RESEARCH ARTICLE

Identification of a potential allosteric site of Golgi α-mannosidase II using computer-aided drug design

Lina Irsheid[1☉], **Thomas Wehler**[1,2☉], **Christoph Borek**[1], **Werner Kiefer**[1], **Ruth Brenk**[2], **Maria Elena Ortiz-Soto**[3], **Jürgen Seibel**[3], **Tanja Schirmeister**[1]*

1 Institute of Pharmacy and Biochemistry, University of Mainz, Mainz, Germany, **2** Department of Biomedicine, University of Bergen, Bergen, Norway, **3** Institute of Organic Chemistry, University of Würzburg, Würzburg, Germany

☉ These authors contributed equally to this work.
* schirmei@uni-mainz.de

Abstract

Golgi α-mannosidase II (GMII) is a glycoside hydrolase playing a crucial role in the *N*-glycosylation pathway. In various tumour cell lines, the distribution of *N*-linked sugars on the cell surface is modified and correlates with the progression of tumour metastasis. GMII therefore is a possible molecular target for anticancer agents. Here, we describe the identification of a non-competitive GMII inhibitor using computer-aided drug design methods including identification of a possible allosteric binding site, pharmacophore search and virtual screening.

Introduction

Selectins are carbohydrate-recognizing proteins and are essential for adhesion of metastasizing cancer cells [1]. Oncogenic cells often show an altered distribution of oligosaccharides on their cell surface. This modified distribution of carbohydrates is linked to disease progression, metastasis and poor clinical outcome. Targeting the glycosylation pathways is a possible approach to identify new anticancer agents [2]. Golgi α-mannosidase II (GMII) is one of these possible targets, since it is a key enzyme in *N*-glycan processing. Its inhibition has been linked to an induced tumour repression [3]. The alkaloid swainsonine, a well-known inhibitor of GMII, has exhibited anti-tumour effects in certain colon, breast, or skin cancer types [2]. However, due to the simultaneous inhibition of the structurally related lysosomal mannosidase, side effects restrict the clinical use of this potent inhibitor [4–6]. GMII belongs to the hydrolase family 38 [7] and catalyzes the cleavage of α-(1,6) and α-(1,3) glycosidic bonds between mannose residues in $GlcNAcMan_5GlcNAc_2$ (GlcNAc: *N*-acetyl-glucosamine; Man: mannose) resulting in the formation of $GlcNAcMan_3GlcNAc_2$ [5].

The catalytic mechanism of this enzyme is well understood and was subject of numerous studies [8–10]. The catalytic site of GMII essentially consists of two aspartate residues and a zinc cation. GMII acts as a retaining glycosidase and cleaves the glycosidic bonds in a two-step mechanism via a covalent glycosyl-enzyme transition state complex, thereby preserving the configuration of the anomeric C-atom [5,10–12].

Over the past two decades, only few novel ligands of glycosidases addressing the catalytic site have been identified. Known irreversible inhibitors are cyclophellitol (1991) [13] or 5-fluoro- (1996) [14] and 2-deoxy-2-fluoro-glycosides (1987) [15]. In a recent study from Overkleeft *et al.* [16], a new class of irreversible glycosidase inhibitors, cyclophellitol cyclosulfates, was introduced. However, since the discovery of the known reversible inhibitors of GMII, swainsonine (1982) [17] and mannostatin (1989) [18], no significant advances have been made with regard to this inhibitor class. This may be due to the unfavourable properties of this binding site, namely its small size and a high degree of negatively charged amino acids.

In this study, we aimed to identify new inhibitors by addressing a potential allosteric site. The GMII enzyme of *Drosophila melanogaster* (dGMII) with a high sequence identity to its human homologue (hGMII), was used as model enzyme [19]. Both dGMII and hGMII are 100% structurally identical in the radius of 6 Å around the central Zn^{2+}-ion and share 41% sequence identity and 59% sequence similarity with regard to the entire protein according to our homology modeling studies. Furthermore, GMII has a high degree of sequence conservation among many eukaryotes [20]. Based on the structure, a potential binding site was identified and ligands were discovered using a combination of a protein-based pharmacophore model and molecular docking. A number of compounds were short-listed for enzymatic testing which resulted in one hit.

Materials and methods

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) [21], 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 [22] were reported along with a list of the residues containing the 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.

Comparison of binding site properties–active site versus 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 0 and 1, whereupon pockets with a score approaching 1 are assumed to be the more druggable ones [23,24].

Receptor preparation and sphere set generation for virtual screening

The crystal structure of the *Drosophila melanogaster* GMII (dGMII) in complex with the inhibitor swainsonine (PDB code: 1HWW [3]) was chosen as the receptor for this study. Subsequently, polar hydrogen atoms were added to the protein and their positions were minimized using the AMBER force field [25] 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 [26] were assigned to the protein atoms. Due to the lack of a reference ligand, the binding site was manually filled with dummy atoms. 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. Grids for storing the binding site information with respect to electrostatic and van der Waals potential as well as the solvent-excluded volume were calculated as described earlier [27,28].

Small molecule preparation

For the compilation of a ligand database suitable for docking and the pharmacophore search a virtual in-house library of 5,574,182 commercially available small molecules was used. The molecules were stored in SMILES format [29] alongside with their calculated physicochemical properties (e.g. molecular weight, number of H-bond donors/acceptors, logP) [30]. The molecules were filtered and converted in a dockable format.

First, the library compounds were filtered for desired lead-like properties (Table 1).

Additionally, all compounds containing unwanted groups were removed. For the remaining molecules, likely 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

Table 1. 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[a]

[a] 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 [30].

https://doi.org/10.1371/journal.pone.0216132.t001

with OMEGA [31] (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/) [32]. Finally, the multiple conformers of each molecule were aligned to their ring systems and stored in a hierarchical (flexibase) format [33].

Pharmacophore search

The pharmacophore model was developed based on 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. At the total, a set of 6 pharmacophore features were created. The hydrophobic feature is given a tolerance radius of 2.1 Å. In the region of E459, two features were represented: one expressed a hydrogen-bond donor or a cationic functionality within a radius of 2.4 Å and the counterpart of this feature, a hydrogen-bond acceptor, had a radius of 1.7 Å. The remaining three features all had radii of 1.0 Å. 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). Only those molecules with at least one low-energy conformation meeting the criteria of the pharmacophore search were kept for docking.

Molecular docking

DOCK 3.6 was used to place the ligands into the potential allosteric binding site [28,32,33]. 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 matching sites were 1.2 Å. Each docking pose was scored for electrostatic and van der Waals complementarity and corrected for partial desolvation [33]. Only the best-scoring representation (correct protomer, tautomer and stereoisomer) for each docked molecule was stored in the output database.

Docking analysis

The docking poses were stored in a MOE database alongside with their corresponding scores in kJ/mol. Subsequently, predicted binding modes were visualized for closer inspection to manually select the most promising candidate compounds. 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 Eq (1):

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

where "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) [34]. 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 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.

Cloning and recombinant expression of the Golgi α-mannosidase II from *Drosophila melanogaster*

A *Drosophila* expression system (DES-Inducible/Secreted Kit) with a pCoHygro selection vector was purchased from Invitrogen (ThermoFisher, USA). $CdCl_2$ and HIS-Select HF Nickel Affinity Gel both were supplied by Sigma-Aldrich (Germany). Affi-Gel Blue Gel was purchased from Bio-Rad (Germany) and the *Drosophila* media and buffers were purchased from Biowest and Sigma-Aldrich, respectively. A QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent Genomics (Germany) was used for the introduction of base substitutions.

The Golgi α-mannosidase II gene was amplified by PCR from its corresponding cDNA of the *D. melanogaster* homologue while adding *NdeI/BglII* and *XhoI/EcoRI* restrictions sites at the 5′and 3′ends, respectively. Apart from the restriction sites introduced, the amplified gene already contained three internal *BglII* and *XhoI* recognition sequences. Thus, the gene was cloned into the vector pMAL-c5X flanked by *NdeI* and *EcoRI* restriction sites and the internal *BglII* and *XhoI* sequences were silenced by synonymous base substitution using the Quik-Change Lightning Site-Directed Mutagenesis Kit. The gene with silenced internal restriction sites was cloned into the expression vector pMT/Bip/V5-HisA between *BglII* and *XhoI* restriction sites based on the DES expression system from Invitrogen (Inducible/secreted kit). Stable cell lines were selected via co-transfection of 3 mL *Drosophila* Schneider 2 (S2) cells (1.5×10^6 cells/mL) with pMT/Bip/V5-HisA and pCoHygro as selection vectors in a 6 well plate. Cells were selected with 300–600 μg/mL hygromycin (final concentration) over 3 weeks at 28°C. Afterwards, cells were assayed for expression in 12 well plates employing different concentrations of $CdCl_2$ (0–15 μM) over 24 and 48 h. The recombinant protein was detected in the growth medium by Western blot. 300 μg/mL Hygromycin was used in further experiments with suspension cultures and enzyme induction was performed with 15 μM $CdCl_2$ over 48 h.

The Golgi α-mannosidase II was purified from the growth medium by immobilized metal affinity chromatography (IMAC) using the HIS-Select HF Nickel Affinity Gel after removal of fetal bovine serum (FBS) by incubation with the Affi-Gel Blue resin. The purified protein was aliquoted and stored at -80°C in 10 mM Tris-HCl (pH 8.3) supplemented with 100 mM NaCl (S1 Fig).

Following SDS-PAGE, proteins were transferred to a PVDF Immobilon-P Membrane (Merck, Germany). The expression level of the HIS-tagged Golgi α-mannosidase II was analyzed through binding and detection of monoclonal anti-HisG alkaline phosphatase conjugated antibodies (Life technologies, Germany).

Enzyme assay

Inhibition assays with dGMII were performed at pH 5.75 in MES buffer (40 mM), supplemented with bovine serum albumin 1.1% (0.1% final concentration in MES buffer) and $ZnSO_4$-solution 1.1% (0.1% final concentration in MES buffer) [35]. The recombinantly expressed and purified enzyme was added to a final concentration of 0.002 mg/mL per well. 4-Methylumbellifer-7-yl-α-D-mannopyranoside (4-MU-Man) [36] was synthesized [37] (S1 File and S2 Fig) and used as substrate with a final concentration of 1 mM per well. The total reaction volume was 50 μL and the mixture was incubated at 37°C for five minutes. The test compounds were dissolved in DMSO so that the DMSO concentration in the wells did not exceed 5% (v/v). The release of 4-methylumbelliferone was measured for 15 min by fluorescence spectroscopy using Greiner 96 Flat Bottom Black Polystyrol plates and a Tecan Reader

Infinite F200 PRO equipped with a filter system (355 nm excitation and 485 nm emission). Assays in order to investigate the selectivity of compounds between different mannosidases were carried out on JBM. Assays concerning the selectivity between the desired mannosidase inhibition and a potential glucosidase inhibition were carried out on the β-glucosidase of sweet almonds using 4-methylumbelliferyl-β-D-glucopyranoside (4-MU-Glc) as substrate. The substrate 4-MU-Glc was synthesized [38] as described in the supplementary information (S1 File and S3 Fig). JBM and β-glucosidase from sweet almonds were purchased from Sigma Aldrich. The enzyme β-glucosidase from sweet almonds was diluted in citrate buffer in order to obtain a concentration of 0.04 mg/mL. The inhibition assays were carried out at a pH of 5.0 in citrate buffer (20 mM) [39]. The substrate concentration was chosen to be 1 mM per well in accordance with the results of several K_M value determinations since this concentration ensures good enzyme activity (K_M = 1.9 ± 0.1 mM). Inhibition assays JBM were performed in the same buffer as for dGMII. The final concentration of the enzyme per well was 0.003 mg/ml. The substrate 4-MU-Man was used with a final concentration of 0.4 mM per well. The conducted assays with JBM and the β-glucosidase from sweet almonds were performed as described for dGMII.

The reproducibility of the assay was assessed by determining the Z'-value as indicator for overall assay quality [40]. This was performed by measuring the residual reaction rate during 15 minutes with swainsonine (50 nM per well) and without swainsonine. The Z'-value was calculated to be 0.68 and therefore indicates good reproducibility since values between 0.5 and 1 are acceptable [40].

The most promising compounds according to the molecular modeling studies were tested at 300 μM final concentration except for compound **24** which was tested at 100 μM per well due to the small amount available. We considered the compounds as inactive when less than 40% of the enzyme activity was inhibited. The inhibition was calculated with the following Eq (2):

$$\text{Inhibition (\%)} = \left(1 - \frac{\text{activity (with inhibitor)}}{\text{activity (without inhibitor)}}\right) \cdot 100 \tag{2}$$

K_M values were calculated using the Michaelis-Menten kinetics Eq (3)

$$v = \frac{v_{max}[S]}{K_M + [S]} \tag{3}$$

whereby v_{max} is the maximal velocity, [S] the substrate concentration and K_M the substrate concentration yielding a velocity of $v_{max}/2$.

IC_{50} values were calculated using following Eq (4):

$$y = \frac{y_{max} - y_{min}}{1 + \left(\frac{[I]}{IC_{50}}\right)^s} + y_{min} \tag{4}$$

whereby y_{max} is defined as 100% enzyme activity, y_{min} as 0% enzyme activity, s is defined as the slope of the linear part of the curves (Hill coefficient), and [I] represents the inhibitor concentration. For the K_M value determination, substrate concentrations ranging from 0.4 to 8.0 mM for dGMII and from 0.008 to 1.6 mM for JBM were used. For the IC_{50} determination of swainsonine, concentrations ranging from 6 to 120 nM were applied. For the IC_{50} determination of compound **18,** the substrate concentrations ranged from 10 to 1000 μM. In order to define the mode of inhibition, the determination of the IC_{50} values was carried out using different substrate concentrations ranging from 0.1 to 3 mM, since in case of non-competitive inhibition the IC_{50} value should be independent from the substrate concentration [41]. The

program GraFit (5.0.13, Erithacus Software, Horley, Surrey, UK (2006)) was used to calculate the IC_{50} values.

Results

The potential allosteric Golgi α-mannosidase II binding site in comparison to the known active site

The protein structure of dGMII was analyzed to detect pockets that could be targeted by allosteric inhibitors. A total of 60 cavities were identified in the protein structure (PDB code: 1HWW [3]). After visual inspection of the most promising sites with regard to the size, the amino acid composition, the accessibility and the distance from the proposed site to the active site, one site has been chosen as potential allosteric site. This site consists of 102 receptor atoms, 29 are hydrophobic and 68 are side chain atoms. In contrast, the catalytic site is composed of 93 receptor atoms, 31 are hydrophobic and 75 are side chain atoms. Approximately 30 Å separate the catalytic site from the potential binding site when measured from the center of each site (Fig 1).

A comparison of the amino acid composition revealed some differences (Fig 2(A) and 2 (B)).

The known catalytic site is compact, hydrophilic and predominantly negatively charged in the region where potential ligands are deeply buried (Fig 2(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 (Fig 2(B)). The active site has a large negatively charged surface due to the acidic residues in its pocket (Fig 2(A)).
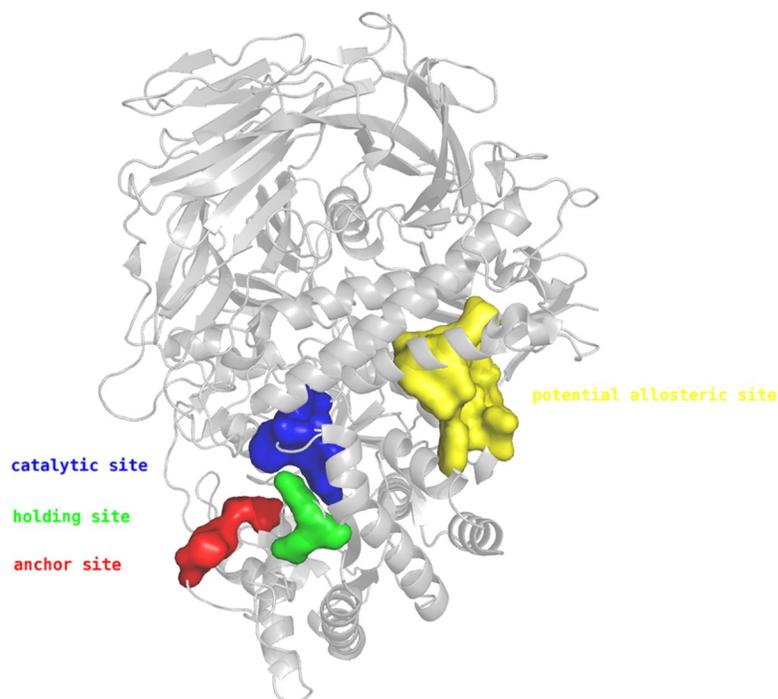

Fig 1. Overview of the different locations of the active site consisting of the catalytic site (blue), holding site (green), anchor site (red) with the potential allosteric site (yellow) within dGMII [3].
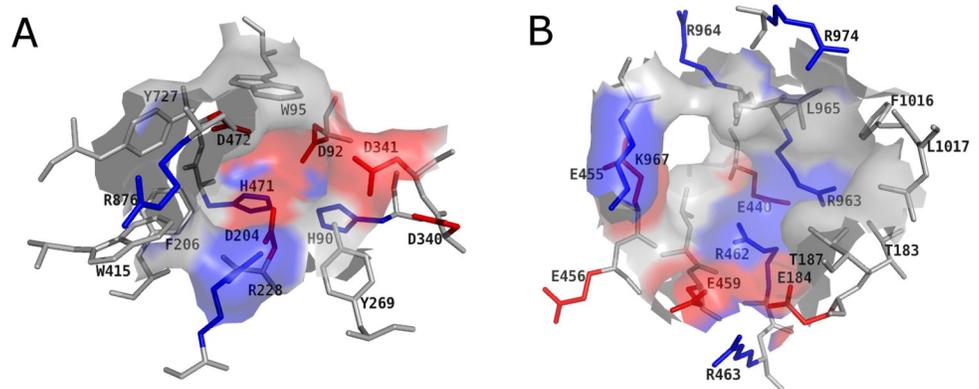
https://doi.org/10.1371/journal.pone.0216132.g001

Fig 2. Colour-coded surface generated for the active site (A) and the identified pocket (B). Acidic and basic residues are coloured red and blue, respectively.

Amino acids such as D204, D340, D92 or D472 may accordingly interact with positively charged binding partners. The GMII is a metallo enzyme 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 [42]. 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, R974 or K967 (Fig 2(B)). Amino acid residues such as E459, E184 or E440 are responsible for the negatively charged surface. In all other regions of the potential allosteric site neutral amino acids are predominant.

The druggability of the potential allosteric binding site identified was evaluated using the DoGSite Scorer considering the following properties: pocket volume, surface area, depth, druggability score and the apolar amino acid ratio (Table 2).

With a druggability score of 0.83, the DoGSite Scorer predicted a high druggability for the potential allosteric site, since scores > 0.5 indicate a druggable binding site [23]. Furthermore, the properties of both sites differed significantly (Fig 2(A) and 2(B)). The larger pocket volume, higher depth, and a higher apolar amino acid ratio of the potential allosteric site are positive predictors of its druggability [23]. Thus, the potential allosteric site could potentially bind drug-like compounds.

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 (Fig 3).

Two pharmacophore features were included to achieve interactions with E459. On the one hand, a hydrogen-bond donor or a cationic functionality was required close to the glutamate.

Table 2. Shape description and druggability prediction of the active and the potential allosteric site of GMII using DoGSiteScorer (PDB code: 1HWW [3]).

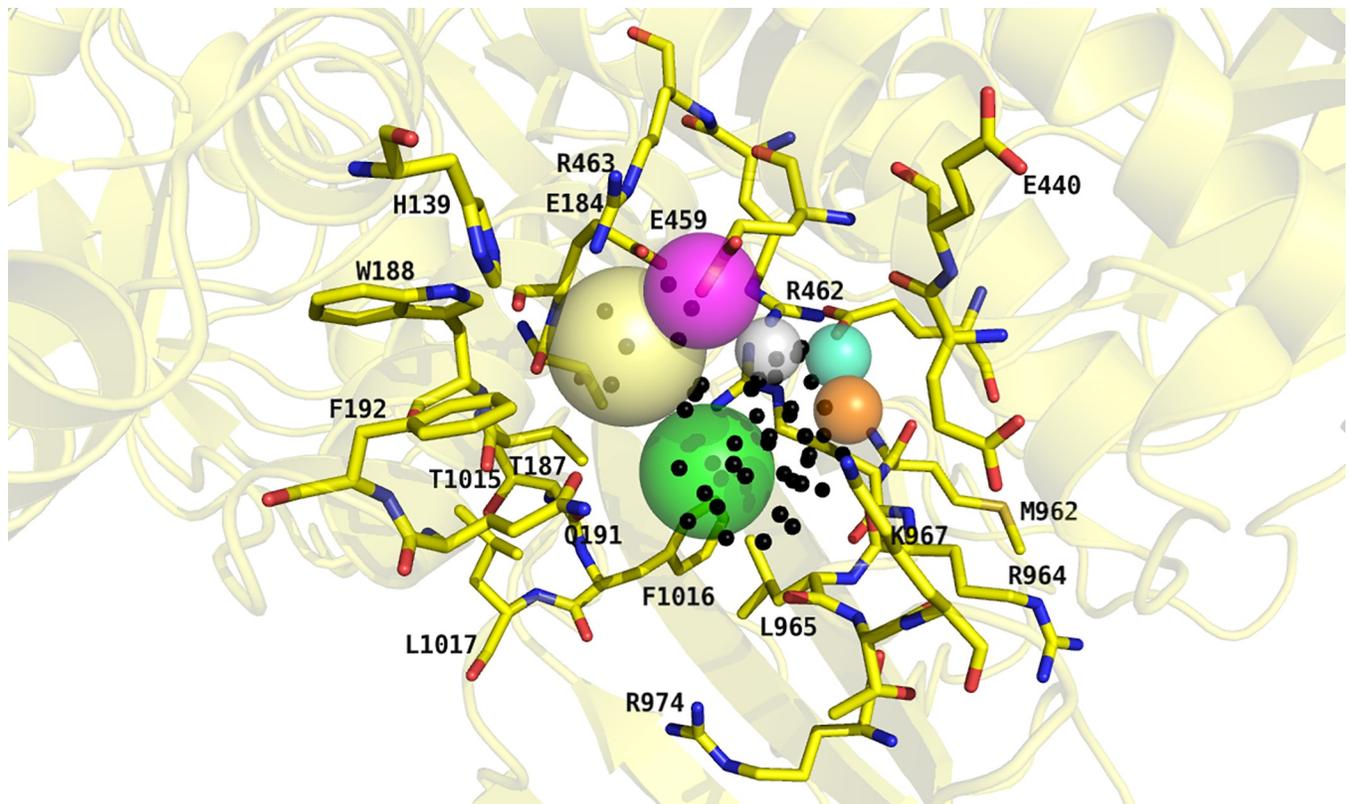|	Volume [Å³]	Surface [Å²]	Depth [Å]	Drugg-ability score	Apolar amino acid ratio
Active site	194.65	229.38	8.97	0.28	0.21
Potential allosteric site	879.26	820.14	22.46	0.83	0.46

Fig 3. 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 yellow 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.

On the other hand, the counterpart of this interaction, the carboxylic acid function of E459, was included as a hydrogen-bond acceptor. Further, a hydrophobic moiety was required close to the side chains of L965 and F1016. Another pharmacophore feature was located between E459 and R462 to either act as hydrogen-bond donor to E459 or as acceptor for 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. 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 for the interaction network to E459 and/or R462, respectively, were determined as being mandatory. Ligand conformations could only pass the pharmacophore search if six out of seven features (with excluded volume) were matched.

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.
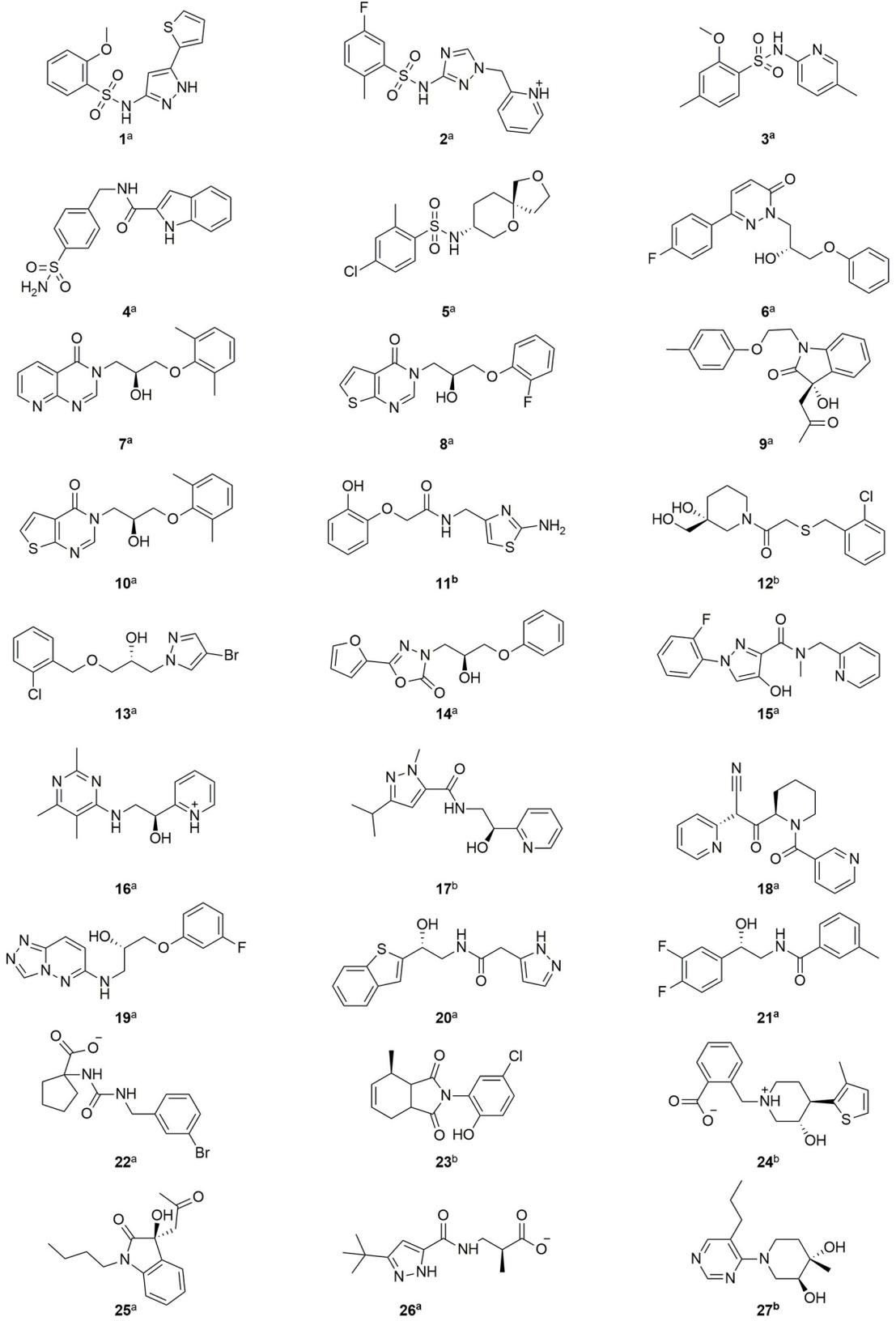
Fig 4. Molecular structures of virtual screening hits with corresponding compound number. [a] Compounds have been purchased from Enamine. [b] Compounds have been purchased from ChemBridge. The molecules are represented in the protonated form in which they were docked.

Virtual screening and docking results

For 36,607 compounds, a docking pose was obtained. Compounds for subsequent biological evaluation were selected according to the following criteria:

- total docking score (ΔG) predicted to be $\leq$ -25 kJ/mol

- predicted ligand efficiency (LE) as defined by Eq (1) $\geq$ 1 kJ/mol per heavy atom

- visual inspection of the predicted pose including an accurate analysis of the postulated protein-ligand interactions: 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 in particularly with E459 und R462.

Based on the described criteria, compounds from different substance classes with scores from -69.11 to -25.66 kJ/mol and LEs reaching from 1.16 to 3.14 kJ$\times$mol$^{-1}\times$HAC^{-1} were finally purchased. In total, 27 out of the 41,815 compounds were selected for testing (Fig 4).

These 27 compounds were either purchased from ChemBridge or Enamine and tested using a fluorometric enzyme assay. For the compounds with stereoisomerism, it was only possible to obtain racemic mixtures, although stereoisomers had been taken into account for the molecular docking.
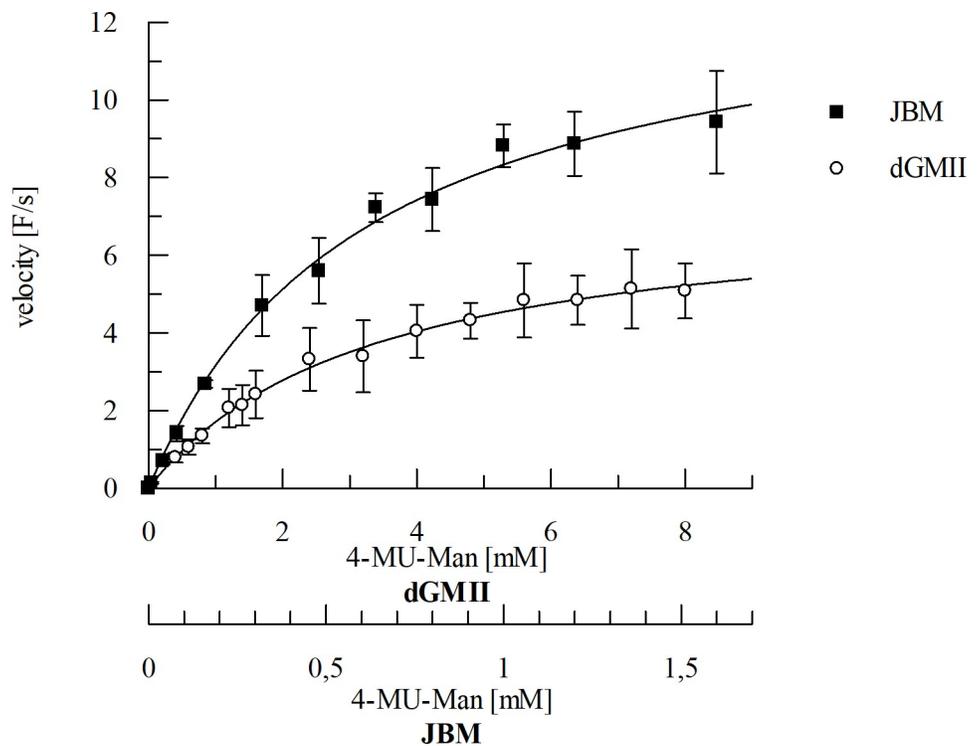

Fig 5. Example of the non-linear fit for the determination of the K_M values of the enzymes dGMII and JBM with 4-MU-Man as substrate. Non-linear regression was applied to the data using the Michaelis-Menten Eq (3).
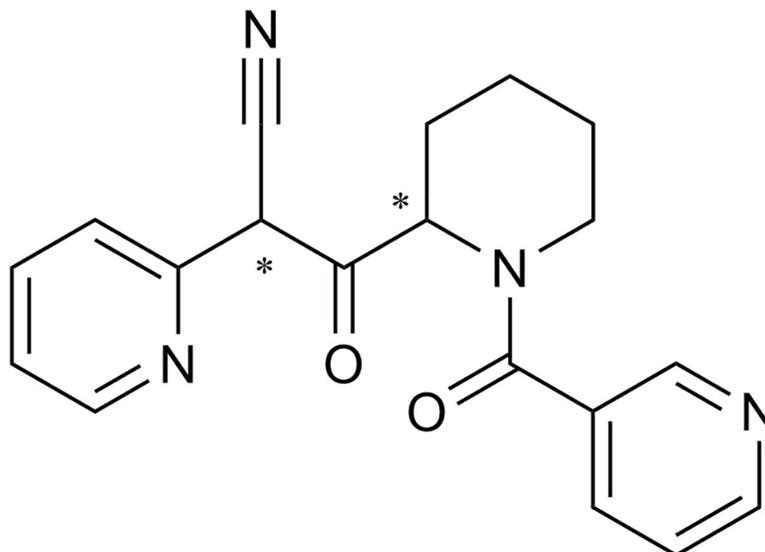
Fig 6. Chemical structure of compound 18.

Results of enzyme assays, determination of K_M and IC_{50} values, and the mode of inhibition

The determined K_M values were 3.6 ± 0.3 mM and 0.6 ± 0.06 mM for dGMII and JBM, respectively (Fig 5).

Two out of the 27 compounds exhibited inhibition > 40% at 300 μM, namely the bis-pyridyl derivative **18** (Fig 6), (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 the observed inhibition of GMII was expected to be caused by an experimental artefact. Compound **20** strongly interfered with the assay due to its auto-fluorescent properties. Independent of the concentration of compound **20**, large RFU values were detected. For compound **18,** the IC_{50} value was determined to be 216.7 ± 44.3 μM. The measured IC_{50} value of swainsonine was 34 ± 1.4 nM and is in the lower nanomolar range as described in the literature [43,44]. Assays with varying substrate concentrations showed the IC_{50} value of compound **18** to be independent of the substrate concentrations indicating non-competitive inhibition (Fig 7 and Fig 8).

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 (**Fig 9**).

The properties of the ligand atoms also match the properties of the dummy atoms created during binding site identification. Furthermore, the dummy atoms display areas that could be addressed during the optimization process (Fig 10).

In order to determine the selectivity of compound **18**, it was also tested for inhibition of the β-glucosidase of sweet almonds and JBM. At 300 μM no significant inhibition of the β-glucosidase of sweet almonds was observed. However, an inhibition of 70% of JBM activity was determined.

Discussion

In this study, we described the identification of a potential allosteric site using the site finder application of MOE v2014.10. The DogSite Scorer predicted a higher druggability for this site
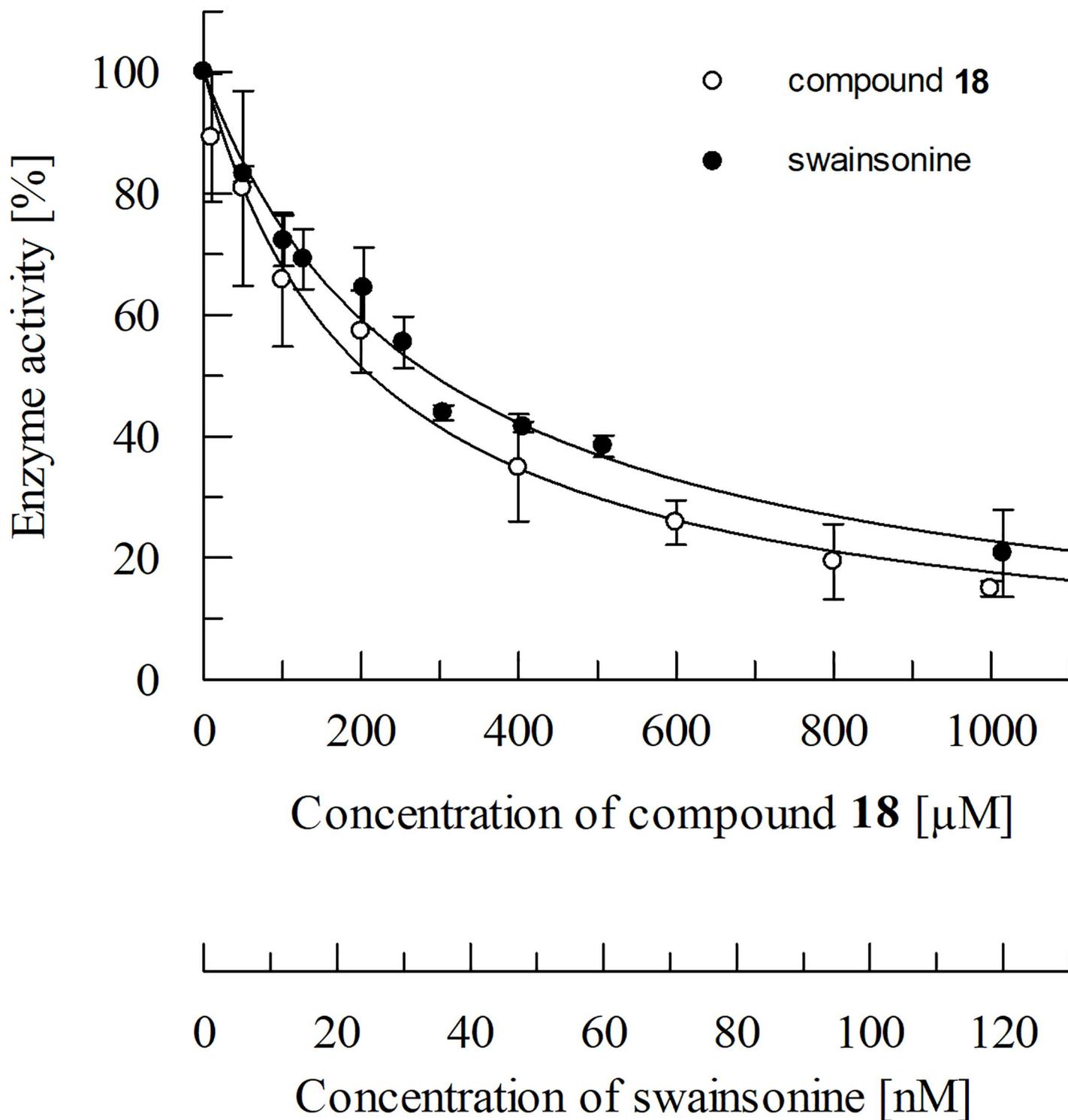
Fig 7. Example of an IC$_{50}$ curve of compound 18 and of swainsonine at a substrate (4-MU-Man) concentration of 1.5 mM. Non-linear regression was applied to the data using Eq (4).

in comparison to the known catalytic site, which is rather compact, hydrophilic and predominantly negatively charged in the region where potential ligands bind. Thus, its ligands are
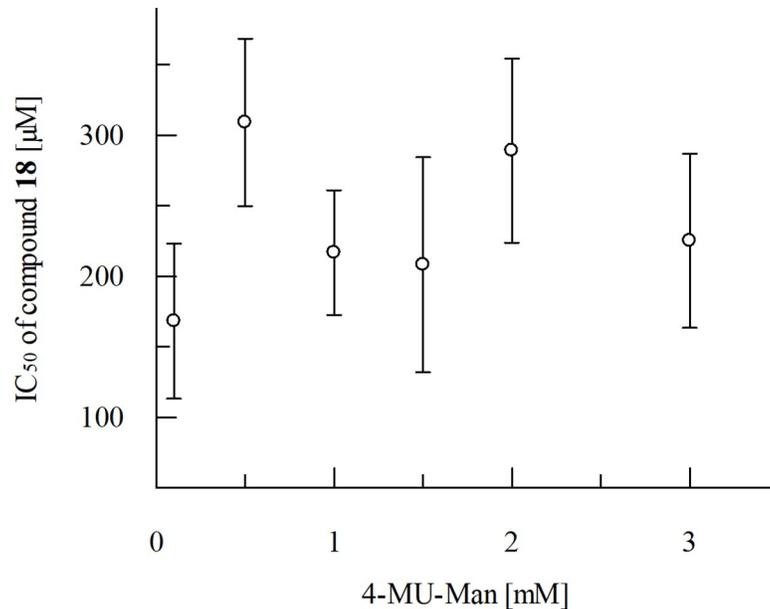
Fig 8. Effect of different substrate concentrations on the IC$_{50}$ values of compound 18 from three independent measurements. The single sample t-Test [45] was applied to calculate the significance of the results. The t-value is 1.01. The value of p is 0.37. The result is not significant at $p > 0.05$. The p-value indicates that there is no statistically significant difference in the IC$_{50}$ values of compound 18 at different substrate concentrations.

deeply buried. Thomas A. Halgren [46] has described the most suitable properties for the differentiation between "druggable", "undruggable" and "difficult" sites. The "undruggable" sites are usually much smaller and on average less protected from the solvent. Furthermore, these sites tend to display the most hydrophilic and the least hydrophobic character. According to these criteria, the catalytic site can be categorised as a "difficult" or "undruggable" site. Using DOCK 3.6, we performed a virtual screening aiming to identify new chemical scaffolds
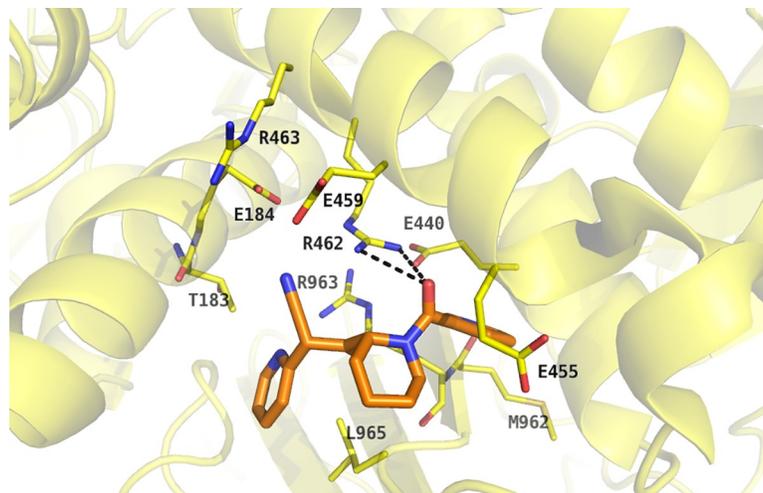

Fig 9. Predicted binding mode for compound 18. The GMII protein (PDB code: 1HWW [3]) is shown as *transparent, yellow-coloured cartoon* with binding site residues being displayed as stick representation with yellow-coloured carbon atoms. Compound 18 is represented with orange coloured carbon atoms and *black dashes* indicate potential hydrogen bonds between the virtual screening hit and the target protein.
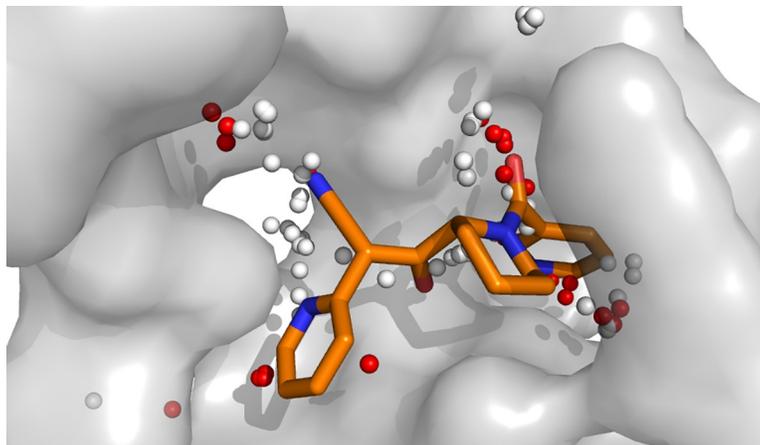
Fig 10. Comparison of the locations of alpha spheres and compound 18. Red alpha spheres indicate hydrophilic type of atoms and white hydrophobic groups favoured by the binding site regions.

possibly interacting with this allosteric binding site. These studies revealed compound **18** as a first weak inhibitor of GMII (Fig 7). Since the IC$_{50}$ value was independent of the substrate concentration (Fig 8) compound **18** is supposed to be a non-competitive inhibitor of GMII. This supports the hypothesis of having identified an allosteric site. According to the postulated binding mode, the polar interaction between the carbonyl oxygen of the amide bond and the protonated side chain of R462 might contribute to the ligand's binding affinity to its target (Fig 9). The commercially available β-glucosidase of sweet almonds was used as an off-target. Its primary structure consists of 542 amino acids compared to 1108 amino acids in the dGMII [47]. According to our homology modeling studies, both dGMII and the β-glucosidase share a low sequence identity of 12.2% and a similarity of 29.3%. Moreover, these enzymes belong to different glycoside hydrolase families with respect to retaining mechanism (GH38 and GH1, respectively) [7]. Compound **18** did not exhibit significant inhibition of the β-glucosidase of sweet almonds but showed inhibition activity against JBM (70% at 300 μM per well). The dGMII protein (PDB entry 1HWW [3]), homologous to hGMII, shares 23% identity and 40% similarity with JBM (PDB entry 6B9O [48]. The superposition of dGMII and JBM did not indicate the presence of a binding site similar to that identified in dGMII within the JBM structure. Similar to dGMII, the primary sequence of JBM suggests that this enzyme belongs to the glycoside hydrolase family 38 [49]. Thus, compound **18** can be considered as a weakly active inhibitor with some selectivity towards mannosidases in general. We identified compound **18** as a starting point for the identification of a novel class of GMII inhibitors which do not bind into the active site. Based on the inhibitory activity observed in the fluorometric assay, the ligand efficiency is approximately 0.20 kcal/mol. This is lower than the often applied threshold level of 0.30 kcal/mol for compounds suited for further optimization [50]. Despite the rather low efficiency of compound **18**, future investigations of the binding mode by e.g. site-directed mutagenesis studies might provide valuable insights into the properties that are required for ligands interacting with the potential allosteric binding site. This knowledge will then be useful to guide the design of analogues with improved affinities. Compound **18** could be structurally optimized in the future to address the amino acids R462, R463, K965, T187, E184 and E459 of the potential allosteric site of dGMII. For example, compound **18** could be optimized by addition of a substituent with cationic properties which enables the formation of strong ionic interactions with E184.

Conclusion

The identified new GMII inhibitor compound **18** ([Fig 6](#)) is a non-competitive inhibitor. This mode of inhibition underlines our proposed model of compounds binding to a potential allosteric pocket. Since the catalytic site has been studied in detail and it has been shown to be challenging to be addressed by molecular modeling [51], we opted for the identification of a potential allosteric site. 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. Thus, the identification of a first weak non-competitive inhibitor represents a first successful step into the development of novel GMII inhibitors.

Supporting information

S1 Table. List of score-ranked virtual screening hits.
(PDF)

S1 File. Synthesis of 4-methylumbellifer-7-yl-α-D-mannopyranoside (4-MU-Man) and Synthesis of 4-methylumbellifer-7-yl-β-D-glucopyranoside (4-MU-Glc).
(DOCX)

S1 Fig. SDS-PAGE of *Drosophila* proteins exported to the culture medium (A). Detection of the α-mannosidase II in the culture medium by western blot (B). Lines 1 to 5 show the expression of the α-mannosidase II in not transfected *Drosophila* cultures (1), Not induced, co-transfected cultures (2), transfected cultures with 5 (3), 10 (4) and 15 μM $CdCl_2$ (5). α-mannosidase II was purified by incubating the culture medium with Affi-Gel Blue resin to eliminate the BSA from fetal bovine serum and subsequently by immobilized metal affinity chromatography (IMAC). Protein Molecular Weight Marker (Thermo Scientific Pierce, 14.4kDa to 116kDa) was used to identify the position of α-mannosidase II (121.54 kDa).
(TIF)

S2 Fig. Synthesis of 4-methylumbellifer-7-yl-α-D-mannopyranoside (4-MU-Man).
(TIF)

S3 Fig. Synthesis of 4-methylumbellifer-7-yl-β-D-glucopyranoside (4-MU-Glc).
(TIF)

Author Contributions

Conceptualization: Lina Irsheid, Thomas Wehler, Christoph Borek, Tanja Schirmeister.

Data curation: Lina Irsheid, Thomas Wehler, Christoph Borek, Maria Elena Ortiz-Soto.

Formal analysis: Lina Irsheid, Thomas Wehler, Christoph Borek, Maria Elena Ortiz-Soto.

Investigation: Lina Irsheid, Thomas Wehler, Maria Elena Ortiz-Soto.

Methodology: Lina Irsheid, Werner Kiefer, Ruth Brenk, Maria Elena Ortiz-Soto, Tanja Schirmeister.

Project administration: Werner Kiefer, Ruth Brenk, Jürgen Seibel, Tanja Schirmeister.

Resources: Jürgen Seibel, Tanja Schirmeister.

Supervision: Werner Kiefer, Ruth Brenk, Jürgen Seibel, Tanja Schirmeister.

Validation: Thomas Wehler.

Visualization: Lina Irsheid, Thomas Wehler.

Writing – original draft: Lina Irsheid, Thomas Wehler, Christoph Borek, Maria Elena Ortiz-Soto.

Writing – review & editing: Ruth Brenk, Jürgen Seibel, Tanja Schirmeister.

References

1. Läubli H, Borsig L. Selectins promote tumor metastasis. Semin Cancer Biol. 2010; 20(3):169–77. https://doi.org/10.1016/j.semcancer.2010.04.005 PMID: 20452433

2. Goss PE, Baker MA, Carver JP, Dennis JW. Inhibitors of Carbohydrate Processing: A New Class of Anticancer Agents. Clin Cancer Res. 1995; 1(9):935–44. PMID: 9816064

3. Van den Elsen JM, Kuntz DA, Rose DR. Structure of Golgi alpha-mannosidase II: a target for inhibition of growth and metastasis of cancer cells. EMBO J. 2001; 20(12):3008–17. https://doi.org/10.1093/emboj/20.12.3008 PMID: 11406577

4. Venkatesan M, Kuntz DA, Rose DR. Human lysosomal α-mannosidases exhibit different inhibition and metal binding properties. Protein Sci. 2009; 18(11):2242–51. https://doi.org/10.1002/pro.235 PMID: 19722277

5. Rose DR. Structure, mechanism and inhibition of Golgi α-mannosidase II. Curr Opin Struct Biol. 2012 Oct; 22(5):558–62. https://doi.org/10.1016/j.sbi.2012.06.005 PMID: 22819743

6. Cheng TJR, Chan TH, Tsou EL, Chang SY, Yun WY, Yang PJ, et al. From natural product-inspired pyrrolidine scaffolds to the development of new human Golgi α-mannosidase II inhibitors. Chem—An Asian J. 2013; 8(11):2600–4.

7. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42(D1).

8. Shah N, Kuntz DA, Rose DR. Golgi α-mannosidase II cleaves two sugars sequentially in the same catalytic site. Proc Natl Acad Sci. 2008; 105(28):9570–5. https://doi.org/10.1073/pnas.0802206105 PMID: 18599462

9. Petersen L, Ardévol A, Rovira C, Reilly PJ. Molecular mechanism of the glycosylation step catalyzed by Golgi α-mannosidase II: A QM/MM metadynamics investigation. J Am Chem Soc. 2010 Jun; 132 (24):8291–300. https://doi.org/10.1021/ja909249u PMID: 20504027

10. Numao S, Kuntz DA, Withers SG, Rose DR. Insights into the Mechanism of Drosophila melanogaster Golgi α-Mannosidase II through the Structural Analysis of Covalent Reaction Intermediates. J Biol Chem. 2003 Nov; 278(48):48074–83. https://doi.org/10.1074/jbc.M309249200 PMID: 12960159

11. Li B, Kawatkar SP, George S, Strachan H, Woods RJ, Siriwardena A, et al. Inhibition of Golgi mannosidase II with mannostatin A analogues: Synthesis, biological evaluation, and structure-activity relationship studies. ChemBioChem. 2004; 5(9):1220–7. https://doi.org/10.1002/cbic.200300842 PMID: 15368573

12. Vasella A, Davies GJ, Böhm M. Glycosidase mechanisms. Curr Opin Chem Biol. 2002; 6(5):619–29. PMID: 12413546

13. Atsumi S, Umezawa K, Iinuma H, Naganawa H, Nakamura H, Iitaka Y, et al. Production, isolation and structure determination of a novel beta-glucosidase inhibitor, cyclophellitol, from Phellinus sp. J Antibiot (Tokyo). 1990; 43(1):49–53.

14. McCarter JD, Withers SG. 5-Fluoro glycosides: A new class of mechanism-based inhibitors of both α- and β-glucosidases. J Am Chem Soc. 1996; 118(1):241–2.

15. Withers SG, Street IP, Bird P, Dolphin DH. 2-Deoxy-2-fluoroglucosides: A Novel Class of Mechanism-Based Glucosidase Inhibitors. J Am Chem Soc. 1987; 109(24):7530–1.

16. Artola M, Wu L, Ferraz MJ, Kuo CL, Raich L, Breen IZ, et al. 1,6-Cyclophellitol Cyclosulfates: A New Class of Irreversible Glycosidase Inhibitor. ACS Cent Sci. 2017; 3(7):784–93. https://doi.org/10.1021/acscentsci.7b00214 PMID: 28776021

17. Tulsiani DRP, Harris TM, Touster O. Swainsonine inhibits the biosynthesis of complex glycoproteins by inhibition of Golgi mannosidase II. J Biol Chem. 1982; 257(14):7936–9. PMID: 6806288

18. Aoyagi T, Yamamoto T, Kojiri K, Morishima H, Nagai M, Hamada M, 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). 1989; 42(6):883–9.

19. Rabouille C, Kuntz DA, Lockyer A, Watson R, Signorelli T, Rose DR, et al. The Drosophila GMII gene encodes a Golgi alpha-mannosidase II. J Cell Sci. 1999; 112:3319–30. PMID: 10504337

20. Henrissat B, Bairoch A. Updating the sequence-based classification of glycosyl hydrolases. Biochem J. 1996; 316(2):695–6.

21. Paul Labute and Martin Santavy. Locating Binding Sites in Protein Structures. J Chem Comput Gr [Cited 23 September 2019] [Internet]. 2007; Available from: https://www.chemcomp.com/journal/sitefind.htm

22. Soga S, Shirai H, Kobori M, Hirayama N. Use of Amino Acid Composition to Predict Ligand-Binding Sites. J Chem Inf Model. 2007 Mar; 47(2):400–6. https://doi.org/10.1021/ci6002202 PMID: 17243757

23. Volkamer A, Kuhn D, Grombacher T, Rippmann F, Rarey M. Combining Global and Local Measures for Structure-Based Druggability Predictions. J Chem Inf Model. 2012 Feb; 52(2):360–72. https://doi.org/10.1021/ci200454v PMID: 22148551

24. Volkamer A, Kuhn D, Rippmann F, Rarey M. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. Bioinformatics. 2012 Aug; 28(15):2074–5. https://doi.org/10.1093/bioinformatics/bts310 PMID: 22628523

25. Cieplak L, Caldwell J, Kollman P. 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. 2001; 22(10):1048–57.

26. Cornell WD, Cieplak P, Bayly CI, Kollman P. Application of Resp Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation. J Am Chem Soc. 1993; 115 (21):9620–31.

27. Brenk R, Vetter SW, Boyce SE, Goodin DB, Shoichet BK. Probing Molecular Docking in a Charged Model Binding Site. J Mol Biol. 2006 Apr; 357(5):1449–70. https://doi.org/10.1016/j.jmb.2006.01.034 PMID: 16490206

28. Mysinger MM, Shoichet BK. Rapid Context-Dependent Ligand Desolvation in Molecular Docking. J Chem Inf Model. 2010 Sep 27; 50(9):1561–73. https://doi.org/10.1021/ci100214a PMID: 20735049

29. Weininger D. Smiles, a Chemical Language and Information-System .1. Introduction to Methodology and Encoding Rules. J Chem Inf Comput Sci. 1988; 28(1):31–6.

30. Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, et al. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. ChemMedChem. 2008 Mar; 3 (3):435–44. https://doi.org/10.1002/cmdc.200700139 PMID: 18064617

31. Hawkins PCD, Skillman AG, Warren GL, Ellingson BA, Stahl MT. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. J Chem Inf Model. 2010 Apr; 50(4):572–84. https://doi.org/10.1021/ci100031x PMID: 20235588

32. Wei BQ, Baase WA, Weaver LH, Matthews BW, Shoichet BK. A model binding site for testing scoring functions in molecular docking. J Mol Bio. 2002; 322(2):339–55.

33. Lorber DM, Shoichet BK. Flexible ligand docking using conformational ensembles. Protein Sci. 2008 Dec; 7(4):938–50.

34. Schrödinger LLC. The {PyMOL} Molecular Graphics System, Version 1.5.0.5. 2013.

35. Kuntz DA, Zhong W, Guo J, Rose DR, Boons GJ. The molecular basis of inhibition of golgi α-mannosidase II by mannostatin A. ChemBioChem. 2009; 10(2):268–77. https://doi.org/10.1002/cbic.200800538 PMID: 19101978

36. Coleman DJ, Kuntz DA, Venkatesan M, Cook GM, Williamson SP, Rose DR, et al. A long-wavelength fluorescent substrate for continuous fluorometric determination of alpha-mannosidase activity: Resorufin alpha-d-mannopyranoside. Anal Biochem. 2010; 399(1):7–12. https://doi.org/10.1016/j.ab.2009.11.039 PMID: 20026005

37. Touisni N, Maresca A, McDonald PC, Lou Y, Scozzafava A, Dedhar S, et al. Glycosyl coumarin carbonic anhydrase IX and XII inhibitors strongly attenuate the growth of primary breast tumors. J Med Chem. 2011; 54(24):8271–7. https://doi.org/10.1021/jm200983e PMID: 22077347

38. Malet C, Viladot JL, Ochoa A, Gallégo B, Brosa C, Planas A. Synthesis of 4-methylumbelliferyl-β-d-glucan oligosaccharides as specific chromophoric substrates of (1 → 3),(1 → 4)-β-d-glucan 4-glucanohydrolases. Carbohydr Res. 1995; 274(C):285–301.

39. Pistia-Brueggeman G, Hollingsworth RI. A preparation and screening strategy for glycosidase inhibitors. Tetrahedron. 2001; 57(42):8773–8.

40. Zhang J-H. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen. 1999; 4(2):67–73. https://doi.org/10.1177/108705719900400206 PMID: 10838414

41. Yung-Chi C, Prusoff WH. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (IC50) of an enzymatic reaction. Biochem Pharmacol. 1973; 22(23):3099–108. PMID: 4202581

42. Yamashita MM, Wesson L, Eisenman G, Eisenberg D. Where metal ions bind in proteins. Proc Natl Acad Sci. 1990; 87(15):5648–52. https://doi.org/10.1073/pnas.87.15.5648 PMID: 2377604

43. Fiaux H, Kuntz DA, Hoffman D, Janzer RC, Gerber-Lemaire S, Rose DR, 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. 2008; 16(15):7337–46.

44. Rabouille C, Kuntz DA, Lockyer A, Watson R, Signorelli T, Rose DR, et al. The Drosophila GMII gene encodes a Golgi alpha-mannosidase II. J Cell Sci. 1999; 112:3319–30. PMID: 10504337

45. T-Test Calculator for a Single Sample [Cited 29 March 2019][Internet]. Available from: https://www.socscistatistics.com/tests/tsinglesample/default.aspx

46. Halgren TA. Identifying and characterizing binding sites and assessing druggability. J Chem Inf Model. 2009; 49(2):377–89. https://doi.org/10.1021/ci800324m PMID: 19434839

47. Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R, et al. UniProt: The universal protein knowledgebase. Nucleic Acids Res. 2017; 45(D1):D158–69. https://doi.org/10.1093/nar/gkw1099 PMID: 27899622

48. Complexes EI, Howard E, Cousido-siah A, Lepage ML, Schneider JØØP, Bodlenner A, et al. Structural Basis of Outstanding Multivalent Effects in Jack Bean a -Mannosidase Inhibition Angewandte. 2018; (789):8002–6.

49. Gnanesh Kumar BS, Pohlentz G, Schulte M, Mormann M, Siva Kumar N. Jack bean -mannosidase: Amino acid sequencing and N-glycosylation analysis of a valuable glycomics tool. Glycobiology. 2014; 24(3):252–61. https://doi.org/10.1093/glycob/cwt106 PMID: 24295789

50. Hopkins AL, Keserü GM, Leeson PD, Rees DC, Reynolds CH. The role of ligand efficiency metrics in drug discovery. Nat Rev Drug Discov. 2014; 13(2):105–21. https://doi.org/10.1038/nrd4163 PMID: 24481311

51. Englebienne P, Fiaux H, Kuntz DA, Corbeil CR, Gerber-Lemaire S, Rose DR, et al. Evaluation of docking programs for predicting binding of Golgi α-mannosidase II inhibitors: A comparison with crystallography. Proteins Struct Funct Genet. 2007; 69(1):160–76. https://doi.org/10.1002/prot.21479 PMID: 17557336