

Scientific paper

Design of a Focused Virtual Library to Explore Cholera Toxin B-site[†]

Črtomir Podlipnik^{a,b,*} and Anna Bernardi^b^a Faculty of Chemistry and Chemical Technology, University Ljubljana, Aškerčeva 6, 1000 Ljubljana, Slovenia^b Dipartimento di Chimica Organica e Industriale, Università degli Studi di Milano, Via Venezian 21, I-20131 Milano, Italy

* Corresponding author: E-mail: crtomir.podlipnik@siol.net

Received: 08-05-2007

[†]Dedicated to Prof. Dr. Jože Škerjanc on the occasion of his 70th birthday

Abstract

A virtual library of cholera toxin (CT) inhibitors has been designed. The library consists of small molecules designed to act as decoys for the toxin's GM1 binding site and thus to prevent binding of CT to the cell membranes of intestinal epithelial cells. Structures of known inhibitors have been taken from different sources and in addition new inhibitors have been developed using structure based molecular design. The information that we report here may help in further design of more effective and metabolically stable Cholera Toxin B-site inhibitor.

Keywords: Cholera toxin, GM1 ganglioside, C-galactosides, virtual screening, molecular docking

1. Introduction

Cholera is a life threatening disease that is caused by *Vibrio cholerae* Cholera toxin (CT). The WHO each year reports a number of pandemic cases of Cholera and other diarrheas especially in developing countries.¹ Cholera Toxin belongs to the AB₅ holotoxin family, which includes CT itself and the *Escherichia coli* heat-labile toxins (LTs) LT-I and LT-II, among others.² The structure and function of AB₅ toxins have been reviewed in detail in several occasions.^{3–6} AB₅ toxins have characteristic struc-

tures where a single catalytically active component, A, is fused with a nontoxic receptor-binding component, a pentamer of B subunits (Figure 1A). The B pentamer is responsible for binding of CT to the GM1 ganglioside on the external membrane of intestinal epithelial cells. This binding is essential for initiation of the threatening action of CT. The function of recognition is retained even in absence of A subunit.^{7,8} The interaction of the oligosaccharide head groups of ganglioside GM1 (Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-4Glcβ1-OH, o-GM1) with the B₅ pentamer of CT is shown in figure 1B.

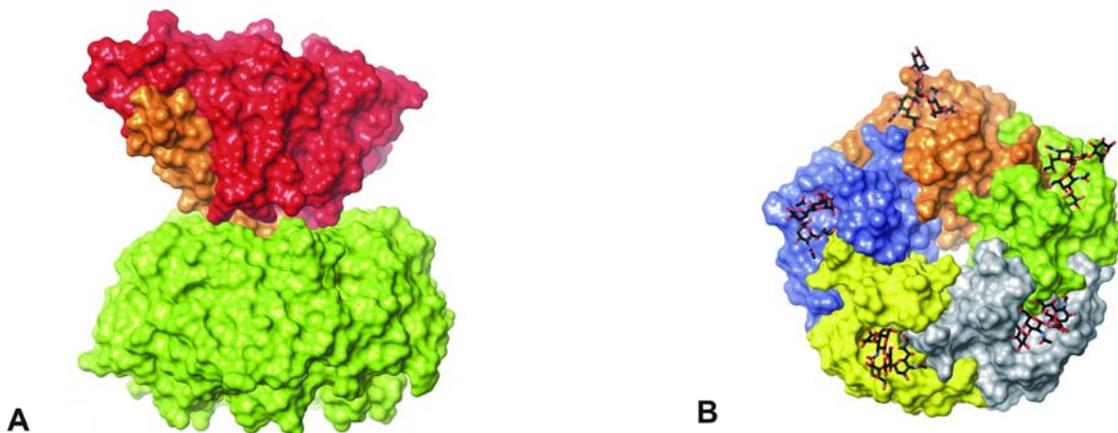


Figure 1: Structure of Cholera Holotoxin (A). Interaction of o-GM1 with Cholera Toxin B-site (B).

The binding event is followed by nicking of A chain and disulphide bond reduction, which yield the two fragments A1 (23.5 K Daltons) and A2 (5 K Daltons). The enzymatic A1 fragment is translocated across the membrane to the cytosol of the host cell, where it catalyses a transfer of an ADP from an NAD^+ to a component of the regulatory adenylyl cyclase mechanism. Permanent activation of adenylyl cyclase by CT results in rising levels of cAMP and the consequent activation of sodium pumps in the lumen of the cell through cAMP dependent kinase pathway, forcing Na^+ out. The electrochemical imbalance is then compensated by driving out Cl^- and H_2O , which results in the cholera symptoms, mainly an enormous loss of fluids, which may lead to death by dehydration.^{9,10}

The bidentate interaction of o-GM1 (compound 1) with CTB binding site is highlighted in figure 2. Binding of the terminal galactose is very specific. The pyranose ring stacks on top of TRP 88 (CH/ π interaction) and form a hydrogen bond network with ASN90, LYS91, GLU51 and GLN61. While the terminal galactose is shielded from the solvent, the rest of the toxin binding site is shallow and solvent exposed. The sialic acid moiety represents the second area of contact and it is placed less specifically than terminal galactose. The sugar ring of sialic acid makes hydrophobic interaction with TYR12, the hydroxyl, and N-acetyl substituents form hydrogen bonds with the protein backbone and the carboxy group interacts with TRP88 through a water molecule.¹¹ In terms of buried protein surface, the terminal Gal and Neu5Ac residues contribute more than 80% of intermolecular contacts.^{3,12}

Given the mechanism of action of CT three strategies are possible to design a prophylactic cure against Cholera:

1. Inhibition of the action of the catalytically active unit A of CT.^{13,14}
2. Prevention of assembly of the AB_5 complex.¹⁵

3. Design of small molecules acting as decoys for the toxin's GM1 binding site and thus preventing binding of CT to the cell membranes of intestinal epithelial cells.^{7,16–24}

In this work we created a virtual library of inhibitors of the CTB site and screened it for CTB binding. Some of these compounds are chemical structures collected from different sources, all containing a galactose residue and known to interact with CTB binding site. In addition we used a C-galactose anchor to (virtually) build a second group of new compounds that potentially bind to the receptor site of CTB. The first group of compounds was used for validation of the screening protocol used to rank the second group. As a control, the NCI diversity library²⁵ was also docked in CTB binding site using the same protocol and the results were compared to those obtained using the focused library of C-galactosides.^{23,26,27}

2. Methods

2.1. Preparation of Protein

The X-ray of the complex between Cholera toxin B-pentamer and o-GM1 (PDB_ID:3CHB) has been used as initial structure in the preparation of CTB receptor site.⁵ This crystal structure represents a B_5 pentamer complexed with five pentasaccharides o-GM1. In the structure each o-GM1 is bound to a receptor site formed by a pair of neighbouring B subunits. So only two B subunits out of five are necessary to model the binding site. Units G and H were selected for our model, because they showed no deficiencies in the X-ray structure. Thus, after manual inspection and cleaning of structure we retained a complex composed of protein chains G and H, o-GM1 ligand and the two water molecules (residue numbers in PDB: 7302 and 7303) that bridge between ligand and protein, all other water molecules were suppressed. Hydrogens were

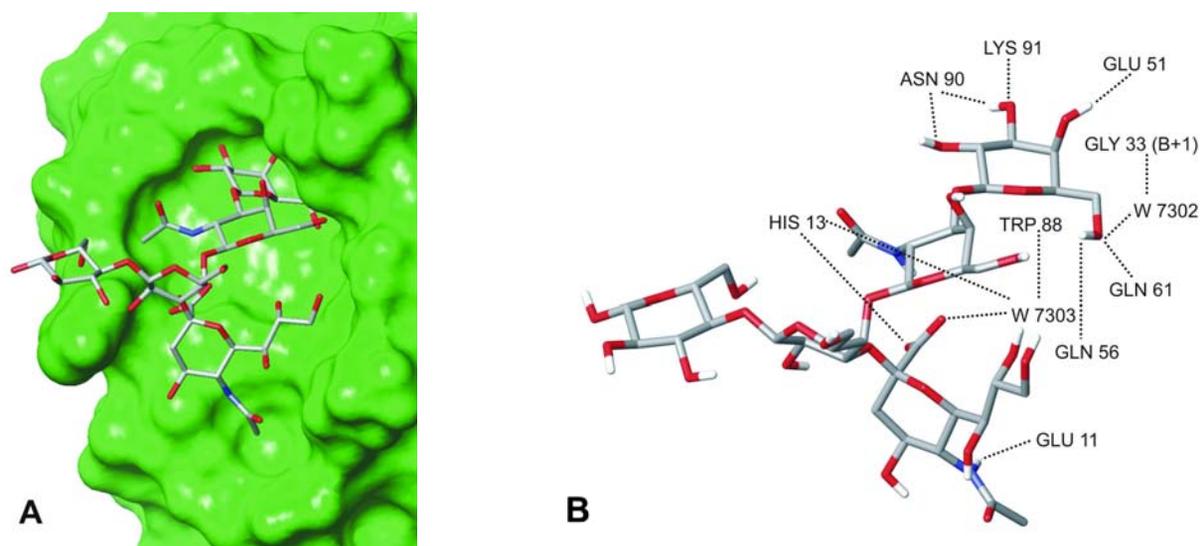


Figure 2: Close view of CTB:o-GM1 interaction (A). Interaction map of CTB:o-GM1 complex (B).

added to the model automatically via the Maestro interface,²⁸ leaving no lone pair and using an explicit all-atom model. In the next phase the two water molecules were temporarily removed from the complex. The multi step Schrodinger's Protein preparation tool (Pprep) has been used for final preparation of receptor model. Pprep neutralizes side chains that are not close to binding cavity and do not participate in salt bridges.²⁸ This step is then followed by restrained minimization of co-crystallized complex, which reorients side chain hydroxyl groups and alleviates potential steric clashes. The two water molecules were then merged to the receptor model. The complex obtained was minimized using the Amber* force field with the Still-Senderowitz parameters for saccharides²⁹ and the Polack-Ribiere Conjugate Gradient (PRCG) algorithm.³⁰ The minimization was stopped either after 5000 steps or after the energy gradient converged below 0.05 kJ mol^{-1} . We defined several shells that allow the ligand to relax inside the receptor site. The first shell consists of the ligand and all polar hydrogens within 5 \AA of the ligand.

2. 2. Virtual Library Design

The virtual library of CT inhibitors contains 264 compounds divided in 8 sublibraries (table 1). Some of the compounds are obtained by conjugation of scaffolds **15** and **16** with a group of selected amines or acids, res-

Table 1: The composition of virtual library of CT inhibitors.

Sublib. ID	Sublibrary content	Compounds
I	pseudo GM1 ligands	10
II	Aminoacids, natural and unnatural	55
III	β - Aminoacids	54
IV	Cinnamic, Maleic and Succinic	32
V	Aspartam	4
VI	Verlinde hydrophobics library	75
VII	m-nitro phenyl galactosides	34
TOTAL		264

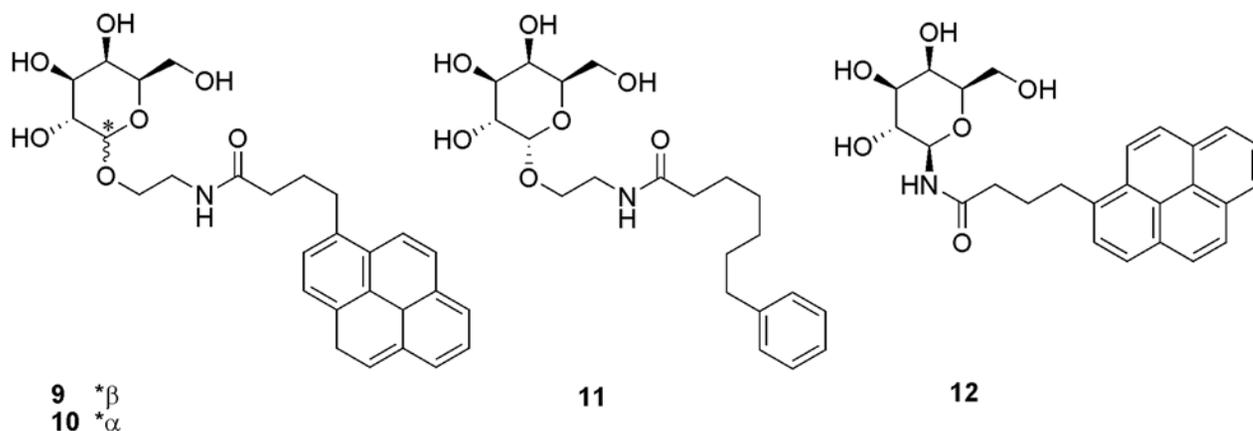


Figure 4: Examples from Galactose library for exploring hydrophobic pocket in receptor binding site of CT.

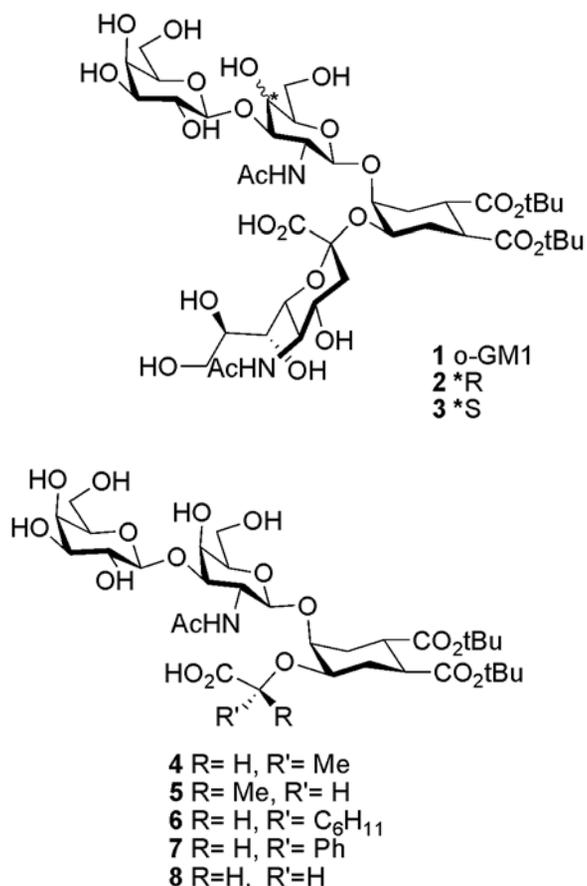


Figure 3: Pseudo GM1 ligands.

pectively. Other compounds are taken from various sources among these are pseudo GM1 ligands, meta nitrophenyl galactoconjugates,²² and compounds from galactose library for exploration of hydrophobic pocket²⁴ in the CTB receptor binding site. The structures of pseudo GM1 have been taken from the material described in studies by Bernardi et al.^{17–20,31–33}

To the first sublibrary (*Sublib-I*) belong the so-called pseudo GM1 mimics (**2–8**). These molecules were rationally designed as functional mimics of α -GM1 with the goal of simplifying the chemical synthesis and improving the metabolic stability (Figure 3).^{16–18,20,33,34}

Sublib-VI contains compounds designed by Verlinde *et al* to explore a hydrophobic pocket in the receptor site of the CT.²⁴ Some examples from this library (compounds **9–12**) are presented in Figure 4.

The last set of known ligands included in the virtual library (*Sublib-VII*) are derivatives of *m*-nitro phenyl galactoside

(MNPG), that have been described as interesting CTB inhibitors operating through a peculiar binding mode, which involves displacement of one of the crystallographic water molecules.²² MNPG **13** and one of its derivatives **14** are presented in Figure 5.

The virtual library of C-galactosides (*Sublib-II-V*) was built from scaffolds **15** and **16** (Figure 6) by conjugation with amines or acids, respectively, through the formation of an amide bond. Some of the elements from the cinnamic acid sublibrary (*Sublib-IV*) **17–23** were selected for synthesis and affinity evaluation (by SPR) (Figure 7), as described elsewhere.²³

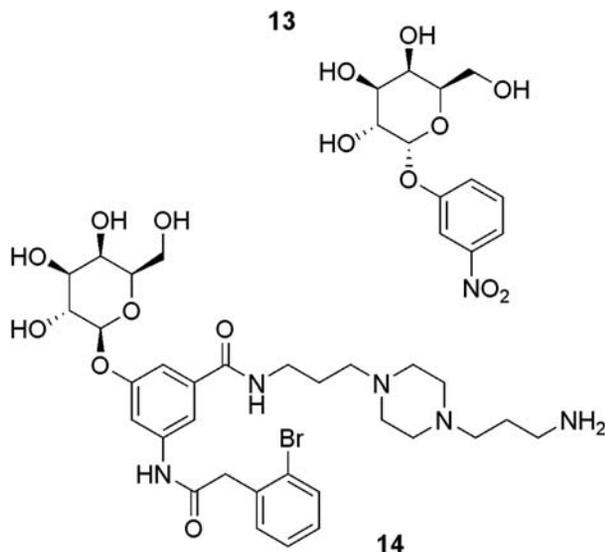


Figure 5: Structure of *m*-nitro phenyl galactoside **13** and one of its derivatives **14**.

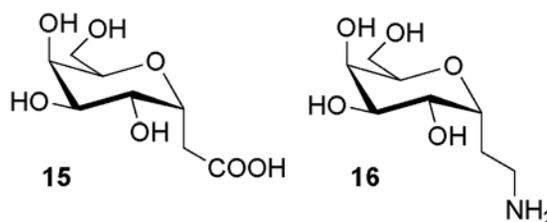


Figure 6: The C-galactoside anchors **15** and **16** used for the development of metabolically stable CT-ligands.

We used ChemAxon's Marvin software for sketching structures and converting them to their 3D representation.³⁵ LigPrep²⁸ was used for final preparation of ligands from both libraries for docking. LigPrep is an utility of the Schrodinger software suit that combines tools for generating 3D structures from 1D (Smiles) and 2D (SDF) representation, searching for tautomers and steric isomers, and performing a geometry minimisation of ligands.

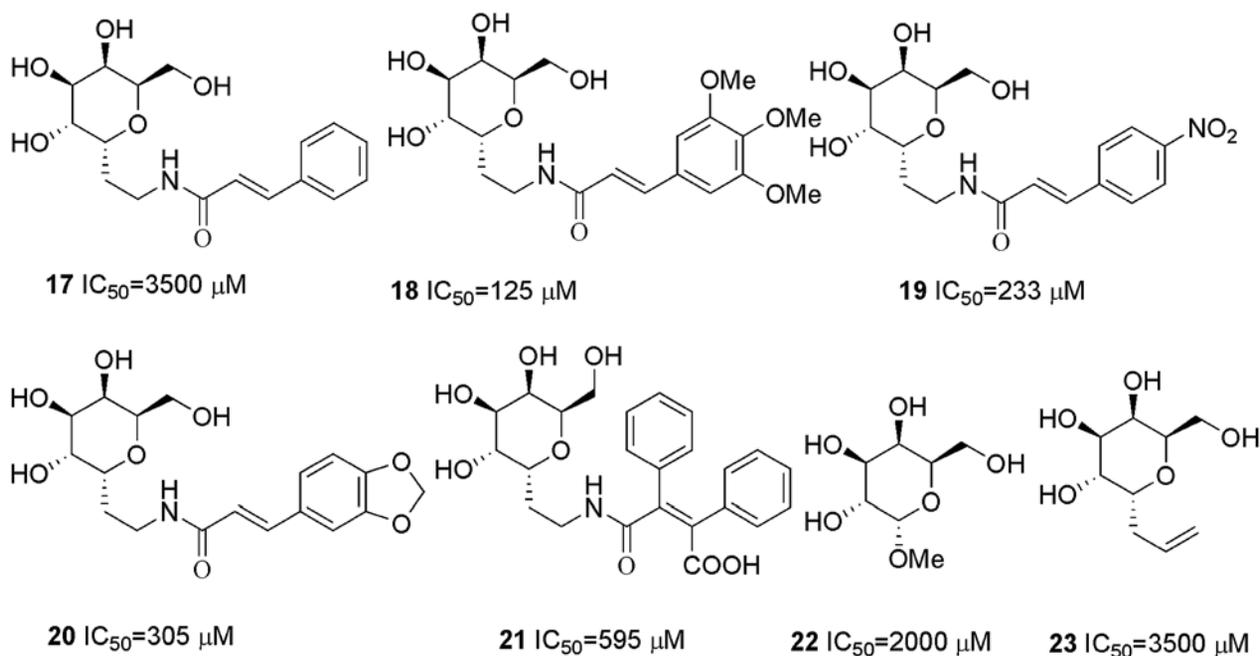


Figure 7: Structure and CT inhibition values of ligands **17–23** (from ref 23).

2. 3. Docking Procedure

The Schrodinger Glide program ver. 4.0 has been used for docking.^{36,37} The best 10 poses and corresponding scores have been evaluated using Glide in single precision mode (Glide SP) for each ligand from the virtual library of CT inhibitors and from NCI diversity library. For each screened ligand, the pose with the lowest Glide SP score have been taken as the input for the Glide calculation in Extra Precision mode (Glide XP).

2. 4. Rescoring using Prime/MM-GBSA Approach

For each ligand, the pose with the lowest Glide score SP (XP) was rescored using Prime/MM-GBSA approach.³⁸ This approach is used to predict the free energy of binding for set of ligands to receptor. The docked poses were minimized using the local optimization feature in Prime, and the energies of complex were calculated using the OPLS-AA force field and GBSA continuum solvent model. The binding free energy ΔG_{bind} is then estimated using equation:

$$\Delta G_{\text{bind}} = E_{\text{R:L}} - (E_{\text{R}} + E_{\text{L}}) + \Delta G_{\text{solv}} + \Delta G_{\text{SA}}, \quad (1)$$

where $E_{\text{R:L}}$ is energy of the complex, $E_{\text{R}} + E_{\text{L}}$ is sum of the energies of the ligand and unliganded receptor, using the OPLS-AA force field, ΔG_{solv} (ΔG_{SA}) is the difference between GBSA solvation energy (surface area energy) of complex and sum of the corresponding energies for the ligand and unliganded protein. Corrections for entropic changes were not applied in this type of free energy calculation.

3. Results and Discussion

3. 1. Virtual Screening of Library of CT Antagonists Using Glide SP and Glide XP

Glide 4.0 in two different modes are used for screening a virtual library of CT inhibitors. The Glide SP (XP) scores for the best 20 ligands are collected in table 2.

Table 2: The "best" 20 ligands ranked by Glide SP and Glide XP score. Structures of compounds 24–46 are shown in Figure 8.

Rank	Glide SP			Glide XP		
	Struct	G-Score	Sublib.	Struct	G-Score	Sublib.
1	2	-12.75	I	1	-12.21	I
2	1	-12.75	I	2	-11.87	I
3	3	-12.61	I	7	-11.61	I
4	24	-11.28	II	3	-11.52	I
5	6	-11.00	I	6	-11.33	I
6	7	-10.89	I	29	-10.74	IV
7	8	-10.55	I	41	-10.71	VI
8	5	-10.23	I	32	-10.66	IV
9	43	-10.18	VI	30	-10.56	IV
10	45	-10.13	I	42	-10.54	VI
11	4	-10.09	I	37	-10.52	VI
12	46	-10.02	V	10	-10.48	VI
13	39	-9.99	VI	36	-10.43	VI
14	41	-9.96	VI	34	-10.43	VI
15	28	-9.90	IV	38	-10.36	VI
16	26	-9.85	III	14	-10.34	VII
17	33	-9.85	VI	35	-10.29	VI
18	25	-9.84	II	33	-10.29	VI
19	44	-9.84	VI	31	-10.21	III
20	27	-9.82	III	4	-10.16	I

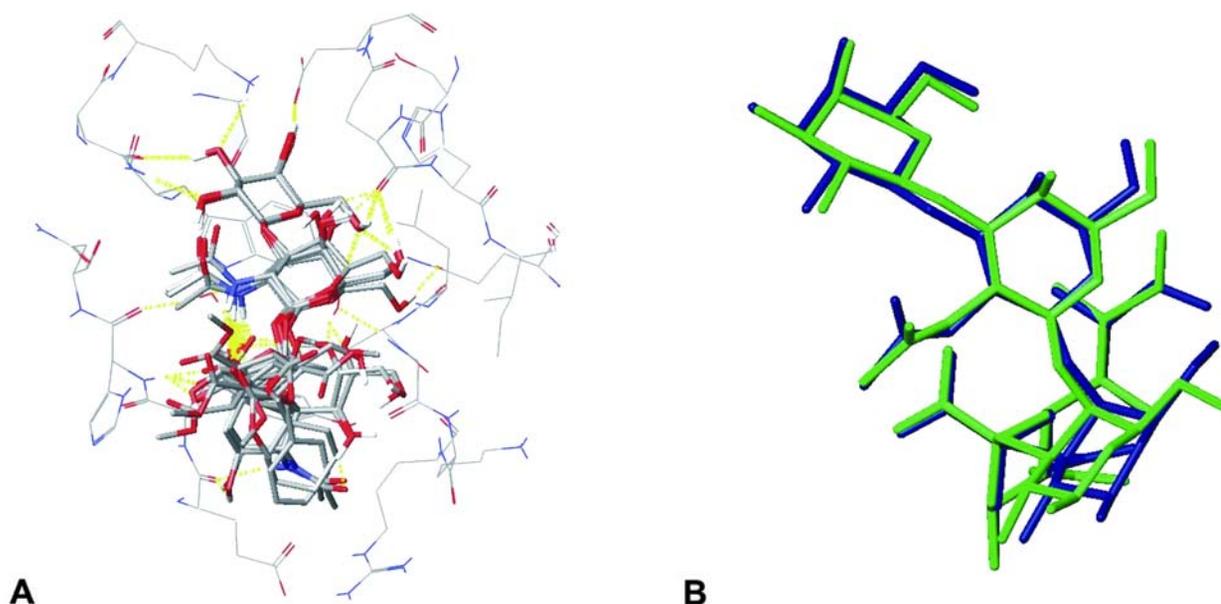


Figure 8: (A) Superposition of pseudo GM1 ligands (1,2,4,6,7) poses within binding site of CT. Poses are result of Glide XP docking. (B) Superposition of pose of o-GM1 obtained with Glide XP docking (blue) and Glide XP docking (green); RMSD (heavy atoms) = 0.77 Å.

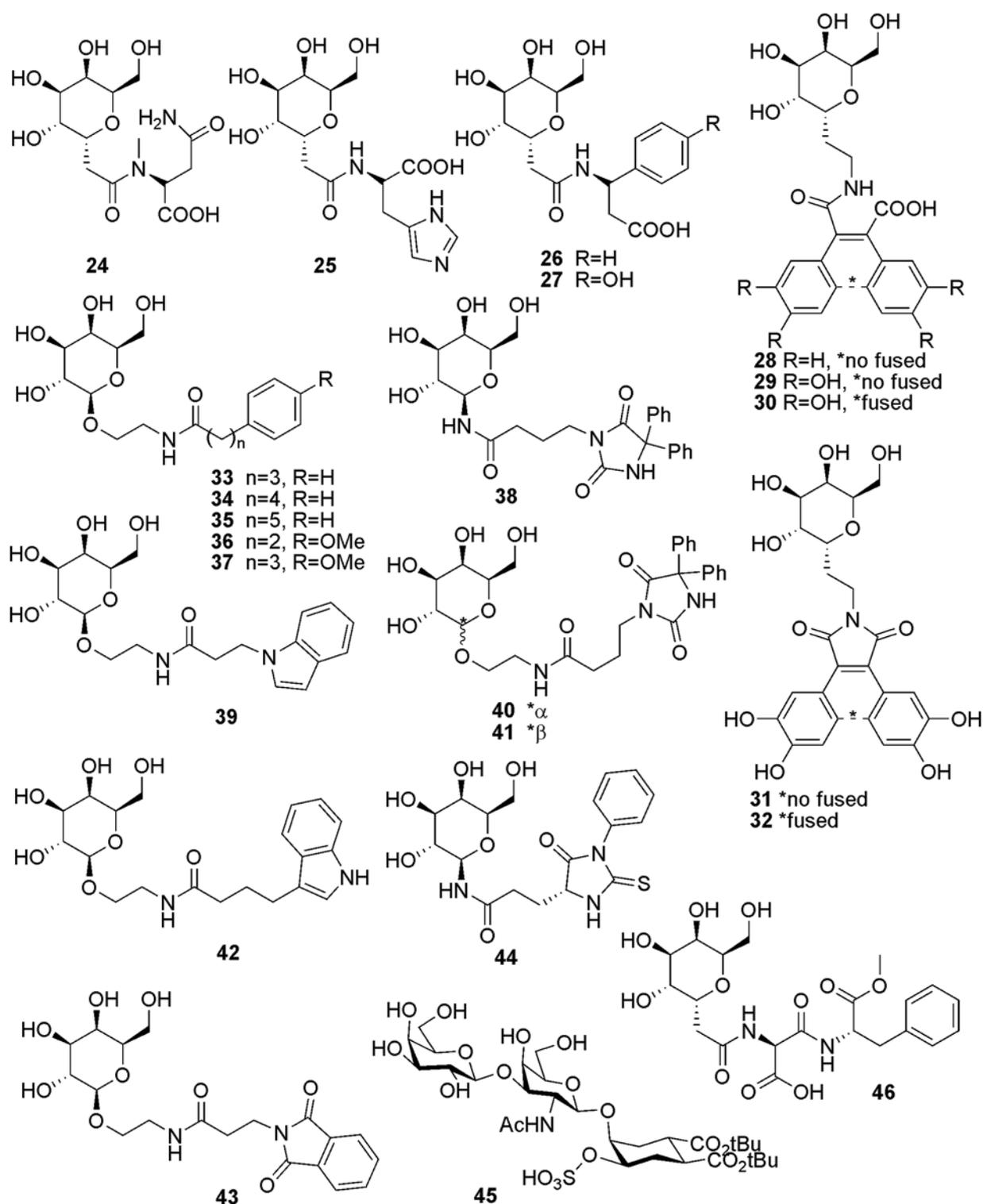


Figure 9: Selected ligands from a virtual library (See table 2).

Structures from table 2 that are not mentioned in previous chapters are shown in figure 9.

Pseudo GM1 ligands (*Sublib-I*) designed by stepwise rationalization of natural o-GM1 are most populated

among “best 20” ligands ranked by Glide SP (XP) score. The binding modes of five superimposed pseudo GM1 ligands within CTB receptor site are shown in figure 8A. In this figure we can observe that all ligands present the

Table 3: Glide score distribution in sublibraries of CT_LIB and NCI diversity database.

Glide SP	I	II	III	IV	V	VI	VII	CT_LIB*	NCI
G Score < -12.0	3	0	0	0	0	0	0	3	0
-11.0 > G Score ≥ -12.0	1	0	0	1	0	0	0	2	0
-10.0 > G Score ≥ -11.0	4	0	0	0	1	0	0	6	0
-9.0 > G Score ≥ -10.0	1	15	34	34	1	36	6	127	0
-8.0 > G Score ≥ -9.0	1	9	13	19	2	33	7	84	6
-7.0 > G Score ≥ -8.0	0	3	6	2	0	5	19	35	32
-6.0 > G Score ≥ -7.0	0	4	1	0	0	0	2	8	397
-5.0 > G Score ≥ -6.0	0	0	0	0	0	0	0	0	1149
G Score ≥ -5.0	0	1	0	0	0	0	0	1	1652

* CT_LIB – Library of CT inhibitors

terminal galactose anchored in a well defined binding pocket. Orientation of carboxylic acid moiety of terminal sialic acid is also important part of the interaction between pseudo GM1 ligands and CTB. All five ligands from figures 8A are in conformation that keep this carboxylic acid in a suitable position to form hydrogen bond interaction with HIS13 backbone. The quality of docking is often measured by ability of the software to reproduce binding mode of ligand that is found in crystal structure. In figure 8B we can see a good agreement between docked pose (Glide XP) and conformation of α -GM1 that is found in the crystal structure of the CT: α -GM1 complex.

In table 3 are collected Glide Score (SP) distributions of various sublibraries (I–VII) of CT inhibitors and of the NCI diversity database. We can observe that the most potent CT inhibitors were found among the pseudo GM1 ligands (*Sublib-I*). The majority (83%) of the ligands from sublibraries II–VII have Glide score between -10.0 and -8.0. On the contrary, there are less than 0.2% ligands from NCI diversity library that have Glide score lower than -8.0. So we may conclude that our virtual library is far more focused to CTB binding site than the NCI diversity library.

3. 2. Building a Model for Prediction of pIC50 Using Glide SP and Glide XP Score

We selected some compounds with known inhibitory activity (pIC50) from virtual library for further analysis and for building a model for prediction of pIC50. Activities (pIC₅₀) of pseudo GM1 ligands are measured with fluorescence titrations^{16,19,20}, SPR (Surface Plasmon Resonance) are used for determination activity of C-galactosides.²³ We included also three compounds from the galactose library designed by Verlinde²⁴ in our set of compound.

It has been seen from the results collected in Table 3 that the pseudo GM1 ligands have significantly better activities compared to the C-galactosides. On the other hand, the synthesis of C-galactosides is simpler than that of pseudo GM1 ligands, and in addition there is space for further optimisation of C-galactoside structures. Among

galactosides the best activity (SPR) was found for structure **18** (IC₅₀ = 125 μ M). The best Glide SP and XP binding modes for compound **18** are presented in figure 10. A similar mode was also observed by the Autodock/Glide procedure described in a previous study.²³

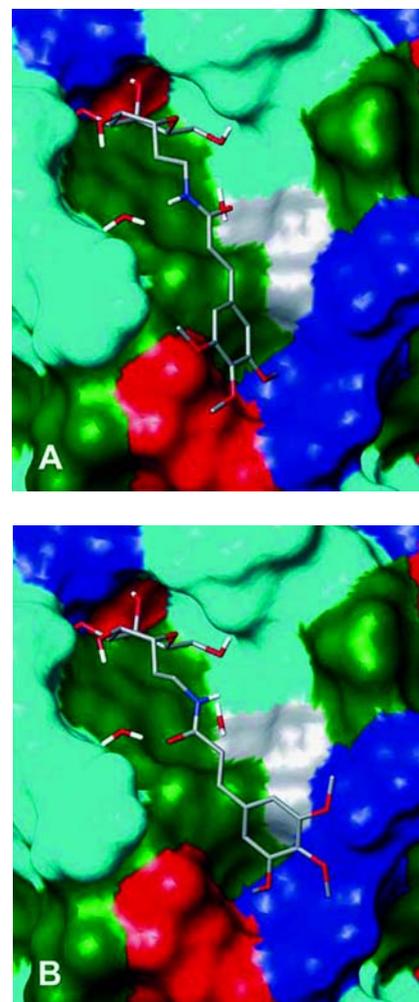


Figure 10: Two binding modes for compound **18** obtained with Glide SP (A) and Glide XP (B). Protein surface is colored by property of residue according color scheme: green – lipophilic; light blue – hydrophilic non charged; red – negatively charged; blue – positively charged; gray – GLY or water.

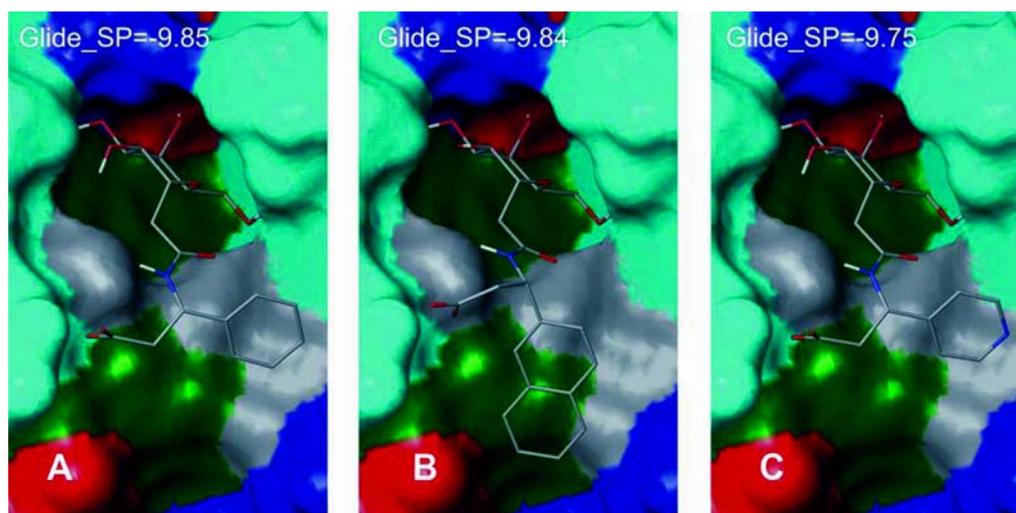


Figure 11: Binding poses of some β amino acid's galactoconjugates. Color scheme is the same as figure 10. Binding poses and scores are result of Glide SP docking protocol.

Another interesting group of C-galactosides that could be promising CTB binders are derivatives of scaffold **15** and β -amino acids. In such structures we can find all three elements that are decisive for effective binding (Galactose anchor, carboxylic group that might interact with HIS13, and hydrophobic pattern that are necessary for unspecific interactions in bottom part of CTB receptor site). The binding poses and scores of some β -amino C-galactoconjugates are presented in figure 11.

We used the data collected in the table 4 for building a models for prediction of pIC50 of the CT inhibitors. We built two linear regression models. In first model we used Glide Score (SP) as a descriptor. The equation (2) of model and the corresponding statistics is.) written below:

$$\text{pIC50} = -1.97 (\pm 1.11) - 0.59 (\pm 0.11) * \text{G-Score (SP)} \\ (r^2 = 0.67; F = 27.4; s^2 = 0.43; n = 15; q^2 = 0.59). \quad (2)$$

Equation 3 represents a model that uses Glide Score (XP) as a descriptor:

$$\text{pIC50} = -2.94 (\pm 1.33) - 0.66 (\pm 0.13) * \text{G-Score (XP)} \\ (r^2 = 0.68; F = 26.04; s^2 = 0.44; n = 15; q^2 = 0.47). \quad (3)$$

The relation between modelled (line) and experimental pIC50 (numbers) for two models described by equations 2 and 3 is shown in figure 12.

Reasonably good agreement between modelled and experimental data are found (Figure 12). There are some weakness of both models. We can see that both models overestimate the activity of compounds **5** and **17**. Glide SP also overestimate pIC50 of compound **11** (*Sublib-VI*). We can observe also some discrepancy in prediction of activity of compounds **1** and **2** using Glide XP model. The underestimation of pIC50 of α -methyl-D-galactoside **22**

Table 4: Relation between computed activity using Glide Score SP (XP) as a descriptor and measured activity for selected ligands. Activity data is collected from different sources referred in the last column.

Entry	No	Glide-Score SP/XP	Exp. pIC50	Calc. pIC50 SP / XP	Lit
1	1	-12.75 / -12.21	5.80	5.54 / 5.17	5
2	2	-12.71 / -11.86	5.89	5.51 / 4.94	17,18,20
3	4	-10.02 / -10.41	3.70	3.97 / 3.98	17,18,20
4	5	-10.23 / -10.21	3.00	4.05 / 3.84	17,18,20
5	6	-10.89 / -11.33	4.43	4.44 / 4.59	17,18,20
6	7	-10.89 / -11.61	5.00	4.44 / 4.78	17,18,20
7	9	-9.35 / -9.49	2.92	3.54 / 3.42	24
8	10	-9.43 / -10.48	4.40	3.58 / 4.03	24
9	11	-9.06 / -8.53	2.49	3.36 / 2.73	24
10	17	-9.25 / -9.98	2.45	3.47 / 3.69	23
11	18	-9.46 / -9.77	3.90	3.60 / 3.55	23
12	19	-9.09 / -9.57	3.63	3.38 / 3.42	23
13	20	-9.03 / -9.92	3.51	3.29 / 3.65	23
14	21	-7.32 / -9.91	3.22	2.34 / 3.65	23
15	22	-7.52 / -6.73	2.65	2.45 / 1.53	23

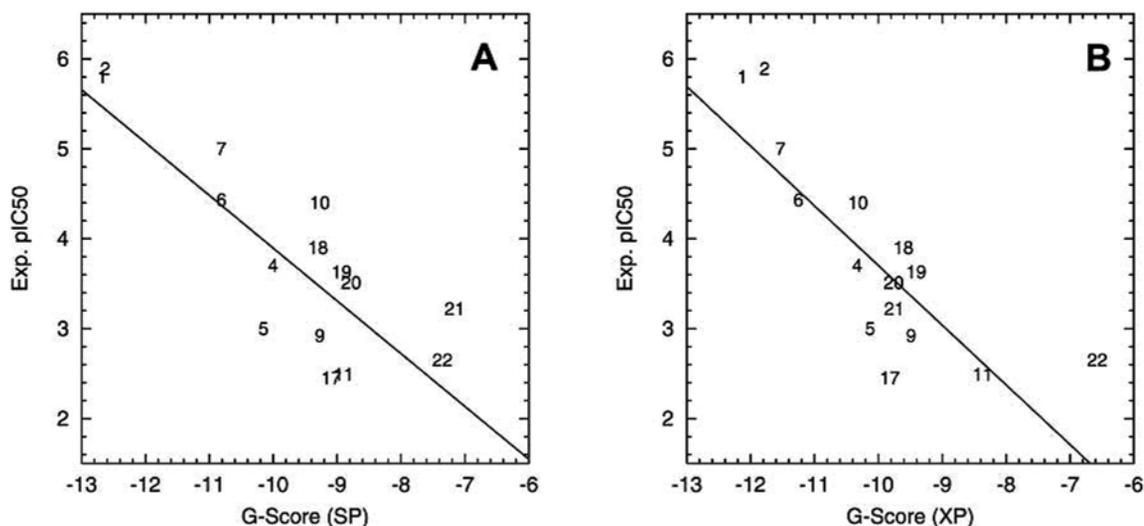


Figure 12: Models for predicting pIC₅₀ of the CT inhibitors based on Glide SP and Glide XP score.

with Glide XP is probably result of more rigorous scoring implemented in this protocol. Generally, using Glide XP did not lead to better predictivity comparing Glide SP in our case.

3. 3. Rescoring using Prime/MM-GBSA

We have used Prime/MM-GBSA protocol for rescoring Glide SP poses of the o-GM1. From the results collected in table 5 we didn't find a correlation between Glide Score and Prime/MM-GBSA energy (Table 5). However, we did find a better correlation between Prime/MM-GBSA and experimental results (Table 6 and Figure 13).

Table 5: Rescoring of different poses of the o-GM1 using Prime/MM-GBSA.

No. pose	G-Score (SP)	PRIME/MM-GBSA	Glide Energy	RMSD (Å)
1	-12.75	-87.91	-85.36	0.78
2	-12.41	-81.48	-81.45	0.84
3	-12.22	-78.59	-81.11	0.66
4	-12.16	-77.42	-78.59	0.66
5	-12.14	-77.98	-80.69	0.96
6	-12.05	-77.63	-79.84	0.94
7	-11.97	-91.09	-78.42	0.87
8	-11.55	-71.96	-72.02	0.99
9	-10.20	-70.34	-81.03	1.07
10	-9.75	-80.53	-75.16	1.07

Table 6: Prime/MM-GBSA energy obtained with rescoring of Glide SP (XP) poses.

Entry	No	Prime/ MM-GBSA energy (kJ/mol)				Exp. pIC ₅₀	Calc. pIC ₅₀	
		SP-1 ^a	SP-2 ^b	XP-1 ^a	XP-2 ^b		SP ^c	XP ^c
1	1	-87.91	-91.09 (7)	-86.35	-87.40 (2)	5.80	5.80	5.78
2	2	-87.31	-87.31 (1)	-86.73	-88.48 (5)	5.89	5.76	5.81
3	4	-64.86	-64.86 (1)	-55.57	-62.68 (5)	3.70	4.55	3.97
4	5	-50.26	-58.07 (5)	-52.14	-58.92 (3)	3.00	3.76	3.77
5	6	-60.27	-60.27 (1)	-62.30	-64.73 (10)	4.43	4.30	4.37
6	7	-60.84	-61.96 (2)	-66.78	-66.78 (1)	5.00	4.33	4.63
7	9	-39.28	-41.04 (9)	-49.09	-49.09 (1)	2.92	3.17	3.59
8	10	-49.73	-56.27 (3)	-57.54	-57.54 (1)	4.40	3.73	4.08
9	11	-37.12	-46.57 (9)	-42.32	-42.32 (1)	2.49	3.05	3.19
10	17	-36.40	-36.40 (1)	-35.13	-35.96 (2)	2.45	3.01	2.76
11	18	-39.40	-43.30 (10)	-42.62	-42.62 (1)	3.90	3.18	3.20
12	19	-35.42	-37.90 (2)	-38.65	-39.47 (4)	3.63	2.96	2.97
13	20	-36.45	-41.11 (5)	-34.36	-39.74 (6)	3.51	3.02	2.71
14	21	-	-29.61 (9)	-27.04	-30.10 (7)	3.22	-	2.29
15	22	-39.59	-40.94 (8)	-39.66	-39.66 (1)	2.65	3.19	3.03

^a Prime/MM-GBSA energy for the best pose ranked by Glide of the individual ligand.

^b The lowest Prime/MM-GBSA energy found among 10 poses of the individual ligand. The number in bracket represents the pose rank of ligand.

^c Prime/MM-GBSA energy of the best pose ranked by Glide SP (XP) score of each compound has been used as a descriptor in generation of models (Equations 4,5). References for experimental data are written in last column of table 5.

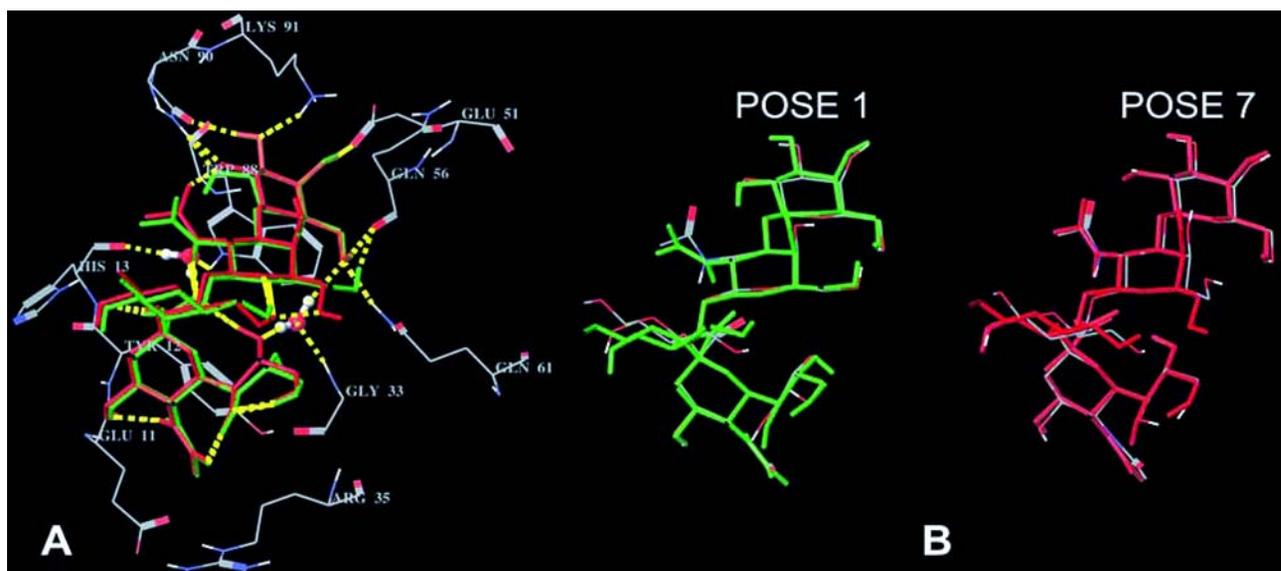


Figure 13: A.) Binding poses of o-GM1 ligand within CT receptor after applying PRIME/MM-GBSA rescoring (green – pose 1 and red – pose 7 from table 5). B.) Reorganisation of o-GM1 ligand after PRIME/MM-GBSA and comparison with predecessor pose of Glide SP (in CPK colors).

We have checked pose 1 (best Glide score) and pose 7 (the lowest Prime/MM-GBSA energy) to find possible rearrangements of ligand in receptor site after Prime/MM-GBSA rescoring that may lead to different ranking of poses (Table 5, Figure 13).

We can observe that rescoring leads to minor changes of the ligand conformations within receptor site (Figure 9 b.). These changes result from minimization of the ligand in receptor's environment. We may point out that rescoring procedure leads to reorientation of hydroxyls due to form intra and inter molecular hydrogen bonds and consequent stabilisation of receptor:ligand complex.

In table 6 are collected results of Prime/MM-GBSA rescoring of poses obtained with Glide for selected CT inhibitors.

Linear regression models that we built using PRIME/MM-GBSA descriptors obtained by rescoring best Glide SP (XP) poses of CT ligands within receptor site, equations (4,5) are written below:

$$\begin{aligned} \text{pIC}_{50} &= 1.05 (\pm 0.51) - 0.054 (\pm 0.009) \\ & \text{*Prime Energy (SP-1)} \\ (r^2 &= 0.74; F = 37.97; s^2 = 0.37; q^2 = 0.67) \end{aligned} \quad (4)$$

$$\begin{aligned} \text{pIC}_{50} &= 0.69 (\pm 0.50) - 0.059 (\pm 0.009) \\ & \text{*Prime Energy (XP-1)} \\ (r^2 &= 0.79; F = 44.17; s^2 = 0.30; q^2 = 0.73) \end{aligned} \quad (5)$$

These two regressions are shown in figure 14.

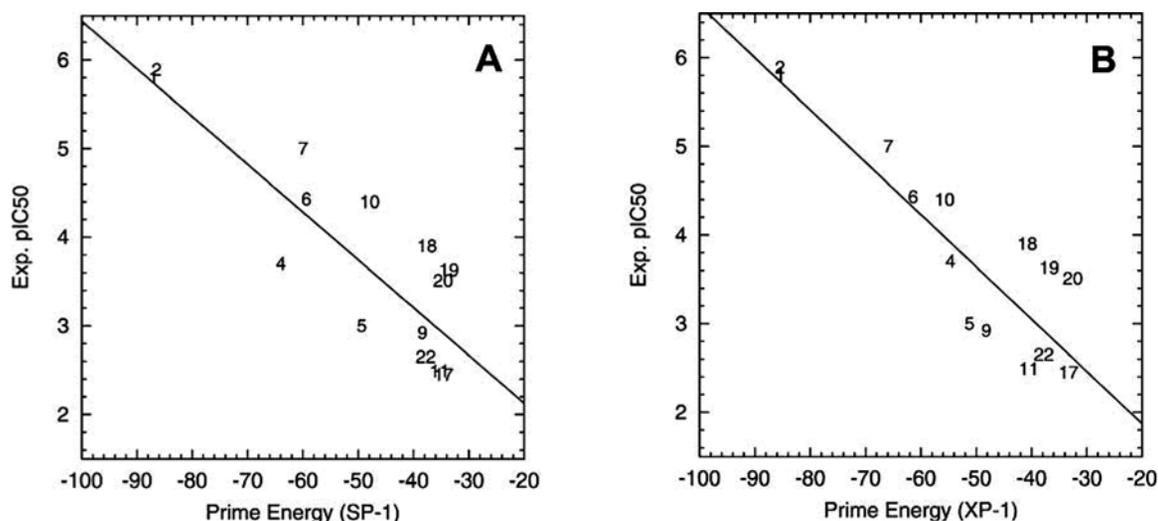


Figure 14: Two models for predicting pIC₅₀ of CT inhibitors using Prime Energy descriptor obtained with Prime/MM-GBSA rescoring of Glide SP (A) and Glide XP (B) poses.

We may observe that models using PRIME/MM-GBSA descriptors are better for predicting activity (pIC50) of CT inhibitors than models using Glide Score SP (XP) as a descriptor (Equations 2,3). We observed the best pIC50 predictivity for the model that uses descriptors that are result of rescoring of the best Glide XP poses.

4. Conclusion

We have compiled a virtual library of Cholera Toxin Inhibitors. Structures used in our study were taken from different sources or they were built from two different galactose scaffolds where galactose serves as an anchor. We tried to get insights into CT:ligand interactions and corresponding inhibitory potency against CT using Glide docking protocol. We also compared our “focused” library with the library of diverse compounds compiled by NCI. In NCI library there are only few compounds (6 of 3236) that have Glide score (SP) less than –8.0. On the other hand in our “focused” library a majority of ligands have Glide score (SP) lower than –8.0. When we analyzed NCI database we observed that “the best” NCI ligands contain a galactose anchor.

We found that pseudo GM1 ligands have generally speaking the lowest Glide Score (SP) among all of the ligands in our library. This result is not surprising and it is due the fact that these ligands are carefully designed with stepwise rationalization and all the structural elements that are important for successful binding were retained. Using more rigorous Glide XP scoring did not affect the trend of our results.

We have also built models for prediction of the inhibitory activity (pIC50) using Glide SP (XP) descriptors. Experimental data were taken from different sources. We have found that these models could be useful to predict the range of activity for new CT ligands. We also found that refinement of poses and consequent rescoring with PRIME/MM-GBSA leads to better predictivity of pIC50.

The information that we expressed in this study may lead to design (synthesis) of more potent CT inhibitor based on C-galactoside scaffolds.

5. Acknowledgments

We thank the European Union for financial support (under contract HPRN-CT-2002-00173, Glycidic Scaffolds Network). This work is partially supported also by found of the research program (P1-0201, Physical Chemistry) provided by Slovenian Ministry of Higher Education, Science and Technology.

6. References

1. WHO 2000, p. 249–256.
2. C. L. Gyles, *Can J. Microbiol.* **1992**, *38*, 734–746.
3. A. Bernardi, Č. Podlipnik, J. Jimenez-Barbero, in: *Protein-Carbohydrate Interactions in Infectious Diseases*, C. A. Bewley, (Ed), RCS Publishing **2006**, p. 73–91.
4. E. A. Merritt, W. G. J. Hol, *Curr. Opin. Struct. Biol.* **1995**, *5*, 165–171.
5. E. A. Merritt, P. Kuhn, S. Sarfaty, J. L. Erbe, R. K. Holmes, W. G. J. Hol, *J. Mol. Biol.* **1998**, *282*, 1043–1059.
6. I. A. Velter, M. Politi, C. Podlipnik, F. Nicotra, *Minirev. Med. Chem.* **2007**, *7*, 159–170.
7. E. Fan, C. J. O’Neil, D. D. Mitchell, M. A. Robien, Z. Zhang, J. C. Pickens, X. J. Tan, K. Korotkov, C. Roach, B. Krumm, C. L. M. J. Verlinde, E. A. Merritt, W. G. J. Hol, *Int. J. Med. Micr.* **2004**, *294*, 217–223.
8. W. I. Lencer, D. Slaslwsy, *Bioch. Biophys. Acta.* **2005**, *1746*, 314–321.
9. J. Moss, M. Vaughan, *J. Biol. Chem.* **1977**, *252*(7), 2455–2457.
10. J. Moss, M. Vaughan, *Annu. Rev. Biochem.* **1979**, *48*, 581–600.
11. W. E. Minke, D. J. Diller, W. G. J. Hol, C. L. Verlinde, *J. Med. Chem.* **1999**, *42*, 778–1788.
12. W. B. Turnbull, B. L. Precious, S. W. Homans, *J. Am. Chem. Soc.* **2004**, *126*, 1047–1054.
13. H. Oi, D. Matsuura, M. Miyake, M. Ueno, I. Takai, T. Yamamoto, M. Kubo, J. Moss, M. Noda, *PNAS* **2002**, *99*, 3042–3046.
14. T. Saito, M. Miyake, M. Toba, H. Okamatsu, S. Simizu, M. Noda, *Microbiol. Immunol.* **2002**, *46*, 249–255.
15. B. T. Hovey, C. L. M. J. Verlinde, E. A. Merritt, W. G. J. Hol, *J. Mol. Biol.* **1999**, *285*, 1169–1178.
16. D. Arosio, S. Baretti, S. Cattaldo, D. Potenza, A. Bernardi, *Bioorganic and Medicinal Chem. Lett.* **2003**, *13*, 3831–3843.
17. A. Bernardi, D. Arosio, L. Manzoni, D. Monti, H. Posteri, D. Potenza, S. Mari, J. Jiménez-Barbero, *Org. Biomol. Chem.* **2003**, *1*, 1–9
18. A. Bernardi, D. Arosio, S. Sonnino, *Neurochemicals research* **2002**, *27*(7/8), 539–545.
19. A. Bernardi, L. Carrettoni, A. G. Ciponte, D. Monti, S. Sonnino, *Bioorg. Med. Chem. Lett.* **2000**, *19*, 2197–2200.
20. A. Bernardi, D. Potenza, A. M. Capelli, A. Garcia-Herrero, F. J. Cañada, J. Jiménez-Berberero, *Chem. Eur. J.* **2002**, *8*(20), 4598–4612.
21. W. E. Minke, J. Pickens, E. A. Merritt, E. Fan, C. L. M. J. Verlinde, W. G. J. Hol, *Acta Cryst.* **2000**, *D56*, 795–804.
22. D. D. Mitchell, J. C. Pickens, K. Korotkov, E. Fan, W. G. J. Hol, *Bioorganic and Medicinal Chem.* **2004**, *12*, 907–920.
23. Č. Podlipnik, I. Velter, B. La Ferla, G. Marcou, L. Belvisi, F. Nicotra, A. Bernardi, *Carb. Res.* **2007**, *432*, 1651–1660.
24. W. E. Minke, F. Hong, C. L. Verlinde, W. G. J. Hol, E. Fan, *J. Biol. Chem.* **1999**, *274*, 33469–33473.
25. *NCI Diversity Dataset*, <http://autodock.scripps.edu/resources/databases>, (accessed: 24. 04. 2007)

26. B. La Ferla, F. Cardona, I. Perdigo, F. Nicotra, *Synlett*. **2005**, 17, 2641–2642.
27. F. Nicotra, *Synthesis of C-glycosides of biological interest*, Springer-Verlag, Berlin Heidelberg **1997**.
28. *Schrodinger L. L. C.*, <http://www.schrodinger.com>, (accessed: 24. 04. 2007)
29. H. Senderowitz, C. Parish, W. C. Still, *J. Am. Chem. Soc.* **1996**, 118, 2078–2086.
30. E. Polak, G. Ribiere, *Revue Francaise Inf. Rech. Oper., Serie Rouge*. **1969**, 16-R1, 35–43.
31. A. Bernardi, A. Checchia, P. Brocca, S. Sonnino, F. Zuccotto, *J. Am. Chem. Soc.* **1999**, 121, 2032–2036.
32. A. Bernardi, M. Galgano, L. Belvisi, G. Colombo, *J. Comput. Aided Mol. Des.* **2001**, 15, 117–128.
33. A. Bernardi, L. Raimondi, F. Zuccotto, *J. Med. Chem.* **1997**, 40, 1855–1862.
34. P. Brocca, A. Bernardi, L. Raimondi, S. Sonnino, *Glycoconjugate Journal* **2000**, 17, 283–299.
35. ChemAxon 2006, Marvin was used for drawing, displaying and characterising chemical structures, substructures and reactions.
36. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, D. E. Shaw, M. Shelley, J. K. Perry, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, 47, 1739–1749.
37. T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, 47, 1750–1759.
38. P. D. Lyne, M. L. Lamb, J. C. Saeh, *J. Med. Chem.* **2006**, 49, 4805–4808.

Povzetek

Generirali smo virtualno knjižnico inhibitorjev kolera toksina. Knjižnica vsebuje večinoma majhne molekule, ki se specifično vežejo na kolera toksin in tako onemogočijo vezavo toksina na ganglozid GM1, ki se nahaja na celičnih membranah epitelnih celic v prebavnem traktu. V knjižnico smo vključili znane inhibitorje zbrane iz različnih virov. Predlagali smo tudi nekaj novih inhibitorjev, ki smo jih razvili s pomočjo metod molekulskega modeliranja. Informacije, ki smo jih podali v tem članku lahko pripomorejo nadaljnemu razvoju bolj učinkovitih in predvsem metabolno stabilnih inhibitorjev kolera toksina.