of the assay in order to have a final volume of 100-200 μL . Small volumes are more difficult to pipette accurately than large volumes. The volume of 50 μL was chosen in order to save both substrate and enzyme. Only small volumes of the inhibitor solutions (2.5 μL) were pipetted in order to not exceed the appropriate DMSO concentration of 5% (v/v) per well. As part of her bachelor thesis, [REDACTED] (JGU, Mainz) tried to improve the Z'-factor by using a single-channel pipette for the addition of all components of the enzyme assay and thus avoiding the pipetting errors that may be caused by the multichannel pipette. The Z'-factor value was in the negative range and thus an indication that the assay is not reliable and makes the screening of compounds impossible. This means that the use of a single-channel pipette has a negative influence on the value of the Z'-factor. A possible reason for this might be the fact that the incubation time varies because the handling of the single-channel pipette requires significantly more time than the use of a multi-channel pipette.

5.2.2 Enzyme assays with JBM

5.2.2.1 K_M value determination

The determined K_M value is 0.4 ± 0.05 mM (see section 4.5.2.2.1). The reference K_M value of JBM in the literature for the substrate pNP-Man is 2.5 mM [62]. Michael Juchum (diploma student, JGU Mainz, 2012) [155] established the photometric assay with JBM and pNP-Man as a substrate and determined a K_M value of 1.05 ± 0.08 mM. Therefore, 4-MU-Man has a slightly higher affinity to JBM than pNP-Man (2.5 fold, see section 5.2.1.1). In the presented work, we decided to replace the photometric assay by a fluorometric assay due to several reasons. First of all, fluorescence measurements are more sensitive than the absorbance measurements (about hundredfold) [174]. Furthermore, the fluorescence assay enables a continuous measurement of the released fluorescence and is not a one-point measurement as described for the photometric assay, where a stop reaction preceded the determination (see 1.7). Thus, for fluorescence measurements it is possible to determine the velocity of an enzyme reaction and to obtain the complete reaction curve. The velocity is defined by the slope of the linear part of the progress curve. In the case of the photometric assay of JBM using pNP-Man as substrate, the reaction is stopped with sodium carbonate in order to convert pNP into the deprotonated phenolate, whose absorption can be measured at 405 nm. To obtain the velocity, a line is drawn through the blank value and the single measured value. It is important to note that the measured values must be within the linear range. Only then, correct results can be obtained. In the case of a continuous assay, the velocity is derived from the slope of the linear portion of the reaction progress curve. As depicted in Figure 86, single determinations can lead to uncertain results.

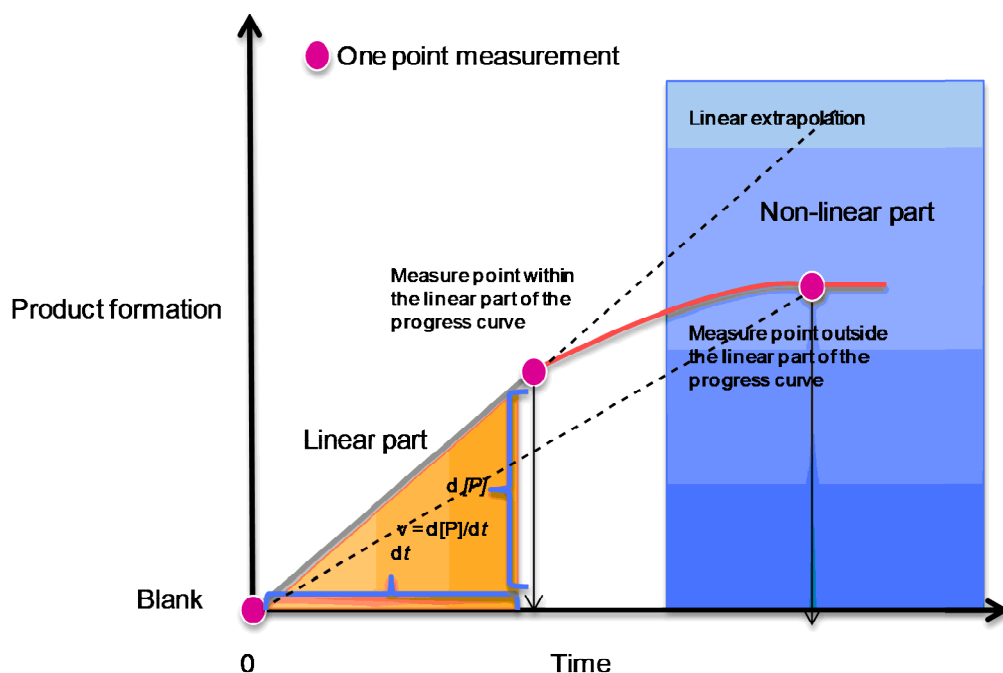


Figure 86: Progress curve of a typical enzyme reaction after Bisswanger H. [174]

5.2.2.2 IC₅₀ determination of swainsonine

Table 25 shows the IC₅₀ values of swainsonine obtained for JBM at different substrate concentrations (0.05 mM, 0.1 mM, 0.2 mM, 0.3 mM and 0.7 mM). These vary between 141 and 278 nM. Michael Juchum [155] (diploma student, [REDACTED], JGU, Mainz, 2012) determined an IC₅₀ value of 690 ± 30 nM using the photometric assay and pNP-Man as substrate [155]. Fiaux *et al.* [29] reported a competitive inhibition mode for swainsonine, an IC₅₀ of 40 nM with 5 min incubation time using pNP-Man as substrate. Tulsiani *et al.* [79] published an IC₅₀ value of 400 nM with incubation time of 5 min using pNP-Man as substrate and reported a competitive mode of action for swainsonine at inhibitor concentration lower or equal to 500 nM. Our results are closer to the results reported by Tulsiani *et al.* [79]. Fiaux *et al.* [29] incubated swainsonine and the enzyme for 5 minutes at 20 °C during their investigations. After the addition of the substrate, the enzymatic reaction was incubated for 20 minutes and then stopped by adding sodium borate buffer. An explanation for the lower IC₅₀-value reported by Fiaux *et al.* [29] could be that the one point measurements were not in the linear range of the progression curve (see Figure 86) due to the incubation time of 20 min and consequently unrealistic data can be obtained.

Assuming competitive inhibition, the K_i value can be calculated using the Cheng-Prusoff equation (see Equation 20, section 3.7.2.4):

$$IC_{50} = \left(1 + \frac{[S]}{K_m}\right) \cdot K_i.$$

At a substrate concentration of 0.1 mM and assuming a competitive mode of inhibition, after an incubation time of 5 min the K_M -value was 0.4 mM and the IC_{50} -value for swainsonine was 160 nM. By applying the equation above, a K_i -value of 128 nM was determined.

However, considering the IC_{50} values at different substrate concentrations, we can postulate that swainsonine is a non-competitive inhibitor of JBM. The IC_{50} values remained nearly unchanged, regardless of the substrate concentration. In the case of competitive inhibition, the opposite is expected. Here, the IC_{50} values are strongly dependent on the substrate concentration.

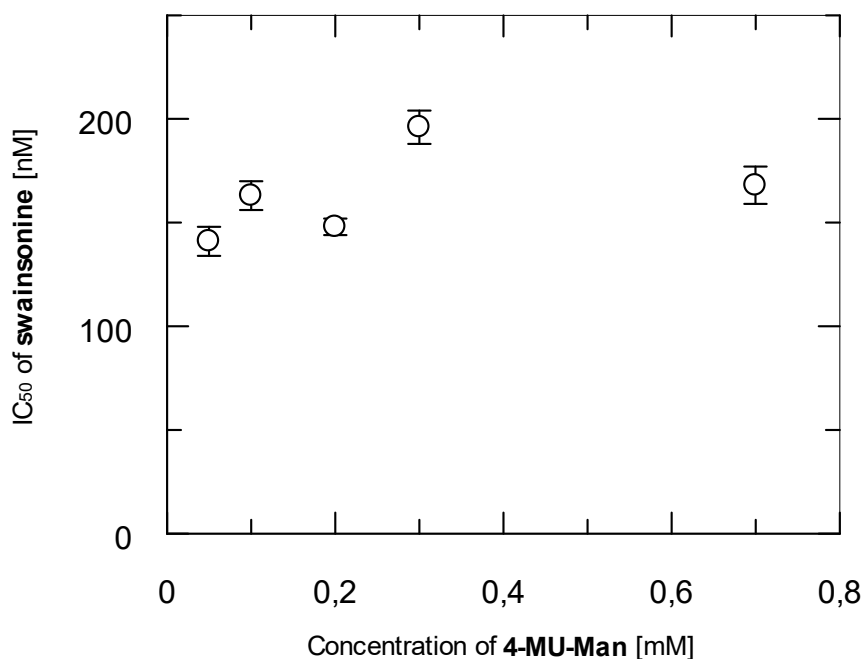


Figure 87: Plot of the determined IC_{50} values of swainsonine at different substrate concentration of 4-MU-Man.

5.2.2.3 K_i value of swainsonine as slow binder

A K_M value of 0.7 ± 0.2 mM was determined for the enzyme reaction of JBM with 4-MU-Man (see section 4.5.2.2.1). The K_M' value of this reaction in the presence of swainsonine is 1.1 ± 0.2 mM (see Table 26). Consequently, the K_M value has not changed considerably in the presence of swainsonine. A comparison of the v_{max} values with and without addition of swainsonine clearly shows that in the presence of swainsonine the v_{max} value decreases (without swainsonine: $v_{max} = 11 \pm 1.2$ F / s, with swainsonine: $v_{max} = 3.7 \pm 0.3$ F / s, see Table 26). Due to the fact that the K_M value does not change and the v_{max} value is reduced, it can be concluded, that under the given assay conditions, swainsonine acts as a non-competitive inhibitor of JBM. Moreover, the non-competitive fit matches the data better than the competitive one (see Figure 78 and Figure 79). The Lineweaver-Burk plot (see Figure 80) indicates non-competitive behaviour due to the position of the intersection of the lines. In this special case, α is equal to 1 and it can be deduced whether the inhibitor preferentially binds to the free enzyme or to the enzyme-substrate complex in the same manner as described in section 5.2.1.4. Swainsonine displays the same affinity to the free enzyme and to the enzyme-substrate complex, so that both, the ESI and the EI complex, are formed. As already described above, in the literature swainsonine is commonly described as a competitive inhibitor of the JBM [64]. Tulsiani *et al.* [79] reported that swainsonine exhibits competitive behaviour at a concentration ≤ 0.5 μ M and non-competitive behaviour at a concentration higher than 0.5 μ M. This switch of inhibition mode was never observed in our enzyme assays. Assuming a non-competitive mode of inhibition, a K_i value of 58 ± 9 nM was determined (see Table 28). Michael Juchum [155] determined swainsonine to be a non-competitive inhibitor of JBM with a K_i value of 400 ± 26 nM. The K_i value determined by Fiaux *et al.* is 40 nM [29]. The determined K_i value of 58 ± 9 nM in the fluorometric assay is in the same range as the one reported in the literature.

K_i and K_i^* value of swainsonine as slow binder

The mean values determined for K_i and K_i^* are 177 ± 11 nM and 127 ± 8 nM, respectively (see Table 28, section 4.5.2.2.3) Therefore, it can be concluded that the isomerized EI complex (E^*I complex) displays a higher affinity between enzyme and inhibitor than the initially formed enzyme-inhibitor complex (EI complex). A comparison with the K_i value of 58 ± 9 nM (Table 28, section 4.5.2.2.3) shows that the determined value assuming swainsonine to be a slow- binder (177 ± 11 nM) (Table 29, section 4.5.2.2.3) is in the same nanomolar range.

5.2.3 Enzyme assays with the β -glucosidase from sweet almonds

5.2.3.1 K_M Values determination

From three independent measurements, a mean value of 2.4 ± 0.5 mM for the K_M and a v_{max} value of 2.4 ± 1.7 F/s were determined (see Table 30, section 4.5.2.3.1). According to these results, the substrate concentration of 4-MU-Glc was fixed at 1 mM per well, since this concentration ensures good enzyme activity. Michael Juchum [155] established the photometric assay first with PNPG-Man as substrate and determined a K_M value of 5.0 ± 0.2 mM. In the literature, the photometric assay is already described and Byers *et al.* [175] reported a K_M value of 2.5 mM. For the reasons mentioned in section 5.2.2.1, the photometric assay was subsequently replaced by a fluorometric assay. Nevertheless, it is noticeable that the β -glucosidase from sweet almond displays approximately the same affinity for PNPG-Man as for 4-MU-Man. This means that the determined IC_{50} values herein can be compared with the literature, since the inhibition is competitive. Isofagomine is a known inhibitor of the β -glucosidase and was synthesized by [REDACTED] (PhD student, JGU, AK [REDACTED], Mainz). The aim of the described assays with β -glucosidase was to reproduce the IC_{50} and K_i values from the literature and preliminary work in order to verify the suitability of the assay.

5.2.3.2 IC_{50} determination of isofagomine

An IC_{50} value of 122.8 ± 25.0 nM was determined for isofagomine (see Figure 82, section 4.5.2.3.2). The mean value of both measurements resulted in an IC_{50} value of 174.7 ± 73.4 nM. The Cheng-Prusoff [116] equation (see Equation 20, section 3.7.2.4) can be used to calculate the K_i value:

$$IC_{50} = \left(1 + \frac{[S]}{K_M} \right) \cdot K_i$$

With the previously determined K_M and IC_{50} values at a substrate concentration of 1 mM a K_i value of 123.3 nM was calculated for isofagomine when assuming competitive inhibition. In the literature, a K_i value of isofagomine against sweet almonds β -glucosidase of 110 nM is described by Sierks *et al.* [52]. Therefore, the determined K_i values in the present study are in the same range as the literature K_i values. It can be concluded that this assay is suitable for its intended purpose to detect the inhibitory activity of test compounds

5.3 Discussion of the virtual screening results

5.3.1 Virtual ligand screening of the catalytic site of the homology model of hGMII

In the following paragraphs, the results of the virtual screening with the homology model of hGMII are discussed.

For this virtual screening, we decided to choose a suitable cut-off for the inhibitory potency for the hit identification. At 1 mM concentration per well of the compound in the assay system, at least 50 % inhibition should be observed. This screening of the active site of the homology model of hGMII resulted in the identification of LiSA 1 ((*R*)-(-)-3-Chloromandelic acid) as a weak inhibitor in the established photometric assay with JBM (see section 3.7.1.1.1). At a testing concentration of 1 mM per well, an inhibition level of 52 ± 3 % was obtained.

It is assumed that potent inhibitors of GMII must be able to mimic the oxocarbenium-like transition state [44], [55], [70]. Therefore, especially compounds with protonated groups at physiological pH which potentially form an electrostatic interaction with the nucleophilic residue Asp289 (corresponding to Asp204 in dGMII) were selected. Considering the interactions of swainsonine formed with dGMII (1HWW), it becomes evident that eight interactions (see Figure 16, section 1.5) constitute the binding mode and are responsible for the high affinity. As already described in section 1.2.1, the active site consists of a central Zn^{2+} -ion, which is coordinated by the four amino acid side-chains of His90, Asp92, Asp204 and His471 (dGMII). With respect to the homology model, coordination of the zinc ion is mediated via His175, Asp177, Asp289 and His569. In X-ray structural data, it has been observed so far that inhibitors and substrate analogues in the catalytic site show at least a fifth, and occasionally a sixth, coordinate to Zn^{2+} [31], [176]–[178]. Thus, the absence of six-fold coordination of Zn^{2+} would maybe also explain the weak binding behaviour of the virtual screening compounds because only monodentate binding to the Zn^{2+} -ion was predicted. It was also attempted to select those molecules with similar properties as known GMII ligands such as swainsonine or deoxynojirimycin. The inhibitory activity of these known compounds is due to a large contribution of hydrophobic interactions involving the aromatic residues, several interactions with the hydrophilic amino acids and Zn^{2+} -coordination.

Englebienne *et al.* [179] described the challenges for the accurate docking into the GMII active site. According to this study, docking accuracy would be connected with "*a proper handling of the zinc ligation and solvation/desolvation by both the conformational sampling*

engine and the scoring function". This might explain the inability of the scoring function of FlexX to differentiate between true binders and non-active compounds. As can be seen from Table 7 (section 4.5.1.1.1), almost all compounds exhibited no inhibitory activity against JBM, the model enzyme used in the assays.

The database was filtered for compounds with drug-like properties (see Table 3, section 3.4.3). This filter step helped to reduce the database size. With regard to the binding site properties and the structure of the known ligand swainsonine, the selection criteria should be specifically adjusted and not be determined by simply applying well-known property filters such as the rule-of-five (RO5) [180]. In particular, the threshold for the molecular weight of the compounds should be limited to only 350 g/mol considering that swainsonine has a molecular weight of only 173.21 g/mol and that the catalytic site is a very small cavity within the GMII protein as described in section 4.3.2. The pharmacophore search can also be seen as possible source of error, since it might be necessary to further tighten the requirements for potential hits passing the search step. Indeed, only four out of seven features were required to be met in the present pharmacophore search. If for example six out of seven features were required to pass the pharmacophore search, maybe the interaction network of swainsonine would have been reflected better in hit compounds.

Possible reasons for the differences between the docking prediction and the experimental validation for specific compounds are discussed in the following. Upon closer inspection of the predicted binding modes, some unfavourable properties were observed. The predicted binding modes of compounds LiSA 17 and LiSA 21 indicate that the hydrophobic parts of the molecules were directed toward the solvent without interacting with hydrophobic residues (see Figure 88). For LiSA 22 and LiSA 13, a similar unfavourable binding mode was postulated. The binding modes of such solvent-exposed substituents would also explain the unfavourable Hyde assessment for these compounds (see Table 7, section 4.5.1.1.1).

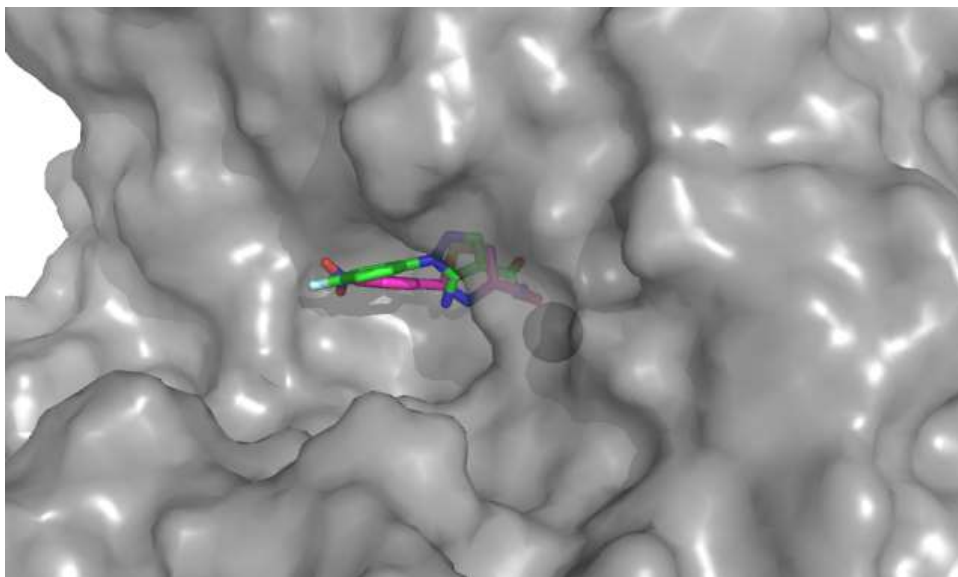
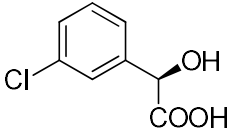
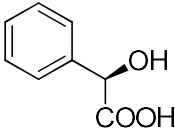
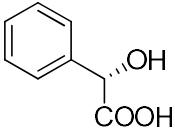


Figure 88: Hydrophobic moieties of LiSA 17 (green carbon atoms) and LiSA 21(magenta carbon atoms). hGMII is shown as grey surface

Collectively, all of the purchased compounds showed no or only weak (LiSA 1) inhibition of JBM. However, since LiSA 1 is very small, it was supposed to be suitable as a first starting point for the synthesis of derivatives. In order to verify the observed inhibition in the photometric assay, Malena dos Santos Guilherme (diploma student 2016, JGU, AK [REDACTED]) [168] tested LiSA 1 and R-(-) as well as S-(+)-mandelic acid in the fluorometric assay (see section 1.7) with dGMII in order to potentially elucidate initial structure-activity relationship results. The following table (see Table 32) summarizes the inhibition of dGMII by the respective mandelic acid derivatives.

Table 32: Inhibition of dGMII by 3- (R) -chloromandelic acid (LiSA 1), (R) -mandelic acid and (S) -mandelic acid at a well concentration of 500 μ M

Compound	Structure	Inhibition of dGMII at 500 μ M in the well (%)
(R)-(-)-3-Chloromandelic acid (LiSA 1)		No inhibition
(R)-(-)-mandelic acid		No inhibition
(S)-(+)-mandelic acid		No inhibition

From these results, it becomes clear that LiSA 1 and the derivatives did not exhibit inhibitory activity in the fluorometric assay. Even at a higher test concentration of 1.2 mM for LiSA 1, no inhibition could be observed [168]. It can be concluded that the observed differences are most likely based on the different assay methods. The fluorometric assay is more sensitive as the photometric assay. In order not to completely reject the potential of LiSA 1 to act as a hGMII inhibitor, derivatives of this compound were synthesized by Malena dos Santos Guilherme [168] and tested in the enzyme assay for inhibition of dGMII at a well concentration of 500 μ M [168]. However, these derivatives exhibited no inhibition of dGMII. The structures of the derivatives together with their corresponding identification codes are depicted in Figure 89.

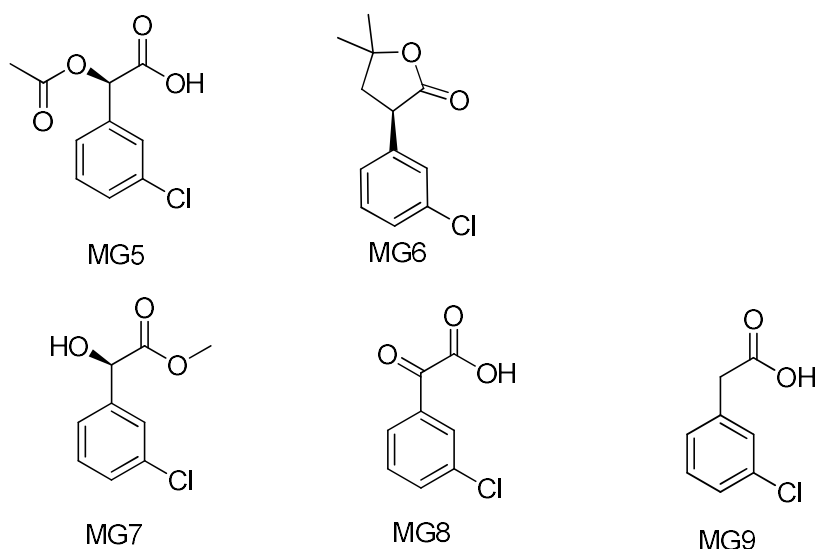


Figure 89: Structures of the derivatives of (R)-(-)-3-Chloromandelic acid

With these results in our hands, we decided to shelve this project. In the fluorometric assay there were no indications for an inhibitory effect of LiSA 1

Thus, for the search of GMII inhibitors, a new approach had to be considered. The virtual screening described above was carried out with the homology model of hGMII based on the 3D structure of dGMII. The purchased compounds were initially tested against JBM. As described in section 4.1.1, hGMII and dGMII share a high sequence identity and similarity. Although this is desired for the development of a reliable homology model, still a possible option for improvement of the docking studies would be to carry out the virtual screening based on a crystal structure rather than utilizing a homology model, which only represents an approximation of the structure. It is widely accepted that docking based on homology models is less successful than docking with a crystal structure. However, in the literature, protein homology models that have been successfully used as receptor for virtual screening approaches resulting in new inhibitors or lead compounds for several protein targets are described. The accuracy of the homology model is of particular importance. Generally, models of proteins which share more than 50% sequence identity to their templates are considered as high-accuracy comparative models and they tend to have about 1 Å RMS error for the main-chain atoms [181]. In the present case, the homology model is at the threshold of a high-accuracy comparative model and can be considered as a medium-accuracy comparative model with its sequence identity of 51%. Medium-accuracy comparative models are based on 30-50% sequence identity and they tend to have about 90% of the main-chain atoms modelled with 1.5 Å RMS error [181]. In medium-accuracy comparative models, side-chain packing, core distortion, and loop modelling errors as well as alignment errors can occasionally occur [181]. It can be concluded that a high quality homology model or a crystal structure should be used as structural basis for the next virtual screening approach.

Furthermore, for a virtual screening approach, a larger database containing more than a drug-like database of two suppliers could be used as starting point. This guarantees a larger diversity of the structures. Besides screening libraries from supplier such as Sigma Aldrich and TimTec, which were used in the present study, it is conceivable to also include further commercial libraries from suppliers like Chembridge (www.chembridge.com), Asinex (www.asinex.com), or Specs (www.specs.net) etc.

5.3.2 Virtual ligand screening of the potential allosteric site of dGMII

We described the identification of a potential allosteric site using the site finder application of MOE v2014.10. The Docksite Scorer predicted a higher druggability for this site in comparison to the known active site. We performed a virtual screening using DOCK 3.6 with the aim to identify new chemical scaffolds possibly interacting with this allosteric binding site. These studies revealed compound **18** as a first weak inhibitor of GMII (see Figure 68, section 4.5.2.1.8). Since the IC_{50} value was independent of the substrate concentration (see Figure 69, section 4.5.2.1.8) compound **18** is supposed to act in a non-competitive manner with regard to the substrate, supporting the assumption that it binds to the identified allosteric site. In the following figure (see Figure 90), the postulated binding mode of compound **18** is depicted.

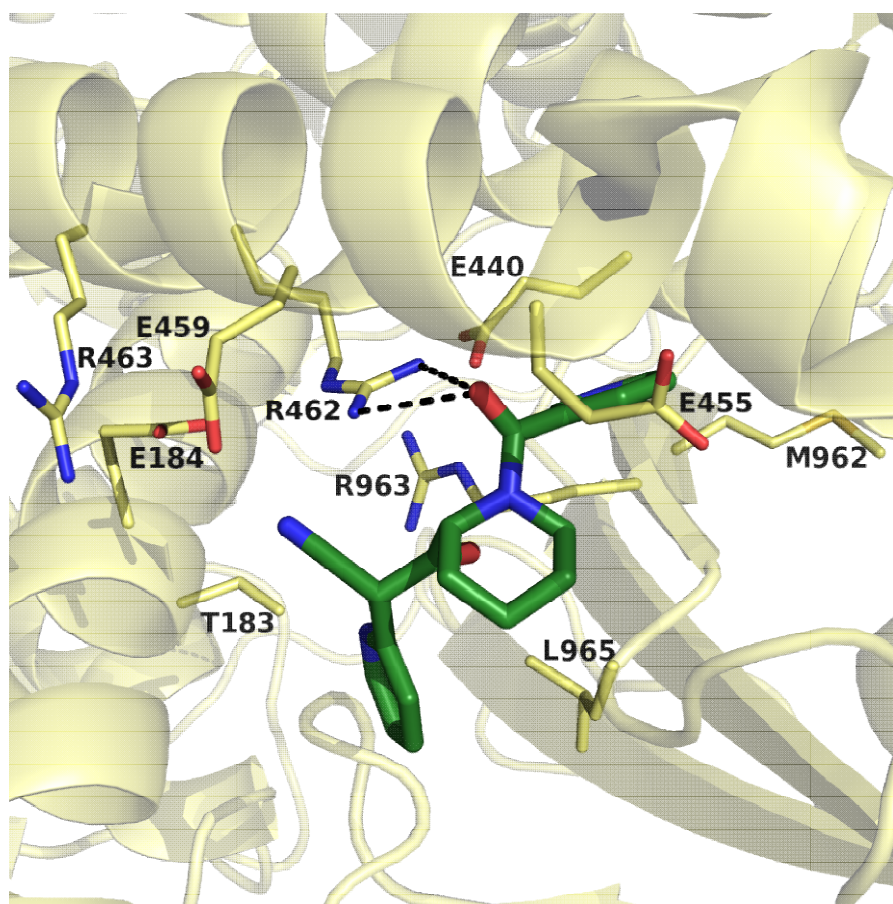
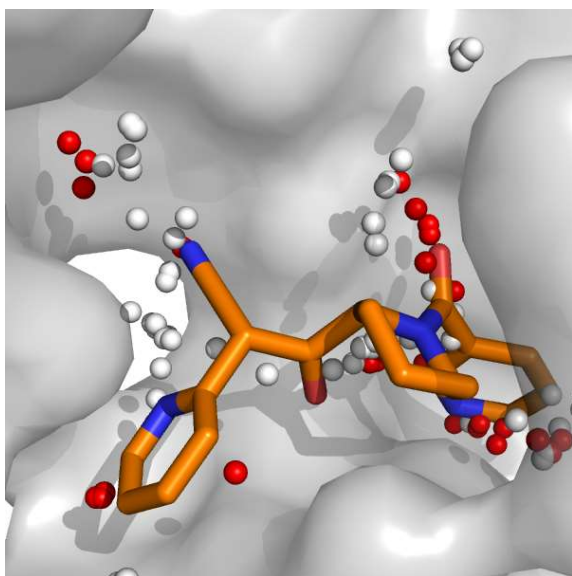


Figure 90: Predicted binding mode for compound **18**. The GMII protein (PDB code 1HWW) is shown as *transparent, yellow-coloured cartoon* with binding site residues being displayed as stick representation. Compound **18** is represented as stick representation with dark-green coloured carbon atoms. *Black dashes* indicate potential hydrogen bonds between the virtual screening hit and the target protein.

In the predicted binding mode of compound **18** a bidentate hydrogen bond between the carbonyl oxygen of the amide bond and the protonated side chain of R462 is postulated (see Figure 90). The properties of the ligand atoms also match the properties of the dummy atoms created during binding site identification (see Figure 91, (A)). Furthermore, positions of the dummy atoms indicate areas that could be addressed during the optimization process. Moreover, the atom positions of compound **18** fit very well with the dummy atoms positions.

(A)



(B)

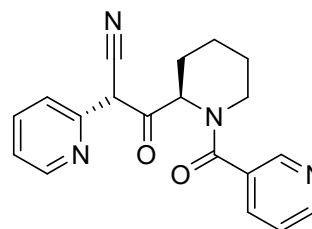


Figure 91: (A) Comparison of the locations of alpha spheres and compound **18** atoms: red alpha spheres indicate regions where a hydrophilic type of atoms is favoured by the binding site, whereas white alpha spheres represent regions where a hydrophobic type of atoms is favoured. (B) Structure of compound **18**

The polar interaction between the carbonyl oxygen of the amide bond and the protonated side chain of R462 seems to be crucial for the postulated binding mode of compound **18** (see Figure 90).

Compound **18** did not exhibit significant inhibition of the off-target β -glucosidase of sweet almonds but showed inhibition activity against JBM (70% at 300 μ M per well). Thus, this compound can be considered as a weakly active inhibitor with some selectivity towards mannosidases in general. Taken together, we identified compound **18** as a starting point for the identification of a novel class of GMII inhibitors which possibly bind into a potential allosteric binding pocket. Based on the inhibitory activity observed in the fluorometric assay, the ligand efficiency is approximately 0.20 (kcal/mol)/HA. This is lower than the often applied threshold level of 0.30 kcal/mol for compounds that are suitable for further lead optimization [182]. Despite the rather low efficiency of compound **18**, future investigations of the binding

mode by e.g. site-directed mutagenesis studies might give valuable insights into the properties that are required for ligands interacting with the potential allosteric binding site. This in turn could guide the design of analogues with improved affinities.

5.3.3 Virtual ligand screening of the catalytic site of dGMII

In the following paragraphs, the results of the virtual screening approach targeting the catalytic site of dGMII are discussed. Before the virtual screening was performed, a redocking of swainsonine was done in order to check the model for its suitability. A RMSD-value of 1.099 Å between the co-crystallized ligand and the docking pose was determined, indicating a successful pose prediction. However, as described by Cole *et al.* [167], numerical measurements such as RMSD should be interpreted with caution. Small ligands, such as swainsonine, "give easily low RMSDs even when randomly placed" [167].

Three out of the thirty purchased compounds exhibited weak inhibition at concentrations of 100 µM per well (compound **32** and compound **33**) and 300 µM per well (compound **48**), respectively. Compound **32** showed an IC₅₀-value of 196.5 ± 13.9 µM at 1 mM substrate concentration per well. In the predicted binding mode of compound **32** a hydrophobic interactions was postulated (see Figure 93). The aromatic ring of the indoline structure of compound **32** formed π-π-stacking interactions to W95. In addition the indoline structure of showed a good shape complementarity to the site. The chlorophenyl group of compound **32** stucked out of the catalytic site. Thus no interactions were predicted for this moiety.

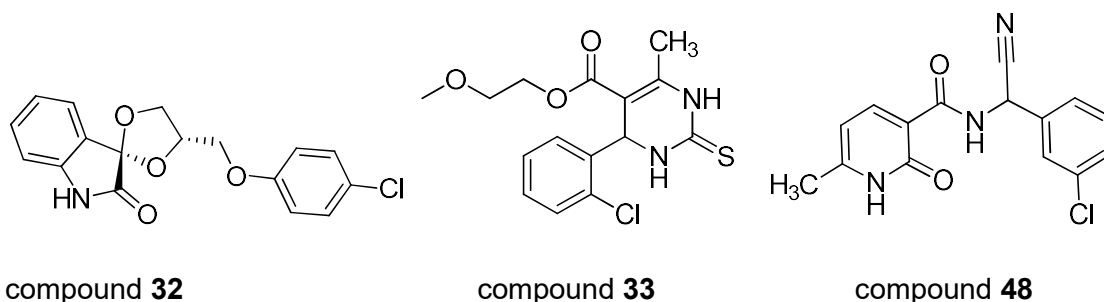


Figure 92: Structures of the virtual screenings hits of the catalytic site off dGMII

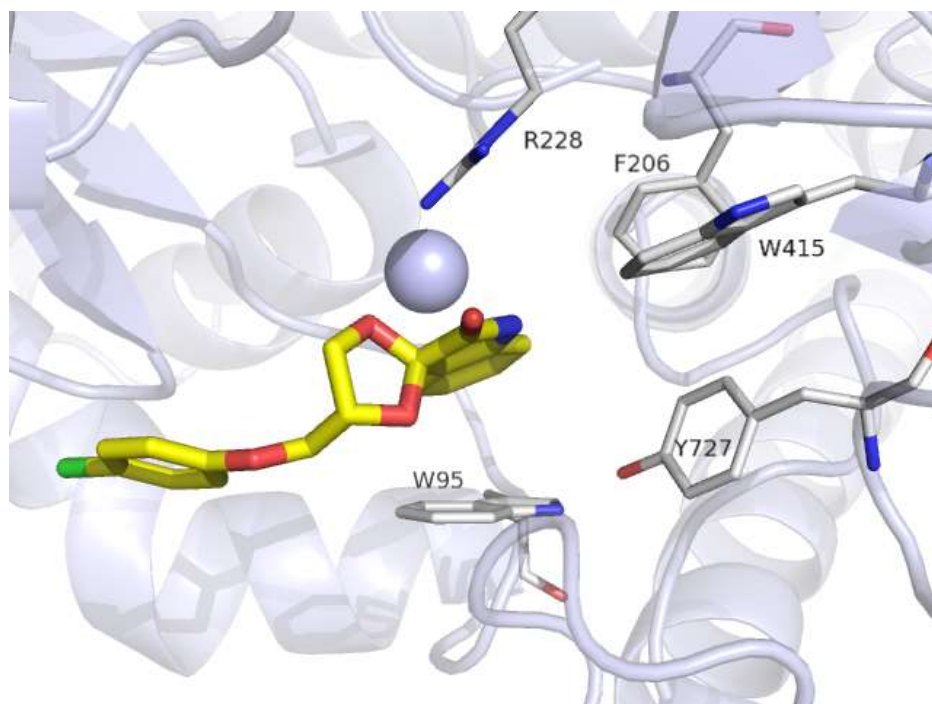


Figure 93: Predicted binding mode for compound **32**. The GMII protein (pdb code 1HWW) is shown as transparent, grey-coloured cartoon with binding site residues being displayed as stick. Compound **32** is represented with yellow coloured carbon atoms.

In the enzymatic assays, the K_M value determined in the absence of compound **32** was 1.0 ± 0.3 mM, while v_{max} was 1.0 ± 0.1 F/s. In the presence of compound **32**, a K_M value of 2.0 ± 0.5 mM and a v_{max} value of 0.4 ± 0.03 F/s were determined (see Table 22, section 4.5.2.1.10). Considering these values, it becomes evident that a significant change in the K_M value could not be observed. In contrast, the v_{max} value is reduced by approximately 50 % in presence of compound **32**. These facts are indicative of a non-competitive mode of inhibition of dGMII. Moreover, this assumption is confirmed by the fact that the non-competitive fit matches the data better than the competitive one (see Figure 72 and Figure 73, section 4.5.2.1.9). Moreover, the Lineweaver-Burk plot is also indicative of non-competitive behaviour of compound **32**. A K_i value of 62.9 ± 7 μ M under the assumption of a non-competitive mode of inhibition could be determined (see Table 23, 4.5.2.1.10).

The mode of inhibition of compound **32** was also tested by the determination of IC_{50} values at different substrate concentrations. In Table 33, the IC_{50} values and K_i values are listed.

Table 33: Effect of different substrate concentrations on the IC₅₀ values of compound **32**

[4-MU-Man] Substrate concentration	IC₅₀
0.1 mM	217.3 ± 26,9 μM
1.0 mM	196.5 ± 13.9 μM
2.0 mM	322.4 ± 50.8 μM

As can be concluded from Table 33, the IC₅₀ values increase from 217.3 ± 26.9 μM (0.1 mM 4-Mu-Man) to 322.4 ± 50.8 μM (2.0 mM [4-Mu-Man]). Considering the values of the IC₅₀ determinations at 0.1 mM and 1.0 mM substrate concentration, no significant deviation can be observed. With these data in our hands, a non-competitive type of inhibition can be assumed for compound **32**.

6 Conclusion and future perspectives

Conclusion and Outlook

In the present study, GMII, an enzyme involved in the *N*-glycosylation pathway, was chosen as target for the discovery of new starting points for the design of anti-cancer drugs. The aim of this work was to find novel inhibitors of GMII with chemical scaffolds. For this purpose, different tools of structure-based drug design were used. In total, three virtual screenings were performed. The docking studies were carried out against the catalytic site of the homology model of hGMII and against the known catalytic site and a potential allosteric site of dGMII. The most promising candidates were purchased and subjected to enzyme inhibition assays yielding several weakly active compounds. For this purpose, several fluorometric assays were established, namely for JBM, dGMII and the β -glucosidase from sweet almonds. Two compounds were identified with micromolar affinity. In the following scheme (see Figure 94), the overall approach of this work is depicted:

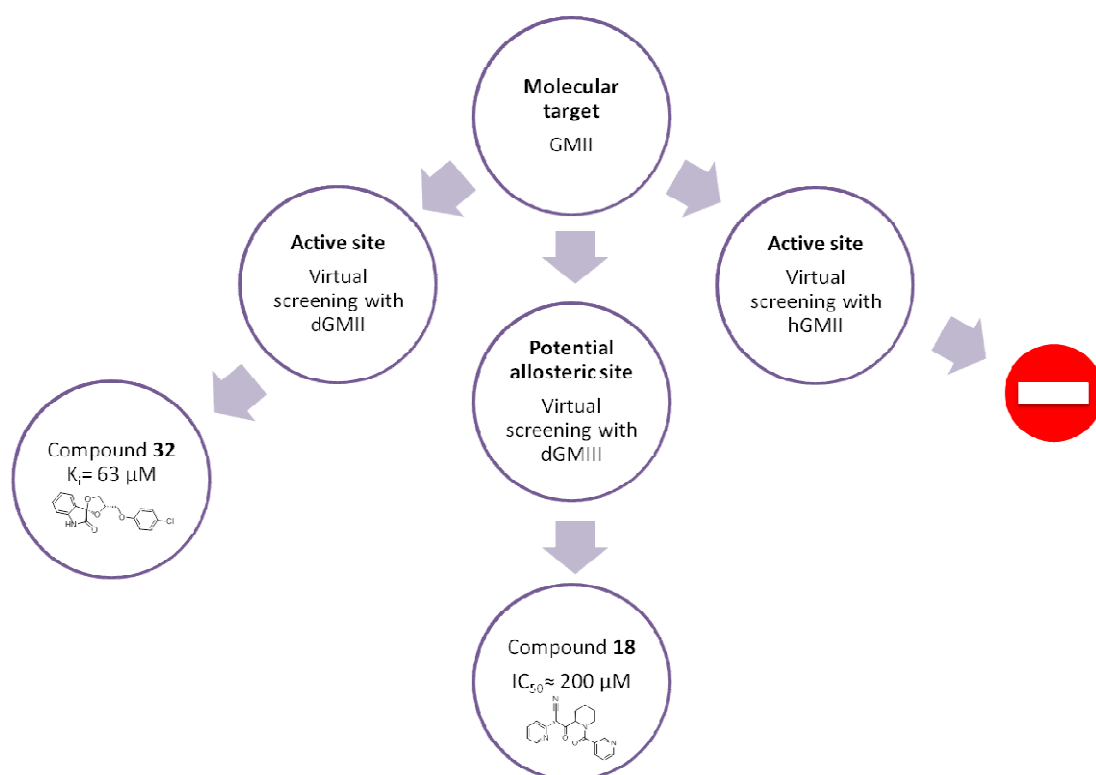


Figure 94: Overall approach of the thesis

Virtual screening against the catalytic site of hGMII

In the case of the virtual screening with the homology model of hGMII, no active compound could be identified. The possible reasons for this are already described in section 5.3.1. The size of the library, the number of the tested compounds and the use of the homology model, which only represents an approximation of the true 3D structure, could be the reasons for the unsuccessful result. The goal of this screening was to find an initial hit compound with potential for structural optimization. Compound LiSA 1, originally found to be weakly active in the photometric assay, showed no inhibitory activity at 1 mM per well in the fluorometric assay and thus we decided not to change the defined activity cut-off. Conclusively, we stopped further pursuing this project and sought a completely new approach for the identification of GMII inhibitors.

Virtual screening against a potential allosteric site of dGMII

Since the catalytic site has been studied in detail and it has been proven to be hard to be addressed by molecular modelling, we had the idea to identify a potential allosteric site [179]. In a second approach, we performed an extensive virtual screening using DOCK 3.6 in order to identify new allosteric inhibitors. The identified new GMII inhibitor, compound **18** (see Figure 90, section 5.3.2), is presumably acting as a non-competitive inhibitor with a potency in the micromolar range ($IC_{50} = 216.7 \pm 44.3 \mu\text{M}$) which was proven by various assays. The determined mode of inhibition confirms our proposed model of compounds that should bind to a potential allosteric pocket. A subsequent step will be the synthesis and the purchase of analogues of compound **18** based on further docking studies in order to elucidate the structure-activity relationship and increase the affinity of the inhibitors. This could be achieved for instance by addressing the acid side chains of Glu459 and/or Glu440. Another possibility might be to modify compound **18** in such a manner that nucleophilic amino acids of the allosteric site such as Thr187 or Glu440 are addressed. Finally, in order to further confirm the existence of the potential allosteric site, co-crystallisation experiments of GMII with the identified inhibitors will be performed.

The rational de novo design of inhibitors targeting this potential allosteric site is more challenging than the optimization of an already known inhibitor class or the screening for novel ligands binding to a site, for which ligand-bound three-dimensional structures are available. This was the major challenge in this project. A robust fluorometric was performed in order to identify novel ligands with weak affinity. Nevertheless, an orthogonal method in order to confirm the initial screening hit is required. For this purpose, a SPR (surface plasmon resonance) or ITC (Isothermal Titration Calorimetry)-based method for example could be established. For one compound purchased on the basis of this virtual screening, the

autofluorescence compromised the evaluation of the results of the fluorometric assay. Thus, a secondary assay is also important for testing such compounds.

Collectively, the identified weak inhibitor compound **18** might have a great potential for optimisation and will represent a good starting point for the design of new allosteric inhibitors of GMII in the anti-cancer therapy.

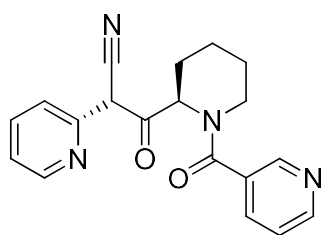
Virtual screening against the catalytic site of dGMII

In a third approach, a virtual screening in order to identify new scaffolds for active site directed inhibitors was performed using DOCK 3.6. It was considered as useful to again perform a virtual screening against the catalytic site using a crystal structure rather than a homology model. Moreover, a larger database of 5,574,182 commercially available small molecules containing more than a drug-like database of two suppliers (10,000 compounds Sigma Aldrich and TimTec, see section 3.4.3) was used. This guaranteed a larger diversity of the structures. As extensively described in this work (see section 3.5.2), the catalytic site is relatively small, hydrophilic and thus challenging for molecular docking approaches. This virtual screening initially yielded several weakly active inhibitors with IC_{50} values ranging from 200 to 400 μ M. Finally, the focus was put onto compound **32** ($IC_{50} = 196.5 \pm 13.9 \mu$ M, $K_i = 62.9 \pm 7 \mu$ M, see section 5.3.3). A non-competitive mode of inhibition of compound **32** was determined (see section 5.3.3), even though this compound was found through the catalytic site screening approach. As described in section 5.2.1.4, the known inhibitor swainsonine acts as a non-competitive inhibitor of dGMII under the given assay conditions. However, swainsonine is bound into the catalytic site of dGMII (pdb-code: 1HWW [162] and 3BLB [172]) and is described in the literature as a competitive inhibitor of JBM at concentrations of ≤ 500 nM [79]. Thus, ITC experiments could clarify the binding stoichiometry of swainsonine and identify whether a potential second site, e.g. the holding site, is addressed as well. Additionally, binding competition experiments could elucidate if the same pocket is addressed by both compounds.

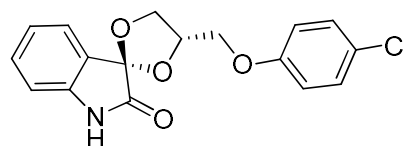
It would be also useful here, to experimentally determine the binding mode via crystallization studies of enzyme-inhibitor complexes. As already described above, a second assay using a different method needs to be established in order to validate the binding of this hit compound

Alternatively, the substrate for the enzymatic assay could be optimized. In section 5.2.1.1, it was described that the pseudosubstrate 4-MU-Man might also bind to the holding site of GMII and not only to the catalytic site. The natural substrate of GMII, $GlcNAcMan_5(GlcNAc)_2$, or a mimic of similar size could also serve as substrate for the enzymatic assays.

Finally, two new weak inhibitors of GMII were identified (compound **18** and compound **32**, see Figure 95).



compound **18**
 $IC_{50} \approx 200 \mu\text{M}$



compound **32**
 $K_i = 63 \mu\text{M}$

Figure 95: Structures of compound 18 and compound 32 and the corresponding inhibition data

The most of the prospective virtual screenings to date yielded only weak hits usually in the range of 1 to 500 μM [166]. With further optimizations, some of those virtual screenings hits led to ligands with improved affinity up to nM-range [166]. Likewise, the inhibitors presented in this work could serve as starting points for the design of new high affinity GMII inhibitors.

7 Bibliography

- [1] P. E. Goss, J. Baptiste, B. Fernandes, M. Baker, and J. W. Dennis, "A phase I study of swainsonine in patients with advanced malignancies.," *Cancer Res.*, vol. 54, no. 6, pp. 1450–7, 1994.
- [2] P. E. Goss, C. L. Reid, D. Bailey, and J. W. Dennis, "Phase IB clinical trial of the oligosaccharide processing inhibitor swainsonine in patients with advanced malignancies," *Clin. Cancer Res.*, vol. 3, no. 7, pp. 1077–1086, 1997.
- [3] I. J. Fidler, "Origin and biology of cancer metastasis," *Cytometry*, vol. 10, no. 6, pp. 673–680, 1989.
- [4] M. F. Leber and T. Efferth, "Molecular principles of cancer invasion and metastasis (Review)," *International Journal of Oncology*, vol. 34, no. 4, pp. 881–895, 2009.
- [5] R. Valastyan, S. Weinberg, "Tumor metastasis: molecular insights and evolving paradigms," *Cell*, vol. 14;147(2), no. 2, pp. 275–292, 2011.
- [6] I. Zeidman, M. McCutcheon, and D. R. Coman, "Factors Affecting the Number of Tumor Metastases Experiments with a Transplantable Mouse Tumor," *Cancer Res.*, vol. 10, no. 6, pp. 357–359, 1950.
- [7] H. S. N. Greene and E. K. Harvey, "The Relationship between the Dissemination of Tumor Cells and the Distribution of Metastases," *Cancer Res.*, vol. 24, no. 5, pp. 799–811, 1964.
- [8] K. Yoshida, T. Fujikawa, a Tanabe, and K. Sakurai, "Quantitative analysis of distribution and fate of human lung cancer emboli labeled with 125I-5-iodo-2'-deoxyuridine in nude mice.," *Surg. Today*, vol. 23, no. 11, pp. 979–83, 1993.
- [9] R. Kannagi, "Carbohydrate-mediated cell adhesion involved in hematogenous metastasis of cancer," *Glycoconj. J.*, vol. 14, no. 5, pp. 577–584, 1997.
- [10] M. L. Phillips *et al.*, "ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex.," *Science (80-.)*, vol. 250, no. 4984, pp. 1130–2, 1990.
- [11] D. Tyrrell *et al.*, "Structural requirements for the carbohydrate ligand of E-selectin," *Proc Natl Acad Sci U S A*, vol. 88, no. 22, pp. 10372–10376, 1991.
- [12] E. L. Berg, M. K. Robinson, O. Mansson, E. C. Butcher, and J. L. Magnani, "A carbohydrate domain common to both sialyl leaand sialyl le(x) is recognized by the endothelial cell leukocyte adhesion molecule ELAM-1," *J. Biol. Chem.*, vol. 266, no. 23, pp. 14869–14872, 1991.
- [13] A. Takada *et al.*, "Adhesion of human cancer cells to vascular endothelium mediated by a carbohydrate antigen, sialyl Lewis A.," *Biochem. Biophys. Res. Commun.*, vol. 179, no. 2, pp. 713–9, 1991.
- [14] G. S. Kansas, "Selectins and Their Ligands: Current Concepts and Controversies," *Blood*, vol. 88, no. 9, pp. 3259–3278, 1996.
- [15] R. Kannagi, M. Izawa, T. Koike, K. Miyazaki, and N. Kimura, "Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis," *Cancer Science*, vol. 95, no. 5, pp. 377–384, 2004.
- [16] E. Jubeli, L. Moine, J. Vergnaud-Gauduchon, and G. Barratt, "E-selectin as a target for drug delivery and molecular imaging," *J. Control. Release*, vol. 158, no. 2, pp. 194–206, 2012.
- [17] H. Läubli and L. Borsig, "Selectins promote tumor metastasis," *Semin. Cancer Biol.*, vol. 20, no. 3, pp. 169–177, 2010.
- [18] M. Sperandio, C. A. Gleissner, and K. Ley, "Glycosylation in immune cell trafficking," *Immunol Rev.*, vol. 230, no. 1, pp. 97–113, 2009.
- [19] I. Brockhausen, "Pathways of O-glycan biosynthesis in cancer cells," *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1473, no. 1, pp. 67–95, 1999.
- [20] C. Whitehouse, J. Burchell, S. Gschmeissner, I. Brockhausen, K. O. Lloyd, and J. Taylor-Papadimitriou, "A transfected sialyltransferase that is elevated in breast cancer and localizes to the medial/trans-Golgi apparatus inhibits the development of core-2-based O-glycans," *J. Cell Biol.*, vol. 137, no. 6, pp. 1229–1241, 1997.
- [21] M. Dalziel *et al.*, "The Relative Activities of the C2GnT1 and ST3Gal-I Glycosyltransferases

- Determine O-Glycan Structure and Expression of a Tumor-associated Epitope on MUC1,” *J. Biol. Chem.*, vol. 276, no. 14, pp. 11007–11015, 2001.
- [22] A. Rivinoja, F. M. Pujol, A. Hassinen, and S. Kellokumpu, “Golgi pH, its regulation and roles in human disease,” *Annals of Medicine*, vol. 44, no. 6, pp. 542–554, 2012.
- [23] J. Ogawa, H. Inoue, and S. Koide, “Expression of alpha-1,3-fucosyltransferase type IV and VII genes is related to poor prognosis in lung cancer.,” *Cancer Res.*, vol. 56, no. 2, pp. 325–329, 1996.
- [24] P. Compain and O. R. Martin, “Carbohydrate mimetics-based glycosyltransferase inhibitors,” *Bioorganic Med. Chem.*, vol. 9, no. 12, pp. 3077–3092, 2001.
- [25] W. Zou, “C -Glycosides and Aza- C -Glycosides as Potential Glycosidase and Glycosyltransferase Inhibitors,” pp. 1363–1391, 2005.
- [26] E. Walker-Nasir, I. Ahmad, M. Saleem, and D. Hoessli, “Glycosyltransferase and Glypiation Inhibitors,” *Curr. Org. Chem.*, vol. 11, no. 7, pp. 591–607, 2007.
- [27] Y. Ito, S. Hagihara, I. Matsuo, and K. Totani, “Structural approaches to the study of oligosaccharides in glycoprotein quality control,” *Current Opinion in Structural Biology*, vol. 15, no. 5, pp. 481–489, 2005.
- [28] D. R. Rose, “Structure, mechanism and inhibition of Golgi α -mannosidase II,” *Curr. Opin. Struct. Biol.*, vol. 22, no. 5, pp. 558–562, Oct. 2012.
- [29] H. Fiaux *et al.*, “Functionalized pyrrolidine inhibitors of human type II α -mannosidases as anti-cancer agents: Optimizing the fit to the active site,” *Bioorganic Med. Chem.*, vol. 16, no. 15, pp. 7337–7346, 2008.
- [30] J. M. H. Van Den Elsen, D. A. Kuntz, and D. R. Rose, “Structure of Golgi α -mannosidase II: A target for inhibition of growth and metastasis of cancer cells,” *EMBO J.*, vol. 20, no. 12, pp. 3008–3017, Jun. 2001.
- [31] S. Numao, D. A. Kuntz, S. G. Withers, and D. R. Rose, “Insights into the Mechanism of *Drosophila melanogaster* Golgi α -Mannosidase II through the Structural Analysis of Covalent Reaction Intermediates,” *J. Biol. Chem.*, vol. 278, no. 48, pp. 48074–48083, Nov. 2003.
- [32] N. Shah, D. A. Kuntz, and D. R. Rose, “Golgi alpha-mannosidase II cleaves two sugars sequentially in the same catalytic site.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 28, pp. 9570–5, 2008.
- [33] C. Rabouille *et al.*, “The *Drosophila* GMII gene encodes a Golgi alpha-mannosidase II,” *J Cell Sci*, vol. 112, pp. 3319–30, 1999.
- [34] S. Howard, S. He, and S. G. Withers, “Identification of the active site nucleophile in jack bean alpha-mannosidase using 5-fluoro-beta-L-gulosyl fluoride,” *J. Biol. Chem.*, vol. 273, no. 4, pp. 2067–2072, 1998.
- [35] B. S. Gnanesh Kumar, G. Pohlentz, M. Schulte, M. Mormann, and N. Siva Kumar, “Jack bean -mannosidase: Amino acid sequencing and N-glycosylation analysis of a valuable glycomics tool,” *Glycobiology*, vol. 24, no. 3, pp. 252–261, 2014.
- [36] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, and B. Henrissat, “The carbohydrate-active enzymes database (CAZy) in 2013,” *Nucleic Acids Res.*, vol. 42, no. D1, 2014.
- [37] A. Bateman *et al.*, “UniProt: The universal protein knowledgebase,” *Nucleic Acids Res.*, vol. 45, no. D1, pp. D158–D169, 2017.
- [38] K. W. Moremen and O. Touster, “Biosynthesis and modification of Golgi mannosidase II in HeLa and 3T3 cells,” *J. Biol. Chem.*, vol. 260, no. 11, pp. 6654–6662, 1985.
- [39] K. W. Moremen and O. Touster, “Topology of mannosidase II in rat liver Golgi membranes and release of the catalytic domain by selective proteolysis,” *J. Biol. Chem.*, vol. 261, no. 23, pp. 10945–10951, 1986.
- [40] L. Petersen, A. Ardévol, C. Rovira, and P. J. Reilly, “Molecular mechanism of the glycosylation step catalyzed by Golgi α -mannosidase II: A QM/MM metadynamics investigation,” *J. Am. Chem. Soc.*, vol. 132, no. 24, pp. 8291–8300, Jun. 2010.
- [41] D. E. Koshland, “Stereochemistry and the mechanism of enzymatic reactions,” *Biol. Rev.*, vol. 28, no. 4, pp. 416–436, 1953.
- [42] W. Zhong *et al.*, “Probing the substrate specificity of golgi α -mannosidase II by use of synthetic oligosaccharides and a catalytic nucleophile mutant,” *J. Am. Chem. Soc.*, vol. 130, no.

- 28, pp. 8975–8983, Jul. 2008.
- [43] L. Petersen, A. Ardèvol, C. Rovira, and P. J. Reilly, “Molecular Mechanism of the Glycosylation Step Catalyzed by Golgi α -Mannosidase II: A QM/MM Metadynamics Investigation,” *J. Am. Chem. Soc.*, vol. 132, no. 24, pp. 8291–8300, Jun. 2010.
- [44] A. Vasella, G. J. Davies, and M. Böhm, “Glycosidase mechanisms,” pp. 619–629, 2002.
- [45] B. Li *et al.*, “Inhibition of Golgi mannosidase II with mannosatin A analogues: Synthesis, biological evaluation, and structure-activity relationship studies,” *ChemBioChem*, vol. 5, no. 9, pp. 1220–1227, 2004.
- [46] W. Y. Jeng *et al.*, “Structural and functional analysis of three β -glucosidases from bacterium *Clostridium cellulovorans*, fungus *Trichoderma reesei* and termite *Neotermes koshunensis*,” *J. Struct. Biol.*, vol. 173, no. 1, pp. 46–56, 2011.
- [47] Y. Kacher *et al.*, “Acid beta-glucosidase: insights from structural analysis and relevance to Gaucher disease therapy,” *Biol. Chem.*, vol. 389, no. 11, pp. 1361–9, 2008.
- [48] M. J. Farrer *et al.*, “Glucosidase-beta variations and Lewy body disorders,” *Park. Relat. Disord.*, vol. 15, no. 6, pp. 414–416, 2009.
- [49] Q. Wang, D. Trimbur, R. Graham, R. A. J. Warren, and S. G. Withers, “Identification of the Acid/Base Catalyst in *Agrobacterium faecalis* β -Glucosidase by Kinetic Analysis of Mutants,” *Biochemistry*, vol. 34, no. 44, pp. 14554–14562, 1995.
- [50] T. M. Gloster, R. Madsen, and G. J. Davies, “Structural basis for cyclophellitol inhibition of a β -glucosidase,” *Org. Biomol. Chem.*, vol. 5, no. 3, pp. 444–446, 2007.
- [51] B. Acta, M. Retardation, and T. Road, “P-Glucosidase inhibition in murine peritoneal macrophages by conduritol-B-epoxide : an in vitro model of the Gaucher cell David S . Newburg *, Shaul Yatziv **, Robert H . McCluer,” vol. 877, pp. 121–126, 1986.
- [52] T. M. Jespersen, W. Dong, M. R. Sierks, T. Skrydstrup, I. Lundt, and M. Bols, “Isogagomin, ein wirksamer neuer Glycosidaseinhibitor,” *Angew. Chemie*, vol. 106, no. 17, pp. 1858–1860, 1994.
- [53] D. L. Zechel *et al.*, “Iminosugar Glycosidase Inhibitors: Structural and Thermodynamic Dissection of the Binding of Isogagomine and 1-Deoxynojirimycin to β -Glucosidases,” *J. Am. Chem. Soc.*, vol. 125, no. 47, pp. 14313–14323, 2003.
- [54] E. Borges de Melo, A. da Silveira Gomes, and I. Carvalho, “ α - and β -Glucosidase inhibitors: chemical structure and biological activity,” *Tetrahedron*, vol. 62, no. 44, pp. 10277–10302, 2006.
- [55] T. D. Heightman and A. T. Vasella, “Recent Insights into Inhibition, Structure, and Mechanism of Configuration-Retaining Glycosidases,” *Angew. Chemie Int. Ed.*, vol. 38, no. 6, pp. 750–770, 1999.
- [56] V. V. Krasikov, D. V. Karelov, and L. M. Firsov, “Alpha-Glucosidases,” *Biochem.*, vol. 66, no. 3, pp. 267–281, 2001.
- [57] V. H. Lillelund, H. H. Jensen, X. Liang, and M. Bols, “Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin,” *Chem. Rev.*, vol. 102, no. 2, pp. 515–553, 2002.
- [58] P. Notes, “Organisch-Chemisches Institut der Universitdt Bonn , Bonn (Germany),” vol. 151, pp. 728–729, 1968.
- [59] S. He, S. G. Withers, S. He, and S. G. Withers, “Protein chemistry and structure : Assignment of Sweet Almond β -Glucosidase as a Family 1 Glycosidase and Identification of Its Active Site Nucleophile,” vol. 272, no. 40, pp. 24864–24867, 1997.
- [60] D. E. Eveleigh and A. S. Perlin, “A proton magnetic resonance study of the anomeric species produced by d-glucosidases,” *Carbohydr. Res.*, vol. 10, no. 1, pp. 87–95, 1969.
- [61] S. M. Snaith, “Characterization of Jack-Bean alpha-D-Mannosidase as a Zinc Metalloenzyme,” *Biochem J*, vol. 147, pp. 83–90, 1975.
- [62] I. Isolation and D. R. Primate, “on the Glycosidases in Jack Bean,” *Time*, vol. 242, no. 23, pp. 5474–5480, 1967.
- [63] S. Howard, C. Braun, J. McCarter, K. W. Moremen, Y. F. Liao, and S. G. Withers, “Human lysosomal and jack bean α -mannosidases are retaining glycosidases,” *Biochem. Biophys. Res. Commun.*, vol. 238, no. 3, pp. 896–898, 1997.
- [64] M. S. Kang and A. D. Elbein, “Mechanism of Inhibition of Jack Bean alpha-Mannosidase by

- Swainsonine.," *Plant Physiol.*, vol. 71, no. 3, pp. 551–554, 1983.
- [65] A. D. Elbein, Y. T. Pan, R. Solf, and K. Vosbeck, "Effect of swainsonine, an inhibitor of glycoprotein processing, on cultured mammalian cells," *J. Cell. Physiol.*, vol. 115, no. 3, pp. 265–275, 1983.
- [66] S. Ogawa and Y. Yuming, "Synthesis of mannosidins A and B from myo-inositol," *Bioorganic Med. Chem.*, vol. 3, no. 7, pp. 939–943, 1995.
- [67] C. U. Kim *et al.*, "Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity," *J. Am. Chem. Soc.*, vol. 119, no. 4, pp. 681–690, 1997.
- [68] N. Katsilambros *et al.*, "A double-blind study on the efficacy and tolerance of a new alpha-glucosidase inhibitor in type-2 diabetics.," *Arzneimittelforschung.*, vol. 36, no. 7, pp. 1136–1138, 1986.
- [69] P. Compain and O. R. Martin, *Iminosugars: From Synthesis to Therapeutic Applications*. 2008.
- [70] D. L. Zechel and S. G. Withers, "Glycosidase mechanisms: Anatomy of a finely tuned catalyst," *Acc. Chem. Res.*, vol. 33, no. 1, pp. 11–18, 2000.
- [71] T. Kajimoto *et al.*, "Inhibitors against glycosidases as medicines.," *ChemBioChem*, vol. 16, no. 9, pp. 7337–7346, 2014.
- [72] S. Inouye, T. Tsuruoka, T. Ito, and T. Niida, "Structure and synthesis of nojirimycin.," *Tetrahedron*, vol. 24, no. 5, pp. 2125–2144, 1968.
- [73] M. H. Yagi M, Kouno T, Aoyagi Y, "The structure of moranoline, a piperidine alkaloid from *Morus* species," *Nippon Nogei Kagaku Kaishi.*, vol. 50, no. 11, pp. 571–572, 1976.
- [74] H. Nojima *et al.*, "Antihyperglycemic effects of N-containing sugars from *Xanthocercis zambesiaca*, *Morus bombycis*, *Aglaonema treubii*, and *Castanospermum australe* in streptozotocin-diabetic mice," *J. Nat. Prod.*, vol. 61, no. 3, pp. 397–400, 1998.
- [75] L. D. Hohenschutz *et al.*, "Castanospermine, A 1,6,7,8-tetrahydroxyoctahydroindolizine alkaloid, from seeds of *Castanospermum australe*," *Phytochemistry*, vol. 20, no. 4, pp. 811–814, 1981.
- [76] R. Saul, J. P. Chambers, R. J. Molyneux, and A. D. Elbein, "Castanospermine, a tetrahydroxylated alkaloid that inhibits β -glucosidase and β -glucocerebrosidase," *Arch. Biochem. Biophys.*, vol. 221, no. 2, pp. 593–597, 1983.
- [77] R. Saul, J. J. Ghidoni, R. J. Molyneux, and a D. Elbein, "Castanospermine inhibits alpha-glucosidase activities and alters glycogen distribution in animals.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 82, no. 1, pp. 93–97, 1985.
- [78] S. Colegate, P. Dorling, and C. Huxtable, "A Spectroscopic Investigation of Swainsonine: An α -Mannosidase Inhibitor Isolated from *Swainsona canescens*," *Aust. J. Chem.*, vol. 32, no. 10, p. 2257, 1979.
- [79] D. Tulsiani, H. Broquist, and O. Touster, "Marked Differences in the Swainsonine Inhibition of Rat Liver Lysosomal-alpha-D-Mannosidase, Rat Liver Golgi Mannosidase II, and Jack Bean alpha-D-Mannosidase," *Arch. Biochem. Biophys.*, vol. 236, no. 1, pp. 427–434, 1985.
- [80] D. R. P. Tulsiani, T. M. Harris, and O. Touster, "Swainsonine inhibits the biosynthesis of complex glycoproteins by inhibition of Golgi mannosidase II," *J. Biol. Chem.*, vol. 257, no. 14, pp. 7936–7939, 1982.
- [81] P. R. Dorling, C. R. Huxtable, and S. M. Colegate, "Inhibition of lysosomal alpha-mannosidase by swainsonine, an indolizidine alkaloid isolated from *Swainsona canescens*," *Biochem. J.*, vol. 191, no. 2, pp. 649–651, 1980.
- [82] A. D. Elbein, P. R. Dorling, K. Vosbeck, and M. Horisberger, "Swainsonine prevents the processing of the oligosaccharide chains of influenza virus hemagglutinin," *J. Biol. Chem.*, vol. 257, no. 4, pp. 1573–1576, 1982.
- [83] M. Hino *et al.*, "Studies of an immunomodulator, swainsonine. I. Enhancement of immune response by swainsonine in vitro," *J Antibiot*, vol. 38, no. 7, pp. 926–935, 1985.
- [84] J. Hamaguchi *et al.*, "Swainsonine reduces 5-fluorouracil tolerance in the multistage resistance of colorectal cancer cell lines," *Mol. Cancer*, vol. 6, 2007.
- [85] F. M. Santos *et al.*, "Increased antitumor efficacy by the combined administration of swainsonine and cisplatin in vivo," *Phytomedicine*, vol. 18, no. 12, pp. 1096–1101, 2011.

- [86] N. You *et al.*, “Swainsonine inhibits growth and potentiates the cytotoxic effect of paclitaxel in hepatocellular carcinoma in vitro and in vivo,” in *Oncology Reports*, 2012, vol. 28, no. 6, pp. 2091–2100.
- [87] D. Singh and G. Kaur, “The antileukaemic cell cycle regulatory activities of swainsonine purified from *Metarhizium anisopliae* fermentation broth,” *Nat. Prod. Res.*, vol. 28, no. 22, pp. 2044–2047, 2014.
- [88] M. J. Humphries, K. Matsumoto, S. L. White, and K. Olden, “Oligosaccharide modification by swainsonine treatment inhibits pulmonary colonization by B16-F10 murine melanoma cells.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 83, no. 6, pp. 1752–6, 1986.
- [89] J. Y. Sun, M. Z. Zhu, S. W. Wang, S. Miao, Y. H. Xie, and J. B. Wang, “Inhibition of the growth of human gastric carcinoma in vivo and in vitro by swainsonine,” *Phytomedicine*, vol. 14, no. 5, pp. 353–359, 2007.
- [90] V. Medicine and A. Northwest, “The Biosynthesis Pathway of Swainsonine, a New Anticancer Drug from Three Endophytic Fungi,” vol. 27, no. 11, pp. 1897–1906, 2017.
- [91] X. R. Liu W, Zhang XB, Li Y, Yang FZ, “Anti-tumor and immune enhancement of swainsonine.,” *NW Pharm. J.*, vol. 6, no. 17, pp. 258–260, 2006.
- [92] P. R. Dorling, C. R. Huxtable, and P. Vogel, “Lysosomal storage in swainsona spp. toxicosis: an induced mannosidosis,” *Neuropathol. Appl. Neurobiol.*, vol. 4, no. 4, pp. 285–295, 1978.
- [93] A. R. Leach and V. J. Gillet, *An introduction to chemoinformatics*. 2007.
- [94] P. D. Lyne, “Structure-based virtual screening: An overview,” *Drug Discov. Today*, vol. 7, no. 20, pp. 1047–1055, 2002.
- [95] J. Mestres, J. Mestres, R. M. a Knegtel, and R. M. a Knegtel, “docking in 3D virtual screening,” *Perspect. Drug Discov. Des.*, pp. 191–207, 2000.
- [96] J. S. Mason, A. C. Good, and E. J. Martin, “3-D pharmacophores in drug discovery.,” *Curr. Pharm. Des.*, vol. 7, no. 7, pp. 567–97, 2001.
- [97] J. Srinivasan *et al.*, “Evaluation of a novel shape-based computational filter for lead evolution: Application to thrombin inhibitors,” *J. Med. Chem.*, vol. 45, no. 12, pp. 2494–2500, 2002.
- [98] G. Klebe, *Wirkstoffdesign Entwurf und Wirkung von Arzneistoffen, 2. Auflage*. 2009.
- [99] P. E. B. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, “The Protein Data Bank,” *Nucleic Acids Res.*, vol. 28, no. 1, pp. 235–242, 2000.
- [100] T. J. A. Ewing, S. Makino, A. G. Skillman, and I. D. Kuntz, “DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases,” *J. Comput. Aided. Mol. Des.*, vol. 15, no. 5, pp. 411–428, 2001.
- [101] G. M. Morris *et al.*, “Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function,” *J. Comput. Chem.*, vol. 19, no. 14, pp. 1639–1662, 1998.
- [102] M. Rarey, B. Kramer, T. Lengauer, and G. Klebe, “A fast flexible docking method using an incremental construction algorithm,” *J. Mol. Biol.*, vol. 261, no. 3, pp. 470–489, 1996.
- [103] V. Schnecke and L. A. Kuhn, “Virtual screening with solvation and ligand-induced complementarity,” in *Perspectives in Drug Discovery and Design*, 2000, vol. 20, pp. 171–190.
- [104] E. C. Meng, B. K. Shoichet, and I. D. Kuntz, “Automated docking with grid-based energy evaluation,” *J. Comput. Chem.*, vol. 13, no. 4, pp. 505–524, 1992.
- [105] M. L. Connolly, “Analytical molecular surface calculation,” *J. Appl. Crystallogr.*, vol. 16, no. 5, pp. 548–558, 1983.
- [106] M. L. Connolly, “Solvent-Accessible Surfaces of Proteins and Nucleic Acids,” *Science (80-)*, vol. 221, no. 4612, p. 709, 1983.
- [107] Susan cain, “Table of of contents,” *Games Virtual Worlds Serious Appl. (VS-Games), 2015 7th Int. Conf.*, vol. 33, no. 3, pp. i–iii, 1999.
- [108] B. K. Shoichet, D. L. Bodian, and I. D. Kuntz, “Molecular docking using shape descriptors,” *J. Comput. Chem.*, vol. 13, no. 3, pp. 380–397, 1992.
- [109] P. Daldrop *et al.*, “Novel ligands for a purine riboswitch discovered by RNA-ligand docking,” *Chem. Biol.*, vol. 18, no. 3, pp. 324–335, 2011.
- [110] H. J. Böhm, “The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure,” *J. Comput. Aided. Mol. Des.*, vol. 8, no. 3, pp. 243–256, 1994.

- [111] M. Stahl and M. Rarey, "Detailed analysis of scoring functions for virtual screening," *J. Med. Chem.*, vol. 44, no. 7, pp. 1035–1042, 2001.
- [112] I. Reulecke, G. Lange, J. Albrecht, R. Klein, and M. Rarey, "Towards an integrated description of hydrogen bonding and dehydration: Decreasing false positives in virtual screening with the HYDE scoring function," *ChemMedChem*, vol. 3, no. 6, pp. 885–897, 2008.
- [113] R. A. Copeland, *Evaluation of Enzyme Inhibitors in Drug Discovery A Guide for Medicinal Chemists and Pharmacologists*. Wiley-Interscience, 2005.
- [114] R. A. Copeland, *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*. Wiley-Interscience, 2000.
- [115] C. I. Drugs, *Assessment of Enzyme Inhibition : A Review with Examples from the Development of Monoamine Oxidase and Cholinesterase Inhibitory Drugs*. 2017.
- [116] C. Yung-Chi and W. H. Prusoff, "Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction," *Biochem. Pharmacol.*, vol. 22, no. 23, pp. 3099–3108, 1973.
- [117] J.-H. Zhang, "A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays," *J. Biomol. Screen.*, vol. 4, no. 2, pp. 67–73, Apr. 1999.
- [118] M. B. Tierno *et al.*, "Development and optimization of high-throughput in vitro protein phosphatase screening assays," *Nat. Protoc.*, vol. 2, no. 5, pp. 1134–1144, 2007.
- [119] Y.-Y. Zhao *et al.*, "Functional roles of N-glycans in cell signaling and cell adhesion in cancer," *Cancer Sci.*, vol. 99, no. 7, pp. 1304–1310, 2008.
- [120] Y. J. Kim and A. Varki, "Perspectives on the significance of altered glycosylation of glycoproteins in cancer," *Glycoconj. J.*, vol. 14, no. 5, pp. 569–576, 1997.
- [121] C. A. Reis, H. Osorio, L. Silva, C. Gomes, and L. David, "Alterations in glycosylation as biomarkers for cancer detection," *J. Clin. Pathol.*, vol. 63, no. 4, pp. 322–329, 2010.
- [122] A. A. Watson, G. W. J. Fleet, N. Asano, R. J. Molyneux, and R. J. Nash, "Polyhydroxylated alkaloids - Natural occurrence and therapeutic applications," *Phytochemistry*, vol. 56, no. 3, pp. 265–295, 2001.
- [123] P. E. Goss, M. A. Baker, J. P. Carver, and J. W. Dennis, "Inhibitors of Carbohydrate Processing: A New Class of Anticancer Agents," *Clin. Cancer Res.*, vol. 1, no. 9, pp. 935–944, 1995.
- [124] M. Artola *et al.*, "1,6-Cyclophellitol Cyclosulfates: A New Class of Irreversible Glycosidase Inhibitor," *ACS Cent. Sci.*, vol. 3, no. 7, pp. 784–793, 2017.
- [125] S. Atsumi *et al.*, "Production, isolation and structure determination of a novel beta-glucosidase inhibitor, cyclophellitol, from *Phellinus* sp.," *J. Antibiot. (Tokyo)*, vol. 43, no. 1, pp. 49–53, 1990.
- [126] J. D. McCarter and S. G. Withers, "5-Fluoro glycosides: A new class of mechanism-based inhibitors of both α - and β -glucosidases," *J. Am. Chem. Soc.*, vol. 118, no. 1, pp. 241–242, 1996.
- [127] S. G. Withers, I. P. Street, P. Bird, and D. H. Dolphin, "2-Deoxy-2-fluoroglucosides: A Novel Class of Mechanism-Based Glucosidase Inhibitors," *J. Am. Chem. Soc.*, vol. 109, no. 24, pp. 7530–7531, 1987.
- [128] S. P. Schröder *et al.*, "Towards broad spectrum activity-based glycosidase probes: synthesis and evaluation of deoxygenated cyclophellitol aziridines," *Chem. Commun.*, 2017.
- [129] T. Aoyagi *et al.*, "Mannostatins A and B: new inhibitors of alpha-D-mannosidase, produced by *Streptoverticillium verticillus* var. *quintum* ME3-AG3: taxonomy, production, isolation, physico-chemical properties and biological activities," *J. Antibiot. (Tokyo)*, vol. 42, no. 6, pp. 883–889, 1989.
- [130] S. P. Kawatkar, D. A. Kuntz, R. J. Woods, D. R. Rose, and G. J. Boons, "Structural basis of the inhibition of Golgi α -mannosidase II by mannostatin A and the role of the thiomethyl moiety in ligand-protein interactions," *J. Am. Chem. Soc.*, vol. 128, no. 25, pp. 8310–8319, 2006.
- [131] A. Nayeem, D. Sitkoff, and S. Krystek, "A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models.," *Protein Sci.*, vol. 15, no. 4, pp. 808–824, 2006.
- [132] "file:///C:/moe2012/html/tutorials/hom_tut.htm." .
- [133] K. Kelly, "3D bioinformatics and comparative protein modeling in MOE," *J. Chem. Comp.*

- Gr.*, vol. autumn edi, 1999.
- [134] R. D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, G. S. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, J. L. J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, G. X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, T. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, and P. A. K. Luchko, S. Gusarov, A. Kovalenko, “AMBER 12.” San Francisco- University of California, 2012.
- [135] P. R. Gerber and K. Müller, “MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry,” *J. Comput. Aided. Mol. Des.*, vol. 9, no. 3, pp. 251–268, 1995.
- [136] C. S. Reddy, K. Vijayasarathy, E. Srinivas, G. M. Sastry, and G. N. Sastry, “Homology modeling of membrane proteins: A critical assessment,” *Comput. Biol. Chem.*, vol. 30, no. 2, pp. 120–126, 2006.
- [137] P. H. Bindu, G. M. Sastry, U. S. Murty, and G. N. Sastry, “Structural and conformational changes concomitant with the E1-E2 transition in H+K+-ATPase: A comparative protein modeling study,” *Biochem. Biophys. Res. Commun.*, vol. 319, no. 2, pp. 312–320, 2004.
- [138] Schrödinger LLC, “The {PyMOL} Molecular Graphics System, Version 1.5.0.5.” 2013.
- [139] B. GmbH, “LeadIT version 2.1.6.” Sankt Augustin, Germany, 2014.
- [140] D. Weininger, “Smiles, a Chemical Language and Information-System .1. Introduction to Methodology and Encoding Rules,” *J. Chem. Inf. Comput. Sci.*, vol. 28, no. 1, pp. 31–36, 1988.
- [141] P. Labute and M. Santavy, “Locating Binding Sites in Protein Structures.”
- [142] S. Soga, H. Shirai, M. Kobori, and N. Hirayama, “Use of Amino Acid Composition to Predict Ligand-Binding Sites,” *J. Chem. Inf. Model.*, vol. 47, no. 2, pp. 400–406, Mar. 2007.
- [143] A. Volkamer, D. Kuhn, T. Grombacher, F. Rippmann, and M. Rarey, “Combining Global and Local Measures for Structure-Based Druggability Predictions,” *J. Chem. Inf. Model.*, vol. 52, no. 2, pp. 360–372, Feb. 2012.
- [144] A. Volkamer, D. Kuhn, F. Rippmann, and M. Rarey, “DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment,” *Bioinformatics*, vol. 28, no. 15, pp. 2074–2075, Aug. 2012.
- [145] L. Cieplak, J. Caldwell, and P. Kollman, “Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: Aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/,” *J Comp Chem*, vol. 22, no. 10, pp. 1048–1057, 2001.
- [146] W. D. Cornell, P. Cieplak, C. I. Bayly, and P. Kollman, “Application of Resp Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation,” *J. Am. Chem. Soc.*, vol. 115, no. 21, pp. 9620–9631, 1993.
- [147] R. Brenk, S. W. Vetter, S. E. Boyce, D. B. Goodin, and B. K. Shoichet, “Probing Molecular Docking in a Charged Model Binding Site,” *J. Mol. Biol.*, vol. 357, no. 5, pp. 1449–1470, Apr. 2006.
- [148] M. M. Mysinger and B. K. Shoichet, “Rapid Context-Dependent Ligand Desolvation in Molecular Docking,” *J. Chem. Inf. Model.*, vol. 50, no. 9, pp. 1561–1573, Sep. 2010.
- [149] R. Brenk *et al.*, “Lessons learnt from assembling screening libraries for drug discovery for neglected diseases,” *ChemMedChem*, vol. 3, no. 3, pp. 435–444, Mar. 2008.
- [150] P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, and M. T. Stahl, “Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database,” *J. Chem. Inf. Model.*, vol. 50, no. 4, pp. 572–584, Apr. 2010.
- [151] B. Q. Wei, W. A. Baase, L. H. Weaver, B. W. Matthews, and B. K. Shoichet, “A model binding site for testing scoring functions in molecular docking,” *J Mol Bio*, vol. 322, no. 2, pp. 339–55, 2002.
- [152] D. M. Lorber and B. K. Shoichet, “Flexible ligand docking using conformational ensembles,” *Protein Sci.*, vol. 7, no. 4, pp. 938–950, Dec. 2008.
- [153] M. M. Mysinger and B. K. Shoichet, “Rapid Context-Dependent Ligand Desolvation in Molecular Docking,” *J. Chem. Inf. Model.*, vol. 50, no. 9, pp. 1561–1573, Sep. 2010.
- [154] A. Heifets and R. H. Lilien, “LigAlign: Flexible ligand-based active site alignment and

- analysis," *J. Mol. Graph. Model.*, vol. 29, no. 1, pp. 93–101, 2010.
- [155] M. Juchum, "Synthese und Testung bizyklischer Acetale als Potentielle Glykosidase Inhibitoren," Johannes Gutenberg Universität Mainz, 2012.
- [156] D. J. Coleman *et al.*, "A long-wavelength fluorescent substrate for continuous fluorometric determination of α -mannosidase activity: Resorufin α -D-mannopyranoside," *Anal. Biochem.*, vol. 399, no. 1, pp. 7–12, 2010.
- [157] L. Misuri *et al.*, "The use of dimethylsulfoxide as a solvent in enzyme inhibition studies: the case of aldose reductase," *J. Enzyme Inhib. Med. Chem.*, vol. 32, no. 1, pp. 1152–1158, 2017.
- [158] P. Heikinheimo *et al.*, "The structure of bovine lysosomal alpha-mannosidase suggests a novel mechanism for low-pH activation," *J. Mol. Biol.*, vol. 327, no. 3, pp. 631–644, 2003.
- [159] S. Caramia, A. G. M. Gatius, F. dal Piaz, D. Gaja, and A. Hochkoeppler, "Dual role of imidazole as activator/inhibitor of sweet almond (*Prunus dulcis*) β -glucosidase," *Biochem. Biophys. Reports*, vol. 10, no. November 2016, pp. 137–144, 2017.
- [160] T. Barrett, C. G. Suresh, S. P. Tolley, E. J. Dodson, and M. a Hughes, "The crystal structure of a cyanogenic beta-glucosidase from white clover, a family 1 glycosyl hydrolase.," *Structure*, vol. 3, pp. 951–960, 1995.
- [161] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings," *Adv. Drug Deliv. Rev.*, vol. 23, pp. 3–25, 1997.
- [162] J. M. van den Elsen, D. A. Kuntz, and D. R. Rose, "Structure of Golgi alpha-mannosidase II: a target for inhibition of growth and metastasis of cancer cells.," *EMBO J.*, vol. 20, no. 12, pp. 3008–3017, 2001.
- [163] Paul Labute and Martin Santavy, "Locating Binding Sites in Protein Structures," *J. Chem. Comput. Gr.*, 2007.
- [164] S. Soga, H. Shirai, M. Kobori, and N. Hirayama, "Use of Amino Acid Composition to Predict Ligand-Binding Sites," *J. Chem. Inf. Model.*, vol. 47, no. 2, pp. 400–406, Mar. 2007.
- [165] M. M. Yamashita, L. Wesson, G. Eisenman, and D. Eisenberg, "Where metal ions bind in proteins.," *Proc. Natl. Acad. Sci.*, vol. 87, no. 15, pp. 5648–5652, 1990.
- [166] K. E. Hevener *et al.*, "Validation of molecular docking programs for virtual screening against dihydropteroate synthase," *J. Chem. Inf. Model.*, vol. 49, no. 2, pp. 444–460, 2009.
- [167] J. C. Cole, C. W. Murray, J. W. M. Nissink, R. D. Taylor, and R. Taylor, "Comparing protein-ligand docking programs is difficult," *Proteins Struct. Funct. Genet.*, vol. 60, no. 3, pp. 325–332, 2005.
- [168] M. dos S. Guilherme, "Synthese und Testung neuer potentieller Golgi-alpha-Mannosidase-II-Inhibitoren," 2016.
- [169] C. Rabouille *et al.*, "The Drosophila GMII gene encodes a Golgi alpha-mannosidase II," *J Cell Sci*, vol. 112, pp. 3319–30, 1999.
- [170] H. A. Strachan, "Expression, purification and characterization of Drosophila Melanogaster orthologous to Drosophila Melanogaster Golgi Mannosidase II," University of Georgia, 2009.
- [171] H. B. Brooks *et al.*, "Basics of Enzymatic Assays for HTS," *Assay Guid. Man.*, pp. 3–4, 2012.
- [172] D. A. Kuntz and D. R. Rose, "Crystal structure of Golgi Mannosidase II in complex with swainsonine at 1.3 Angstrom resolution," 2008. .
- [173] D. A. Kuntz, C. A. Tarling, S. G. Withers, and D. R. Rose, "Structural analysis of Golgi α -mannosidase II inhibitors identified from a focused glycosidase inhibitor screen," *Biochemistry*, vol. 47, no. 38, pp. 10058–10068, 2008.
- [174] H. Bisswanger, "Enzyme assays," *Perspect. Sci.*, vol. 1, no. 1–6, pp. 41–55, 2014.
- [175] M. P. Dale, H. E. Ensley, K. Kern, K. A. R. Sastry, and L. D. Byers, "Reversible Inhibitors of β -Glucosidase," *Biochemistry*, vol. 24, no. 14, pp. 3530–3539, 1985.
- [176] N. Shah, D. A. Kuntz, and D. R. Rose, "Comparison of Kifunensine and 1-Deoxymannojirimycin Binding to Class I and II α -Mannosidases Demonstrates Different Saccharide Distortions in Inverting and Retaining Catalytic Mechanisms," *Biochemistry*, vol. 42, no. 47, pp. 13812–13816, 2003.
- [177] L. M. Kavlekar *et al.*, "5-Thio-D-glycopyranosylamines and their amidinium salts as potential transition-state mimics of glycosyl hydrolases: Synthesis, enzyme inhibitory activities, X-ray crystallography, and molecular modeling," *Tetrahedron Asymmetry*, vol. 16, no. 5, pp. 1035–

- 1046, 2005.
- [178] D. A. Kuntz, A. Ghavami, B. D. Johnston, B. M. Pinto, and D. R. Rose, "Crystallographic analysis of the interactions of *Drosophila melanogaster* Golgi α -mannosidase II with the naturally occurring glycomimetic salacinol and its analogues," *Tetrahedron Asymmetry*, vol. 16, no. 1, pp. 25–32, 2005.
- [179] P. Englebienne *et al.*, "Evaluation of docking programs for predicting binding of Golgi α -mannosidase II inhibitors: A comparison with crystallography," *Proteins Struct. Funct. Genet.*, vol. 69, no. 1, pp. 160–176, 2007.
- [180] P. J. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting," *Adv. Drug Deliv. Rev.*, 1997.
- [181] D. Baker, D. Baker, and A. Sali, "Protein Structure Prediction and Structural Genomics," vol. 93, no. 2001, pp. 2–5, 2008.
- [182] A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees, and C. H. Reynolds, "The role of ligand efficiency metrics in drug discovery," *Nat. Rev. Drug Discov.*, vol. 13, no. 2, pp. 105–121, 2014.