JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

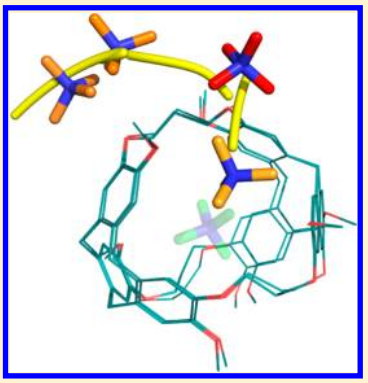
Ligand Binding Pathway Elucidation for Cryptophane Host–Guest Complexes

Christopher C. Roberts and Chia-en A. Chang*

Department of Chemistry, University of California, Riverside, California 92521, United States

Supporting Information

ABSTRACT: Modeling binding pathways can provide insight into molecular recognition, including kinetic mechanisms, barriers to binding, and gating effects. This work represents a novel computational approach, Hopping Minima, for the determination of conformational transitions of single molecules as well as binding pathways for molecular complexes. The method begins by thoroughly sampling a set of conformational minima for a molecular system. The natural motions of the system are modeled using the normal modes of the sampled minima. The natural motions are utilized to connect conformational minima and are finally combined to form association/binding pathways in the case of molecular complexes. We provide an implementation and example application of the method using alanine dipeptide and a set of chemical host–guest systems: two cryptophane hosts with two guest cations, trimethylammonium and tetramethylammonium. Our results demonstrate that conformational transitions can be modeled and extended to find binding pathways as well as energetic information relevant to the minimum conformations involved. This approach has advantages over simulation-based methods for studying systems with slow binding processes and can help design molecules with preferred binding kinetics.



INTRODUCTION

The mechanisms of molecular association are critical to the understanding of many problems in chemistry, biology, and medicine.[1−4] In addition to studying the final ligand–receptor bound complexes, investigating binding processes helps elucidate fundamental mechanisms, including allostery, induced fit, and gated control associations. An understanding of the binding pathway of a molecular complex can lend thermodynamic and kinetic information, which is relevant in medical and pharmaceutical research, and can aid in the development of new or improved drugs and drug carriers. The kinetic behavior, in particular, is relevant to the efficacy of drugs *in vivo*.[5,6] This study focuses on chemical host–guest systems as approachable ligand–receptor models to gain insight into the thermodynamics and kinetics of binding.[7,8] Supramolecular systems are ideal systems for method development and optimization because they are simple but preserve all the important characteristics of more complicated ligand–protein systems.

The standard free energy change of binding is an alternative way of expressing its equilibrium constant, $K_{eq} = \exp(-\Delta G^{\circ}/RT)$, which in turn is the ratio of the rate constants for association (k_{on}) and dissociation (k_{off}), $K_{eq} \approx k_{on}/k_{off}$.[9−11] A free energy barrier is coupled with association or dissociation rates, which may result from entropy loss, a desolvation penalty, and conformational changes. Important and elegant work has been done on the gated theories and statistical thermodynamics of noncovalent association.[2,12−14] However, real molecular systems are usually quite complicated, and the theoretical studies and analytical solutions may not be able to provide direct links to real molecular systems of practical interest. Computational modeling is a powerful tool used to explain fundamental processes of molecular binding.

Computational methods capable of establishing binding pathways traditionally rely on various forms of molecular dynamics (MD) or Monte Carlo (MC) simulations.[15−25] Standard simulations are computationally expensive, and in some cases impractical, for this determination because of the relatively long time scales under which binding processes typically occur.[8] In addition, many repeated simulations are required to adequately sample possible binding pathways.[26] Dedicated hardware has been engineered to make this direct simulation route more practical, but at a high monetary cost.[27,28] This has been more commonly addressed, at the cost of accuracy, through the use of accelerated dynamics simulations or through a reduction in system complexity with coarse-grained models and/or Brownian dynamics simulations.[29−40] The computation of ligand binding free energy barriers is even more challenging. One popular method is to apply steered MD simulations to gradually "pull" a ligand from the ligand binding site to the solvent or use simulations with various sampling techniques to obtain the potential of mean force (PMF).[20,22,23,41−48] Other methods apply metadynamics or the adaptive biasing force (ABF) methods to bias the simulations.[31,38,49,50] These methods generate a free energy profile along a ligand dissociation/association pathway; thus, an energy barrier may be observed from the PMF plot. However, because the pathways are not known beforehand, one may need

Received: November 20, 2012
Published: March 5, 2013

to perform very thorough sampling.[18,45,51−55] Moreover, both the ligand and receptor may undergo conformational changes during ligand association, so adequate sampling to compute an accurate binding free energy landscape can be difficult.

Because transition/intermediate states are too rare, standard unbiased dynamics-based simulations have difficulty finding accurate pathways and free energy landscapes. Many methods, for example transition path sampling (TPS), nudged elastic band (NEB), and string methods, have been developed to determine proper transition pathways between adjacent stable states without making *a priori* assumption about transition paths.[24,49,56−65] They provide accurate coordinates and energy barriers of transition states and can be applied to study chemical reactions or conformation transitions. However, applying these rigorous and computationally intensive methods to protein−ligand association remains a daunting task; especially binding pathways which cover large configuration space and involve several states.[66]

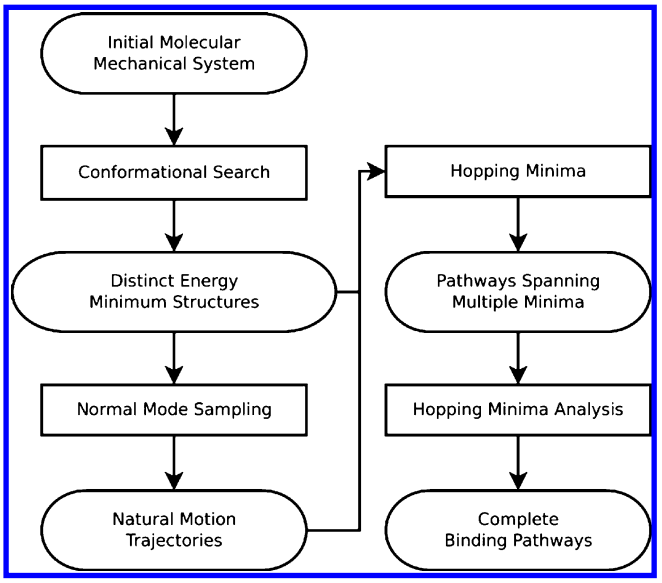
This article describes a new methodological approach called Hopping Minima for the determination of conformational transitions and association/dissociation pathways by connecting local energy minima using the molecular system's natural motions. The natural motions of a molecular system are modeled from their normal modes, which are exaggerated into the coordinated motions of the molecule or host−guest complex. The method is similar to the superposition approach, termed the reaction path Hamiltonian superposition approach (RPHSA), developed by Strodel and Wales; however, there are several key differences between the two methods.[67] Both methods begin with using a conformational search engine to find local energy minima and utilize harmonic approximations. Our objective is to rapidly identify connected intermediate states to determine binding pathways from numerous local energy minima. Because we aim to work on much more complex molecular systems, we have developed an automated algorithm to bridge minima in the HM method using internal coordinates. The RPHSA method obtains a more detailed free energy surface (FES), but due to the concern of computation time, our method provides a simpler approximation for the FES of a ligand−receptor binding pathway.

We present a proof of concept implementation of the method which uses the conformational search program Tork to perform conformational searches and a set of programs to carry out the detection of conformational transitions and building of the binding pathways.[68−71] The method is exemplified by utilizing our implementation to connect distinct conformations of alanine dipeptide. Alanine dipeptide has served as a standard model system for various computational works, including the development of new sampling methods, benchmarking of force fields, and determination of transition paths between two stable states.[56,58,66,67] Unlike existing works that constructed detailed free energy surfaces of a transition pathway, our work shows a quick method to illustrate possible pathways of conformational changes. We then extend the implementation to a set of host−guest systems: two cryptophane hosts with two cation guests. The cryptophanes are a class of spherically shaped organic molecular capsules with hollow internal cavities. They are commonly studied in the context of molecular recognition because of their propensity to bind atoms and small molecules.[72−75] For example, a ^{129}Xe-cryptophane biosensor was designed for potential use as a magnetic resonance imaging (MRI) contrast agent.[76] A well-studied member of the cryptophanes, cryptophane-E, has proven to form supramolecular complexes with small neutral and cationic guest molecules. Cryptophane-ES, a derivative of cryptophane-E with the methoxy gating arms replaced with thiomethyl gates, has been shown to have similar binding characteristics.[72] The association and dissociation rates of a guest binding to cryptophane-E can be 10^3 to 10^4 times faster than with the same guest binding to cryptophane-ES, but the net binding affinities (ΔG) are essentially the same. However, it is not clear how such a tiny difference in the arms of the host can significantly affect the kinetics but not the thermodynamics. We therefore applied the method presented in this work to crytophanes-E and -ES with trimethylammonium and tetramethylammonium cation guests to reveal interesting guest binding pathways and shed light on the nature of the interactions between cryptophanes and their guests. The work also characterizes changes in free energy, configurational entropy, and enthalpy of minimum conformations along ligand−receptor association pathways

METHOD

Overview. The procedure, termed "Hopping Minima," has been summarized as follows (Scheme 1). An extended conformational search procedure is performed on the molecular system. Free energies of the individual minima are calculated using the Mining Minima generation 2 (M2) program. Natural motions are then sampled using a modified M2 program.[68] The second derivative (Hessian) matrix of the potential energy function is computed and diagonalized to obtain eigenvalues and eigenvectors in Bond-Angle-Torsion (BAT) coordinates for each local minimum.[70,71] The eigenvalues are sorted, and only eigenvectors corresponding to the first three to 10 small eigenvalues are considered. Soft modes, or normal modes with a small eigenvalue, represent the most relevant natural motions of a molecular system for our purposes, as they have the largest associated ranges of vibration. The system's coordinates are scanned along the normal modes, and snapshots are saved as a

Scheme 1. Overview of the Hopping Minima Method[a]



[a]Rectangles represent distinct steps in the method, while ovals represent input and output structures and trajectories of structures. Arrows direct the work flow through the method.

trajectory during the scanning. Transitions between conformational minima are determined by structurally comparing the saved trajectories to the minimum conformations. If a motion approaches multiple local energy minima, the trajectory is saved as a molecular path. For molecular complexes, multiple paths can be combined to reveal possible ligand binding pathways. If local energy minima are numerous, several steps and analyses are carried out in order to determine tractable pathways and/or pathways that connect more popular (low energy) local energy minima.

This method relies on the principles of the energy landscape theory.[77] As such, only the local energy minimum conformations are explicitly sampled in the conformational search procedure. We rely on the natural motion sampling to capture the high energy transition state structures.

Finding Local Energy Minima. Tork is a conformational search utility that distorts a molecular system's initial conformation along internal coordinates, minimizing along the way in order to find new minima. This process is automated but conceptually similar to previous manual methods of exploring conformational space.[71] The results of a given search are dependent on the conformation used to seed the search. Thus, many conformations, including both high and low energy conformations, should be used as input to thoroughly sample the conformational space of a molecular system.

Unlike the free energy calculations where one can examine if the integration has converged, it is less straightforward to tell whether a conformational search is complete. Conformational search should be targeted to structures in and around conformations of relevance, if the algorithm allows. Searches should be performed until relevant conformational states are well represented. Because we aim to study association pathways, many local energy minima should be available in addition to the global energy minimum. For complexes, we are most notably interested in those local energy minima in which the ligand is in the intermediately bound and unbound states. To more efficiently find local energy minima, multiple intermediate partially bound or unbound conformations are used as initial structures for successive runs of conformational search. Root mean square deviation (RMSD) based alignment of all conformations can help reveal areas around the host with sparse ligand population.[78] Minimum conformations with a guest molecule near the sparsely populated areas can be used to start additional searches. Because the search will result in the same conformations multiple times, repeat conformations are filtered out. Eliminating repeats is nontrivial when a molecular system has symmetries, so a symmetry detection algorithm is applied.[79] For example, for our Me_4N^+ guest, simply rotating a methyl group by 120° is considered the same conformation.

Free Energy Calculation. The M2 method is used to compute the free energy for all conformations in a system via a modified harmonic approximation. The M2 method calculates the configuration integral as a sum of contributions from the conformational states, which is Boltzmann weighted by each conformation.[68] The configuration integral for each local minimum is calculated upon the harmonic approximation. The M2 method builds upon the harmonic approximation by correcting for anharmonicities in BAT coordinates.[70,71] To compute the binding free energy, the calculation is carried out for three species: the complex, free host, and free guest. These values are then subtracted to find the absolute binding free energy for the complex. For the HM method, the free energy is calculated for all sampled conformational states of the complex only with no Boltzmann sum applied.

Natural Motion Sampling. Natural motions are modeled using the normal modes of the host−guest complex. Although the mass of the atom is not considered and it is not equivalent to classical normal modes in mechanics, we term each set of eigenvalue and eigenvectors a "normal mode." Internal BAT coordinates are built for the structures, instead of using standard Cartesian coordinates. With a set of local energy minima of intermediate states available, the second derivative (Hessian) matrix of the potential energy function is computed for each local minimum and diagonalized to obtain eigenvalues and eigenvectors. The eigenvalues are sorted, and Gaussian integrands associated with the harmonic approximation are constructed.[70,71] The molecular system's coordinates are scanned along the soft normal modes in increments of 0.05 radians or Angstroms until the eigenvectors reach the nine standard deviation (9σ) point of the Gaussian. The scanned conformations are recorded and the results output as a trajectory. In this study, we scan the lowest five normal modes to obtain five paths of natural motions for a given local energy minimum.

Identifying Paths. There are two RMSD-based methods for comparing similarity in molecular structure between natural motion trajectories and conformational minima: Cartesian coordinate RMSD and dihedral angle RMSD. Cartesian coordinate RMSD compares the absolute position of atoms between the natural motion trajectory and conformational minima. Alignment of all structures to the same reference minimum is required. Because of this alignment, two atomic selections are required: an alignment selection and an RMSD comparison selection. Cartesian coordinate RMSD is appropriate for all types of systems, whether a single molecule or complex of molecules. Dihedral angle RMSD compares all, or a subset of, dihedral angles containing only heavy atoms. This comparison is fast relative to standard Cartesian coordinate RMSD because alignment of the structures is not required. Dihedral angle RMSD is appropriate for single molecules only, as it is unable to capture relative positions and orientations of multiple molecules.

For the alanine dipeptide, dihedral RMSD and Cartesian RMSD are both employed using alignment and RMSD selections of all heavy atoms and amide hydrogens. For the cryptophane complexes, we selected atoms in the two nine-membered rings, the least flexible portions of the host molecule, for alignment.

The Hopping Minima application performs the Cartesian RMSD comparison as follows: for a natural motion trajectory t_i calculated from minimum conformation m_i, each frame in t_i, $t_{i,f}$, is aligned to m_i. All conformational minima, m_j $(j \neq i)$, are then aligned to m_i, and the Cartesian RMSD comparison is performed between $t_{i,f}$ and m_j. The dihedral RMSD comparison similarly cycles through all $t_{i,f}$ and calculates the dihedral RMSD between minima m_j $(j \neq i)$ and $t_{i,f}$. For either type of comparison, if the RMSD value is within a user-supplied tolerance, then m_i and m_j are considered to have a conformational transition described by t_i, and it is saved as a separate path trajectory.

RMSD tolerance is the main variable that affects resultant conformational transitions. RMSD tests with more lenient tolerance values result in numerous pathways, and most of them are not of interest. Therefore, we introduced several techniques to our implementation to reduce redundancy in

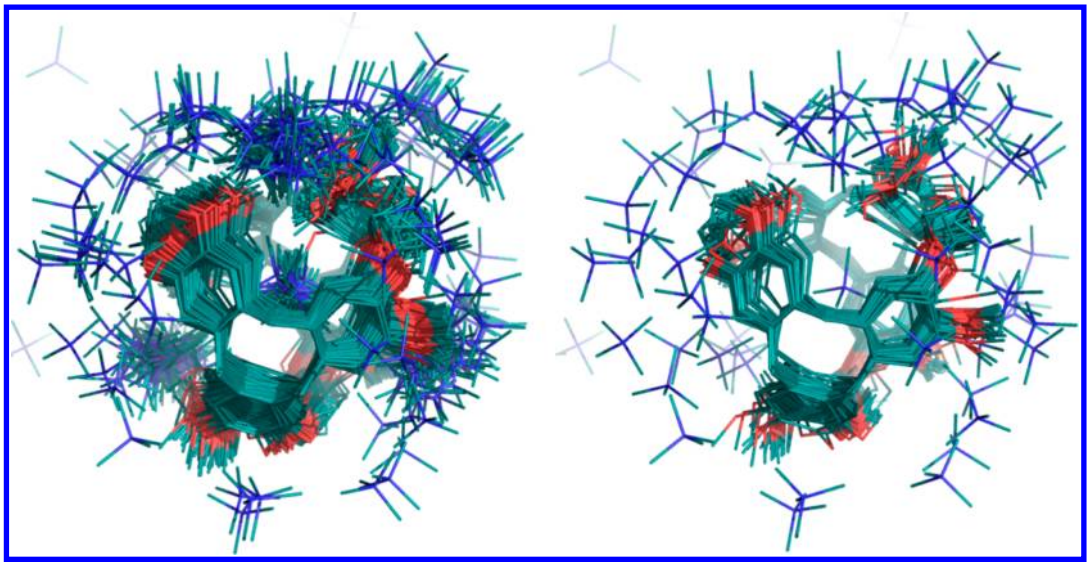


Figure 1. Thinning filter. (Left) An aligned set of conformational minima for a host−guest complex, showing the many bound and unbound ligand states. (Right) The set of minima after thinning, showing the regionally representative conformational minima. A thinning cutoff of 1.5 Å was used; thus no ligand is within a 1.5 Å radius from another.

minimum conformations and to filter out highly similar or unlikely pathways. These techniques are provided to help the user reduce noise in the results of calculations with lenient RMSD tolerance values.

Thinning. After a thorough conformational search, numerous distinct conformations, including high energy conformations, may densely populate spatially small regions near an equilibrium basin. Therefore, before building any paths or binding pathways, low-energy regionally representative conformations are determined in order to limit the search results (Figure 1). A thinning cutoff parameter is provided by the user that dictates that no two ligands will be within the cutoff distances from one another. The thinning step starts with sorting the conformations based on their computed free energy. The conformations are sequentially analyzed. Each minimum conformation's position is compared with its predecessors. If any previous conformation is within a thinning cutoff, then the conformation will be ignored; otherwise, the new conformation will be reported as a new representative conformation.

Scan Filtering. As multiple normal modes are scanned for each conformational minimum, paths formed from a given minimum may be very similar, resulting in similar paths intersecting the same local minima. Thus, if two or more motions, sampled from the same minimum conformation, yield similar paths, we will only keep the one built from the lowest normal mode.

Combining Multiple Paths. To illustrate guest association processes and possible binding pathways for host−guest systems, individual minima-connecting paths of host guest systems will be combined to form a more complete and longer binding pathway. A script was written that aids in determining combinations of relevant paths for the formation of binding pathways. This script prompts the user for minima of interest and returns information on paths involving those minima. The output is sorted by the potential energy of the path and provides information such as maximum, minimum, and average potential energy of the natural motion paths, and the index of local energy minima that the path intersects. This allows the user to inspect paths that may possibly join to form a binding pathway and quickly select paths of low potential energy. Note that high potential energy suggests that a natural motion may contain steric clashes or overly distorted structures. Although these paths are often not relevant, occasionally a rare and high potential energy pathway may be of interest. Thus, we do not explicitly exclude the high energy paths.

A strategy that we found useful was to construct the pathway of unbinding. Any paths that linked a bound state conformation to an intermediate state conformation were used as the starting point for construction. These intermediate state conformations were then used to find paths leading to further unbound conformations. This process is repeated until there are no additional linking paths.

Computational Details. The alanine dipeptide model was parametrized using the AMBER FF03 force field.[80] The initial structure of the free cryptophane-E host was obtained from the Cambridge Structural Database, ID SEDPOG.[81] As only the experimental structure for the cryptophane host was available, all guest and complex structures were created manually. We created the sulfurous derivative, cryptophane-ES, using the cryptophane-E initial structure as a template and manually constructed the ligands using Avogadro.[82] The DREIDING force field was used to describe all complex structures with parameters assigned by the program Vdock and partial charges computed with the program Vcharge with the VC/2004 parameter set.[83−85] The solvation energy, W, is computed with the generalized Born model using the water dielectric constant of 80.0 for the alanine dipeptide system and the tetrachloroethane dielectric constant of 8.42 for the complexes.[86] For the cryptophane hosts with an open cavity, the dielectric cavity radius of each atom is set to the mean of the size of chlorine (∼1.8 Å) and the atom's van der Waals radius.[87] All local energy conformations were energy-minimized by the conjugate gradient method and then the Newton−Raphson method until the energy gradient was $<10^{-3}$ kcal/mol/Å. Once the free cryptophane hosts and free guest molecules were fully constructed, we created the complex systems by combining each combination of host and guest and minimizing the structures using the aforementioned minimization methods. When running the Hopping Minima application, the alanine dipeptide required no filtering methods. RMSD tolerances of

0.2 and 0.3 Å and radians were used to explore conformational transitions of the alanine dipeptide. For the four complexes, the thinning cutoff was set to 1.5 Å, except for cryptophane-E in complex with trimethylammonium, where a value of 2.5 Å was used. The RMSD tolerance was set to 2.5 Å for all complex systems. All calculations were performed as single processes on an Intel Xeon 2.67 GHz quad core processor computer using Ubuntu Linux 10.10. Each conformational search run with Tork took about 10 min, with 10 to 15 runs performed in total on each complex system and at least six on each free host structure. All conformational minima were obtained for the alanine dipeptide after three Tork search runs. The free energy calculations took 6−8 h for the 160−255 cryptophane-E host−guest complexes, 16 h for the 144−174 cryptophane-ES host−guest complexes, 4 h for the 120 free cryptophane-E host conformations, and 6 h for the 197 free cryptophane-ES host conformations. The alanine dipeptide and ligand free energies were calculated in less than 1 min due to having few minimum conformations. RMSD-based identification of conformational transitions took less than 20 min to complete for all systems. Each manual path construction time for the complexes took an hour or less to complete.

RESULTS AND DISCUSSION

Because alanine dipeptide has been used as a model system by many computational methods, we first applied the HM method to the system as an example to demonstrate the procedures of the method. Five consecutive conformational searches were performed, at which point the conformational search converged. Additional searches were performed on the identified minima to ensure that the conformational space was thoroughly sampled. The HM application was applied using the 16 conformations that were found. Utilizing either RMSD calculation method, with no thinning of the minima, five unique conformational transitions were found for the system that connect two minimum states within 0.2 Å. Larger RMSD tolerance values revealed conformational transitions between additional minima. As seen in Figure 2, an RMSD tolerance value of 0.3 Å yielded two conformational transitions that were combined to form a transition spanning three minima. This transition is driven by rotation around the ψ and ϕ dihedral angles. The first motion primarily rotates the ψ angle, while the second motion rotates the ψ and ϕ angles in concert. These results demonstrate that for small single molecule systems, a single normal mode can accurately describe specific conformational transitions.

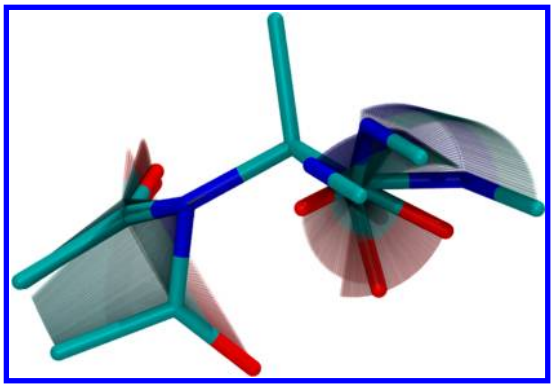


Figure 2. Conformational transitions of the alanine dipeptide. Two conformational transitions connect three local energy minimum conformations of the alanine dipeptide. The major dihedral angles of rotation are the ϕ and ψ angles of the alanine backbone. The first motion primarily rotates the ψ angle, while the second motion rotates the ψ and ϕ angles in concert.

Table 1 shows the binding free energies, entropies, and mean energy components for the four cryptophane systems. We first computed the binding free energy on each system using the M2 method to validate the molecular mechanical force field parameters and partial charges model employed. For easier comparison, we separated the systems into two groups based on the guest molecules. The relative binding free energies, $\Delta\Delta G$, are in a reasonable range when compared with experiments, although the absolute binding free energies are not as accurate as other reported values.[68,88,89] All complexes are stabilized upon binding primarily by van der Waals interactions, as assessed experimentally with a neutral ligand.[73] However, with the force field utilized for these systems, the cryptophane-ES system overstabilized the guest molecules through van der Waals interaction, resulting in a free energy calculation that did not perfectly match experiments. While the electrostatic interactions between the hosts and guests are significantly favorable, the desolvation penalty negates this impact. Because the free energies of only local minimum conformations are available to the HM method, the free energy barrier to binding cannot be known exactly. However, with multiple constructed binding pathways for a complex, one can analyze the free energies of each conformational minimum along the pathway to rule out unlikely paths that contain unfavorable conformational transitions.

The results of the conformational search of free hosts, cryptophane-E and cryptophane-ES, were analyzed for trends in gating, flexibility, and the degree each host exhibited structural preorganization for binding. Identical conformational search procedures yielded more conformational minima for cryptophane-ES relative to cryptophane-E. However, cryptophane-E had more conformations in the low energy states relative to its global minimum, while cryptophane-ES had a higher population of conformations in the high energy states. For example, cryptophane-E had 30 out of 120 total conformations within $10RT$ (where R is the gas constant and T is the temperature, 300 K), or ∼6 kcal/mol, of its global minimum conformation. Cryptophane-ES had 14 out of 197 conformations within $10RT$. As seen in Figure 3, cryptophane-E displays a distinctly more open cavity structure relative to cryptophane-ES. Cryptophane-E conformations also show a higher propensity for gating arms to be in the "open" position, while the gating arms of cryptophane-ES tend to be in the closed position. Cryptophane-ES does exhibit these "open" conformations, but they are higher energy transition state structures to binding. The closed arm conformations from cryptophane-ES may be due to the larger radius of the sulfur atom than the oxygen atom, resulting in stronger intramolecular interactions between the two arms. The distribution of partial charges is also different in cryptophane-ES and cryptophane-E, which contributes further to the overall shape of the cavity.

Analysis of the Binding Pathways. Upon analysis of the binding pathways, we found that the cryptophane-E host was preorganized for binding and had little structural movement during the course of the binding process. Cryptophane-ES, unexpectedly and on the contrary, underwent large structural changes to accommodate the entry of the ligand to the interior of the host structure. The consistency in the entropic penalties

Table 1. Free Energies and Mean Energies for Cryptophane-Cation Systems[a]

Guest	Host	ΔG_{cal}	$\Delta\Delta G_{cal}$	ΔG_{exp}	$\Delta\Delta G_{exp}$	$-T\Delta S_{conf}$	ΔE	ΔW_{GB}
Me_3NH^+	Cryp. E	-8.5	0	-4.3	0	10.6	-27.5	9.1
	Cryp. ES	-9.3	-0.8	-4.1	0.2	12.2	-43.3	22.3
Me_4N^+	Cryp. E	-9.5	0	-7.6	0	12.4	-30.3	8.8
	Cryp. ES	-12.4	-2.9	-7.3	0.3	11.9	-41.7	17.9

[a]The calculated binding free energy (ΔG_{cal}) and experimental binding free energy (ΔG_{exp}) are presented for comparison, along with relative binding free energies for each ($\Delta\Delta G$). The calculated entropy term ($-T\Delta S_{conf}$), contribution to solvation free energy (ΔW_{GB}), and mean potential energy contributions (ΔE) for electrostatics, van der Waals, and bonded terms. All data are in units of kcal/mol.

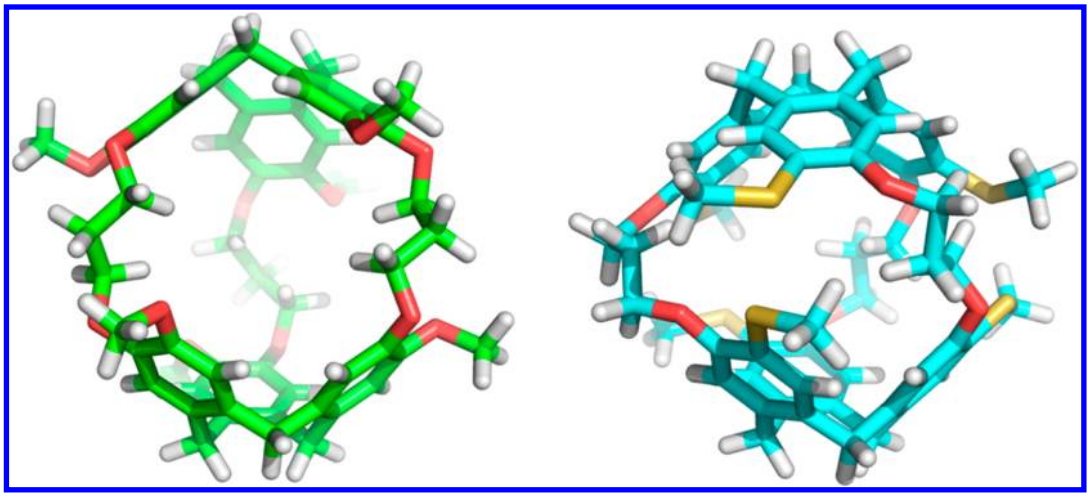


Figure 3. Structural comparison of cryptophane-E and cryptophane-ES. The global energy minimum conformation of cryptophane-E (left, green) has distinctly open gating arms and an open cavity while those of cryptophane-ES (right, cyan) have a more compact structure with closed gating arms.

matched these flexibility trends, staying constant for cryptophane-E and fluctuating for cryptophane-ES before dropping for both systems in the final step of the binding process. Interestingly, the guest molecule quickly lost configurational entropy, mainly from translation/rotation entropy loss while approaching the host surface. However, it did not fully gain favorable intermolecular attractions before reaching the final bound states, resulting in unfavorable free energy intermediates (Figure 4). In the following sections, we discuss the results of the four host−guest systems in more detail. While many high potential energy and high free energy pathways can be sampled for each system, only those with reasonable free energy barriers are discussed and presented.

Trimethylammonium and Cryptophane-E. Two low free energy barrier binding pathways were found for the complex of cryptophane-E with the trimethylammonium guest (Figure 5, Figure S1). Of the four complexes, this system exhibited the lowest free energy transition state and the most favorable enthalpy upon binding. Rather than approaching the host directly from free space, the pathways had ligands that first associated with the surface of the host molecule, then approached the window from the surface. The initial contacts provide intermolecular attractions. However, these interactions are not large enough to overcome the configurational entropy loss. (Figure 4a).

Trimethylammonium and Cryptophane-ES. Three binding pathways were sampled for the complex of cryptophane-ES with trimethylammonium (Figure 6, Supporting Information Figures S2 and S3) The three pathways sampled binding processes through each of the three host windows and approached the binding site from free space. While the conformational search found guests on the host surface, our method did not connect any pathways from these local energy minima. The gating arms were in the closed position when the guest approached the surface; therefore the guest may need to return to free space and rebind to induce gate opening. Positive enthalpies were observed upon association, which dropped dramatically upon binding (Figure 4b). The intermediate state conformations, specifically those conformations with ligands near the gating arms of the host, suffered van der Waals repulsions and/or structural perturbations that resulted in penalties in the calculated enthalpy and free energy.

Tetramethylammonium and Cryptophane-E. One association pathway was determined for cryptophane-E with tetramethylammonium, involving a ligand path to the window of the host starting from the surface of the host molecule (Figure 7). The pathway had a negative enthalpy upon

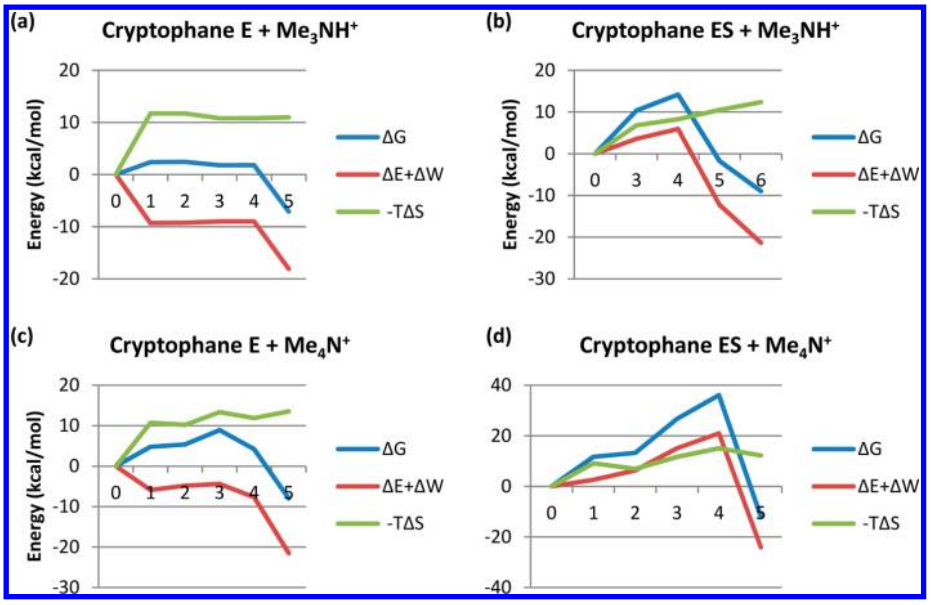


Figure 4. Free energy distributions for pathways developed for the cryptophane−cation systems. Energetic breakdown for the lowest free energy barrier binding pathways developed for the four test systems. The *x* axis represents the intermediate minima state in the binding pathway. All values are calculated relative to a reference energy ($X = 0$), which is calculated as the free energy of the host and ligand when not interacting. The final intermediate step is the bound state.

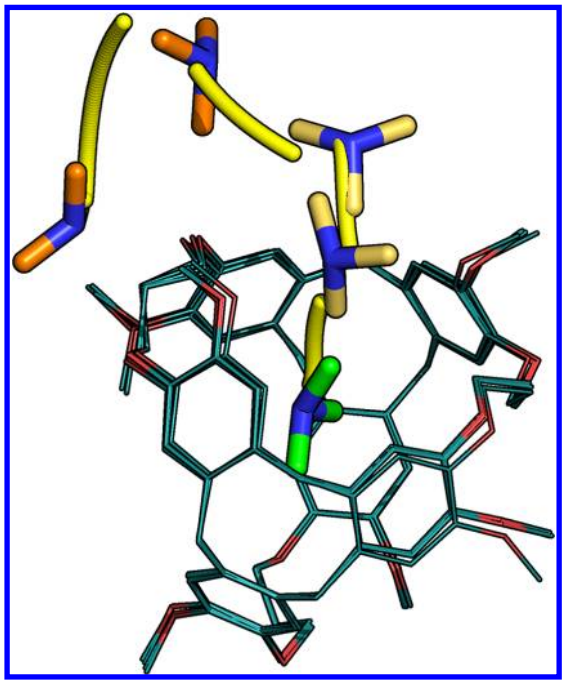


Figure 5. Hopping Minima resulting binding pathway for cryptophane-E with trimethylammonium guest. Four natural motion paths, represented by yellow traces, connect five distinct minimum states, with free energies of 664.2, 664.3, 663.6, 663.6, and 654.7 kcal/mol, in order of decreasing distance from the center of the host.

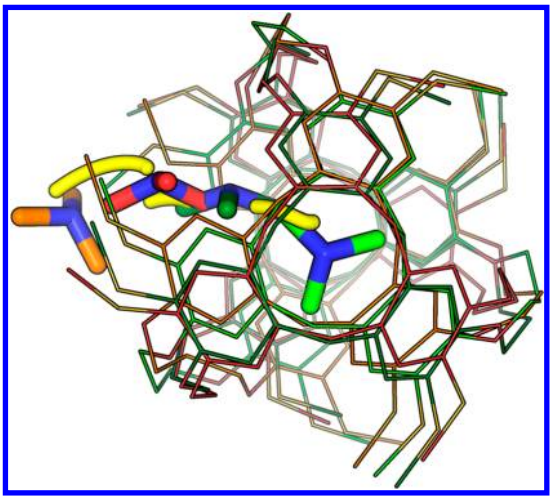


Figure 6. Hopping Minima resulting binding pathway for cryptophane-ES with trimethylammonium guest. Three natural motion paths, represented by yellow traces, connect four distinct minimum states, which are colored for clarity. The orange, red, dark green, and bright green conformations correspond to states with free energies of 652.5, 656.3, 640.4, and 633.2 kcal/mol, respectively.

association with significant entropic compensation (Figure 4c). The last step of the binding process, connecting the tetramethylammonium in the window to the center of the host cavity, was not found. Because the guest has a larger size that limits its flexibility while in the window and final bound state, the trajectories sampled following the normal mode motions show small ranges of vibration. Therefore, the Hopping Minima method could not connect the two minima using the parameters applied for building other paths. However, the final step is not complicated as it only has one direction to move forward; thus it does not affect the illustration of the binding pathway.

Tetramethylammonium and Cryptophane-ES. Two binding pathways were generated for the complex of cryptophane-ES with the tetramethylammonium guest. Both sampled pathways approached the binding site from free space (Figure 8, Supporting Information Figure S4). Analysis of the minima involved in these paths revealed free energy values for the intermediate state conformations that deviate from that expected based on experimental free energy barrier values. The results suggest that correctly computing energies for intermediate states may require more accurate force field

dx.doi.org/10.1021/ct301023m | J. Chem. Theory Comput. 2013, 9, 2010−2019

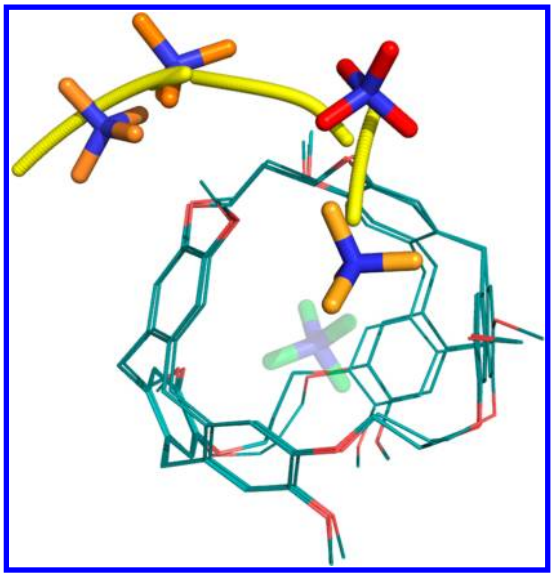


Figure 7. Hopping Minima resulting association pathway for cryptophane-E with tetramethylammonium guest. Three natural motion paths, represented by yellow traces, connect four distinct minimum states, with free energies of 698.5, 699.1, 702.6, and 697.9 kcal/mol, in order of decreasing distance from the host window. The lowest free energy bound state ligand conformation, with a value of 685.8 kcal/mol, is included to visualize free energy trends.

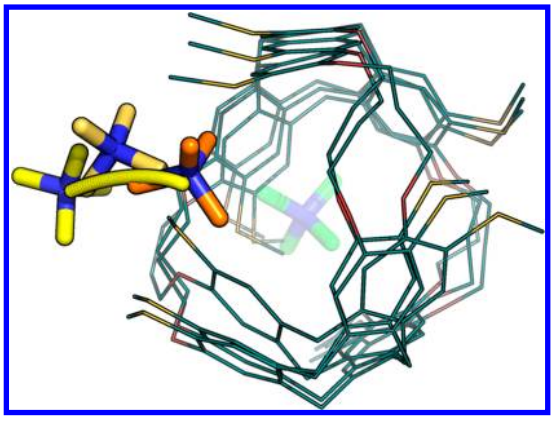


Figure 8. Hopping Minima resulting association pathway for cryptophane-ES with tetramethylammonium guest. A single natural motion path, represented by a yellow trace, connects three distinct minimum states, with free energies of 687.66, 695.99, and 701.74 kcal/mol, in order of decreasing distance from the host window. The lowest free energy bound state ligand conformation is included to visualize free energy trends.

parameters than those in the free and bound conformations. Our results again show that positive changes in enthalpies and free energies for the intermediate conformations were due to steric clashes resulting in van der Waals repulsion and unfavorable structural perturbation. High repulsion (>40 kcal/mol) was observed in several minima hopping paths that approached the host window. The repulsive potentials were more pronounced in this system than in cryptophane-ES with trimethylammonium due to the increased size of the ligand in this system.

GENERAL COMMENTS

The alanine dipeptide test system demonstrated the efficacy and speed of the HM method. After parametrization of the system, only a few minutes were required to carry out all required calculations, and multiple transitions between several different conformations were found. Thus, the HM method may be an attractive and competitive option for quickly exploring conformational transitions in single molecules without the need for time-consuming dynamics or the need to specify starting and ending conformational states.

The results of modeling binding pathways of the cryptophane-guest systems are promising, provide insight into molecular recognition, and show potential for use in a wide range of applications, such as designing drug carriers and inhibitors with preferred binding kinetics. Actual guest binding processes can take milliseconds to seconds; therefore, applying classical molecular dynamics simulations to sample binding pathways can be very expensive computationally. As with all methods that rely on molecular mechanical force fields, our method is dependent on the force field for its accuracy. Although we validated our force field parameters by reproducing experimental binding affinities using the M2 free energy calculation method, the free energy profile found from Hopping Minima, particularly for the cryptophane-ES systems, show the sensitivity of the method to the choice of force field parameters. Our models reveal that when a guest is passing the window of a host, the intermediate states have very small spatial restrictions, and small errors in force field parameters may result in a local minimum with unrealistically large computed free energy.

While the AMBER FF03 molecular mechanical force field, DREIDING force field, and VC/2004 partial charge models were utilized for the series of test systems, any molecular mechanical model can be employed.[90−94] Other implicit solvent models, in addition to the generalized Born model, can be used as well.

Thorough conformational search is a key to successfully finding binding pathways and building binding free energy landscapes. This study used the Tork conformational search tool, which has advantages in efficiently finding conformations for chemical hosts with macrocyclic rings. Nevertheless, any conformational search method used together with a duplicate conformation filtering algorithm that accounts for symmetry can be used. For example, other conformational search methods, such as the basin-hopping approach or commercial packages, also can be used.[95−97] The method can also be scaled up, with regards to system size. For larger molecular systems, such as protein−ligand complexes, significant speed increases can be obtained by fixing regions of the protein that are not critical to ligand binding while allowing the rest of the atoms to be flexible.

For the cryptophane host−guest systems, our experiments were set up such that the RMSD intersection tests between conformational minima and natural motions used only the distance between the center nitrogen atom of the ligands. This measurement works efficiently for small molecules with symmetry, such as the guests in this work. For more complicated ligand structures, more atoms should be included in the RMSD selection. Additionally, host atoms could be added to the selection to capture host conformational transitions that occur during ligand translation. Because the Hopping Minima method is based on connecting distinct local energy minima, occasionally binding processes from the free to intermediate bound states are sampled, but a complete binding pathway to the final bound state could not be found. Molecular dynamics simulations could be employed to sample the final

binding event for the route to bridge the gap. Molecular dynamics or existing methods for transition path samplings, such as the RPHSA, string, and TPS methods, can be further applied to the distinct local energy minima found in a binding pathway to smooth the transitions between each energy state. If one is interested in applying steered or targeted molecular dynamics in explicit solvent, the binding pathway could be used as a physically based reaction coordinate, thus avoiding the potential for introducing unnatural bias to the simulation.

CONCLUSIONS

The presented method allows for the computational modeling of molecular transitions as well as binding pathways for host−guest systems. Investigating binding pathways helps elucidate fundamental mechanisms, including allostery, conformational changes, and gated control associations, and can guide molecular design. Our study also illustrated that the gating effects and molecular distortion contribute mainly to slow association rate in guest binding to the cryptophane-ES system. Our analysis reveals the balance or imbalance between enthalpy and configurational entropy changes during a ligand binding process. The method relies on multiple filters and analysis procedures and a thorough minimum conformational search of a host−guest complex. The method has a distinct computational cost advantage over existing methods. Calculations performed on the alanine dipeptide provided evidence of the efficacy of the method to capture conformational transitions. Using four cryptophane complexes as example systems, the method demonstrated the effectiveness of sampling multiple binding pathways for each system and deepened our understanding of cryptophane-guest associations.

ASSOCIATED CONTENT

Supporting Information

Additional information regarding paths constructed for the host−guest systems presented in this article. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Tel.: (951) 827-7263. E-mail: chiaenc@ucr.edu.

Author Contributions

This study was performed through the contributions of both authors equally. The manuscript was written through contributions of both authors and both have given approval to the final version of the manuscript.

Funding Sources

This work was supported by grants from the National Science Foundation (MCB-0919586).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Myungshim Kang for helpful discussions, Rizi Ai for preparing input files, and Verachem for the use of software.

ABBREVIATIONS

HM, Hopping Minima; RMSD, root-mean-square deviation; MD, molecular dynamics; MC, Monte Carlo; PMF, potential of mean force; ABF, adaptive biasing force; RPHSA, reaction path Hamiltonian superposition approach; FES, free energy surface; M2, Mining Minima generation 2; BAT, Bond-Angle-Torsion

REFERENCES

(1) Lauffenburger, D. A.; Linderman, J. *Receptors: Models for Binding, Trafficking, and Signaling*; Oxford University Press: Oxford, England, 1996; p 376.
(2) Janin, J. *Proteins* **1997**, *28*, 153−161.
(3) Swinney, D. C. *Label-Free Technologies for Drug Discovery*; Wiley: New York, 2011; pp 284−302.
(4) Schreiber, G.; Haran, G.; Zhou, H.-X. *Chem. Rev.* **2009**, *109*, 839−860.
(5) Markgren, P. O.; Lindgren, M. T.; Gertow, K.; Karlsson, R.; Hämäläinen, M.; Danielson, U. H. *Anal. Biochem.* **2001**, *291*, 207−218.
(6) Dong, B.; Zhou, Q.; Zhao, J.; Zhou, A.; Harty, R. N.; Bose, S.; Banerjee, A.; Slee, R.; Guenther, J.; Williams, B. R. G.; Wiedmer, T.; Sims, P. J.; Silverman, R. H. *J. Virol.* **2007**, *78*, 8983−8993.
(7) Helgeson, R. C.; Hayden, A. E.; Houk, K. N. *J. Org. Chem.* **2010**, *75*, 570−575.
(8) Rieth, S.; Hermann, K.; Wang, B.-Y.; Badjić, J. D. *Chem. Soc. Rev.* **2010**, *40*, 1609−1622.
(9) Berry, R. S.; Rice, S. A.; Ross, J. *J. Phys. Chem.*; Oxford University Press: New York, 2000; p 1080.
(10) Dill, K. A.; Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*; Garland Science: New York, 2010; p 720.
(11) Wales, D. *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*; Cambridge University Press: Cambridge, England, 2003; p 681.
(12) Schlosshauer, M.; Baker, D. *J. Phys. Chem. B* **2002**, *106*, 12079−12083.
(13) Cai, L.; Zhou, H.-X. *J. Chem. Phys.* **2011**, *134*, 105101.
(14) Northrup, S. H.; Zarrin, F.; McCammon, J. A. *J. Phys. Chem.* **1982**, *86*, 2314−2321.
(15) Simunovic, M.; Zagrovic, B.; Tomić, S. *J. Mol. Recognit.* **2011**, *24*, 854−861.
(16) Vashisth, H.; Abrams, C. F. *Biophys. J.* **2008**, *95*, 4193−4204.
(17) Zuckerman, D. *Statistical Physics of Biomolecules: An Introduction*; CRC Press: Boca Raton, FL, 2010.
(18) Held, M.; Noé, F. *Eur. J. Cell Biol.* **2012**, *91*, 357−364.
(19) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. *J. Am. Chem. Soc.* **2009**, *131*, 11811−11818.
(20) Buch, I.; Giorgino, T.; De Fabritiis, G. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10184−10189.
(21) Ahmad, M.; Gu, W.; Helms, V. *Angew.Chem., Int. Ed.* **2008**, *47*, 7626−7630.
(22) Huang, D.; Caflisch, A. *PLoS Comput. Biol.* **2011**, *7*, e1002002.
(23) Ramadugu, S. K.; Kashyap, H. K.; Ghirlanda, G.; Margulis, C. *Glycobiology* **2011**, *21*, 1477−1478.
(24) Peters, B. *Mol. Simul.* **2010**, *36*, 1265−1281.
(25) Elber, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 151−156.
(26) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13118−13123.
(27) Shaw, D. E.; Bowers, K. J.; Chow, E.; Eastwood, M. P.; Ierardi, D. J.; Klepeis, J. L.; Kuskin, J. S.; Larson, R. H.; Lindorff-Larsen, K.; Maragakis, P.; Moraes, M. A.; Dror, R. O.; Piana, S.; Shan, Y.; Towles, B.; Salmon, J. K.; Grossman, J. P.; Mackenzie, K. M.; Bank, J. A.; Young, C.; Deneroff, M. M.; Batson, B. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis - SC '09*; ACM Press: New York, 2009; p 1.
(28) Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. . *Am. Chem. Soc.* **2011**, *133*, 9181−9183.
(29) Durrant, J. D.; McCammon, J. A. *BMC Biol.* **2011**, *9*, 71.
(30) Pal, S.; Fichthorn, K. A. *Chem. Eng. J.* **1999**, *74*, 77−83.
(31) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919−11929.
(32) Chang, C.-E. A.; Trylska, J.; Tozzini, V.; McCammon, J. A. *Chem. Biol. Drug Des.* **2007**, *69*, 5−13.

(33) Spaar, A.; Dammer, C.; Gabdoulline, R. R.; Wade, R. C.; Helms, V. *Biophys. J.* **2006**, *90*, 1913−1924.
(34) Frembgen-Kesner, T.; Elcock, A. H. *Biophys. J.* **2010**, *99*, L75−7.
(35) Kim, Y. C.; Hummer, G. *J. Mol. Biol.* **2008**, *375*, 1416−1433.
(36) Mereghetti, P.; Gabdoulline, R. R.; Wade, R. C. *Biophys. J.* **2010**, *99*, 3782−3791.
(37) Huber, G. A.; McCammon, J. A. *Comput. Phys. Commun.* **2010**, *181*, 1896−1905.
(38) Wang, J.-C.; Pal, S.; Fichthorn, K. *Phys. Rev. B* **2001**, *63*, 085403.
(39) Carmichael, S. P.; Shell, M. S. *J. Phys. Chem. B* **2012**, *116*, 8383−8393.
(40) Zheng, W. *J. Chem. Phys.* **2012**, *136*, 155103.
(41) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(42) Gan, W.; Roux, B. *Proteins* **2009**, *74*, 996−1007.
(43) Wang, J.; Zhang, K.; Lu, H.; Wang, E. *Biophys. J.* **2006**, *91*, 866−872.
(44) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559.
(45) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.
(46) Elenewski, J. E.; Hackett, J. C. *J. Phys. Chem. B* **2010**, *114*, 11315−11322.
(47) Zhang, D.; Gullingsrud, J.; McCammon, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 3019−3026.
(48) Isralewitz, B.; Gao, M.; Schulten, K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224−230.
(49) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(50) Chipot, C.; Hénin, J. *J. Chem. Phys.* **2005**, *123*, 244906.
(51) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234−2246.
(52) Lettieri, S.; Zuckerman, D. M. *J. Comput. Chem.* **2012**, *33*, 268−275.
(53) Lettieri, S.; Mamonov, A. B.; Zuckerman, D. M. *J. Comput. Chem.* **2011**, *32*, 1135−1143.
(54) Lv, C.; Zheng, L.; Yang, W. *J. Chem. Phys.* **2012**, *136*, 044103.
(55) Zheng, L.; Yang, W. *J. Chem. Theory Comput.* **2012**, *8*, 810−823.
(56) Ren, W.; Vanden-Eijnden, E.; Maragakis, P. *J. Chem. Phys.* **2005**, *123*, 134109.
(57) Ren, W.; Vanden-Eijnden, E. *J. Chem. Phys.* **2007**, *126*, 164103.
(58) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2007**, *446*, 182−190.
(59) Branduardi, D.; Gervasio, F. L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.
(60) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9978.
(61) Bolhuis, P.; Chandler, D.; Dellago, C.; Geissler, P. *Annu. Rev. Phys. Chem.* **2002**, 53.
(62) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. *Adv. Phys.* **2002**, *123*, 1−78.
(63) Fischer, S.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 252−261.
(64) Maragliano, L.; Vanden-Eijnden, E.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 2589−2594.
(65) Noé, F.; Fischer, S. *Curr. Opin. Struct. Biol.* **2008**, *18*, 154−162.
(66) Pan, A. C.; Sezer, D.; Roux, B. *J. Phys. Chem.* **2008**, *112*, 3432−3440.
(67) Strodel, B.; Wales, D. J. *Chem. Phys. Lett.* **2008**, *466*, 105−115.
(68) Chang, C.-E.; Gilson, M. K. *J. Am. Chem. Soc.* **2004**, *126*, 13156−13164.
(69) Chang, C.-E.; Gilson, M. K. *J. Comput. Chem.* **2003**, *24*, 1987−1998.
(70) Chang, C.-E.; Potter, M. J.; Gilson, M. K. *J. Phys. Chem. B* **2003**, *107*, 1048−1055.
(71) Isogai, Y. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 414−418.
(72) Garcia, C.; Humilière, D.; Riva, N.; Collet, A.; Dutasta, J.-P. *Org. Biomol. Chem.* **2003**, *1*, 2207.
(73) Canceill, J.; Cesario, M.; Collet, A.; Guilhem, J.; Lacombe, L.; Lozach, B.; Pascard, C. *Angew. Chem., Int. Ed.* **1989**, *28*, 1246−1248.
(74) Holman, K. T. *Encyclopedia of Supramolecular Chemistry*; Dekker: New York, 2004; pp 340−348.
(75) Spence, M. M.; Rubin, S. M.; Dimitrov, I. E.; Ruiz, E. J.; Wemmer, D. E.; Pines, A.; Yao, S. Q.; Tian, F.; Schultz, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10654−10657.
(76) Aaron, J. A.; Chambers, J. M.; Jude, K. M.; Di Costanzo, L.; Dmochowski, I. J.; Christianson, D. W. *J. Am. Chem. Soc.* **2008**, *130*, 6942−6943.
(77) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545−600.
(78) Liu, P.; Agrafiotis, D. K.; Theobald, D. L. *J. Comput. Chem.* **2010**, *31*, 1561−1563.
(79) Chen, W.; Huang, J.; Gilson, M. K. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1301−1313.
(80) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(81) Allen, F. H. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2002**, *58*, 380−388.
(82) Avogadro: an open-source molecular builder and visualization tool. http://avogadro.openmolecules.net/ (accessed Sep 20, 2012).
(83) Kairys, V.; Gilson, M. K. *J. Comput. Chem.* **2002**, *23*, 1656−1670.
(84) Gilson, M. K.; Gilson, H. S. R.; Potter, M. J. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1982−1997.
(85) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897−8909.
(86) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.
(87) Luo, R.; Head, M. S.; Given, J. A.; Gilson, M. K. *Biophys. Chem.* **1999**, *78*, 183−193.
(88) Moghaddam, S.; Yang, C.; Rekharsky, M.; Ko, Y. H.; Kim, K.; Inoue, Y.; Gilson, M. K. *J. Am. Chem. Soc.* **2011**, *133*, 3570−3581.
(89) Chen, W.; Chang, C.-E.; Gilson, M. K. *Biophys. J.* **2004**, *87*, 3035−3049.
(90) Muddana, H. S.; Gilson, M. K. *J. Chem. Theory Comput.* **2012**, *8*, 2023−2033.
(91) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(92) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(93) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671−690.
(94) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(95) Kolossváry, I.; Guida, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 5011−5019.
(96) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111−5116.
(97) Wales, D. J. *Science* **1999**, *285*, 1368−1372.