Dynamic and Rotational Analysis of Cryptophane Host−Guest Systems: Challenges of Describing Molecular Recognition

Paul D. Kirchhoff,*,†,‡ Jean-Pierre Dutasta,§ André Collet,§ and J. Andrew McCammon‡

Contribution from the Stéréochimie et Interactions moléculaires, École Normale Supérieure de Lyon, 46 allée d'Italie, F-69364 Lyon Cedex 07, France, and Department of Chemistry and Biochemistry and Department of Pharmacology, University of California at San Diego, La Jolla, California 92093-0365

Received May 4, 1998. Revised Manuscript Received September 15, 1998

Abstract: Cryptophanes are aromatic hosts which bind a variety of guests. Here, we describe three 25 ns molecular dynamics simulations of a particular cryptophane in water. Simulations have been conducted on the uncomplexed cryptophane, the cryptophane−tetramethylammonium ion (TMA^+) complex, and the cryptophane−neopentane (NEO) complex. TMA^+ and NEO are both tetrahedral species and are nearly isomorphic. In the current study, we examine how the presence of these guests influences motions of the host. Also examined are the preferred orientations and the motions of the guests relative to the cryptophane. This study demonstrates some of the many challenges of describing molecular recognition.

Introduction

Host−guest association is central to many areas of chemistry and biochemistry. It is important in areas ranging from the practical, such as pharmaceutical research, to the fundamental, such as clarifying the principles of molecular recognition.[1−8] In much of host−guest chemistry, the practical and fundamental are intertwined. Studies of quaternary ammonium cations binding to artificial receptors have been used to better understand the complexation of acetylcholine and related biomolecules to their natural receptors.[9−14]

Cryptophanes are a group of cagelike host molecules that have been well characterized in organic and aqueous solvents.[15−22] They form reversible complexes with a wide variety of guests. The binding constant of a guest with a particular cryptophane depends on a number of factors, including size, shape, and net charge. However, the amount that each of these factors contributes to the binding constant of a guest may vary between complexes, even those with similar overall stabilities. Because of the wide variety of complexes, cryptophanes are useful in studying issues of molecular recognition.

The applicability of cryptophanes for the study of molecular recognition was recently demonstrated by their use in solving a decades-old question of chirality. Using computational techniques to characterize complexes with a chiral cryptophane, the absolute configuration of bromochlorofluoromethane (CHFClBr) was determined to be (R)-(−) and (S)-(+).[23] This conclusion was subsequently verified by independent assignment based on Raman optical activity studies.[24,25]

In this study, the analysis of a new, 25 ns simulation of a cryptophane−NEO complex is compared to the previously reported simulations of the cryptophane and the cryptophane−TMA^+ complex. TMA^+ and NEO were chosen as guests in the simulations because they are nearly isomorphic and have been used in experimental studies of cryptophane host−guest complexes. The ability of the guests to influence the motions of the host is examined, as well as details of interactions and motions

† Current address: Molecular Simulations Inc., 9685 Scranton Road, San Diego, CA 92121-3752.
‡ University of California at San Diego.
§ École Normale Supérieure de Lyon.

(1) Mordasini Denti, T. Z.; van Gunsteren, W. F.; Diederich, F. *J. Am. Chem. Soc.* **1996**, *118*, 6044−6051.
(2) Houk, K. N.; Nakamura, K.; Sheu, C.; Keating, A. E. *Science* **1996**, *273*, 627−629.
(3) Sheu, C.; Houk, K. N. *J. Am. Chem. Soc.* **1996**, *118*, 8056−8070.
(4) Fox, T.; Thomas, B. E., IV; McCarrick, M.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 10779−10783.
(5) Böhm, H.-J.; Klebe, G. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2588−2614.
(6) Wallimann, P.; Mattei, S.; Seiler, P.; Diederich, F. *Helv. Chim. Acta* **1997**, 80, 2368−2390.
(7) Nakamura, K.; Sheu, C.; Keating, A. E.; Houk, K. N.; Sherman, J. C.; Chapman, R. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1997**, *119*, 4321−4322.
(8) Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359−1472.
(9) Dougherty, D. A.; Stauffer, D. A. *Science* **1990**, *250*, 1558−1560.
(10) Méric, R.; Lehn, J.-M.; Vigneron, J.-P. *Bull. Soc. Chim. Fr.* **1994**, *131*, 579−583.
(11) Dougherty, D. A. *Science* **1996**, *271*, 163−168.
(12) Mecozzi, S.; West, A. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10566−10571.
(13) Murayama, K.; Aoki, K. *J. Chem. Soc., Chem. Commun.* **1997**, 119−120.
(14) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303−1324.
(15) Canceill, J.; Cesario, M.; Collet, A.; Guilhem, J.; Lacombe, L.; Lozach, B.; Pascard, C. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 1246−1248.
(16) Collet, A.; Dutasta, J.-P.; Lozach, B.; Canceill, J. *Top. Curr. Chem.* **1993**, *165*, 103−129.
(17) Collet, A.; Dutasta, J.-P.; Lozach, B. *Adv. Supramol. Chem.* **1993**, *3*, 1−35.
(18) Garel, L.; Lozach, B.; Dutasta, J.-P.; Collet, A. *J. Am. Chem. Soc.* **1993**, *115*, 11652−11653.
(19) Garel, L.; Vezin, H.; Dutasta, J.-P.; Collet, A. *J. Chem. Soc., Chem. Commun.* **1996**, *6*, 719−720.
(20) Collet, A. In *Cryptophanes*; Vögtle, F., Ed.; Comprehensive Supramolecular Chemistry, Vol. 2; Pergamon: New York, 1996; Chapter 11, pp 325−365.
(21) Garel, L.; Dutasta, J.-P.; Collet, A. *New J. Chem.* **1996**, *20*, 1265−1271.
(22) Bartik, K.; Luhmer, M.; Dutasta, J.-P.; Collet, A.; Reisse, J. *J. Am. Chem. Soc.* **1998**, *120*, 784−791.
(23) Costante-Crassous, J.; Marrone, T. J.; Briggs, J. M.; McCammon, J. A.; Collet, A. *J. Am. Chem. Soc.* **1997**, *119*, 3818−3823.
(24) Costante, J.; Ehlinger, N.; Perrin, M.; Collet, A. *Enantiomer* **1996**, *1*, 377−386.
(25) Costante, J.; Hecht, L.; Polaravapu, P. L.; Collet, A.; Barron, L. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 885−887.

10.1021/ja981526a CCC: $18.00
Published on Web 12/29/1998

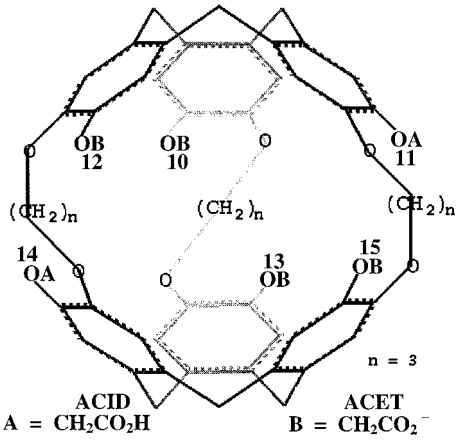
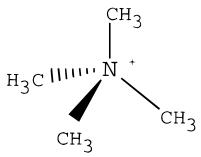


Figure 1. Structure of the cryptophane used in this study. The arrangement of the ACET and ACID groups on the cryptophane is shown along with system residue numbers.

Table 1. Structures of the Two Guests Used in This Study Along with a Number of Their Parameters

TMA^+

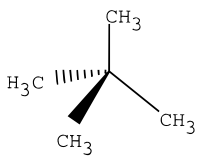


OPLS Atom Type	Epsilon (kcal/mol)	Sigma (Å)
101 N3	0.170	3.25
290 CT	0.066	3.50
295 HC	0.030	2.50

Pair	Equilibrium Separation (Å)
101 N3 – 290 CT	1.471
290 CT – 295 HC	1.090
295 N3 – 295 HC	2.103

OPLS Atom Type	Charge
101 N3	0.208
290 CT	–0.270
295 HC	0.156

NEO



OPLS Atom Type	Epsilon (kcal/mol)	Sigma (Å)
294 CT	0.066	3.50
290 CT	0.066	3.50
295 HC	0.030	2.50

Pair	Equilibrium Separation (Å)
294 CT – 290 CT	1.526
290 CT – 295 HC	1.090
294 CT – 295 HC	2.151

OPLS Atom Type	Charge
294 CT	0.656
290 CT	–0.347
295 HC	0.061

of the guests relative to the cryptophane host. Despite the similarity in shape of the guests, notable differences are seen in their complexes with the cryptophane. The current comparisons modify earlier conclusions and demonstrate some of the many challenges and subtleties of describing molecular recognition.

Methods

A representation of the cryptophane used in this study is displayed in Figure 1. The cryptophane is a hexaacid derivative of cryptophane-E.[18] It consists of two cyclotriveratrylene groups, each comprising three aromatic rings connected by methylene bridges. An acetic acid group is joined to each aromatic ring through an ether linkage. The two cyclotriveratrylene groups are joined by three phenolic propyl linkers to form a roughly spherical molecule with a cavity. The propyl linkers along with the acetic acid groups form pores through which a guest must pass to bind with the cryptophane. Representations of two guests which bind within the cryptophane cavity and were used in this study are displayed in Table 1 along with a number of the parameters used for these guests.

This study is based on the analysis of three separate molecular dynamics simulations. The cryptophane host, the cryptophane−TMA^+ complex, and the cryptophane−NEO complex were each simulated for 25 ns in explicit water systems. The cryptophane host simulation is an extension of a 20 ns simulation reported earlier.[26] The simulation of the cryptophane−TMA^+ complex was also reported recently;[27] however, the simulation of the cryptophane−NEO complex is new. Consistency has been maintained between the three simulations. The same parameters have been used to describe the cryptophane and solvent, and the equilibration and molecular dynamics protocol is also the same. For completeness, the methods used for these simulations will be outlined here; many of the details can be found in the previous papers.[26,27]

The molecular dynamics simulations were conducted at 298 K in the NPT ensemble using ARGOS.[28] The cryptophane host was modeled in exactly the same way for each of the simulations with four of its six acid protons removed. The arrangement of the two remaining protons is illustrated in Figure 1, resulting in one of the cryptophane pores containing an ACET/ACET pair and the other two pores containing ACET/ACID pairs. ACID and ACET refer to the protonated and deprotonated acidic side chains, respectively. As will be addressed later in this study, sampling in the cryptophane is sensitive to the choice of the acid proton arrangement. Other details of the cryptophane parameters are left to ref 26.

Some of the parameters used for the TMA^+ and NEO guests are listed in Table 1. (Atomic charges for NEO were determined in a fashion similar to those used for the TMA^+ which is described in ref 27.) From the table, it may be seen that the two guests are very similar, with TMA^+ and NEO being nearly isomorphic. The equilibrium separation between one of the hydrogen atoms and the central atom of each guest is a good measure of the size difference between the two species. For TMA^+, the equilibrium separation is 2.10 Å. NEO is slightly larger, with an equilibrium separation of 2.15 Å. The main difference between the two guests is the overall charge of +1 on TMA^+ versus the overall neutral charge on NEO. The charges, however, are distributed symmetrically over the guests and no standard hydrogen-bonding interactions are possible between either guest and the cryptophane.

Results and Discussion

Conformational Analysis. Conformational analysis was conducted on the cryptophane molecule in each of the simulations to determine how the presence of the guests affects sampling of the host. As in the earlier studies, the rate at which unique conformations were sampled with respect to simulation time was computed by analyzing the cryptophane's dihedral angles.[26] Only angles for the host were examined.

Conformational analysis was conducted on the "whole" cryptophane and on two subsets of the cryptophane. The "exo" region was defined as consisting of the ACID and ACET groups up through and including the ether oxygens. The remainder of the cryptophane was defined as the "cage" region. The fluctuations of the cage region may be particularly relevant to the binding of a guest. For the whole and cage regions of the cryptophane, the equivalence of generated structures was also considered. Two structures were considered equivalent if the molecule could be rotated so that the corresponding rotameric states were the same.

Plots for the sampling of unique conformations with equivalent structures removed versus simulation time are shown for the cage region in Figure 2 for each of the systems. One would expect that the presence of one large guest within the cavity as opposed to several smaller guests, such as water molecules, would make the cryptophane molecule, and especially its cage

(26) Kirchhoff, P. D.; Bass, M. B.; Hanks, B. A.; Briggs, J. M.; Collet, A.; McCammon, J. A. *J. Am. Chem. Soc.* **1996**, *118*, 3237−3246.
(27) Kirchhoff, P. D.; Dutasta, J.-P.; Collet, A.; McCammon, J. A. *J. Am. Chem. Soc.* **1997**, *119*, 8015−8022.
(28) Straatsma, T. P.; McCammon, J. A. *J. Comput. Chem.* **1990**, *11*, 943−951.

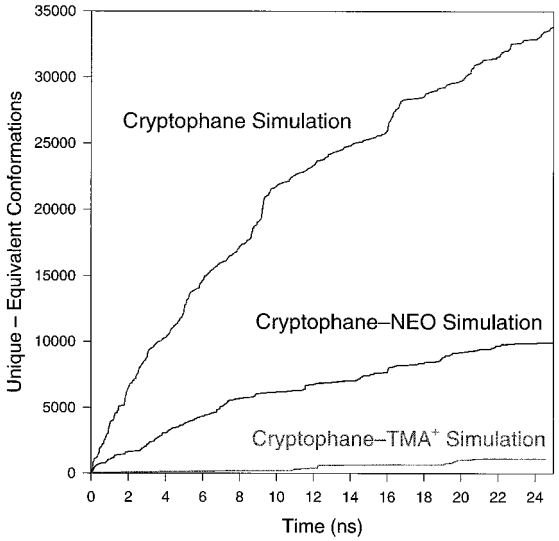


Figure 2. Unique conformations sampled with simulation time for the cage region of the cryptophane. Equivalent structures have been removed.

region, more rigid. The increased rigidity would be expected to decrease the rate at which unique conformations were sampled. As can be seen in Figure 2, this is exactly what is observed. The cryptophane molecule in both of the complexes showed significantly fewer unique conformations than the uncomplexed cryptophane.

What would not necessarily be expected is the large difference in sampling of the cryptophane molecule in the two complexes. The cryptophane molecule in the host−NEO simulation showed significantly more conformations than the cryptophane molecule in the host−TMA^+ simulation. The NEO guest is actually slightly larger than TMA^+, so it is clear that steric interactions are not the only factors influencing sampling in the cryptophane host.

Results for the conformational analysis are displayed in Table 2 for the three simulations of each of the regions. As can be seen in the table (and Figure 2), the cage region of the host−TMA^+ and host−NEO simulations showed about 3% and 29%, respectively, of the number of conformations sampled in the host simulation. For the whole cryptophane, the host−TMA^+ and host−NEO simulations showed about 20% and 89%, respectively. In the exo region, the host−TMA^+ and host−NEO simulations showed about 74% and 113%, respectively. The 13% increase in sampling of the exo region appears to be significant. Analysis presented later in this study suggests that the presence of NEO may actually increase sampling of the cryptophane's carboxyl groups relative to the host simulation.

It is interesting to note that the conformational analysis suggests that the binding of TMA^+ results in a greater entropic penalty to the host than the binding of NEO. By the 15th ns (data not shown), sampling has converged such that the uncomplexed host is sampling 5 times more conformations than the host in the host−TMA^+ simulation but only 1.1 times more than the host in the host−NEO simulation. Assuming that these reflect the converged number of distinct conformations available to the host and that these conformations are of similar energy, binding of the guest would result in an unfavorable change in the entropy of $\Delta S_{conf} = R \ln(1/5)$ for TMA^+, but only $\Delta S_{conf} = R \ln(1/1.1)$ for NEO. Of course, this is only the change in conformational entropy for the host and does not include the overall entropic change for the binding process.

Pore and Cavity Calculations. Pore and cavity radii were determined for the cryptophane host molecule excluding the presence of guest molecules.[26] The analysis was conducted every 0.05 ps for the host simulation and every 0.1 ps for the host−guest simulations. The radii were determined to further describe the structural character of the systems and to gain an indication of how the presence of a guest affects pore and cavity fluctuations. This analysis also provided a means to determine effective radii for the TMA^+ and NEO guests for comparisons with the calculated pore and cavity radii.

Histograms of cavity and pore radii are displayed in Figures 3 and 4, respectively, with plots of cumulative times through 5, 10, 15, 20, and 25 ns for each of the three simulations. These plots show the distribution and convergence of the cavity and pore radii. Figure 5 displays plots of cavity and pore radii as a function of simulation time for each of the three simulations. The data are plotted at 5 ps intervals. The range and average radii for all structures analyzed in each of the simulations are displayed on the corresponding graphs of Figure 5.

Histograms of the cavity radii are basically as one would expect on comparing the host to the host−guest simulations. The presence of one large guest (TMA^+ or NEO) as opposed to zero to five smaller guests (water molecules)[26] causes the cavity radii to have a much narrower distribution which appears to converge more rapidly. This makes sense from purely steric considerations. The presence of one large guest would restrict fluctuations of the cryptophane host more so than a variable number of small guests which can move independently of one another and even exchange with the surrounding solvent. As a result, the cryptophane cavity is unable to collapse as much when TMA^+ or NEO is bound and the host is forced to sample a narrower range of cavity radii around a larger average radius. The cavity radius distributions for the simulations containing TMA^+ or NEO would also tend to converge more rapidly since a narrower range of cavity radii is available.

Comparison of the cavity radius histograms for the two host−guest simulations is not necessarily as one would expect given the results of the conformational analysis. The distributions of the cavity radii for the two host−guest simulations are very similar despite the fact that the cage region of the host−NEO simulation shows approximately 9 times more conformations than the host−TMA^+ simulation. This is somewhat unexpected since atoms comprising the cage region as defined in the conformational analysis are the primary atoms which form the cryptophane cavity and therefore have the greatest influence on the cavity radii.

One explanation for the apparent inconsistency is that because the guests are nearly isomorphic, the cage region of the host shows very similar conformations in the two simulations, resulting in very similar distributions of cavity radii. However, because of the overall charge of +1 on the TMA^+ guests, transitions between cage conformations are energetically less favorable and occur less often than when NEO is bound. As a result, the cage region in the host−TMA^+ simulation shows fewer conformations in a given amount of simulation time than in the host−NEO simulation.

It is also interesting to notice from the cavity results that the minimum cavity radius of the host−NEO simulation is 0.05 Å larger than the minimum cavity radius of the host−TMA^+ simulation. Likewise, the maximum cavity radius is 0.05 Å larger in the host−NEO simulation. The effective radii of NEO and TMA^+ are 2.15 and 2.10 Å, respectively, as defined by the equilibrium separation between one of the hydrogen atoms and the central atom of each guest. The effective radius for TMA^+ was used in the earlier studies to predict the effect of pore fluctuations on binding kinetics.[26,27] These results further

Table 2. Results of Conformational Analysis

	no. sampled (25 ns)					
	host		host−TMA^+		host−NEO	
region	unique	unique − equivalent	unique	unique − equivalent	unique	unique − equivalent
whole	1 121 331	1 120 863	230 585	223 940	995 480	993 205
exo	197 568		146 245		222 494	
cage	58 050	33 855	1 615	1 130	18 648	9 965

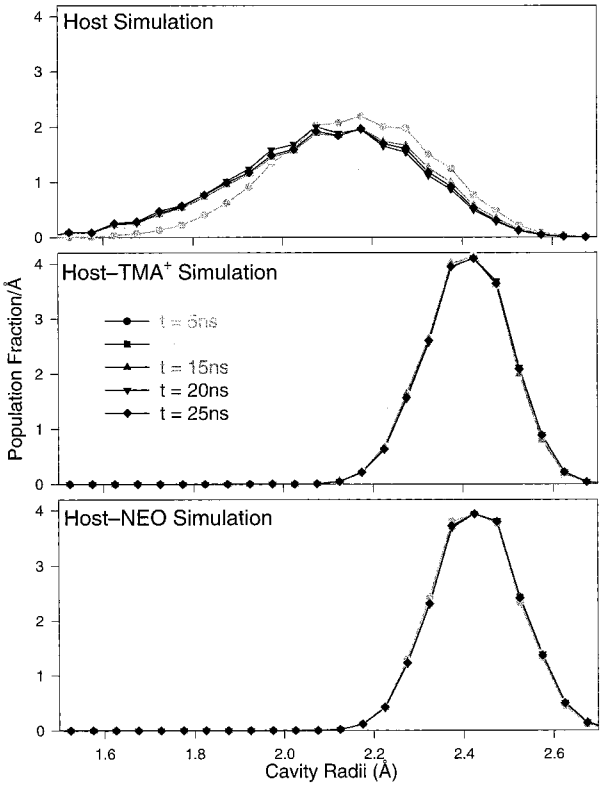


Figure 3. Distribution and convergence of the cavity radii. Histograms of cavity radii are displayed at cumulative times through 5, 10, 15, 20, and 25 ns for each of the three simulations.

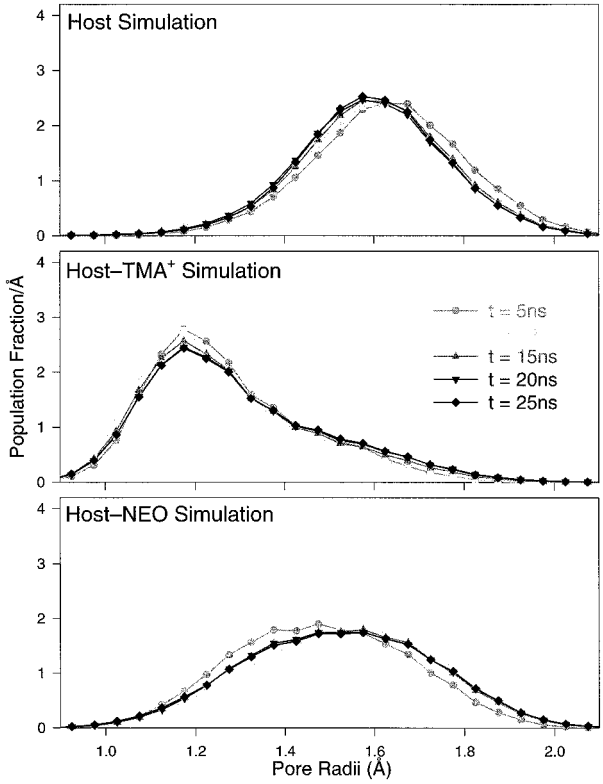


Figure 4. Distribution and convergence of the pore radii. Histograms of pore radii are displayed at cumulative times through 5, 10, 15, 20, and 25 ns for each of the three simulations.

validate the current definition of the effective radius for comparison with the cavity and pore radii as determined in these studies.

Histograms of the pore radii, shown in Figure 4, and pore radii as a function of simulation time (Figure 5) are very interesting and somewhat unexpected. The host simulation had an average pore radius of 1.57 Å. The average pore radius for the host−TMA^+ simulation has been shifted to a significantly smaller value of 1.27 Å. The average pore radius for the host−NEO simulation lies intermediate to the other two at 1.49 Å. An interesting aspect of these histograms is the more uniform distribution of pore radii for the host−NEO simulation as compared to the other two simulations.

The results of the pore calculations for the three simulations can be explained in the following way. The cryptophane pores are formed by the propyl linkers and the carboxyl groups. In the absence of a single large bound guest as in the host simulation, the propyl linkers and carboxyl groups have a greater range of motion than when a large guest is bound. For this reason, in the host simulation, the cryptophane shows a larger range of pore (and cavity) radii than in the other simulations. As seen in Figure 4, there is a more uniform distribution for the pore radii in the host−NEO simulation; however, a larger range of pore radii was sampled in the host simulation as listed in Figure 5.

When a single large guest is bound, steric factors will tend to decrease the range of motion of the propyl linkers since they must all remain fairly extended in a similar configuration for the cryptophane to accommodate the guest. As a result, the host−guest simulations show a smaller range of pore and cavity radii. The decrease in average pore radius for the host−guest simulations is probably due in part to the steric influence of the guest on the cryptophane linkers. However, the largest contribution to the decrease in average pore radius in the host−TMA^+ simulation is probably due to electrostatic interactions between the cryptophane and TMA^+. This will be discussed further in the next section of this paper.

The more uniform distribution of pore radii for the host−NEO simulation as compared to the other two simulations may be related to the observation that more conformations were sampled for the exo region in the host−NEO simulation than the other two simulations. What appears to be happening is that cryptophane atoms comprising the exo region interact more strongly with TMA^+ and to a lesser extent bound water molecules than with the bound NEO molecule. Because of the weaker interactions, the carboxyl groups have greater mobility in the host−NEO simulation than in the other two simulations. As a result, the exo region shows more conformations and the pore radii have a more uniform distribution in the host−NEO simulation.

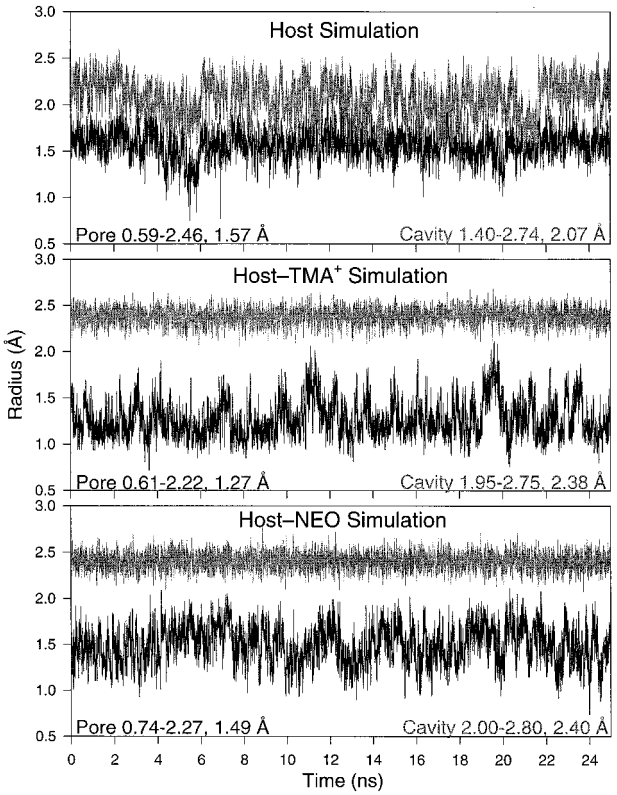


Figure 5. Cavity and pore radii as a function of simulation time for each of the three simulations. Data are plotted at 5 ps intervals. The range and average radii for all structures analyzed are displayed on the corresponding graphs.

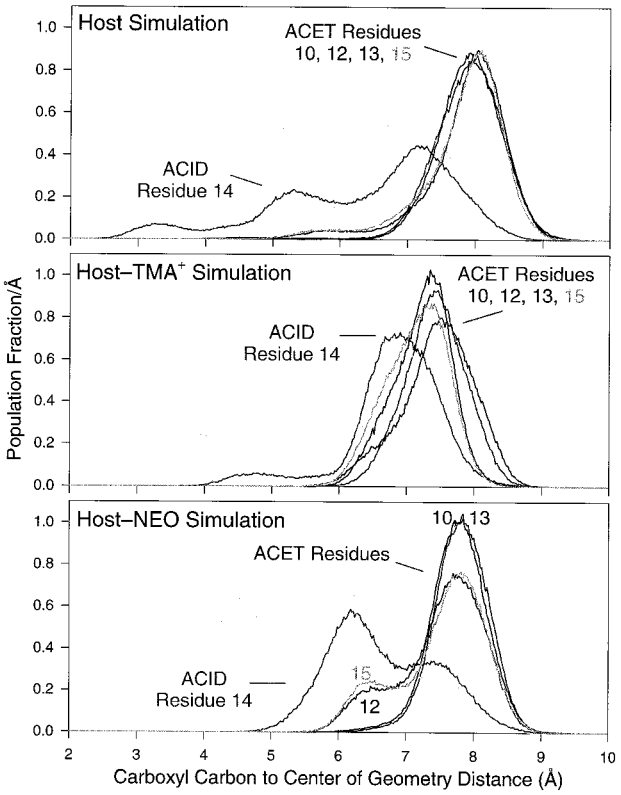


Figure 6. Histograms of the carboxyl carbon to the center of geometry distances for the three simulations. The plots are labeled for comparison to Figure 1.

Distance Histograms. To gain insight into how interactions between the cryptophane and the guests occur, a number of distance histograms were calculated. For each structure analyzed, the center of geometry based on the cap and linker regions of the host (atoms making up the ARYL and PRPL fragments) was determined. Histograms were generated for the distance between the center of geometry of a structure and each of the six carboxyl carbons. These histograms are displayed in Figure 6 for the three simulations. The plots are labeled for comparison to Figure 1.

There are a number of interesting features in these plots. One notable feature is the effect of TMA^+ and NEO on the range of motion of the ACET groups. In both host−guest simulations, the presence of the large guest causes a slight reduction in the average ACET distance from the center of the cryptophane. For the host−TMA^+ complex, the change in the average ACET distance is not as significant as one might expect from a purely Coulombic anion−cation interaction. These results suggest that the stabilization of this cryptophane−guest complex is not due simply to cation−anion (TMA^+−ACET) type interactions. A rough estimate of the Coulombic energy change due to the inward shift of the carboxylate groups by 0.5 Å can be obtained by assuming that the cationic charge is concentrated at the center of the cryptophane and that the effective dielectric constant is 80; the energy change is then about 0.2 kcal/mol.

As noted before in a comparison of the host and host−TMA^+ systems, there is a greater range of motion of the neutral ACID groups as compared to the charged ACET for both the host and host−guest systems. The ACID groups can also come closer to the center of the cryptophane. (The presence of TMA^+ and NEO does limit how closely the ACID groups approach the cryptophane center. This is likely due simply to the presence of the relatively large single guest as opposed to several smaller water molecules which are capable of moving independently of one another.) This makes sense in terms of solvation effects. The ACET groups, being charged, interact more strongly with the surrounding water than do the neutral ACID groups. Interaction with the surrounding water molecules tends to suppress the fluctuations of the ACET groups while keeping them extended away from the cryptophane cavity.

Another interesting feature of these plots is the significant difference in ACET histograms between residues 12 and 15 as compared to residues 10 and 13 in the host−NEO simulation. This difference can also be seen to a much lesser extent in the host simulation, in the shoulder for residues 12 and 15 around 6 Å. Looking at Figure 1, we can easily explain this difference. The ACET residues 12 and 15 form the one ACET/ACET pore. The other two pores are ACID/ACET pores formed by the 11/13 and 14/10 pairs. The fact that this difference can be observed suggests that the placement of the two acid protons may have a more significant effect on the dynamics of the cryptophane than one might expect.

To further investigate what is occurring in these simulations, histograms for the distances of the ether oxygens to the cryptophane center of geometry were generated for each of the simulations. Plots of these histograms are displayed in Figure 7. There are twelve ether oxygens in the cryptophane with six connecting the carboxyl groups to the aromatic rings and six connecting the propyl linkers to the aromatic rings. Plots for the ether oxygens connecting the propyl linkers are shown with heavy lines in Figure 7.

The histograms for the carboxyl ether oxygens show basically the same trends for the three simulations as did the histograms for the carboxyl carbons. The interesting portion of these graphs involves the ether oxygen histograms for the propyl linkers. The average ether oxygen distances for the propyl linkers in the host−TMA^+ simulation are smaller than in the host and host−

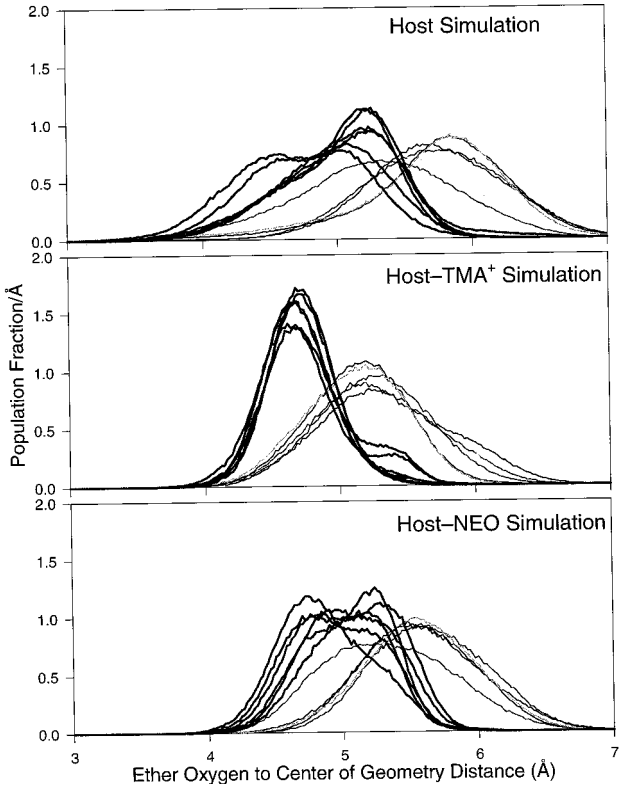


Figure 7. Histograms of the distances of the ether oxygens to the cryptophane center of geometry for each of the simulations. Plots for the ether oxygens connecting the propyl linkers are shown with heavy lines.

NEO simulations. Also, in the host and host−NEO simulations, the ether oxygen histograms for the propyl linkers are broader and more disordered than that of the host−TMA^+ simulation.

The results for the conformational analysis and for the histograms of Figures 6 and 7 can be tied together and explained in the following way. Without a single large guest bound, the cryptophane host has a greater range of internal motions. The greater range of motions appears clearly in the histograms especially for the ACID groups. When a single large guest is bound, the range of internal motions of the cryptophane may be significantly reduced as appears to be the case from the histograms. The number of unique conformations available to the cryptophane on an energetic basis would also likely be significantly reduced. How rapidly the remaining unique conformations would be sampled would depend on the interaction between the host and the particular guest.

Given that TMA^+ and NEO are nearly isomorphic, it is reasonable that they will have the same type of steric interaction with the cryptophane. Excluding electrostatic interactions, the cryptophane of the TMA^+ and NEO complexes could be expected to have approximately the same number of conformations available on an energetic basis. However, TMA^+ does differ from NEO, having an overall charge of +1. It is reasonable to expect the positively charged TMA^+ guest to have more favorable electrostatic interactions with the negatively charged cryptophane than the neutral NEO. Electrostatic interactions between TMA^+ and cryptophane would likely dampen the motions of the nearby, negatively charged ether oxygens. The result would be consistent with the histograms of Figure 7.

The presence of the +1 charge on TMA^+ may or may not significantly reduce the number of conformations available to the host on an energetic basis as compared to the host in the cryptophane−NEO complex. However, the charge would be expected to reduce the rate at which unique conformations were sampled with simulation time. Histograms involving large motions of the cryptophane would tend to converge more slowly for the cryptophane−TMA^+ simulation than for the cryptophane−NEO simulation. This appears to be what is occurring in the carboxyl carbon histograms of Figure 6. Without a single large guest bound, many conformations are available and even after 25 ns of simulation the histograms for the host simulation have not fully converged. Differences between residues 12 and 15 as compared to residues 10 and 13 are only slightly visible. On the binding of a single large guest, the number of conformations available to the host on an energetic basis is significantly reduced. However, with the charge on TMA^+, the rate at which conformations are sampled is also greatly reduced and convergence of the histograms for the host−TMA^+ simulation is incomplete. In the host−NEO simulation, the number of conformations for the cryptophane is significantly reduced but the rate of sampling is not. As a result, the host−NEO histograms are the best converged of the three simulations.

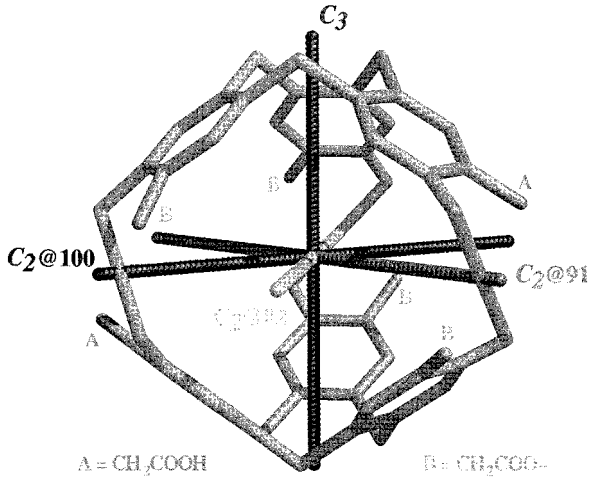


Figure 8. Illustration of the symmetry axes for the cryptophane.

Defining Unit Vectors for the Rotational Analysis. Rotational analysis was conducted by following rotation along the symmetry axes for the molecules studied. For TMA^+ and NEO shown in Table 1, each molecule has four C_3 axes of rotation running along the carbon−nitrogen and carbon−carbon bonds of the molecules, respectively. To conduct the analysis, unit vectors were calculated for each of the axes from the corresponding atomic coordinates.

Ignoring the protonation state, the cryptophane also has four unique axes of rotation. They are a C_3 axis, which runs through the center of the cryptophane caps, and three C_2 axes, which run perpendicular to the C_3 axis. Each C_2 axis passes through the central carbon of one of the propyl linkers. Figure 8 illustrates where these axes lie in relation to one another and the cryptophane.

The C_3 axis for the cryptophane was defined by finding the center of geometry of each cap based on atoms comprising the ARYL fragments of that cap. The C_3 axis was then defined by the unit vector running between these two points. The C_2 axes for the cryptophane required calculating the center of geometry for the molecule based on atoms comprising the ARYL and PRPL fragments. The C_2 axes were then defined by the unit vector running between the center of geometry of the cryptophane and the central carbon of the corresponding propyl linker. These axes are identified in this study by the system

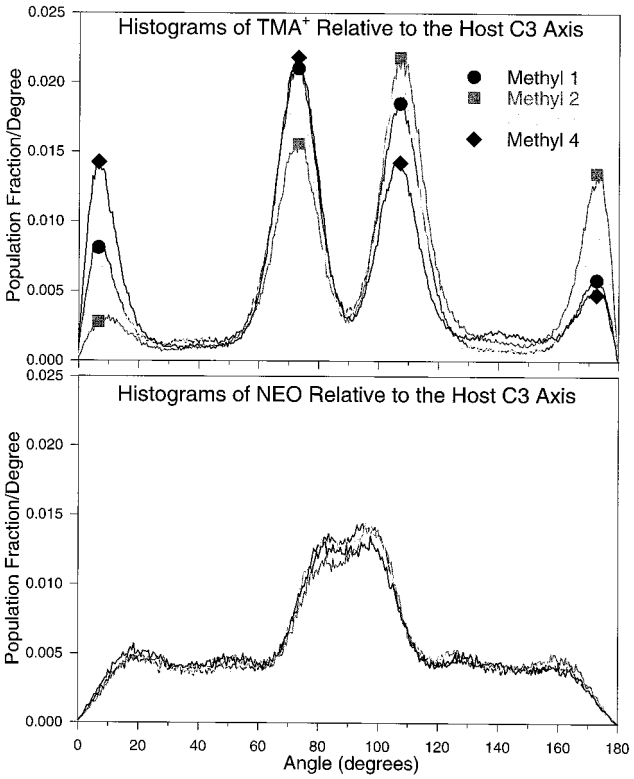


Figure 9. Histograms for each of the guest C_3 axes relative to the host C_3 axis.

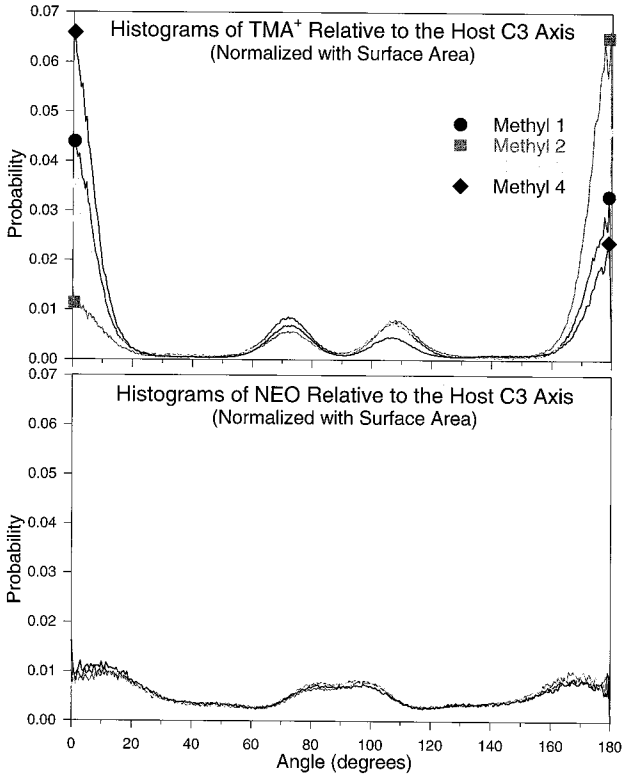


Figure 10. Histograms of Figure 9 expressed in terms of probability.

atom numbers of the central carbon on the corresponding propyl linkers (82, 91, and 100).

It should be noticed from Figure 8 that the three C_2 axes are not equivalent due to the protonation state of the cryptophane. Two of the C_2 axes (91 and 100) pass through pores containing an ACID/ACET pair. The other C_2 axis (82) passes through the one pore containing an ACET/ACET pair. This difference in the C_2 axes appears in the rotational analysis.

Rotational analysis was conducted on the cryptophane−TMA^+ and cryptophane−NEO simulations using structures saved at 0.1 ps intervals throughout the 25 ns simulations. Unit vectors were determined for each of the host and guest symmetry axes and stored as a function of simulation time. The stored unit vectors were then used in a number of rotational analyses.

Preferred Orientation of the Guest Relative to the Host C_3 Axis. It is of interest to determine what the preferred orientation of the guest is relative to the cryptophane in the host−guest complex and whether the preferred orientation differs between the two guests studied. To make this determination, angles between each of a guest's C_3 axes and the cryptophane's C_3 axis were calculated for a given structure of the host−guest complex from the saved unit vectors, using the dot product relation. This was done for all of the saved structures of both host−guest simulations. The angles obtained were sorted into bins of equal width. Once all the angles were sorted, each bin was divided by the total number of structures analyzed and by the bin width in degrees. The result was a histogram for each of the guest C_3 axes relative to the host C_3 axis. These histograms which are in units of population fraction per degree versus the angle in degrees are displayed for TMA^+ and NEO in Figure 9.

The histograms for TMA^+ are quite different from those of NEO. The TMA^+ histograms have a series of four fairly well defined peaks at approximately 10, 70, 110, and 170°. This is basically what is expected if the guest prefers to have one of its axes aligned with the host C_3 axis. If the guest always had one of its axes perfectly aligned with the host C_3 axis, there would be peaks at 0, 70.5, 109.5, and 180° due to the tetrahedral structure of the guests studied. Also, because of the tetrahedral structure of the guest, there would be a 1:3:3:1 ratio of the areas under the peaks. Integrating each of the TMA^+ histograms from 0 to 45, 45 to 90, 90 to 135, and 135 to 180° and then averaging result in an overall ratio of 1.0:2.8:2.7:1.0 for the areas under the peaks. This suggests that the preferred orientation of TMA^+ within the cryptophane is such that one of the guest axes is usually aligned or nearly aligned with the host C_3 axis.

In contrast, the histograms for NEO have a more uniform distribution than those for TMA^+. Integrating the histograms for NEO and averaging as was done on the TMA^+ histograms result in an overall ratio of 1.1:2.3:2.4:1.0 for the areas under the peaks. This suggests that the preferred orientation of NEO within the cryptophane is such that the alignment of its axes with the host C_3 axis is less rigidly defined than it is for TMA^+.

For generation of the histograms of Figure 9, bins of equal width were used throughout the entire range of angles. This results in the histograms being in terms of population fraction per degree. These histograms can be expressed in terms of probability by taking into account that the same bin width corresponds to a larger surface area near the equator than at the poles. The histograms of Figure 9 expressed in terms of probability are displayed in Figure 10. The probability distribution for NEO is nearly uniform compared to the distribution for TMA^+.

The histograms of Figures 9 and 10 reveal two important results. One is that TMA^+ has a more well defined, preferred orientation within the cryptophane than does NEO, even though TMA^+ and NEO are nearly isomorphic. The second important result from these histograms is that, even after 25 ns of simulation, the four methyl groups of TMA^+, which are equivalent, have not had equivalent sampling within the cryptophane. This can be seen as the difference in peak height

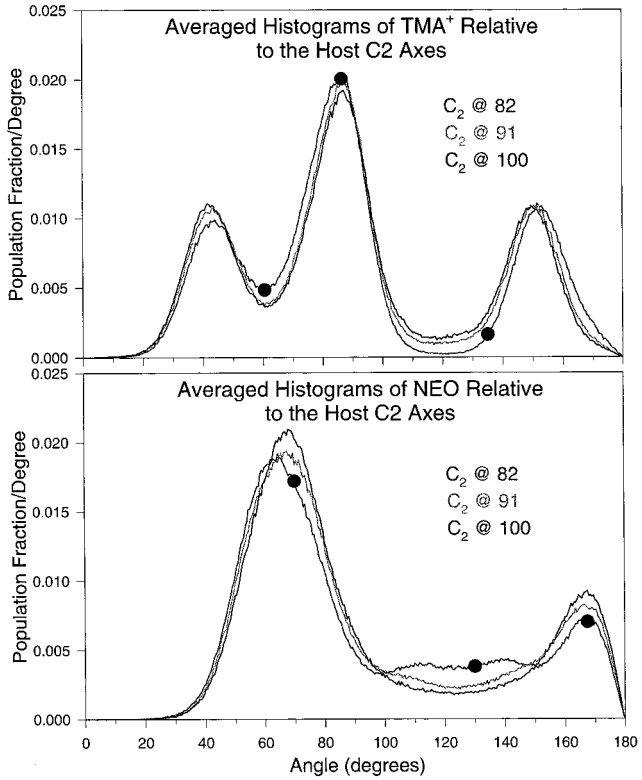


Figure 11. Averaged histograms for each of the guest C_3 axes relative to the host C_2 axes.

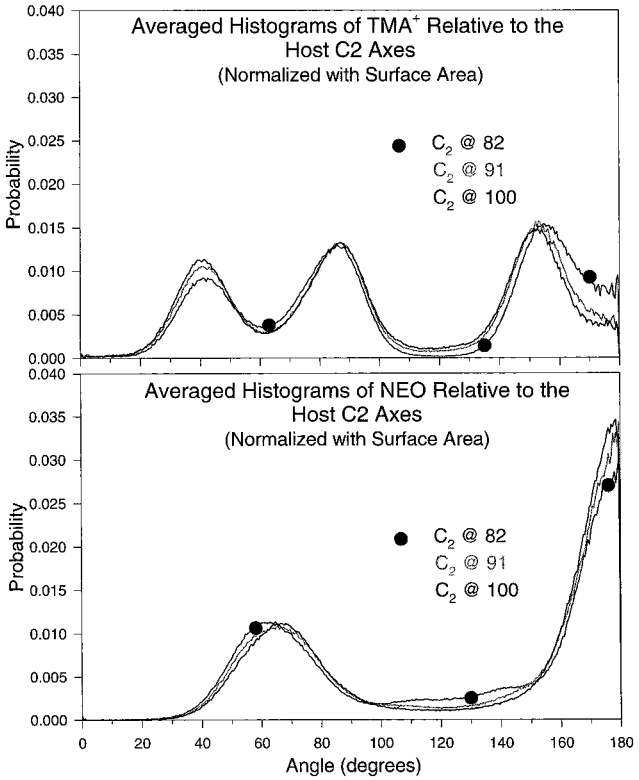


Figure 12. Histograms of Figure 11 expressed in terms of probability.

for the four TMA^+ C_3 axes. Together, these results suggest that NEO tumbles relative to the host C_3 axis at a faster rate than TMA^+.

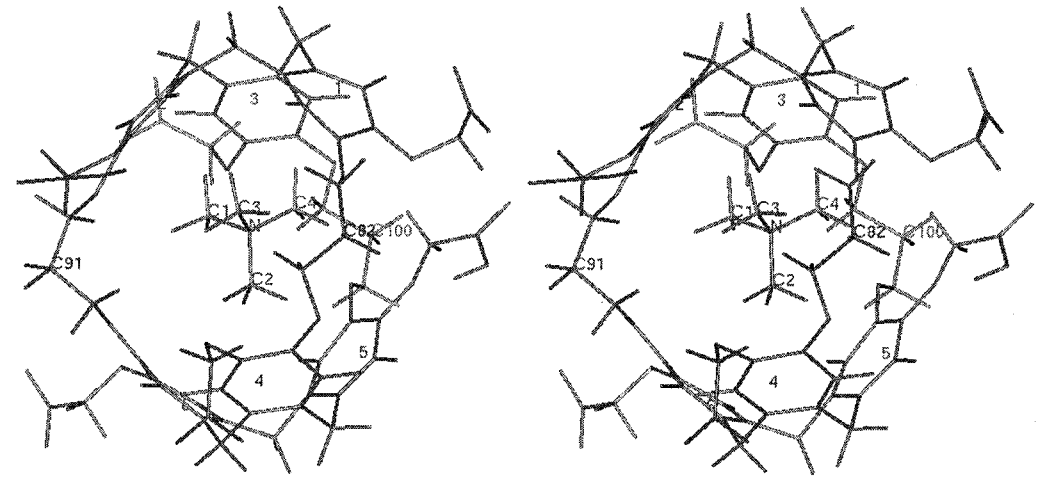
Preferred Orientation of the Guest Relative to the Host C_2 Axes. As in the comparisons to the host C_3 axis, the preferred orientation of the guest relative to the cryptophane C_2 axes was determined. Histograms were calculated for each of the guest's four C_3 axes relative to each of the cryptophane's three C_2 axes, resulting in a total of twelve histograms for each guest. The four histograms for the guest C_3 axes relative to each host C_2 axis were averaged. The averaged histograms are displayed in Figure 11 as population fraction per degree versus the angle in degrees. These histograms expressed in terms of probability are displayed in Figure 12. As can be seen in these histograms, the unique C_2 axis which passes through a pore containing an ACET/ACET pair differs from the other two C_2 axes which pass through pores containing ACID/ACET pairs.

The histograms for TMA^+ differ significantly from those of NEO. The TMA^+ histograms have a series of three peaks at approximately 40, 90, and 150°. Integrating each of the TMA^+ histograms from 0 to 60, 60 to 120, and 120 to 180° and then averaging result in an overall ratio of 1.0:2.1:1.0 for the areas under the peaks. In contrast, the histograms for NEO have only two peaks at approximately 70 and 170°. Integrating each of the NEO histograms from 0 to 120 and 120 to 180° and then averaging result in an overall ratio of 2.8:1.0 for the areas under the peaks. These histograms indicate a significant difference in the preferred orientation of the guests relative to the host C_2 axes.

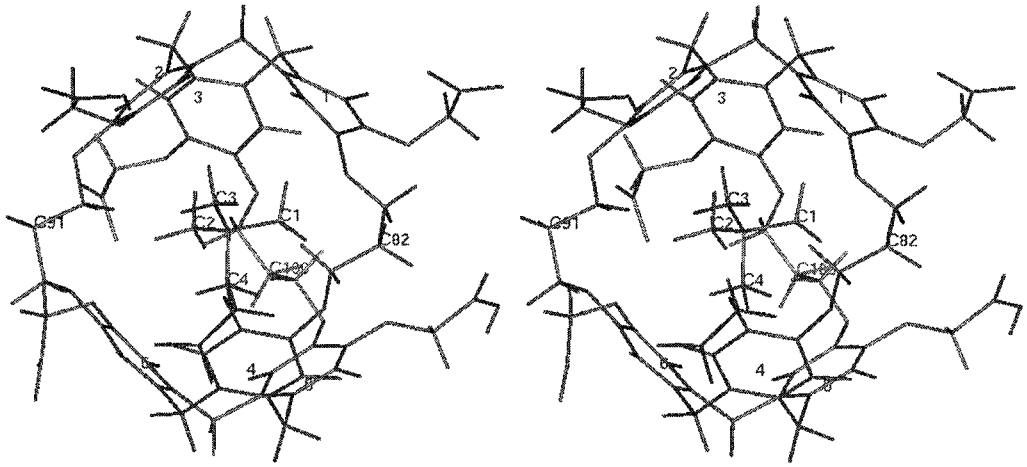
Overall Preferred Orientation of the Guest Relative to the Host. To illustrate the difference in preferred orientation of the guests relative to the host, representative structures for the two complexes are shown in Figure 13 as stereoviews. Angles between the various axes in the two structures in Figure 13 are given in Table 3. These structures were chosen to emphasize the difference in the preferred guest orientations relative to the host C_2 axes. As can be seen in Figure 13, NEO is oriented such that one of its methyl groups is almost directly centered in each of the pores. As a result, each host C_2 axis has one NEO methyl group at approximately 165° with the remaining three methyl groups at angles ranging from 53 to 89°. Having a methyl group centered in each of the cryptophane pores presumably minimizes steric conflicts between the host and the guest. This observation is consistent with what has been seen experimentally using X-ray diffraction on crystals of a cryptophane−chloroform complex.[15] In the cryptophane−chloroform complex, each of the chlorine atoms of the chloroform is approximately centered in one of the host pores.

The preferred orientation of TMA^+ relative to cryptophane also directs the methyl groups toward the host pores. However, for TMA^+ the methyl groups tend not to be directly centered within the host pore. At least in this model, the methyl groups appear to be shifted due to electrostatic interactions between the hydrogen atoms of TMA^+ and the linker ether oxygens of the cryptophane. As a result, each host C_2 axis has one TMA^+ methyl group at approximately 150°, one at approximately 40°, and the remaining two methyl groups at angles ranging from 85 to 94°.

Tumbling Rate of the Guest Relative to the Host. Because TMA^+ and NEO are nearly isomorphic, the histograms of the guests relative to the host suggest that NEO is tumbling within the cryptophane at a faster rate than TMA^+. This is especially apparent in the histograms of the guests relative to the host C_3 axis. (If TMA^+ and NEO were very different from one another, it would be difficult to determine whether the differences were due to tumbling rates of the guests or their preferred orientation within the cryptophane.) To obtain a more quantitative value for the rates at which the guests are tumbling relative to the host, the average duration of time a guest's C_3 axis spent aligned with a host axis was determined. The values calculated here are very sensitive to how alignment of a guest axis with a host



Cryptophane–TMA^+



Cryptophane–NEO

Figure 13. Stereoplots of representative structures for the two host−guest complexes. Angles between the various axes in these two structures are given in Table 3. These structures were chosen to emphasize the difference in the preferred guest orientations relative to the host C_2 axes.

Table 3. Angles Between the Various Axes in the Two Representative Structures Displayed in Figure 13

	methyl			
	1	2	3	4
TMA^+ Angles Relative to the Host				
C3	111	1.94	108	112
C2 @ 82	148	88.5	41.6	87.2
C2 @ 91	44.3	84.9	79.7	152
C2 @ 100	84.8	93.5	150	39.5
NEO Angles Relative to the Host				
C3	102	96.6	114	3.05
C2 @ 82	59.7	170	70.2	82.5
C2 @ 91	167	66.4	61.2	89.4
C2 @ 100	63.6	52.6	161	87.5

axis is defined. These results are only semiquantitative and are only meant for comparison between the two guests in this study.

The rate at which the guests tumble relative to the host C_3 axis was determined in the following way. A particular C_3 axis of the guest was considered to be aligned with the host C_3 axis if it was within 50° of the axis. The total amount of time a particular methyl group (C_3 axis) of the guest spent within a range of 0−50° from the host C_3 axis was determined for the 25 ns of simulation. Also for the same methyl group, the number of times it left and then returned to this range of 0−50° was determined. From these two numbers, the average duration the methyl group spent within that 50° range was calculated. This was done for each of the guest's four methyl groups (C_3 axes). The average durations were also determined for the methyl groups within the 50° range from 130 to 180° from the host C_3 axis. The overall average durations for the alignment of a guest C_3 axis with the host C_3 axis were 3.2 ps for TMA^+ and 1.2 ps for NEO. Clearly, NEO is tumbling relative to the cryptophane C_3 axis at a faster rate than TMA^+.

The rate at which the guests tumble relative to the host C_2 axes was determined in a similar fashion. A particular guest C_3 axis was considered to be aligned with a particular host C_2 axis if it was within a 50° range of 130−180° of that axis. The average duration for the alignment of a guest C_3 axis with a host C_2 axis was determined for each of the guest's four methyl groups for each of the host's three C_2 axes. The overall average durations for the alignment of a guest relative to the host C_2 axes were 8.3 ps for TMA^+ and 2.9 ps for NEO. As in the comparison to the host C_3 axis, NEO is tumbling relative to the cryptophane C_2 axes at a faster rate than TMA^+.

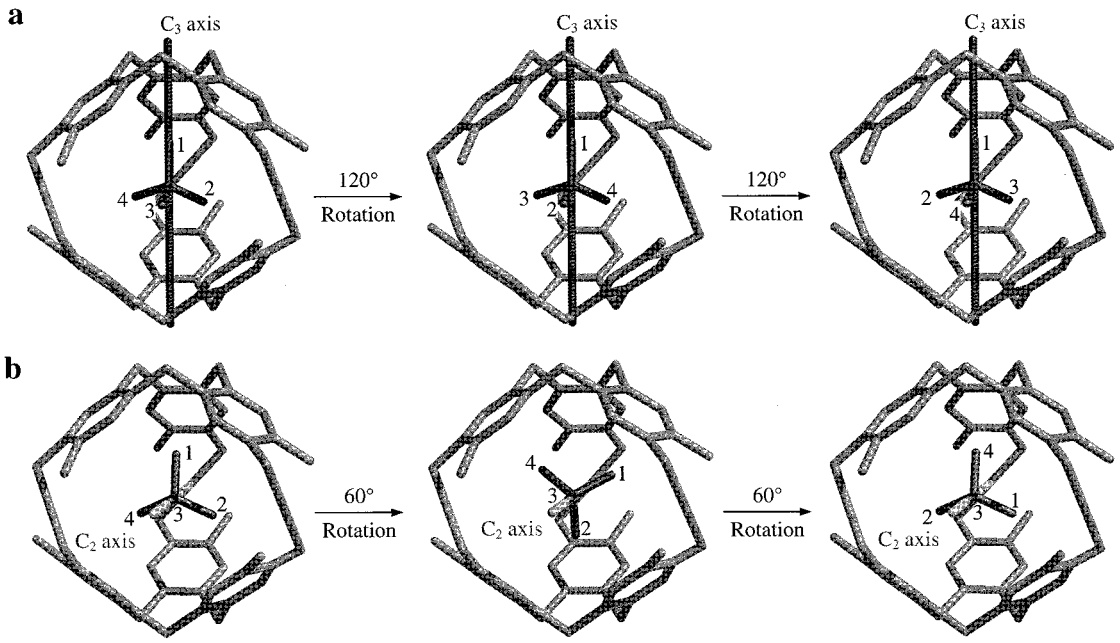


Figure 14. Illustration of guest rotation about (a) the host C_3 axis and (b) one of the host C_2 axes while maintaining the preferred orientation of the guest. To simplify the illustration, a guest axis is shown as being aligned with the host C_3 axis and one of the host C_2 axes in parts a and b, respectively.

These results indicate that NEO has less hindered rotation within the cryptophane cavity than TMA^+. This is consistent with the interpretation of the distance histograms presented in this study. The interpretation is that TMA^+ has more favorable electrostatic interactions with the host than NEO, causing a greater dampening of the motions of the host in the cryptophane−TMA^+ complex than in the cryptophane−NEO complex.

The tumbling rates also indicate that both guests rotate faster about the host C_2 axes than about the host C_3 axis. This is what one might expect. Rotation about the host C_3 axis would require rotation in 120° increments to maintain the preferred orientation of the guest. In a 120° rotation about the host C_3 axis, a guest methyl group would have to be dislodged from each of the three host pores. Rotation about a host C_2 axis, however, could be done in 60° increments while still maintaining the preferred orientation of the guest. In a 60° rotation about a host C_2 axis, a methyl group aligned with the host C_3 axis would displace one methyl group from a pore, moving that second methyl group to alignment with the host C_3 axis in the opposite direction. These motions are illustrated in Figure 14.

Concluding Remarks

Three simulations were conducted utilizing a small, fairly rigid cryptophane host. Simulations of the uncomplexed cryptophane (water molecules loosely bound) and cryptophane−TMA^+ and cryptophane−NEO host−guest complexes were examined. The TMA^+ and NEO guests are both tetrahedral species and are nearly isomorphic. The main difference between the two guests is the overall charge of +1 on TMA^+ versus the overall neutral charge of NEO. However, no overall dipole exists for either guest owing to their tetrahedral symmetry.

Despite the similarity between TMA^+ and NEO, the complexes they form with the cryptophane are very different. TMA^+ greatly reduces the conformational sampling of the host, but NEO may actually enhance sampling in some regions of the host as compared to the uncomplexed cryptophane. TMA^+ and NEO have similar preferred orientations relative to the host C_3 axis; however, TMA^+ has a more well defined orientation. The guests have significantly different orientations relative to the host C_2 axes. The rates at which the guests tumble relative to the cryptophane are also significantly different, with NEO tumbling much faster than TMA^+.

This study demonstrates some of the many challenges in the understanding of molecular recognition. The cryptophane can be viewed as a small, fairly rigid receptor site, and TMA^+ and NEO, as similar ligands. As seen in this study, small differences in ligands can result in significant differences in sampling of the receptor and the motions of the bound ligand. Significant differences in the sampling and motions of the system from subtle changes in ligands present significant challenges to the prediction of binding orientation and binding free energies, both of which are key aspects of molecular recognition.

Acknowledgment. The authors thank Dr. T. P. Straatsma for the use of ARGOS and for his advice, Dr. Michael Gilson for providing the surfacing code adapted for use in the pore and cavity determinations, and Dr. Michael Bass for providing guest atomic charges. P.D.K. thanks Heather Carlson, Adrian Elcock, Volkhard Helms, Philippe Hunenberger, Michael Potter, and other members of the McCammon group for their suggestions, comments, and support of this study. This work was supported in part by grants from the NSF and the NSF Supercomputer Centers MetaCenter Program.

JA981526A