

Accepted Manuscript

From Carbohydrates to Drug-Like Fragments: Rational Development of Novel α -Amylase Inhibitors

Jamil Al-Asri, Erika Fazekas, Gábor Lehoczki, Andrej Perdih, Cornelia Görick, Matthias F. Melzig, Gyöngyi Gyémánt, Gerhard Wolber, Jérémie Mortier

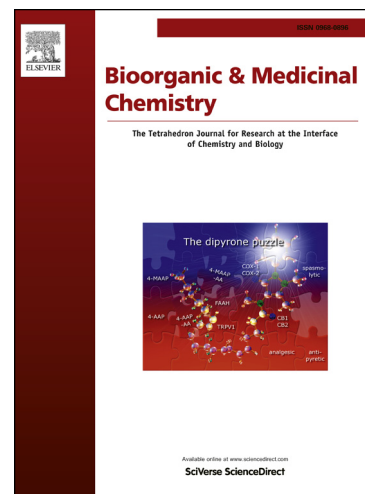
PII: S0968-0896(15)30030-4
DOI: <http://dx.doi.org/10.1016/j.bmc.2015.09.007>
Reference: BMC 12554

To appear in: *Bioorganic & Medicinal Chemistry*

Received Date: 1 July 2015
Revised Date: 31 August 2015
Accepted Date: 3 September 2015

Please cite this article as: Al-Asri, J., Fazekas, E., Lehoczki, G., Perdih, A., Görick, C., Melzig, M.F., Gyémánt, G., Wolber, G., Mortier, J., From Carbohydrates to Drug-Like Fragments: Rational Development of Novel α -Amylase Inhibitors, *Bioorganic & Medicinal Chemistry* (2015), doi: <http://dx.doi.org/10.1016/j.bmc.2015.09.007>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



From Carbohydrates to Drug-Like Fragments: Rational Development of Novel α -Amylase Inhibitors

Jamil Al-Asri^a, Erika Fazekas^b, Gábor Lehoczki^b, Andrej Perdih^c, Cornelia Görick^a, Matthias F. Melzig^a, Gyöngyi Gyémánt^b, Gerhard Wolber^{a*}, Jérémie Mortier^{a*}

^a Freie Universität Berlin, Institute of Pharmacy, Department of Pharmaceutical & Medicinal Chemistry, Königin-Luise Str. 2-4, 14195 Berlin, Germany

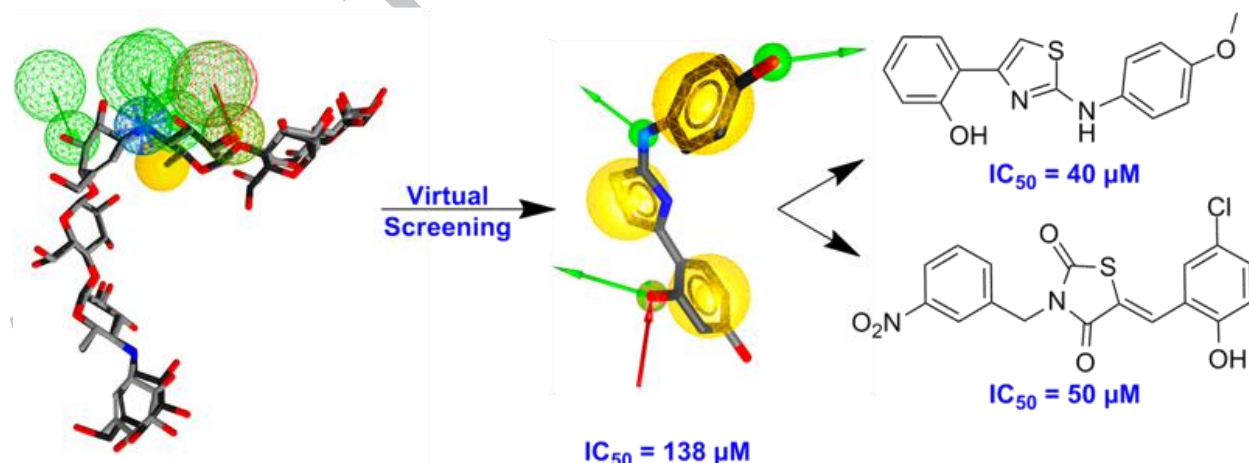
^b University of Debrecen, Department of Inorganic & Analytical Chemistry, Egyetem ter 1, POBox 21, H-4032 Debrecen, Hungary

^c National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia

Abstract

Starch catabolism leading to high glucose level in blood is highly problematic in chronic metabolic diseases, such as type II diabetes and obesity. α -Amylase catalyses the hydrolysis of starch, increasing blood sugar concentration. Its inhibition represents a promising therapeutic approach to control hyperglycaemia. However, only few drug-like molecule inhibitors without sugar moieties have been discovered so far, and little information on the enzymatic mechanism is available. This work aims at the discovery of novel small α -amylase binders using a systematic *in silico* methodology. 3D-pharmacophore-based high throughput virtual screening of small compounds libraries was performed to identify compounds with high α -amylase affinity. Twenty-seven compounds were selected and biologically tested, revealing IC_{50} values in the micromolar range and ligand efficiency higher than the one of the bound form of acarbose, which is used as a reference for α -amylase inhibition.

Graphical abstract



* Corresponding authors:

GW: +49 30 838 52686 - gerhard.wolber@fu-berlin.de

JM: +49 30 838 58090 - jeremie.mortier@fu-berlin.de

1. Introduction

Metabolic diseases like obesity and type II diabetes are characterized by high levels of blood glucose.¹ One of the most attractive targets for the development of novel therapeutic agents and controlling the glucose level in the blood is α -amylase.^{2, 3} α -Amylases (1,4- α -D-glucan-4-glucanohydrolase, EC 3.2.1.1) are endoamylases broadly distributed throughout the microbial, plant and animal kingdoms,^{4, 5} which catalyze hydrolysis of starch and glycogen.⁶ In humans, α -amylase is expressed as two isoforms, secreted from salivary glands and pancreas, respectively. These enzymes play a role in digestion of polysaccharides like starch, the main source of glucose in human diet.⁷⁻⁹ Human salivary α -amylase (HSA) initiates the hydrolysis of α -(1,4) glycosidic bonds in the dietary starch into smaller oligosaccharides.¹⁰ Further digestion of carbohydrates takes place in the gut under control of human pancreatic α -amylase (HPA).¹¹ Hence, hydrolysis of polymeric starch by HSA and HPA produces maltose, maltotriose and other oligomers, which are then degraded into glucose by α -glucosidase.¹²

Therefore, α -amylase inhibitors can control the breakdown of dietary starch into smaller oligomers,¹³ with the purpose of delaying glucose absorption to a rate the body can handle and hence decreasing the post-prandial hyperglycemia.^{2, 14, 15} α -Amylase inhibition was recently observed with natural products extracted from Hungarian sour cherries.¹⁶ But peptide-based compounds were the first reported highly potent α -amylase inhibitors.¹⁷⁻¹⁹ Carbohydrate-based inhibitors have been studied intensively as α -amylase inhibitors, showing a similar range of potency as the peptide-based inhibitors.⁷ Acarbose, a natural metabolite of *Streptomyces* sp., has been used in the treatment of type II diabetes and was shown to inhibit α -amylase and α -glucosidase but with undesirable gastrointestinal disturbances.²⁰ These side effects arise as a consequence of maltose fermentation which accumulates due to α -glucosidase inhibition. Therefore, a non-carbohydrate-based α -amylase inhibitor is expected to be better agent to restrain postprandial hyperglycemia since it would not lead to abdominal accumulation of maltose.²¹

In humans, α -amylases secreted from pancreas and saliva are two very closely related isoforms with 97 % sequence identity (92 % in the catalytic domains).²²⁻²⁴ The active site is located at a large V-shaped cavity in the catalytic domain to accommodate the natural substrate, starch.²⁵ Subsites in this binding cleft are denoted with the numbers from -4 to +3, according to Davies et al.^{7, 26, 27} The cleavage of the natural substrate takes place between subsites -1 and +1, where catalytic residues Asp197, Glu233 and Asp300 play a central role in the hydrolysis mechanism.¹²

In this work, we aim at identifying novel small non-sugar drug-like molecules binding to α -amylase and inhibiting its activity. To reach this objective, a structure-based approach was developed using 3D-pharmacophore models and virtual screening.²⁸

2. Results and Discussion

2.1 Development and Validation of the 3D-Pharmacophore Model

First, all inhibitors co-crystallized with α -amylase available from the Protein Data Bank (PDB)^{29, 30} were collected in order to build a pharmacophore model that compiles all essential features required for an

optimal ligand-enzyme interaction in the catalytic pocket. Due to their different binding areas further away from the enzymatic active site, peptide-based inhibitors were excluded from this preliminary analysis.^{25, 31} Assuming the cleavage site of α -amylase (at subsites -1 and +1) is a central domain for substrate binding, particular attention was paid to interactions observed between the enzyme and the investigated ligands in this region. The study of carbohydrate-based inhibitor binding modes shows that one water molecule (HOH-746) is highly conserved among available crystal structures in the PDB. This molecule often stabilizes co-crystallized ligands in the enzyme by bridging H-bonds between enzyme and inhibitor. Superposing the studied crystal structures (Supplementary data, Fig. S1) revealed that HOH-746 is located nearly at the same position in all investigated macromolecules. This interaction was included in our 3D pharmacophore model to retain this particular environment of the pocket, surrounded by residues Glu233 and Lys200 (Figure 1).

The reliability of the developed model (i.e. ability to discriminate between inhibitors and non-inhibitors) was assessed and improved using a dataset of 19 active compounds and 55 known inactive compounds assembled from the literature and the ChEMBL database (see supplementary data, Tables S1 and S2).³² Multi-conformational screening of these compounds was carried out and the signal-noise ratio was determined using a Receiver Operating Characteristics (ROC) curve.³³ Iterative development and validation steps yielded a 3D-pharmacophore model with good early enrichment factor value ($EF_{1\%} = 3.9$). This model exhibited an excellent sensitivity in retrieving 63 % (12 out of 19) of the active compounds while being highly restrictive, i.e. excluding 98 % (54 out of 55) of inactive compounds.

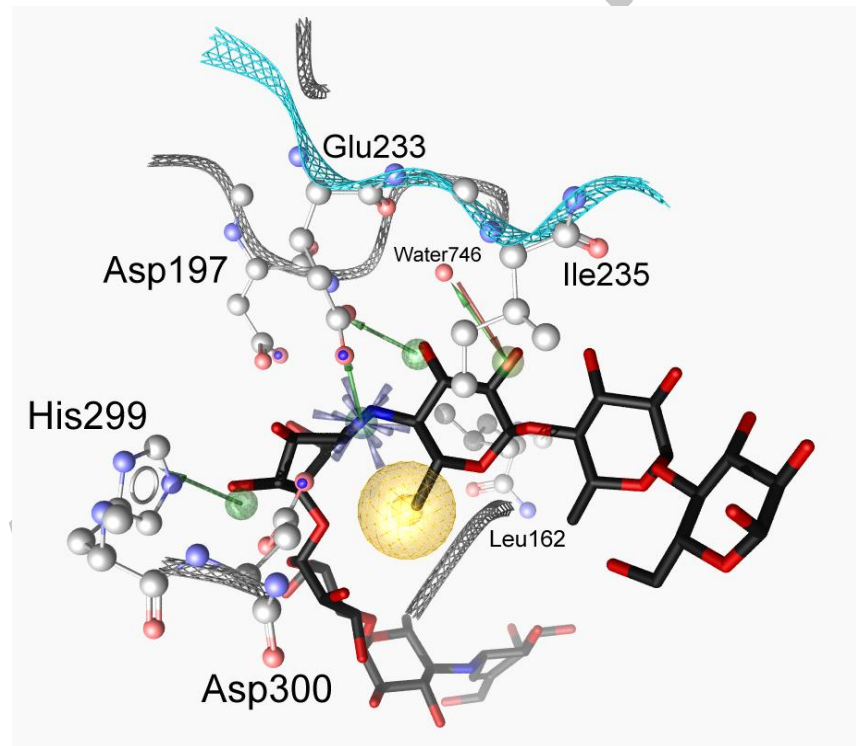


Figure 1. Acariostatin II03 inhibitor ($K_i \sim 14$ nM, PDB: 3OLE)⁷ aligned with the developed 3D pharmacophore that contains the following chemical features: 3 H-bond donors (green), 1 H-bond

acceptor (red), 1 positive ionizable interaction (blue star), 1 hydrophobic contact (yellow sphere), and 25 exclusion volumes (not shown for clarity).

The chemical features of the final model mainly represent ligand-enzyme interactions detected at the center of the catalytic cavity, where the catalytic triad Asp197-Glu233-Asp300 is located. This model consists of one positive ionizable interaction (PI) with the carboxylic groups of the catalytic triad, one hydrophobic contact (HYD) with Ile235 and, deeper in the active site, Leu162, one H-bond acceptor (HBA) with the water molecule HOH746, and four H-bond donors (HBD) (Figure 1). Among all HBD, two were set as essential features: The one involving the oxygen OE2 from the side chain of Glu233 and the second with the oxygen of the water molecule. Regarding the optional HBD, one involves the nitrogen atom NE2 of His299 and, the second, the carbonyl oxygen OE1 of Glu233. Exclusion volumes were defined as forbidden areas to represent the steric constraints of the pocket.

A further validation was conducted for the developed model by screening drugs and biological compounds available from the databases Derwent World Drug Index 2005 (WDI, www.thomsonreuters.com), MDDR2009 (www.symyx.com), and DrugBank.³⁴ This validation step was carried out for three reasons: (i) to investigate the ability of the model to retrieve commercial hits either with the same or different biological activity, (ii) to examine the efficiency of the model by ranking the 12 previously retrieved active compounds when screened altogether with drug databases (iii) to give a first impression on the restrictiveness of the model by specifying the total number of the retrieved hits from these three databases. Among the twelve biologically known actives previously retrieved during the first validation, eleven compounds were ranked within the first 100 molecules, according to their pharmacophore fit score calculated with LigandScout 3.1.³⁵ About 17 % of the compounds retrieved and ranked within these first 100 belonged to the trestatin family, such as acarbose, and contained an acarviosine scaffold. This means that the developed model can retrieve active compounds and prioritize α -amylase inhibitors over non inhibitors. This model was considered suitably restrictive as the overall amount of recovered hits from WDI, MDDR and DrugBank was about ~ 1.4 % (1,969 out of 141,233 compounds). The main values from this validation step are summarized in Table 1.

Table 1 : Results of the computational validation of the final model by screening drugs and biological compounds databases

Database	Database size	Retrieved hits from		EF _{1%}
		Database	Known actives	
WDI2005	64,255	1,141 (1.77%)	12	58.0
MDDR2009	72,383	752 (1.03%)	12	63.2
DrugBank	4,595	76 (1.65%)	12	58.1
Total	141,233	1,969 (1.39%)		

EF_{1%}: early enrichment factor

2.2 Virtual Screening

After validation of the 3D pharmacophore, the model was used for virtual screening of 1,762,189 drug-like compounds from commercial libraries (Analyticon, Asinex, Life chemicals, Chembridge, Specs, Bionet, and Prestwick). The 3D model picked 5,748 hits, which is 0.32 % of the full library. Retrieved hits were ranked using the pharmacophore fit and the first 3,000 compounds were retained. Subsequently, 2D descriptors were used (i.e. HBD \leq 2-5, HBA \leq 10, Mol. Wt. \leq 700 and ClogP \leq 5) to exclude non drug-like compounds and select 2,292 molecules for further processing. In order to reduce the amount of inhibitors to handle prior to biological experiments, two different strategies were applied: virtual docking and structural clustering. On one hand, software GOLD 5.1³⁶ was used to dock all selected virtual hits into the α -amylase active site (PDB code 3OLE).⁷ With LigandScout, poses were minimized inside the enzymatic pocket (using the MMFF94 force field)³⁷ and prioritized based on their ability to geometrically fulfill the features compiled in the 3D-pharmacophore. Careful visual inspection was undertaken for each compound conformation with the highest fit scores and the 30 most promising structures were selected for the next step. On the other hand, all 2,292 virtual hits were subjected to structural clustering with the software Jklustor 5.8.0.³⁸ The 17 candidates with the highest structural diversity were selected for further investigation. Based on their availability at the vendors, 9 virtual hits were purchased for biological testing (Figure 2).

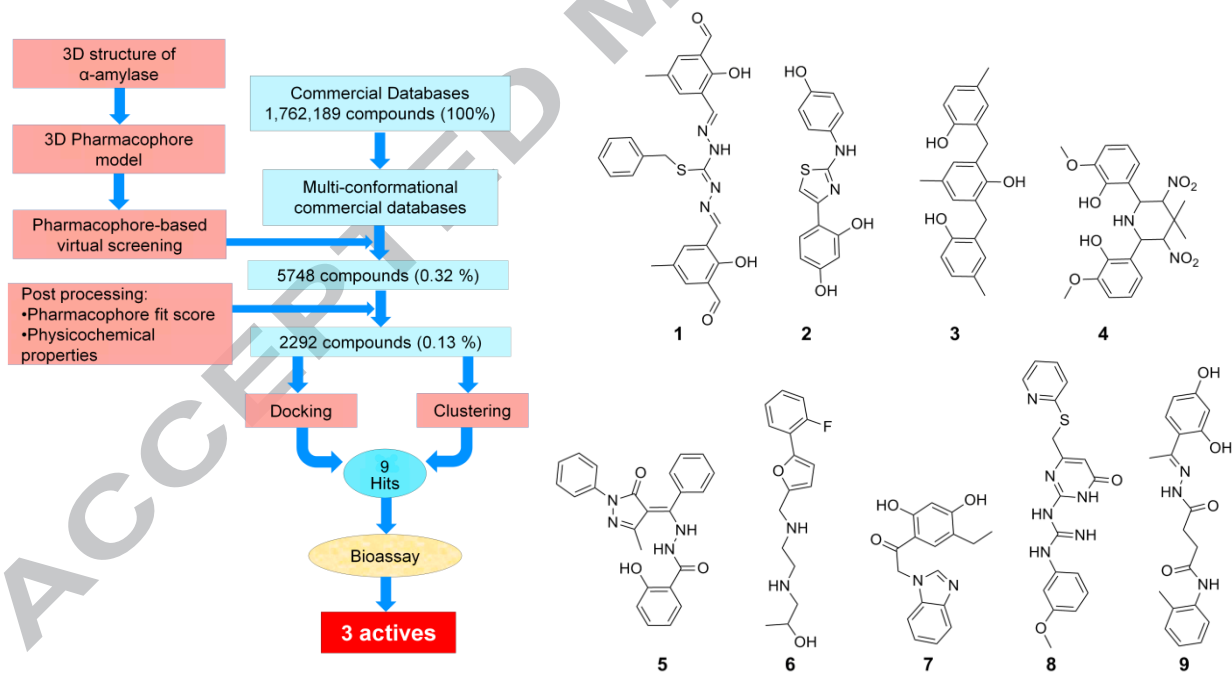


Figure 2. Virtual screening workflow (left) and selected molecules (right)

Kinetic assay was carried out using 2-chloro-4-nitrophenyl-4-O- β -D-galactopyranosylmaltoside (GalG₂CNP) as a substrate for HAS.³⁹ The released chromogenic product chloronitrophenol (CNP) was

continuously monitored at 400 nm with respect to time. Using this method, the inhibitory effect of compounds **2 - 9** was evaluated (**1** was discarded due to its low solubility). In this assay, compounds **2**, **3**, and **4** revealed promising inhibition potencies against HSA with $IC_{50} \leq 300 \mu\text{M}$ (Table 2). Compounds **2** and **3** demonstrate competitive inhibition with IC_{50} values of 138 μM and 200 μM , respectively. These inhibitors are characterized by three ring-backbones. Compound **2** possesses *N*,4-diphenylthiazole-2-amine scaffold. While compound **4** retains a dinitro-piperidine fragment substituted with hydroxyphenyl rings, compound **3** is composed of tri-*p*-cresol rings. Although less potent than the reference inhibitor acarbose ($IC_{50} = 0.5 \mu\text{M}$), these compounds bear a promising ligand efficiency (LE) due to their smaller size. Defined as the ratio between potency and their number of heavy atoms (HA), LE can be determined using the equation $LE = (1.37/HA) \times pIC_{50}$.⁴⁰⁻⁴² The active compounds identified by virtual screening own higher LE compared to the bound acarbose (Table 2). These small molecules are therefore considered promising compounds that can be used as a structural query to lead the development of novel potent α -amylase inhibitors.

Table 2 Biological results and ligand efficiency values for the tested compounds

Compound	Mol.wt	HA	GalG2CNP, HSA	
			IC_{50} (μM)	Ligand Efficiency
Bound acarbose*	948.92	65	0.5	0.13
1	489.56	35	NT	-
2	300.33	21	140	0.25
3	348.43	26	200	0.19
4	447.43	32	300	0.15
5	413.44	31	ND	-
6	294.36	21	1000	0.19
7	296.32	22	ND	-
8	382.44	27	330	0.18
9	355.38	26	850	0.16

HA stands for number of heavy atoms, ND stands for inactive at 200 μM and NT means not tested due to insolubility.

* Bound acarbose (pseudo hexasaccharide) in the active site of HSA (PDB: 1MFV)⁴³ was used for LE calculations. All experiments were repeated three times and the obtained IC_{50} did not vary more than 5 %

Re-docking of the active compounds in the binding site of α -amylase was conducted to rationalize their structure-activity relationships (SAR) using their potential binding modes and comparing them to the co-crystallized inhibitor acarviosin O113 ($K_i \sim 14 \text{ nM}$, PDB code of enzyme 3OLE).⁷ The ideal accommodation in subsites -1 and +1 of compound **2** can explain the inhibitory effect of this small molecule. Indeed, the three ring-backbone of this inhibitor fully occupies subsites -2, -1, and +1, and partially +2. This interaction with the catalytic core of the cavity is similar to the one of acarviosine, the

subunit of acarviostatin II03 and acarbose holding a valienamine and a 4,6-dideoxyglucose fragments. This essential scaffold of acarbose and other trestatin family inhibitors interacts with the central subsites -1 and +1. Our study shows that the hydroxyl group of resorcinol in compound **2** can form two H-bonds with His305 and the catalytic residue Asp300, in subsite -1. The side chain of Glu233, another catalytic residue in subsite +1, can form a H-bond with the amino group linking to the thiazole and phenolic rings of compound **2**. An additional H-bond can be formed between the phenolic hydroxyl group and Tyr151 near subsite +2. Interestingly, in the proposed binding mode of compound **2**, the thiazole moiety occupies subsite -1, just like valienamine, and stabilizes the inhibitor via hydrophobic contact with the side chain of Tyr62. Lipophilic residues Leu165, Trp58 near subsite -2, on one hand, and Ile235 and Leu162 near subsite +1, on the other hand, can form hydrophobic contacts with the aromatic resorcinol and phenolic rings, respectively. These interactions can explain the good stabilization of compound **2** in the catalytic core of α -amylase (Figure 3).

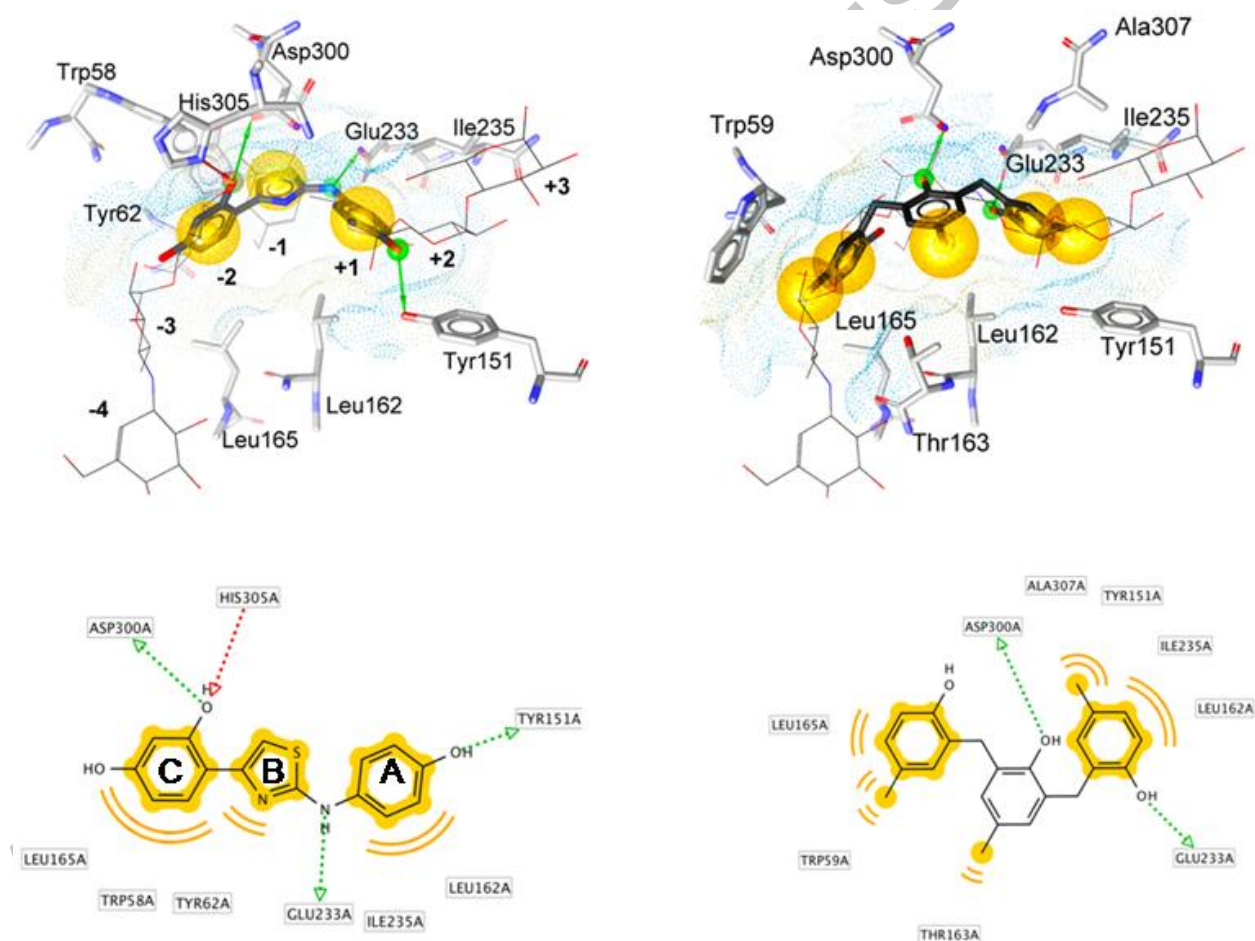


Figure 3. Plausible binding modes for compounds **2** (left) and **3** (right) as thick black sticks in 3D superposed to the co-crystallized acarviostatin II03 as thin lines (above) and in 2D (below). Color codes: Yellow spheres for hydrophobic contacts; green and red arrows for H-bond donors and acceptors, respectively.

Similarly, our docking study shows that compound **3** also occupies the catalytic core (subsites -1 and +1) and protrudes to the nearby subsites. Superposition with acarviostatin II03 highlights a similar binding mode. The predicted conformation of **3** suggests that two H-bonds can be formed between the catalytic residues Asp300 and Glu233 and the hydroxyl groups of two *p*-cresol rings. Methyl group of the central *p*-cresol fragment of **3** can point to subsite -3 and form hydrophobic contact with the side chain of Thr163, stabilizing the inhibitor in the pocket. The other two *p*-cresol rings can form hydrophobic contacts in subsite -2 with Leu165 and Trp59, subsite +1 with Leu162 and Ile235, and near subsite +2 with Tyr151. Although compound **3** is spanning from subsites -2 to +2, its inhibitory potency ($IC_{50} = 200 \mu\text{M}$) was lower compared to compound **2** ($IC_{50} = 138 \mu\text{M}$). This can be due to (i) the fewer amount of H-bonds that could be detected compared to the binding mode of **2**, and (ii) the thiazole ring of **2** that is better accommodated than the *p*-cresol of **3**, and could play an important role in stabilization of the ligand in subsite -1 (Figure 3).

Compound **4** shows a weaker inhibitory potency ($IC_{50} \sim 300 \mu\text{M}$) even though the docking study shows it can occupy subsites -3 to +2. In our model, the catalytic residue Asp300 forms a positive ionizable interaction with the amino group of the piperidine moiety, contributing to the stabilization of this ligand in the cavity. This inhibitor is stabilized inside the pocket by forming hydrophobic contacts with lipophilic residues Thr163, Tyr62, Leu165, Leu162, and Ile235. Nevertheless, the presence of the nitro groups seems to affect its binding and the lack of adequate interaction with the catalytic residues can play a crucial role in the weaker inhibitory potency of **4**.

2.3 Similarity Search

Compound **2** represents the most active competitive inhibitor with the highest ligand efficiency identified in this virtual screening. Interestingly, **2** is also an allosteric inhibitor of the fructose 1,6-bisphosphatase ($IC_{50} = 11 \mu\text{M}$)⁴⁴, an enzyme involved in gluconeogenesis. Thus, this chemical scaffold has the potential to act as dual inhibitor and to decrease glucose levels in blood. Therefore, this molecule was considered as a promising scaffold for the design of non-carbohydrate anti-diabetic molecule, and used as template to search for similar molecules with improved α -amylase inhibition.

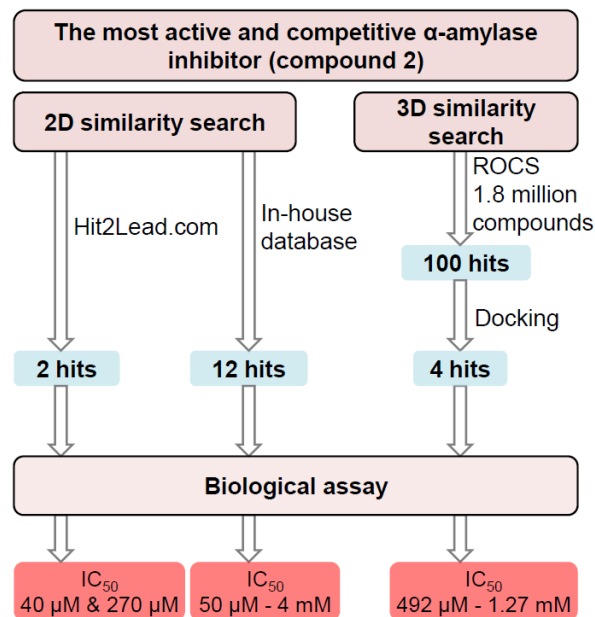


Figure 4. Overview of the similarity search approaches using compound **2** as query.

For the selection of chemical analogues, 2D and 3D similarity approaches were considered in parallel, (Figure 4). The first similarity search was conducted using the software ROCS 3.1.2⁴⁵ performing a high throughput 3D alignment of our library of 1,762,189 commercially available molecules. The 100 molecules with the best alignments (similarity with the lead query in terms of shape and chemical features) were selected and docked into the α -amylase active site (PDB: 3OLE).⁷ A pharmacophore model of compound **2** was used to prioritize the compounds able to create similar interactions with the enzyme. This step resulted in the selection of four compounds (**22** - **25**,

Table 3). The second approach was a 2D similarity search. This was performed with (i) an in-house collection of molecules, from which 12 compounds were selected, and (ii) on the other side with the Chembridge library, from which 2 compounds were purchased (molecules **10** - **21** & **26** - **27**,

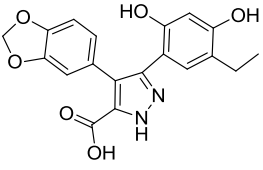
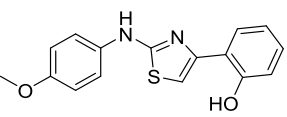
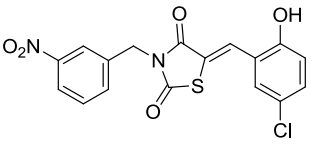
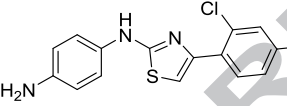
Table 3).

Overall, 18 analogues of **2** were selected for further biological evaluation. IC_{50} values ranging from 40 μ M to 4 mM were measured (

Table 3). Interestingly, minor structural changes were shown to have a strong impact on the inhibitory potency of these compounds (Figure 5). For instance, *p*-dimethylamino (**14**, $IC_{50} \sim 100 \mu M$) or *m*-acetyl (**15**, $IC_{50} \sim 100 \mu M$) substitutions in ring A increase the inhibition by about 1.4-folds compared to compound **2** ($IC_{50} = 138 \mu M$). In contrast, the presence of the *p*- or *m*-carboxyl group (**10** - **12**, IC_{50} range: 423 - 680 μM) diminishes the potency up to 5-folds while the *p*-sulfamoyl group (**13**, $IC_{50} = 165 \mu M$) shows an inhibitory potency in the same range as compound **2**.

Table 3 : Inhibitory potency (IC_{50}) of analogues of compound **2**

Inhibitor	Structure	IC_{50} (μM)	Inhibitor	Structure	IC_{50} (μM)
10		680	19		62
11		642	20		260
12		423	21		700
13		165	22		492
14		100	23		735
15		100	24		1043
16		4000	25		1274

Inhibitor	Structure	IC ₅₀ (μ M)	Inhibitor	Structure	IC ₅₀ (μ M)
17		4000	26		40
18		50	27		270

Interestingly, a dramatic decrease in activity is observed when ring B is substituted with a pyrazole ring or in the absence of the NH-bond between ring A and ring B (**16** & **17**, IC₅₀ ~ 4 mM). If a *p*-methoxy group is flanking ring A, a 3-fold improvement in inhibition is observed (compound **26**, IC₅₀ = 40 μ M). Comparing substitutions in ring C between **2** and compounds **22** – **27** shows that the hydroxyl-group in *ortho*-position is important for activity, since those that lack this substituent displays weaker inhibitory potencies.

Substitutions in the 5-benzylidenethiazolidin-4-one scaffold of **18** - **21** exert different effects on α -amylase activity. The *m*-carboxyphenyl group of **19** (IC₅₀ = 62 μ M) decreases the potency compared to the *m*-nitrobenzyl group in **18** (IC₅₀ = 50 μ M). However, **18** and **19** show more than 2-fold increase in potency compared to **2**. Substitution in the previously mentioned scaffold with a *m*-carboxyphenylamino moiety (**20**, IC₅₀ = 260 μ M) or 4-acetylpiperazinyl ring combined with no NH-bond between rings A and B (**21**, IC₅₀ ~ 700 μ M) results in a 5- and 14-folds decrease in α -amylase inhibition, respectively.

Because of their small size, all analogues **10** - **27** have a higher LE than the positive control (LE = 0.13). The best inhibitor in the series **26** even displays an efficiency of 0.30 (

Table 4).

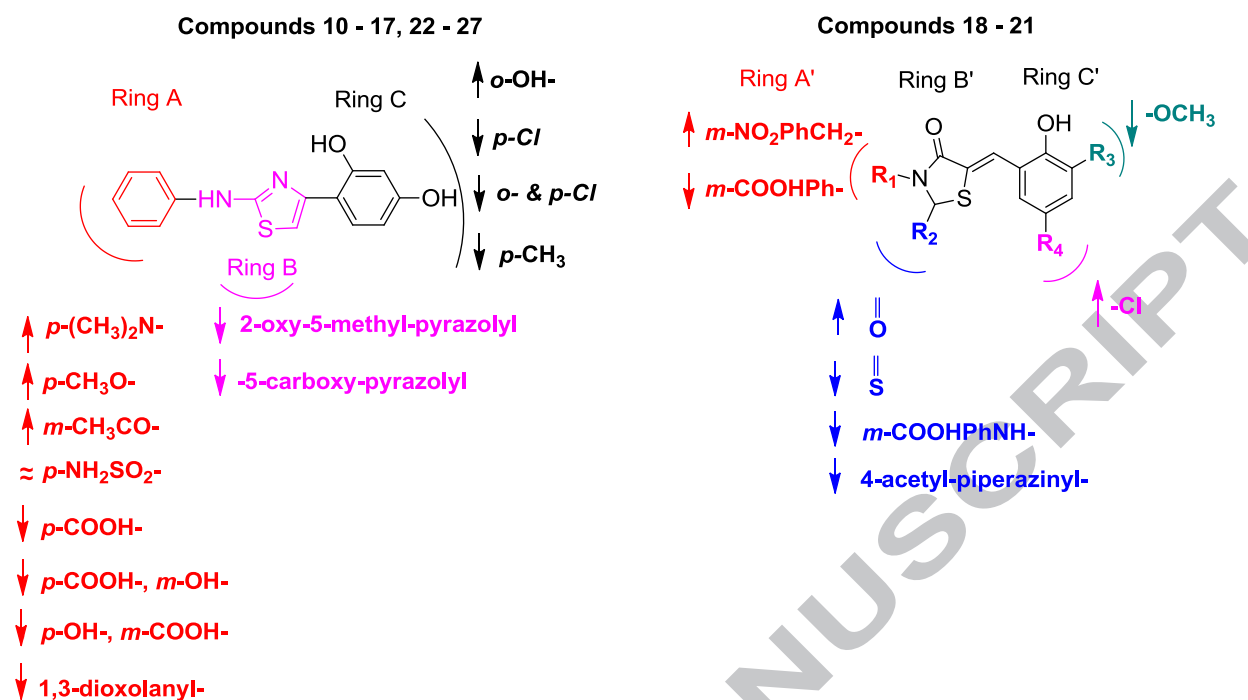


Figure 5. Overview of the structure-activity relationships for analogues of compound **2**

To rationalize the structure-activity relationships (SARs) of this chemical series, the binding conformations inside the α -amylase cavity were investigated by docking each molecule in the enzyme active site. All compounds were docked into the catalytic active site (PDB entry 3OLE)⁷ using GOLD 5.1.^{36, 46} Plausible binding conformations were prioritized based on their ability to fulfill the pharmacophore features extracted from the bound conformation of **2**, and therefore to create similar interactions with the same residues in the catalytic center (Figure 6). More details about the predicted poses are given in supplementary information, Fig. S3 and S4.

Since compounds **10** – **15** and **22** - **27** share the *N*,4-diphenylthiazole-2-amine scaffold with **2**, they can form similar interactions. However, the presence of the strongly electron withdrawing carboxyl group of **10** – **12**, near to the carboxylic amino acid Glu240 in subsite +2, can explain the weaker activity of these compounds. Compound **12**, flanked by a carboxyl group in *meta*-position, shows better inhibitory potency than **10** and **11** with *para*-carboxyl groups located further away from Glu240. Interestingly, compared to **2**, compounds **13** – **15** (IC_{50} 100 – 165 μM) and **26** (IC_{50} = 40 μM) have bulkier groups flanking ring A and hence exert better potency.

The SARs of compounds sharing a 5-benzylidenethiazolidin-4-one scaffold (**18** – **21**) suggest that the H-bond between the 4-carbonyl oxygen of thiazolidin-4-one of **18** and Arg195 in subsite -1 is important for α -amylase inhibition, as **19** - **21** do not interact with this residue. Moreover, the chlorine substituent of **18** can form a hydrophobic contact with Ile235 in subsite +1 that stabilizes the ligand binding.

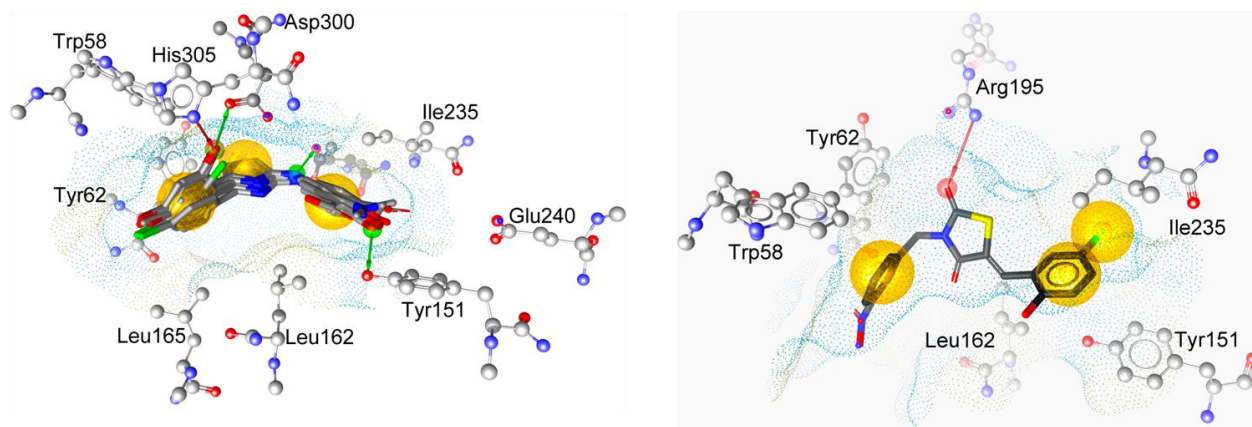


Figure 6. Predicted binding poses of compound 2 analogues. Left: superposition of compounds 2, 11, 12, 14, 15, 22 – 27, right: compound 18. Yellow spheres represent hydrophobic contacts, red and green arrows show H-bond acceptors and donors, and yellow and blue dots represent hydrophobic and lipophilic surfaces of the receptor binding pocket.

Investigation of the docking poses of **16** and **17** reveals that the pyrazole ring is located near to the charged subsite -1 (including Asp197, Glu233 and Asp300) and can form favorable interactions with this region of the protein. However, a substitution by a methyl group (**16**), or a carboxyl group (**17**) can destabilize both ligands, which explains a weaker inhibition potency for such compounds.

Table 4: Evaluation of identified inhibitors that showing $IC_{50} \leq 300 \mu M$

Inhibitor	IC_{50} (μM)	activity ratio	inhibition type	HA	LE	LE ratio
2	138	NA	competitive	21	0.25	NA
3	200	NA	competitive	26	0.19	NA
4	300	NA	ND	32	0.15	NA
13	165	0.83	ND	24	0.22	0.88
14	100	1.38	competitive	23	0.24	0.96
15	100	1.38	Mixed	23	0.24	0.96
18	50	2.76	competitive	26	0.23	0.92
19	62	2.22	ND	22	0.24	0.96
20	260	0.53	ND	27	0.18	0.72
26	40	3.45	ND	20	0.30	1.2
27	270	0.51	ND	21	0.23	0.92

NA: not applicable; ND: not determined; HA: heavy atoms; LE: ligand efficiency; Activity ratio = IC_{50} of compound **2** / IC_{50} of the analogue compound; LE ratio = LE of the analogue / LE of compound **2**.

3. Conclusions

Using a stepwise and rational structure-based 3D-pharmacophore approach, nine small molecules from a library of about one million commercial compounds were selected for biological testing. Three out of eight tested compounds revealed promising inhibitory potency ($IC_{50} \leq 300 \mu\text{M}$) and ligand efficiency ($LE \geq 0.15$) for further optimization. Using molecular docking, binding modes of the most active compounds were elucidated. It was shown that suitable orientations in the active site and occupancy of the productive subsites (-1, +1) were key parameters for an optimal inhibition.

Compounds **2** and **3** represent the most promising ligands identified in the pharmacophore-based virtual screening and are among the first small drug-like molecules reported as α -amylase inhibitors so far. While this project was in development, two other groups reported potent small molecules for α -amylase inhibition⁴⁷⁻⁴⁹. Interestingly, one of these compounds, a phloroglucinol derivative DDBT,⁴⁷ is a natural product with a very analogous scaffold to compound **3**. This similarity confirms the rational of our strategy and the strength of our computer-aided methodology.

In a recent perspective study, Zhu et al.⁵⁰ report that a LE value of 0.3 for a drug-like compound (molecular weight < 500 or including 35 heavy atoms maximum) corresponds to about 10 nM activity, which is practically unrealistic for initial hit identification. Similarly, the authors state that LE values of 0.32, 0.25, and 0.19 are recommended for compounds with $HA \leq 18$, $HA = 19 - 25$, and $HA = 26 - 35$, respectively. As stated in

Table 4, all inhibitors identified in this work display LE values higher than the one of the bound acarbose and, in agreement with LE values suggested by Zhu et al. in 2013.⁵⁰ Inhibitors **2** ($HA = 21$), **3** ($HA = 26$), and **26** ($HA = 20$) have LE values of 0.25, 0.19 and, 0.30 and hence are interesting for further chemical optimization.

In an evaluation of the quality of the analogues of compound **2**, LE ratios and activity improvement were compared (Table 4). Only compound **26** display a LE ratio with a value of 1.2-fold better than the primary lead query **2**. Other ligands showed improved potencies with IC_{50} values 100 μM , 100 μM , 50 μM , 62 μM , and 40 μM for **14**, **15**, **18**, **19** and **26** respectively, while preserving their efficiency (LE ratio around 1). Thus, searching for analogues of **2** led to the identification of compounds with good activity ratio (

Table 4). Among them, compound **26** display a 3-fold activity improvement compared to **2**.

Thus, using the available structural information, a pharmacophore-based virtual screening of commercial compounds was carried out, leading to the identification of small molecules and fragments with drug-like properties. Then, a second optimization step was conducted by searching for analogues of the best inhibitor, from which the SARs were analyzed with the support of molecular docking studies. All inhibitors discovered in this work are reported here for the first time as α -amylase inhibitors. Consequently, the structural simplicity of the identified chemical structures and their low molecular weight is an important step in the design of new potent α -amylase inhibitors for the treatment of chronic metabolic diseases such as type II diabetes and obesity.

4. Materials and methods

4.1 Pharmacophore creation and validation

Crystal structures (3OLD, 3OLE, 3OLG, 3OLI, 3IJ9, 3IJ7, 2QV4, 3BAJ, 3BAY, 1XH2, 1XD0, 1XD1, 1XH0, 1XCX, 1XCW, 1CPU, 1MFV, and 1OSE) of the HPA, HSA and porcine pancreatic α -amylases (PPA) co-crystallized with sugar-based inhibitors deposited in PDB^{29, 30} were investigated using the platform LigandScout 3.1.^{35, 51, 52} Binding modes of co-crystallized inhibitors were analyzed in detail and their chemical interactions were translated into 3D pharmacophore models. Then, the resulting pharmacophores were aligned and used as a basis to develop a unique query for virtual screening. Pharmacophore models created and developed were validated using 19 biologically known active and 55 inactive α -amylase inhibitors collected from ChEMBL database⁵³ and literature (Supplementary data, Tables S1 and S2). All these collected molecules were built and minimized using MOE 2010.10.⁵⁴ 3D coordinates were generated using CORINA 3.4.⁵⁵ The command-line tool idbgen was used to generate database files to be used as inputs for screening in LigandScout 3.1.³⁵ The software Omega⁵⁶ was used to generate 25 conformations per compound. Based on the best early enrichment factor (EF_{1%}) and AUC_{100%}, one pharmacophore was selected for virtual screening of commercial chemical compounds. Commercial libraries were obtained from Analyticon (30,352),^{57, 58} Asinex (465,543),⁵⁹⁻⁶¹ Life chemicals (372,071),⁶² Chembridge (646,018),⁶³ Specs (204,362),⁶⁴ Bionet (42,660),⁶⁵ and Prestwick (1,183).⁶⁶ All libraries were computationally processed using Standardizer 5.8.0⁶⁷ to correct protonation states. Subsequently, the command-line program idbgen was used to convert libraries into ligand-database format (ldb) to be used as inputs for virtual screening in LigandScout 3.1. The software Omega,⁵⁶ with FAST parameters was used to generate up to 25 conformations per molecule. For shape- and feature-based similarity search by ROCS 3.1.2,⁴⁵ commercial libraries were transformed into 3D coordinates using CORINA 3.4⁵⁵ and then the software OMEGA 2.4.6⁶⁸ was used to create 25 conformations per molecule as oeb.gz format as inputs for ROCS.⁴⁵

4.2 Molecular docking

All docking studies were performed using PDB entry 3OLE (resolution 1.55 Å)⁷ using GOLD 5.1.^{36, 46} Validation of docking was performed by reproducing the binding mode of the co-crystallized inhibitor, acarviostatin II03, with RMSD between the heavy atoms of the original co-crystallized ligand and the docked conformation ligand of 0.8 Å (Supplementary data, Fig. S2). Docking volume was defined by selecting all residues within 10 Å around the co-crystallized inhibitor. Default docking parameters were used. All docking experiments were carried out using GoldScore⁶⁹ as scoring function with 100 % search efficiency. 100 Docking poses were generated for each compound. All conformations were imported into LigandScout 3.1 and minimized with MMFF94 force field³⁷ before analysis. 3D Pharmacophores were created in a volume within 7 Å sphere around the co-crystallized ligand. Visualization, analysis and illustration were done using LigandScout 3.1. JKlustor 5.8.0³⁸ was used to perform structural similarity clustering for compounds obtained from virtual screening.

4.3 Biological testing

α -Amylase (EC 3.2.1.1) from human saliva (Type IXA) was purchased from SIGMA Aldrich (Steinheim, Germany). It gave a single band on sodiumdodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and possessed no α - and β -glycosidase activity. The substrate 2-chloro-4-nitrophenyl-4-O- β -D-galactopyranosyl-maltoside (GalG₂CNP) was purchased from Sorachim, Switzerland. Liberation of CNP from GalG₂CNP was monitored continuously with a UV-VIS spectrophotometer (JASCO V550) under temperature-controlled conditions (37°C) at 400 nm. Kinetic experiments were carried out based on previously published protocol,³⁹ at 37°C in 50 mM MES buffer pH 6.0 containing 5 mM Ca(OAc)₂, 51.5 mM NaCl and 152 mM NaN₃. Inhibitor compounds were dissolved in DMSO and dilutions for the latter tests were made in the same solvent. The substrate (0.75 - 4 mM) and inhibitor were mixed together and the reaction was initiated by adding HSA (2 nM) to the incubation medium (total volume of 500 μ l). Measurements without inhibitor were carried out at a final concentration of 2.5 % DMSO. The increase of absorbance of CNP liberated by HSA was measured continuously at 400 nm using the Parallel Kinetics Analysis program of a JASCO V550 spectrophotometer. Concentration-response plots were used to determine the effects of the inhibitor on the enzymatic reaction and IC₅₀ values of inhibitors. These experiments were performed at constant enzyme and substrate concentrations. Fractional activity (Y axis) was plotted as a function of inhibitor concentration (X axis). The data were fit using a standard four-parameter logistic nonlinear regression analysis of Grafit software.⁷⁰ The type of inhibition was determined by Lineweaver-Burk plot (Supplementary data, Fig. 5 and Fig. 6). All experiments were repeated three times and the obtained IC₅₀ did not vary more than 5 %.

Acknowledgments

Authors would like to thank Dr. Bernd Rupp from the Leibnitz Institute of Pharmacology, Berlin for virtually screening WDI and MDDR. The position of JM is funded by the *Deutsche Forschung Gemeinschaft* (WO 1955/1-1) and, the position of J.A., by *Erasmus Mundus ECW program*.

References

1. Low, L.C., *Pediatr. Diabetes* **2010**, *11*, 212.
2. McCue, P.; Kwon, Y.I.; Shetty, K., *Asia Pac. J. Clin. Nutr.* **2005**, *14*, 145.
3. Subramanian, R.; Asmawi, M.Z.; Sadikun, A., *Acta Biochim. Pol.* **2008**, *55*, 391.
4. Gupta, R.G., P.; Mohapatra, H.; Goswami, V.K.; Chauhan, B., *Process Biochem.* **2003**, *38*, 1599.
5. Muralikrishnaa, G., Nirmalab, M., *Carbohydr. Polym.* **2005**, *60*, 163.
6. Brayer, G.D.; Sidhu, G.; Maurus, R.; Rydberg, E.H.; Braun, C.; Wang, Y.; Nguyen, N.T.; Overall, C.M.; Withers, S.G., *Biochemistry* **2000**, *39*, 4778.
7. Qin, X.; Ren, L.; Yang, X.; Bai, F.; Wang, L.; Geng, P.; Bai, G.; Shen, Y., *J. Struct. Biol.* **2011**, *174*, 196.
8. Butterworth, P.J.; Warren, F.J.; Ellis, P.R., *Starch-Starke* **2011**, *63*, 395.
9. Groot, P.C.; Bleeker, M.J.; Pronk, J.C.; Arwert, F.; Mager, W.H.; Planta, R.J.; Eriksson, A.W.; Frants, R.R., *Genomics* **1989**, *5*, 29.
10. Nikitkova, A.E.; Haase, E.M.; Scannapieco, F.A., *Appl. Environ. Microbiol.* **2013**, *79*, 416.
11. Lehmann, U.; Robin, F., *Trends Food Sci. Technol.* **2007**, *18*, 346.
12. Li, C.; Begum, A.; Numao, S.; Park, K.H.; Withers, S.G.; Brayer, G.D., *Biochemistry* **2005**, *44*, 3347.
13. Kajaria, D.; Ranjana; Tripathi, J.; Tripathi, Y.B.; Tiwari, S., *J. Adv. Pharm. Technol. Res.* **2013**, *4*, 206.
14. Preuss, H.G.; Echard, B.; Bagchi, D.; Stohs, S., *Int. J. Med. Sci.* **2007**, *4*, 209.
15. Thilagam, E.; Parimaladevi, B.; Kumarappan, C.; Mandal, S.C., *J. Acupunct. Meridian Stud.* **2013**, *6*, 24.
16. Homoki, J.R.; Nemes, A.; Fazekas, E.; Gyémánt, G.; Balogh, P.; Gál, F.; Al-Asri, J.; Mortier, J.; Wolber, G.; Babinszky, L., *Food Chem.* **2016**, *194*, 222.
17. Svensson, B.; Fukuda, K.; Nielsen, P.K.; Bonsager, B.C., *Biochim. Biophys. Acta* **2004**, *1696*, 145.
18. Heyl, D.L.; Tobwala, S.; Lucas, L.S.; Nandanie, A.D.; Himm, R.W.; Kappler, H.J.; Blaney, E.J.; Groom, J.; Asbill, J.; Nzoma, J.K.; Jarosz, C.; Palamma, H.; Schullery, S.E., *Protein Pept. Lett.* **2005**, *12*, 275.
19. Porcelli, F.; Olivieri, C.; Masterson, L.R.; Wang, Y.; Veglia, G., *J. Biol. Inorg. Chem.* **2011**, *16*, 1197.
20. Krentz, A.J.; Bailey, C.J., *Drugs* **2005**, *65*, 385.
21. Uchida, R.; Nasu, A.; Tokutake, S.; Kasai, K.; Tobe, K.; Yamaji, N., *Chem. Pharm. Bull. (Tokyo)* **1999**, *47*, 187.
22. Lo Piparo, E.; Scheib, H.; Frei, N.; Williamson, G.; Grigorov, M.; Chou, C.J., *J. Med. Chem.* **2008**, *51*, 3555.
23. Brayer, G.D.; Luo, Y.; Withers, S.G., *Protein Sci.* **1995**, *4*, 1730.
24. Ramasubbu, N.; Paloth, V.; Luo, Y.; Brayer, G.D.; Levine, M.J., *Acta Crystallogr., Sct. D: Biol. Crystallogr.* **1996**, *52*, 435.
25. Nahoum, V.; Roux, G.; Anton, V.; Rouge, P.; Puigserver, A.; Bischoff, H.; Henrissat, B.; Payan, F., *Biochem. J.* **2000**, *346*, 201.
26. Davies, G.J.; Wilson, K.S.; Henrissat, B., *Biochem. J.* **1997**, *321*, 557.
27. Motyan, J.A.; Gyemant, G.; Harangi, J.; Bagossi, P., *Carbohydr. Res.* **2011**, *346*, 410.
28. Mortier, J.; Rakers, C.; Frederick, R.; Wolber, G., *Curr Top Med Chem* **2012**, *12*, 1935.
29. Berman, H.; Henrick, K.; Nakamura, H., *Nat. Struct. Biol.* **2003**, *10*, 980.
30. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E., *Nucleic Acids Res.* **2000**, *28*, 235.
31. Machius, M.; Vertesy, L.; Huber, R.; Wiegand, G., *J. Mol. Biol.* **1996**, *260*, 409.
32. Gaulton, A.; Bellis, L.J.; Bento, A.P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J.P., *Nucleic Acids Res.* **2012**, *40*, D1100.
33. Triballeau, N.; Acher, F.; Brabet, I.; Pin, J.P.; Bertrand, H.O., *J. Med. Chem.* **2005**, *48*, 2534.

34. Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J., *Nucleic Acids Res.* **2006**, *34*, D668.
35. *LigandScout 3.1*, available from *Inteligand GmbH, Mariahilferstrasse 74B, 1070 Vienna, Austria*, www.inteligand.com.
36. *GOLD 5.1*, available from *Cambridge Crystallographic Data Centre, Cambridge, UK*, www.ccdc.cam.ac.uk.
37. Halgren, T.A., *J. Comput. Chem.* **1999**, *20*, 730.
38. *JKlustor 5.8.0*, available from *ChemAxon, Budapest, Hungary*, www.chemaxon.com.
39. Kandra, L.; Zajacz, A.; Remenyik, J.; Gyemant, G., *Biochem. Biophys. Res. Commun.* **2005**, *334*, 824.
40. Hopkins, A.L.; Keseru, G.M.; Leeson, P.D.; Rees, D.C.; Reynolds, C.H., *Nat. Rev. Drug Discov.* **2014**, *13*, 105.
41. Shultz, M.D., *Bio. Med. Chem. Lett.* **2013**, *23*, 5980.
42. Schultes, S.; De Graaf, C.; Haaksma, E.E.J.; De Esch, I.J.P.; Leurs, R.; Kramer, O., *Drug Discov. Today Technol.* **2010**, *7*, 157.
43. Ramasubbu, N.; Ragunath, C.; Mishra, P.J., *J. Mol. Biol.* **2003**, *325*, 1061.
44. Heng, S.; Gryncel, K.R.; Kantrowitz, E.R., *Bioorg. Med. Chem.* **2009**, *17*, 3916.
45. *ROCS 3.1.2*, available from *OpenEye Scientific Software: Santa Fe, New Mexico, USA*, www.eyesopen.com.
46. Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R., *J. Mol. Biol.* **1997**, *267*, 727.
47. Kawamura-Konishi, Y.; Watanabe, N.; Saito, M.; Nakajima, N.; Sakaki, T.; Katayama, T.; Enomoto, T., *J. Agric. Food Chem.* **2012**, *60*, 5565.
48. Patil, V.S.; Nandre, K.P.; Ghosh, S.; Rao, V.J.; Chopade, B.A.; Bhosale, S.V., *Bioorg. Med. Chem. Lett.* **2012**, *22*, 7011.
49. Patil, V.S.; Nandre, K.P.; Ghosh, S.; Rao, V.J.; Chopade, B.A.; Sridhar, B.; Bhosale, S.V., *Eur. J. Med. Chem.* **2013**, *59*, 304.
50. Zhu, T.; Cao, S.Y.; Su, P.C.; Patel, R.; Shah, D.; Chokshi, H.B.; Szukala, R.; Johnson, M.E.; Hevener, K.E., *J. Med. Chem.* **2013**, *56*, 6560.
51. Wolber, G.; Langer, T., *J. Chem. Inf. Model.* **2005**, *45*, 160.
52. Wolber, G.; Dornhofer, A.A.; Langer, T., *J. Comput. Aided Mol. Des.* **2006**, *20*, 773.
53. Gaulton, A.; Bellis, L.J.; Bento, A.P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J.P., *Nucleic Acids Res* **2012**, *40*.
54. *MOE 2010.10*, available from *Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite # 910, Montreal, QC, Canada*, www.chemcomp.com.
55. *CORINA 3.4*, available from *Molecular Networks GmbH, Erlangen, Germany*, www.molecular-networks.com.
56. Kirchmair, J.; Wolber, G.; Laggner, C.; Langer, T., *J. Chem. Inf. Model.* **2006**, *46*, 1848.
57. *MEGx: Natural Products Screening Compounds*, *Analyticon Discovery GmbH, Potsdam, Germany*, www.ac-discovery.com, downloaded in Sept. 15, 2011.
58. *NATx: Semi-Synthetic Screening Compound Library*, *Analyticon Discovery GmbH, Potsdam, Germany*, www.ac-discovery.com, downloaded in Sept. 15, 2011.
59. *Asinex Synergy Library*, *Asinex Ltd., Moscow, Russia*, www.asinex.com, downloaded in Sept. 15, 2011.
60. *Asinex Platinum Collection*, *Asinex Ltd., Moscow, Russia*, www.asinex.com, downloaded in Sept. 15, 2011.
61. *Asinex Gold Collection*, *Asinex Ltd., Moscow, Russia*, www.asinex.com, downloaded in Sept. 15, 2011.
62. *Life Chemical Compound Collection for HTS*, *Life Chemical Inc., Niagara-on-the-Lake, Ontario, Canada*, www.lifechemicals.com, downloaded in Sept. 15, 2011.
63. *Chembridge Screening Library*, *Chembridge Corporation, San Diego, CA, USA*, www.chembridge.com, downloaded in Sept. 15, 2011.
64. *Specs Screening Compounds*, *Specs, Delft, The Netherlands*, www.specs.net, downloaded in Sept. 15, 2011.
65. *Bionet Screening Compounds*, *Key Organics Ltd.: Camelford, UK*, www.keyorganics.net, downloaded in Sept. 15, 2011.

66. *The Prestwick Chemical Library, Prestwick Chemical: Illkirch, France, www.prestwickchemical.com, downloaded in Sept. 15, 2011.*
67. *Standardizer 5.8.0, available from ChemAxon, Budapest, Hungary, www.chemaxon.com.*
68. *OMEGA 2.4.6, available from OpenEye Scientific Software, Santa Fe, New Mexico, USA, www.eyesopen.com.*
69. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D., *Proteins* **2003**, 52, 609.
70. *GraFit 2.1, available from Erithacus Software Ltd., Horley, U.K, www.erithacus.com.*