

# Factor Analysis of Conformations and NMR Signals of Rotaxanes: AIMD and Polarizable MD Simulations

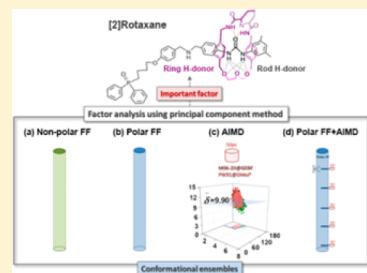
Pingying Liu,<sup>†,‡</sup> Wei Li,<sup>\*,†</sup> Zigui Kan,<sup>†</sup> Hui Sun,<sup>†</sup> and Jing Ma<sup>\*,†</sup>

<sup>†</sup>School of Chemistry and Chemical Engineering, Key Laboratory of Mesoscopic Chemistry of MOE, Nanjing University, Nanjing 210093, People's Republic of China

<sup>‡</sup>School of Materials Science and Engineering, Jingdezhen Ceramic Institute, Jingdezhen 333403, People's Republic of China

## Supporting Information

**ABSTRACT:** The interlocked  $\langle$ rod | ring $\rangle$  structures of pseudorotaxanes and [2]rotaxanes are usually maintained by the complex hydrogen-bonding (H-bonding) network between the rod and ring. Ab initio molecular dynamics (AIMD) using generalized energy-based fragmentation approach and polarizable force field (polar FF)-based molecular dynamics (MD) simulations were performed to investigate the conformational changes of mechanically interlocked systems and to obtain the ensemble-averaged NMR chemical shifts. Factor analysis (FA) demonstrates that the ring H-donor (2,6 pyridinedicarboxamide group) plays an important role in the ring–rod recognition. In comparison to the conventional fixed-charge force field, the polarization effect is crucial to account for the H-bonding interactions in supramolecular systems. In the hybrid scheme, the polar FF-based MD simulations are used to generate different initial states for the AIMD simulations, which are able to give better prediction of ensemble-averaged NMR signals for chemically equivalent amide protons. The magnitude of the deshielding shift of NMR signal is correlated with the length of hydrogen bond. The polar FF model with variable charges shows that the dipole–dipole interactions between the flexible diethylene glycol chain of ring and polar solvents induce the upfield shifts of NMR signals of rod H-donors and the directional distribution of the neighboring  $\text{CH}_3\text{CN}$  solvents.



## 1. INTRODUCTION

The mechanically interlocked molecules, such as rotaxanes, have potential applications in building molecular devices and nanoscale machines. A [2]rotaxane is composed of an axle-shaped molecule (the “rod”) surrounded by a macrocyclic compound (the “ring”). The rod is usually terminated by bulky groups (called “stoppers”) to prevent disassembly of the interlocked structures. The two-component [2]rotaxane is also named as a  $\langle$ rod | ring $\rangle$  system. In the preparation of rotaxane, the precursor is a pseudorotaxane without bulky terminals or with only one stopper (Figure 1).

Noncovalent interactions between the components, such as H-bonding, electrostatic, cation- $\pi$ , van der Waals, and hydrophobic interactions, are the driving forces for the efficient synthesis and assembly of rotaxanes.<sup>1–3</sup> Some different spectroscopic techniques are available for the experimental detection of intermolecular interactions.<sup>4–6</sup> Among them, nuclear magnetic resonance (NMR) spectroscopy is one of the useful methods for detecting the interplay of H-bonding interactions with the structures of supramolecular systems.<sup>7–9</sup> It has been revealed that the magnitude of the shift of NMR signal is directly correlated with the distance of hydrogen bond (HB).<sup>10–12</sup>

However, NMR chemical shift is extremely sensitive to conformational changes in different chemical environments. With an increase in the conformational freedom, the unambiguous assignment of NMR signals becomes very difficult. The ensemble-averaged NMR signals of amide protons sampling

from Car–Parrinello molecular dynamics simulations for N-methylacetamide are in good agreement with the experimental chemical shifts.<sup>13</sup> Being limited by the expensive computational cost, ab initio molecular dynamics (AIMD) simulations were mostly performed on small model systems.

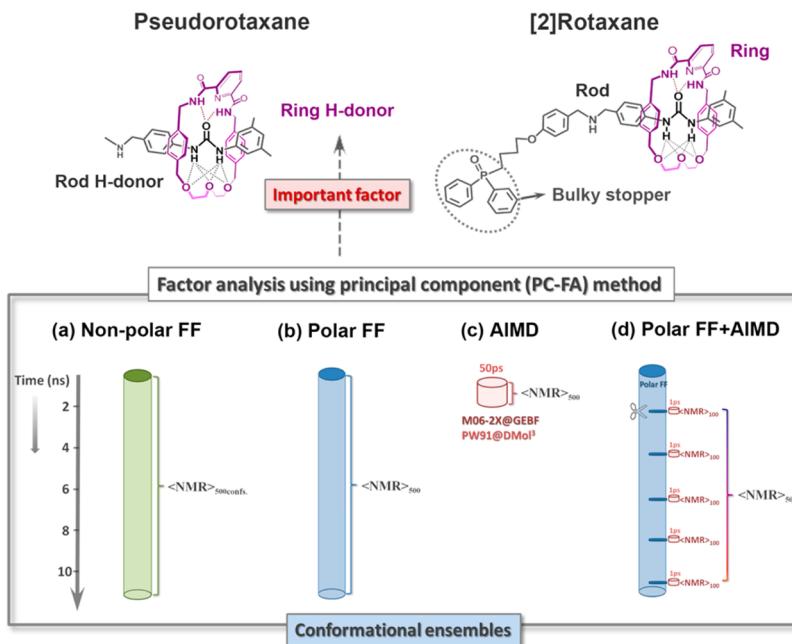
Recently, the fragment-based methods have been used to treat large-sized systems. Developments of fragment-based methods have been introduced in recent reviews.<sup>14–16</sup> We have employed generalized energy-based fragmentation (GEBF) approach<sup>17–20</sup> to study the geometry and binding strength of [2]rotaxane at the density functional theory (DFT)/M06-2X and second-order Møller–Plesset perturbation theory (MP2) levels, respectively.<sup>21</sup> In the GEBF approach, the ground-state energy of a large molecule is assembled from the corresponding energies of overlapping subsystems, each of which is embedded in the background charges of the atoms outside of the subsystems.<sup>17–19</sup> The GEBF-based AIMD method was also implemented to investigate dynamic behaviors of large systems on the order of tens of picoseconds, in which the forces on nuclei were obtained by GEBF-based quantum mechanics (QM) calculations.<sup>20,22</sup> In the present work, GEBF-based AIMD simulations will be applied to collect the conformational ensemble of pseudorotaxane for the NMR calculations.

Received: October 15, 2015

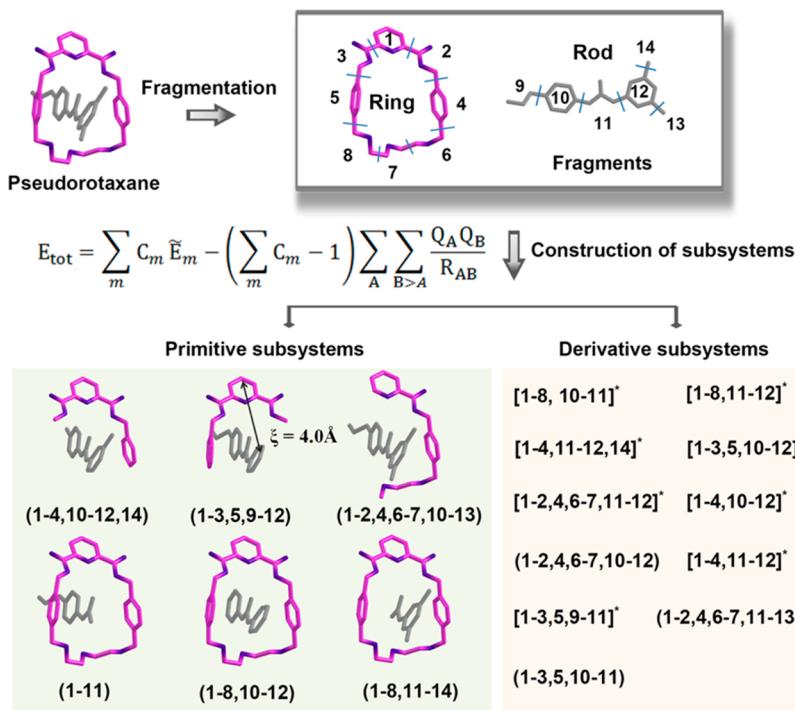
Revised: January 12, 2016

Published: January 12, 2016





**Figure 1.** Schematic illustration of pseudorotaxane and [2]rotaxane and different theoretical models for collecting the conformational ensembles, including (a) nonpolarizable force field (nonpolar FF) model, (b) polarizable force field (polar FF) model, (c) ab initio MD (AIMD), and (d) polar FF +AIMD simulations, respectively.



**Figure 2.** Fragmentation scheme and the resultant fragments and subsystems of pseudorotaxane.  $C_m$  is the coefficient of the  $m$ th embedded subsystem. Each overlapping subsystem, which is denoted by the square bracket and the asterisk (e.g.,  $[1-8,10-44]^*$ ), has the coefficient of  $C_m = -1$ . All the other subsystems are given the coefficient of  $C_m = 1$  and are represented by the parentheses such as  $(1-4,10-12,14)$ .

Force field-based molecular dynamics (FF-based MD) simulations allow the search of much larger conformational spaces due to their cheaper computational costs compared to AIMD ones. To accurately capture the physics of hydrogen bonds, force field incorporating electronic polarizability is necessary in some cases. Efforts have been devoted to treating the polarization effect by developing various polarizable force fields (labeled as polar FF).<sup>23–28</sup> There are many successful

applications of polarizable force fields, such as induced dipole model,<sup>29–31</sup> Drude-oscillators model,<sup>32–35</sup> and fluctuating charge model.<sup>36,37</sup> Alternatively, we can alter the charge distribution through the first principle calculations of partial charges and dipoles “on-the-fly” in response to the fluctuation of chemical environment.<sup>28,38–41</sup> Until now, there are few literature reports on the investigation of polarization effect in supramolecular systems. Our QM-based polarizable model is hence

used to study the conformational changes of pseudorotaxane and [2]rotaxane in both vacuum and mixed  $\text{CH}_3\text{CN}-\text{CHCl}_3$  solution (Figure 1).

The trajectories collected from MD simulations usually consist of very large data sets. There has been an increasing interest to develop methods to extract the essential information from a series of snapshots over the simulation time. Principal component analysis (PCA) has been employed to elucidate the representative averaging structures and the dominant deformations and to extract the important information on biomolecular processes such as protein folding or molecular recognition from MD trajectories.<sup>42–53</sup> Furthermore, factor analysis (FA) is a generalization of PCA, being used to explain the covariances or correlations between the variables. Factor analysis using principal component (PC-FA) method is a simple and non-parametric method of extracting relevant information from huge data sets and reducing complex data sets to lower dimensions. PC-FA technique has been widely used in various fields, such as NMR spectroscopy,<sup>54,55</sup> the interactions of acid–base and metal ions with the macromolecule in solution,<sup>56</sup> prediction of material properties,<sup>57</sup> etc.

In the present work, PC-FA method is applied to the analysis of the AIMD and FF-based MD trajectory data of pseudorotaxane and [2]rotaxane in a simple way. The attempt at applying PC-FA technique to the data mining of MD conformational ensembles may deepen our understanding of the relationship between the H-bonding interactions and NMR chemical shifts for supramolecular systems. The insight gained in this work may have important implications for the molecular design, synthesis, and construction of interesting molecular interlocked architectures in the future.

## 2. COMPUTATIONAL METHODS

**2.1. Ab Initio Molecular Dynamics Simulations: M06-2X@GEBF-AIMD.** Within the Born–Oppenheimer approximation, the AIMD forces acting on the nuclei are estimated via directly solving the quantum electronic Schrödinger equations. Here, we adopted GEBF-based AIMD<sup>20,22</sup> approach, implemented in Low Scaling Quantum Chemistry (LSQC)<sup>58,59</sup> program, to study conformational dynamics of supramolecular systems.

The central idea of the GEBF method is to decompose a large system into small fragments and to obtain energy and properties of the whole molecule from a combination of the corresponding electrostatically embedded subsystems.<sup>17–20</sup> Figure 2 demonstrates the flowchart of GEBF method applied on the pseudorotaxane. The total energy of the target system can be evaluated with the expression

$$E_{\text{tot}} = \sum_m^M C_m \tilde{E}_m - \left( \sum_m^M C_m - 1 \right) \sum_A \sum_{B>A} \frac{Q_A Q_B}{R_{AB}} \quad (1)$$

where  $M$  is the number of subsystems, and  $C_m$  and  $\tilde{E}_m$  are the coefficient and energy (including the self-energy of point charges) of the  $m$ th embedded subsystem, respectively.  $Q_A$  and  $Q_B$  are the point charges of atoms A and B, respectively, and  $R_{AB}$  is the distance between atoms A and B. As shown in Figure 2, for each fragment, a primitive subsystem centered on this fragment can be formed by adding environmental fragments within a given distance threshold ( $\xi = 4.0 \text{ \AA}$ ). To eliminate the overcounting of some multifragment (or one-fragment) terms, one should construct a series of smaller subsystems (called derivative subsystems). Figure 2 presents all the fragments and subsystems

of the pseudorotaxane. In this work, the energies and analytic energy gradients of subsystems are obtained with M06-2X functional.

To speed up the GEBF energy gradients calculations in each MD step, we decomposed the pseudorotaxane into 14 fragments, that is, eight ones for ring and six ones for rod. Each six-membered aromatic ring is defined as an isolated fragment. The urea group of rod, that is, binding site, is also set as a fragment. Eleven derivative subsystems were constructed. The single point energy difference of the pseudorotaxane between GEBF-M06-2X approach and conventional M06-2X method is 0.19 kcal/mol with the 6-31G (d) basis set. The GEBF-based AIMD calculation on pseudorotaxane was implemented at the same level.

AIMD simulations were performed within the canonical (NVT) ensemble. The system temperature was kept around 298.15 K by using a Nosé–Hoover chain of thermostat. The trajectories were propagated with the modified Beeman algorithm with a time step of 1 fs. A 10 ps MD run was then performed to collect the trajectory after the equilibrium stage was achieved.

Our previous study of the performance of six different functionals (B3LYP, CAM-B3LYP, PBEPBE, mPW1PW91, M06-2X, and  $\omega$ B97X-D) in treating interlocked [2]rotaxane indicated that M06-2X binding energies were close to the corresponding GEBF-MP2 results.<sup>21</sup> For comparison, the Born–Oppenheimer molecular dynamics (BOMD) simulation furnished in Dmol<sup>3</sup> module of Material Studio software<sup>60</sup> was also performed on pseudorotaxane using the Perdew–Wang (PW91) generalized gradient approximation (GGA). Double numerical basis sets including polarization (DNP) functions were used in the calculations. The total simulation time is 50 ps. As expected, PW91@DMol<sup>3</sup> cannot describe the H-bonding interactions very well, with the results shown in Figure S1 and Figure S2 of Supporting Information. So we mainly use the GEBF-M06-2X-based AIMD to test the performance of the classical MD simulations based on the fixed charge or polarizable force fields.

**2.2. Force Field-Based Molecular Dynamics Simulations.** The force field-based MD simulations were performed by using the Discover module of Materials Studio package.<sup>60</sup>

**Fixed-Charge Force Field.** In the force-field method, the electrostatic interaction is usually described in terms of Coulomb interaction between the charged centers  $q_i$  and  $q_j$

$$E_{\text{electrostatic}} = \sum_{i>j} \frac{q_i q_j}{\epsilon r_{ij}} \quad (2)$$

where  $q_i$  and  $q_j$  are the atomic partial charges, and  $r_{ij}$  is the distance between the  $i$ th and  $j$ th atoms. In conventional polymer consistent force field (PCFF), the predefined atomic partial charges were fixed during the geometry optimization and MD simulation. Thus, the fixed charge model, called nonpolar FF, fails to describe the variations of electron density in a variable chemical environment.

**Polarizable Force Field.** In this work, the QM calculation is introduced to update the partial charge of each atom in the interlocked systems along the MD trajectory. The atomic charges are obtained from DFT calculations and then taken as the input of electrostatics parameters (eq 2) for the subsequent MD simulation.<sup>28,38–41</sup>

To keep the energy conservation along the QM-based MD simulations, the atomic charges of the target molecule would be renewed by the natural population analysis (NPA) charges at M06-2X/6-31G (d, p) level if the relative deviation of total

energy between the two successive MD steps is below 10%. If not, the partial charges would be fixed in the next time step. The details are shown in Figure S3 (schematic illustration of procedure) and Figure S4 (time evolution of potential energy) of Supporting Information. The other MD simulation setups are almost same as the nonpolarizable MD simulation ones. The charge-variable polar FF was implemented in the MD simulations for both pseudorotaxane and [2]rotaxane in vacuum and in mixed CH<sub>3</sub>CN–CHCl<sub>3</sub> solvents, respectively. From those simulations, the polarization effects are found to be crucial to account for H-bonding interactions in pseudorotaxane and [2]rotaxane.

**Polar Force Field and ab Initio Molecular Dynamics Scheme.** The combination of AIMD and polarizable FF was also implemented in this work. Because of the expensive computational cost of AIMD, it is difficult to run long-time AIMD simulations. It is possible to use less expensive polar FF-based MD simulations to yield different initial states for the AIMD simulations, which is labeled as polar FF+AIMD (Figure 1d). In the polar FF+AIMD scheme, several parallel short-time AIMD simulations are performed with the polar FF MD snapshots taken as the initial configurations. In the present work, five parallel AIMD runs were performed for 1 ps. To equilibrate the systems, a 200 fs AIMD run was performed. 100 snapshots were taken from 1 ps AIMD simulation time. A total of 500 snapshots were selected for ensemble-averaged NMR signals. The polar FF+AIMD MD scheme gives better prediction of ensemble-averaged NMR signals for chemically equivalent amide protons in the ring of pseudorotaxane. Our purpose is just to demonstrate the feasibility of the polar FF+AIMD method. One can use longer AIMD simulations to obtain more reliable results.

**2.3. Factor Analysis Using Principal Component Method.** Principle component (PC) method creates variables that are linear combinations of the original variables. A vector of observations for the *i*th subject reads

$$X_i = \begin{pmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{ip} \end{pmatrix} \quad (3)$$

$S$  is the variance–covariance matrix, which is as expressed as

$$S = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})(X_i - \bar{x})' \quad (4)$$

where  $\bar{x}$  is the mean value of the set  $X$ . We will have  $p$  eigenvalues ( $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p$ ) and corresponding eigenvectors ( $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_p$ ) for this  $S$  matrix.

Then the variance-covariance matrix  $S$  can be re-expressed in the following form as a function of the eigenvalues and the eigenvectors

$$\begin{aligned} S &= \sum_{i=1}^p \lambda_i e_i e_i' \\ &\cong \sum_{i=1}^m \lambda_i e_i e_i' \\ &= (\sqrt{\lambda_1} e_1 \sqrt{\lambda_2} e_2 \dots \sqrt{\lambda_m} e_m) \begin{pmatrix} \sqrt{\lambda_1} e_1' \\ \sqrt{\lambda_2} e_2' \\ \vdots \\ \sqrt{\lambda_m} e_m' \end{pmatrix} \\ &= LL' \end{aligned} \quad (5)$$

The idea of the PC method is to approximate the above expression. Instead of summing from 1 to  $p$ , now we would sum it from 1 to  $m$  ( $m < p$ ), ignoring the last  $p-m$  terms in the sum. The factor loadings are calculated by

$$\hat{l}_{ij} = \hat{e}_{ij} \sqrt{\hat{\lambda}_i} \quad (6)$$

This will form the factor loading matrix  $L$  in the factor analysis (FA). According to the factor loadings, the principal component scores can be obtained.

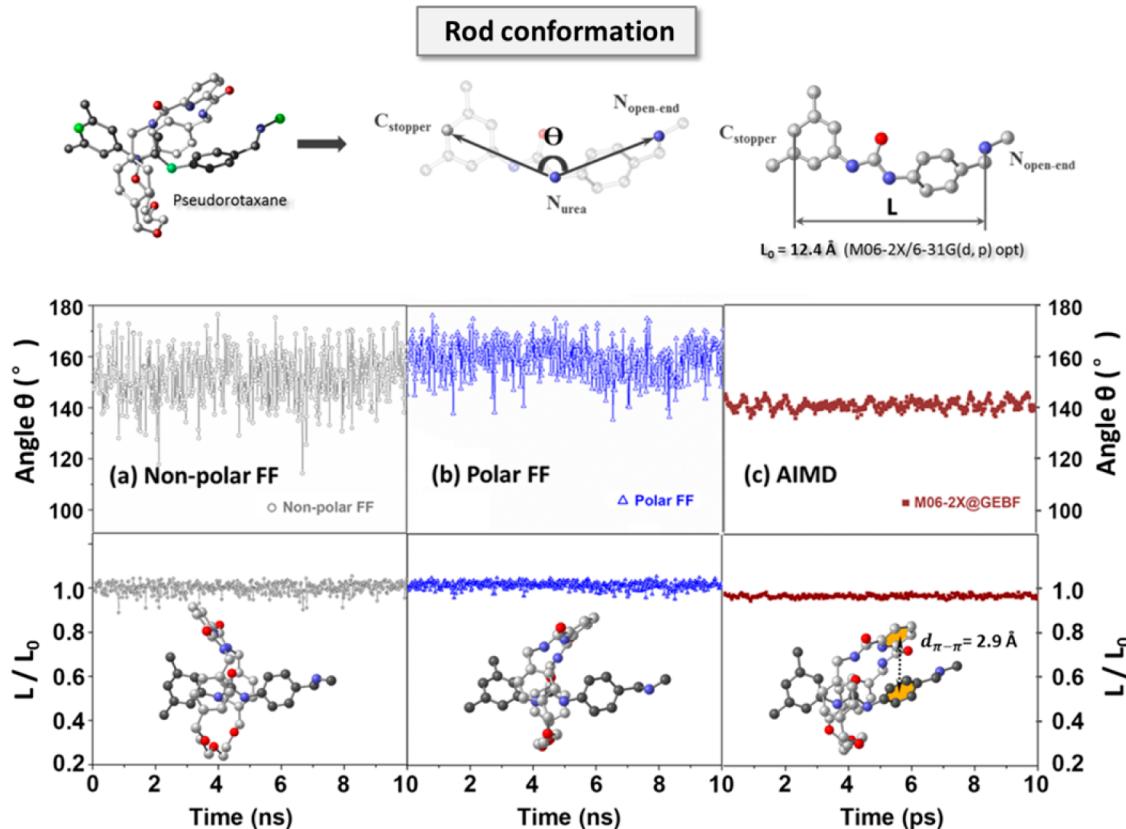
The first principal component (PC1) that accounts for most of the variance in the original variables is constructed from the eigenvector corresponding to the largest eigenvalue. In the present work, the PC1 score of ring-rod contacting geometry is used to evaluate the importance of the factors responsible for ring-rod interaction. The analysis of the data was performed using the SPSS software.<sup>61</sup>

**2.4. Solvent Models.** The solvent effects on the conformational changes and NMR signals were taken into account by using the explicit solvent model in which the solute was surrounded by 3000 explicit solvent molecules. MD simulations of a solvated solute molecule (urea-based pseudorotaxane or [2]rotaxane) in CH<sub>3</sub>CN and CHCl<sub>3</sub> solvent molecules were performed using the conventional nonpolar FF and polar FF, respectively. A periodic cubic box of 69.3 Å, containing 1777 CH<sub>3</sub>CN and 1223 CHCl<sub>3</sub> solvent molecules, was created. The density was set to 1.13 g/cm<sup>3</sup>. The temperature was maintained at 298.15 K by using an Andersen thermostat method. The nonbonded van der Waals and Coulomb interaction were treated using the atom-based and Ewald truncation methods, respectively, and the cutoff distances for both nonbonded interactions were set to be 15.5 Å. The equations of the motion were integrated by the velocity Verlet algorithm with a time step of 1 fs.

**2.5. Computation of NMR Chemical Shifts.** The NMR isotropic chemical shieldings were calculated with Gaussian 09 software<sup>62</sup> using the gauge invariant/including atomic orbital (GIAO) method.<sup>63–65</sup> The influences of DFT functionals and basis sets on the predicted NMR chemical shifts have been investigated in our previous work.<sup>21</sup> The calculated <sup>1</sup>H NMR chemical shifts using six different functionals were close to each other. In this work, the M06-2X functional with the 6-31G (d, p) basis set was used.

For a direct comparison with the experiment, the calculated NMR isotropic chemical shieldings must be converted to chemical shifts by the following expression.

$$\delta_{\text{cal}} = \sigma_{\text{stand}} - \sigma_{\text{cal}} + \delta_{\text{stand}} \quad (7)$$



**Figure 3.** Fluctuations of bent angle of rod,  $\theta$ , and relative end-to-end distance,  $L/L_0$ , with the time evolutions from (a) nonpolar FF, (b) polar FF, and (c) AIMD simulations.

where  $\delta_{\text{cal}}$  and  $\delta_{\text{stand}}$  are the chemical shifts of target protons and the corresponding standard, respectively, and  $\sigma_{\text{cal}}$  and  $\sigma_{\text{stand}}$  are the isotropic shieldings of target protons and the corresponding standard, respectively. In this work, acetonitrile has been used as standard ( $\sigma_{\text{stand}} = 30.05 \text{ ppm}$ ,  $\delta_{\text{stand}} = 1.94 \text{ ppm}$ ).

To test the convergence behavior of the calculated NMR chemical shifts, Figure S1a displays the MD-averaged chemical shifts for four amide protons using 50, 100, 250, and 500 configurations sampled with even intervals for pseudorotaxane in vacuum. It is shown that statistical averaging over 250 configurations is necessary to get reliable results for all the selected methods. As shown in Figure 1, we adopt 500 configurations to get the ensemble-averaged values of NMR signals.

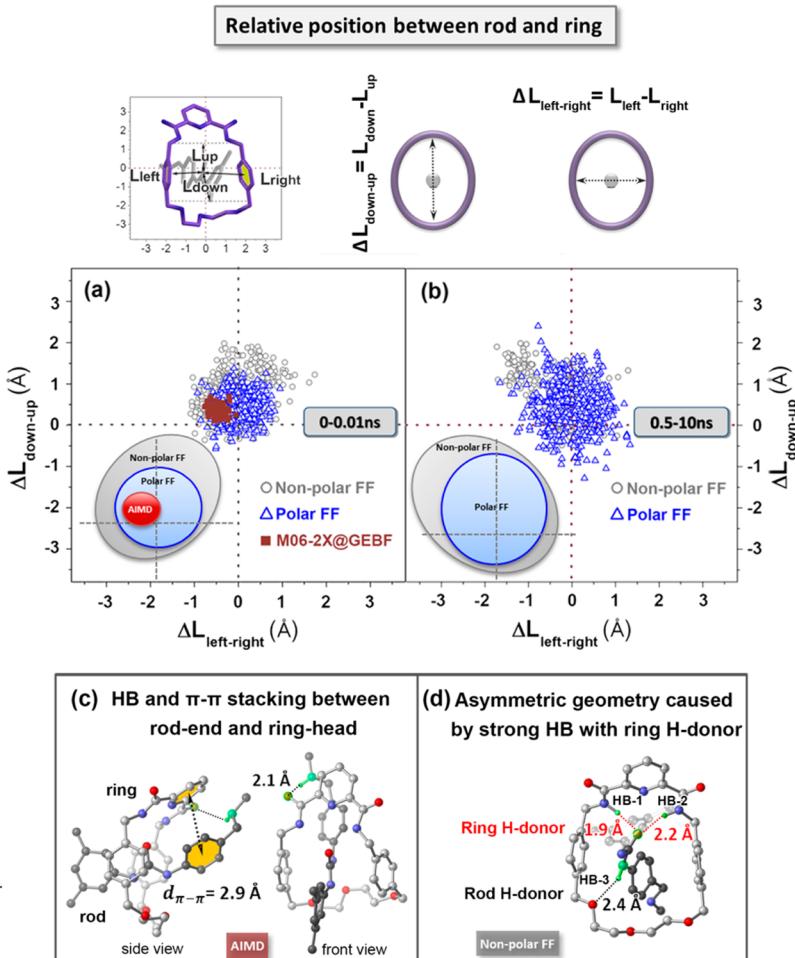
## RESULTS AND DISCUSSION

**3.1. Conformational Ensembles of Pseudorotaxane in Vacuum.** To investigate the conformational changes of the pseudorotaxane in vacuum, the superimpositions of different conformations sampled from different snapshots are shown in Figure S1b of Supporting Information. The overlaid snapshots of the pseudorotaxane illustrate that the conformation of polyether chain in the ring is very flexible. The head of ring, consisting of the 2,6-pyridinediamide group, is swinging backward and forward during the MD simulation.

To further illustrate the details of conformational changes, the bent angle ( $\theta$ ) and the end-to-end length ( $L$ ) are defined to estimate the folded degree of the rod, as shown in Figure 3. The bent angle,  $\theta$ , is the angle between C<sub>stopper</sub>–N<sub>urea</sub> and N<sub>urea</sub>–N<sub>open-end</sub> vectors. The little fluctuation of bent angle with the time

evolution in the AIMD simulation indicates that the rod is rigid in pseudorotaxane. The nonpolar FF model shows the largest fluctuation of the bent angle in the range between 120° and 178°. A smaller variation of  $\theta$  is observed by using the polar FF model with the value ranging from 140° to 178°. In addition, we defined a parameter,  $L/L_0$ , to illustrate the stretching (when  $L/L_0 > 1$ ) or shortening ( $L/L_0 < 1$ ) of rod in pseudorotaxane, where  $L_0$  is the “natural” end-to-end length of rod (12.4 Å) obtained from the geometry optimization at the M06-2X/6-31G (d, p) level. The end-to-end length ( $L$  or  $L_0$ ) is measured as a distance from the N atom of open end (N<sub>open-end</sub>) to the C atom of the terminal 3,5-dimethylphenyl group (C<sub>stopper</sub>) in rod. According to AIMD trajectory of  $L/L_0$ , the rod is slightly folded with the shortening of the end-to-end length. The folded geometry of the rod is ascribed to the  $\pi$ – $\pi$  interaction and the H-bonding interaction between the open-end of rod and the 2,6-pyridinedicarboxamide head of ring. The separation between the center of the benzene rings of the ring head and the rod is ~2.9 Å.

The relative motions of the rod to the cavity center of ring were also analyzed by the geometrical parameters  $L_{\text{up}}$ ,  $L_{\text{down}}$ ,  $L_{\text{left}}$ , and  $L_{\text{right}}$ , as shown in Figure 4. The vertical lengths  $L_{\text{up}}$  and  $L_{\text{down}}$  are estimated from carbonyl C atom of the urea group in guest (rod) to the middle of two methyl C atoms of the xylol rings in host. The horizontal lengths  $L_{\text{left}}$  and  $L_{\text{right}}$  are measured as the distances from carbonyl C atom of binding site to two centroids of the xylol rings. The difference between  $L_{\text{down}}$  and  $L_{\text{up}}$  (labeled as  $\Delta L_{\text{down-up}}$ ) and the difference between  $L_{\text{left}}$  and  $L_{\text{right}}$  (labeled as  $\Delta L_{\text{left-right}}$ ) are adopted to illustrate the perpendicular and horizontal movements of guest in the cavity of host, respectively. The stronger ring-rod interaction leads to the narrower



**Figure 4.** Scatterplots of ring-rod relative locations that sampled from different MD simulations with time durations of (a) 0–0.01 ns and (b) 0.5–10 ns, respectively. Representative snapshots of (c) AIMD and (d) nonpolar FF-based MD trajectories are also presented to show the origin of relative motion of the ring and rod.

distribution of ring-rod relative location (or vice versa). It can be seen from Figure 4a that the AIMD predicts a localized distribution of ring-rod relative location ( $\Delta L_{\text{left-right}} < 1.0 \text{ \AA}$ ,  $\Delta L_{\text{down-up}} < 1.0 \text{ \AA}$ ). The polar FF model provides a more localized distribution than that of nonpolar FF model (Figure 4b). The variation of  $\Delta L_{\text{left-right}}$  or  $\Delta L_{\text{down-up}}$  in the polar FF MD is  $\sim 1.5 \text{ \AA}$ , but it is  $\sim 3.0 \text{ \AA}$  in the nonpolar FF MD. The typical snapshot of AIMD in Figure 4c shows that there are cooperative H-bonding and  $\pi-\pi$  stacking interactions between the open-end of rod and the head of ring, resulting in the up-left orientation of ring-rod location. The N–H···O H-bonding interaction between the carbonyl oxygen on the ring and the amino group on the rod pulls the rod close to the upper parts of macrocyclic component. In the fixed charge FF-based MD simulation, the location of the rod is far from the ring center. Such an asymmetric ring-rod arrangement is induced by the much stronger H-bonding interactions (called HB-1) between the amide group of ring and the urea moiety of rod than the other H-bonding interactions (HB-2, HB-3) between ring and rod (Figure 4d). The H···O distance of HB-1 is  $1.9 \text{ \AA}$ , shorter than those of the other two HBs (HB-2:  $2.2 \text{ \AA}$ ; HB-3:  $2.4 \text{ \AA}$ ).

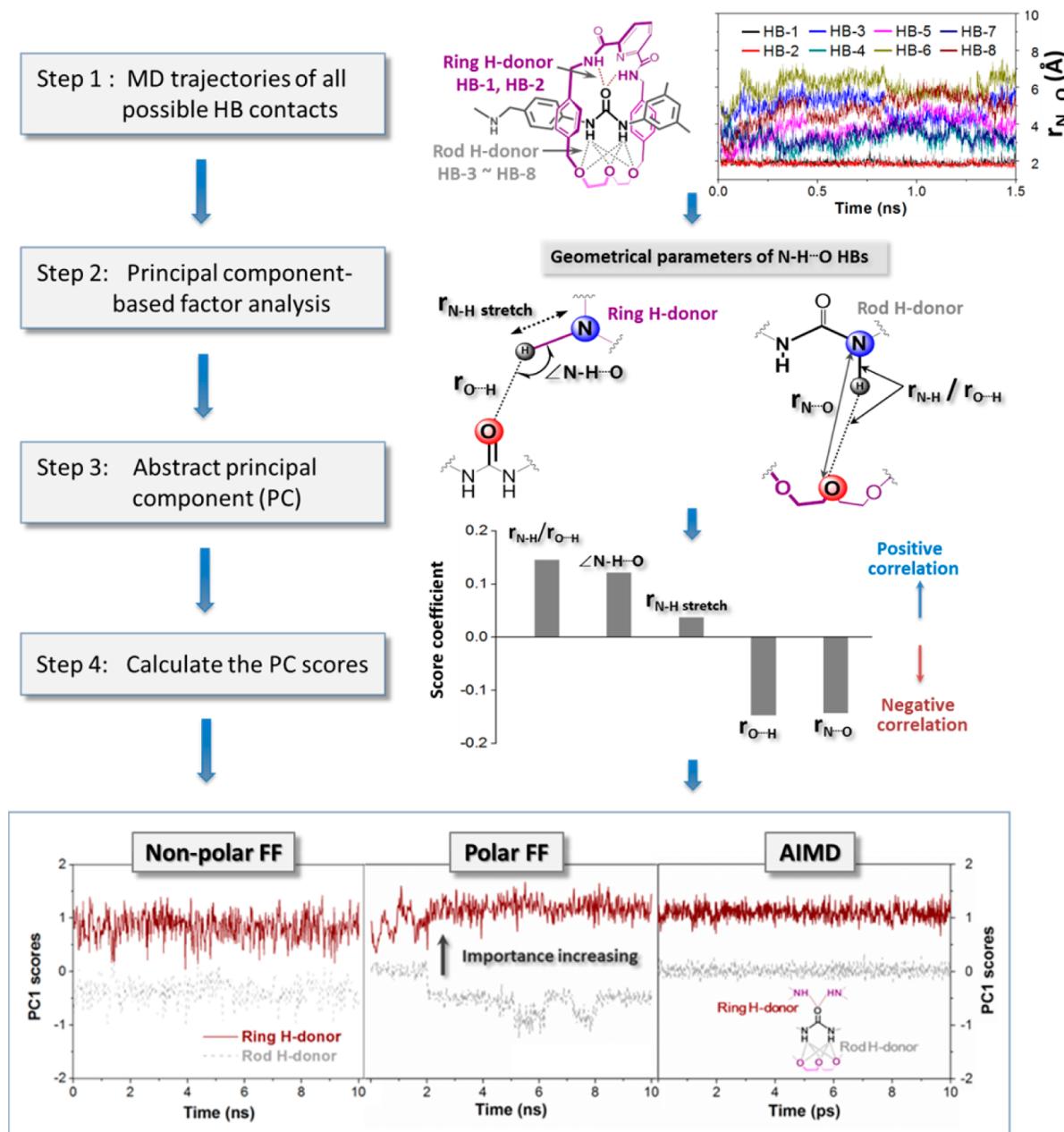
### 3.2. Factor Analysis on Ring-Rod H-Bonding Network.

In most cases, the cooperative H-bonding interactions are employed to assemble the pseudorotaxanes and rotaxanes. A HB forms when two electronegative atoms, such as nitrogen and

oxygen, interact with the same hydrogen. The hydrogen is normally covalently attached to one atom, called the H-donor, but interacts electrostatically with the other, the H-acceptor. In the present work, the ring-rod assembly is controlled by means of a complex H-bonding network, consisting of eight possible N–H···O contacts, as shown in Figure 5.

As described by Jeffrey, the strength of H-bonding interaction is generally associated with five geometrical parameters.<sup>4</sup> Taking the N–H···O type HB as an example, the definitions of the five geometrical parameters, that is, N–H···O angle ( $\angle \text{N–H} \cdots \text{O}$ ), H···O and N···O bond lengths ( $r_{\text{H} \cdots \text{O}}$  and  $r_{\text{N} \cdots \text{O}}$ ), lengthening of N–H bond length ( $r_{\text{N–H stretch}}$ ), and N–H length versus H···O length ( $r_{\text{N–H}} / r_{\text{H} \cdots \text{O}}$ ), are illustrated in Figure 5. The lengthening of N–H bond length is the bond length change before ( $1.008 \text{ \AA}$ ) and after the formation of HB. Ab initio and FF-based MD simulations have produced an immense quantity of trajectory data relating to H-bonding interactions. To reveal the key factor responsible for ring-rod recognition from the complex H-bonding network, PC-FA method is applied to analyze the MD trajectories. The flowchart of PC-FA method is sketched in Figure 5.

The first step of factor analysis is to collect the initial 200 000 H-bonding geometrical parameters in all MD trajectories. Such a large data set is prepared as the input of SPSS software. Tables S1–S3 in Supporting Information show the details of PC-FA



**Figure 5.** Flowchart and analysis results of factor analysis using principal component (PC-FA) method. The definition of the hydrogen bond (HB) geometrical parameters is also given.

treatment, such as the correlation matrix, the initial eigenvalues, the extraction sums of squared loadings, and component score coefficient matrix.

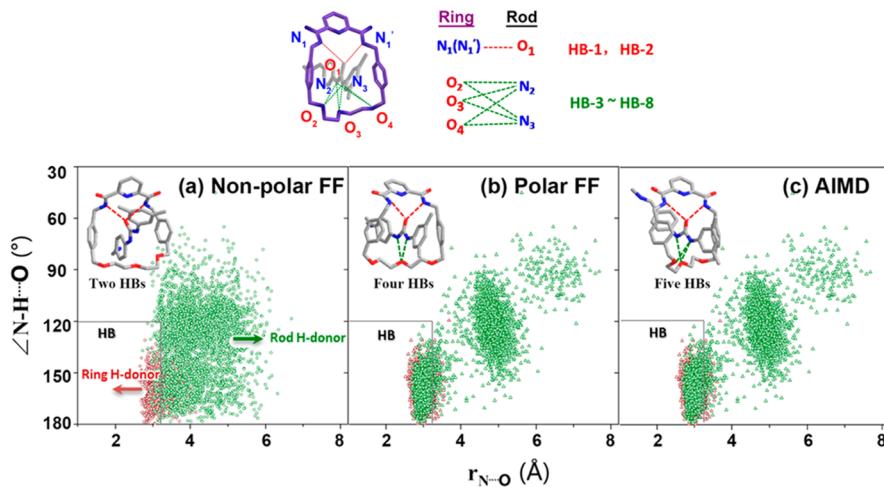
To evaluate the strength of H-bonding interaction, the obtained scores of the PC1 can be calculated by the following equation.

$$\text{PC1score} = 0.146 \times r_{\text{N}-\text{H}}/r_{\text{H}\cdots\text{O}} + 0.122 \times \angle\text{N}-\text{H} - \text{O} + 0.038 \times r_{\text{N}-\text{H}\text{stretch}} - 0.147 \times r_{\text{H}\cdots\text{O}} - 0.143 \times r_{\text{N}\cdots\text{O}} \quad (8)$$

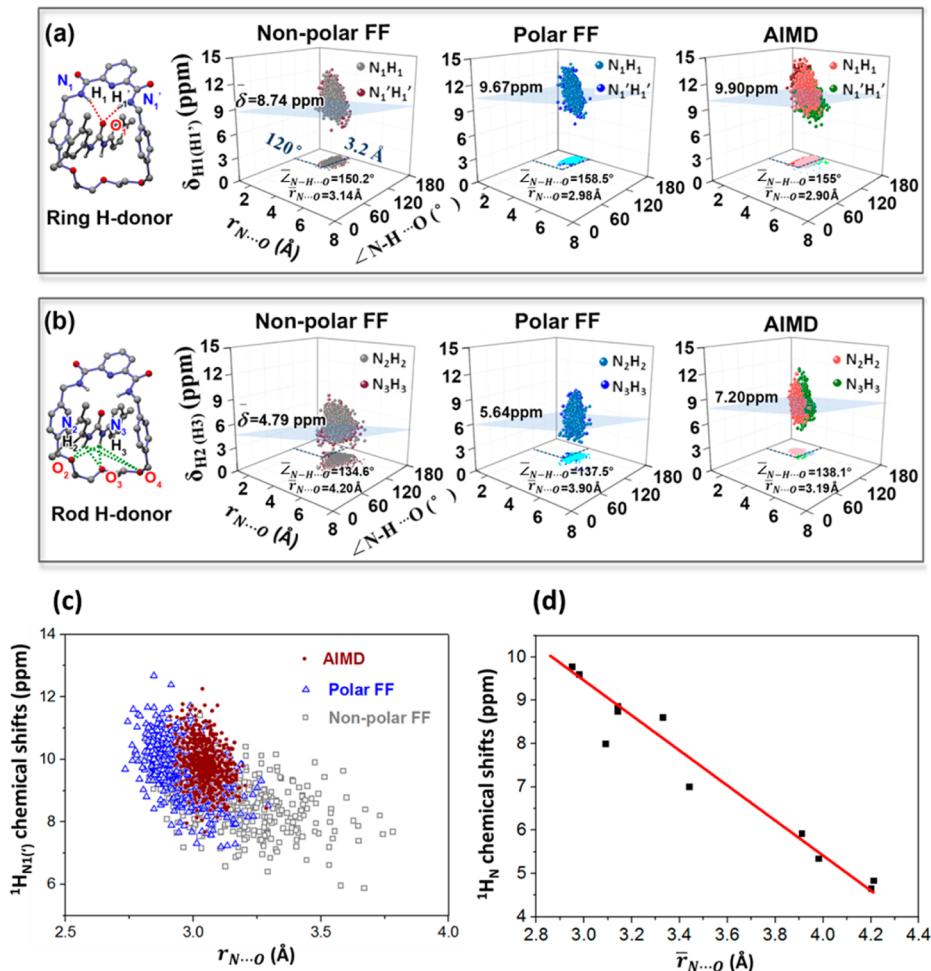
where the PC1 score is positively correlative with the three parameters,  $r_{\text{N}-\text{H}}/r_{\text{H}\cdots\text{O}}$ ,  $\angle\text{N}-\text{H}\cdots\text{O}$ ,  $r_{\text{N}-\text{H}\text{stretch}}$ , and negatively correlative with the other two parameters,  $r_{\text{H}\cdots\text{O}}$  and  $r_{\text{N}\cdots\text{O}}$ . The

result is consistent with the general perception of H-bonding interaction.

Finally, Figure 5 compares the factor analysis results of AIMD and FF-based MD simulations for pseudorotaxane in vacuum. The PC1 score of ring H-donor (red in Figure 5) is obtained by averaging the two PC1 scores of HB-1 and HB-2, which describe the binding between amide group (in ring) and carbonyl oxygen (in rod). The PC1 score of rod H-donor (gray in Figure 5) is the average value of six PC1 scores of the ring-rod binding between the two amide units of urea moiety (in rod) and the oxygen atoms of polyether chain in ring (HB-3 - HB-8). The higher averaged PC1 score means the more important role played in ring-rod interaction. It can be seen that the ring H-donor plays a more important role in ring-rod recognition than the rod H-donor for pseudorotaxane in vacuum. The PC1 score of ring H-



**Figure 6.** Scatterplots of HB geometrical parameters ( $N\cdots O$  length vs  $N-H\cdots O$  angles), and the representative H-bonding configurations taken from (a) nonpolar FF-based, (b) polar FF-based, and (c) AIMD simulations.

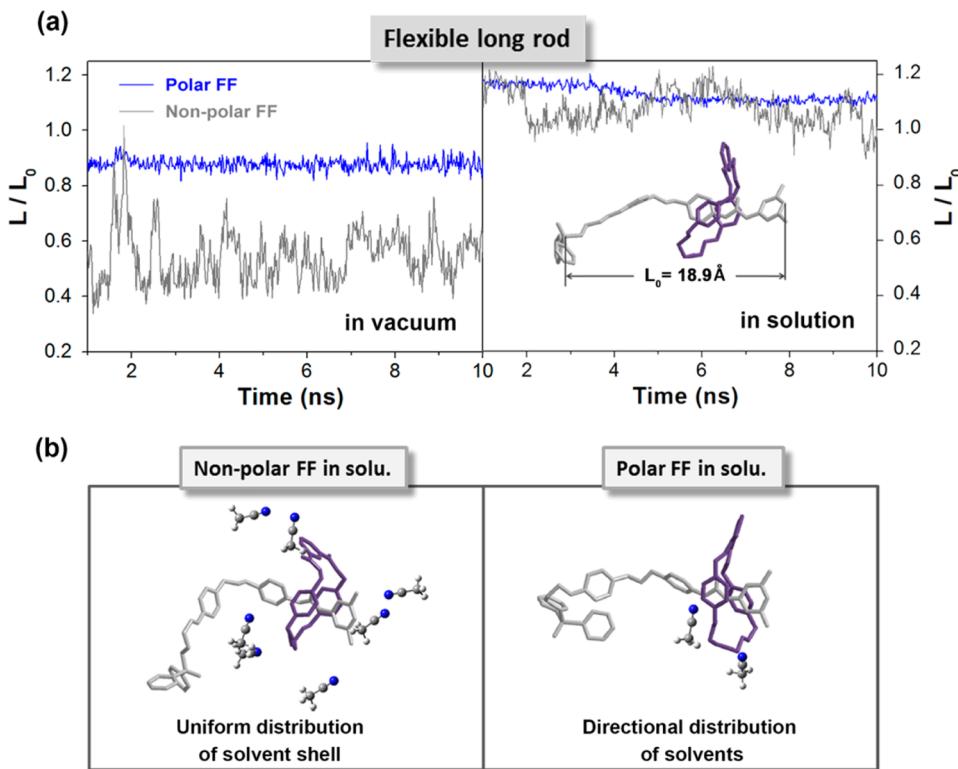


**Figure 7.** Calculated  $^1\text{H}$  NMR chemical shifts and ensemble-averaged  $^1\text{H}$  NMR chemical shifts for amide protons of (a) ring H-donor and (b) rod H-donor using different conformational ensembles, which show good correlations of (c) ring H-donor amide-proton chemical shifts with  $N\cdots O$  lengths,  $r_{N\cdots O}$ , and (d) ensemble-averaged amide-proton chemical shifts with the average lengths,  $\bar{r}_{N\cdots O}$ .

donor is  $\sim 1.0$ , whereas that of rod H-donor is close to or below zero.

To give more details, Figure 6 depicts the distributions of  $N\cdots O$  bond length ( $r_{N\cdots O}$ ) and  $N-H\cdots O$  angle ( $\angle N-H\cdots O$ ) of 500 MD snapshots, which are extracted from different levels of MD

simulations. The two proton donors in the ring are labeled as  $N_1$  and  $N_1'$  (ring H-donor), and three proton acceptors are labeled as  $O_2$ ,  $O_3$ , and  $O_4$ . For the urea group in the rod (called rod H-donor), two proton donors are labeled as  $N_2$  and  $N_3$ , and the proton acceptor is  $O_1$ . Both polar FF and AIMD can lead to the



**Figure 8.** (a) Fluctuations of end-to-end lengths of rod for [2]rotaxane; and (b) representative snapshots of solute–solvent interactions using nonpolar and polar FF MD simulations.

similar localized distributions of N–H···O geometries, different from the discrete distribution of nonpolar FF.

In the present work, a HB is counted when the N···O length is less than 3.2 Å and the N–H···O angle is larger than 120°, as depicted in Figure 6. According to the statistical analysis of the numbers of H-bonds, the most probable number of H-bonds and the representative H-bonding configurations are illustrated in Figure 6. Nonpolar FF model predicts a weakly binding picture (Figure 6a). Results of polar FF (Figure 6b) and AIMD (Figure 6c) indicate a tight ring-rod recognition interaction by means of a cooperative H-bonding network composed by four and five HBs, respectively. Figure S5 also shows the fluctuations of the numbers of HBs with time evolution using different MD simulations. In comparison with the PW91-based AIMD results (Figure S2 and Figure S5), the M06-2X-based AIMD is more suitable for describing the H-bonding interactions in the supramolecular system.

**3.3. Ensemble-Averaged NMR Signals.** In supramolecular chemistry, NMR spectroscopy is a powerful technique for inspection of the site-specific binding interactions, such as H-bonding interactions. The relationship between the  $^1\text{H}$  chemical shifts of H-bonded protons ( $\delta_{\text{H}}$ ) and the geometrical parameters of HBs have been widely used in experimental assignment. The  $^1\text{H}$  NMR chemical shifts of proton donor were calculated using the conformations sampled from different ensembles. The deshielding effect is observed with an overall shift of electron density from the proton donor (–NH) to the acceptor (O). Consequently, the NMR signals are shifted to lower magnetic field. This can be used as an indicator for the formation of a HB.

The dependence of the calculated  $^1\text{H}$  NMR chemical shifts on the N···O bond lengths and N–H···O angles for pseudorotaxane has been investigated. Figure 7 illustrates the calculated NMR chemical shifts (at the M06-2X/6-31G (d, p) level) of amide

protons based on 500 MD snapshots. The MD-ensemble averaged  $^1\text{H}$  chemical shifts ( $\bar{\delta}$ ) of amide protons are also listed in Figure 7. Taking the AIMD predicted NMR values (9.90 ppm) as the reference, the amide-hydrogen chemical shift of ring H-donor is shifted by ca. 0.23 and 1.16 ppm upfield for the polar FF ( $\bar{\delta} = 9.67 \text{ ppm}$ ) and nonpolar FF ( $\bar{\delta} = 8.74 \text{ ppm}$ ) models, respectively (Figure 7a). The polar FF with variable charges gives better agreement with the AIMD in describing the H-bonding interactions. The amide-proton chemical shifts of rod H-donor (Figure 7b) are shifted to the higher field (which corresponds to lower chemical shifts) than those of ring H-donor (Figure 7a). For example, the AIMD averaged NMR chemical shift for the rod H-donor is 7.20 ppm, which is 2.7 ppm lower than that (9.90 ppm) of ring H-donor. This partially indicates that a stronger H-bonding interaction (deshielding effect) of ring H-donor than that of rod H-donor. Such a conclusion is consistent with the result of previous factor analysis on H-bonding interaction from MD trajectories. Another reason is that ring H-donors are amidic hydrogens, while rod H-donors belong to a urea derivative, and urea hydrogens normally have lower chemical shifts than amides. The predicted NMR chemical shifts are in correlation with the lengths of HBs, as exemplified by the N···O distances,  $r_{\text{N}\cdots\text{O}}$ , between the H-donors and H-acceptors. Figure 7c plots the correlation of all the calculated chemical shifts with the  $r_{\text{N}\cdots\text{O}}$  data, while Figure 7d correlates the ensemble-averaged chemical shifts with the average N···O lengths,  $\bar{r}_{\text{N}\cdots\text{O}}$ .

With the aim of generating different initial configurations for AIMD simulations, we combined the polar FF with AIMD conformational ensemble (Figure 1d). The factor analysis was also done for the polar FF+AIMD trajectories (Figure S6). Again, the ring H-donor is found to be very important to hold the interlocked structure of pseudorotaxane. The averaged NMR signals for the chemically equivalent amide protons are predicted

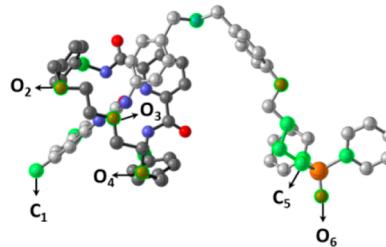
to be 10.10 ppm for ring H-donor  $N_1H_1$  and 10.30 ppm for another  $N_1'H_1'$  by using the combined polar FF+AIMD scheme. In comparison with the hybrid scheme, the 10 ps GEBF-M06-2X-based AIMD yields a larger difference in the NMR signals for the equivalent  $N_1H_1$  and  $N_1'H_1'$  ( $N_1H_1: \delta = 8.97$  ppm;  $N_1'H_1': \delta = 10.76$  ppm). The parallel AIMD simulations with different initial structures give better performance for the chemically equivalent amide protons. Longer time AIMD runs in the hybrid polar FF+AIMD are required to get converged values of ensemble averaging.

**3.4. Solvent Effects.** A systematic investigation of solvent influence on geometry and NMR chemical shift was performed on the pseudorotaxane and [2]rotaxane, respectively. The 10 ns MD simulations of a single solute molecule solvated in 1777  $CH_3CN$  and 1223  $CHCl_3$  were performed using nonpolar FF and polar FF, respectively. For comparison, the 10 ns MD simulations of a single solute molecule in vacuum were also performed.

The changes of rod conformations can be seen from the fluctuations of the end-to-end lengths for pseudorotaxane (Figure S7 in Supporting Information) and [2]rotaxane (Figure 8) in vacuum and in solution. Same as that defined in subsection 3.1, the end-to-end length of pseudorotaxane is estimated from the distance between the N atom of open end and the C atom of the terminal 3,5-dimethylphenyl group. The length is  $\sim 12$  Å whether in vacuum or in solution, indicating a rigid conformation for pseudorotaxane (Figure S7). However, the [2]rotaxane with a long rod exhibits a different picture upon solution. The end-to-end length of [2]rotaxane is then measured as a distance from the P atom of the phosphine oxide group to the C atom of the terminal 3,5-dimethylphenyl group. As shown in Figure 8a, the fluctuation of end-to-end lengths of nonpolar FF is larger than that of polar FF, particularly in vacuum. The largest difference between folded ( $L/L_0 < 1$ ) and extended ( $L/L_0 > 1$ ) geometries in vacuum is  $\sim 13$  Å by using nonpolar FF model. To clarify the difference of rod conformation between the polar and nonpolar FF, five snapshots, selected from polar FF-based MD trajectory at 2, 4, 6, 8, and 10 ns, respectively, are used to analyze the changes of atomic partial charges of [2]rotaxane in vacuum (Table 1). The atomic partial charges are calculated on the different snapshots by using M06-2X functional with the 6-31G (d, p) basis set. In the case of two bulky stopper groups, a large electrostatic repulsion exists between C atom ( $-0.73$  e at 2 ns) of the terminal 3,5-dimethylphenyl group and O atom ( $-1.12$  e) of the phosphine oxide group in polar FF MD simulation. But in the nonpolar FF model, the partial charges are fixed all the time (C:  $-0.16$  e, O:  $-0.01$  e). The electrostatic repulsion between the terminal C and O atoms is weaker in nonpolar FF than that in polar FF. Thus, the strong repulsion interaction between the terminal groups prohibits the folding of the long rod of [2]rotaxane.

It displays from Figure 8 that a uniform solvent shell (with  $\sim 6\text{--}7$  solvent molecules) is distributed around the solute in nonpolar FF. But, the polar FF shows a directional distribution, with two solvents coming close to the flexible diethylene glycol tail of ring. The representative configurations of the neighboring solvents are shown in Figure 8b. A large difference between the nonpolar FF and polar FF results can be also found in the distributions of the N···O lengths and N–H···O angles of ring H-donor and rod H-donor for pseudorotaxane (Figure S9) and [2]rotaxane (Figure 9a). For nonpolar FF (left side of Figure 9a), the solvent effect on the H-bonding geometry on solvents is small. However, for polar FF (right side), the distributions of

**Table 1. Differences in Atomic Partial Charges ( $q$ , in units of  $e$ ) between Variable Charge Polar FF Model and the Fixed Charge Non-Polar Model of [2]Rotaxane in Vacuum**



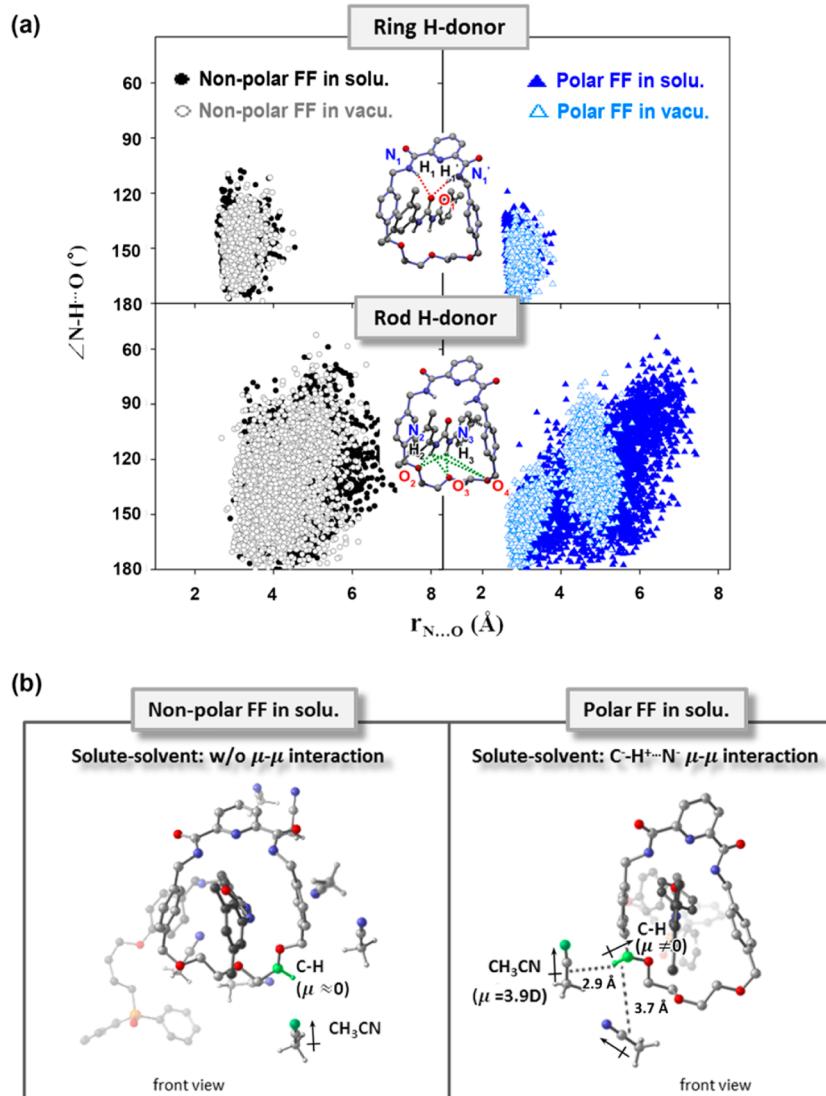
atom	nonpolar FF	polar FF				
		2 ns	4 ns	6 ns	8 ns	10 ns
rod						
C <sub>1</sub>	-0.16	-0.73	-0.71	-0.72	-0.71	-0.71
C <sub>5</sub>	-0.16	-0.89	-0.96	-0.90	-0.94	-0.91
O <sub>6</sub>	-0.01	-1.12	-1.11	-1.10	-1.11	-1.08
ring						
O <sub>2</sub>	-0.27	-0.61	-0.60	-0.61	-0.59	-0.59
O <sub>3</sub>	-0.27	-0.59	-0.59	-0.59	-0.62	-0.58
O <sub>4</sub>	-0.27	-0.60	-0.61	-0.63	-0.64	-0.64

geometrical parameters of rod H-donor in Figure 9a are broader in solution than those in vacuum. That means the conformational changes of HBs are predicted to be larger in solution by using polar FF. The factor analysis illustrates that the ring H-donor also dominates the ring-rod recognition in solution (Figure S10). As shown in Figure 9b, in nonpolar FF model, the partial charges of C (+0.03 e) and H (+0.05 e) atoms are close to each other, while the outward C–H bond is nonpolar with  $\mu \approx 0$ . There is no specific dipole–dipole ( $\mu\text{--}\mu$ ) electrostatic interaction between solute and solvent. Therefore, the  $CH_3CN$  solvents are distributed homogeneously around the solute. In contrast, in polar FF, the dipole–dipole interactions exist between polar  $CH_3CN$  solvents and the polar C–H bonds (e.g.,  $q_C = -0.13$  e;  $q_H = +0.22$  e at 10 ns) of flexible diethylene glycol chain of ring. Those induce the directional and inhomogeneous distribution of solvents close to the macrocyclic component.

Table 2 lists the calculated MD ensemble-averaged chemical shifts ( $\bar{\delta}$ ) for amide protons in ring and rod for [2]rotaxane as well as the experimental NMR signals in mixed  $CH_3CN$ – $CHCl_3$  solution. The polar FF provides better predicted NMR chemical shifts of rod H-donor than the nonpolar FF model. The last column of Table 2 lists the differences of NMR signals between those calculated in solution ( $\bar{\delta}_{solu}$ ) and in vacuum ( $\bar{\delta}_{vacu}$ ). The solvent molecules induce a slight downfield shift (nonpolar FF: +0.21 ppm, polar FF: +0.28 ppm) for  $-N_1H_1$  ( $N_1'H_1'$ ) of strong ring H-donor. For the weak rod H-donor, the solvent molecules induce a greater upfield shift for polar FF (−0.60 to ca. −0.83 ppm) than nonpolar FF (−0.27 to ca. −0.34 ppm), indicating a weaker H-bonding interaction of rod H-donor. The observation is consistent with the analysis of HB geometry in Figure 9.

#### 4. CONCLUSIONS

The conformational changes of pseudorotaxane and [2]rotaxane in vacuum and in solution have been investigated by MD simulations based on ab initio methods and polarizable force fields. The dynamic changes of structures, ring-rod relative movements, and HB geometrical parameters were traced in different conformational ensembles and different environments. A hydrogen-bonding network between ring and rod permitted a



**Figure 9.** (a) Distributions of N...O lengths and N–H...O angles for pseudorotaxane and [2]rotaxane in vacuum and in solution. (b) Representative snapshots of solute–solvent interactions using nonpolar and polar FF MD simulations in solution.

**Table 2. Ensemble-Averaged  $^1\text{H}$  NMR Chemical Shifts ( $\bar{\delta}$  in ppm) Using FF-Based Molecular Dynamics Simulations and Solvent Effect ( $\bar{\delta}_{\text{solu}} - \bar{\delta}_{\text{vacu}}$ ) on the Averaged Values,  $\bar{\delta}$ , for Amide Protons of [2]Rotaxane**

	in vacuum		in solution			$\bar{\delta}_{\text{solu}} - \bar{\delta}_{\text{vacu}}$	
	nonpolar.	polar.	nonpolar.	polar.	exp. <sup>a</sup>	nonpolar.	polar.
ring H-donor							
N <sub>1</sub> H <sub>1</sub> /N <sub>1</sub> 'H <sub>1</sub> '	8.79	9.58	9.00	9.86	9.17	+0.21	+0.28
rod H-donor							
-N <sub>2</sub> H <sub>2</sub>	4.77	5.98	4.43	5.15	6.46	-0.34	-0.83
-N <sub>3</sub> H <sub>3</sub>	4.65	5.23	4.38	4.63	6.29	-0.27	-0.60

<sup>a</sup>The experimental data were taken from ref 66.

tight binding to avoid dethreading for pseudorotaxane and [2]rotaxane whether in vacuum or in solution. Factor analysis indicated that the ring H-donor (2,6-pyridinedicarboxamide group) played a more important role responsible for the ring–rod interaction than the rod H-donor. The ensemble-averaged values of the  $^1\text{H}_\text{N}$  chemical shifts were calculated at different theoretical levels. The correlation between the deshielding degree of  $^1\text{H}_\text{N}$  chemical shifts and lengths of HBs has been revealed in both vacuum and solution. The hybrid polar FF+AIMD approach

provided different initial states for AIMD, improving the performance of AIMD in predicting NMR signals of chemically equivalent amide protons of ring H-donor. In polar FF MD simulations, the dipole–dipole interactions between the polar CH<sub>3</sub>CN solvents and the local C–H bond of the ring caused the directional and inhomogeneous distribution of solvents close to the flexible diethylene glycol group of ring. The understanding of the interplay between the solvent–solute interactions and the

guest–host H-bonding interactions is useful for the rational design of novel interlocked systems.

## ■ ASSOCIATED CONTENT

### S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jpca.5b10085](https://doi.org/10.1021/acs.jpca.5b10085).

Convergence behaviors of calculated  $^1\text{H}$  chemical shifts of amide protons for the pseudorotaxane in vacuum and structural variations with the time evolutions. Comparison of PW91-based AIMD with the GEBF-M06-2X AIMD and the force field based MD simulations. Flowchart and potential energies of polar-FF MD simulations. Results of principal component-based factor analysis. The number of H-bonds as a function of MD simulation time. First principle component scores of ring-rod contacting geometries using polar FF+AIMD method. Fluctuations of end-to-end lengths of rod for pseudorotaxane. Variation of number of solvents around the solute for solvated [2]rotaxane. Analysis of HB geometry of pseudorotaxane. PC1 score plots for pseudorotaxane and [2]rotaxane in vacuum and in solution. ([PDF](#))

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [wli@nju.edu.cn](mailto:wli@nju.edu.cn).

\*E-mail: [majing@nju.edu.cn](mailto:majing@nju.edu.cn). Phone: (86)25-89687408.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program (Grant No. 2011CB808604) and the National Natural Science Foundation of China (Grant Nos. 21290192, 21273102, 21473087, and 21333004). We are grateful to the High Performance Computing Center of Nanjing University for doing the quantum chemical calculations in this paper on its IBM Blade cluster system.

## ■ REFERENCES

- (1) Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. Noncovalent Synthesis Using Hydrogen Bonding. *Angew. Chem., Int. Ed.* **2001**, *40*, 2382–2426.
- (2) Cooke, G.; Rotello, V. M. Methods of Modulating Hydrogen Bonded Interactions in Synthetic Host-Guest Systems. *Chem. Soc. Rev.* **2002**, *31*, 275–286.
- (3) Schneider, H. J. Binding Mechanisms in Supramolecular Complexes. *Angew. Chem., Int. Ed.* **2009**, *48*, 3924–3977.
- (4) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, U.K., 1997.
- (5) Desiraju, G. R. Hydrogen Bridges in Crystal Engineering: Interactions without Borders. *Acc. Chem. Res.* **2002**, *35*, 565–573.
- (6) Pluth, M. D.; Raymond, K. N. Reversible Guest Exchange Mechanisms in Supramolecular Host-Guest Assemblies. *Chem. Soc. Rev.* **2007**, *36*, 161–171.
- (7) Brunner, E.; Sternberg, U. Solid-State NMR Investigations on the Nature of Hydrogen Bonds. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 21–57.
- (8) Brown, S. P.; Spiess, H. W. Advanced Solid-State NMR Methods for the Elucidation of Structure and Dynamics of Molecular, Macromolecular, and Supramolecular Systems. *Chem. Rev.* **2001**, *101*, 4125–4155.
- (9) Chierotti, M. R.; Gobetto, R. Solid-State NMR Studies of Weak Interactions in Supramolecular Systems. *Chem. Commun.* **2008**, 1621–1634.
- (10) Eckert, H.; Yesinowski, J. P.; Silver, L. A.; Stolper, E. M. Water in Silicate-Glasses - Quantitation and Structural Studies by  $\text{H}-1$  Solid Echo and MAS-NMR Methods. *J. Phys. Chem.* **1988**, *92*, 2055–2064.
- (11) Bertolasi, V.; Gilli, P.; Ferretti, V.; Gilli, G. Intramolecular O–H...O Hydrogen Bonds Assisted by Resonance. Correlation between Crystallographic Data and  $^1\text{H}$  NMR Chemical Shifts. *J. Chem. Soc., Perkin Trans. 2* **1997**, 945–952.
- (12) Harris, T. K.; Zhao, Q.; Mildvan, A. S. NMR Studies of Strong Hydrogen Bonds in Enzymes and in a Model Compound. *J. Mol. Struct.* **2000**, *552*, 97–109.
- (13) Dračinský, M.; Möller, H. M.; Exner, T. E. Conformational Sampling by Ab Initio Molecular Dynamics Simulations Improves NMR Chemical Shift Predictions. *J. Chem. Theory Comput.* **2013**, *9*, 3806–3815.
- (14) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632–672.
- (15) Collins, M. A.; Bettens, R. P. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.* **2015**, *115*, 5607–5642.
- (16) Raghavachari, K.; Saha, A. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015**, *115*, 5643–5677.
- (17) Li, S. H.; Li, W.; Fang, T. An Efficient Fragment-Based Approach for Predicting the Ground-State Energies and Structures of Large Molecules. *J. Am. Chem. Soc.* **2005**, *127*, 7215–7226.
- (18) Li, W.; Li, S. H.; Jiang, Y. S. Generalized Energy-Based Fragmentation Approach for Computing the Ground-State Energies and Properties of Large Molecules. *J. Phys. Chem. A* **2007**, *111*, 2193–2199.
- (19) Hua, W. J.; Fang, T.; Li, W.; Yu, J. G.; Li, S. H. Geometry Optimizations and Vibrational Spectra of Large Molecules from a Generalized Energy-Based Fragmentation Approach. *J. Phys. Chem. A* **2008**, *112*, 10864–10872.
- (20) Li, S.; Li, W.; Ma, J. Generalized Energy-Based Fragmentation Approach and its Applications to Macromolecules and Molecular Aggregates. *Acc. Chem. Res.* **2014**, *47*, 2712–2720.
- (21) Liu, P.; Li, W.; Liu, L.; Wang, L.; Ma, J. Theoretical Study on Conformation Dynamics of Three-Station Molecular Shuttle in Different Environments and its Influence on NMR Chemical Shifts and Binding Interactions. *J. Phys. Chem. A* **2014**, *118*, 9032–9044.
- (22) Zhang, L.; Li, W.; Fang, T.; Li, S. Ab Initio Molecular Dynamics with Intramolecular Noncovalent Interactions for Unsolvated Polypeptides. *Theor. Chem. Acc.* **2015**, in press.
- (23) Halgren, T. A.; Damm, W. Polarizable Force Fields. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236–242.
- (24) Rick, S. W.; Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev. Comp. Chem.* **2002**, *18*, 89–146.
- (25) Warshel, A.; Kato, M.; Pisliakov, A. V. Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* **2007**, *3*, 2034–2045.
- (26) Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Molecular Modeling and Dynamics Studies with Explicit Inclusion of Electronic Polarizability. Theory and Applications. *Theor. Chem. Acc.* **2009**, *124*, 11–28.
- (27) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. Polarization Effects in Molecular Mechanical Force Fields. *J. Phys.: Condens. Matter* **2009**, *21*, 333102.
- (28) Wang, X.; Yan, T.; Ma, J. Polarizable Force Fields based on Physical Models and Quantum Chemical Calculations. *Int. J. Quantum Chem.* **2015**, *115*, 545–549.
- (29) Leverenz, H. R.; Gao, J.; Truhlar, D. G. Using Multipole Point Charge Distributions to Provide the Electrostatic Potential in the Variational Explicit Polarization (X-Pol) Potential. *Theor. Chem. Acc.* **2011**, *129*, 3–13.

- (30) Patel, S.; Mackerell, A. D., Jr.; Brooks, C. L. CHARMM Fluctuating Charge Force Field for Proteins: II Protein/Solvent Properties From Molecular Dynamics Simulations Using a Nonadditive Electrostatic Model. *J. Comput. Chem.* **2004**, *25*, 1504–1514.
- (31) Wang, J.; Cieplak, P.; Li, J.; Wang, J.; Cai, Q.; Hsieh, M.; Lei, H.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. *J. Phys. Chem. B* **2011**, *115*, 3100–3111.
- (32) Lamoureux, G.; MacKerell, A. D.; Roux, B. T. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185–5197.
- (33) Vorobyov, I. V.; Anisimov, V. M.; MacKerell, A. D. Polarizable Empirical Force Field for Alkanes Based on the Classical Drude Oscillator Model. *J. Phys. Chem. B* **2005**, *109*, 18988–18999.
- (34) Yu, H. B.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D.; Roux, B. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 774–786.
- (35) Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; Mackerell, A. D., Jr. Force Field for Peptides and Proteins based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*, 5430–5449.
- (36) Chelli, R.; Procacci, P. A Transferable Polarizable Electrostatic Force Field for Molecular Mechanics based on the Chemical Potential Equalization Principle. *J. Chem. Phys.* **2002**, *117*, 9175–9189.
- (37) Chelli, R.; Schettino, V.; Procacci, P. Comparing Polarizable Force Fields to Ab Initio Calculations Reveals Nonclassical Effects in Condensed Phases. *J. Chem. Phys.* **2005**, *122*, 234107.
- (38) Jiang, N.; Ma, J. Conformational Simulations of Aqueous Solvated Alpha-conotoxin GI and its Single Disulfide Analogues Using a Polarizable Force Field Model. *J. Phys. Chem. A* **2008**, *112*, 9854–9867.
- (39) Jiang, N.; Ma, J. Influence of Disulfide Connectivity, Electrostatics, and Hydrophobicity on the Conformational Variations of  $\alpha$ -Conotoxin GI Single–Disulfide Analogues: Simulations with Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 11241–11250.
- (40) Zhao, J.; Wang, X.; Jiang, N.; Yan, T.; Kan, Z.; Mendes, P. M.; Ma, J. Polarization Effect and Electric Potential Changes in the Stimuli-Responsive Molecular Monolayers Under an External Electric Field. *J. Phys. Chem. C* **2015**, *119*, 22866–22881.
- (41) Kan, Z.; Yan, X.; Ma, J. Conformation Dynamics and Polarization Effect of  $\alpha,\alpha$ -Trehalose in a Vacuum and in Aqueous and Salt Solutions. *J. Phys. Chem. A* **2015**, *119*, 1573–1589.
- (42) Czarnik-Matusewicz, B.; Kim, S. B.; Jung, Y. M. A Study of Urea-dependent Denaturation of beta-Lactoglobulin by Principal Component Analysis and Two-dimensional Correlation Spectroscopy. *J. Phys. Chem. B* **2009**, *113*, 559–566.
- (43) Riccardi, L.; Nguyen, P. H.; Stock, G. Free-Energy Landscape of RNA Hairpins Constructed via Dihedral Angle Principal Component Analysis. *J. Phys. Chem. B* **2009**, *113*, 16660–16668.
- (44) Ivanov, P. M. Conformations of Some Large-Ring Cyclodextrins Derived from Conformational Search with Molecular Dynamics Simulations and Principal Component Analysis. *J. Phys. Chem. B* **2010**, *114*, 2650–2659.
- (45) Ivanov, P. Conformations of some Lower-Size Large-Ring Cyclodextrins Derived from Conformational Search with Molecular Dynamics and Principal Component Analysis. *J. Mol. Struct.* **2012**, *1009*, 3–10.
- (46) Yamamoto, N. Hot Spot of Structural Ambivalence in Prion Protein Revealed by Secondary Structure Principal Component Analysis. *J. Phys. Chem. B* **2014**, *118*, 9826–9833.
- (47) Sherer, E. C.; Harris, S. A.; Soliva, R.; Orozco, M.; Laughton, C. A. Molecular Dynamics Studies of DNA A-Tract Structure and Flexibility. *J. Am. Chem. Soc.* **1999**, *121*, 5981–5991.
- (48) Wlodek, S. T.; Clark, T. W.; Scott, L. R.; McCammon, J. A. Molecular Dynamics of Acetylcholinesterase Dimer Complexed with Tacrine. *J. Am. Chem. Soc.* **1997**, *119*, 9513–9522.
- (49) Alakent, B.; Doruker, P.; Camurdan, M. C. Application of Time Series Analysis on Molecular Dynamics Simulations of Proteins: A Study of Different Conformational Spaces by Principal Component Analysis. *J. Chem. Phys.* **2004**, *121*, 4759–4769.
- (50) Daidone, I.; D'Abromo, M.; Di Nola, A.; Amadei, A. Theoretical Characterization of Alpha-Helix and Beta-Hairpin Folding Kinetics. *J. Am. Chem. Soc.* **2005**, *127*, 14825–14832.
- (51) Pérez, A.; Blas, J. R.; Rueda, M.; López-Bes, J. M.; de la Cruz, X.; Orozco, M. Exploring the Essential Dynamics of B-DNA. *J. Chem. Theory Comput.* **2005**, *1*, 790–800.
- (52) Lou, H. F.; Cukier, R. I. Molecular Dynamics of Apo-Adenylate Kinase: A Principal Component Analysis. *J. Phys. Chem. B* **2006**, *110*, 12796–12808.
- (53) Meyer, T.; Ferrer-Costa, C.; Perez, A.; Rueda, M.; Bidon-Chanal, A.; Luque, F. J.; Laughton, C. A.; Orozco, M. Essential Dynamics: A Tool for Efficient Trajectory Compression and Management. *J. Chem. Theory Comput.* **2006**, *2*, 251–258.
- (54) Koons, J. M.; Ellis, P. D. Applicability of Factor-Analysis in Solid-State NMR. *Anal. Chem.* **1995**, *67*, 4309–4315.
- (55) Yilmaz, A.; Nyberg, N. T.; Jaroszewski, J. W. Metabolic Profiling Based on Two-Dimensional-J-Resolved  $^1\text{H}$  NMR Data and Parallel Factor Analysis. *Anal. Chem.* **2011**, *83*, 8278–8285.
- (56) Casassas, E.; Tauler, R.; Marques, I. Interactions of  $\text{H}^+$  and Cu(II) Ions with Poly(adenylic Acid): Study by Factor Analysis. *Macromolecules* **1994**, *27*, 1729–1737.
- (57) Park, W. B.; Singh, S. P.; Kim, M.; Sohn, K. S. Phosphor Informatics based on Confirmatory Factor Analysis. *ACS Comb. Sci.* **2015**, *17*, 317–325.
- (58) Li, W.; Chen, C.; Zhao, D.; Li, S. LSQC: Low Scaling Quantum Chemistry Program. *Int. J. Quantum Chem.* **2015**, *115*, 641–646.
- (59) Li, S.; Li, W.; Fang, T.; Ma, J.; Hua, W.; Hua, S.; Jiang, Y. LSQC Program, version 2.2; Nanjing University; Nanjing, China, 2012.
- (60) Materials studio, Version 4.0; Accelrys, Inc: San Diego, CA, 2006.
- (61) Statistical Package for Social Science, version 22.0 for Windows; SPSS Inc: Chicago, IL.
- (62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian 09, Revision B.01; Gaussian, Inc: Wallingford, CT, 2010.
- (63) McWeeny, R. Perturbation Theory for the Fock-Dirac Density Matrix. *Phys. Rev.* **1962**, *126*, 1028–1034.
- (64) Ditchfield, R. Self-Consistent Perturbation Theory of Diamagnetism. I. Gauge-Invariant LCAO Method for N.M.R. Chemical Shifts. *Mol. Phys.* **1974**, *27*, 789–807.
- (65) Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient Implementation of the Gauge-Independent Atomic Orbital Method for NMR Chemical-Shift Calculations. *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.
- (66) Liu, L.; Liu, Y.; Liu, P.; Wu, J.; Guan, Y.; Hu, X.; Lin, C.; Yang, Y.; Sun, X.; Ma, J.; Wang, L. Phosphine Oxide Functional Group Based Three-Station Molecular Shuttle. *Chem. Sci.* **2013**, *4*, 1701–1706.