Supporting Information

Synthesis of a Large-Cavity Carbazole Macrocycle for Size-Dependent Recognition

Ming Dong, a Xiu Liu, a Zhi-Yuan Zhang, a Chengmao Yu, ab Bochao Huo, b and Chunju Li* ab

a. Tianjin Key Laboratory of Structure and Performance for Functional Molecules, College of Chemistry, Tianjin Normal University, Tianjin 300387, P. R. China.

E-mail: cjli@shu.edu.cn

b. Center for Supramolecular Chemistry and Catalysis and Department of Chemistry, Shanghai University, Shanghai 200444, P. R. China.
Table of Contents

1. General Method.................................................................S3
2. Synthesis............................................................................S3-S10
3. Mass spectrum and 2D ROESY spectrum of the host-guest mixture.........S11
4. Job’s plots........................................................................S12
5. $^1$H NMR spectra of traditional macrocycles with 3·Br.....................S13-S15
6. Additional $^1$H NMR spectra of host-guest mixtures.........................S16-S17
7. Determination of the association constants.....................................S18-S26
8. Energy-minimized structure of 3·Cl⊂1......................................S27-S28
9. Crystal data of 1....................................................................S29
10. Reference............................................................................S30
1. General Method

All reagents and solvents were commercially available and used without further purification, unless otherwise noted. $^1$H, $^{13}$C and 2D NMR spectra were recorded on Bruker Avance III 400 MHz. Mass spectra were determined on a Bruker Daltonics AutoflexIII LRF200-CID and Agilent 6520 q-TOF LC-MS. The single crystal X-ray data were measured by direct methods using Bruker SMART APEX II. Association constants ($K_a$) were examined by $^1$H NMR titration methods, where nonlinear curve-fitting method was used to obtain the $K_a$ values through the following equation:

$$\delta = \delta_0 + \Delta \delta (0.5/[H]_0) \left(\frac{[G]}{[H]_0} + 1/K - \left(\frac{[G]}{[H]_0} + 1/K\right)^2 - 4[H]_0[G]\right)^{0.5}$$

2. Synthesis

2.1 2,7-bis(2,4-dimethoxy-phenyl)carbazole (M).

A mixture of 2,7-dibromocarbazole (0.91 g, 2.8 mmol), 2,4-dimethoxybenzene-boronic acid (1.3 g, 7.0 mmol), [1,1'-Bis(diphenylphosphino) ferrocene] dichloro-palladium (II) (0.20 g, 0.27 mmol) and $K_2CO_3$ (0.77 g, 5.6 mmol) in 100 mL dioxane and water (v/v = 5:1) in a flask was stirred at 90 °C for 12 hours. After evaporating the solvents, resulting mixture was extracted with dichloromethane (3 × 50 mL) and then washed with water and brine successively. The organic layer was dried over anhydrous Na$_2$SO$_4$ and evaporated. The residue was purified by column chromatography with pure dichloromethane as eluent to afford 0.80 g 2,7-bis(2,4-dimethoxy-phenyl)carbazole (M) in 65% yield. $^1$H NMR (400 MHz, CDCl$_3$) δ (ppm): 8.06 (d, $J = 8.4$ Hz, 2H), 8.03 (s, 1H), 7.56 (s, 2H), 7.36 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H), 6.61 (m, 4H); $^{13}$C NMR (100 MHz, CDCl$_3$) δ (ppm): 160.1, 157.5, 139.9, 135.8, 131.7, 124.3, 121.9, 121.3, 119.6, 111.4, 104.6, 99.1, 55.6, 55.4.
Figure S1. $^1$H NMR spectrum (400 MHz, CDCl$_3$) of M.

Figure S2. $^{13}$C NMR spectrum (100 MHz, CDCl$_3$) of M.
2.2 Boc-1.

Boc₂O (0.22 g, 1.5 mmol), DMAP (0.12 g, 1.0 mmol) and 2,7-bis(2,4-dimethoxy-phenyl)carbazole (0.44 g, 1.0 mmol) were added into a dry round-bottomed flask and dissolved in dry acetone (10 mL). After stirring for 24 hours, H₂O (5.0 mL) was added. The resulting mixture was extracted with dichloromethane (3 × 50 mL). The combined organic layers were washed with water and brine successively. The combined organic layers were dried over Na₂SO₄ and evaporated under reduced pressure to afford Boc-2,7-bis(2,4-dimethoxy-phenyl)carbazole (Boc-M) which was directly used for the next step without purification. ¹H NMR (400 MHz, CDCl₃) δ (ppm): 8.47 (d, J = 1.2 Hz, 2H), 7.96 (d, J = 8.0 Hz, 2H), 7.51 (dd, J = 8.0 and 1.6 Hz, 2H), 7.37 (d, J = 8.8 Hz, 2H), 6.61 (m, 4H), 3.88 (s, 6H), 3.84 (s, 6H), 1.73 (s, 9H); ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 160.2, 157.5, 151.1, 138.9, 137.2, 131.7, 124.8, 124.3, 124.2, 118.7, 117.2, 104.7, 99.1, 83.5, 55.6, 55.4, 28.4(3C).

The prepared Boc-M was dissolved in 200 mL 1,2-dichloroethane, and paraformaldehyde (75 mg, 2.5 mmol) and boron trifluoride diethyl etherate (10 μL, 80 μmol) were added. The reaction mixture was stirred at 25 °C for 30 minutes. After quenching by 100 mL water, the organic phase was separated and washed with saturated NaHCO₃ solution and brine. The organic layer was dried over anhydrous Na₂SO₄ and evaporated. The residue was purified by column chromatography on silica gel (eluent: 5/1, v/v, petroleum ether/dichloromethane gradually changing to 2/5) to afford Boc-I (0.24 g, 43%). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 8.40 (s, 6H), 7.87 (d, J = 8.0 Hz, 6H), 7.41 (dd, J = 8.0 and 1.6 Hz, 6H), 7.07 (s, 6H), 7.59 (s, 6H), 3.96 (s, 6H), 3.90 (s, 18H), 3.82 (s, 18H), 1.49 (s, 27H); ¹³C NMR (100 MHz, CDCl₃) δ
(ppm): 157.6, 155.7, 150.9, 138.7, 137.2, 124.7, 124.0, 123.2, 121.3, 118.6, 117.3, 96.0, 83.2, 55.9, 55.8, 28.1, 26.8.

Figure S3. $^1$H NMR spectrum (400 MHz, CDCl$_3$) of Boc-M.

Figure S4. $^{13}$C NMR spectrum (100 MHz, CDCl$_3$) of Boc-M.
Figure S5. $^{1}$H NMR spectrum (400 MHz, CDCl$_3$) of Boc-1. Peaks labeled with ↓ are related to the solvent residue of aliphatic hydrocarbons.

Figure S6. $^{13}$C NMR spectrum (100 MHz, CDCl$_3$) of Boc-1.
2.3 Synthesis of host 1.

To the solution of Boc-1 (0.50 g, 0.30 mmol) in toluene (20 mL) was added column chromatography silica gel (2.0 g) and refluxed for 12 hours. After filtration, the filtrate was evaporated. The residue was purified by column chromatography on silica gel (eluent: 1/3, v/v, petroleum ether/dichloromethane) to afford host 1 (0.38 g, 93%). \(^1\)H NMR (400 MHz, CDCl\(_3\)) \(\delta\) (ppm): 7.83 (d, \(J = 8.0\) Hz, 6H), 7.80 (s, 3H), 7.34 (d, \(J = 1.2\) Hz, 6H), 7.23 (dd, \(J = 8.0\) and 1.2 Hz, 6H), 7.06 (s, 6H), 6.55 (s, 6H), 4.00 (s, 6H), 3.82 (s, 18H), 3.75 (s, 18H); \(^{13}\)C NMR (100 MHz, CDCl\(_3\)) \(\delta\) (ppm): 157.5, 155.8, 140.0, 135.8, 132.3, 123.5, 121.5, 121.2, 121.1, 119.2, 111.5, 96.0, 55.9, 55.7, 28.2; MS (m/z): [M+H]\(^+\) calcd. for C\(_{87}H_{76}N_3O_{12}\)\(^+\), 1354.54; found, 1354.65.
Figure S7. $^1$H NMR spectrum (400 MHz, CDCl$_3$) of 1. Peaks labeled with ↓ are related to the solvent residue of aliphatic hydrocarbons.

Figure S8. $^{13}$C NMR spectrum (100 MHz, CDCl$_3$) of 1.
Figure S9. NOESY NMR spectrum (400 MHz, CDCl₃) of 1.

Figure S10. Experimental (black) and calculated (red) mass spectrum of 1.
3. Mass spectrum and 2D ROESY spectrum of the host-guest mixture.

Figure S11. ROESY NMR spectrum (400 MHz, CDCl$_3$, mixing time 160 ms) of 1 + 3·Br. [1] = 2.0 mM, [3·Br] = 40 mM.

Figure S12. Experimental (black) and calculated (red) mass spectrum of 1 + 3·Cl.
4. Job’s plots

**Figure S13.** Job’s plot obtained by plotting the chemical shift change (Δδ) of the N–H proton of host 1 in ¹H NMR spectra by varying the ratio of the host and quaternary ammonium guests (2⁺–6⁺) against the mole fraction of host 1. The total concentration of the host and the guest is fixed: [Host] + [Guest] = 1.0 mM.

**Figure S14.** Job’s plot obtained by plotting the chemical shift change (Δδ) of the N–CH₂ protons of the guests in ¹H NMR spectra by varying the ratio of the host and quaternary ammonium guests (2⁺–6⁺) against the mole fraction of guest. The total concentration of the host and the guest is fixed: [Host] + [Guest] = 1.0 mM.
5. $^1$H NMR spectra of traditional macrocycles with 3·Br.

**Figure S15.** $^1$H NMR spectra (400 MHz, 1.0 mM, CDCl$_3$) of dibenzo-24-crown-8, 3·Br and their 1:1 mixture.

**Figure S16.** $^1$H NMR spectra (400 MHz, 1.0 mM, CDCl$_3$) of calix[6]arene, 3·Br and their 1:1 mixture.
**Figure S17.** $^1$H NMR spectra (400 MHz, 1.0 mM, CDCl$_3$) of per-ethylated pillar[6]arene, 3·Br and their 1:1 mixture.

**Figure S18.** $^1$H NMR spectra (400 MHz, 1.0 mM, D$_2$O) of β-cyclodextrin, 3·Br and their 1:1 mixture.
Figure S19. $^1$H NMR spectra (400 MHz, 1.0 mM, D$_2$O) of cucurbit[7]uril, 3·Br and their 1:1 mixture.
6. Additional $^1$H NMR spectra of host-guest mixtures

**Figure S20.** Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 1.0 mM) of (a) 2$^+$, (b) 1 + 2$^+$, (c) 3$^+$, (d) 1 + 3$^+$, (e) 4$^+$, (f) 1 + 4$^+$, (g) 5$^+$, (h) 1 + 5$^+$, (i) 6$^+$, (j) 1 + 6$^+$, (k) 7$^+$, (l) 1 + 7$^+$, (m) 8$^+$, and (n) 1 + 8$^+$. All of the counter anions are Br$^-$ ions.
Figure S21. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K, 1.0 mM) of host 1 in the presence of quaternary ammonium guests. (a) 1, (b) 1 + 2$^+$, (c) 1 + 3$^+$, (d) 1 + 4$^+$, (e) 1 + 5$^+$, (f) 1 + 6$^+$, (g) 1 + 7$^+$ and (h) 1 + 8$^+$. All of the counter anions are Br$^-$ ions.
7. Determination of the association constants

Figure S22. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 2·Br. From bottom to top, the concentration of 2·Br was 0–60 mM.

Figure S23. Non-linear curve-fitting for the complexation between 1 and 2·Br in CDCl$_3$. $R^2 = 0.9935$
**Figure S24.** Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 3·Cl. From bottom to top, the concentration of 3·Cl was 0–60 mM.

**Figure S25.** Non-linear curve-fitting for the complexation between 1 and 3·Cl in CDCl$_3$. R$^2$= 0.9933
Figure S26. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 3·Br. From bottom to top, the concentration of 3·Br was 0–60 mM.

Figure S27. Non-linear curve-fitting for the complexation between 1 and 3·Br in CDCl$_3$. $R^2 = 0.9983$
Figure S28. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 3·I. From bottom to top, the concentration of 3·I was 0–60 mM.

Figure S29. Non-linear curve-fitting for the complexation between 1 and 3·I in CDCl$_3$. R$^2$ = 0.9988
Figure S30. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 4·Br. From bottom to top, the concentration of 4·Br was 0–60 mM.

Figure S31. Non-linear curve-fitting for the complexation between 1 and 4·Br in CDCl$_3$. $R^2 = 0.9981$
**Figure S32.** Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 5·Br. From bottom to top, the concentration of 5·Br was 0–60 mM.

**Figure S33.** Non-linear curve-fitting for the complexation between 1 and 5·Br in CDCl$_3$. $R^2$= 0.9993
Figure S34. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 6·Br. From bottom to top, the concentration of 6·Br was 0~60 mM.

Figure S35. Non-linear curve-fitting for the complexation between 1 and 6·Br in CDCl$_3$. $R^2= 0.9995$
Figure S36. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 7·Br. From bottom to top, the concentration of 7·Br was 0–60 mM.

Figure S37. Non-linear curve-fitting for the complexation between 1 and 7·Br in CDCl$_3$. $R^2=0.9997$
Figure S38. Partial $^1$H NMR spectra (400 MHz, CDCl$_3$, 298 K) of 1 (1.0 mM) titrated by 8·Br. From bottom to top, the concentration of 8·Br was 0–60 mM.

Figure S39. Non-linear curve-fitting for the complexation between 1 and 8·Br in CDCl$_3$. R$^2$= 0.9996
8. Energy-minimized structure of $3\cdot Cl\subset 1$.

**Figure S40.** Energy-minimized structure of $3\cdot Cl\subset 1$ at the level of B3LYP-D3/6-31G+(d, p) by Gaussian 09 program.[S1] The green atom represents the Cl anion, and the red dashes represent multiple C–H⋯π interactions with H⋯ring center distances of 2.41-3.55 Å, the black dashes represent double C–H⋯O hydrogen bonds with H⋯O distances of 2.93 and 3.12 Å, the yellow dashes represent C–H⋯Cl$^-$ hydrogen bonds with H⋯Cl distances of 2.13 and 2.20 Å.
Figure S41. Energy-minimized structures of 3·X⊂1 at the level of B3LYP-D3/6-31G+(d,p), X = (a) Cl, (b) Br, (c) I, and (d) PF₆⁻.

Table S1. The distances between N atoms and anion central atoms in the energy-minimized structure of 3·X⊂1 complexes (X = Cl, Br, I, and PF₆⁻). The N₁Host and N₂Host atoms are two carbazole N atoms directed into the cavity.

<table>
<thead>
<tr>
<th></th>
<th>N₁Host-Anion</th>
<th>N₂Host-Anion</th>
<th>Ncation-Anion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cl⁻</td>
<td>3.108</td>
<td>3.087</td>
<td>3.770</td>
</tr>
<tr>
<td>Br⁻</td>
<td>4.120</td>
<td>3.331</td>
<td>4.132</td>
</tr>
<tr>
<td>I⁻</td>
<td>4.627</td>
<td>3.827</td>
<td>4.449</td>
</tr>
<tr>
<td>PF₆⁻</td>
<td>5.572</td>
<td>4.466</td>
<td>5.414</td>
</tr>
</tbody>
</table>
9. **Crystal data of 1 (CCDC Number: 2057790).**

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification code</td>
<td>zyl_syn_0m_sq</td>
</tr>
<tr>
<td>Empirical formula</td>
<td>C_{348}H_{296}N_{12}O_{48}</td>
</tr>
<tr>
<td>Formula weight</td>
<td>5413.95</td>
</tr>
<tr>
<td>Temperature</td>
<td>273(2) K</td>
</tr>
<tr>
<td>Wavelength</td>
<td>0.77485 Å</td>
</tr>
<tr>
<td>Crystal system, space group</td>
<td>Tetragonal, I4/acd</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>a = 65.655(6), b = 65.655(6), c = 18.9177(18) Å, α = 90°, β = 90°, γ = 90°</td>
</tr>
<tr>
<td>Volume</td>
<td>81546(17) Å³</td>
</tr>
<tr>
<td>Z, Calculated density</td>
<td>8, 0.882 g/cm³</td>
</tr>
<tr>
<td>Absorption coefficient</td>
<td>0.070 mm⁻¹</td>
</tr>
<tr>
<td>F(000)</td>
<td>22816</td>
</tr>
<tr>
<td>Crystal size</td>
<td>0.120 × 0.110 × 0.100 mm</td>
</tr>
<tr>
<td>Theta range for data collection</td>
<td>1.353 to 25.529 deg.</td>
</tr>
<tr>
<td>Limiting indices</td>
<td>-72&lt;=h&lt;=73, -72&lt;=k&lt;=72, -21&lt;=l&lt;=21</td>
</tr>
<tr>
<td>Reflections collected / unique</td>
<td>317878 / 14654 [R(int) = 0.2024]</td>
</tr>
<tr>
<td>Completeness to theta = 25.529</td>
<td>99.6 %</td>
</tr>
<tr>
<td>Absorption correction</td>
<td>Semi-empirical from equivalents</td>
</tr>
<tr>
<td>Max. and min. transmission</td>
<td>0.993 and 0.993</td>
</tr>
<tr>
<td>Refinement method</td>
<td>Full-matrix least-squares on F²</td>
</tr>
<tr>
<td>Data / restraints / parameters</td>
<td>14654 / 1014 / 990</td>
</tr>
<tr>
<td>Goodness-of-fit on F²</td>
<td>1.061</td>
</tr>
<tr>
<td>Final R indices [I&gt;2sigma(I)]</td>
<td>R1 = 0.1171, wR2 = 0.2802</td>
</tr>
<tr>
<td>R indices (all data)</td>
<td>R1 = 0.1756, wR2 = 0.2985</td>
</tr>
<tr>
<td>Extinction coefficient</td>
<td>0.00075(3)</td>
</tr>
<tr>
<td>Largest diff. peak and hole</td>
<td>0.368 and -0.327 e.A⁻³</td>
</tr>
</tbody>
</table>
Reference: