DOI: 10.1002/chem.200700661

Improved Cyclodextrin-Based Receptors for Camptothecin by Inverse Virtual Screening

Andreas Steffen,[a] Carolin Thiele,[b] Simon Tietze,[c] Christian Strassnig,[b] Andreas Kämper,[a] Thomas Lengauer,[a] Gerhard Wenz,*[b] and Joannis Apostolakis*[c]

Abstract: We report the computer-aided optimization of a synthetic receptor for a given guest molecule, based on inverse virtual screening of receptor libraries. As an example, a virtual set of β-cyclodextrin (β-CD) derivatives was generated as receptor candidates for the anticancer drug camptothecin. We applied the two docking tools AutoDock and GlamDock to generate camptothecin complexes of every candidate receptor. Scoring functions were used to rank all generated complexes. From the 10% top-ranking candidates nine were selected for experimental validation. They were synthesized by reaction of heptakis-[6-deoxy-6-iodo]-β-CD with a thiol compound to form the hepta-substituted β-CDs. The stabilities of the camptothecin complexes obtained from solubility measurements of five of the nine CD derivatives were significantly higher than for any other CD derivative known from literature. The remaining four CD derivatives were insoluble in water. In addition, corresponding mono-substituted CD derivatives were synthesized that also showed improved binding constants. Among them the 9-H-purine derivative was the best, being comparable to the investigated hepta-substituted β-CDs. Since the measured binding free energies correlated satisfactorily with the calculated scores, the applied scoring functions appeared to be appropriate for the selection of promising candidates for receptor synthesis.

Keywords: bioinformatics · computer chemistry · cyclodextrins · drug delivery · molecular recognition

Introduction

Synthetic receptors are molecules that specifically bind guest molecules. In general, they cannot rival proteins in terms of binding affinity and specificity. However, they do exhibit numerous advantages over natural receptors that make them interesting candidates for diagnostic, therapeutic, analytical, and separation purposes.[1] Their three-dimensional structure can be more stable at high temperatures and non-physiological pH conditions. Their comparably low molecular weights and their better tolerability by the human immune system make them interesting candidates for the delivery of drugs. On the other hand, the development of synthetic receptors is hampered by tedious synthesis, by time consuming trial-and-error searches, and problems with solubility in water.[2] In this paper we demonstrate how the computational technique of virtual screening can support synthetic receptor design.

Virtual screening with protein-ligand docking tools, such as FlexX,[3] Gold,[4] Glide,[5] AutoDock,[6] or Dock,[7] is well established in the field of computer-aided drug design for identification of novel ligands for a given protein target.[8] In general, these docking tools consist of two components: 1) the conformational sampling of the guest molecules within the receptor binding site and 2) the scoring function for ranking the different conformations of the complex. In contrast to time-consuming molecular dynamic simulations, docking tools are designed to be fast and to provide solutions in the range of seconds to minutes. This efficiency

[a] A. Steffen, Dr. A. Kämper, Prof. Dr. T. Lengauer
Computational Biology and Applied Algorithmics
Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85
66123 Saarbrücken (Germany)

[b] C. Thiele, C. Strassnig, Prof. Dr. G. Wenz
Organische Makromolekulare Chemie, Saarland University
Geb. C4.2, 66123 Saarbrücken (Germany)
Fax: (+49) 681 302 3909
E-mail: g.wenz@mx.uni-saarland.de

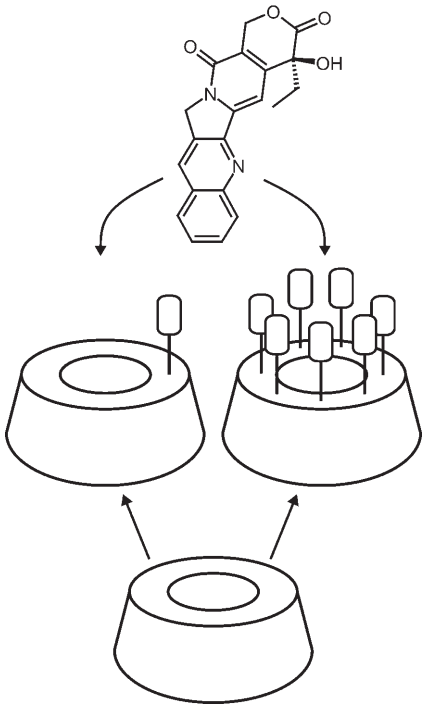
[c] S. Tietze, Dr. J. Apostolakis
Institute for Informatics
Research and Educational Unit for Bioinformatics
and Practical Informatics
Ludwig-Maximilians-Universität, Amalienstr. 17
80333, Munich (Germany)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

allows for virtual screening. Virtual screening can be understood as the virtual analog of high-throughput screening.[9] A library of candidate molecules are docked against the target and ranked according to the predicted binding affinity. The most promising molecules are subsequently submitted to experimental validation. Several successful studies have been reported in which virtual screenings identified relevant ligands leading to new promising drugs.[10]

De Jong et al. reported the identification of novel ligands for a given synthetic receptor by application of the docking tool Dock,[7] originally developed for drug design.[11] Docking was performed by placement of energetically minimized ligand structures into a β-CD dimer. Despite the fact that conformational flexibilities of both the receptor and the ligand had not been taken into account, nine out of 30 proposed ligands were found to bind to the receptor with high affinity. In this work we address the opposite problem of looking for synthetic receptors that bind a given ligand with high affinity, since this task is highly relevant for the complexation and controlled delivery of drugs.

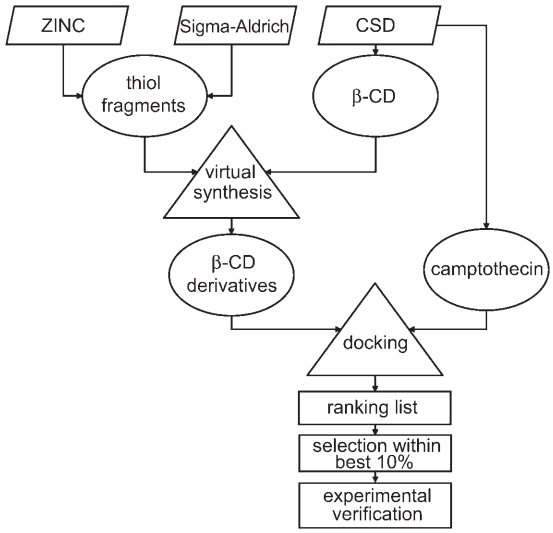
Camptothecin (Scheme 1) was chosen here as the ligand for the design of a tailored receptor by means of computational methods, as it represents a promising class of antineoplastic agents that exhibit a broad spectrum of activity against several types of cancer, including colorectal and ovarian cancer.[12] Camptothecin inhibits topoisomerase I, which is a nuclear enzyme involved in the relaxation of DNA during cell replication and transcription.[13] Unfortunately, its high therapeutic potential is hampered by low solubility and stability.[14] Some attempts have been made to circumvent these difficulties by means of pharmaceutical formulations and inclusion in cyclodextrins (CDs).[14–19]



Scheme 1. Structure of camptothecin and schematic receptor design by attachment of one or seven building blocks to β-CD.

CDs and their derivatives are an interesting class of organic hosts[20,21] that have been applied in various pharmaceutical formulations.[22] Their hydrophobic cavity combined with a hydrophilic exterior designates their application for solubilizing small hydrophobic molecules, such as drugs, in water.[23] In this way, CDs can increase the plasma level of the complexed drugs and thus increase their therapeutic effect. Recently, we demonstrated that CDs are able to recognize the thicknesses of included monomeric[24] and polymeric[25,26] guest molecules. Selectivities and affinities of CDs can be increased by chemical modifications of the CDs.[27]

In this study we focus on the development of β-CD derivatives specifically tailored for solubilization of camptothecin. We chose a straightforward synthesis procedure for regioselective modification of β-CD by nucleophilic displacement reactions of 6-*O*-iodo or 6-*O*-tosyl-β-CDs with various thiols.[28] First, a virtual library of candidate receptors was generated. Each candidate of this library was then docked onto camptothecin. This approach is referred to as inverse virtual screening as the docking direction is inverted[29] with respect to common virtual screenings in drug design.[9] Similar to the normal virtual screening scenario, we applied scoring functions to rank the different candidates. Top-ranking candidate receptors were synthesized and experimentally tested (see Scheme 2).



Scheme 2. Design of the study.

Results

A virtual library (1846 entities) of 6-*O*-mono- and 6-*O*-hepta-substituted β-CD derivatives was generated from the β-CD core and thiol building blocks. The structure of the complex between camptothecin and the different derivatives

was predicted and the derivatives were ranked according to the score of the complex (see Computational Methods). [Scores are used as heuristic estimates of $\Delta G°$. By convention, the lower the score, the more favorable the interaction; thus, the first rank corresponds to the complex exhibiting the lowest score.] Usually, protein-ligand docking tools explore the conformational space of ligands, for example, drug molecules, while treating the protein as rigid (in its crystal structure conformation) during the simulated binding. For our work this simplification was not appropriate. The conformations of the virtually generated β-CD derivatives had to be generated during the docking process. Since the so-called flip-flop hydrogen bonds between secondary hydroxyl groups of neighboring glucose units restrict the flexibility of the β-CD core,[30] only the side-chains were considered as flexible in the present study, whereas the β-CD core was kept as given in the crystal structure (see Computational Methods). Conversely, camptothecin was described by a single conformation, since it is rather rigid. Consequently, we performed an inverse docking in which the receptor conformation was optimized in the field of the rigid ligand.[29,31]

Table 1. Building blocks selected by virtual screening of corresponding β-CD derivatives.

	IUPAC Name	CAS-No.	mono derivative	hepta derivative	Method[a]
1	1-methyltetrazole-5-thiol	13183-79-4	**21**	**11**	AD
2	2-aminoethanethiol	60-23-1	**22**	**12**	GD
3	2-mercaptoacetic acid	68-11-1	**23**	**13**	GD
4	2-mercaptoethanesulfonate	3375-50-6	**24**	**14**	GD
5	2-mercaptopropanoic acid	79-42-5	**25**	**15**	GD
6	3-mercaptopropane-1,2-diol	96-27-5	**26**	**16**	GD
7	3-mercaptopropanoic acid	107-96-0	**27**	**17**	GD
8	9*H*-purine-6-thiol	50-44-2	**28**	**18**	AD
9	pyridine-2-thiol	2637-34-5	**29**	**19**	AD, GD

[a] AD = AutoDock, GD = GlamDock.

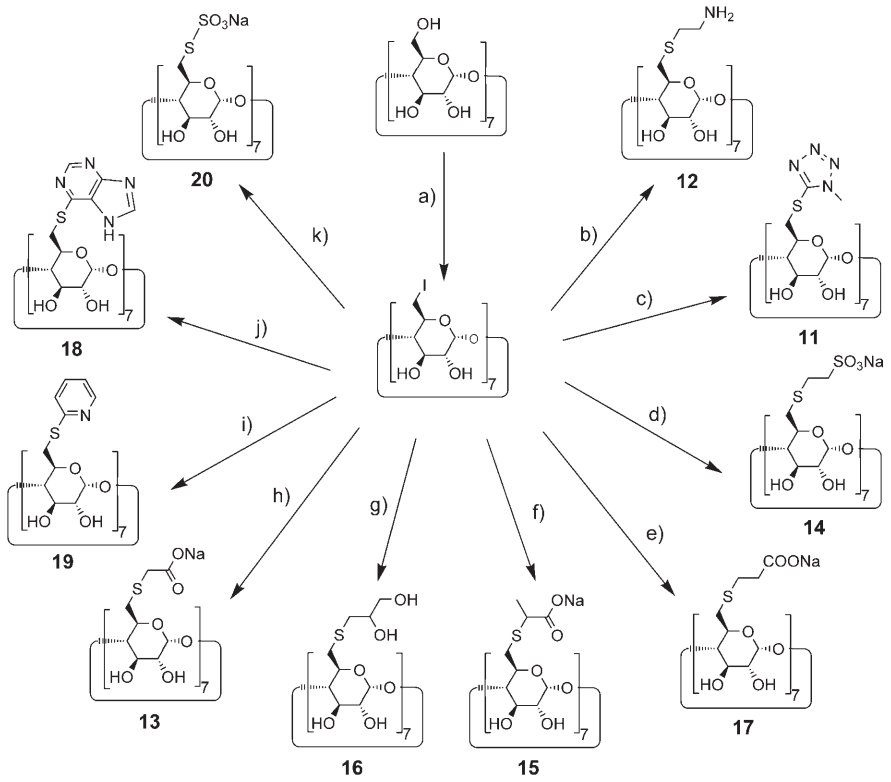
The virtual screening was performed through the use of two docking tools, namely AutoDock[6,32–34] and GlamDock.[35] AutoDock uses a grid-based energy evaluation procedure based on the Amber force field. Four algorithms are implemented for optimizing the complex conformation. The Lamarckian genetic algorithm was chosen for our study, since it had been shown to be most effective and reliable.[6] The current version of GlamDock relies on a Monte Carlo procedure based on matching of functional groups of the ligand with favorably interacting probes in the binding site combined with local minimization.[36,37] The scoring function for the optimization β-CD derivatives conformations is a continuous-gradient approximation to the docking version of ChemScore.[38,39] For ranking the compounds we additionally introduced a size-penalizing term (see Computational Methods). GlamDock has been validated on benchmark sets of literature data. It performed comparably or better than the state-of-the-art methods on the Kellenberger data set.[40]

For experimental verification only those compounds were considered that were found by at least one docking tool within the top 10% of the respective ranking lists. All potential candidates were visually inspected. Promising β-CD derivatives were selected for synthesis and further experimental investigation. Furthermore, the building blocks for synthesizing the β-CD derivative had to be commercially available. AutoDock favored β-CD derivatives with aromatic and hydrophobic side-chains, whereas GlamDock in combination with the size penalty mainly suggested derivatives forming hydrogen bonds to camptothecin (Table 1). The predicted affinity scores for heptakis derivatives were generally more favorable than for the corresponding mono derivatives with both docking programs.

Nine heptakis-β-CD derivatives were synthesized by nucleophilic displacement reactions in good yields (see Scheme 3). Four of them (**11**, **15**, **18**, **19**) were insoluble in water. For a closer investigation of the molecular interactions and for obtaining an estimate of binding affinities of all substituents, we also synthesized the corresponding mono derivatives **21**–**28** (see Scheme 4), which were all soluble in water. Furthermore the heptakis-substituted thiosulfate β-CD (compound **20**) was synthesized. This molecule had an unfavorable predicted binding energy and served as a negative test.

The binding constants K of all synthesized β-CD derivatives were determined from the solubility isotherms. The increase of the solubility of camptothecin with increasing concentration of the CD derivatives is demonstrated in Figure 1. The binding constants K were derived from the slope and the solubility of camptothecin.[19] To assure comparability we additionally measured the binding constants of the native β-CD, hydroxypropyl-β-CD (HP-β-CD) and randomly methylated β-CD (RDM-β-CD), which were already investigated by Kang et al.[19] It should be noted that the value of K obtained for RDM-β-CD ($186\,M^{-1}$) significantly differed from its literature value ($909.7\,M^{-1}$).[19] This difference might be caused by different experimental protocols and different substitution patterns of the randomly methylated β-CDs.

Out of the nine synthesized receptors from the virtual screening, five exhibit binding constants K clearly superior to the ones of the native β-CD and the two other known CD derivatives[19] (see Table 2 and Figure 1). Heptakis-[6-deoxy-6-(2-sulfanylethanesulfonic acid)]-β-CD (**14**) showed the highest value of K with $7496\,M^{-1}$. Since receptors **11**, **15**, **18**, and **19** were insoluble in water, the corresponding mono derivatives **21**–**29** were investigated. Among them, mono-[6-deoxy-6-(6-sulfanyl-9H-purine)]-β-CD (**28**) showed the strongest binding affinity with $3629\,M^{-1}$, which is in the range of the hepta-substituted CD derivatives. As predicted, the negative test example (compound **20**) exhibits a comparably low binding affinity with a K value of $370\,M^{-1}$.



Scheme 3. Synthesis of hepta substituted CD derivatives. a) 1) PPh_3, I_2, DMF 2) CH_3ONa, CH_3OH; b) **2**, NH_4HCO_3, DMF/H_2O; c) **1**, NEt_3, DMF; d) **4**, NEt_3, DMSO; e) 1) methyl ester of **7**, NEt_3, DMF 2) NaOH; f) 1) methyl ester of **5**, NEt_3, DMF 2) NaOH; g) **6**, NEt_3, DMF; h) 1) methyl ester of **3**, NEt_3, DMF 2) NaOH; i) **9**, NEt_3, DMF; j) **8**, NEt_3, DMF; k) sodium thiosulfate, DMSO.

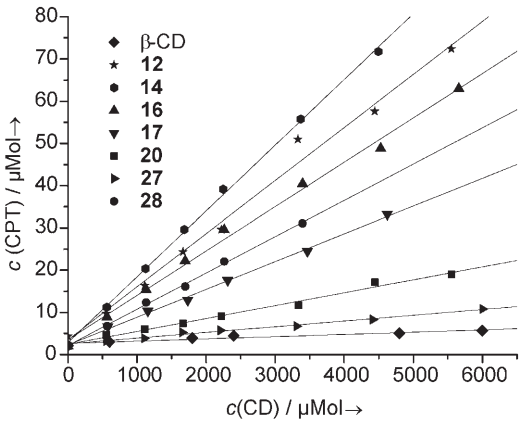


Figure 1. Dependence of the solubility of camptothecin (CPT) on the concentration of the β-CD derivatives.

Comparison of the binding free energies $\Delta G°$ for the majority of the mono-substituted β-CD derivatives **21**–**27** and **29** with unsubstituted β-CD shows a stabilization energy of around $\Delta\Delta G° = -2\ kJ\,mol^{-1}$ caused by one building block. The same comparison for the hepta-substituted β-CDs **12**–**17** results in $\Delta\Delta G° = -(5–9)\ kJ\,mol^{-1}$ due to seven building blocks. This may suggest that only three to four out of seven building blocks are involved in binding, possibly for steric reasons. Remarkably, one 6-sulfanyl-9H-purine building block in compound **28** leads to an exceptionally strong stabilization of $\Delta\Delta G° = -7\ kJ\,mol^{-1}$.

In Figures 2 and 3 the AutoDock and the GlamDock affinity scores, respectively, are plotted versus the experimentally determined values of the binding affinities. For compounds **11**, **15**, **18**, and **19** no binding free energy could be experimentally determined due to insolubility in water. The correlation coefficient for AutoDock is $r = 0.57$ (residual standard error of the regression $2.4\ kJ\,mol^{-1}$), for GlamDock $r = 0.82$ (residual standard error of the regression $1.6\ kJ\,mol^{-1}$). Compound **14** is an evident outlier for both docking tools, but particularly in the case of AutoDock. If this compound is omitted, the correlation coefficient for AutoDock increases to 0.78 (residual standard error of the regression $1.7\ kJ\,mol^{-1}$).

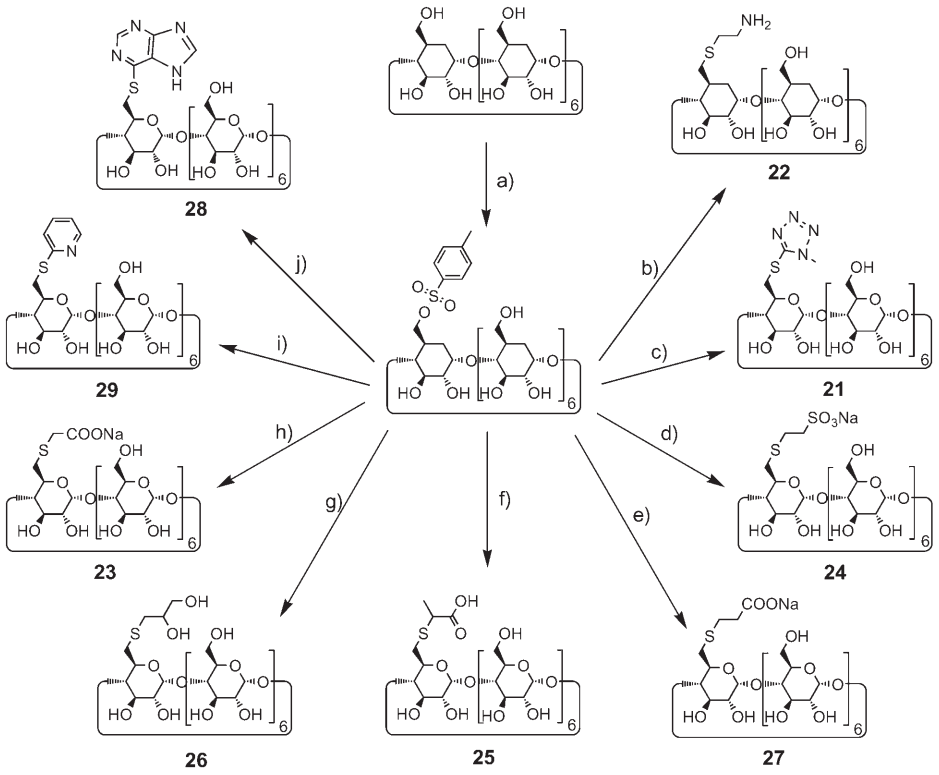
Discussion

Due to the flexibility and larger size of artificial receptors, virtual screening of receptors (inverse screening) is in general more complex than virtual screening of ligands.[11] For a given complex the predicted binding free energy $\Delta G°$ (score) consists in principle of three components [Eq. (1)] in which $\Delta G°_R$ is the change of energy in the receptor molecule, $\Delta G°_L$ is the change of energy in the ligand upon complexation, and $\Delta G°_{RL}$ is the interaction energy of the complex.

$$\Delta G°(\text{predicted}) = \Delta G°_R + \Delta G°_L + \Delta G°_{RL} \quad (1)$$

In ligand screening the larger receptor structure is normally treated as rigid and thus $\Delta G°_R$ cancels. The estimated binding energy of the system depends only on the interaction energy between ligand and receptor $\Delta G°_{RL}$ and to a small extent on the change of the internal energy of the flexible ligand $\Delta G°_L$.

In inverse screening the receptors were treated flexible, whereas the guest molecule was kept rigid ($\Delta G°_L = 0$). Due to the large size of the receptor the change of its internal energy ($\Delta G°_R$) contributes predominantly to the binding energy. Self-inclusion of the receptor can lead to low-energy conformations of the receptors with little interaction to the guest molecule. This is shown in Figure 4, which depicts a



Scheme 4. Synthesis of mono substituted CD derivatives. a) Tos-Cl, NaOH, ACN/H_2O; b) **2**, NH_4HCO_3, DMF/H_2O; c) **1**, NEt_3, DMF; d) **4**, NEt_3, DMF; e) 1) methyl ester of **7**, NEt_3, DMF 2) NaOH; f) 1) methyl ester of **5**, NEt_3, DMF 2) NaOH; g) **6**, NEt_3, DMF; h) 1) methyl ester of **3**, NEt_3, DMF 2) NaOH; i) **9**, NEt_3, DMF; j) **8**, NEt_3, DMF.

complex with a predicted favorable score. Camptothecin lies on top of the receptor. One of the hydrophobic side chains is buried in the CD cavity and leads to a favorable internal energy (ΔG^{o}_{R}), which compensates the poor intermolecular interactions (ΔG^{o}_{RL}). This leads to well-scoring complexes that show little interaction between ligand and receptor. Furthermore, the average interaction of a system increases quadratically with the number of its atoms, and therefore receptors with large substituents are generally scored more favorably than smaller receptors.

There are at least three different approaches to address this type of problem within the paradigm of fast virtual screening:

1) Score the complexes only according to the interaction between receptor and ligand ΔG^{o}_{RL}.
2) Add a size-dependent term to the ranking function, which simply depends on the number of atoms, to penalize large complexes.
3) Constrain the docking to allow only conformations with the camptothecin in the binding site of the receptor.

In the first approach it is important to consider the intramolecular receptor energy ΔG^{o}_{R} during the conformational sampling to avoid physically unreasonable conformations of the receptor. However, the proportionality of interactions to the size of the receptor remains and leads to better scoring of unspecifically interacting hydrophobic receptors. The second approach reduces this problem, but is highly empirical and requires the definition of more or less arbitrary weights for the size term. Finally, in the last option, conformations as shown in Figure 4 are explicitly forbidden, even though they may correspond to the most probable structure of the complex.

In the current work, we chose two different combinations of these approaches. In the screening with GlamDock we used approaches 1) and 2) by explicitly adding a term penalizing the size of the receptor for the ranking, and only used the intermolecular interaction energy ΔG^{o}_{RL} for scoring. For AutoDock we used approach 3), as the sampling region of the receptor is limited in such a way that camptothecin is always within the binding cleft of the derivatives.

Overall, the results show that these two approaches have their own advantages and disadvantages. The AutoDock approach led to the selection of receptors that were highly hydrophobic, and could therefore not be measured. On the other hand, it also led to the identification of compound **28**, which is the only mono derivative that can rival the heptakis-substituted derivatives in terms of binding affinity. The GlamDock approach proposed receptors with smaller and more hydrophilic side chains in general, which show improved binding affinity compared to β-CD. Furthermore the scores correlate reasonably well with the experimental binding affinities. It is interesting to note that one derivative (compound **14**) appears to be an outlier for both scoring functions. Both, AutoDock and GlamDock significantly underpredict its binding affinity. Nevertheless, in spite of the uncertainties of structure prediction, and the modeling itself, the overall results suggest that at least the tendency of binding affinity is reproduced. For AutoDock a residual standard error of 9.11 kJ mol^{-1} was reported in literature for a set of 30 protein–ligand complexes.[6] Furthermore, with regression methods a cross-validated correlation coefficient of 0.89 and a standard deviation of 2.38 kJ mol^{-1} were reported for a set of 218 complexes between β-CD and different guest molecules (CODESSA-PRO descriptors).[41] This correlation is better than those achieved in the present work ($r = 0.82$), while the average error is even higher than in the one here (1.6 kJ mol^{-1}).

Table 2. Binding constants K and binding free energies $\Delta G°$ for camptothecin in 0.02 M HCl.

	Compound	K [M^{-1}]	$\Delta G°$ [kJ mol^{-1}]
	β-CD	202 ± 30	−13.2 ± 0.5
	HP-β-CD	223 ± 32	−13.4 ± 0.4
	RDM-β-CD	186 ± 12	−12.9 ± 0.2
11	heptakis-[6-deoxy-6-(1-methyl-5-sulfanyl-tetrazole)]-β-CD	insoluble	–
12	heptakis-[6-deoxy-6-(2-aminoethylsulfanyl)]-β-CD	4821 ± 572	−21.0 ± 0.3
13	heptakis-[6-deoxy-6-(2-sulfanyl acetic acid)]-β-CD	1450 ± 177	−18.0 ± 0.3
14	heptakis-[6-deoxy-6-(2-sulfanylethanesulfonic acid)]-β-CD	7496 ± 2002	−22.1 ± 0.7
15	heptakis-[6-deoxy-6-(2-sulfanylpropanoic acid)]-β-CD	insoluble	–
16	heptakis-[6-deoxy-6-(3-sulfanylpropane-1,2-diol)]-β-CD	4106 ± 475	−20.6 ± 0.3
17	heptakis-[6-deoxy-6-(3-sulfanylpropanoic acid)]-β-CD	3134 ± 364	−19.9 ± 0.3
18	heptakis-[6-deoxy-6-(6-sulfanyl-9*H*-purine)]-β-CD	insoluble	–
19	heptakis-[6-deoxy-6-(2-sulfanylpyridine)]-β-CD	insoluble	–
20	heptakis-[6-deoxy-6-sulfanylsulfonyloxysodium)]-β-CD	370 ± 48	−14.7 ± 0.3
21	mono-[6-deoxy-6-(1-methyl-5-sulfanyltetrazole)]-β-CD	465 ± 55	−15.2 ± 0.3
22	mono-[6-deoxy-6-(2-aminoethylsulfanyl)]-β-CD	498 ± 69	−15.4 ± 0.3
23	mono-[6-deoxy-6-(2-sulfanyl acetic acid)]-β-CD	493 ± 61	−15.4 ± 0.3
24	mono-[6-deoxy-6-(2-sulfanylethanesulfonic acid)]-β-CD	431 ± 56	−15.0 ± 0.3
25	mono-[6-deoxy-6-(2-sulfanylpropanoic acid)]-β-CD	419 ± 53	−15.0 ± 0.3
26	mono-[6-deoxy-6-(3-sulfanylpropane-1,2-diol)]-β-CD	531 ± 79	−15.6 ± 0.4
27	mono-[6-deoxy-6-(3-sulfanylpropanoic acid)]-β-CD	569 ± 68	−15.7 ± 0.3
28	mono-[6-deoxy-6-(6-sulfanyl-9*H*-purine)]-β-CD	3629 ± 1567	−20.3 ± 1.1
29	mono-(2-mercaptopyridine)-β-CD	641 ± 53	−16.0 ± 0.2

chains leads to higher binding affinity. This effect is illustrated in Figure 5 in which the hydrophobic parts of the cysteaminyl side chains of compound **12** show a good shape complementarity and hydrophobic interactions to the camptothecin ring system (dashed green lines). In addition, camptothecin is also able to interact specifically by forming directional hydrogen bonds. The complex exhibits three intermolecular hydrogen bonds (dashed red lines) of the ammonium groups to hydrogen bond acceptor atoms of camptothecin.

On the other hand, polar interacting groups pay a relatively high desolvation penalty in aqueous solution and are most probably not the main driving

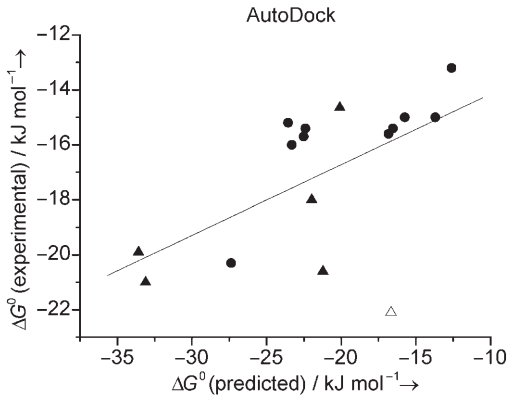


Figure 2. Predicted binding free energies (AutoDock) plotted versus the experimental binding free energies. The mono derivatives are depicted by filled circles, the heptakis derivatives are shown as filled triangles, except compound **14** shown as open triangle.

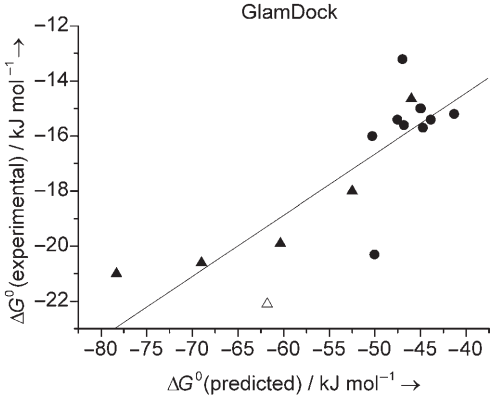


Figure 3. Predicted binding free energies (GlamDock) plotted versus the experimental binding free energies. The mono derivatives are depicted by filled circles, the heptakis derivatives are shown as filled triangles, except compound **14** shown as open triangle.

Regression methods are not applicable for our situation as no training data for generating the regression was available. However, the comparison suggests that our results on this system are at the upper bound of what can be achieved with simple modeling approaches.

To exemplify the interactions involved in the complex formation of camptothecin and the described β-CD derivatives, we show predicted complex structures of compounds **12** and **18** in Figures 5 and 6, respectively. The molecular structure of camptothecin offers several possibilities for intermolecular interactions. The large hydrophobic area of camptothecin facilitates dispersive interactions. Consequently, an enlargement of the hydrophobic CD cavity by hydrophobic side-

force behind complex formation for the regarded system. Additionally we could show that aromatic building blocks, for example, purine in compound **28** and, to a smaller extent, pyridine in compound **29** increase complex stability. This result might be best explained by the occurrence of π-stacking (dashed pink line) between camptothecin and the heterocycle (see Figure 6).

In general, hydrophobic interactions are the main driving force behind affinity, while polar interactions are more responsible for the specificity of the interaction. While a general size effect can be observed in the data, specific effects are evident, as mono-substituted compounds exist that bind better than heptakis-substituted compounds and vice versa.

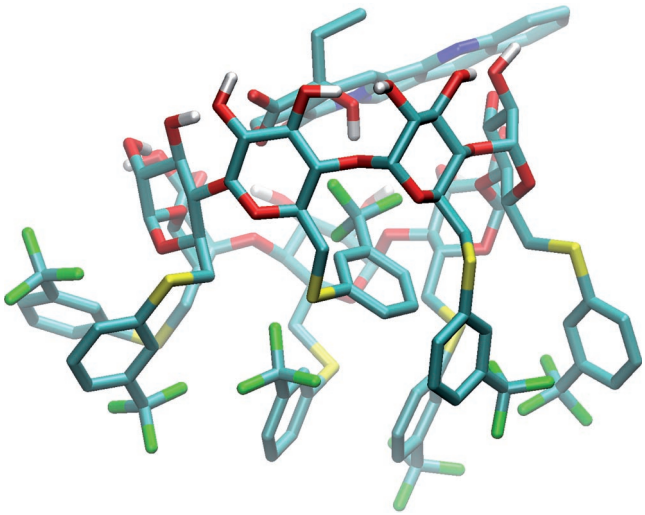


Figure 4. GlamDock docking result for a candidate receptor (heptakis-{6-deoxy-6-[3-(trifluoromethyl)benzenesulfanyl]}-β-CD) and camptothecin with a predicted low binding free energy. Hydrogen atoms are omitted for clarity.

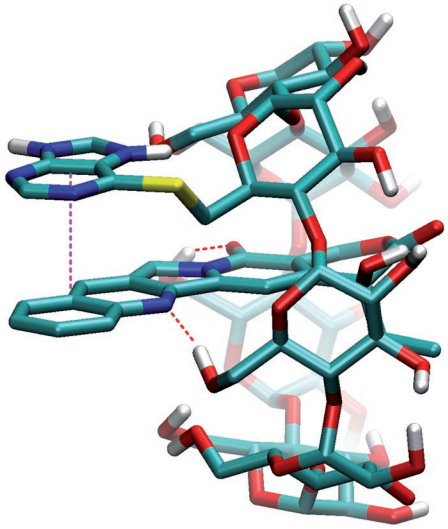


Figure 6. The figure shows the generated complex structure of compound **28** to camptothecin (AutoDock). The dashed pink line depicts a possible π-stack interaction. Hydrogen bonds are shown as dashed red lines. Apolar hydrogen atoms are omitted for clarity.

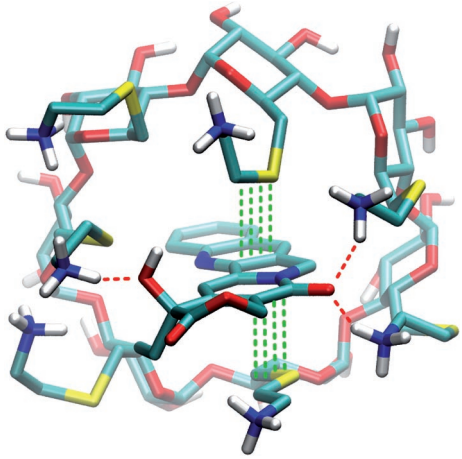


Figure 5. The figure shows the generated complex structure of compound **12** to camptothecin (GlamDock). Hydrogen bonds are depicted by dashed red lines, strong hydrophobic interactions are shown as dashed green lines. Apolar hydrogen atoms are omitted for clarity.

Compound **28** binds better than all other mono-substituted derivatives and better than some heptakis-substituted compounds. It is recognized by both affinity predictions as the best of the mono-derivatized compounds. Inversely, an extension of the cavity does not necessarily result in an increased binding affinity of the complex. The heptakis-substituted thiosulfate β-CD derivative (compound **20)**, for example, exhibits a rather weak binding free energy of $-14.62\ \mathrm{kJ\,mol^{-1}}$ as predicted by both docking tools.

Conclusion

We have investigated a rational optimization approach to synthetic receptor design. Our approach is complementary to the work of de Jong et al.,[11] who described the identification of new ligands for a given CD host. Our results indicate that inverse virtual screening can support the identification of improved receptors for a given ligand and might open up novel possibilities for the tailored design of drug delivery systems. Finally, we would like to point out that this approach is not limited to CD derivatives. The rules for generating the virtual library of hosts can be arbitrarily expanded to other host–guest systems of interest.

Experimental Section

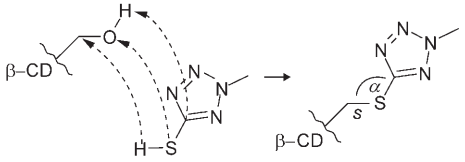
Computational methods: The structures of all compounds were saved in the MOL2 file format. For the preparation of the camptothecin structure the crystal structure of an iodoacetyl derivative of camptothecin was obtained from the Cambridge Structural Database (CSD)[42] (ID: CAMPTC10).[43] The iodoacetyl group was replaced by a hydrogen atom in order to construct the unmodified camptothecin molecule. Missing hydrogen atoms were added with SYBYL 6.7. Subsequently a force-field optimization was performed with the MMFF94s force field[44] until gradient convergence ($0.005\ \mathrm{kcal\,mol^{-1}\,\AA^{-1}}$).

For the preparation of the β-CD core structure, the crystal structure of β-CD with resolved deuterium positions, derived from neutron diffraction data, was obtained from the CSD (ID: BUVSEQ03).[45] The D atoms were changed to H atoms. All remaining atom types were inspected and corrected according to the SYBYL atom type rules if necessary. Water molecules present in the crystal structure were completely removed.

Building blocks containing one thiol group were extracted from the Sigma-Aldrich catalog (605 entities) and the ZINC database[46] (318 entities) by means of standard substructure search interfaces allowing a molecular weight range of 0–200 $\mathrm{g\,mol^{-1}}$ in order to limit the sizes. Low-energy conformations were generated with CORINA,[47] and reasonable protonation states for pH 7 were assigned to all compounds, that is, acidic groups were deprotonated and aliphatic amines were protonated.

The virtual library (1846 entities) of 6-*O*-mono- and 6-*O*-hepta-substituted β-CDs was generated from the β-CD core and each of the thiol build-

ing blocks with the help of a PYTHON script. Each building block was transformed in three-dimensional space such that its thiol group was superimposed onto either one or all seven primary oxygen atoms of the β-CD structure. The oxygen atoms were virtually substituted by the sulfur atoms of the building blocks. Excess hydrogen atoms were removed and the lengths and the angles α of the C–S bonds were adjusted to standard values (see Scheme 5). Rotatable torsion angles were optimized during docking.



Scheme 5. Schematic drawing of the virtual synthesis of the β-CD library.

Docking and scoring was performed with AutoDock 3.05 and GlamDock 1.0 on a cluster of thirty Intel P4 Xeon 3.06 GHz CPUs. The conformational search was performed for the β-CD derivatives, whereas camptothecin was kept rigid. For AutoDock the grid maps were generated for camptothecin with AutoDock tools[48] by defining a $50\times50\times50$ Å cube around camptothecin with grid spacing of 0.375 Å. The number of energetic evaluations was set to five million, the number of GA runs to 100, and the maximal possible number of torsions to 30. For all other parameters default values were used. For GlamDock the docking protocol consisted of five single docking runs each consisting of 650 Monte Carlo minimization (MCM) steps, with 15 steps of Levenberg–Marquardt[49] minimization in torsion space[50] at each MCM step. A maximum of 40 poses were finally post-minimized by 150 steps of Levenberg–Marquardt.

For AutoDock docking and ranking was performed based on the overall score for AutoDock (Dock Score). In the case of GlamDock the scoring function for docking considers the internal energies of the receptors, whereas for ranking a size penalizing variant of the fitness score for GlamDock without internal energy was used. GlamDock does not constrain the receptor to dock around the ligand. It explicitly allows conformations where the ligand lies on top of the CD ring (see Figure 4). Such conformations are mainly stabilized by the internal energy of the receptor, which on the average scales quadratically with its size (the number of atoms). The size penalty has the effect of identifying more specifically interacting complexes; however, this is a clearly empirical approach of ranking the screening results and does not necessarily correlate with binding affinity. The reason for the size penalty was that initially both virtual screening results contained mainly large hydrophobic receptors on top ranks.

Materials: Compounds **2**, **6**, **7**, **8**, **9** and DMF (absolute) were purchased from Fluka, **1** and **5** from Aldrich, **4** from Sigma, **3** from Merck, camptothecin from TCI Europe. β-CD, randomly methylated β-CD (RDM-β-CD, DS = 1.8 per glucose unit) and hydroxypropyl-β-CD (HP-β-CD, DS = 0.9 per glucose unit) were donated by Wacker. The purification of the synthesized compounds was performed by nanofiltration with a Berghof BM-5 membrane (cut-off molecular weight 500 Da) and MilliQ water. Syringe filters from Roth (CME, 0.22 μm) were used for separation of insoluble materials before UV measurement.

General procedure for the synthesis of heptakis-6-*O*-β-CD derivatives (non-acidic thiol building blocks): A solution of heptakis-[6-deoxy-6-iodo]-β-CD[51] (2.86 g, 1.5 mmol) in DMF (15 mL) was mixed with triethylamine (3.66 mL, 26.3 mmol) and the thiol compound (26.3 mmol). After stirring for 3 d at 60 °C under N_2, the product was concentrated in vacuo, precipitated by addition of ethanol or acetone, filtered, dried in vacuo, and further purified by nanofiltration.

General procedure for the synthesis of mono-6-*O*-β-CD derivatives (non-acidic thiol building blocks): A solution of mono-6-*O*-(*p*-toluolsulfonyl)]-β-CD[52] (1.93 g, 1.5 mmol) in DMF (20 mL) was mixed with triethylamine (2.10 mL, 15.0 mmol) and the thiol compound (15.0 mmol). After stirring for 3 d at 60 °C under N_2 the product was worked up as described above.

General procedure for the synthesis of heptakis-6-*O*-β-CD derivatives (acidic thiol building blocks): A solution of heptakis-[6-deoxy-6-iodo]-β-CD[51] (2.86 g, 1.5 mmol) in DMF (20 mL) was mixed with triethylamine (3.66 mL, 26.3 mmol) and the methyl ester of the thiol compound (26.3 mmol). After stirring for 3 d at 60 °C under N_2 the product was concentrated in vacuo, precipitated by addition of ethanol or acetone, and filtered. This product was stirred in 1 M NaOH (50 mL) for 18 h and further purified by nanofiltration.

General procedure for the synthesis of mono-6-*O*-β-CD derivatives (acidic thiol building blocks): A solution of mono-[6-deoxy-6-(*p*-toluolsulfonyl)]-β-CD[52] (1.93 g, 1.5 mmol) in DMF (20 mL) was mixed with triethylamine (2.10 mL, 15.0 mmol) and the methyl ester of the thiol compound (15.0 mmol). After stirring for 3 d at 60 °C under N_2 the product was worked up as described above.

Determination of binding constants: The binding constants K and the corresponding binding free energies $\Delta G°$ of the camptothecin complexes were determined with the solubility method described by Higuchi and Lach[53] and Kang et al.[19] Solutions of the CD derivatives (0–6 mM) in 0.02 M HCl (5 mL) were stirred with an excess of camptothecin at 25 °C for 18 h. The concentration of dissolved camptothecin was determined spectrophotometrically after filtration from an extinction coefficient of camptothecin ε(370 nm, water/DMSO 1:1 v/v) = 42282 M^{-1} cm^{-1}. Due to the overlap of the absorbances of camptothecin and β-CD derivative **28** in the aqueous phase, an aqueous camptothecin solution (5 mL) was extracted with trichloroethane/trifluoroacetic acid 75:1 v/v (5 mL). The camptothecin concentration in the organic phase was determined spectrophotometrically with ε(370 nm, trichloroethane/TFA 75:1 v/v) = 13490 M^{-1} cm^{-1} taken as being equal to the original concentration in the aqueous phase. The solubility of camptothecin was plotted versus the concentration of the CD derivative (see Figure 1). The binding constant K was calculated from the zero-point solubility in the solvent $[G]_0$ 2.11 ± 0.24 μM and the slope B according to Equation (2).[53] Binding free energy was calculated according to $\Delta G° = -RT \ln K$.

$$K = \frac{B}{(1-B)[G]_0} \qquad (2)$$

Acknowledgements

We thank A. Engelke and J. Büch for technical support. Financial support from the Deutsche Forschungsgemeinschaft through the projects KA 1804/1 and AP 101/1 is gratefully acknowledged.

[1] T. Schrader, A. D. Hamilton, *Functional Synthetic Receptors*, Wiley-VCH, Weinheim, **2005**.
[2] G. V. Oshovsky, D. N. Reinhoudt, W. Verboom, *Angew. Chem.* **2007**, *119*, 2418; *Angew. Chem. Int. Ed.* **2007**, *46*, 2366.
[3] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, *J. Mol. Biol.* **1996**, *261*, 470.
[4] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727.
[5] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739.
[6] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639.
[7] T. J. A. Ewing, I. D. Kuntz, *J. Comput. Chem.* **1997**, *18*, 1175.
[8] N. Brooijmans, I. D. Kuntz, *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335.
[9] B. K. Shoichet, *Nature* **2004**, *432*, 862.
[10] G. Klebe, *Drug Discovery Today* **2006**, *11*, 580.

[11] M. R. de Jong, R. M. Knegtel, P. D. Grootenhuis, J. Huskens, D. N. Reinhoudt, *Angew. Chem.* **2002**, *114*, 1046; *Angew. Chem. Int. Ed.* **2002**, *41*, 1004.
[12] C. H. Takimoto, J. Wright, S. G. Arbuck, *Biochim. Biophys. Acta Gene Struct. Expression* **1998**, *1400*, 107.
[13] W. J. Slichenmyer, E. K. Rowinsky, R. C. Donehower, S. H. Kaufmann, *J. Natl. Cancer Inst.* **1993**, *85*, 271.
[14] B. B. Lundberg, *Anti-Cancer Drug Des.* **1998**, *13*, 453.
[15] T. G. Burke, A. E. Staubus, A. K. Mishra, H. Malak, *J. Am. Chem. Soc.* **1992**, *114*, 8318.
[16] B. Ertl, P. Platzer, M. Wirth, F. Gabor, *J. Controlled Release* **1999**, *61*, 305.
[17] S. M. Sugarman, Y. Y. Zou, K. Wasan, K. Poirot, R. Kumi, S. Reddy, R. PerezSoler, *Cancer Chemother. Pharmacol.* **1996**, *37*, 531.
[18] R. Cortesi, E. Esposito, A. Maietti, E. Menegatti, C. Nastruzzi, *Int. J. Pharm.* **1997**, *159*, 95.
[19] J. Kang, V. Kumar, D. Yang, P. R. Chowdhury, R. J. Hohl, *Eur. J. Pharm. Sci.* **2002**, *15*, 163.
[20] G. Wenz, *Angew. Chem.* **1994**, *106*, 851; *Angew. Chem. Int. Ed. Engl.Angew. Chem. Int. Ed.* **1994**, *33*, 803.
[21] M. V. Rekharsky, Y. Inoue, *Chem. Rev.* **1998**, *98*, 1875.
[22] M. E. Davis, M. E. Brewster, *Nat. Rev. Drug Discovery* **2004**, *3*, 1023.
[23] K. A. Connors, *Chem. Rev.* **1997**, *97*, 1325.
[24] A. Müller, G. Wenz, *Chem. Eur. J.* **2007**, *13*, 2218.
[25] G. Wenz, B. H. Han, A. Müller, *Chem. Rev.* **2006**, *106*, 782.
[26] G. Wenz, C. Gruber, B. Keller, C. Schilli, T. Albuzat, A. Müller, *Macromolecules* **2006**, *39*, 8021.
[27] T. Kitae, T. Nakayama, K. Kano, *J. Chem. Soc. Perkin Trans. 2* **1998**, 207–212.
[28] V. A. Karginov, S. M. Hecht, N. Fahmi, K. Aliben (Pinnacle Pharmaceuticals, USA; Innovative Biologics, Inc.; Advanced Biosystems, Inc.) WO 2006001844, **2006**, [*Chem. Abstr.* **2006** 144:82369].
[29] A. Kämper, J. Apostolakis, M. Rarey, C. M. Marian, T. Lengauer, *J. Chem. Inf. Model.* **2006**, *46*, 903.
[30] C. Betzel, W. Saenger, B. E. Hingerty, G. M. Brown, *J. Am. Chem. Soc.* **1984**, *106*, 7545.
[31] A. Steffen, A. Kämper, T. Lengauer, *J. Chem. Inf. Model.* **2006**, *46*, 1695.
[32] J. P. Rogers, A. E. Beuscher, M. Flajolet, T. McAvoy, A. C. Nairn, A. J. Olson, P. Greengard, *J. Med. Chem.* **2006**, *49*, 1658.
[33] D. A. Evans, S. Neidle, *J. Med. Chem.* **2006**, *49*, 4232.
[34] D. Q. Wei, R. Zhang, Q. S. Du, W. N. Gao, Y. Li, H. Gao, S. Q. Wang, X. Zhang, A. X. Li, S. Sirois, K. C. Chou, *Amino Acids* **2006**, *31*, 73.
[35] a) M. Karasz, R. Koerner, M. Marialke, S. Tietze, J. Apostolakis, *Proceedings of the 15th European Symposium on Quantitative Structure Activity Relationships & Molecular Modeling* (Turkey), **2004**, p. 493; b) S. Tietze, J. Apostolakis, *J. Chem. Inf. Model.*, DOI:10.1021/ci7001236.
[36] R. Abagyan, M. Totrov, D. Kuznetsov, *J. Comput. Chem.* **1994**, *15*, 488.
[37] J. Apostolakis, A. Pluckthun, A. Caflisch, *J. Comput. Chem.* **1998**, *19*, 21.
[38] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, R. P. Mee, *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425.
[39] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, R. D. Taylor, *Proteins: Struct. Funct. Genet.* **2003**, *52*, 609.
[40] E. Kellenberger, J. Rodrigo, P. Muller, D. Rognan, *Proteins: Struct. Funct. Bioinf.* **2004**, *57*, 225.
[41] A. R. Katritzky, D. C. Fara, H. Yang, M. Karelson, T. Suzuki, V. P. Solov'ev, A. Varnek, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 529–541.
[42] F. H. Allen, *Acta. Crystallogr. Sect. B.* **2002**, *58*, 380.
[43] A. T. McPhail, G. A. Sim, *J. Chem. Soc. B* **1968**, 923.
[44] T. A. Halgren, *J. Comput. Chem.* **1999**, *20*, 730.
[45] V. Zabel, W. Saenger, S. A. Mason, *J. Am. Chem. Soc.* **1986**, *108*, 3664.
[46] J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Comput. Sci.* **2005**, *45*, 177.
[47] J. Sadowski, J. Gasteiger, *Chem. Rev.* **1993**, *93*, 2567.
[48] M. F. Sanner, *J. Mol. Graphics Modell.* **1999**, *17*, 57.
[49] A. S. Deo, I. D. Walker, *J. Intell. Robot. Syst.* **1995**, *14*, 43.
[50] C. Bystroff, *Protein Eng.* **2001**, *14*, 825.
[51] P. R. Ashton, R. Königer, J. F. Stoddart, D. Alker, V. D. Harding, *J. Org. Chem.* **1996**, *61*, 903.
[52] J. Defaye, S. Crouzy, N. Evrard, H. Law (Centre National De La Recherche Scientifique, Fr.), WO 9961483, **1999** [*Chem. Abstr.* **1999** 132:24077].
[53] T. Higuchi, J. L. Lach, *J. Am. Pharm. Assoc.* **1954**, *43*, 349.
[54] F. Guillo, B. Hamelin, L. Jullien, J. Canceill, J. M. Lehn, L. Derobertis, H. Driguez, *Bull. Soc. Chim. Fr.* **1995**, *132*, 857.
[55] B. Ekberg, L. I. Andersson, K. Mosbach, *Carbohydr. Res.* **1989**, *192*, 111.

Received: May 2, 2007
Published online: July 3, 2007

