Electronic supporting information for:

Cycloaddition to an anthracene derived macrocyclic receptor with supramolecular control of regioselectivity

Jeremiah J. Gassensmith, Jeffrey M. Baumes, Jens Eberhard and Bradley D. Smith*

Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, USA. *smith.115@nd.edu

Contents

Synthesis and Characterization of adduct families 2 and 3 ..... S2

Procedure to synthesize 2a S2

NMR elucidation of 2a: S3

UV-Vis and Fluorescence spectra of 2a S6

Procedure to synthesize 2b S6

Procedure to synthesize 2c S7

Procedure to synthesize 2d S8

General procedure to synthesize adduct family 3 ..... S10

X-ray Crystallography S12

Crystallographic summary of 1a: S12

Crystallographic summary of 3b: *S14*

Computational Studies **S17**

Modeling summary of 2a *S17*

Modeling summary of pre-reaction complex *S17*

References **S19**

Synthesis and Characterization of adduct families 2 and 3

Macrocycles **1a** and **1b** were synthesized according to previously reported procedures.[1]

Procedure to synthesize 2a

Macrocycle **1a** (10 mg, 0.012 mmol) was dissolved in $CDCl_3$ (2 mL) with *N*-ethylmaleimide (1.4 mg, 0.012 mmol). The solution was heated at 50 °C overnight to produce **2a** quantitatively and without need for further purification. Alternatively, after dissolving the compounds in chloroform, the solvent could be removed under vacuum and heated to 40 °C overnight to produce **2a** quantitatively and without need for further purification. λ_{max} (abs, $CHCl_3$) = 242 nm, 262 nm, 359 nm, 378 nm, 399 nm; λ_{max} (em, $CHCl_3$; λ_{ex}: 250 nm) = 404 nm, 427 nm, 453 nm; δ_H (600 MHz, $CDCl_3$) 8.45 (s, 2 H), 8.43 (dd, *J*=7.0, 3.2 Hz, 2 H), 8.36 (s, 2 H), 8.29 (dd, *J*=7.0, 3.2 Hz, 2 H), 8.10 (s, 2 H), 8.08 (dd, *J*=6.5, 3.5 Hz, 2 H), 7.54 - 7.46 (m, 4 H), 7.43 (d, *J*=10.0 Hz, 2 H), 7.34 (dd, *J*=6.5, 3.2 Hz, 2 H), 6.99 (d, *J*=9.7 Hz, 2 H), 6.83 (t, *J*=2.9 Hz, 2 H), 6.34 (dd, *J*=15.0, 9.7 Hz, 2 H), 5.76 (dd, *J*=14.4, 10.3 Hz, 2 H), 5.00 (d, *J*=14.7 Hz, 2 H), 4.81 (br. s., 2 H), 4.32 (d, *J*=14.1 Hz, 2 H), 3.00 (s, 2 H), 1.43 (s, 18 H), 0.44 (q, *J*=7.6 Hz, 2 H), -1.61 ppm (t, *J*=7.3 Hz, 3 H);

δ_C (150 MHz, $CDCl_3$) 180.4, 165.6, 165.1, 153.4, 135.4, 134.3, 133.2, 133.0, 131.1, 130.8, 130.6, 130.0, 129.5, 129.3, 129.2, 127.6, 127.5, 125.2, 124.9, 124.8, 124.7, 119.4, 45.9, 38.6, 35.8, 35.4, 35.3, 31.2, 29.7, 11.1 ppm; *m/z* (FAB) 970 [$(M+H)^+$], 971 (31).

NMR elucidation of 2a:

Evidence that **2a** is the 1,4-endo isomer includes:

A. Comparison of chemical shifts with literature compounds.[2]

B. COSY connectivities. (Figure S3)

C. ROESY correlation between protons (*a*) and (*j*). (Figure S2)

D. Extreme upfield chemical shifts of *N*-ethyl signals of (*p*) and (*q*). (Figure S1)

E. Downfield chemical shifts of NH and proton (*s*). (Figure S1)

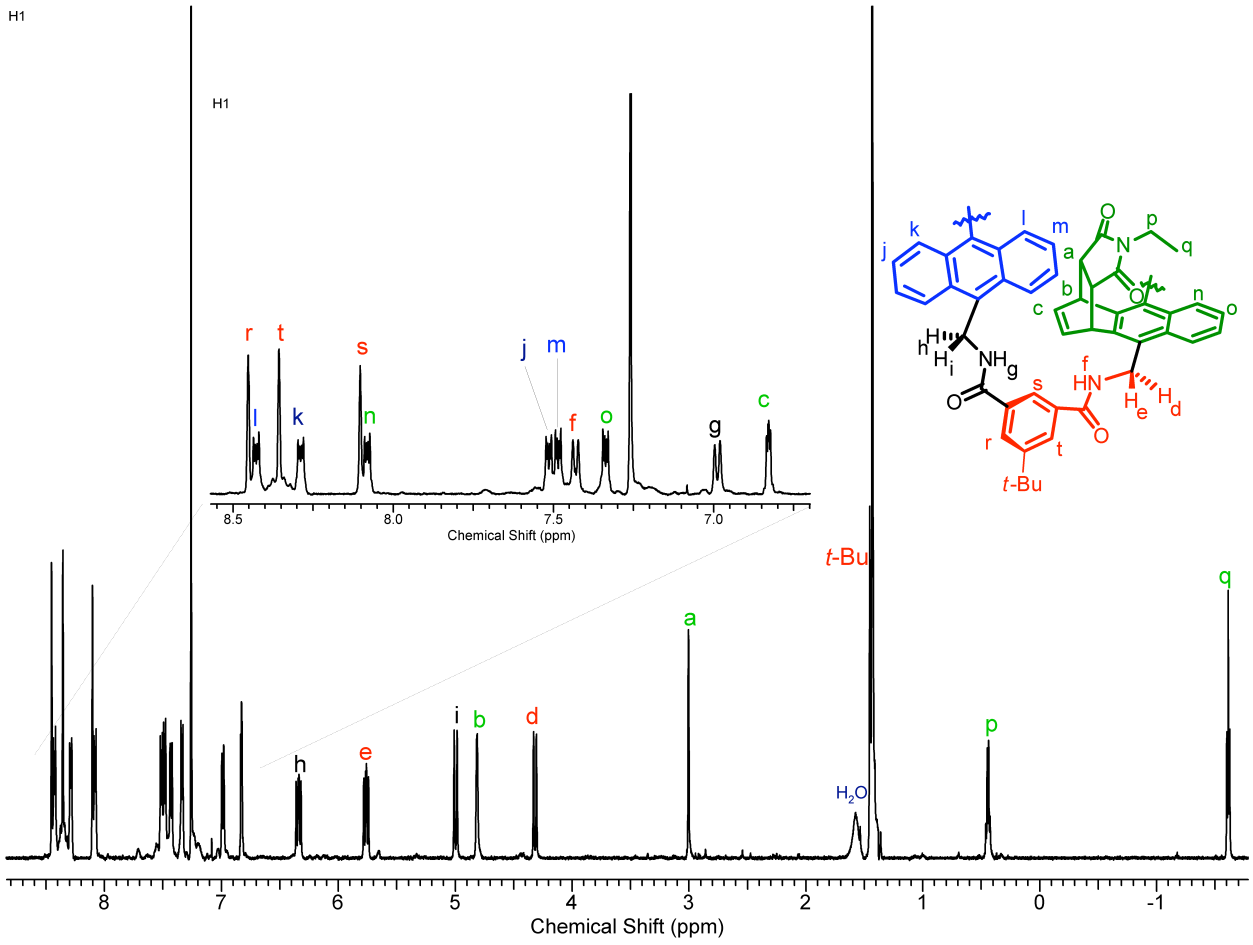

Figure S1: ^{1}H NMR of 2a in $CDCl_3$.

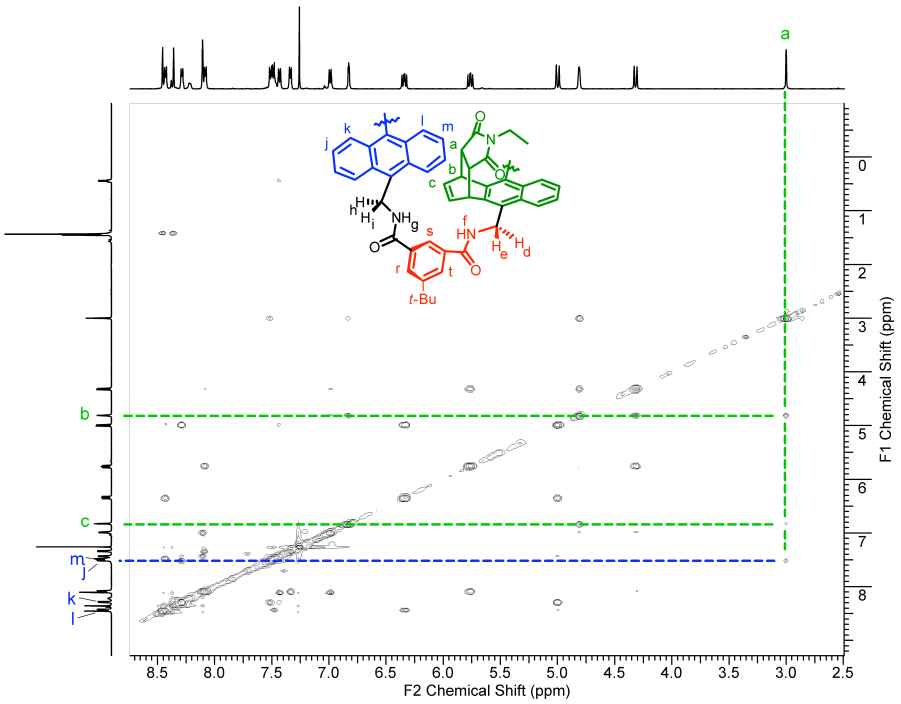

Figure S2: ROESY spectrum of 2a in $CDCl_3$.

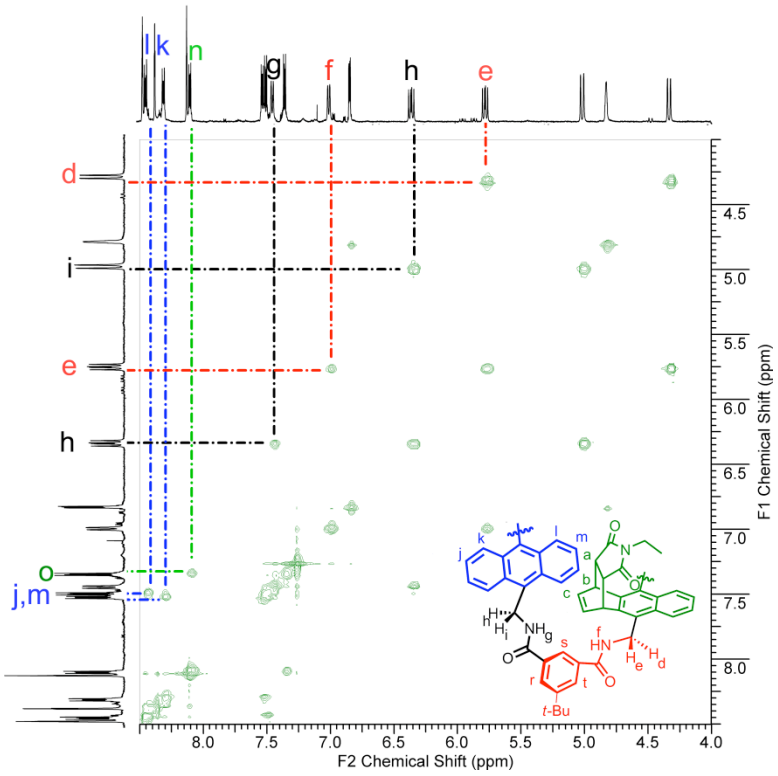

Figure S3: COSY spectrum of 2a in $CDCl_3$.

UV-Vis and Fluorescence spectra of 2a

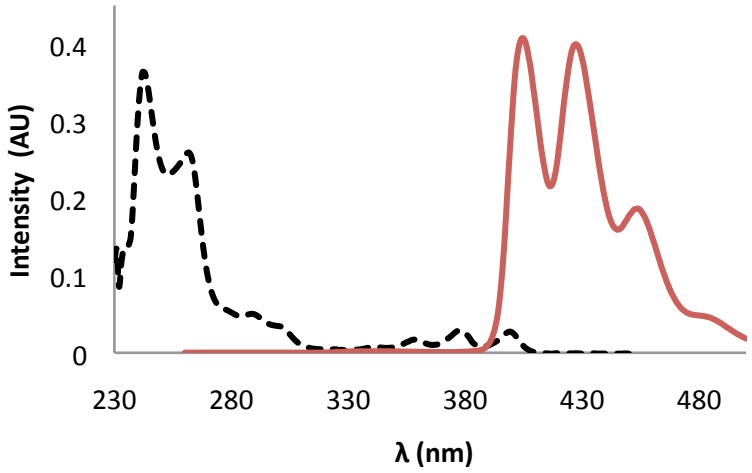

Figure S4: UV-Vis (dotted black line) and Fluorescence (solid maroon line) spectra of 2a in $CHCl_3$. Fluorescence spectrum was obtained by excitation at 250 nm.

Procedure to synthesize 2b

Macrocycle **1b** (10 mg, 0.014 mmol) was dissolved in $CDCl_3$ (1 mL) with *N*-ethylmaleimide (1.7 mg, 0.014 mmol). The solution was heated at 50 °C overnight, then the solvent was removed and the residue purified by column chromatography using silica gel with chloroform:ethyl acetate as eluent (1:3) to give adduct **2b** in 90 % yield. δ_H (600 MHz, $CDCl_3$) 9.29 (d, *J*=10.3 Hz, 2H), 8.57 (dd, *J*=9.2, 1.3 Hz, 2 H), 8.47 (dd, *J*=9.1, 1.2 Hz, 2 H), 8.43 – 8.40 (m, 6 H), 8.10 (dd, *J*=6.5, 3.3 Hz, 2 H), 8.07 (t, *J*=7.9 Hz, 2 H), 7.54 (dd, *J*=7.0, 3.1 Hz, 2 H), 7.46 (dd, *J*=6.9, 3.2 Hz, 2 H), 7.33 (dd, *J*=6.5, 3.2 Hz, 2 H), 6.78 (dd, *J*=4.6, 3.0 Hz, 2 H), 6.38 (dd, *J*=15.1, 10.5 Hz, 2 H), 5.83 (dd, *J*=14.6, 10.9 Hz, 2 H), 5.16 (d, *J*=15.1 Hz, 2 H), 4.88 (br. s., 2 H), 4.42 (dd, *J*=16.0, 1.5 Hz, 2 H), 2.91 (br. s., 2 H), 0.29 (q, *J*=7.4 Hz, 2 H), -1.80 ppm (t, *J*=7.5 Hz, 3 H); δ_C (150

MHz, $CDCl_3$) 180.1, 163.5, 162.9, 149.2, 149.0, 138.6, 135.1, 134.4, 131.0, 130.9, 130.8, 130.1, 128.9, 127.3, 127.2, 126.1, 125.7, 125.2, 125.1, 125.0, 124.9, 46.1, 38.6, 35.2, 35.1, 30.3, 29.7 ppm; *m/z* (MALDI-TOF) 860 [$(M+H)^+$], 860 (100).

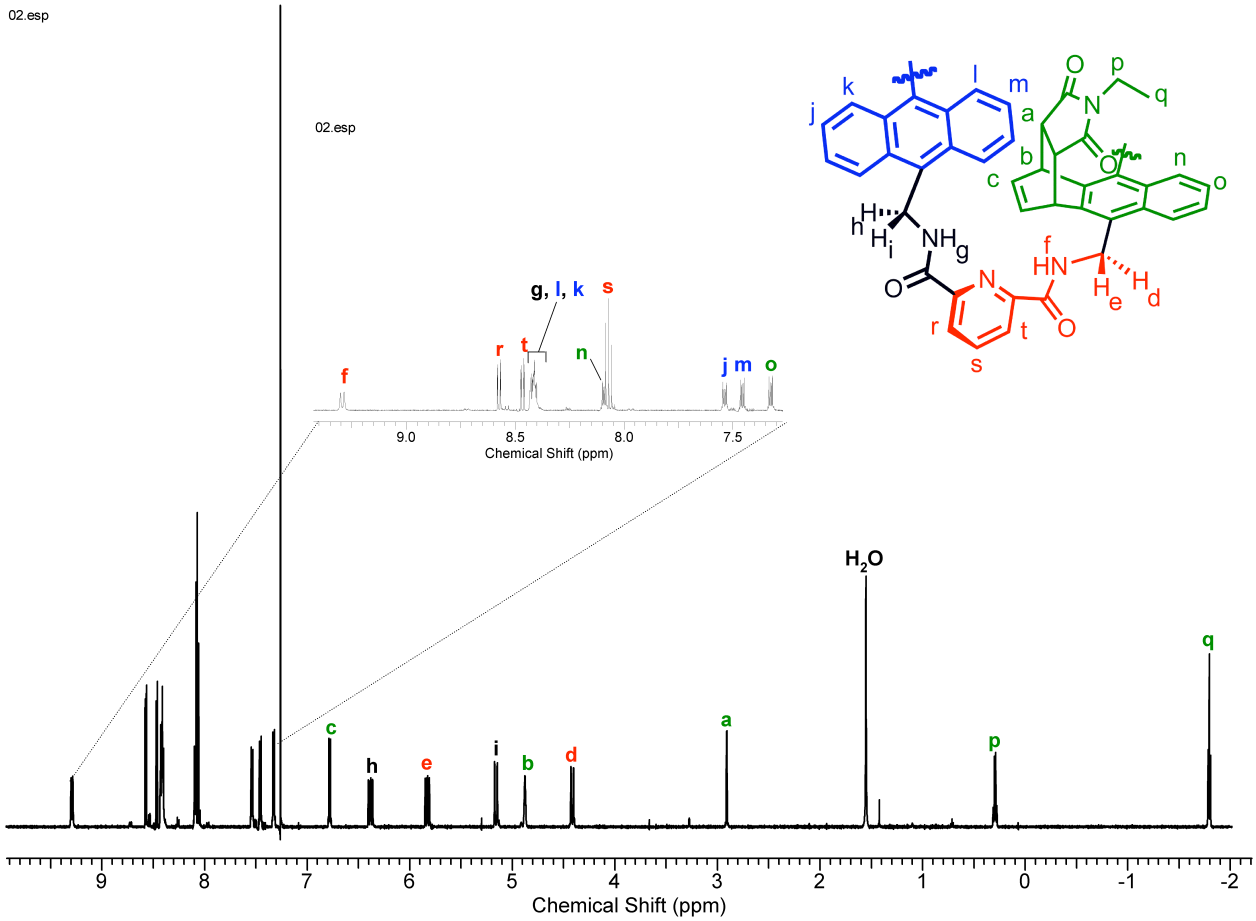

Figure S5: ^{1}H NMR of 2b in $CDCl_3$.

Procedure to synthesize 2c

Macrocycle **1a** (10 mg, 0.012 mmol) and maleic anhydride (1.1 mg, 0.012 mmol) were in dissolved in toluene (2 mL) and heated at 80 °C overnight. The solvent was removed and the residue purified by column chromatography using silica gel with chloroform:ethyl acetate as eluent (1:4). The first fraction collected was unreacted macrocycle followed by adduct **2c** in 70

% yield. δ_H (400 MHz, $CDCl_3$) 8.40 (s, 2 H), 8.30 - 8.36 (m, 4 H), 8.27 (s, 2 H), 8.07 (dd, *J*=6.0, 3.3 Hz, 2 H), 7.71 (s, 2 H), 7.49 (dd, *J*=6.8, 2.8 Hz, 4 H), 7.44 (dd, *J*=6.5, 3.0 Hz, 2 H), 6.97 (d, *J*=8.8 Hz, 2 H), 6.84 (t, *J*=3.8 Hz, 2 H), 6.35 (d, *J*=8.6 Hz, 2 H), 6.17 (dd, *J*=14.4, 8.8 Hz, 2 H), 5.67 (dd, *J*=13.8, 8.8 Hz, 2 H), 5.05 (d, *J*=14.6 Hz, 2 H), 4.89 (br. s., 2 H), 4.40 (d, *J*=14.6 Hz, 2 H), 3.45 (s, 2 H), 1.42 (s, 18 H); ^{13}C NMR spectrum not acquired due to poor solubility; *m/z* (HRMS FAB) calc: 943.4071, obs: 943.4042 $[(M+H)^+ C_{60}H_{55}O_7N_4]$.

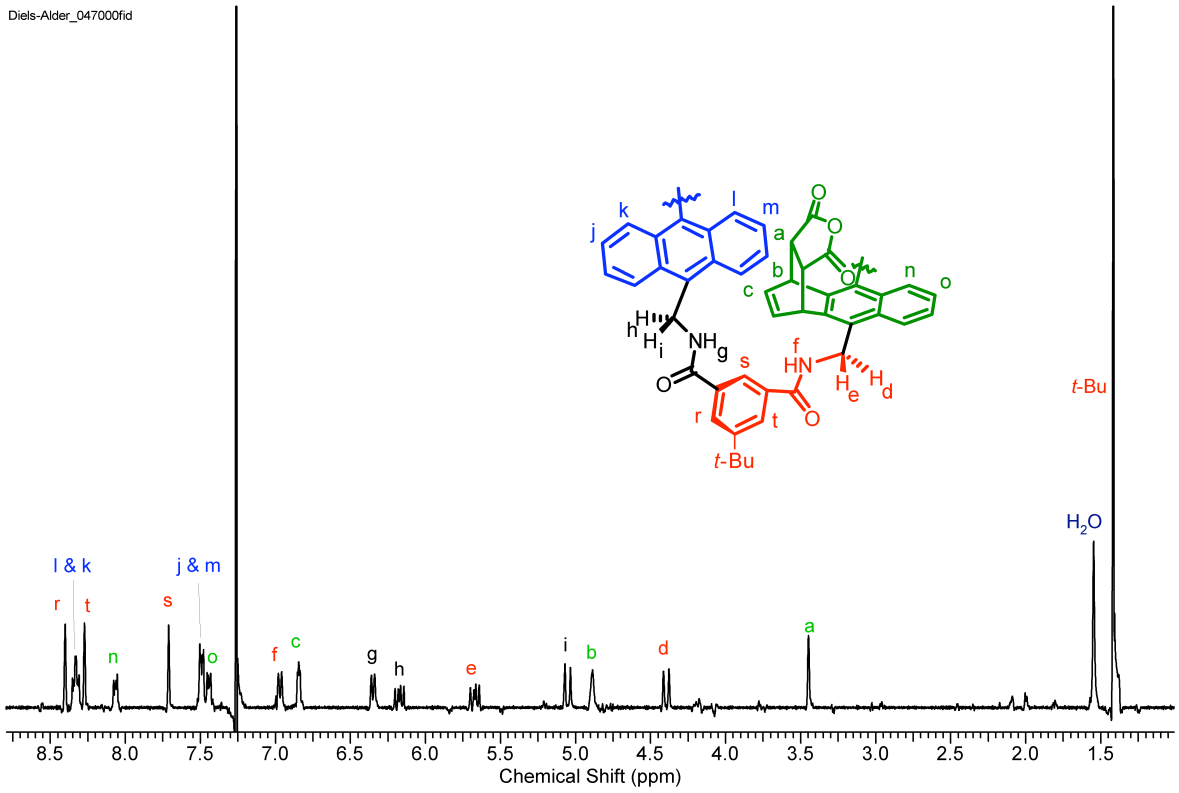

Figure S6: ^{1}H NMR of 2c in $CDCl_3$.

Procedure to synthesize 2d

Macrocycle **1b** (15 mg, 0.020 mmol) and maleic anhydride (2.0 mg, 0.020 mmol) were dissolved in $CDCl_3$ (1 mL) and the solution heated at 50 °C overnight. The solvent was removed

and the residue purified by column chromatography using silica gel with chloroform:ethyl acetate as eluent (1:3) to give **2d** in 68 % yield. δ_H (600 MHz, $CDCl_3$) 8.72 (d, *J*=9.8 Hz, 2 H), 8.54 (dd, *J*=7.9, 1.2 Hz, 2 H), 8.43 (dd, *J*=9.2, 1.3 Hz, 2 H), 8.39 (dd, *J*=7.4, 3.4 Hz, 2 H), 8.26 (dd, *J*=6.9, 3.3 Hz, 2 H), 8.07 – 8.04 (m, 4 H), 7.97 (d, *J*=10.7 Hz, 2 H), 7.50 (dd, *J*=6.9, 3.2 Hz, 2 H), 7.45 (dd, *J*=6.9, 3.2 Hz, 2 H), 7.42 (dd, *J*=6.6, 3.2 Hz, 2 H), 6.79 (dd, *J*=4.6, 3.0 Hz, 2 H), 6.22 (dd, *J*=15.1, 10.2 Hz, 2 H), 5.80 (dd, *J*=14.9, 11.0 Hz, 2 H), 5.14 (d, *J*=15.3 Hz, 2 H), 4.92 (br. s., 2 H), 4.41 (d, *J*=14.8 Hz, 2 H), 3.28 ppm (br. s., 2 H); ^{13}C NMR spectrum not acquired due to poor solubility; *m/z* (MALDI-TOF) 833 [$(M+H)^+$], 833 (40).

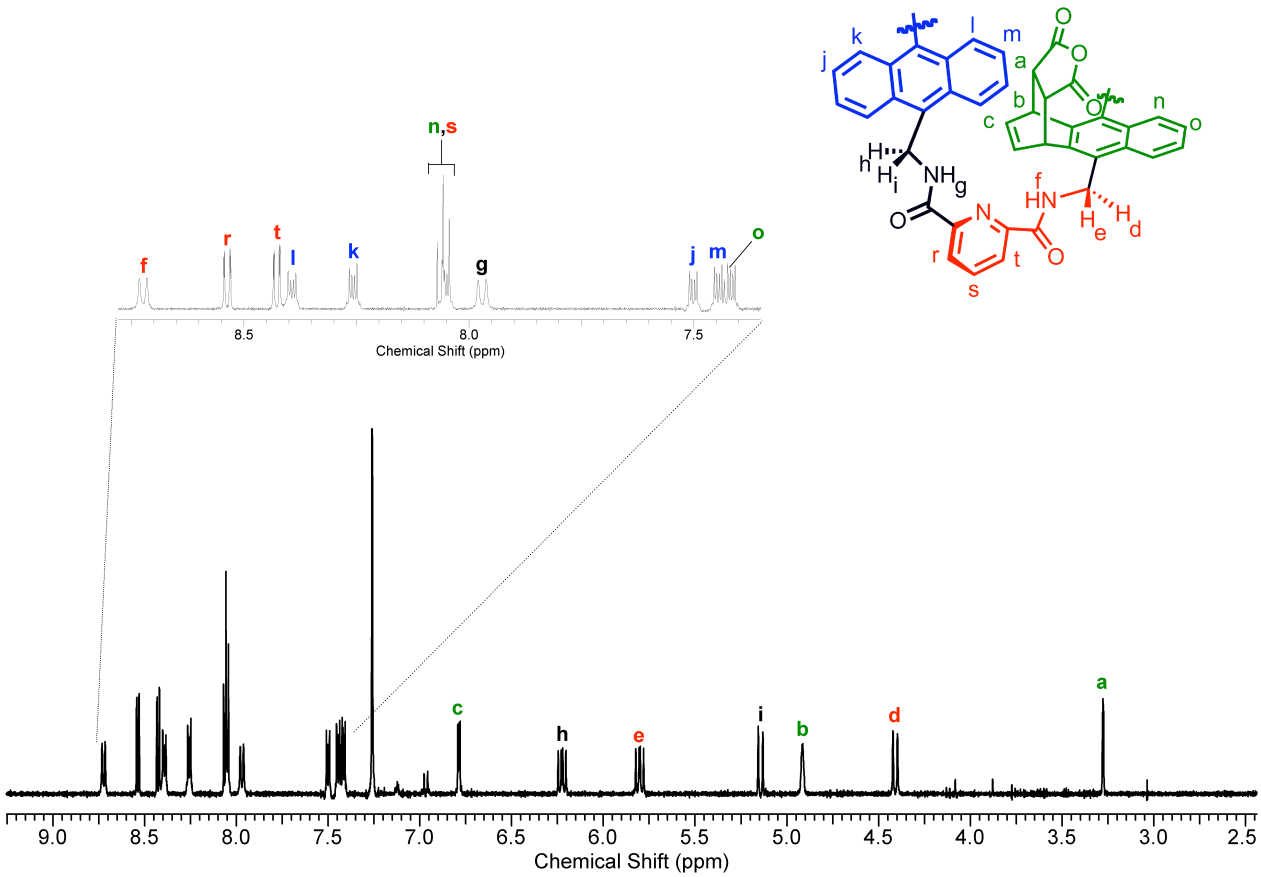

Figure S7: ^{1}H NMR of 2d in $CDCl_3$.

General procedure to synthesize adduct family 3

Macrocycle **1a** or **1b** (0.027 mmol) was dissolved in $CDCl_3$ (1 mL) with 1 mol % Methylene Blue and transferred into a NMR tube. A long, stainless steel needle was placed into the NMR tube and a steady stream of oxygen gas was allowed to bubble through the solution. The NMR tube was fixed 10 cm away from a commercially obtained 40W compact fluorescent lamp (spiral type GE) and irradiated for 20 minutes to produce adduct family **3** quantitatively and without need for further purification. Spectroscopic data for **3a:** δ_H (300 MHz, $CDCl_3$) 8.33 (d, *J*=1.4 Hz, 4 H), 7.58 (s, 2 H), 7.40 (d, *J*=5.5, 3.3 Hz, 8 H), 7.21 (q, *J*=5.5, 3.1 Hz, 8 H), 6.85 (t, *J*=5.5 Hz, 4 H), 4.86 (d, *J*=5.5 Hz, 8 H), 1.40 (s, 18 H); δ_C (75 MHz, $CDCl_3$) 166.6, 153.2, 138.5, 133.0, 129.9, 128.0, 121.3, 119.2, 81.5, 38.1, 35.2, 31.1 ppm; *m/z* (HRMS FAB) calc: 909.3863, obs: 909.3855 [$(M+H)^+$ $C_{57}H_{53}O_8N_4$]. Spectroscopic data for **3b:** δ_H (400 MHz, $CDCl_3$) 8.45 (d, *J*=7.8 Hz, 4 H), 8.38 (t, *J*=5.8 Hz, 4 H), 8.10 (t, *J*=8.1 Hz, 2 H), 7.40 (dd, *J*=5.5, 3.0 Hz, 8 H), 7.23 (dd, *J*=5.5, 3.3 Hz, 8 H), 4.84 ppm (d, *J*=6.0 Hz, 8 H); δ_C (125 MHz, $CDCl_3$) 163.1, 148.0, 139.2, 139.1, 127.7, 125.3, 121.1, 80.9, 37.2 ppm; *m/z* (MALDI-TOF) 821 [$(M+Na)^+$], 821 (20).

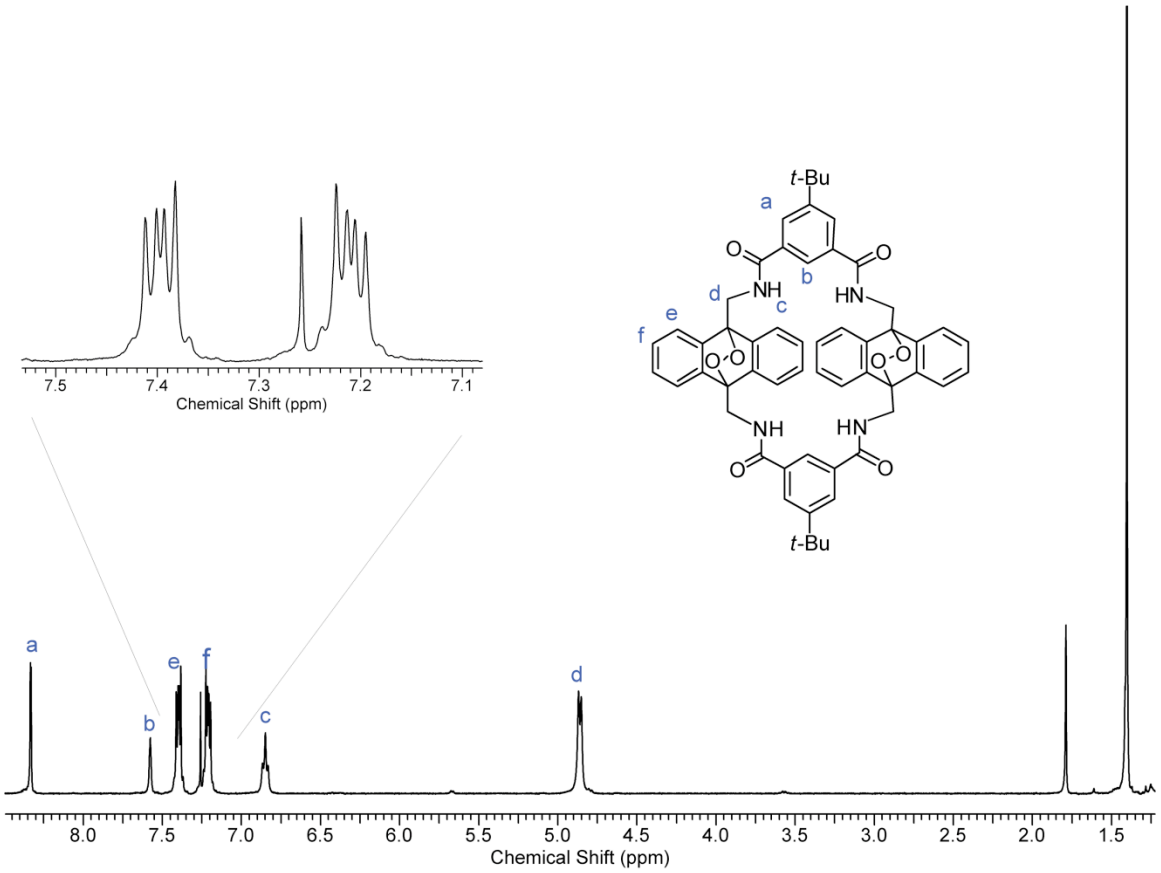

Figure S8: ^{1}H NMR of 3a in $CDCl_3$. Protons *e* and *f* may be transposed.

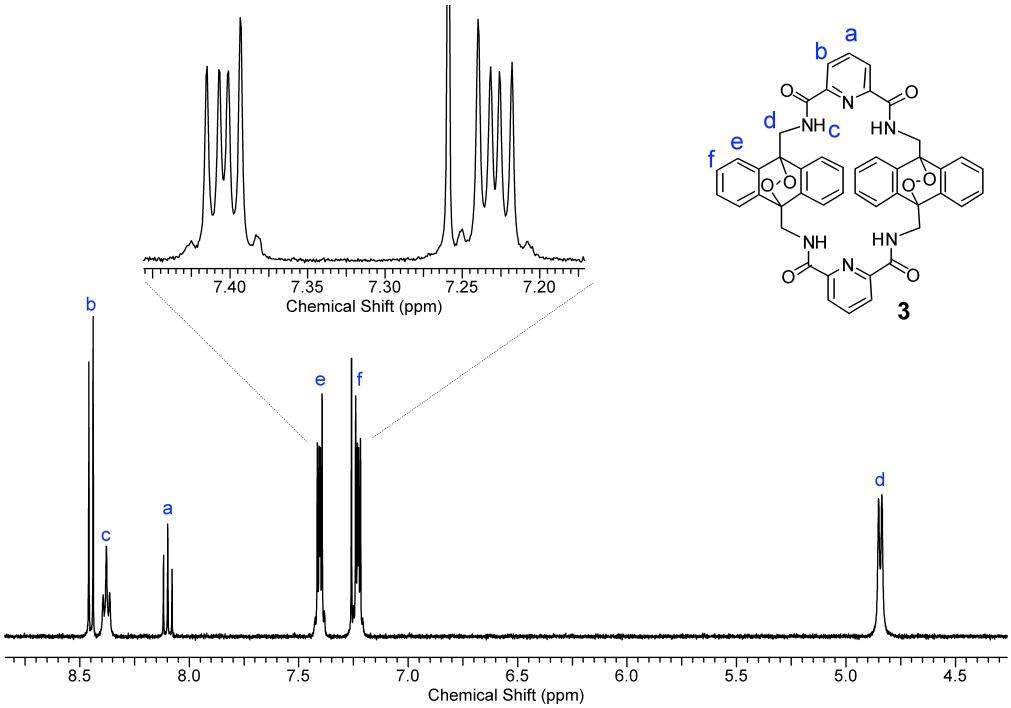

Figure S9: ^{1}H NMR of 3b in $CDCl_3$. Protons *e* and *f* may be transposed.

X-ray Crystallography

Crystallographic summary of 1a:

Single crystals were obtained as follows: a solution of **1a** in 25:75 chloroform : ethyl acetate was prepared in a 2 mL vial and allowed to slowly evaporate. The crystals grew as yellow needles. Crystallographic summary: triclinic, C65 H69 Cl3 N4 O8, FW = 1140.59, P-1, Z = 2 in a cell of dimensions a = 12.5937(5) Å, b = 13.2158(5) Å, c = 18.7174(8) Å, α = 73.655(2) °, ß = 88.986(2) °, γ = 79.997(2) °, V = 2942.3(2) Å^3, D_{calc} = 1.287 mg/m^3, F(000) = 1204.0. 44027 reflections measured and 10733 were independent (R_{int} = 0.0465). The structure was refined on F^2 to *wR2* = 0.1952, conventional *R1* = 0.0649 [7239 reflections with I > 2σ (I)], and a goodness of fit = 1.091 for 755 refined parameters.

Four chemical moieties were found in the asymmetric unit, macrocycle **1a**, two molecules of ethyl acetate and a single disordered chloroform with a site occupancy of 1:4. The structure was solved satisfactorily by direct methods but the partially occupied chloroform did not refine well with a chlorine atom possessing a Ueq(max)/Ueq(min) of 7.85; however, the position and geometry of the bonds leave little doubt as to the atom type. Attempts to split the chlorines into two positions on the recommendations in the list file proved impossible to refine.

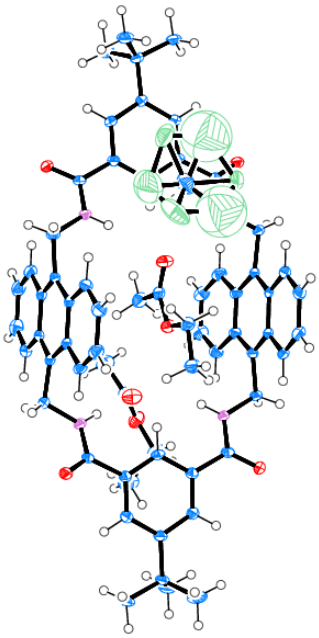

Figure S10: ORTEP of 1a; ellipsoids at 50 % probability.

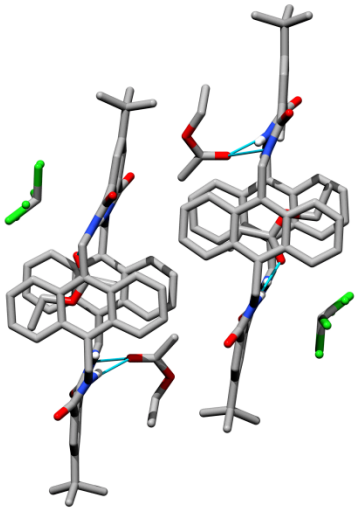

Figure S11: Packing of 1a is showing relative positions of ethyl acetate and chloroform in unit cell.

Crystallographic summary of 3b:

Single crystals were obtained as follows: a solution of **3b** in chloroform was prepared in a 2 mL vial and allowed to slowly evaporate. The crystals grew as yellow needles. Crystallographic summary: monoclinic, C50 H42 Cl12 N6 O9, FW = 1296.30, P 21/c, Z = 4 in a cell of dimensions a = 17.4237(13) Å, b = 18.9917(14) Å, c = 17.0898(12) Å, α = 90 °, 102.159(3) °, γ = 90 °, V = 5528.2(7) $Å^3$, D_{calc} = 1.557 mg/m^3, F(000) = 2640.0. 78882 reflections measured and 7318 were independent (R_{int} = 0.1208). The structure was refined on F^2 to *wR2* = 0.1908, conventional *R1* = 0.0649 [7318 reflections with I > 2σ (I)], and a goodness of fit = 1.084 for 574 refined parameters.

Six chemical moieties were found in the asymmetric unit, four highly disordered molecules of chloroform, one molecule of water and macrocycle **3b**. The crystal readily solved in the P

21/c space group using direct methods. After convergence, a significant amount of residual electron density remained in the difference Fourier map. Attempts to model this as four molecules of chloroform did not allow for adequate refinement, thus PLATON/squeeze was used to correct the data. Three of the chloroforms in the asymmetric unit were part of a packing arrangement of stratified layers, poorly bonded to the separate macrocycles (see Figure S8). One solvent of chloroform, which was highly disordered, appeared to sit in the pocket of the macrocycle, which forms a basket like structure. The hydrogens of all the amides and water were found on the difference Fourier map. Both hydrogens attached to the water have elongated H---O bonds (1.1 Å and 1.2 Å) though neither appear to form typical strong hydrogen bonds to anything on the basis of the O---H---O angles (157.9° O---H---O to amide of another macrocycle in the unit cell and 126.4° O---H---O to endoperoxide).

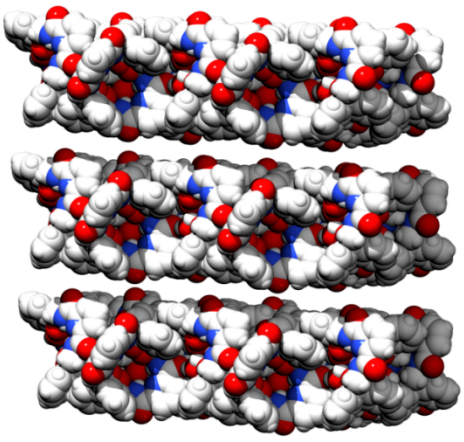

Figure S12: Packing arrangement of 3b showing large solvent voids where significant amount of disordered solvent was removed with PLATON/squeeze.

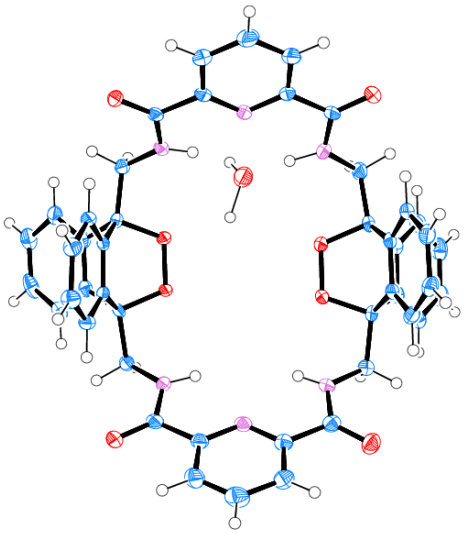

Figure S13: ORTEP of 3b; ellipsoids at 50% probability.

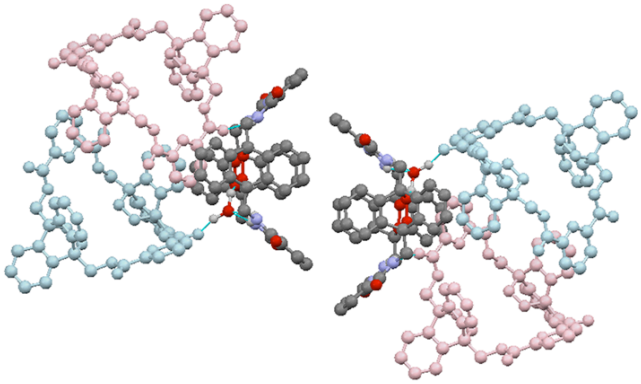

Figure S14: Packing arrangement showing cavity formed by cofacial "basket shaped" molecules. Two highly disordered molecules of chloroform which occupied this space could not be refined and were removed with PLATON Squeeze.

Computational Studies

Modeling summary of 2a

The crystal structure of **1a** served as the macrocycle template. Removal of the ethyl acetate and insertion of *N*-ethylmaleimide 1,4-addition product was done in Gaussian 03.[3] A semi-empirical optimized structure using the AM1 model was imported into the Maestro 8 (Schrodinger) software package.[4] The imported structure was then subjected to a Monte Carlo conformational search using the MacroModel subset. Conformers were found using the OPLS_2005 force field parameters with high quality fit of the stretch, bend, and torsion parameters. A fixed dielectric chloroform model was used in the simulation. The conjugate gradient minimizer (maximum iterations set at 2000) was used with a threshold energy of 50 $kJmol^{-1}$. Using the torsional sampling method to examine 1000 conformers, gave a total of 735 interactions finding 68 unique conformations. The global minimum energy conformer of **2a** was found 49 times.

Modeling summary of pre-reaction complex

The crystal structure of **1a** served as the macrocycle template. Removal of the ethyl acetate and insertion of *N*-ethylmaleimide guest was done in Gaussian 03[3]. A semi-empirical optimized structure using the AM1 model was imported into the Maestro 8 (Schrodinger) software package.[4] The imported structure was then subjected to a Monte Carlo conformational search using the MacroModel subset. Conformers were found using the OPLS_2005 force field parameters with high quality fit of the stretch, bend, and torsion parameters. A fixed dielectric chloroform model was used in the simulation. The conjugate gradient minimization method (maximum iterations set at 2000) was used with a threshold

energy of 50 $kJmol^{-1}$. Using the torsional sampling method to examine 1000 conformers, gave a total of 698 interactions finding 158 unique conformations. The energetically lowest pre-reaction conformer was found 5 times.

References

[1] J. J. Gassensmith, E. Arunkumar, L. Barr, J. M. Baumes, K. M. DiVittorio, J. R. Johnson, B. C. Noll and B. D. Smith, *J. Am. Chem. Soc.*, 2007, **129**, 15054-15059.

[2] M. Yoshizawa, M. Tamura and M. Fujita, *Science*, 2006, **312**, 251-254.

[3] *Gaussian 03*, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford, CT, 2004.

[4] *Maestro*, version 8.0; Schrödinger, LLC: New York, NY, 2005.