Molecular Recognition

DOI: 10.1002/anie.200602412

A Self-Assembled Pyrrolic Cage Receptor Specifically Recognizes β-Glucopyranosides**

*Oscar Francesconi, Andrea Ienco, Gloriano Moneti, Cristina Nativi, and Stefano Roelens**

The molecular recognition of carbohydrates has been acknowledged as a subject of paramount importance in chemistry and biology.[1] Despite the endeavor dedicated to the research in this field in the last decade,[2] full knowledge of recognition events and control over recognition processes have yet to be achieved. By capitalizing on noncovalent interactions,[3] encouraging results have been obtained, mostly in organic solvents, with synthetic receptors that rely on hydrogen bonding; however, high selectivity remains the most ambitious goal yet to be achieved. Among the plethora of synthetic receptors reported so far,[4] cage structures endowed with various hydrogen-bonding groups have been successfully explored.[5–8] The best results have been obtained with cage receptors that employ a concerted arrangement of amidic hydrogen bonds for the multipoint binding of mono- and oligosaccharides.[5] Alternative hydrogen-bonding groups may be conveniently employed: for example, amino and hydroxy groups have been shown to be complementary hydrogen-bonding partners, both geometrically and coordinatively, thus giving rise to molecular recognition and self-assembly.[9] Likewise, pyrroles are well-established hydrogen-bonding donors, which have been largely employed for anion binding,[10] but appear to be yet unexplored for the recognition of carbohydrates. We thought that combining amino and pyrrole groups in organized architectures may result in effective receptors, provided that steric, geometric, coordinative, and functional requirements for the recognition of carbohydrates could be met through careful structural design. We describe herein the first cage receptor featuring pyrrole residues for the multipoint binding of carbohydrates, which spontaneously forms in quantitative yield by the one-pot self-assembly of five components and exhibits a specificity toward β-D-glucose and its glycosides that marks a significant step ahead in the selective recognition of monosaccharides.

When a 2:3 mixture of 1,3,5-tris(aminomethyl)-2,4,6-triethylbenzene (**1**)[4g] and pyrrole-2,5-dicarboxaldehyde (**2**)[11] in methanol was stirred overnight at room temperature, a single compound was unexpectedly obtained in quantitative yield and unambiguously identified as the hexaimine macrobicyclic cage **3** by NMR spectroscopic as well as ESI and high-resolution (HR) mass spectrometric (MS) analysis (Scheme 1).[12] Reversible imine condensation was driven toward the complete formation of a single product by precipitation, as **3** is very poorly soluble in methanol. Solubility is not, however, the only driving factor, as the corresponding reaction of the triaminoethyl homologue of **1** gave only an intractable polymeric material by precipitation. Clearly, **3** is the thermodynamically favored product that arises from condensation of five reacting molecules, which self-assemble through the concerted formation of six imine bonds.

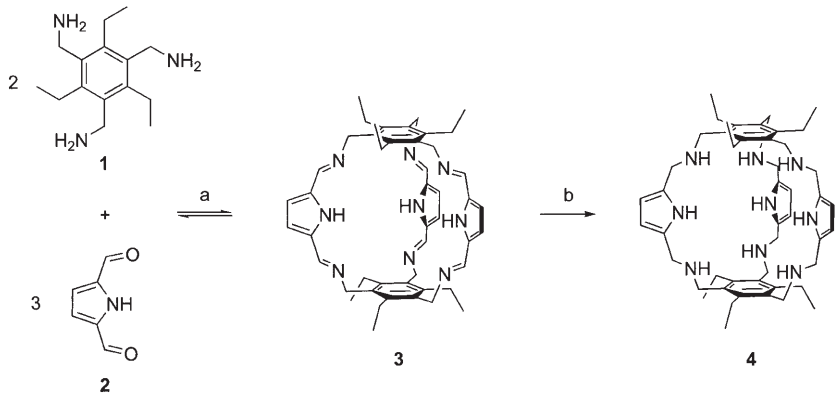
A one-pot reduction of **3** with $NaBH_4$ gave the corresponding macrobicyclic hexaamine **4** in essentially quantitative yield. In contrast to **3**, compound **4** is freely soluble in lipophilic solvents. The ESI MS and HRMS spectra of **4** confirmed the identity of the cage, whereas the ^1H and ^{13}C NMR spectra displayed signals in agreement with a highly symmetrical structure. The cause of the observed spectroscopic simplicity is evident from the single-crystal X-ray structure of **4** (see Supporting Information), which shows a nearly perfect C_{3h} symmetry of the cage,[13] apparently persistent in solution, with all the amine groups pointing inward, the ethyl groups pointing outward, and the pyrrole rings facing the cavity (Figure 1).

Although the pyrrole rings are somewhat tilted, a roughly spherical cavity is envisaged from the projection: the size of the cavity, whose diameter between the benzene rings is

[*] Dr. O. Francesconi, Prof. C. Nativi, Dr. S. Roelens
Dipartimento di Chimica Organica
CNR and Università di Firenze
Via della Lastruccia, 13
50019 Sesto Fiorentino, Firenze (Italy)
Fax: (+39) 055-457-3570
E-mail: stefano.roelens@unifi.it
Dr. A. Ienco
Istituto di Chimica dei Composti Organometallici
CNR, Via Madonna del Piano
50019 Sesto Fiorentino, Firenze (Italy)
Prof. G. Moneti
Centro Interdipartimentale di Spettrometria di Massa
Università di Firenze, V.le G. Pieraccini, 6
50139 Firenze (Italy)

[**] Financial support by Ente Cassa di Risparmio di Firenze for the acquisition of a 400-MHz NMR spectrometer is gratefully acknowledged. High-field NMR experiments were performed at the Magnetic Resonance Center, University of Florence. M. Lucci is gratefully acknowledged for his kind assistance.

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



Scheme 1. Synthesis of the macrobicyclic cage **4**. a) MeOH, 12 h, room temperature; b) $NaBH_4$, MeOH/$CHCl_3$ (3:1), 30 min, room temperature.

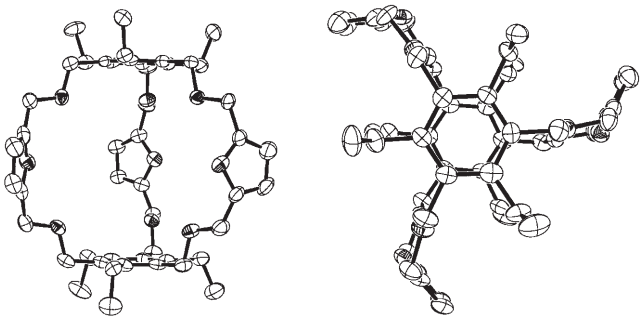


Figure 1. ORTEP projections of the X-ray crystal structure of **4**. Side view (left); top view (right). Ellipsoids are shown at the 50% probability level. Nitrogen atoms are represented as shaded ellipsoids. Solvent molecules and hydrogen atoms are omitted for clarity.

8.4 Å, and the arrangement of the amino groups appear well suited for guest binding.

Binding experiments were thus performed by ^{1}H NMR spectroscopic analysis in $CDCl_3$ with octyl-β-D-glucopyranoside (OctβGlc) as a soluble glycosidic guest. The spectra obtained by varying the **4**/OctβGlc molar ratio at a constant total concentration of reactants are shown in Figure 2.

The disappearance of the signals of **4** and the appearance of a new set of signals testified to the formation of a host–guest complex in a slow-exchange regime with the free species on the NMR timescale. Note that the single signal for the three equivalent pyrrolic NH protons is split into three nonequivalent singlets, whereas the pyrrolic CH signal splits into three strongly coupled nonequivalent signals, which show that the C_{3h} symmetry is lost upon complexation. Both slow exchange and desymmetrization, together with the marked downfield shift of the NH signals, point to the formation of a hydrogen-bonded complex with the glucoside at least partially included in the cavity. Indeed, a separate set of signals is observed for the glucose moiety as well, upfield with respect to that of the free glucoside and consistent with the shielding effect of the benzene rings; the glucose protons H3, H4, and H5 experience the largest shifts, thus suggesting inclusion from the side opposite the glycosidic chain (see Supporting Information). A 1:1 stoichiometry was inferred from complete disappearance of the free host for just over a stoichiometric reactant ratio and from the maximum intensity of the signals of the complex observed for a 1:1 molar ratio; the corresponding association constant (K_a) of $4.83(8)\times10^4\,M^{-1}$, which corresponds to an affinity of 20.7 μM of **4** for OctβGlc, was obtained with excellent agreement from four independent experiments at different reactant concentrations. Quite strikingly, the same experiment performed with the α anomer (OctαGlc) did not show any evidence of binding, thus indicating that **4** is able to bind the β anomer exclusively. Although very high β/α selectivities have been reported,[5d] to our knowledge, exclusive binding of one anomer in the recognition of glycosides is unprecedented for synthetic receptors. We believe that improved performance results from both a precise size fit of the β-glucoside and a closer complementarity of the amino/pyrrole versus the amide group in hydrogen bonding to the glucose moiety.

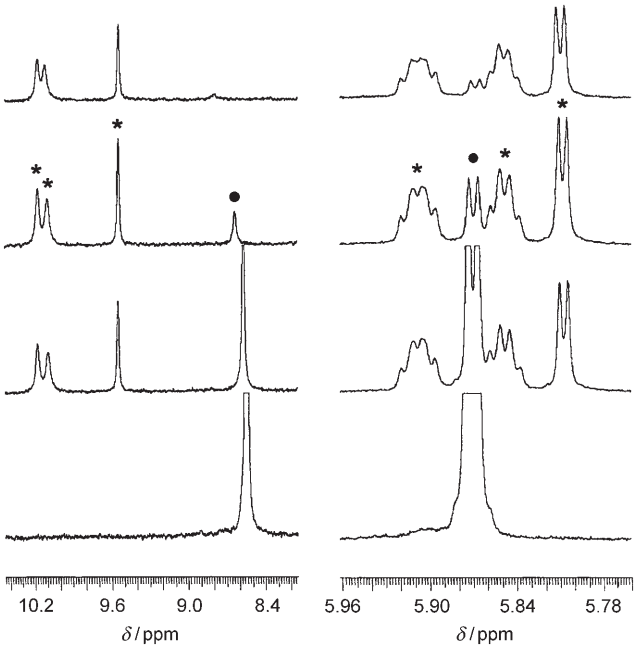
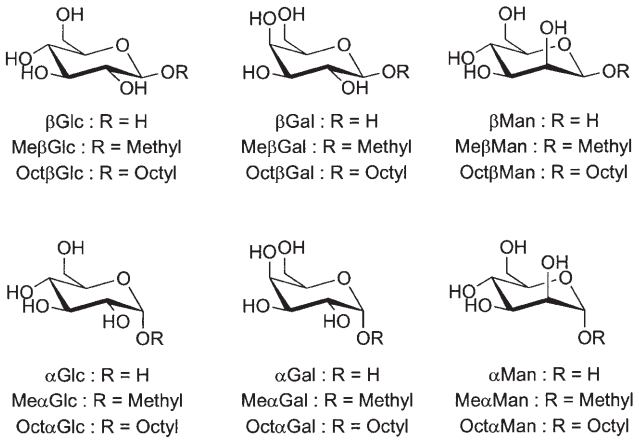


Figure 2. ^{1}H NMR spectra (400 MHz, 25 °C, $CDCl_3$) of mixtures of **4** and OctβGlc for a varying molar ratio (bottom to top: 1:0, 2:1, 1:1, 1:2) at constant total concentration of reactants. Only the pyrrole NH (left) and CH (right) signals of **4** are shown. Asterisks (*) correspond to [**4**·OctβGlc]; circles (●) indicate free **4**.

Further evidence supports the binding ability of **4**. Methyl-β-D-glucopyranoside (MeβGlc) is insoluble in $CDCl_3$. When solid MeβGlc was shaken with a millimolar solution of **4** in $CDCl_3$, the solid partially dissolved and the spectrum of the resulting solution unambiguously showed that over 40% of the cage was present in the complexed form. Bound **4** was increased to 50% in CCl_4 and to 75% when the experiment was performed in C_6D_6, thus proving that the cage receptor is capable of bringing insoluble β-glucosides (but not α-glucosides) into lipophilic solvents of low polarity. Indeed,

MeαGlc was not appreciably dissolved in any of the above solvents. Most remarkably, β-D-glucose (βGlc) itself could be dissolved in benzene, with up to nearly 20% of bound receptor observed, whereas αGlc could not. Thus, the possibility that the observed β/α selectivity may be steric in origin, caused by the bulky octyl group, can be ruled out.

The behavior of the cage receptor toward α- and β-octylglycosides of biologically relevant monosaccharides, namely, galactopyranosides (Gal) and mannopyranosides (Man; Scheme 2), was further tested in $CDCl_3$. Although



Scheme 2. Monosaccharides used in recognition experiments.

interaction between partners was observed by a shift of some signals of both the host and guest, the presence of a separate set of signals for the complex species was not detected in any case. Competitive experiments feeding **4** with equimolar mixtures of OctβGlc and each of the selected glycosides showed that for a 1:1:1 ratio of reactants, the fraction of **4** bound to OctβGlc decreased, with respect to that observed in the absence of competitors, by significantly less than 10% in the most adverse case. Experimental evidence demonstrated that none of the tested glycosides could effectively compete with OctβGlc for **4**.

Interestingly, the hexaimine cage **3** did not exhibit the same binding ability of **4**. Addition of OctβGlc to a solution of **3** in $CDCl_3$ did not show evidence of complexation, but rather induced slow re-equilibration of the cage to oligomeric products. Likewise, mixing **1** and **2** in the presence of OctβGlc as a template gave substantial amounts of oligomeric iminic products, together with lower yields of **3**. Apparently, the iminic cage does not bind to OctβGlc, and therefore using the latter as a template hampers, rather than assists, the formation of the cage. This contrasting behavior is most likely related to the geometrical restrictions imposed by the iminic double bond, which must lie coplanar to the conjugated pyrrole ring, rather than to the basicity/coordinative properties of the imine nitrogen atom, considering that the cage size is essentially identical.

Independent experimental support was desirable to validate the binding affinity results. Unfortunately, crystals suitable for X-ray analysis could not be obtained for any of the complexes of **4** with glucosides. However, ESI MS analysis provided clear-cut evidence that was in full agreement with the NMR binding studies. Two equimolar solutions of **4** with OctβGlc and OctαGlc, respectively, of the same concentration were submitted to positive-ion-mode ESI MS analysis under the same conditions. Although a peak for the $[\mathbf{4}\cdot\text{Oct}\beta\text{Glc}+\text{H}]^+$ complex was present in the spectrum of the former with comparable abundance with respect to the $[\mathbf{4}+\text{H}]^+$ peak, the peak of the complex could only be detected at the noise level in the spectrum of the latter (Figure 3). The

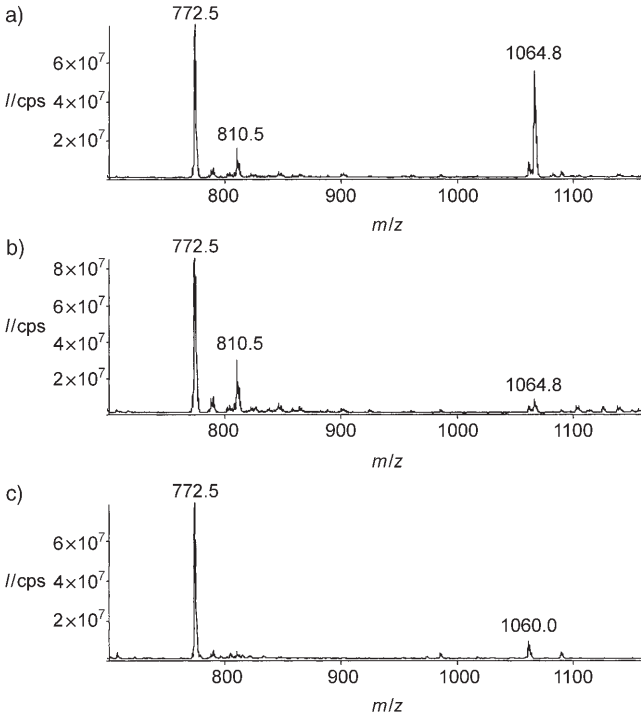


Figure 3. ESI MS spectra of a) **4** + OctβGlc (both 0.2 mM); b) **4** + OctαGlc (both 0.2 mM); and c) **4** + β-D-glucopyranoside pentaacetate (both 0.2 mM). Solvent: $CHCl_3/CH_3CN$ (1:1); ESI voltage: 6 kV; sampling cone potential: 56 V. *m/z*: 772.5 $[\mathbf{4}+\text{H}]^+$, 794.6 $[\mathbf{4}+\text{Na}]^+$, 810.5 $[\mathbf{4}+\text{K}]^+$, 1064.8 $[\mathbf{4}\cdot\text{Oct}\beta(\alpha)\text{Glc}+\text{H}]^+$.

latter spectrum appeared to be unaffected by three subsequent twofold increases of the concentration of the OctαGlc injected. In addition, the spectrum of an equimolar mixture of **4** and β-D-glucopyranoside pentaacetate, run for comparison under the same conditions, revealed a complete absence of the peak of the complex, thus showing that in the absence of free hydroxy groups, binding to the glucose moiety in the gas phase does not occur. Clearly, OctαGlc, which differs from OctβGlc only by the stereochemistry at C1, was not bound to **4** to a significantly larger extent than β-D-glucopyranoside pentaacetate.

Conclusive evidence was obtained by collision-induced dissociation (CID) experiments run on a triple quadrupole mass spectrometer. A scan of the intensity of the $[\mathbf{4}\cdot\text{Oct}\beta\text{Glc}+\text{H}]^+$ and $[\mathbf{4}+\text{H}]^+$ ions that originate from the ion of the complex selected at *m/z* 1065 with increasing potential gave the profiles shown in Figure 4b, which crossed for an energy value of 13.9 eV and corresponds to the energy

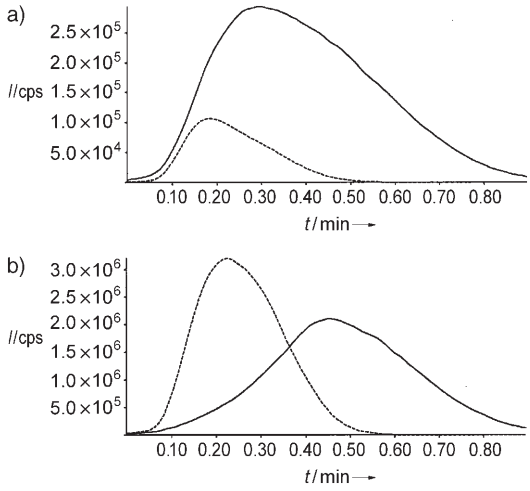


Figure 4. CID MS/MS analysis of the complex detected as the $[M+H]^+$ ion at m/z 1065. Products: 1065 ($[M+H]^+$, dotted line), 772 ($[\mathbf{4}+H]^+$, solid line). Solvent: $CHCl_3/CH_3CN$ (1:1); ESI voltage: 6 kV; sampling cone potential: 56 V; signal acquisition: 0.90 min, 89 scans over collision energies from 6 to 50 eV in 0.5 eV steps; collision-gas pressure: $P = 2.64\times10^{-5}$ torr. a) **4** + OctαGlc (0.2 mM each); b) **4** + OctβGlc (0.2 mM each).

required to dissociate 50% of the complex under the specific experimental conditions. The CID profiles that originated from the $[\mathbf{4}\cdot\text{OctαGlc}+H]^+$ ion under identical conditions (Figure 4a) did not exhibit any crossing point in the whole range of collision energies investigated, thus proving that spontaneous dissociation of the complex is prevalent for [**4**·OctαGlc] and that dissociation as a result of collisions is negligible at all concentrations, as identical results were obtained for four different concentrations of the OctαGlc injected (see Supporting Information). Exclusive recognition of the β anomer, observed in solution, was thus confirmed in the gas phase.

In summary, we have described a new synthetic cage receptor, quantitatively obtained by a one-pot reduction of a spontaneously self-assembled macrobicyclic cage, which forms by condensation of five component molecules through the thermodynamically controlled formation of six imine bonds. The receptor specifically recognizes the β anomer of D-glucose and its alkyl glucosides with complete β/α selectivity and effectively discriminates β monosaccharides of the gluco series from both the α and β anomers of the galacto and manno series. The discovery is bound to have an impact on the understanding of the structural features and the rational design of synthetic receptors for the molecular recognition of carbohydrates.[14]

Received: June 15, 2006
Revised: July 6, 2006

Keywords: cage compounds · carbohydrates · molecular recognition · self-assembly · synthetic receptors

[1] B. Ernst, G. W. Hart, P. Sinaÿ, *Carbohydrates in Chemistry and Biology*, *Part I*, *Vol. 2*, Wiley-VCH, Weinheim, **2000**; B. Ernst, G. W. Hart, P. Sinaÿ, *Carbohydrates in Chemistry and Biology*, *Part II*, *Vol. 4*, Wiley-VCH, Weinheim, **2000**.

[2] a) A. Lutzen in *Highlights in Bioorganic Chemistry* (Eds.: C. Schmuck, H. Wennemers), Wiley-VCH, Weinheim, **2004**, pp. 109–119; b) "Host–Guest Chemistry. Mimetic Approaches to Study Carbohydrate Recognition": *Topics in Current Chemistry*, *Vol. 218* (Ed.: S. Penadés), Springer, Heidelberg, **2001**; c) A. P. Davis, R. S. Wareham, *Angew. Chem.* **1999**, *111*, 3160–3179; *Angew. Chem. Int. Ed.* **1999**, *38*, 2978–2996.

[3] Covalent interactions have also been exploited in the recognition of carbohydrates; see, for example: T. D. James, K. R. A. S. Sandanayake, S. Shinkai, *Angew. Chem.* **1996**, *108*, 2038–2050; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1910–1922.

[4] For some recent examples, see: a) J. Bitta, S. Kubik, *Org. Lett.* **2001**, *3*, 2637–2640; b) S.-i. Tamaru, S. Shinkai, A. B. Khasanov, T. W. Bell, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4972–4976; c) K. Ladomenou, R. P. Bonar-Law, *Chem. Commun.* **2002**, 2108–2109; d) Y.-H. Kim, J.-I. Hong, *Angew. Chem.* **2002**, *114*, 3071–3074; *Angew. Chem. Int. Ed.* **2002**, *41*, 2947–2950; e) O. Rusin, K. Lang, V. Kral, *Chem. Eur. J.* **2002**, *8*, 655–663; f) M. Segura, B. Bricoli, A. Casnati, E. M. Munoz, F. Sansone, R. Ungaro, C. Vicent, *J. Org. Chem.* **2003**, *68*, 6296–6303; g) A. Vacca, C. Nativi, M. Cacciarini, R. Pergoli, S. Roelens, *J. Am. Chem. Soc.* **2004**, *126*, 16456–16465; h) M. Mazik, H. Cavga, P. G. Jones, *J. Am. Chem. Soc.* **2005**, *127*, 9045–9052; i) C. Schmuck, M. Schwegmann, *Org. Lett.* **2005**, *7*, 3517–3520; j) H. Abe, Y. Aoyagi, M. Inouye, *Org. Lett.* **2005**, *7*, 59–61.

[5] a) T. Velasco, G. Lecollinet, T. J. Ryan, A. P. Davis, *Org. Biomol. Chem.* **2004**, *2*, 645–647; b) T. J. Ryan, G. Lecollinet, T. Velasco, A. P. Davis, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4863–4866; c) G. Lecollinet, A. P. Dominey, T. Velasco, A. P. Davis, *Angew. Chem.* **2002**, *114*, 4267–4270; *Angew. Chem. Int. Ed.* **2002**, *41*, 4093–4096; d) A. P. Davis, R. S. Wareham, *Angew. Chem.* **1998**, *110*, 2397–2401; *Angew. Chem. Int. Ed.* **1998**, *37*, 2270–2273.

[6] a) R. Welti, Y. Abel, V. Gramlich, F. Diederich, *Helv. Chim. Acta* **2003**, *86*, 548–562; b) R. Welti, F. Diederich, *Helv. Chim. Acta* **2003**, *86*, 494–503.

[7] T. Ishi-i, M. A. Mateos-Timoneda, P. Timmerman, M. Crego-Calama, D. N. Reinhoudt, S. Shinkai, *Angew. Chem.* **2003**, *115*, 2402–2407; *Angew. Chem. Int. Ed.* **2003**, *42*, 2300–2305.

[8] R. Liu, W. Clark Still, *Tetrahedron Lett.* **1993**, *34*, 2573–2576.

[9] a) S. Hanessian, M. Simard, S. Roelens, *J. Am. Chem. Soc.* **1995**, *117*, 7630–7645; b) O. Ermer, A. Eling, *J. Chem. Soc. Perkin Trans. 2* **1994**, 925–944.

[10] J. L. Sessler, S. Camiolo, P. A. Gale, *Coord. Chem. Rev.* **2003**, *240*, 17–55.

[11] C. E. Loader, H. J. Anderson, *Synthesis* **1978**, 295–297.

[12] For some recent examples of the efficient assembly of macrocyclic and supramolecular hosts of complex structure by imine formation, see: a) X. Liu, Y. Liu, G. Li, R. Warmuth, *Angew. Chem.* **2006**, *118*, 915–918; *Angew. Chem. Int. Ed.* **2006**, *45*, 901–904; b) M. J. MacLachlan, *Pure Appl. Chem.* **2006**, *78*, 873–888; c) J. Gawronski, K. Gawronska, J. Grajewski, M. Kwit, A. Plutecka, U. Rychlewska, *Chem. Eur. J.* **2006**, *12*, 1807–1817; d) N. Kuhnert, A. Lopez-Periago, G. M. Rossignolo, *Org. Biomol. Chem.* **2005**, *3*, 524–537; e) E. A. Katayev, G. Dan Pantos, M. D. Reshetova, V. N. Khrustalev, V. M. Lynch, Y. A. Ustynyuk, J. L. Sessler, *Angew. Chem.* **2005**, *117*, 7552–7556; *Angew. Chem. Int. Ed.* **2005**, *44*, 7386–7390.

[13] For recent highlight articles on C_3-symmetric receptors and ligands, see: a) C. Moberg, *Angew. Chem.* **2006**, *118*, 4838–4840; *Angew. Chem. Int. Ed.* **2006**, *45*, 4721–4723; b) M. Albrecht, *Angew. Chem.* **1999**, *111*, 3671–3674; *Angew. Chem. Int. Ed.* **1999**, *38*, 3463–3465.

[14] See Supporting Information for general methods and materials, synthetic experimental details, NMR spectra, and CID MS spectra.

