Teodorovic, M. V. Rangaishenvi, H. C. Brown, *J. Org. Chem.* **1992**, *57*, 2379 – 2386; c) K. A. Parker, M. W. Ledeboer, *J. Org. Chem.* **1996**, *61*, 3214 – 3217; d) E. J. Corey, R. K. Bakshi, S. Shibata, *J. Am. Chem. Soc.* **1987**, *109*, 5551 – 5553; e) J. Bach, R. Berenguer, J. Garcia, T. Loscertales, J. Vilarrasa, *J. Org. Chem.* **1996**, *61*, 9021 – 9025; f) K. Matsumura, S. Hashiguchi, T. Ikariya, R. Noyori, *J. Am. Chem. Soc.* **1997**, *119*, 8738 – 8739.
[12] J. K. Stille, B. L. Groh, *J. Am. Chem. Soc.* **1987**, *109*, 813 – 817.
[13] W. C. Still, C. Gennari, *Tetrahedron Lett.* **1983**, *24*, 4405 – 4408.
[14] D. A. Evans, J. A. Bartroli, T. L. Shih, *J. Am. Chem. Soc.* **1981**, *103*, 2127 – 2129.
[15] D. A. Evans, S. L. Bender, J. Morris, *J. Am. Chem. Soc.* **1988**, *110*, 2506 – 2526.
[16] T. D. Penning, S. D. Djuric, R. A. Haack, V. J. Kalish, J. M. Myashiro, B. W. Rowell, S. S. Yu, *Synth. Commun.* **1990**, *20*, 307 – 312.
[17] D. B. Dess, J. C. Martin, *J. Am. Chem. Soc.* **1991**, *113*, 7277 – 7287.
[18] A. S.-Y. Lee, Y.-J. Hu, S.-F. Chu, *Tetrahedron* **2001**, *57*, 2121 – 2126.
[19] Y. Watanabe, T. Nakamura, H. Mitsumoto, *Tetrahedron Lett.* **1997**, *38*, 7407 – 7410.
[20] For example, see: J.-M. Duffault, J. Einhorn, A. Alexakis, *Tetrahedron Lett.* **1991**, *32*, 3701 – 3704.

Modular Building Blocks for Amino Acid Recognition in Peptides

Mark Wehner and Thomas Schrader*

The selective recognition of short peptide sequences is a key concept in many natural regulatory processes. Thus, the natural antibiotic vancomycin binds tightly to the C-terminal D-Ala-D-Ala fragment in peptides, which is used for the construction for bacterial cell walls.[1] Numerous cell – cell recognition events rely on the recognition of the specific Arg-Gly-Asp sequence.[2]

The first attempts to mimic Nature's potent peptide hosts with artificial structures, confined the conformational freedom of a host molecule by creating a cleft (Rebek and co-workers), a macrocycle (Still and co-workers), or even a cavity (Still and co-workers); the specific binding sites were often taken from (non)natural amino acids.[3] Several groups have created receptor molecules for important secondary structures found in peptides and proteins. Thus, α-helical[4] and β-sheet[5] portions of polypeptides can be recognized by synthetic ligands with a complementary hydrogen-bond-donor and acceptor pattern. In recent years, considerable progress has been achieved with a combinatorial approach.[6]

Various site-specific proteases, such as thrombin, trypsin, and many others provide a shallow groove for the efficient recognition of the backbone of the peptide to be cleaved, in its extended conformation. This array is combined with a specific binding pocket for the side chains of the target amino acid, and thus defines the cleavage site. We asked ourselves if such a rational design could also be used for artificial peptide recognition. Is it possible to create a set of modules, each of which recognizes a certain amino acid in a peptidic environment or, even better, a short peptide sequence?

We recently reported the stabilization of small peptides in their β-sheet conformation by external ligands.[7] N-acylated 3-aminopyrazoles were shown to interact with every hydrogen-bond-donor (D) and acceptor (A) available at the top face of a dipeptide. Two consecutive amino acids can be clamped together with one of these heterocycles, which causes the formation of three almost-linear hydrogen bonds in a DAD sequence (Figure 1). This recognition process simultaneously fixes the peptide in the thermodynamically favorable β-sheet conformation. Thus, small soluble models of this important secondary structure have been prepared.

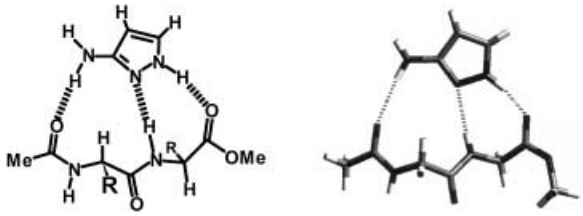


Figure 1. β-Sheet stabilization with aminopyrazoles—three-point binding of the top face of a dipeptide by the DAD binding sites of the aminopyrazole; left: Lewis structure; right: results of molecular-mechanics calculation (Cerius², Molecular Simulations, Dreiding 2.21).

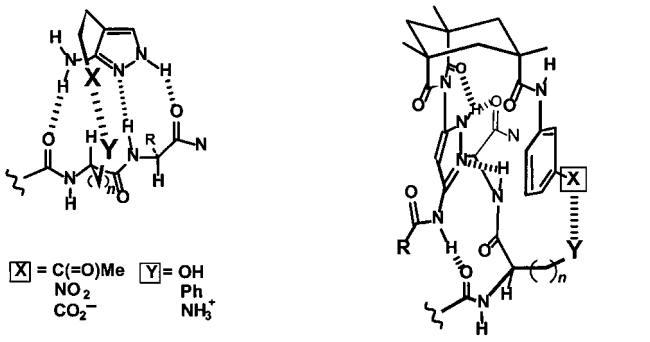
In a β sheet, the N–H and C=O bonds of the backbone point up and down, whereas the amino acid residues (R) are extended horizontally away from the peptide in a well-defined geometry. This preorganized arrangement was our starting point for a new modular concept of peptide receptors: If it were possible to attach to the aminopyrazole a rigid U-shaped substituent with a properly placed binding site at its tip, this could reach down to the side chain of the respective amino acid and lead to an additional, specific noncovalent interaction. With an interchangeable tip, various binding sites could be introduced into the basic peptide receptor, which would lead to a modular set of building blocks, selective for the typical classes of amino acids. To our knowledge, no example of such a rationally designed set of peptide receptors exists to date.

For the U-shaped spacer, we chose the framework of Kemp's triacid, because it is exceptionally rigid and its chemistry is well developed.[8] Two acid groups were designated to carry the monoacylated 3,5-diaminopyrazole unit by way of an imide functionality, while the third could be coupled to an aniline derivative with the correct binding site in the *m*-position. Scheme 1 shows the general structure of the building blocks: both the imide and the neighboring aminopyrazole unit are in the same plane, locked together by an intramolecular hydrogen bond.

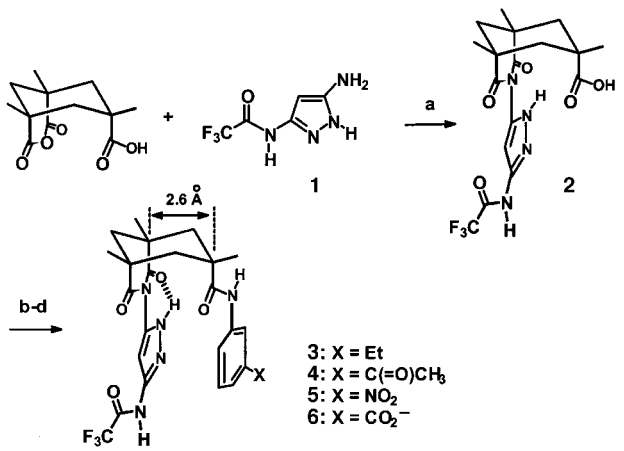
During the synthesis, care must be taken to chemoselectively address the three amino functionalities of 3,5-diaminopyrazole, without an extensive use of protecting groups. We begin with imide formation between mono(trifluoroacetylated) diaminopyrazole **1** and Kemp's acid anhydride (Scheme 2). The imide intermediate **2** must be *N*-Boc pro-

[*] Prof. Dr. T. Schrader, M. Wehner
Fachbereich Chemie der Universität
Hans Meerwein Strasse
35032 Marburg (Germany)
Fax: (+49) 6421-28-28917
E-mail: schradet@mailer.uni-marburg.de

  1433-7851/02/4110-1751 $ 20.00+.50/0



Scheme 1. Representation of the modular building blocks for peptide recognition; right: the U-turn formed by Kemp's triacid brings the receptor tip X close to the characteristic group Y of the amino acid in the peptide.



Scheme 2. Synthesis of hosts **3**–**6**; a) toluene, 110 °C (−H_2O); b) Boc_2O, 90 °C; c) *m*-aminobenzene derivative, PyCloP, 50 °C, $CHCl_3$; d) silica gel, (−Boc); a)–d) 17–22 % overall yield, Boc = *tert*-butoxycarbonyl, PyCloP = chlorotris(pyrrolidin-1-yl)phosphonium hexafluorophosphate.

tected for the subsequent attachment of the aniline derivative. The most critical step involves introduction of the aniline derivative carrying the binding site. Although this only requires formation of an amide bond, the product of this coupling reaction contains two aromatic rings at a distance of 2.6 Å and less. We were successful with PyCloP, but only at 50 °C in chloroform. The Boc group was removed during the subsequent chromatographic purification, which leads directly to the desired products **3**–**6**.

Following this synthetic method, *m*-ethylaniline, *m*-aminoacetophenone, *m*-aminobenzoic acid benzyl ester, and even *m*-aminonitrobenzene could be coupled to the imide intermediate **2**. The benzyl ester was finally removed by hydrogenolysis, and the resulting free acid converted into the tetrabutylammonium benzoate **6**—the binding site for basic, cationic amino acid residues. The *m*-aminoacetophenone derivative **4** is a host molecule for OH- and SH-containing amino acids, *m*-aminonitrobenzene **5** should recognize electron-rich aromatic amino acids, and *m*-ethylaniline **3** serves as a reference compound with a comparable steric demand, but which is suitable only for recognition of the peptide backbone.

Compounds **3**–**6** are some of the first Kemp's triacid derivatives with a double aromatic substitution.[9] As both aromatic moieties are unsymmetrical, two conformers may exist, which should, in principle, differ in their spectroscopic properties. However, in the ^1H and the ^{13}C NMR spectra of **3**–**6**, only one set of resonance signals is seen. NOESY spectra of **4** reveal numerous contacts between the alkyl substituents of the cyclohexyl skeleton, but only one between the two aromatic planes, which should be in close contact with each other. Correspondingly, in the productive conformation only one H–H separation is calculated to be less than 4.5 Å, which is exactly the one observed in the NOESY experiment (3.2 Å, Figure 2).

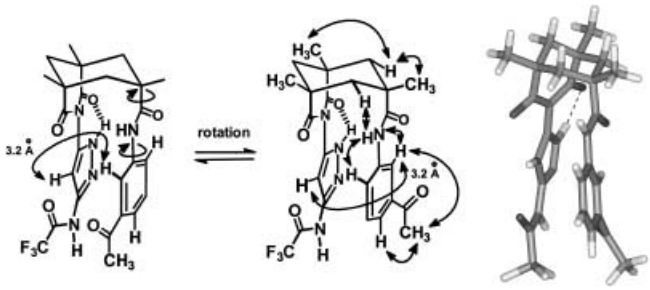


Figure 2. NOE contacts in host **4** and possible rotation around the *trans* amide bond. The productive conformation is on the right; double arrows indicate reciprocal NOE cross peaks. The two conformers shown are by far the lowest-energy structures of all the geometries found during the conformational search.[15]

A self-association experiment with **4** was conducted in chloroform in the concentration range between 10^{-2} M and 10^{-3} M. In this range, which was later used for the NMR spectroscopic titrations, no complexation-induced shifts could be observed for any of the signals arising from the host protons. Clearly the sterically demanding U-turn effectively prevents dimerization of the receptor molecules.

From the simple acylated aminopyrazoles, we knew that preformed dipeptides in an extended conformation gave the highest association constants. However, contrary to all previous experiments, our new reference compound **3** produced no complexation-induced shift after it had been added to N/C-protected divaline (Ac-Val-Val-OMe). It seems that two of the valine isopropyl groups are sterically too demanding to allow complexation with the aminopyrazole-modified Kemp's acid. This view is also supported by force-field calculations.

Pleasingly, exchange of valine for alanine restored the expected downfield shift of the signal arising from the top-face NH proton of the peptide. A Job plot[10] (Figure 3) confirmed the desired 1:1 stoichiometry, and NMR spectroscopic titrations in $CDCl_3$ furnished association constants below $100\,M^{-1}$ (Ac-Ala-Ala-OMe/Ac-Phe-Ala-OMe).[11] These numbers are relatively small, but they must be lower than those for valine-containing dipeptides, as alanine has a much higher tendency to adopt an α-helix conformation than to form a β sheet. For comparison, we titrated the two alanine-containing dipeptides mentioned above with the best β-sheet ligands developed previously by us, and obtained identical values (Table 1).

In molecular-mechanics calculations (MacroModel 7.0, OPLS-AA force-field), the backbone of a serine-containing

  1433-7851/02/4110-1752 $ 20.00+.50/0 *Angew. Chem. Int. Ed.* **2002**, *41*, No. 10

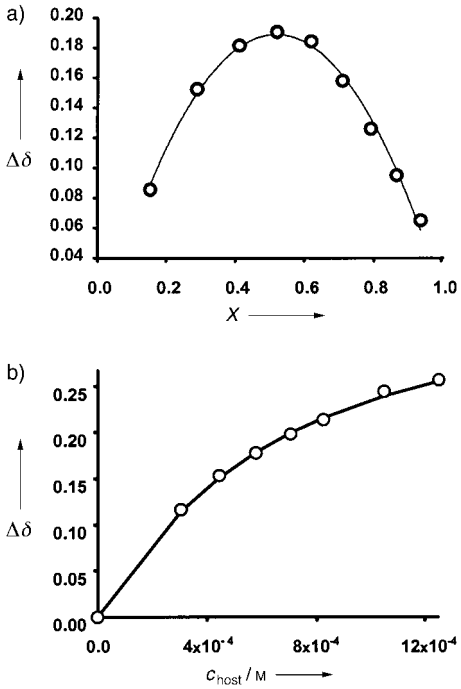


Figure 3. a) Job plot for complex formation between **6** and Ac-Orn-Ala-OMe. b) NMR spectroscopy titration curve for complex formation between **6** and Ac-Orn-Ala-OMe.

Table 1. Association constants K_a [M^{-1}] for complex formation between hosts **3**–**6** and various dipeptides.[a]

Host	Standard	Selective binding	Special cases
triflAMP[b]	Ala – Ala: 50[c]	–	Phe – Ala: 40
reference **3**	Orn – Ala: 280	Ala – Ala: 70	Phe – Ala: $\leq$ 40
XH-binder **4**	Ala – Ala: 80	Ser – Ala: 900	Ser – Val: no shifts
arene-binder **5**	–	Phe – Ala: 350	Ala – Phe: no shifts
cation-binder **6**	Propylamine: 490	Orn – Ala: 2360	–

[a] The association constants were determined by NMR spectroscopic titrations in $CDCl_3$. All dipeptides were protected as *N*-acetyl methyl esters. Stoichiometries were always 1:1, unless otherwise stated. [b] TriflAMP = 3-trifluoroacetylamino-5-methylpyrazole. [c] Errors were determined as standard deviations from the nonlinear regressions; they usually varied between 10% and 40% (exceptions: very weak binding).

dipeptide forms a complex with the trifluoroacetylated aminopyrazole by three-point binding (Figure 4). The serine hydroxy group comes close to the acetophenone C=O group, and a relatively strong, new hydrogen bond is formed.

First we tested our receptor module for serine: NMR spectroscopic titration with dialanine gave a K_a value of $80\,M^{-1}$, in good agreement with the reference compound. When Ac-Ser-Ala-OMe was bound by **4**, however, a 11-fold increase in K_a to $900\,M^{-1}$ was observed. A Job plot confirmed that a 1:1 complex was indeed formed.[12] In conclusion, the introduction of an additional hydrogen-bond acceptor in the rigid framework of the receptor module resulted in a substantial increase in free binding enthalpy for a serine-containing dipeptide, but not for dialanine.

As a representative example of electron-rich, aromatic amino acids, we tested phenylalanine in a peptidic environment. With reference compound **3** and Ac-Phe-Ala-OMe a K_a value of $\leq 40\,M^{-1}$ was measured, which again indicates

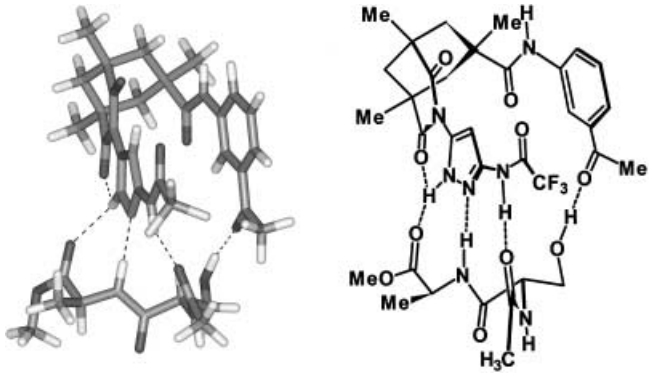


Figure 4. Result of the Monte Carlo simulation of the productive complex geometry in sequence-selective binding of serinyl – alanine-containing dipeptides by **4**.

steric repulsion of the benzyl group by the Kemp's acid derivative. When our *m*-nitroaniline-based receptor module for aromatic amino acids **5** was titrated with Ac-Phe-Ala-OMe, the nonlinear regression of binding curves for several proton signals unanimously converged at a K_a value of $350\,M^{-1}$. Again, the complex stoichiometry was calculated to be 1:1, by means of a Job plot. This time, the ninefold increase in K_a value was accompanied by marked upfield shifts for the signals arising from the aromatic protons of host and guest, which indicates π-stacking interactions. We could thus show that a nitroarene in the correct place in the rigid receptor considerably strengthened the complex of **5** with a dipeptide containing phenylalanine.

In protein recognition, basic amino acids often form ion pairs with carboxylate groups, for example, with aspartates or glutamates. These interactions are generally much stronger in apolar media than simple hydrogen bonds. To avoid the possibility that such an ion pair could override the backbone recognition by the aminopyrazole, we chose a counterion for the basic amino acid which itself forms a strong ion-pair-reinforced hydrogen bond with the alkylammonium ion. Therefore, the short, basic amino acid, ornithine, was incorporated as the acetate salt into the alanine-containing dipeptide Ac-Orn-Ala-OMe.[13] The benzoate group in **6** thus had to compete with the acetate (Ac) counterion of ornithine in the dipeptide. During the NMR spectroscopic titration, an unusually high number of proton signals shifted by $\Delta\delta_{sat}$ values of up to 1 ppm (Figure 5). The broad ammonium group signal and the signal arising from its neighboring methylene group, as well as those from both *ortho* protons next to the host benzoate anion, showed pronounced complexation-induced shifts, which indicates the formation of the new ion pair. Interesting enough, the signal for the top-face NH proton of the peptide also shifted drastically, by about 1 ppm. The nonlinear regression of various binding curves of host and guest signals gave association constants around $2400\,M^{-1}$; 1:1 stoichiometry was again proven by a Job plot (Figure 3). As a control experiment, we titrated **6** with *n*-propylammonium acetate and found a K_a value of $490\,M^{-1}$ for the formation of the salt bridge. The difference between these two numbers must come from the backbone recognition of the peptide by **6**, and provides experimental evidence that both

  1433-7851/02/4110-1753 $ 20.00+.50/0

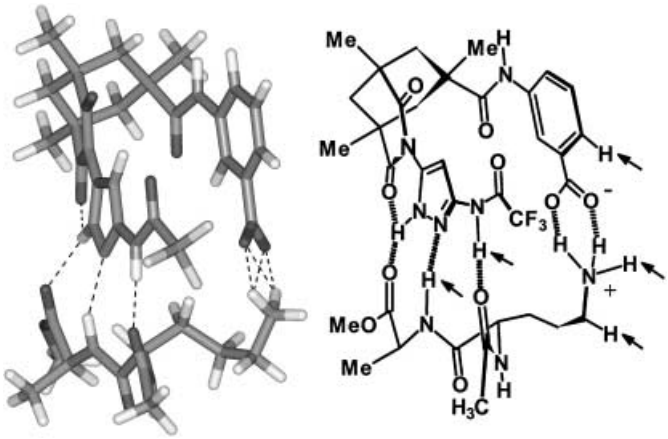


Figure 5. Left: Monte Carlo optimized structure of **6**@Ac-Orn-Ala-OMe in chloroform (1000 steps). Right: arrows indicate the atoms in host and guest for which complexation-induced shifts of proton signals are observed during the titration.

recognition processes operate simultaneously in the aggregate formation between **6** and Ac-Orn-Ala-OMe. This peptide was also titrated with reference compound **3**, which furnished a relatively low binding constant of about $280\,\text{M}^{-1}$. In this case, however, it is not clear if the aminopyrazole reacts with the backbone of the peptide or with the alkylammonium acetate ion pair. In any case, a preference of $\approx 9{:}1$ was established for the selective recognition of ornithine by the anionic receptor module **6**. Thus, introduction of a properly placed benzoate anion into the rigid receptor module led to an additional ion-pair-reinforced hydrogen bond with a representative basic amino acid.

Direct information about the peptide conformation can be drawn from Karplus analyses of the H(N),H(α-C) coupling constants, which correlate with the characteristic torsion angle θ.[14] We examined complexes of the three hosts with additional binding sites, with their best peptide-binding partners. Addition of hosts **3**–**6** to the peptide solution invariably leads to sharper resonance signals and, in most cases, increased ^3J values by comparison to the free peptides, which approach the calculated values for the pure complexes.

To conclude, we have shown that peptide receptors can be developed by rational design; these receptors are selective for some of the main classes of amino acids in a peptidic environment. We are currently working on the selective recognition of acidic and apolar amino acids in peptides. In some cases, complex formation is so sensitive towards steric factors that the same receptor module completely rejects the same dipeptide with an inverted sequence. Thus, Ac-Ala-Phe-OMe is not complexed at all. Modeling studies indicate that only one mutual orientation of the binding partners is productive, the other one seems to be sterically hindered. The sequence selectivity of the new receptor modules is currently under investigation. For a better comparison, we kept the other amino acid as an alanine in most cases. It will be interesting to find out if larger groups than the small methyl group are tolerated at all. The titration of Ac-Ser-Val-OMe with our serine binder **4** gives an initial indication of this. No complexation-induced shifts are produced, although the same dipeptide with a C-terminal alanine was bound with a high association constant of $900\,\text{M}^{-1}$.

In the future, we will combine two or more of these fragments by specific covalent linkers, to achieve sequence-selective binding of tetra- and pentapeptides in polar solution. With these receptor modules, it should be possible to dock onto specific essential-recognition motifs in peptides and proteins. This approach could ultimatively be used for protein tagging, protection, or allosteric inhibition of enzymes.

Received: December 4, 2001 [Z18326]

[1] D. H. Williams, B. Bardsley, *Angew. Chem.* **1999**, *111*, 1264–1286; *Angew. Chem. Int. Ed.* **1999**, *38*, 1172–1193.

[2] R. Haubner, D. Finsinger, H. Kessler, *Angew. Chem.* **1997**, *109*, 1440–1456; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1374–1389.

[3] a) M. Famulok, K.-S. Jong, G. Deslongchamps, J. Rebek, Jr., A*ngew. Chem.* **1991**, *103*, 880–882; *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 858–860; b) S. S. Yoon, W. C. Still, *J. Am. Chem. Soc.* **1993**, *115*, 823–824; c) C.-T. Chen, H. Wagner, W. C. Still, *Science* **1997**, *279*, 851–853; d) M. A. Hossain, H. J. Schneider, *J. Am. Chem. Soc.* **1998**, *120*, 11208–11209.

[4] Review: J. P. Schneider, J. W. Kelly, *Chem. Rev.* **1995**, *95*, 2169–2187; a) D. S. Kemp, T. P. Curran, J. G. Boyd, T. J. Allen, *J. Org. Chem.* **1991**, *56*, 6683–6697; b) M. W. Peczuh, A. D. Hamilton, *J. Am. Chem. Soc.* **1997**, *119*, 9327–9328; c) Y. Mito-Oka, S. Tsukiji, T. Hiraoka, N. Kasagi, S. Shinkai, I. Hamachi, *Tetrahedron Lett.* **2001**, *42*, 7059–7062.

[5] Overview: J. S. Nowick, *Acc. Chem. Res.* **1999**, *32*, 287–296; H. L. Schenck, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 4869–4870.

[6] a) W. C. Still, *Acc. Chem. Res.* **1996**, *29*, 155–163; b) H. Wennemers, M. Conza, M. Nold, P. Krattiger, *Chem. Eur. J.* **2001**, *7*, 3342–3347; c) T. Braxmeier, M. Demarcus, T. Fessmann, S. McAteer, J. D. Kilburn, *Chem. Eur. J.* **2001**, *7*, 1889–1898.

[7] C. Kirsten, T. Schrader, *J. Am. Chem. Soc.* **1997**, *119*, 12061–12068.

[8] J. Rebek, Jr., *Angew. Chem.* **1990**, *102*, 261–272; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 245–255.

[9] S. Shinoda, H. Tsukube, Y. Nishimura, I. Yamazaki, A. Osuka, *J. Org. Chem.* **1999**, *64*, 3757–3762.

[10] a) P. Job, *C. R. Hebd. Seances Acad. Sci.* **1925**, *180*, 928; b) M. T. Blanda, J. H. Horner, M. Newcomb, *J. Org. Chem.* **1989**, *54*, 4626.

[11] a) H. J. Schneider, R. Kramer, S. Simova, U. Schneider, *J. Am. Chem. Soc.* **1988**, *110*, 6442; b) C. S. Wilcox in *Frontiers in Supramolecular Chemistry* (Ed.: H. J. Schneider), Verlag Chemie, Weinheim, **1991**, p. 123.

[12] Unfortunately, the NMR signal for the hydroxy proton remained extremely broad throughout the entire titration, so that no direct evidence for this additional hydrogen bond could be obtained.

[13] The short side chain of ornithine can be reached more easily by the benzoate unit of the host than that of lysine (Figure 5).

[14] a) M. Delepierre, C. M. Dobson, F. M. Poulsen, *Biochemistry* **1982**, *21*, 4756; b) V. F. Bystrov, *Prog. Nucl. Magn. Reson. Spectrosc.* **1976**, *10*, 41.

[15] MacroModel 7.0, Schrödinger Inc., 2000. Force field: Amber* and OPLS-AA.

  1433-7851/02/4110-1754 $ 20.00+.50/0