Structural Fluctuations of a Cryptophane Host: A Molecular Dynamics Simulation

P. D. Kirchhoff,[†] M. B. Bass,[‡] B. A. Hanks,[§] J. M. Briggs,[†] Andre Collet,[⊥] and J. A. McCammon*,[†]

Contribution from the Department of Chemistry and Biochemistry and Department of Pharmacology, University of California at San Diego, La Jolla, California 92093-0365, Amgen, Inc., 1840 DeHavilland Drive, Thousand Oaks, California 91320-1789, Department of Chemistry, Southwestern University, Georgetown, Texas 78626, and Stereochimie et Interactions moleculaires, Ecole Normale Superieure de Lyon, 46, allee d'Italie, F-69364 Lyon Cedex 07, France

Received November 9, 1995[⊗]

Abstract: Cryptophanes are aromatic hosts which bind a variety of guests. Here, we describe a 20 ns molecular dynamics simulation of a particular cryptophane in water. This cryptophane features three pores which open onto a cavity where the guests bind. The molecular dynamics simulation in combination with a surfacing algorithm provides information on the frequency with which these pores open wide enough to admit guest molecules of any given size. We discuss these fluctuations and their possible consequences for binding kinetics.

Introduction

Host−guest chemistry has attracted a great deal of attention for many years. It has practical importance in such areas as environmental remediation[1,2] and fundamental importance in such areas as clarifying the principles of molecular recognition.[3−7]

Cryptophanes are a group of interesting cagelike host molecules that have been well-characterized in terms of the binding of guests of various sizes and net charges, in organic and aqueous solvents.[8−10] The cryptophane complexes are reversibly formed, and their stabilities (i.e., the magnitudes of their binding constants) depend on variety of factors. Factors which do not directly involve the cryptophane, such as the solvent−guest interactions, may convey a significant contribution in the formation of cryptophane host−guest complexes. Even so, experimental data support the idea that host−guest and host−solvent interactions are dominant. For instance, it is likely that in aqueous solvents the host−solvent interactions will play a major role. The cryptophane cavity, despite its supposed hydrophobic character, is large enough to accommodate several water molecules which would require displacement prior to the binding of another guest.

Moreover, experiments indicate that the complexation and release of guest molecules is comparatively slow (barriers in the range of 10−17 kcal/mol), and static structures suggest that, for some guests, passage through a cryptophane pore is sterically hindered. This implies that fluctuations of the pores may be important to the binding kinetics of guest molecules. In this study, molecular dynamics simulations have been conducted to obtain structural and conformational sampling information on a particular cryptophane in water. The fluctuations of the cryptophane structure (and particularly of the pores), as well as the properties of water molecules inside, are characterized, and the implications for the binding kinetics are discussed.

A representation of the cryptophane molecule used in this study is displayed in Figure 1. This molecule is a hexaacid derivative of cryptophane-E, and its binding properties for ammonium guests have recently been reported.[11] It contains two cyclotriveratrylene groups, each comprising three aromatic rings connected by methylene bridges. Each aromatic ring is bonded to an acetic acid group through an ether linkage. The two cyclotriveratrylene groups are joined together 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.

Methods

Atomic Parameters. An all-atom representation was employed for this system. Internal bonded parameters were obtained from the AMBER parameterization.[12] Table 1 lists bond and angle terms necessary for the system but not present in the AMBER parameterization, along with the bond length, angle, and force constant values used for these terms. The additional values were obtained by comparison of bond and angle terms for similar atom types already represented in the AMBER parameterization.

Nonbonded parameters were obtained from two sources. The Lennard−Jones parameters were obtained from the OPLS parameter-

[†] University of California at San Diego.
[‡] Amgen, Inc.
[§] Southwestern University.
[⊥] Ecole Normale Superieure de Lyon.
[⊗] Abstract published in *Advance ACS Abstracts,* March 15, 1996.

(1) Dang, L. X. *J. Am. Chem. Soc.* **1995**, *117*, 6954−6960.
(2) Schulz, W. W.; Bray, L. A. *Sep. Sci. Technol.* **1987**, *22*, 191−214.
(3) Marrone, T. J.; Merz, J. K. M. *J. Am. Chem. Soc.* **1992**, *114*, 7542−7549.
(4) Marrone, T. J.; Merz, J. K. M. *J. Am. Chem. Soc.* **1995**, *117*, 779−791.
(5) Axelsen, P. H. *Isr. J. Chem.* **1994**, *34*, 159−163.
(6) Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1993**, *115*, 9907−9919.
(7) Nakamura, K.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 1853−1854.
(8) Canceill, J.; Cesario, M.; Collet, A.; Guilhem, J.; Lacombe, L.; Lozach, B.; Pascard, C. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 1246−1248.
(9) Collet, A.; Dutasta, J.-P.; Lozach, B.; Canceill, J. *Top. Curr. Chem.* **1993**, *165*, 103−129.
(10) Collet, A.; Dutasta, J.-P.; Lozach, B. *Adv. Supramol. Chem.* **1993**, *3*, 1−35.
(11) Garel, L.; Lozach, B.; Dutasta, J.-P.; Collet, A. *J. Am. Chem. Soc.* **1993**, *115*, 11652−11653.
(12) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.

0002-7863/96/1518-3237$12.00/0

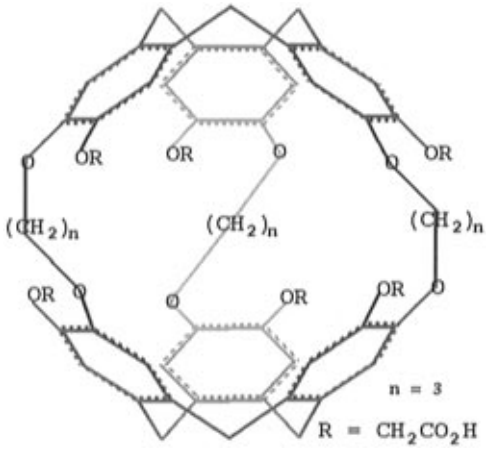

Figure 1. The structure of the cryptophane host used in this study.

Table 1. Additional Bond and Angle Terms Created for the Cryptophane System[a]

bond	bond length (Å)	force constant [(kcal/mol)/Å^2]
CA−OS	1.364	450

angle	measurement, deg	force constant [(kcal/mol)/rad^2]
CA−OS−CT	113	35
CA−CT−CA	114	80
CA−CA−OS	120	70
OS−CT−C	110	80
CT−C−OH	117	70
O−C−OH	126	80

proper angle	measurement, deg	force constant (kcal/mol)	periodicity
CT−OS−CA−CA	180	0.90	2

improper angle	measurement, deg	force constant (kcal/mol)	periodicity
CA−CA−CA−OS	180	10.5	2

[a] See Table 2 for comparison of atom types.

ization with all-atom extensions.[13,14] It has been suggested by Jorgensen's group that 1−4 interactions be scaled by 1/8.[13] In this study, all 1−4 interactions not involving polar-polar interactions, have been scaled by 1/8. The 1−4 polar−polar interactions, namely the 1−4 acid oxygen to ether oxygen interactions, have been scaled by 1/2. This was done because it has been observed that when using OPLS parameters with molecular dynamics simulations, internal hydrogen bonds often form five-membered rings which are unrealistically stable.[15]

Atomic charges were derived through ab initio calculations since OPLS charges suitable for this system were not available. The atomic charges were derived in the following way. First, due to the large size of the cryptophane, the molecule was divided into 15 fragments. Four unique fragments identified as ACET, ACID, ARYL, and PRPL resulted from these divisions. Auxiliary groups were added at the division points to form four complete molecules. Figure 2 displays the resulting molecules. Atoms making up the fragments are numbered for comparison with Table 2.

Calculations were then conducted on each of the resulting molecules individually. For each molecule, the geometry was optimized using MNDO[16] as implemented in MOPAC6.[17] Using the resulting geometries, wave functions were obtained with the 6-31G(d) basis set using

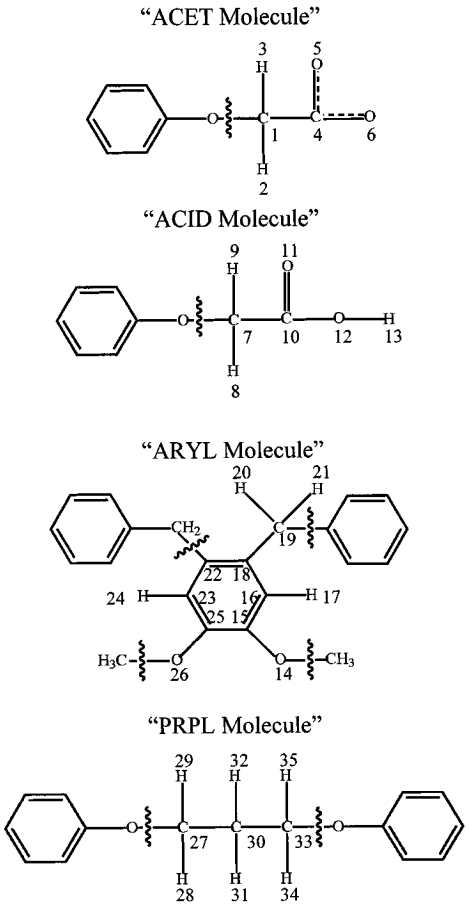

Figure 2. Molecular fragments used in the determination of atomic charges for the cryptophane. Figure labels are present for comparison with Table 2.

Table 2. Nonbonded Parameters Used for the Cryptophane System

OPLS atom type	ϵ (kcal/mol)	σ (Å)	charge	Figure 2 labels
				ACET
291 CT	0.066	3.50	0.170	1
295 HC	0.030	2.50	−0.020	2, 3
147 C	0.105	3.75	0.861	4
18 O2	0.210	2.96	−0.832	5, 6
				ACID
291 CT	0.066	3.50	0.177	7
295 HC	0.030	2.50	0.100	8, 9
147 C	0.105	3.75	0.860	10
149 O	0.210	2.96	−0.640	11
148 OH	0.170	3.00	−0.800	12
24 HO	0.000	0.00	0.530	13
				ARYL
302 OS	0.170	3.00	−0.480	14, 26
301 C	0.070	3.55	0.305	15, 25
230 CA	0.070	3.55	−0.442	16, 23
231 HC	0.030	2.42	0.250	17, 24
230 CA	0.070	3.55	0.140	18, 22
291 CT	0.066	3.50	−0.410	19
295 HC	0.030	2.50	0.145	20, 21
				PRPL
291 CT	0.066	3.50	0.055	27, 33
295 HC	0.030	2.50	0.096	28, 29, 34, 35
291 CT	0.066	3.50	−0.120	30
295 HC	0.030	2.50	0.060	31, 32

Gaussian92 (6-31G(d)//MNDO).[18] Atomic-centered charges were obtained from an electrostatic potential fit to the wave function using the CHELPG method.[19] Once the calculations were complete, the auxiliary groups and their charges were ignored. Charges on atoms at the division points were scaled so that the entire cryptophane molecule, when assembled, would have an integer charge. This adjustment

(13) Jorgensen, W. L.; Tirado-Rivers, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(14) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4768−4774.
(15) Smith, P. E.; Dang, L. X.; Pettitt, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 67−73.
(16) Thiel, W.; Dewar, M. J. S. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.
(17) Stewart, J. J. P. QCPE 455, Mopac 6.0.

resulted in a ca. 10% maximum change in charge. In addition to scaling at the division points, atomic charges on chemically equivalent positions were averaged. All-atom OPLS charges were used for the atoms of the central methylene of the PRPL fragment.[20] Table 2 lists the atom types, Lennard−Jones parameters, and atomic charges used for this study.

Molecular System. For solubility reasons, experimental studies conducted in water involve the cryptophane molecule titrated until four of the six acetic acid groups are ionized.[11] Likewise, in this study, the cryptophane molecular (Figure 1) was modeled with four of its acid protons removed, giving the system an overall charge of −4. To form the molecular system, this model was immersed in a cubic periodic box of SPC/E water[21] with dimensions of 35 × 35 × 35 Å, containing 1387 SPC/E water molecules. Molecular dynamics simulations were conducted on the system using ARGOS V6.0.[22] In all calculations, a 14.0 Å long-range nonbonded cutoff (forces evaluated once every 10 steps) and a 9.0 Å short-range nonbonded cutoff (forces evaluated every step) to the energy expression were used. The nonbonded pair list was evaluated every 10 steps. By using a 14.0 Å cutoff, all intramolecular nonbonded interactions were included. Use of an Ewald or other method to eliminate cutoffs entirely would be desirable, but would not be expected to alter the results presented here substantially because the host molecule is fairly symmetrical and a rather large value is used for the long-range cutoff. SHAKE was used to fix all bond lengths to their equilibrium values.

Equilibration of the System. Coordinates of the cryptophane molecule displayed in Figure 1 with $n = 5$ (cryptophane-O) were provided by Dr. Andr Aubry. This structure was modified by removing two methylene carbons and their hydrogens from each of the three pentane linkers, thereby forming the cryptophane structure with three propane linkers. Four of the six acid protons were also removed. The protons were removed in such a way that one of the cryptophane pores contained an ACET/ACET pair and the other two pores contained ACET/ACID pairs. The resulting structure was then partially relaxed with steepest descent energy minimization and placed in the aqueous system described above. Water molecules whose oxygens were within 2.0 Å of a heavy atom of the cryptophane were removed from this initial model.

Equilibration of the system was conducted by first relaxing the solvent energy with steepest decent while the solute was held fixed. Second, molecular dynamics was conducted on the solvent at 298 K for 20 ps with velocity reassignment every 0.2 ps while the solute was held fixed. Third, molecular dynamics was conducted on the entire system for 5 ps intervals at temperatures of 100, 200, and 298 K with velocity reassignment every 0.2 ps. Finally, molecular dynamics was conducted on the entire system at 298 K for 10 ps with velocity reassignment every 0.5 ps followed by a 20 ps simulation with no velocity reassignment. During the equilibration, the short-range cutoff for solvent−solvent interactions was 7 Å through the 100 K run, 8 Å for the 200 K run, and 9 Å for the remainder of the equilibration. A time step of 2.0 fs was used in all equilibration molecular dynamics.

Molecular Dynamics Simulations. Molecular dynamics was performed at 298 K in the NPT ensemble. The temperature was maintained by weak coupling to a temperature bath by separate solute and solvent temperature scaling using a 0.1 ps coupling constant. The pressure was maintained at 1.0×10^5 Pa, using a compressibility of 4.5×10^{-10} m^2/N and a pressure relaxation time constant of 0.5 ps. The time step during dynamics was 2.0 fs. An analysis run was conducted for 20 ns with velocity reassignment at the beginning of each nanosecond. Conformational analysis was conducted at every step, pore and cavity calculations at every 0.05 ps, and all other properties at every 0.5 ps.

Results and Discussion

General Structural Information. The cryptophane remains roughly spherical throughout the 20 ns simulation. This is illustrated in Figure 3 which displays stereoplots of the cryptophane at 2 ns intervals. The surfacing in Figure 3 represents van der Waals radii as drawn by QUANTA.[23]

Coordinates for the water molecules were saved along with the cryptophane structure every 0.5 ps during the first nanosecond and every 0.05 ps during the nineteenth and twentieth nanoseconds. This information was used along with the cavity radii (calculated as described below) to determine the number of water molecules within the cavity at each of these structures as well as the average number within the cavity. For each structure, its center of geometry based on the cap and linker regions (atoms making up the "ARYL" and "PRPL" fragments) was determined. A water molecule was considered to fall within the cryptophane's cavity if its oxygen atom lay within the cavity radius from the center of geometry. As determined in this manner, the number of water molecules within the cryptophane cavity ranged from 0 to 5, with an average number of 2.1. Figure 4 displays stereoplots of sample structures containing 0−5 water molecules within the cryptophane cavity. The surfacing in Figure 4 represents van der Waals radii as drawn by QUANTA.[23]

Conformational Analysis. Because the cryptophane molecule has a few dozen single bonds for which transitions among distinct rotameric states might occur, comprehensive sampling of the solute conformations cannot be expected in current molecule dynamics simulations. However, given the fact that the cryptophane is a somewhat rigid molecule with its six aromatic rings, short side chains, and high intramolecular connectivity, it may be possible to sample a fairly representative set of the conformations available to the cryptophane at ambient temperatures.

The rate at which unique conformations are being sampled with simulation time was computed by analyzing the cryptophane's dihedral angles. To do this, each angle was divided into wells. Divisions between wells occur at the maxima in the dihedral potential function for each angle. At each molecular dynamics step, every angle was assigned an integer designating which well the angle currently occupied. In this way, the conformation of the cryptophane at each molecular dynamics step was defined by a string of integer values. Included with the integer string was the molecular dynamics step at which it was created. Excluding the indistinguishability of certain atoms and symmetry considerations, two integer strings defined the same conformation only if they were identical.

For the cryptophane with four of the six acid protons removed and modeled in the AMBER parameterization, there are a total of 412 proper and improper dihedral angles. Of these 412 dihedral angles, 212 did not undergo even a single rotameric transition during the course of the entire 20 ns simulation and therefore did not contribute to the formation of unique conformations. These angles largely involve atoms making up the two cyclotriveratrylene groups, and it was expected that they would not undergo rotameric transitions. Excluding conformations involving transitions of these 212 angles on the basis of energetics significantly reduces the number of conformations available to the cryptophane.

In order to remove some of the remaining improper dihedral angles and to simplify the conformational analysis, angles

(18) Gaussian 92, Revision C.4, Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Fill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J., Pittsburgh, PA, 1992.

(19) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.

(20) Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L. *J. Phys. Chem.* **1994**, *98*, 13077−13082.

(21) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(22) Straatsma, T. P.; McCammon, J. A. *J. Comput. Chem.* **1990**, *11*, 943−951.

(23) Biosym Technologies/Molecular Simulations Inc. 1994, 9685 Scranton Road, San Diego, CA 92121-4778.

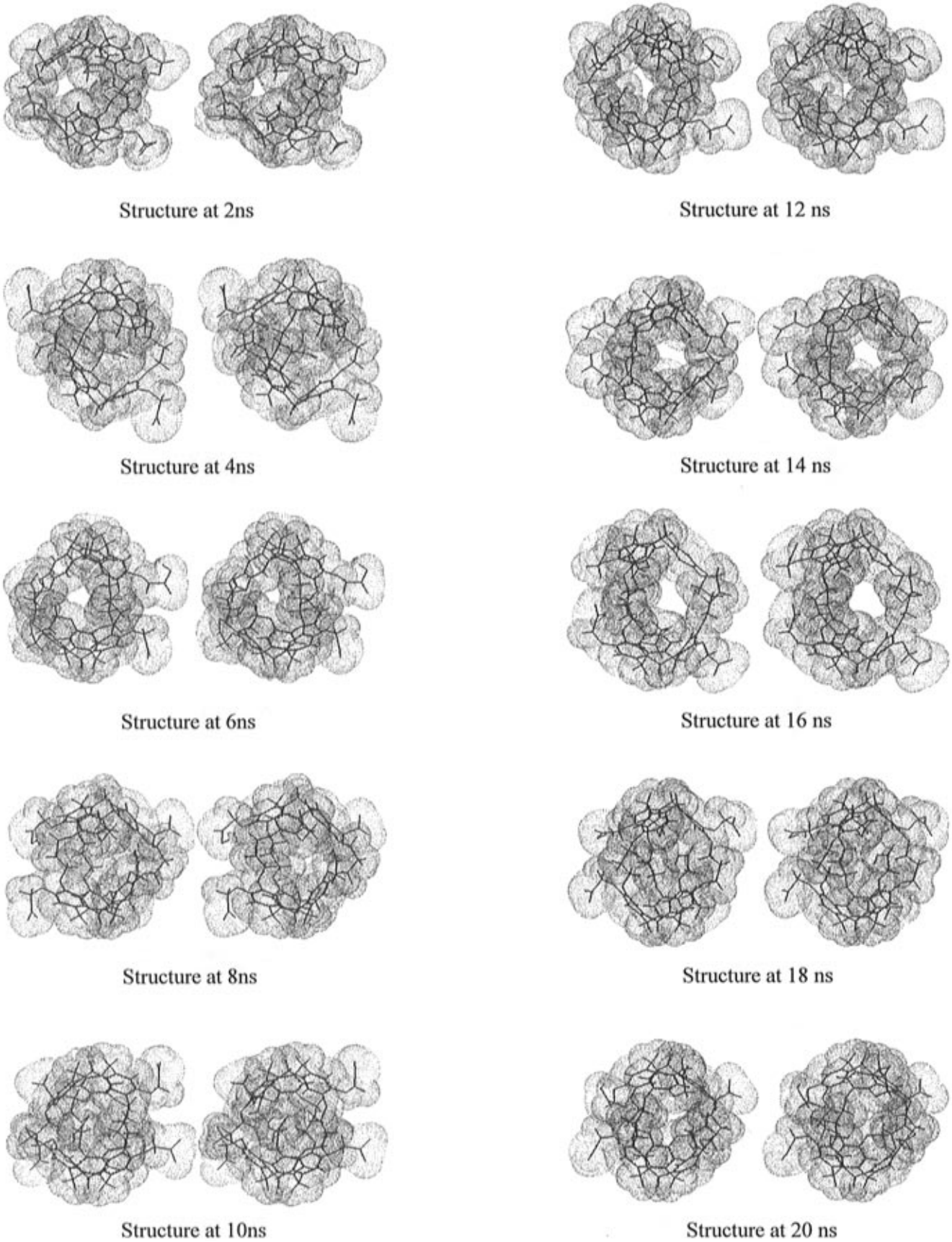

Figure 3. Stereoview snapshots of the cryptophane at 2 ns intervals throughout the course of the 20 ns simulation. Surfacing represents van der Waals radii as drawn by QUANTA.[23]

involving protons were excluded from consideration. This further significantly reduces the number of conformations available to the cryptophane. After removal of the 212 inactive angles and 146 angles involving protons, only 54 active angles remain. This includes some effective redundancies, since, for example, $O-C_1-C_4-O_5$ and $O-C_1-C_4-O_6$ of Figure 2 were considered independently. Defined in this fashion, there are 2.66×10^{19} conformations available to the cryptophane. In

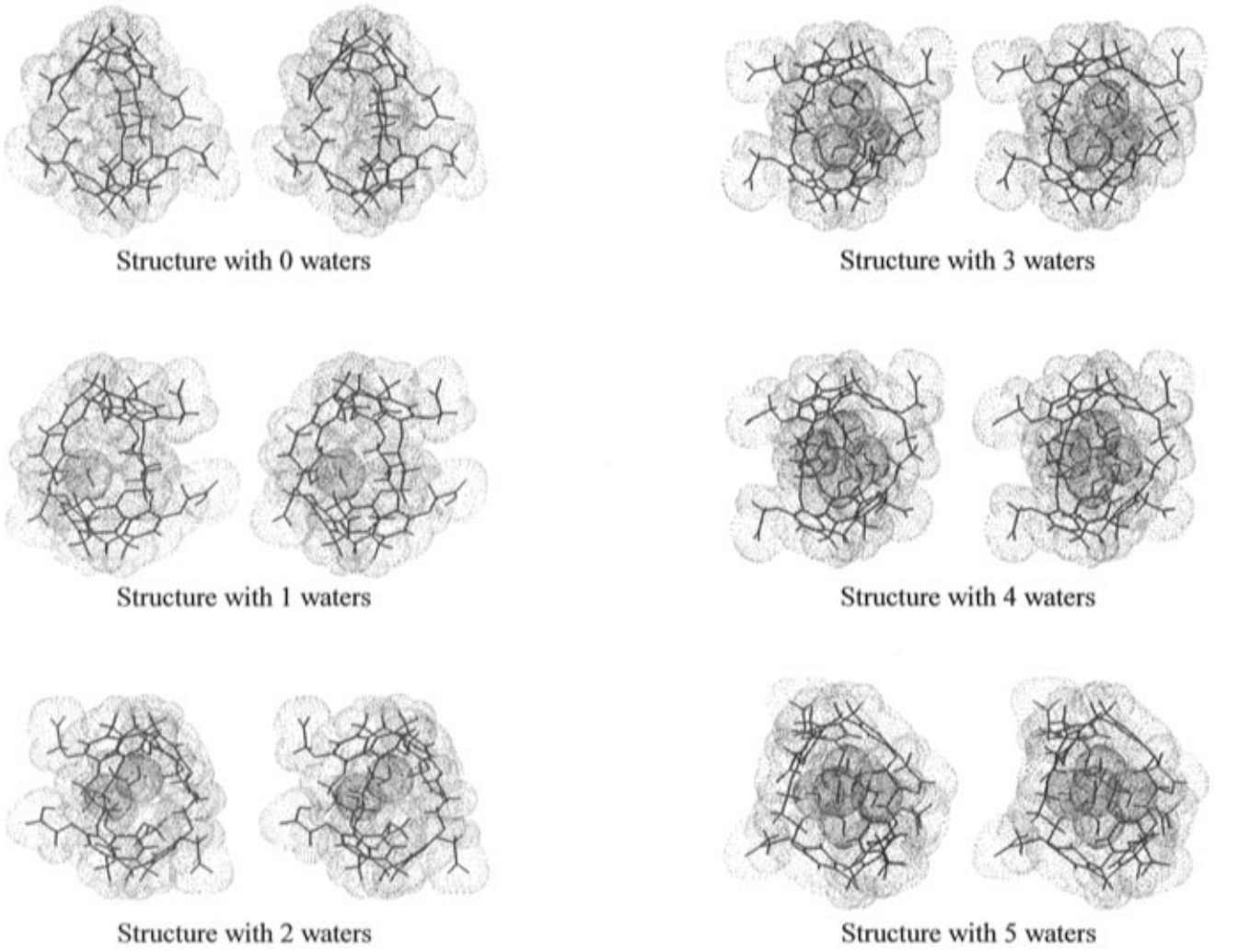

Figure 4. Stereoplots of sample structures containing 0−5 water molecules within the cryptophane cavity. Surfacing represents van der Waals radii as drawn by QUANTA.[23] Water molecules have been given a darker surface to differentiate them from the cryptophane.

addition to including some redundancies, many of these conformations will not be available to the cryptophane on an energetic basis.

To determine the rate at which unique conformations were being sampled with simulation time, the conformational analysis was conducted in a top-down fashion. Starting at the beginning of the simulation, unique conformations and the simulation time in which they were generated were saved. Conformations which were not unique, having been sampled earlier in the simulation, were removed. The result was a list of unique conformations and the time in which they were first sampled.

In addition to conducting conformational analysis on the "whole" cryptophane, conformational analysis was also conducted separately on two regions 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 guest molecules.

Table 3 lists the number of angles of period 2 and period 3, the number of possible conformations, and the number of unique conformations sampled during the 20 ns simulation for the cryptophane and each of the two regions. Not surprisingly, the number of unique conformations sampled was only a small fraction of the number of conformations that are theoretically available in each of the three cases excluding energetic considerations. These numbers give an indication of what is

Table 3. Results of Conformational Analysis

region	period 2	period 3	no. of possible conformations	no. sampled unique	no. sampled unique − equivalent
whole	36	18	2.66×10^{19}	890 042	889 624
exo	24	6	1.22×10^{10}	154 868	
cage	12	12	2.18×10^{9}	49 280	29 749

happening within the cryptophane. As the number of unique conformations sampled in the cage and exo regions approach the number of conformations possible for each of the two regions, their product will approach the number of unique conformations sampled for the whole cryptophane. However, when only a very small number of the conformations available are being sampled, the sum of the number of unique conformations sampled in the two regions, compared to the number of unique conformations sampled for the whole cryptophane, gives an indication of the correlation of the rotameric states that occur in the two regions. For a system in which there is very high correlation between the two regions, the sum of the unique conformations sampled in the two regions should be greater than the number of unique conformations sampled for the whole system. For a system in which there is very high correlation within one or both of the regions but relatively little correlation between the two regions, the sum of the unique conformations sampled in the two regions should be less than the number of unique conformations sampled for the whole system. This is what was observed, and it is exactly what one would expect

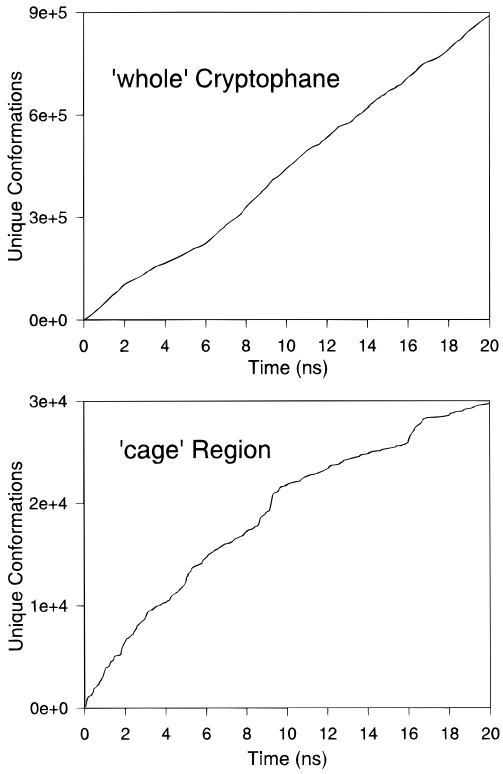

Figure 5. Unique conformations sampled with simulation time for the whole cryptophane and the cage region. Equivalent structures have been removed.

for the cryptophane and the way the cage and exo regions have been defined. Motions within the cage region should be highly correlated, motions within the exo region should have little correlation, and there should be relatively weak coupling of motions between the two regions.

In conducting the conformational analysis, the equivalence of the generated structures was considered. Two structures were considered equivalent if the molecule could be rotated so that the corresponding rotameric states were the same. The results for the number of unique conformations sampled for the cryptophane and the cage region with the equivalent conformations removed are also displayed in Table 3. Not unexpectedly, there were very few equivalent conformations for the cryptophane and relatively many for the cage region. This results from two factors. The whole cryptophane simply has more angles, and therefore the probability that two structures are equivalent is greatly reduced. Also, the correlation of motions within the cage region is very high, but correlation of motions within the exo region is very low.

Plots of unique conformations sampled with simulation time for the whole cryptophane and the cage region are displayed in Figure 5. For the whole cryptophane, unique conformations are still being rapidly sampled even after 20 ns of simulation and show no indication of slowing. The number of unique conformations being sampled with simulation time for the cage region, however, is beginning to slow. The difference in slope between these two graphs results from the same factors discussed in the previous paragraph.

Pore and Cavity Calculations. Pore and cavity radii were determined for the cryptophane every 0.05 ps throughout the 20 ns simulation. The radii were determined to further describe the cryptophane structurally and to make possible predictions about guests which could bind in whole or part within the cryptophane. Pore radii were also determined to gain an indication of how fluctuations in the pore radius might affect binding rates for guests of various sizes.

Pore and cavity radii were obtained for each cryptophane structure analyzed by generating a series of Lee and Richards[24] like probe-accessible dot surfaces using a series of probe radii. The method used, and the type of analysis described below, are similar to those used in a recent study of structural fluctuations in an enzyme that may allow substrate to penetrate to a buried active site.[25] The surfaces were generated using an all-atom representation with atomic radii based on one half the OPLS Lennard−Jones σ values used in the molecular dynamics simulation. The actual values used for the atomic radii in the surface generation were carbon 1.81 Å, oxygen 1.50 Å, and hydrogen 1.25 Å. All water molecules were excluded from these calculations. Only the coordinates of the cryptophane were used.

The probe-accessible dot surface for a given probe radius was then generated by evenly distributing 182 "sphere" points around the center of each atom at a distance of the atomic radius plus the probe radius for that surface. To create the surface, every sphere point of every atom was checked to see if it lay within a distance of the atomic radius plus the probe radius of any other atom in the cryptophane. Sphere points which did lie within this distance were eliminated. The remaining sphere points made up the surface which could consist of one continuous surface or a number of discrete surface clusters.

For each probe-accessible dot surface generated, the number of discrete clusters of surface points was determined. A cluster of points was considered discrete if the points were separated from other points outside the cluster by a distance relative to the spacing of the 182 sphere points. The value used for this distance was twice the greatest nearest neighbor distance of the sphere points distributed on a sphere with a radius equal to the carbon atomic radius plus the probe radius used in the generation of that surface.

The pore and cavity radii for a given cryptophane structure were then determined by following the formation of discrete clusters of surface points as the probe radius was varied. The pore radius was determined by increasing the probe radius incrementally until two discrete clusters of surface points were formed: one cluster of surface points forming the outer surface of the cryptophane and the other clusters forming the internal cavity surface. This radius defines the maximum size of a probe which can enter into the cavity for a given structure of the cryptophane. No differentiation was made between the three pores. Only the radius for the largest pore of each cryptophane structure was determined. As the probe radius was further increased, a point was reached where the internal cavity surface no longer had surface points. This radius defined the size of the internal cavity for a given structure of the cryptophane.[26]

Figure 6 illustrates the method used to determine the pore and cavity radii. The structure used in this figure was calculated to have a pore radius of 1.97 Å and a cavity radius of 2.43 Å. The surfacing in Figure 6 was produced using the "solvent accessible surface" routine in QUANTA[23] which uses a Connolly surface calculation. The surfaces in Figure 6a−c were produced with probe radii of 1.8, 2.0, and, 2.5 Å, respectively. As can be seen in Figure 6a, a single discrete set of surface

(24) Lee, B.; Richard, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.

(25) Gilson, M. K.; Straatsma, T. P.; McCammon, J. A.; Ripoll, D. R.; Faerman, C. H.; Axelsen, P. H.; Silman, I.; Sussman, J. L. *Science* **1994**, *263*, 1276−1278.

(26) For very small probe radii, especially in the all-atom representation, discrete clusters of surface points can form which do not represent the outer and cavity surfaces. Therefore as a matter of practice, calculations were started with a large probe radius which was incrementally decreased to determine the cavity and then the pore radius for a given cryptophane structure.

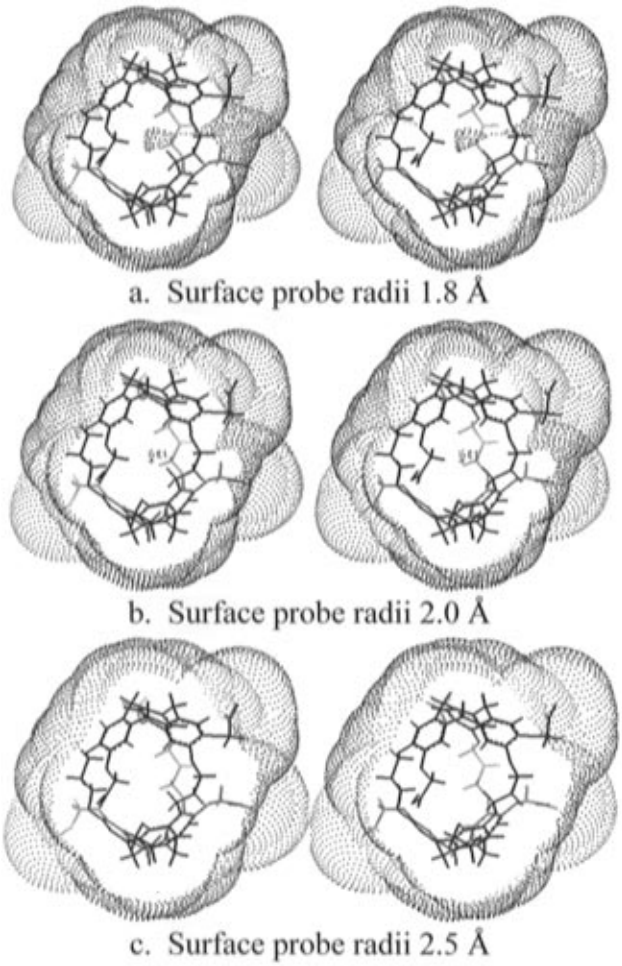

Figure 6. Illustration of the method used to determine pore and cavity radii for the cryptophane. The structure used in this figure was calculated to have a pore radius of 1.97 Å and a cavity radius of 2.43 Å. The surfacing was produced using the "solvent accessible surface" routine in QUANTA[23] which uses a Connolly surface calculation.

points exist, in Figure 6b, two sets of discrete surface points exist, and in Figure 6c, once again a single discrete set of surface points exist.

During the course of the simulation, the calculated distributions of pore and cavity radii converge within about 5−10 ns. Figure 7 displays the normalized pore and cavity radii distributions as a function of simulation time. The pore distribution appears to converge somewhat more quickly than the cavity distribution.

For the structures sampled in the course of the 20 ns simulation, the pore had an average radius of 1.57 Å, a minimum of 0.59 Å, and a maximum of 2.46 Å. The cavity had an average radius of 2.07 Å, a minimum of 1.40 Å, and a maximum of 2.74 Å.[27] Figure 8 displays stereoplots for the maximum and minimum, pore, and cavity structures sampled during the course of the 20 ns simulation. The surfacing in Figure 8 represents van der Waals radii as drawn by QUANTA.[23] Figure 9 displays plots of pore and cavity radii as a function of

(27) This algorithm will fail for structures in which the pore radius is larger than the cavity radius. Of the 400 000 structures for which the pore and cavity radii were determined, this occurred 4420 times. Average radii values reported here completely exclude structures for which the algorithm failed. For all other calculations involving cryptophane radii, missing pore values were replaced by averaging pore values directly before and after the missing value, and likewise, missing cavity values were replaced by averaging cavity values directly before and after the missing value. Generally only one or two structures failed sequentially with a maximum number of 12 structures failing sequentially.

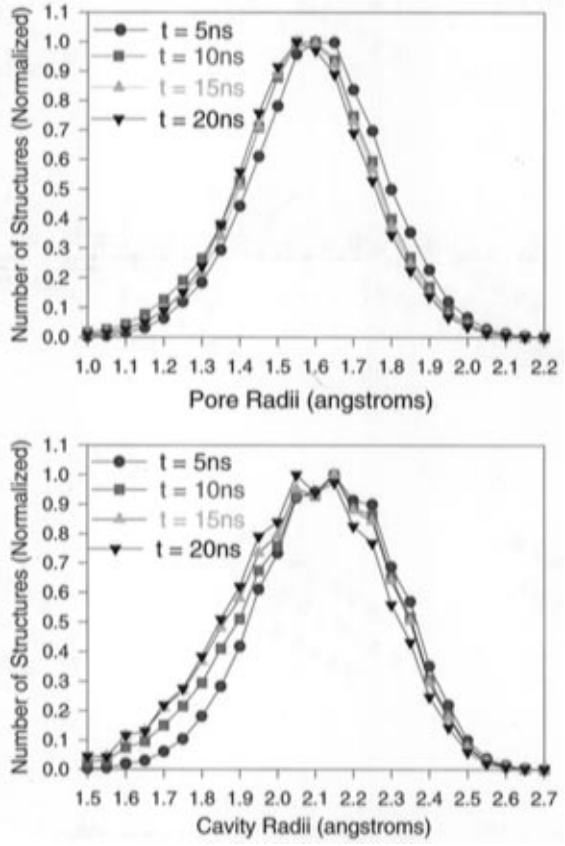

Figure 7. Convergence of pore and cavity radii with simulation time. The plots represent distributions of radii at increasing durations of simulation time. Each plot has been normalized to its maximum bin value.

simulation time. The data are plotted at 5 ps intervals for the 20 ns simulation.

To gain a rough indication of the ease or difficulty with which the cryptophane could admit a given guest into its cavity, percentage plots of structures with pore and cavity radii less than the probe radius were generated as a function of the probe radius. These plots are displayed in Figure 10. The vertical line on the plot is drawn at a probe radius of 2.10 Å. This radius represents the effective radius of a tetramethyl ammonium ion (TMA^+), which is known to bind within the cryptophane.[11] The effective radius of TMA^+ used here was defined as the distance between the nitrogen at the geometric center of the molecule and one of the methyl protons. This radius is smaller than the largest distance from the nitrogen to the van der Waals surface of the molecule, reflecting the fact that the aspherical guest will tend to present less than its maximum cross section as it enters a pore.

From these plots, 46% of the structures sampled could have contained a TMA^+ ion within the cavity; however, only 0.075% of the structures had at least one pore large enough to admit a TMA^+ ion. This indicates that the binding rates of a number of the guests of interest will be influenced by the fluctuations of the cryptophane pores.

Stochastic Gating. A number of guest molecules are known to bind in part or whole within the cavity of the cryptophane.[8−11] The steady-state binding rate of these guest would presumably be nearly diffusion limited if it were not for the fact that the guest must pass through a fluctuating pore to reach the cavity of the cryptophane. The effect of the cryptophane pore fluctuations on binding rate of a particular guest can be analyzed at

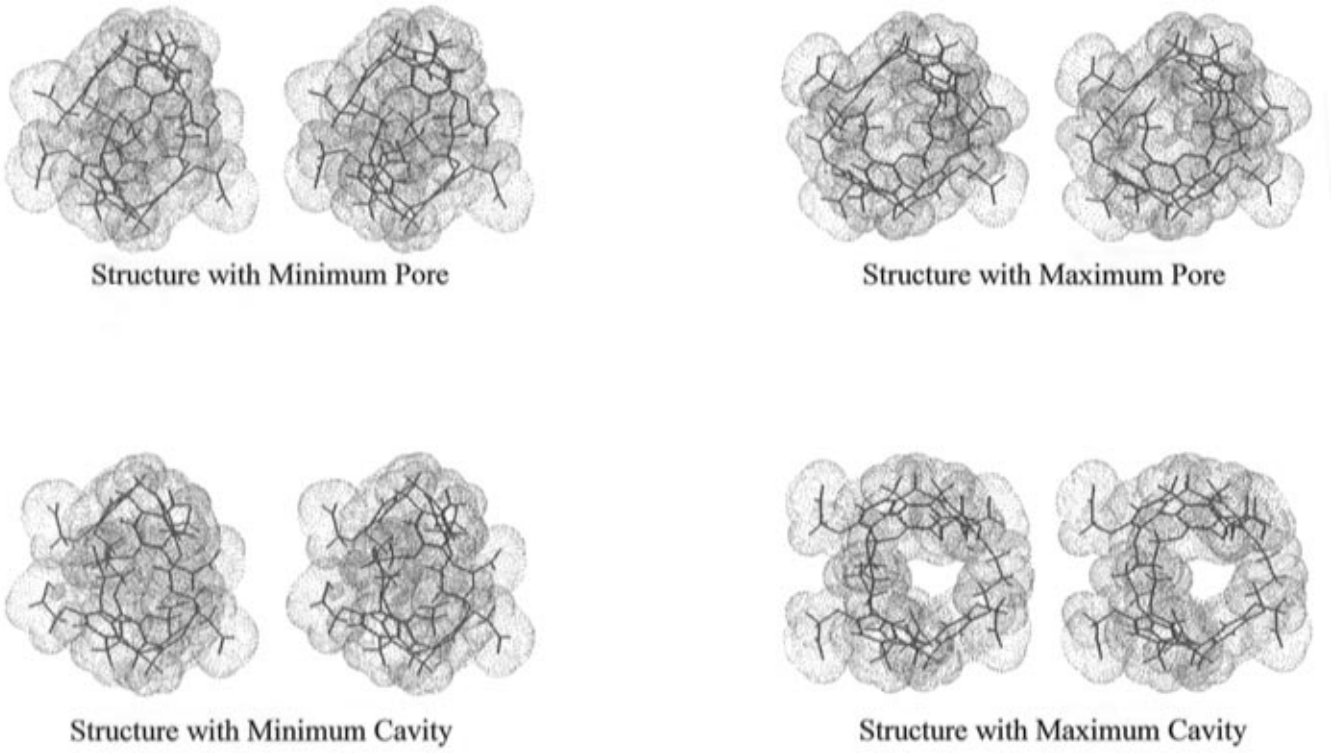

Figure 8. Stereoplots for the minimum and maximum, pore, and cavity structures sampled during the course of the 20 ns simulation. Surfacing represents van der Waals radii as drawn by QUANTA.[23]

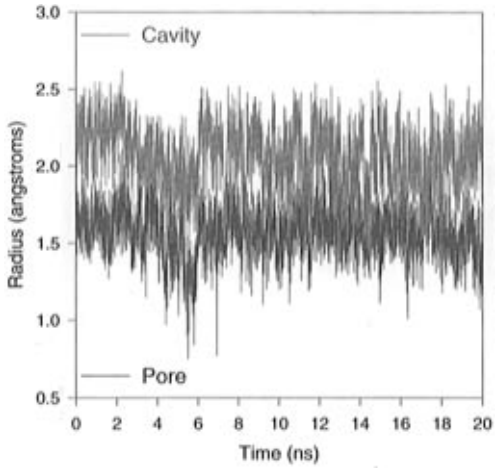

Figure 9. Pore and cavity radii as a function of simulation time plotted at 5 ps intervals for the 20 ns simulation. The pore had mean, minimum, and maximum values of 1.57, 0.59, and 2.46 Å, respectively. The cavity mean, minimum, and maximum values of 2.07, 1.40, and 2.74 Å, respectively. These values are based on all pore and cavity radii determined for the structures sampled at 0.05 ps intervals throughout the entire 20 ns simulation.

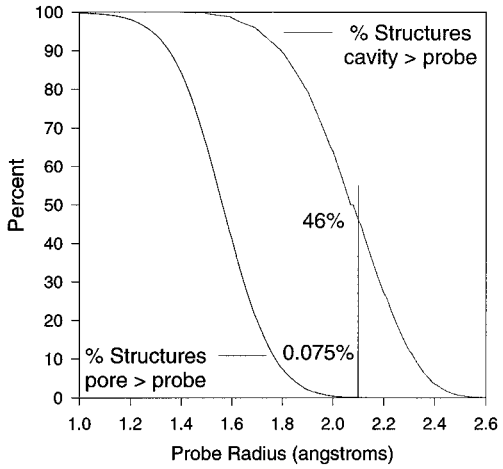

Figure 10. Percentage plots of structures with pore and cavity radii less than the probe radius as a function of probe radii. The vertical line is drawn at a probe radius of 2.10 Å which is the effective radius of a TMA^+ ion in the AMBER representation. The percentages of structures with pore and cavity radii greater than the probe radius of 2.10 Å are indicated on the graph.

least qualitatively as a stochastically gated reaction. There are two limiting cases for a diffusion-influenced, stochastically gated reaction.[28]

In one limiting case, the opening and closing of the gate (e.g. the fluctuation of the pore radii) is slow compared to the characteristic time for diffusion of the cryptophane and the guest. In this case (case I), the overall rate constant is the steady-state association rate constant for the reaction with the gate fixed open, multiplied by the equilibrium probability that the gate is open. In the other limiting case (case II), the opening and closing of the gate is fast compared to the characteristic time for diffusion of the cryptophane and the guest. In this case the overall rate constant is simply the steady-state association rate constant for the system with the gate fixed open. This is due to the fact that the gate is very likely to open during any diffusional encounter.

To gain insight into the effect of guest size on the binding rate, the characteristic time for diffusion of the cryptophane and guest was compared to a gating period defined for the cryptophane based on the fluctuations of its pore radii. The analysis here is qualitative. A more detailed analysis based on the theory of ref 28 does not seem warranted in view of the approximations made for the parameters of the theory as applied here.

Both the characteristic time for diffusion and the gating period are functions of the guest size. These functions are displayed

(28) Szabo, A.; Shoup, D.; Northrup, S. H.; McCammon, J. A. *J. Chem. Phys.* **1982**, *77*, (9), 4484−4493.

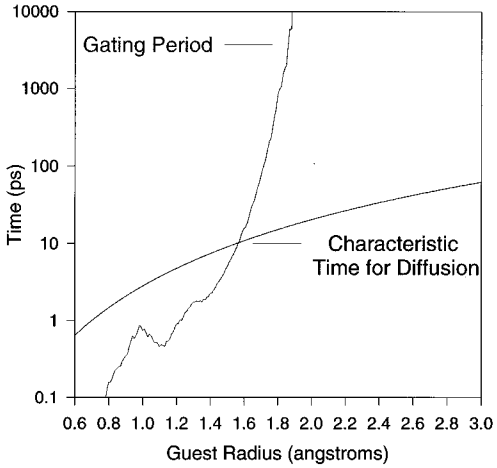

Figure 11. Influence of pore fluctuations on binding kinetics as a function of guest radius. The gating period and τ_d intersect at a guest radius of approximately 1.6 Å.

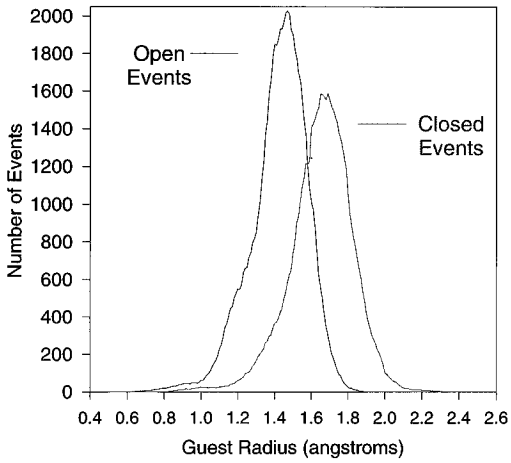

Figure 12. The number of open and closed events as a function of guest radius.

in Figure 11. The characteristic time for diffusion (τ_d) is

$$\tau_d = R^2/D_{rel} \tag{1}$$

where R is the host/guest encounter distance which was approximated as the radius of the guest and D_{rel} is the relative diffusion coefficient of the guest and the cryptophane. The argument for using the guest radius for R is that this provides a reasonable estimate of the distance apart the solutes have to diffuse to end an encounter. D_{rel} is the sum of the diffusion coefficients of the cryptophane and the guest. The diffusion coefficients were determined by using the molecular radii of the cryptophane and the guest in the Stokes−Einstein relation with stick boundary conditions

$$D = kT/6\pi\eta a \tag{2}$$

where k is Boltzmann's constant, T is absolute temperature, η is the solvent viscosity, and a is the molecular radius of the cryptophane or the guest.

The characteristic time for diffusion is plotted against the radius of the guest in Figure 11. The value of 7.84 Å was used for the cryptophane radius throughout the plot. This is the average radius of the cryptophane defined as the average distance between the center of geometry of the cryptophane and the carboxylate oxygen atoms. The average radius was obtained from cryptophane structures saved at 5 ps intervals throughout the 20 ns of simulation. The average temperature of the simulation, 308 K, was used for T. For η a value of 0.719 cp was used, which is the viscosity of water at 308 K.[29]

Defining a gating period for the entry of guests into the cryptophane was not completely straightforward for a number of reasons. First, the pore radius obtained for each cryptophane structure was the maximum value for the three pores. Also, the gating behavior as described here was based on a sequence of structures of the cryptophane in the absence of any guest, except for solvent molecules. As a guest actually approaches the cryptophane, it may perturb the structure of the host molecule, increasing or decreasing the "open" and "closed" times. For this reason, structures intermediate between open and closed were defined. A structure was initially defined as intermediate if its pore radius was within 5% of the probe radius. Structures with a pore radius greater than the probe radius plus 5% of the probe radius were defined as open. Structures with a pore radius less than the probe radius minus 5% of the probe radius were defined as closed. The intermediate structures were then defined as being either open or closed depending on the state of the five structures immediately preceding and following the intermediate structure. In this manner, a series of open and closed structures were obtained which have been smoothed to reduce errors associated with the determination of the pore radii and minor fluctuations of the pores.

Finally a guest will require a certain amount of time to pass through a cryptophane pore and ultimately bind within the cavity. Defining the typical amount of time is somewhat arbitrary, but will depend in part on the size of the guest. To remain consistent with the way in which characteristic time for diffusion was determined, the following Einstein relation was used:

$$6tD_{rel} = \langle |r_i(t) - r_i(0)|^2 \rangle \tag{3}$$

where t is the typical amount of time required for the guest to pass through a cryptophane pore and D_{rel} is the relative diffusion coefficient obtained from the Stokes−Einstein relation. The square of the radius of the guest was used for the mean square displacement, $\langle |r_i(t) - r_i(0)|^2 \rangle$. With the amount of time required for a guest of particular size to pass through a cryptophane pore determined, an "open event" can be defined. An open event was defined as an unbroken series of open structures corresponding to a length of simulation time of at least t. A "gating event" was then defined as the time between the end of one open event and the beginning of the next open event. The gating period was defined as the sum of gating event time divided by the number of open events. The gating period during the course of the 20 ns simulation has been determined and plotted as a function of guest radius in Figure 11.

As is apparent from Figure 11, the gating period defined in this way is not a simple monotonic increasing function. This is due in part to the discrete nature of the analysis. As the guest radius increases, the sum of gating event time increases sporadically. Open events of long duration break into open events of shorter duration as the guest radius exceeds the pore radius of individual structures within the open event. Open events of small duration disappear completely as t exceeds their length of duration. The actual number of open events increased rapidly up to a maximum number at a guest radius of 1.5 Å and then rapidly decreases with increasing guest radius. The number of open and closed events is plotted as a function of guest radius in Figure 12.[30]

(29) *CRC Handbook of Chemistry and Physics*; 65th ed.; CRC Press, Inc.: Boca Raton, FL, 1984; pp F−37.

As can be seen in Figure 11, values for the gating period and the characteristic time for diffusion intersect at a guest radius of approximately 1.6 Å. In this region, the gating period is rapidly increasing relative to the characteristic time for diffusion. Therefore, guests with atomic radii larger than 1.6 Å will tend toward case I of a diffusion-influenced, stochastically gated reaction. That is, the overall rate constant for the binding of the guest within the cryptophane will be the steady-state association rate constant for binding of the guest with the cryptophane pores fixed open, multiplied by the equilibrium probability that the pores are open. The pore size and fluctuations affect the binding rates of guests with radii larger than 1.6 Å.

Guests with atomic radii smaller than 1.6 Å will tend toward case II of a diffusion-influenced, stochastically gated reaction. That is, the overall rate constant for the binding of the guest within the cryptophane will be the steady-state association rate constant for binding of the guest with the cryptophane pores fixed open. The pore size and fluctuations do not affect the binding rates of guests with radii smaller than 1.6 Å.

Concluding Remarks

In this study, molecular dynamics simulations have been used to obtain structural and conformational sampling information on the cryptophane in an aqueous solution. The extent to which a cryptophane cavity can be solvated by water has been explored since the presence of solvent molecules within the cryptophane cavity will likely influence the binding of another guest molecule. The main purpose of this study, however, has been to analyze the expected dependence of binding kinetics upon the size of the guests. This analysis is largely qualitative in nature. Even so, it illustrates how the binding of guests depends on the existence of structural fluctuations in the cryptophane molecule and how the time scales of these fluctuations and the diffusional displacements may influence the kinetics of binding of the larger guests.

(30) A "closed event" was defined as an unbroken series of closed structures corresponding to a length of simulation time of at least *t*.

Acknowledgment. We thank Dr. T. P. Straatsma for the use of ARGOS and many useful discussions, Dr. Michael Gilson for providing the surfacing code adapted for use in the pore and cavity determinations and for his advice, and Dr. Andr Aubry for providing the coordinates of the cryptophane-O. Finally, we thank Dr. Tami Marrone, 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 NSF and the NSF Supercomputer Centers MetaCenter Program.

JA953772H