[5] a) S. J. Holder, R. C. Hiorns, N. A. J. M. Sommerdijk, S. J. Williams, R. G. Jones, R. J. M. Nolte, *Chem. Commun.* **1998**, 1445; b) N. A. J. M. Sommerdijk, S. J. Holder, R. C. Hiorns, R. G. Jones, R. J. M. Nolte, *Macromolecules* **2000**, *33*, 8289.
[6] B. M. Discher, Y. Won, D. S. Ege, J. C-M. Lee, F. S. Bates, D. E. Discher, D. A. Hammer, *Science* **1999**, *284*, 1143.
[7] a) J. Zipfel, P. Lindner, M. Tsianou, P. Alexandridis, W. Richtering, *Langmuir* **1999**, *15*, 2599; b) J. Zipfel, J. Berghausen, G. Schmidt, P. Lindner, P. Alexandridis, M. Tsianou, W. Richtering, *Phys. Chem. Chem. Phys.* **1999**, *1*, 3905.
[8] K. Schillén, K. Bryskhe, Y. S. Mel'nikova, *Macromolecules* **1999**, *32*, 6885.
[9] a) C. Nardin, T. Hirt, J. Leukel, W. Meier, *Langmuir* **2000**, *16*, 1035; b) C. Nardin, S. Thoeni, J. Widmer, M. Winterhalter, W. Meier, *Chem. Commun.* **2000**, 1433.
[10] M. Maskos, J. R. Harris, *Macromol. Rapid Commun.* **2001**, *22*, 271.
[11] D. D. Lasic, *Liposomes: From Physics to Applications*, Elsevier, Amsterdam, **1993**.
[12] U. Seifert, *Adv. Phys.* **1997**, *46*, 13.
[13] R. G. Laughlin, *The Aqueous Phase Behavior of Surfactants*, Academic Press, San Diego, **1994**.
[14] H. Hoffmann, W. Ulbricht, *Recent Res. Dev. Phys. Chem.* **1998**, *2*, 113.
[15] A. V. Kabanov, T. K. Bronich, V. A. Kabanov, K. Yu, A. Eisenberg, *J. Am. Chem. Soc.* **1998**, *120*, 9941.
[16] M. A. Cohen Stuart, N. A. M. Besseling, R. G. Fokkink, *Langmuir* **1998**, *14*, 6846.
[17] F. Ilhan, T. H. Galow, M. Gray, G. Clavier, V. M. Rotello, *J. Am. Chem. Soc.* **2000**, *122*, 5895.
[18] S. Zhou, C. Burger, B. Chu, M. Sawamura, N. Nagahama, M. Toganoh, U. E. Hackler, H. Isobe, E. Nakamura, *Science* **2001**, *291*, 1944.
[19] a) E. W. Kaler, A. K. Murthy, B. E. Rodriguez, J. A. N. Zasadzinski, *Science* **1989**, *245*, 1371; b) E. W. Kaler, K. L. Herrington, A. K. Murthy, J. A. Zasadzinski, *J. Phys. Chem.* **1992**, *96*, 6698; c) K. L. Herrington, E. W. Kaler, D. D. Miller, J. A. Zasadzinski, S. Chiruvolu, *J. Phys. Chem.* **1993**, *97*, 13792; d) M. T. Yatcilla, K. L. Herrington, L. L. Brasher, E. W. Kaler, S. Chiruvolu, J. A. Zasadzinski, *J. Phys. Chem.* **1996**, *100*, 5874; e) H. T. Jung, B. Coldren, J. A. Zasadzinski, D. J. Iampietro, E. W. Kaler, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1353.
[20] S. A. Safran, P. Pincus, D. Andelman, *Science* **1990**, *248*, 354.
[21] a) H. J. Watzke, *Prog. Colloid Polym. Sci.* **1993**, *93*, 15; b) O. Regev, A. Khan, *J. Colloid Interface Sci.* **1996**, *182*, 95.
[22] a) E. F. Marques, O. Regev, A. Khan, M. G. Miguel, B. Lindman, *J. Phys. Chem. B* **1998**, *102*, 6746; b) E. F. Marques, O. Regev, A. Khan, M. G. Miguel, B. Lindman, *J. Phys. Chem. B* **1999**, *103*, 8353; c) E. F. Marques, *Langmuir* **2000**, *16*, 4798.
[23] a) R. Joannic, L. Auvray, D. D. Lasic, *Phys. Rev. Lett.* **1997**, *78*, 3402; b) I. Szleifer, O. V. Gerasimov, D. H. Thompson, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1032.
[24] L. Luo, A. Eisenberg, *J. Am. Chem. Soc.* **2001**, *123*, 1012.
[25] L. Luo, A. Eisenberg, *Langmuir* **2001**, *17*, 6804.
[26] a) L. Zhang, A. Eisenberg, *Polym. Adv. Technol.* **1998**, *9*, 677; b) L. Zhang, A. Eisenberg, *Macromolecules* **1999**, *32*, 2239.
[27] R. W. O'Brien, R. J. Hunter, *Can. J. Chem.* **1981**, *59*, 1878.
[28] P. G. Hartley, P. J. Scales, *Langmuir* **1998**, *14*, 6948.
[29] M. Satoh, E. Yoda, T. Hayashi, J. Komiyama, *Macromolecules* **1989**, *22*, 1808.
[30] a) E. Uchida, Y. Uyama, Y. Ikada, *Langmuir* **1994**, *10*, 1193; b) H.-J. Jacobasch, F. Simon, P. Weidenhammer, *Colloid Polym. Sci.* **1998**, *276*, 434; c) K. Emoto, Y. Nagasaki, K. Kataoka, *Langmuir* **1999**, *15*, 5212.
[31] O. E. Philippova, D. Hourdet, R. Audebert, A. R. Khokhlov, *Macromolecules* **1997**, *30*, 8278.
[32] a) J. R. Lakowicz, *Principle of Fluorescence Spectroscopy*, Plenum, New York, **1983**; b) T. Cao, W. Yin, J. L. Armstrong, S. E. Webber, *Langmuir* **1994**, *10*, 1841; c) T. Kawamoto, Y. Morishima, *Langmuir* **1998**, *14*, 6669.

A Method To Identify and Screen Libraries of Guests That Complex to a Synthetic Host**

Menno R. de Jong, Ronald M. A. Knegtel, Peter D. J. Grootenhuis, Jurriaan Huskens,* and David N. Reinhoudt*

In supramolecular host–guest chemistry, the de novo design of a host molecule often does not lead to selectivity for the desired guest.[1] Alternatively, host–guest recognition can be reformulated from the perspective of a known host molecule that requires the identification of suitable guest molecules. This would allow the application of existing host molecules in other fields, for example, for sensing other guest molecules, or for carrying other drugs in artificial drug-carrier systems. The identification of guests is a well-known problem in medicinal chemistry, where lead discovery aims at the identification of possible binders to an enzyme, for example. In this field, one approach to overcome the limitations of de novo design is the development of rapid (virtual) screening techniques.

Docking is a computational method often used to identify new leads or suggest possible binding modes of known guests.[2] When used for the identification of leads, docking starts from a simplified representation of the binding pocket of a protein or receptor, for example, one derived from a crystal structure. Subsequently, a large number of potential ligands are fitted into this binding site and the docking algorithm finds guests that yield a favorable interaction energy with the host. The success of this approach in identifying potential binders for a receptor inspired us to investigate whether it would also allow the rapid screening of large numbers of guests for synthetic receptors.

In supramolecular chemistry, cyclodextrins (Figure 1a) are often used because of their unique ability to complex a variety of small organic guests in water.[3] Previously, we reported the binding properties of cyclodextrin dimer **1** (Figure 1b) with cholates and the fluorescence-signaling behavior of a closely related dansyl-appended dimer.[4] Recently, a number of cyclodextrin dimers have been reported and these often exhibit enhanced binding with suitable guests,[5] some of them as stable as enzyme–ligand interactions. This renders these

[*] Dr. Ir. J. Huskens, Prof. Dr. Ir. D. N. Reinhoudt, Dr. M. R. de Jong, Prof. Dr. P. D. J. Grootenhuis
Laboratory of Supramolecular Chemistry and Technology
MESA+ Research Institute, University of Twente
PO Box 217, 7500 AE Enschede (Netherlands)
Fax: (+31) 53-489-4645
E-mail: SMCT@ct.utwente.nl
Dr. R. M. A. Knegtel
Vertex Pharmaceuticals (Europe) Ltd.
88 Milton Park, Abingdon, Oxfordshire OX14 4RY
(United Kingdom)

[**] We are grateful to the Council for Chemical Sciences of the Netherlands Organization for Scientific Research (NWO-CW) and the Technology Foundation STW for financial support in the framework of the CW/STW program Technical Chemistry (project number 349-4213, MRdJ).

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

  1433-7851/02/4106-1004 $ 17.50+.50/0

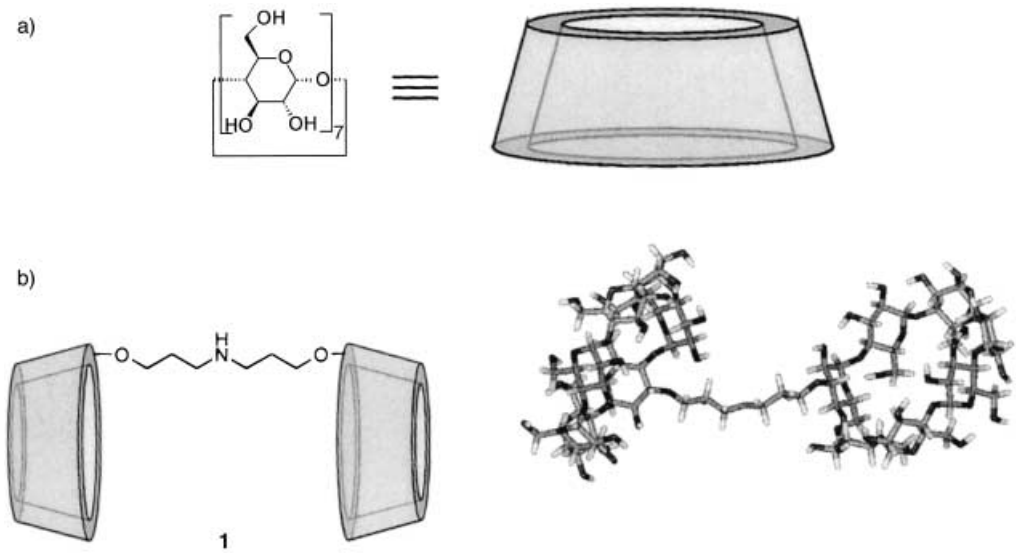


Figure 1. a) Structure of β-cyclodextrin; b) schematic representation (left) and energy-minimized structure (right) of cyclodextrin dimer **1**.

cyclodextrin derivatives interesting candidates for comparing the computational and experimental assessment of binding properties. Computational methods have been used to identify binding modes for known guests into cyclodextrins.[6] In the present study we have performed a docking procedure with one synthetic host and many guest molecules, and have subsequently tested its success score by experimental verification of the complexation strength by using a facile and rapid fluorescence assay.

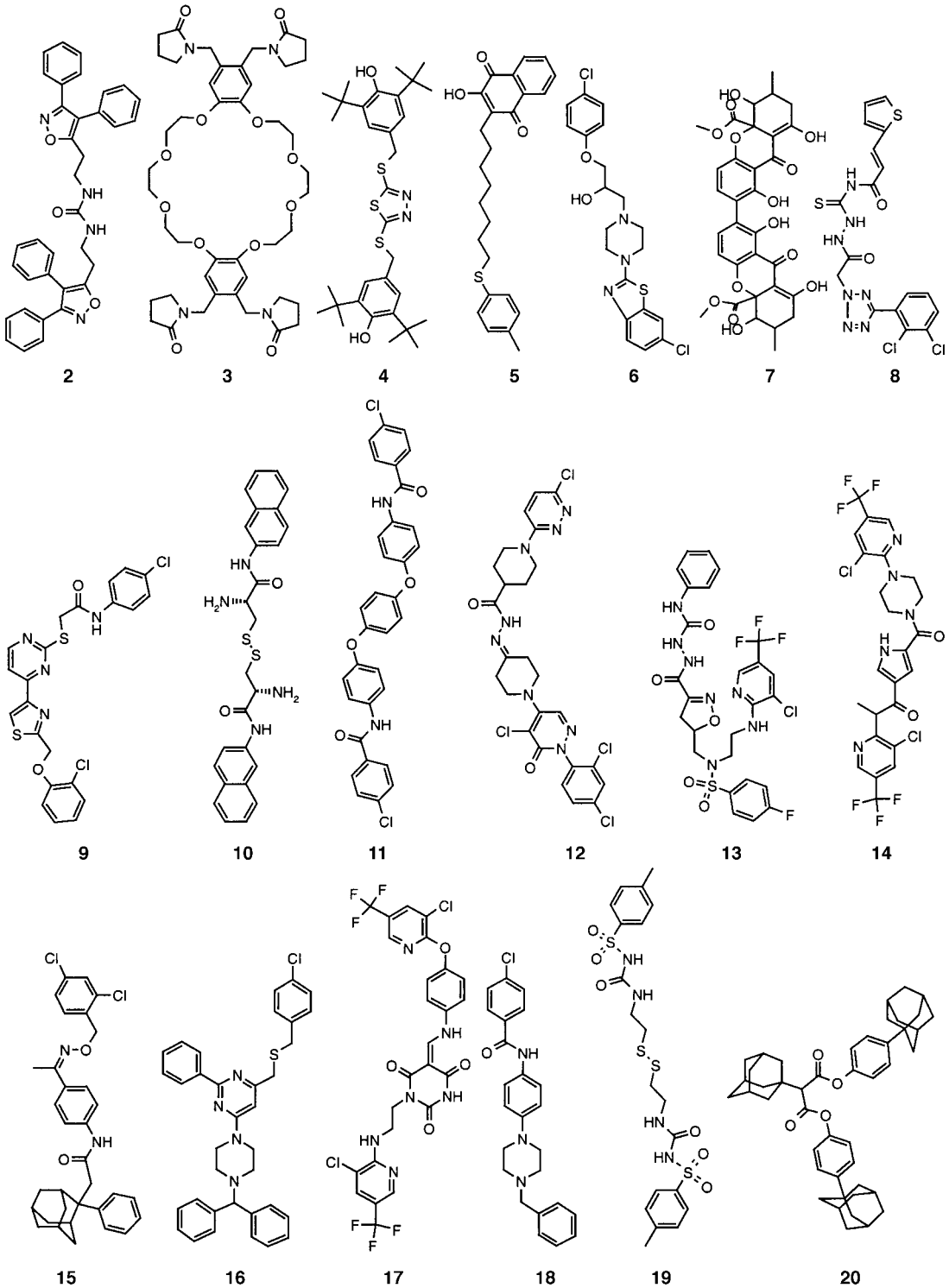
As there is no crystal structure of the cyclodextrin dimer (**1**), its conformation used in the docking procedure was generated by a molecular mechanics study of **1** based on a crystal structure of native β-cyclodextrin and an extended conformation of the flexible alkyl spacer (Figure 1b).[7] This conformation was used for docking with the Available Chemicals Database (ACD) and the Comprehensive Medicinal Chemistry Database (CMCD). Ligands with molecular weights below 200 D and compounds with more than two formal charges or reactive functionalities were excluded, and the remaining $\sim 10^5$ molecules from the ACD and $\sim 10^4$ molecules from the CMCD were docked in **1**. The highest ranking structures, about 500 from the ACD and 200 from the CMCD, were visually inspected. Among the highest ranking docking solutions for molecules from the ACD, several are of sufficient length to position hydrophobic functionalities in both β-cyclodextrin cavities. In particular, phenyl, cyclohexyl, adamantyl, *tert*-butyl, and complex aliphatic ring systems (steroid- or opiate-like) were found to be located in the β-cyclodextrin ring interiors. In some cases, additional hydrogen bonding to hydroxy groups of the cyclodextrin rings was also suggested. The interaction with the linker moiety was usually minimal.

For the majority of ligands retrieved from the CMCD, binding to both cyclodextrins appeared to be less common. This is to be expected since the CMCD contains drug molecules that typically have relatively low molecular weights (<600) and corresponding small molecular sizes. Although traditionally applied in virtual screening of protein guests,[2] molecular database docking appears to perform well as a computational tool to identify small molecule guests for synthetic receptors. After visual inspection of docking hits, a total of 93 ligands from the ACD and 125 ligands from the CMCD were selected for experimental evaluation (subject to commercial availability). A selection of these hits had to be made since the experimental evaluation not only required commercial availability, but also high purity of the compound and solubility in water (or at least in methanol). Many compounds from the CMCD are not commercially available in pure form. Thus, this number was reduced to 30 guests: 19 from the ACD (Scheme 1), and 11 from the CMCD (Scheme 2).[8]

For a rapid indication as to whether a guest binds or not, we performed a fluorescence competition assay in which the guests to be screened compete for binding to **1** with a fluorophore of known binding strength. We employed the polarity probe 2-*p*-toluidinylnaphthalene-6-sulfonate (TNS), the fluorescence intensity of which increases strongly upon binding to cyclodextrin or cyclodextrin dimers.[5e, 9] Its binding constant to **1** ($5.3 \times 10^4\,\text{M}^{-1}$) was determined by means of a fluorescence titration. The binding constants of most known guests in native β-cyclodextrin are in the range of $10^2 - 10^4\,\text{M}^{-1}$; only a few stronger binding guests are known. Therefore, a threshold value of $10^5\,\text{M}^{-1}$ for the association constant was taken as an indication that both cavities cooperate in guest binding. The screening was performed by the addition of a concentrated methanolic solution of the competing guest to an aqueous solution of **1** and TNS.[10] The results are summarized in Figure 2.

We found that three guests from the ACD (**5**, **8**, and **13**) and six from the CMCD (**21**–**26**) bind strongly to dimer **1**, since competition fluorescence titrations revealed association constants $> 10^5\,\text{M}^{-1}$.[11] Given the absence of terms that account for entropic effects from the force field score applied in this study and the use of rigid molecular structures, little correlation is expected between predicted binding enthalpies and measured binding constants. The correlation between predicted and

  1433-7851/02/4106-1005 $ 17.50+.50/0



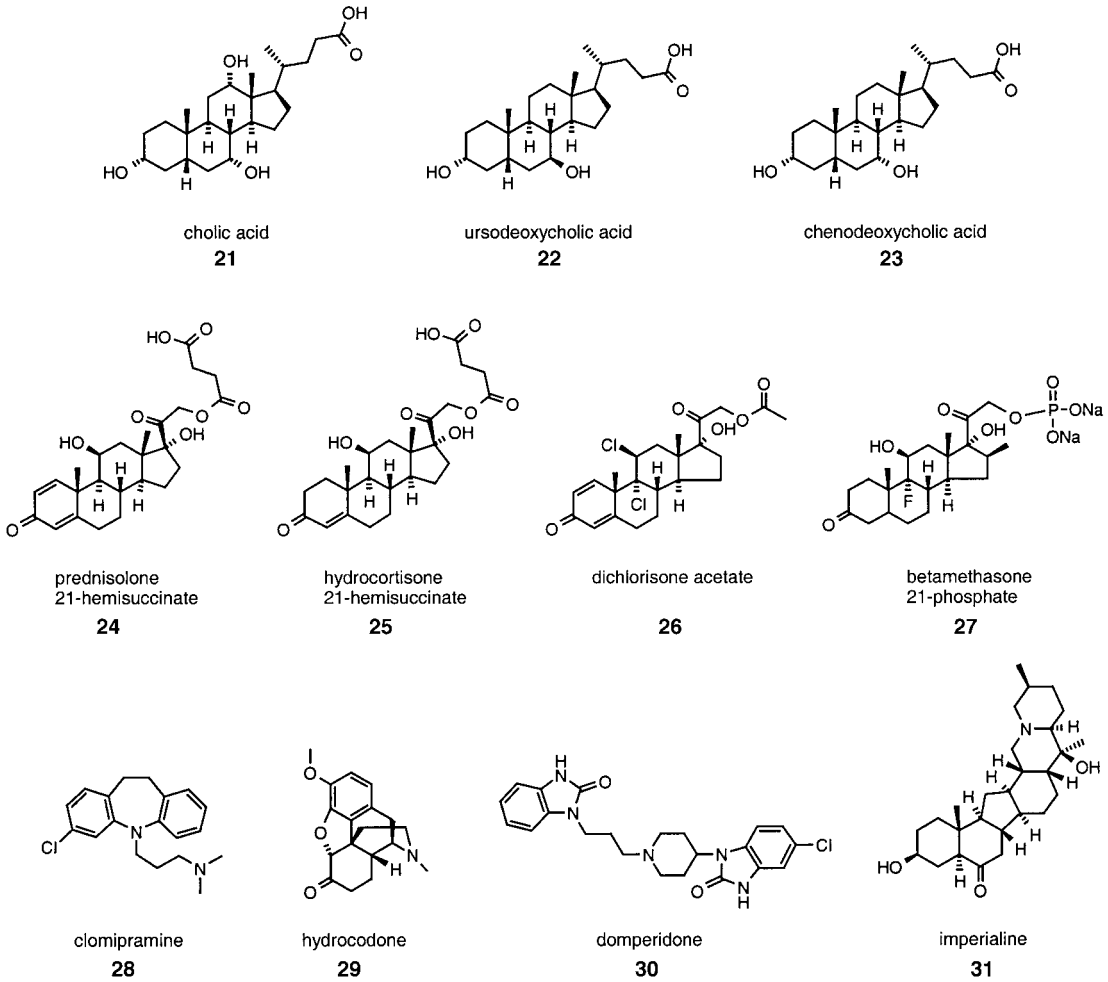
Scheme 1. Compounds from the Available Chemicals Database (ACD) selected for screening.

measured binding constants also remains a problem for molecular docking applications in medicinal chemistry.[12] In principle, however, this does not prohibit virtual screening from identifying suitable ligands for receptors.

The hit rate of 9 out of 30 (30%) compares very favorably to the results obtained with docking of ligands to bioreceptors, for which hit rates are usually $<10\%$. The favorable hit rate obtained when using a cyclodextrin receptor with respect to similar applications in medicinal chemistry can be explained from the relative structural simplicity of the synthetic receptor.[13] A summary of the various stages that led to the identification of the nine strongly bound guests is shown in Figure 3.

Molecular docking to synthetic receptors is probably subject to similar sources of error as virtual screening of molecule databases with biological receptors.[14] The scoring

  1433-7851/02/4106-1006 $ 17.50+.50/0



Scheme 2. Compounds from the Comprehensive Medicinal Chemistry Database (CMCD) selected for screening.

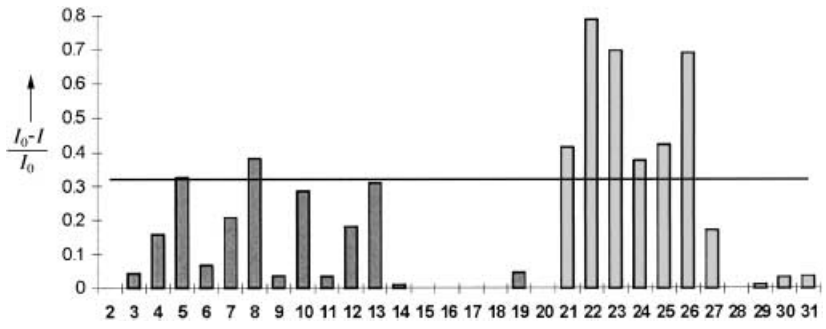


Figure 2. Overview of the screening of guests found by the docking procedure. The solid line represents the calculated response for a K value of the competing guest of $10^5\,\text{M}^{-1}$.

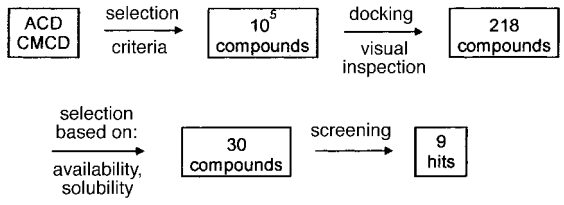


Figure 3. Stages of the docking procedure.

function, the computational implementation, and the degree of conformational sampling of ligands and receptor influence the results. However, it is encouraging that a relatively simple protocol, which involves rigid ligand conformations and force-field scoring without correction for entropic effects, is able to identify novel ligands for cyclodextrin-derived receptors with such an impressive degree of accuracy.

This work demonstrates the potential for virtual screening methodologies outside the confines of pharmaceutical lead identification. In future, similar computational procedures as the one described here may also provide a useful tool in the design of synthetic receptors for a given guest.

Experimental Section

Computational methodology: To identify small organic molecules that bind both cyclodextrin cavities of the dimer, the molecular docking program DOCK 4.0[15] was applied. Two chemical databases supplied by MDL Information Systems Inc. (San Leandro, USA) were considered for docking. The Available Chemicals Database (ACD, version 2000.1) provides a source of commercially available small organic molecules, while the Comprehensive Medicinal Chemistry Database (CMCD, version 98.1) is a source of currently marketed drugs and other molecules of medical importance. Three-dimensional structures for potential ligands from the ACD and CMCD were generated by using the CORINA program[16] (version 1.81). The linear shape of the dual cyclodextrin receptor molecule

suggests preferred binding of molecules in an extended conformation. It was therefore decided to use directly the extended ligand conformations commonly generated by the CORINA program without allowing for additional ligand flexibility during docking. After filtering out ligands with molecular weights below 200, more than two formal charges, and reactive functionalities, a total of 70867 molecules from the ACD and 7497 molecules from the CMCD were docked against the β-cyclodextrin dimer.

Docking of rigid ligand molecules involves their superposition onto site points in the receptor molecule. Docking site points, covering the binding pockets of both cyclodextrin host molecules, were generated with SPHGEN[17] and limited to a total of 25 by means of visual inspection with Quanta 98 (Molecular Simulations Inc., San Diego, USA). The highest ranking molecules were visually inspected by using Quanta 98 and web tools that apply the Chime plug-in (MDL, San Leandro USA). For the ACD, three sets of the 500 best scoring molecules with total charges of 0, 1, and 2 were visually inspected. For the smaller CMCD, the total number of molecules visually inspected was 200. Since the force-field-based scoring function used by DOCK 4.0 does not take entropic contributions to host–guest binding into account (i.e. solvation, flexibility) further manual filtering was applied to the docked solutions. Visual inspection implied removal of docking solutions that placed charged or polar functionality in the hydrophobic interior of the cyclodextrin ring or exposed large hydrophobic groups to the solvent. Furthermore, duplicates (i.e. different salts of the same compound) were removed and molecules that bind to both cyclodextrin units of the dimer were preferred over smaller molecules. Those ligands that display good complementarity to the cyclodextrin dimer were selected for testing. More information on the calculations is given in the Supporting Information.

Received: September 3, 2001 [Z17844]

[1] P. Timmerman, K. G. A. Nierop, E. A. Brinks, W. Verboom, F. C. J. M. van Veggel, W. P. van Hoorn, D. N. Reinhoudt, *Chem. Eur. J.* **1995**, *1*, 132–143.

[2] a) D. Kuntz, *Science* **1992**, *257*, 1078–1082; b) P. A. Walters, M. T. Stahl, M. A. Murcko, *Drug Discovery Today* **1998**, *3*, 160–178; c) H. Kubinyi, *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 16–27; d) C. J. Manly, S. Louise-May, J. D. Hammer, *Drug Discovery Today* **2001**, *6*, 1101–1110, and references therein.

[3] a) G. Wenz, *Angew. Chem.* **1994**, *106*, 851–888; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 803–822; b) J. Szejtli, T. Osa, *Comprehensive Supramolecular Chemistry*, *Vol. 3* (Eds.: J. L. Atwood, J. E. D. Davies, D. D. MacNicol, F. Vögtle, J.-M. Lehn), Pergamon, Oxford, **1996**.

[4] M. R. De Jong, J. F. J. Engbersen, J. Huskens, D. N. Reinhoudt, *Chem. Eur. J.* **2000**, *6*, 4034–4040.

[5] a) K. Fujita, S. Ejima, T. Imoto, *Chem. Commun.* **1984**, 127–128; b) R. Breslow, S. Chung, *J. Am. Chem. Soc.* **1990**, *112*, 9659–9660; c) R. Breslow, B. L. Zhang, *J. Am. Chem. Soc.* **1996**, *118*, 8495–8496; d) F. Venema, A. E. Rowan, R. J. M. Nolte, *J. Am. Chem. Soc.* **1996**, *118*, 257–258; e) F. Venema, H. F. M. Nelissen, P. Berthault, N. Birlirakis, A. E. Rowan, M. C. Feiters, R. J. M. Nolte, *Chem. Eur. J.* **1998**, *4*, 2237–2250.

[6] a) F. M. Menger, M. J. Sherrod, *J. Am. Chem. Soc.* **1988**, *110*, 8606–8611; b) V. Tran, M. M. Delage, A. Buleon, *J. Inclusion Phenom. Mol. Recognit. Chem.* **1993**, *14*, 271–284; c) F. Melani, N. Mulinacci, A. Romani, G. Mazzi, F. F. Vinceri, *Int. J. Pharm.* **1998**, *166*, 145–155; d) S. Ishikawa, S. Hada, S. Neya, N. Funasaki, *J. Phys. Chem. B* **1999**, *103*, 1208–1215.

[7] The MM calculation was run with the CHARMM 22.0 and 23.0 programs with Quanta 97 (Molecular Simulations Inc., Burlington, MA, USA). Parameters were taken from Quanta 97 and point charges were assigned with the charge template option in Quanta/CHARMM. The structure of β-cyclodextrin was taken from the Cambridge Crystallographic Database. Two cyclodextrin units were manually connected through a spacer in an *all-trans* conformation. CHARMM: a) B. R. Brooks, R. E. Bruccoleri, B. D. Olafsen, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187–217; b) F. A. Momany, V. J. Klimkowski, L. Schäfer, *J. Comput. Chem.* **1990**, *11*, 654–662; c) F. A. Momany, R. Rone, H. Kunz, R. F. Frey, S. Q. Newton, L. Schäfer, *J. Mol. Struct. (THEOCHEM)* **1993**, *105*, 1–18.

[8] The relatively large number of steroids found in the selection of the CMCD stems from the combination of a relatively large availability of this class of compounds in this database (5.9% vs. 1.7% in the ACD) and the lack of commercial sources for the other classes. In contrast, the selection from the ACD generally shows more symmetrical structures that may be expected for this dimeric host, also because the ACD has larger molecules that have a greater chance of interacting with both cavities in the docking study, thus reducing the probability of picking up the smaller steroids from the ACD.

[9] Y. Liu, C. T. Wu, G. P. Xue, J. Li, *J. Inclusion Phenom. Macrocyclic Chem.* **2000**, *36*, 95–100.

[10] From the independently calculated binding constant of TNS ($K = 5.3 \times 10^4\,\text{M}^{-1}$) in dimer **1** and the fluorescence intensities of bound and unbound TNS, a host concentration of 2 μM and concentrations of 10 μM for TNS and the guest to be screened were calculated to be suitable for identifying guests with binding constants $>10^5\,\text{M}^{-1}$. Fluorescence spectra were recorded before and after adding a solution of the guest (6.1×10^{-4} M, 50 μL) in methanol to an aqueous solution (3 mL) of **1** (2.0×10^{-6} M) and TNS (1.0×10^{-5} M).

[11] Fluorescence competition titrations were performed by adding aliquots of a solution of the guest (6.1×10^{-4} M) in methanol to an aqueous solution (3 mL) of **1** (2.0×10^{-6} M) and TNS (1.0×10^{-5} M) finally reaching a maximum of 3% v/v of methanol. After each addition, the fluorescence spectrum was recorded. The binding constants for **21**, **22**, and **23** have been reported before.[4] The others are: **5**: $1.0 \times 10^5\,\text{M}^{-1}$; **8**: $1.9 \times 10^5\,\text{M}^{-1}$; **13**: $1.6 \times 10^6\,\text{M}^{-1}$; **24**: $3.4 \times 10^5\,\text{M}^{-1}$; **25**: $4.3 \times 10^5\,\text{M}^{-1}$; **26**: $3.6 \times 10^6\,\text{M}^{-1}$.

[12] R. M. A. Knegtel, P. D. J. Grootenhuis, *Perspect. Drug Discovery Des.* **1998**, *9–11*, 99–114.

[13] Finding suitable inhibitors for enzymes requires optimal matching of shape and polarity between a ligand and a complex protein surface. In the case of the cyclodextrin dimer used in this study, the receptor has a simple concave shape with a predominantly hydrophobic inner surface while the entire dimer is in an optimally accessible, extended conformation. This essentially reduces the docking problem to finding ligands of the correct shape and hydrophobicity to fit inside one or both cyclodextrin cavities. Therefore, a simple force field scoring method can be expected to produce reasonable results. Additional manual filtering is still required, as it would be in applications related to drug discovery, to remove those complexes in which the force field score is dominated by an overestimation of electrostatic interactions.

[14] a) T. Lengauer, M. Rarey, *Curr. Opin. Struct. Biol.* **1996**, *6*, 402–406; b) R. M. A. Knegtel, M. Wagener, *Proteins* **1999**, *37*, 334–345.

[15] T. J. A. Ewing, I. D. Kuntz, *J. Comput. Chem.* **1997**, *18*, 1175–1189.

[16] J. Sadowski, J. Gasteiger, G. Klebe, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000–1008.

[17] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. L. Langridge, T. E. Ferrin, *J. Mol. Biol.* **1982**, *161*, 269–288.

  1433-7851/02/4106-1008 $ 17.50+.50/0