



The self-assembly of calix[4]arene derivatives based on an A-T base pairing

Cheng-Chu Zeng, Ya-Ling Tang, Qi-Yu Zheng, Li-Jun Huang, Bin Xin and Zhi-Tang Huang*

Center for Molecular Science, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100080, China

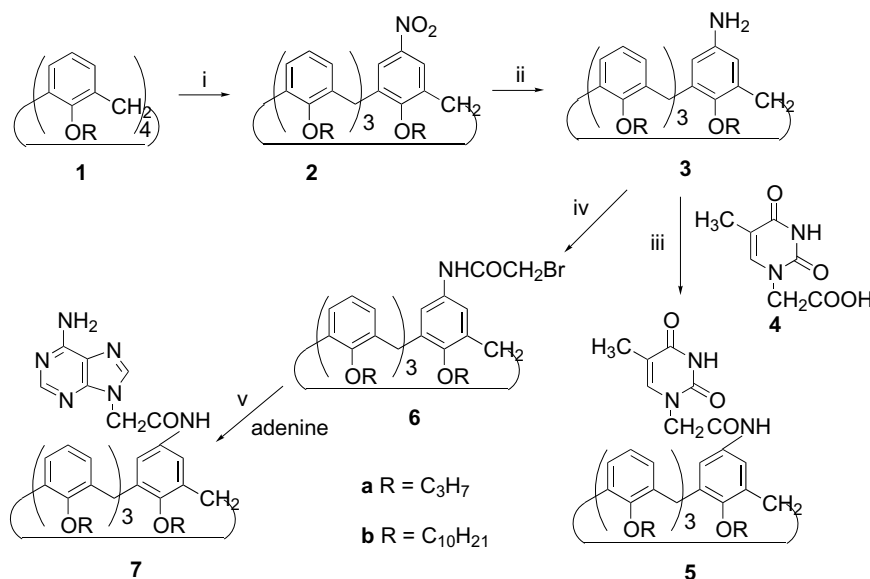
Received 16 May 2001; accepted 5 July 2001

Abstract—Calix[4]arene derivatives with one adenine or thymine sidearm have been synthesized and their self-association characteristics which depend on solvent, temperature and concentrations are described. Mainly Watson–Crick A-T base pairing in these systems was detected using ^1H NMR and ESI-MS. © 2001 Elsevier Science Ltd. All rights reserved.

The search for effective scaffolds remains one of the main challenges in supramolecular chemistry.¹ Based on the self-complementary principle, a large number of structures able to undergo assembly have been designed and synthesized in the past few decades.² Surprisingly, however, the use of the DNA base pairing interaction, the best known complementarity in nature, has remained largely unexplored and no significant progress has been made using it in the study of molecular

assembly.³ This is partly due to the weak interactions between the bases.⁴

Calix[*n*]arenes, cyclic oligomers derived from the condensation of phenol and formaldehyde, are unique and versatile building blocks in supramolecular chemistry because they possess different cavities and conformations fine-tunable by chemical modification both at the upper and lower rims.⁵ Progress has been made towards



Scheme 1. Reagents and conditions: (i) HNO_3 , CH_3COOH , CH_2Cl_2 , rt, 30 min; (ii) $\text{SnCl}_2 \cdot 2\text{H}_2\text{O}$, $\text{C}_2\text{H}_5\text{OH}$, reflux; (iii) DCC, HOBT, DMF, rt; (iv) BrCH_2COBr , CH_2Cl_2 , rt; (v) NaH, DMF, rt.

* Corresponding author. Tel.: +86-10-62544082; fax: +86-10-62559373; e-mail: huangzt@public.bta.net.cn

the molecular recognition of calix[*n*]arene derivatives with organic molecules and metal ions.⁶ Self-assembly of calix[*n*]arene compounds has been achieved by non-covalent intermolecular recognition or complexation with organic molecules and metal cations.^{2,7} For example, a molecular capsule has been obtained while a top-rim derivatized calix[4]arene dimerized by hydrogen-bonding of ureas.⁸ Our interest in the construction of multi-cavity receptors and their inclusion properties has led us to investigate the synthesis of adenine- and thymine-bearing calix[4]arene derivatives. We envisaged that the DNA-like adenine–thymine interaction would give rise to a calix[4]arene assembly. The study may also allow us to shed light on recognition processes in chemical and biological systems and to develop novel biological targeting agents.

The synthesis of base-containing calix[4]arenes is outlined in Scheme 1. The monoamino calix[4]arenes **3** were obtained from mononitration of calix[4]arenes **1**, followed by reduction using tin dichloride in boiling ethanol. The coupling reactions between **3** and thymine-1-ylacetic acid **4** in the presence of DCC afforded the monothymine-substituted calix[4]arenes **5**. The same synthetic strategy did not apply to the adenine analogs **7** because adenine-9-ylacetic acid is sparingly soluble in DMF and DMSO. Instead, compounds **7** were prepared from the substitution reactions of bromoacylated monoamino calix[4]arenes **6** with adenine with the aid of sodium hydride. The introduction of four alkyl groups at the lower rim improved the solubility of the calix[4]arene derivatives⁵ and thus facilitated the study.⁹

The spectroscopic data and elemental analysis were in agreement with the structure of the products.¹⁰ It is worth noting that both the thymine- and adenine-containing calix[4]arene products **5** and **7** tended to undergo self-association through intermolecular hydrogen bonding in solution, which was evidenced by a ¹H NMR spectroscopic study. For example, a significant downfield shift ($\Delta\delta = 1.56$ – 1.62 ppm) of the signal of the thymine imido-H of compound **5** was observed when the concentration increased from 5 to 22 mM in CDCl₃. The signals of the amino protons of the adenine of **7** were downfield shifted ($\Delta\delta = 0.26$ – 0.40 ppm) similarly upon increase of the concentration. Such a distinct concentration dependence of the NMR signals of the thymine and adenine base residues reflected, respectively, strong T–T and A–A self-association through intermolecular hydrogen bonding.¹¹ Intermolecular hydrogen bond formation was further confirmed by variable-temperature ¹H NMR spectra. A downfield shift of 0.75–0.60 ppm and of 0.16 ppm was observed, respectively, corresponding to the thymine imido-H of **5** and adenine amino-H of **7**, when the temperature decreased from 326 to 300 K. The change of solvent from CDCl₃ to CD₃SOCD₃, a polar solvent able to minimize the intermolecular hydrogen bonding, resulted in a shift of the imido-H peak of **5a** from 8.15 to 11.26 ppm and of the amino-H peak of **7a** from 5.92 to 7.17 ppm. It should be noted that, except for the observation of the peak shifts of the thymine imido-H and of the adenine amino-H mentioned above, the

proton signals of the rest of the molecule remained almost constant under the conditions employed. This implies the formation of discrete self-assembly of **5** or **7** rather than a polymeric one in CDCl₃.¹²

Molecular assembly between compounds **5** and **7** was studied by monitoring by ¹H NMR at 300 K when a thymine-bearing calix[4]arene **5** solution in CDCl₃ (10 mM) was treated with an equal amount of the adenine-containing calix[4]arene **7**. From the resulting spectra (Fig. 1), we found that both the signals of the imido proton of the thymine residue of **5a** and the amino protons of the adenine moiety of **7a** moved downfield, with $\Delta\delta$ being 1.32 and 0.33 ppm, respectively. Importantly, the peak corresponding to the proton at C-2 of the adenine of **7a** also shifted downfield ($\Delta\delta$ 0.24 ppm), while the signal due to the proton at C-8 of the adenine of **7a** remained almost unchanged ($\Delta\delta$ 0.04 ppm). This clearly indicated that the calix[4]arene derivatives **5a** and **7a** undergo assembly mainly through Watson–Crick A–T base pairing recognition (Scheme 2).^{3,13} The situation of assembly between **5b** and **7b** was similar.

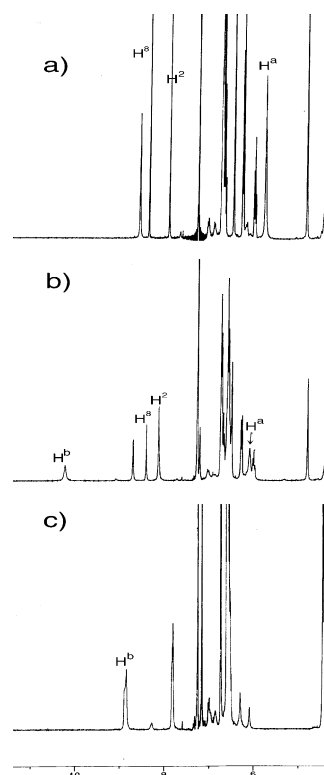
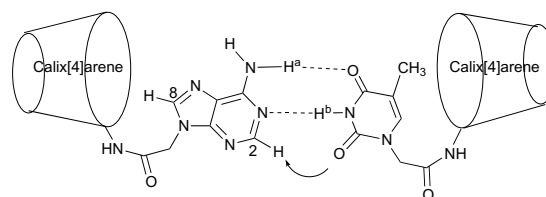


Figure 1. Partial ¹H NMR spectra of (a) **7a**; (b) the 1:1 mixture of **5a** and **7a** and (c) **5a**. Spectra were recorded in CDCl₃ on a 300 MHz spectrometer. [**5a**] = [**7a**] = [**5a**+**7a**] = 10 mM.



Scheme 2. Watson–Crick pairing.

The interaction of A and T has been reported to proceed through both Watson–Crick and Hoogsteen type recognition,¹⁴ with the former being the sole recognition mechanism in nucleic acids. The precise reason for the formation of mainly Watson–Crick A–T base pairing between **5** and **7** is unclear, although it may be due to less steric interaction between the two calix[4]arene moieties in such a complex.

Encouraged by the results from the ¹H NMR study, we investigated the molecular assembly of **5** and **7** further using mass spectroscopy. The MALDI-TOF mass spectrum of compounds **5** and **7**, and of the mixture of **5** and **7** in CHCl₃ did not give any peaks other than individual molecular peaks. When electron spray ionization was applied, however, peaks of complex species were detected. For instance, a mixture of equimolar **5a** and **7a** in CH₂Cl₂/CH₃CN/AcOH gave an intense **5a**·**7a**·H⁺ complex peak at 1556.8, while a mixture of **5b** and **7b** gave a peak at 2341.7 corresponding to **5b**·**7b**·H⁺.

In summary, we have synthesized calix[4]arene derivatives bearing one thymine or one adenine moiety, and they have been shown to undergo self-association in solution. Effective Watson–Crick A–T recognition leads to molecular assembly between the adenine- and thymine-containing calix[4]arenes. We have also demonstrated that electron spray ionization mass spectrometry is a powerful means to study molecular assembly based on base pairing interactions.

Acknowledgements

We thank the National Natural Science Foundation of China, the Major State Basic Research Development Program of China (Grant No. G2000078100) and the Chinese Academy of Sciences for the financial support.

References

- (a) Lehn, J.-M. *Supramolecular Chemistry*; VCH: Weinheim, 1995; (b) *Comprehensive Supramolecular Chemistry*; Lehn, J.-M.; Atwood, J. L.; Davies, J. E. D.; MacNicol, D. D.; Vögtle, F., Eds.; Pergamon: New York, 1996.
- (a) Conn, M. M.; Rebek, Jr., J. *Chem. Rev.* **1997**, *97*, 1647; (b) Linton, B.; Hamilton, A. D. *Chem. Rev.* **1997**, *97*, 1669; (c) Higler, I.; Timmerman, P.; Verboom, W.; Reinhoudt, D. N. *Eur. J. Org. Chem.* **1998**, 2689.
- (a) Schall, O. F.; Gokel, G. W. *J. Am. Chem. Soc.* **1994**, *116*, 6089; (b) Sessler, J. L.; Wang, R. *J. Am. Chem. Soc.* **1996**, *118*, 9809.
- (a) Sessler, J. L.; Magda, D.; Furuta, H. *J. Org. Chem.* **1992**, *57*, 818; (b) Kyogoku, Y.; Lord, R. C.; Rich, A. *Natl. Acad. Sci. USA* **1967**, *57*, 250.
- (a) Gutsche, C. D. *Calixarene Revisited*; Royal Society of Chemistry: London, 1998; (b) Gutsche, C. D. *Calixarene*; Royal Society of Chemistry: London, 1989; (c) Ikeda, A.; Shinkai, S. *Chem. Rev.* **1997**, *97*, 1713; (d) Böhmer, V. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 715.
- Ludwig, R. *Fresenius J. Anal. Chem.* **2000**, 367, 103.
- (a) Vreekamp, R. H.; van Duynhoven, J. P. M.; Hubert, M.; Verboom, W.; Reinhoudt, D. N. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1215; (b) Hajek, F.; Graf, E.; Hosseini, M. W.; Delaigue, X.; Cian, A. D.; Fisher, J. *Tetrahedron Lett.* **1996**, *37*, 1401.
- Hamann, B. C.; Shimizu, K. D.; Rebek, Jr., J. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1326.
- Sessler, J. L.; Wang, R. *J. Org. Chem.* **1998**, *63*, 4079.
- 5a**: Mp 222–224°C; ¹H NMR (CDCl₃, 300 MHz): 0.85–1.01 (m, 12H, CH₂CH₃), 1.83–1.94 (m, 11H, CH₂CH₃+T-CH₃), 3.12 and 4.43 (ABq, 4H, *J*=13.2 Hz, ArCH₂Ar), 3.16 and 4.47 (ABq, 4H, *J*=13.2 Hz, ArCH₂Ar), 3.76–3.88 (m, 8H, OCH₂), 4.45 (s, 2H, COCH₂), 6.50 (s, 6H, Ar-H), 6.58 (t, 1H, *J*=7.32 Hz, ArH), 6.67 (d, 2H, *J*=7.32 Hz, ArH), 6.87 (s, 2H, ArH), 7.23 (s, 1H, T-H), 8.34 (s, br, 1H, CONH), 9.77 (s, br, 1H, T-NH); ¹³C NMR (CDCl₃): 10.2, 10.3, 12.3, 23.1, 23.1, 23.2, 29.7, 30.9, 51.5, 76.6, 76.7, 111.1, 120.2, 121.5, 122.0, 128.0, 128.2, 128.2, 131.0, 134.4, 134.9, 135.4, 135.9, 141.0, 151.3, 154.0, 156.3, 156.8, 164.0, 164.3; MS (MALDI-TOF): *m/z* 773.39 (M⁺). IR (KBr): 1680 cm⁻¹, 3302 cm⁻¹; anal. calcd for C₄₇H₅₅N₃O₇: C, 72.92; H, 7.17; N, 5.46. Found: C, 72.99; H, 7.18; N, 5.19. **7a**: Mp 276–278°C; ¹H NMR (CDCl₃, 300 MHz): 0.91–1.02 (m, 12H, CH₂CH₃), 1.82–1.95 (m, 8H, CH₂CH₃), 3.09 and 4.39 (ABq, 4H, *J*=13.2 Hz, ArCH₂Ar), 3.11 and 4.41 (ABq, 4H, *J*=13.2 Hz, ArCH₂Ar), 3.72–3.86 (m, 8H, OCH₂), 4.84 (s, 2H, COCH₂), 6.09 (t, 1H, *J*=7.47 Hz, ArH), 6.18 (s, 2H, A-NH₂), 6.33 (d, 2H, *J*=7.47 Hz, Ar-H), 6.54 (s, 2H, Ar-H), 6.61–7.02 (m, 6H, Ar-H), 7.91 (s, 1H, A-C₂-H), 8.33 (s, 1H, A-C₈-H), 8.79 (s, 1H, CONH); ¹³C NMR (CDCl₃): 10.2, 10.4, 23.1, 23.2, 23.3, 30.9, 31.0, 48.6, 76.6, 77.0, 119.3, 120.0, 121.2, 121.9, 127.8, 128.3, 128.4, 131.0, 134.7, 135.1, 135.5, 135.6, 141.0, 149.8, 153.0, 153.6, 155.8, 156.3, 156.9, 163.3; MS (MALDI-TOF): *m/z*: 783.66 (M⁺). IR (KBr): 1655 cm⁻¹, 3390 cm⁻¹; anal. calcd for C₄₇H₅₄N₆O₅: C, 72.10; H, 6.95; N, 10.73. Found: C, 71.85; H, 7.35; N, 10.35.
- Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984.
- Brody, M. S.; Schalley, C. A.; Rudkevich, D. M.; Rebek, Jr., J. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1640.
- (a) Newmark, R. A.; Cantor, C. R. *J. Am. Chem. Soc.* **1968**, *90*, 5010; (b) Engel, J. D.; Von Hippel, P. H. *Biochemistry* **1974**, *13*, 4143.
- (a) Aoyama, Y.; Onishi, H.; Tanaka, Y. *Tetrahedron Lett.* **1990**, *31*, 1177; (b) Ling, R.; Yoshida, M.; Mariano, P. S. *J. Org. Chem.* **1996**, *61*, 4439–4449.