

# Many-Body Optimization Using an Ab Initio Monte Carlo Method

Ned C. Haubein, Scott A. McMillan, and Linda J. Broadbelt\*

Department of Chemical Engineering, Northwestern University, Evanston, Illinois 60208-3120

Received May 31, 2002

Advances in computing power have made it possible to study solvated molecules using ab initio quantum chemistry. Inclusion of discrete solvent molecules is required to determine geometric information about solute/solvent clusters. Monte Carlo methods are well suited to finding minima in many-body systems, and ab initio methods are applicable to the widest range of systems. A first principles Monte Carlo (FPMC) method was developed to find minima in many-body systems, and emphasis was placed on implementing moves that increase the likelihood of finding minimum energy structures. Partial optimization and molecular interchange moves aid in finding minima and overcome the incomplete sampling that is unavoidable when using ab initio methods. FPMC was validated by studying the boron trifluoride-water system, and then the method was used to examine the methyl carbenium ion in water to demonstrate its application to solvation problems.

## 1. INTRODUCTION

Increases in both the power of computer hardware and the efficiency of computational chemistry algorithms have made it possible to move from the study of simple, model systems to the study of more realistic systems. One area where significant gains have been achieved is the study of molecules in the presence of a liquid solvent. A number of methods have been developed to calculate properties of solute/solvent systems, including group contribution methods, continuum approximations of the solvent, and direct inclusion of many discrete solvent molecules.

The simplest, but most limited, of these approaches is group contribution methods. Group contribution methods are able to predict the free energy of solvation of a molecule given molecular connectivity information.<sup>1,2</sup> While these methods are fast and reliable for molecules similar to those for which they were parametrized, they are not suitable for prediction of the free energy of solvation of species which are not similar to those against which the method was parametrized. Ionic species are a particularly good example of the types of species to which group contribution methods are not readily applicable. Additionally, no information about the geometry of the solute is obtained.

To obtain information about the geometry of the solvated species, the solute must be optimized in the presence of a continuum or discrete representation of the solvent. Continuum models<sup>3</sup> have the advantage of being computationally efficient but are limited in that they can only describe nonspecific solute–solvent interactions. Both specific and nonspecific interactions can be accounted for in a discrete fashion by either including a large number of explicit solvent molecules in the calculation or by adding a small number of explicit solvent molecules to a solute in the presence of a continuum solvent.<sup>4</sup> However, two difficulties have limited the degree to which either method is successful.

The first challenge encountered when using a discrete description of the solvent is the location of the global minimum in energy for a many-body system. Many-body

systems often have complicated potential energy surfaces which exhibit many local and shallow minima. Conventional optimization algorithms do not handle either situation well as they only proceed to the nearest minimum and may require many steps to converge to the minimum on a flat surface. One way to handle these complicated systems is through the use of Monte Carlo and simulated annealing methods, which often yield good results if accurate evaluations of the energy are available. A related area of research is conformational searching where one is interested in finding minimum energy configurations of flexible molecules. Many of these studies have combined the use of Monte Carlo techniques with energy minimizations using classical force fields.<sup>5–8</sup> Accurate force fields are readily available if studying organic or biological molecules, but reliable force fields are less available when the molecule in question contains inorganic or metal atoms. In these cases, ab initio calculations are more reliable and may be the only viable option.

The second difficulty encountered when using a discrete solvent description is that adding many solvent molecules to the system can make the calculation time increase by an order of magnitude or more, depending on the number of molecules and the method used to evaluate the energy. This is the primary reason that discrete descriptions of the solvent have been used in conjunction with classical Monte Carlo and molecular dynamics methods rather than the more computationally demanding ab initio methods.<sup>9,10</sup> To alleviate some of the computational burden, classical-quantum mechanical hybrid methods such as the ONIOM method have been developed. In these calculations, the solute can be treated quantum mechanically, and the discrete solvent molecules are treated using a classical force field. This allows one to achieve an accurate quantum mechanical description of the solute while avoiding the extremely long calculation times associated with applying ab initio methods to both the solute and solvent. One limitation of the hybrid methods<sup>11</sup> is that solvent polarization is only captured if a polarizable force field is used, and these force fields are not always readily available. Ab initio calculations do not suffer from this limitation.

\* Corresponding author phone: (847) 491-5351; e-mail: broadbelt@northwestern.edu.

There have been a small number of investigations that have demonstrated the utility of combining Monte Carlo and ab initio methods to investigate solute/solvent systems by looking at small clusters of molecules. Previous ab initio Monte Carlo research includes the study of lithium–hydrogen clusters,<sup>12,13</sup> fluoride–water clusters,<sup>14</sup> the water dimer,<sup>15</sup> hydrogen chloride–water clusters,<sup>16</sup> and lithium chloride–water clusters.<sup>17</sup>

In the present research, a first principles Monte Carlo (FPMC) method has been developed with the specific goal of finding the global minimum in many-body systems. Emphasis was placed on implementing moves that increase the likelihood of finding minimum energy structures. In particular, partial optimizations and molecular interchange moves have been implemented as described in the Computational Methods section. FPMC was validated by studying the boron trifluoride–water system and then applied to examine the methyl carbenium ion in water to demonstrate its application to solvation problems.

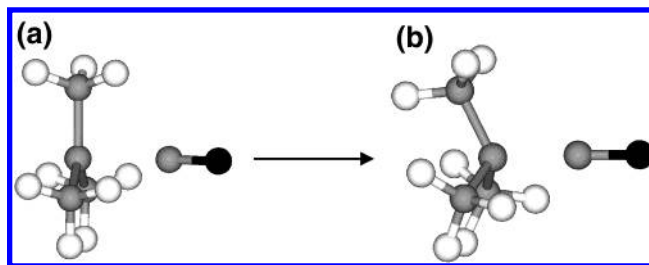
## 2. COMPUTATIONAL METHODS

FPMC is a parallel, object-oriented Perl program. Energy evaluations and optimizations are carried out using a separate ab initio quantum chemistry package, which is controlled by FPMC. Perl was chosen due to the ease with which this language handles ASCII text files and communicates with other programs. In conventional Monte Carlo schemes, use of a more scientifically oriented language such as FORTRAN or C is preferred because computational efficiency is of the utmost importance, and large amounts of time are spent evaluating the energy of the system. In this case, the vast majority of the computational demand is due to the external energy evaluation program, and the increased computational cost incurred by using a scripting language to control this program is far outweighed by the convenience it affords.

The primary functions of FPMC are to perform molecular moves, set up the input files for and collect the output from the energy evaluation program, accept or reject moves, adjust the temperature during simulated annealing calculations, and stop the calculation if any of the termination criteria are met. In a manner similar to FPMC functions most other Monte Carlo programs, but unlike traditional implementations, FPMC includes two novel types of moves and interfaces with an arbitrary, external energy evaluation program that includes capabilities for ab initio energy evaluations. Each of these functions will now be described in more detail.

**Molecular Moves.** A molecular system consists of a set of discrete molecules which are stored separately by their nuclear coordinates. The type of move performed on a given molecule is randomly selected with user-defined probabilities. The basic moves are random translation of or rotation about the center of mass of a rigid molecule. The two novel types of moves implemented within FPMC are partial optimization and molecular interchange followed by full optimization. Both types of moves are specifically designed to aid in the search for the global minimum in a many-body system.

Partial optimization moves consist of the optimization of one molecule in a system while the nuclear coordinates of all other molecules remain fixed. There are two consequences of partial optimization. First, this move results in internal relaxation of the selected molecule, allowing it to respond

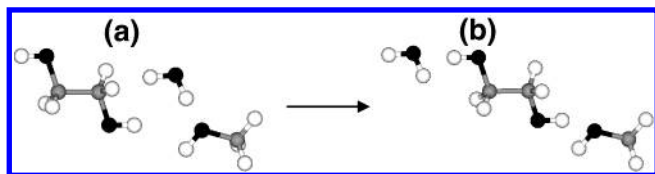


**Figure 1.** A partial optimization move allows a planar *tert*-butyl cation (a) to react to the presence of the CO and assume a bent configuration (b). Carbon atoms are shown in gray, oxygen atoms in black, and hydrogen atoms in white.

to its current environment. This relaxation could also be achieved by implementing intramolecular moves, but these moves are often rejected because atoms are moved to overlapping positions. Configurational bias<sup>18</sup> and concerted rotation<sup>19</sup> could also achieve this type of transformation, but partial optimization has the advantage of being easy to implement. In addition, partial optimizations are accepted almost any time they are attempted and are therefore very efficient relative to conventional intramolecular moves. Figure 1a,b illustrates the partial optimization of a *tert*-butyl cation in the presence of a carbon monoxide molecule as an example of a partial optimization. In Figure 1a the *tert*-butyl cation is planar, as it is expected to be in isolation; after optimizing this molecule in the presence of CO, the *tert*-butyl cation is bent, as expected for a *tert*-butyl acylium cation, shown in Figure 1b. This system could never reach the global minimum, the *tert*-butyl acylium cation, if the *tert*-butyl cation were not allowed to respond to the presence of the carbon monoxide molecule. This type of transformation would be difficult to achieve using only intramolecular moves because it involves the concerted movement of multiple atoms.

The second consequence of a partial optimization move is that the optimized molecule undergoes translation and rotation relative to the frozen molecules during the course of the optimization process. This movement leads to refinement of the structure of the entire system and normally moves the system closer to a minimum energy configuration. Since the most significant reduction in the energy and change in the geometry will occur at the beginning of the optimization, the number of optimization steps can be restricted, usually to fewer than 10 steps, to decrease the computational cost. Partial optimization moves should move a system quickly toward a local minimum.

Moving efficiently toward a local minimum is only one part of finding the global minimum for a system. The other important aspect is the ability to move across energy barriers to explore minima other than the closest one. A novel way to achieve these large displacements in phase space is through molecular interchange moves. These moves swap the center of mass of two molecules within the same intermolecular system, unlike a similar move implemented in Gibbs-ensemble Monte Carlo that exchanges two molecules between simulation boxes.<sup>20</sup> As shown in Figure 2a,b, molecular interchange moves can significantly change the overall configuration of a system in one move. Interchanging the ethylene glycol and water molecules in Figure 2a results in a completely different configuration of the system where new interactions are present, shown in Figure 2b. This transfor-



**Figure 2.** A molecular interchange move swaps the positions of the ethylene glycol and water molecules so that the configuration of the system is different before (a) and after (b) the move. Carbon atoms are shown in gray, oxygen atoms in black, and hydrogen atoms in white.

mation would be difficult to generate using only translation and rotation moves. As might be expected, these moves are often rejected because the molecules have not adjusted to their new environments in any way. Therefore, a molecular interchange move is always followed by a small number of full optimization steps to remove the initial strain in the system and increase the probability of accepting the interchange move.

Both partial optimization and molecular interchange moves are useful in searching for minimum energy configurations. One limitation of both types of moves is that they do not satisfy microscopic reversibility since there is no way to return to the previous state. This limitation does not affect the results of the method as presently implemented because the goal is to find low energy configurations. Maintaining microscopic reversibility is only important if one is collecting statistics and computing ensemble averages, and this method cannot practically achieve the number of iterations needed to acquire good statistics at present. One could restore microscopic reversibility by constructing a reverse move which uses a saved version of the configuration prior to an optimization. This problem has not been addressed to date as it is not a critical issue with respect to the current goals.

**Energy Program Interface.** FPMC currently interfaces with Gaussian 98<sup>21</sup> to evaluate the electronic energy of a given molecular configuration and carry out partial optimizations. This interface consists of setting up calculations, running Gaussian, and evaluating the output. These tasks are fairly straightforward, and FPMC can be modified to interface with any of the common ab initio packages with only minor modifications.

All of the calculations in this study were carried out using density functional theory. Becke's<sup>22</sup> exchange functional and Lee, Yang, and Parr's<sup>23</sup> correlation functional (BLYP) were used so that validation against previous theoretical calculations could be carried out. A 6-31G\*\* basis set was used for all atoms because it offers a good compromise between accuracy and computational speed. Conventional optimizations were carried out for comparison to the FPMC results and used the GDIIS algorithm<sup>24,25</sup> to aid convergence. FPMC partial or full optimizations used the default optimization parameters.

**Acceptance of Moves.** The acceptance/rejection of FPMC moves is based on the Boltzmann factor of the energy difference between the new and old configurations, where the energy used in the Boltzmann factor is the electronic energy of the molecule without inclusion of the zero point vibrational energy. The extra cost associated with calculating the vibrational frequencies makes this impractical. The primary drawback of using the electronic energy rather than the free energy when deciding whether to accept or reject

**Table 1.** Intermolecular B–F Distances (Å) in BF<sub>3</sub>–xHF Complexes<sup>28</sup>

basis set	BF <sub>3</sub> –HF	BF <sub>3</sub> –2HF
6-31G**	2.327	
6-31+G**	2.645	2.430
6-311G**	2.447	2.165
6-311++G**	2.681	2.491

moves is that the simulation will eventually converge to minimum electronic energy positions instead of the more thermodynamically relevant free energy minimum. This is analogous to the approach used in many quantum chemical methods which are aimed at locating minima on the electronic energy surface.

**Simulated Annealing.** The final approach that FPMC uses to find minimum energy structures is simulated annealing,<sup>26</sup> whereby the temperature used in the Boltzmann factor is lowered throughout the calculation. While this method does not ensure that a global minimum is found, it is generally accepted as a means to help find the global minimum. The cooling rate utilized by FPMC is identical to that used by Keshari and Ishikawa:<sup>12</sup>

$$T_{\text{new}} = \gamma T_{\text{old}} \quad (1)$$

In addition, the maximum displacement/rotation,  $\Delta R_{\text{max}}$ , during translation/rotation is lowered to reduce the number of rejected steps at low temperatures without sacrificing large, exploratory moves at higher temperatures. Again, the schedule used by Keshari and Ishikawa<sup>12</sup> is employed:

$$\Delta R_{\text{max,new}} = \gamma^{1/2} \Delta R_{\text{max,old}} \quad (2)$$

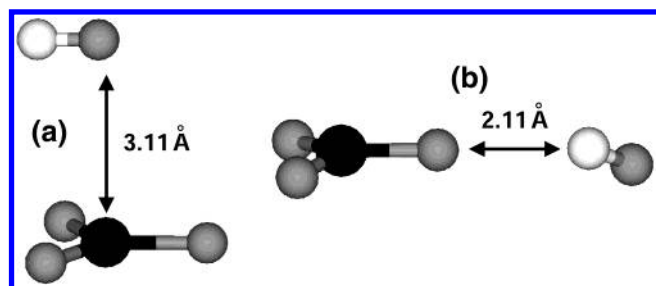
In both eqs 1 and 2,  $\gamma$  is a user-defined parameter which controls the rate at which both temperature and displacements are lowered.

**Termination Criteria.** There are a number of termination criteria which can halt a FPMC calculation. One criterion is that the temperature drops below the minimum temperature, indicating the end of a simulated annealing run. Other criteria include exceeding a user-defined maximum number of attempted steps or a maximum number of consecutive rejected steps. Finally, if a partial geometry optimization converges to a minimum energy structure, the calculation is terminated.

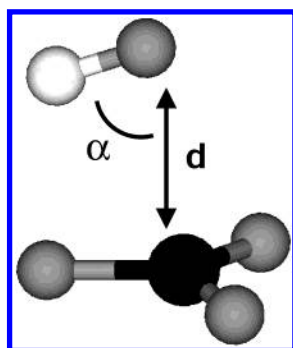
## RESULTS AND DISCUSSION

**Validation.** The first application of FPMC was the study of the BF<sub>3</sub>–HF and BF<sub>3</sub>–2HF systems. These systems include boron, an element missing from most force fields, and serve as validation for the method as they have been studied both experimentally and theoretically. Rotational spectroscopy has been carried out on the BF<sub>3</sub>–HF system<sup>27</sup> and has shown that the fluorine atom of the HF molecule is near the C<sub>3</sub> axis of the BF<sub>3</sub> molecule at a distance of 2.54 Å with the proton lying off axis. Rapid vibrational averaging leads to a structure which is effectively a symmetric top. Density functional theory calculations have found structures which agree with this result. The intermolecular B–F distance changes with the basis set employed, as shown in Table 1, but all calculations showed the correct overall structure as described by the experimental investigation.





**Figure 3.** Both accurate (a) and poor (b) starting geometries were used in FPMC and optimization calculations for  $\text{BF}_3\text{-2HF}$ . Boron atoms are shown in black, fluorine atoms in gray, and hydrogen atoms in white.



**Figure 4.** Configuration of  $\text{BF}_3\text{-HF}$  system with intermolecular B-F distances ( $d$ ) and B-F-H angles ( $\alpha$ ) listed in Table 3.

**Table 2.** FPMC Parameters Used in Calculations Involving  $\text{BF}_3$

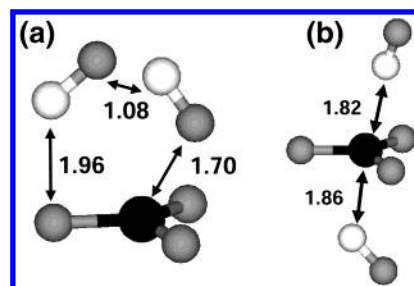
parameter	$\text{BF}_3\text{-HF}$		$\text{BF}_3\text{-2HF}$	
	accurate	poor	accurate	poor
initial guess	accurate	poor	accurate	poor
max translation (Å)	0.2	0.4	1.0	1.0
max rotation (degrees)	10	15	30	90
optimization %	0	15	15	20
swap %	0	0	0	15
initial temp (K)	600	1000	1000	1000
cooling rate	1.0	0.97	0.97	0.97

**Table 3.** FPMC and Optimization Results for  $\text{BF}_3\text{-HF}^a$

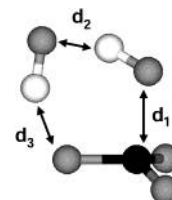
calculation	$d$ (Å)	$\alpha$ (degrees)	rel energy (kcal/mol)
direct optimization	2.33	83.1	0.00
FPMC (accurate initial guess)	2.29	85.8	0.04
FPMC (poor initial guess)	2.36	79.5	0.17

<sup>a</sup> Note:  $d$  and  $\alpha$  are defined in Figure 4.

FPMC calculations were carried out on this system using both accurate and poor starting geometries, shown respectively in Figure 3a,b. The controlling parameters for these two calculations are shown in Table 2. The geometry calculated by FPMC for both starting geometries is shown in Figure 4. The specific geometrical parameters and electronic energies relative to the electronic energy of the optimum geometry for all of the calculations on this system are listed in Table 3. The intermolecular B-F distances of 2.29 Å and 2.36 Å compare favorably to the optimum value of 2.33 Å for this system. The slight variance in the intermolecular B-F distance and the B-F-H angle for the FPMC calculations compared to the optimum values and the extremely small energy difference between their energies indicate that the optimum structure for this system lies on a flat area of the potential energy surface. These results demonstrate that FPMC can locate low energy configurations



**Figure 5.** Both accurate (a) and poor (b) starting geometries were used in FPMC and optimization calculations.



**Figure 6.** Optimal configuration of  $\text{BF}_3\text{-2HF}$  system with intermolecular distances listed in Table 4.

**Table 4.** FPMC and Optimization Results for  $\text{BF}_3\text{-2HF}^a$

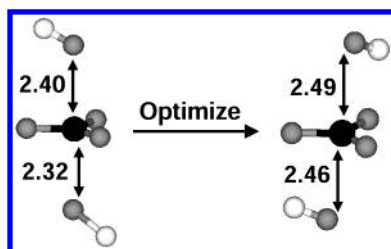
calculation (initial guess)	$d_1$ (Å)	$d_2$ (Å)	$d_3$ (Å)	rel energy (kcal/mol)
FPMC (accurate)	1.89	1.48	1.70	0.12
FPMC (poor)	1.87	1.51	1.64	0.09
direct optimization (accurate)	1.87	1.50	1.64	0.00
direct optimization (poor)	unable to converge optimization			

<sup>a</sup> Note:  $d_1$ ,  $d_2$ , and  $d_3$  are defined in Figure 6.

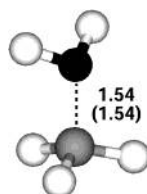
given either accurate or poor initial guesses. This is a very simple system, of course, so additional validation calculations were performed using  $\text{BF}_3\text{-2HF}$ , as previous theoretical calculations were also performed on this system.<sup>28</sup>

As in the case of  $\text{BF}_3\text{-HF}$ , both accurate and poor initial guesses for  $\text{BF}_3\text{-2HF}$  were tested to confirm that FPMC could find good geometries even if the initial guess is poor. Figure 5a,b shows the starting geometries for the poor and accurate guesses, respectively. Figure 6 and Table 4 outline the geometry determined by FPMC for these species. In both cases, FPMC found low energy structures resembling a ring-like structure, which represents excellent agreement with previous theoretical calculations. Direct optimization of the accurate initial guess also resulted in the ring-like structure, but direct optimization of the poor initial guess proved difficult. Optimizations using very small step sizes and involving direct calculation of the Hessian matrix at each step failed to converge to optimum geometries. A modified poor starting guess, where the fluorine atoms from both HF molecules were directed toward the boron, led to a geometry which was 10.5 kcal/mol higher in energy than the ring-like structure. The modified starting geometry and resulting minimum energy structure are shown in Figure 7. This case illustrates the utility of a Monte Carlo search scheme in reducing the significance of the initial guess, which is important for clusters containing many entities where the number of initial configurations increases dramatically.

**Discrete Solvation.** The motivation for developing FPMC was to aid in the placement of discrete solvent molecules. The solvation of a methyl cation by water was examined with both FPMC and a conventional optimization method.



**Figure 7.** A converged geometry was not found upon optimization of the poor FPMC starting geometry in the  $\text{BF}_3\text{--}2\text{HF}$  system. Instead, a slightly modified, but still poor, geometry was optimized and located the minimum shown, which was 10.5 kcal/mol higher in energy than the optimal geometry shown in Figure 6.



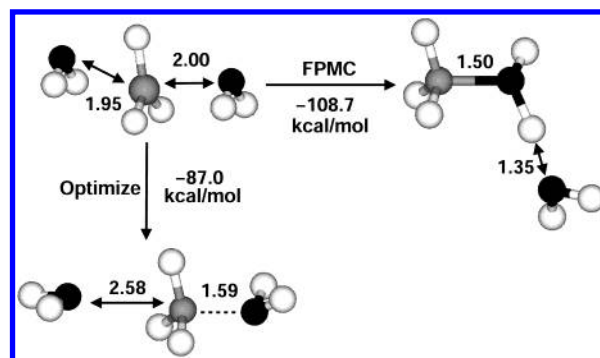
**Figure 8.** Configuration of the  $\text{CH}_3^+\text{--H}_2\text{O}$  system is shown with the C–O distance determined by FPMC indicated and the optimized C–O distance shown in parentheses. The FPMC result is only 0.04 kcal/mol higher than the electronic energy of the optimized structure. The carbon atom is shown in gray, the oxygen atom is shown in black, and the hydrogen atoms are shown in white.

This system was chosen because previous researchers<sup>29</sup> have studied this system and provided geometric information for up to four water molecules. Using FPMC, a series of calculations with one to four solvating water molecules was performed.

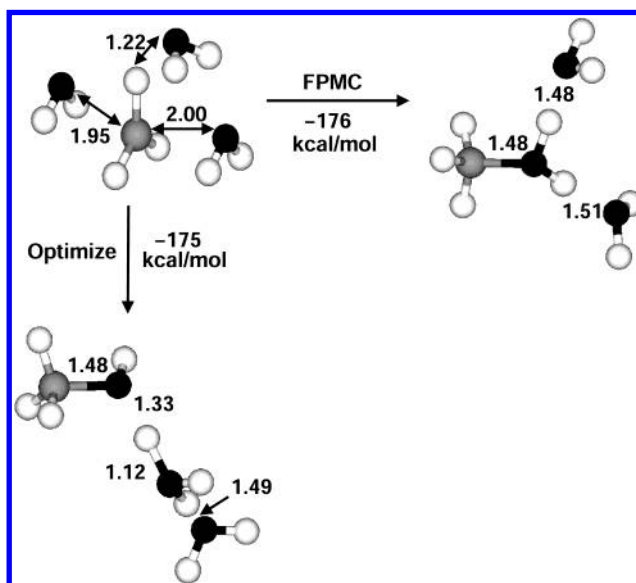
First, a single water molecule was used to solvate the methyl cation. Both optimization and FPMC calculations were run using the same starting geometry and converged to the same final configuration, shown in Figure 8. The FPMC electronic energy was only 0.04 kcal/mol higher than the optimized energy. This result is in agreement with previous studies which found that this system optimizes to form a protonated methanol molecule.

As additional water molecules are included, the number of solvation sites also increases. A second water molecule likely occupies one of two solvation sites, either opposite the first water molecule on the other side of the methyl cation, or in a hydrogen bonding position with the first water. Starting from the first proposed structure, conventional optimization leads to a minimum energy configuration with the same basic structure as the initial guess, but the FPMC method finds the second proposed site, as shown in Figure 9. The hydrogen bonding configuration is clearly lower in energy (by 21.7 kcal/mol), and this calculation again demonstrates the ability of FPMC to reduce the importance of the starting geometry. Optimization of the hydrogen bonding configuration leads to a structure with an electronic energy only 0.13 kcal/mol lower than the FPMC result, indicating that FPMC has already done an excellent job of refining the structure. These results also agree with the previous study which found that the energy of the system was lower when the second water molecule is hydrogen bonded with the first water molecule.

Continuing the series, a third water molecule would likely occupy one of three possible positions: opposite the first two water molecules on the opposite side of the methyl

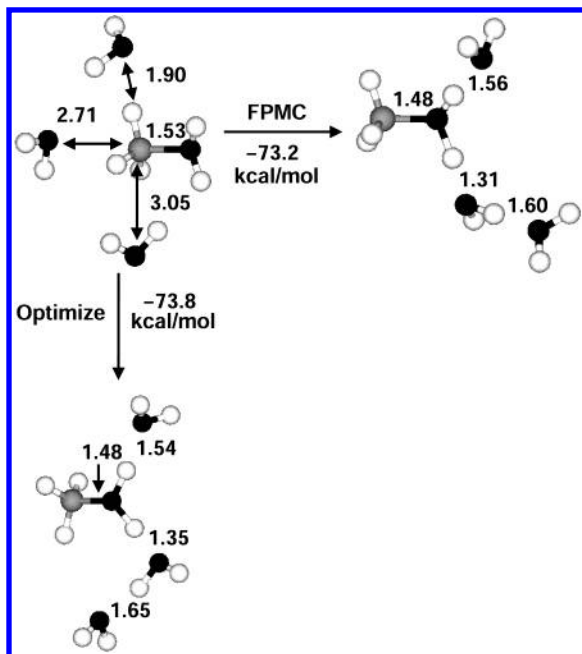


**Figure 9.** Configuration of the  $\text{CH}_3^+\text{--}2\text{H}_2\text{O}$  system is shown. Beginning with the same starting geometry, FPMC and optimization lead to very different system configurations, with FPMC finding a geometry of much lower electronic energy.



**Figure 10.** Configuration of the  $\text{CH}_3^+\text{--}3\text{H}_2\text{O}$  system is shown. Beginning with the same starting geometry, FPMC and optimization lead to very different system configurations, with FPMC finding a geometry that has a slightly lower electronic energy.

cation, a hydrogen bonding position with the first water molecule, or a hydrogen bonding position with the second water molecule. To test the ability of FPMC to overcome poor initial guesses, none of these likely positions were used as the starting geometry. Instead, the three water molecules were positioned such that a water molecule was placed on each side of the methyl cation and the third water molecule was placed above the methyl cation. The starting geometry and the resulting geometries for both optimization and FPMC calculations are shown in Figure 10. While the optimization of the two-water system retained the two water molecules on opposite sides, optimization of the three-water system resulted in migration of the second water molecule around the methyl cation. It is likely that interaction between the second and third water molecules is responsible for this difference. The results in Figure 10 indicate that the global minimum involves the third water molecule hydrogen bonding to the first water molecule, although the minimum involving hydrogen bonding between the second and third water molecules is only slightly higher in energy. Subsequent optimization of the FPMC result leads to a decrease in the electronic energy of only 0.94 kcal/mol and confirms that this is the lower energy configuration. The previous study

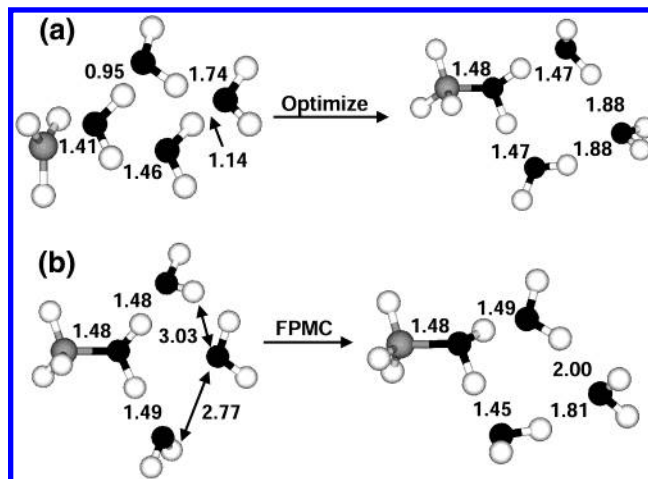


**Figure 11.** Configuration of the  $\text{CH}_3^+-4\text{H}_2\text{O}$  system is shown. Beginning with the same starting geometry, FPMC and optimization lead to similar configurations, with FPMC finding a geometry that has a slightly higher electronic energy.

does not provide enough information to identify the preferred configuration of three water molecules. When coupled with the results for the methyl cation with two solvating water molecules, these results indicate that the methyl cation in aqueous solution prefers being solvated on one side.

Finally, a fourth water molecule was added to the solvated methyl cation. Again, a poor initial guess was used as the starting geometry for both a FPMC run and a geometry optimization, and the results of these calculations are shown in Figure 11. Both the optimization and the FPMC method lead to the same overall configuration, which resembles the three-water configuration with the fourth water molecule hydrogen bonded to one of the perimeter water molecules. Additional FPMC runs on this system located other structures which were also based on the three-water configuration with the fourth water hydrogen bonded in various configurations. All of these structures had similar energies, indicating that the energy of the system is weakly dependent on the hydrogen bonding of the fourth water molecule. This can also be seen in Figure 11 where the FPMC and conventional optimization calculations located two different configurations that only differed in energy by 0.6 kcal/mol, at 73.2 kcal/mol and 73.8 kcal/mol below the starting geometry, respectively.

Additional FPMC and optimization calculations with starting geometries different than that in Figure 11 were also carried out. Many of these calculations resulted in structures uncovered using the starting geometry in Figure 11, but both FPMC and conventional optimization calculations located a ring-like configuration which is lower in energy than the structures in Figure 11. Starting and ending geometries for these calculations are shown in Figure 12. If these results are referenced to the starting geometry in Figure 11, the conventional optimization led to a structure 74.4 kcal/mol lower in energy, and the FPMC calculation led to a structure 75.2 kcal/mol lower, which makes it the lowest energy structure found for the four-water system. The results shown



**Figure 12.** Starting and optimized geometries for both FPMC and conventional optimization which lead to a low energy, ring-like configuration. The electronic energy of the conventional optimization was 74.4 kcal/mol lower than the starting geometry in Figure 11, and the FPMC geometry was the lowest energy structure found at 75.2 kcal/mol lower.

in Figures 11 and 12 further suggest that the methyl cation prefers to be solvated on one side. They also indicate that a number of nearly energetically equivalent positions exist for the placement of the fourth water molecule, as evidenced by the number of local minima with energies close to that of the global minimum. It is anticipated that FPMC would have located all geometries seen in Figures 11 and 12 if more complete sampling were within the current computational limits of the method. It is believed that these are the first calculations carried out on this system.

#### 4. CONCLUSIONS

A number of methods exist to account for the presence of solvent molecules in a molecular model, but the direct inclusion of discrete solvent molecules is necessary to uncover geometrical details about solvated molecules. The accurate placement of solvent molecules around a solute in a computational investigation presents two challenges. The first is the difficulty in finding the configuration which minimizes the energy when many local minima exist. Conventional Monte Carlo methods aid in finding these structures but are limited by the availability of force fields. The second difficulty is that the computational demand greatly increases when many solvent molecules are included in ab initio calculations. A first principles Monte Carlo (FPMC) method has been developed to address these issues. Novel Monte Carlo moves, namely partial optimization and molecular interchange followed by full optimization, have been introduced to efficiently locate minimum energy structures, thereby reducing the number of required energy evaluations. FPMC calculations have been compared to optimization results for  $\text{BF}_3\text{--HF}$ ,  $\text{BF}_3\text{--}2\text{HF}$ , and methyl cation solvated by one to four water molecules. It has been demonstrated that FPMC can find minimum energy configurations in each of these cases, even when provided with a very poor initial guess. The primary limitation of the method at this point is the relatively small number of energy evaluations that are practically achievable, but the size of the systems accessible by the method will increase as computing power continues to grow.



## REFERENCES AND NOTES

- (1) Guthrie, J. P. A Group Equivalents Scheme for Free Energies of Formation of Organic Compounds in Aqueous Solution. *Can. J. Chem.* **1992**, *70*, 1042–1054.
- (2) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group Contributions to the Thermodynamic Properties of Non-Ionic Organic Solutes in Dilute Aqueous Solution. *J. Solution Chem.* **1981**, *10*(8), 563–595.
- (3) Tomasi, J.; Persico, M. Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994**, *94*(7), 2027–2094.
- (4) Pliego, J. R.; Riveros, J. M. The Cluster-Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species. *J. Phys. Chem. A* **2001**, *105*(30), 7241–7247.
- (5) Saunders, M. Stochastic Exploration of Molecular Mechanics Energy Surfaces. Hunting for the Global Minimum. *J. Am. Chem. Soc.* **1987**, *109*(10), 3150–3152.
- (6) Li, Z.; Scheraga, H. A. Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611–6615.
- (7) Ferguson, D. M.; Raber, D. J. A New Approach to Probing Conformational Space with Molecular Mechanics: Random Incremental Pulse Search. *J. Am. Chem. Soc.* **1989**, *111*(12), 4371–4378.
- (8) Robertson, D. H.; Brown, F. B. Determination of the Structure of Mixed Argon–Xenon Clusters Using a Finite-Temperature, Lattice-Based Monte Carlo Method. *J. Chem. Phys.* **1989**, *90*(6), 3221–3229.
- (9) Levy, R. M.; Gallicchio, E. Computer Simulations with Explicit Solvent: Recent Progress in the Thermodynamic Decomposition of Free Energies and in Modeling Electrostatic Effects. *Annu. Rev. Phys. Chem.* **1998**, *49*, 531–567.
- (10) Börjesson, U.; Hünenberger, P. H. Explicit-Solvent Molecular Dynamics Simulation at Constant pH: Methodology and Application to Small Amines. *J. Chem. Phys.* **2001**, *114*(22), 9706–9719.
- (11) Maseras, F.; Morokuma, K. IMOMM: A New Integrated Ab Initio + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* **1995**, *16*, 1170–1179.
- (12) Keshari, V.; Ishikawa, Y. Ab Initio Monte Carlo Simulated Annealing Method. *Chem. Phys. Lett.* **1994a**, *218*, 406–412.
- (13) Keshari, V.; Ishikawa, Y. First-Principles Monte Carlo Simulated Annealing Study of the Structures and Properties of Hydrogenated Lithium Clusters. *Int. J. Quantum Chem.* **1994b**, *Suppl.* 28, 541–551.
- (14) Vaughn, S. J.; Akhmatskaya, E. V.; Vincent, M. A.; Masters, A. J.; Hillier, I. H. Monte Carlo Simulation of  $F^-(H_2O)_4$  Using an Ab Initio Potential. *J. Chem. Phys.* **1999**, *110*(9), 4338–4346.
- (15) Bandyopadhyay, P.; Ten-No, S.; Iwata, S. Ab Initio Monte Carlo Simulation Using Multicanonical Algorithm: Temperature Dependence of the Average Structure of Water Dimer. *Mol. Phys.* **1999**, *96*(3), 349–358.
- (16) Bacelo, D. E.; Binning, R. C., Jr.; Ishikawa, Y. Ab Initio Monte Carlo Simulated Annealing Study of  $HCl(H_2O)_n$  ( $n = 3, 4$ ) Clusters. *J. Phys. Chem. A* **1999**, *103*(24), 4631–4640.
- (17) Bacelo, D. E.; Ishikawa, Y. Theoretical Study of Microscopic Solvation of LiCl in Water Clusters:  $LiCl(H_2O)_n$  ( $n = 1–4$ ). *Chem. Phys. Lett.* **2000**, *319*, 679–686.
- (18) Siepmann, J. I.; Frenkel, D. Configurational Bias Monte Carlo—A New Sampling Scheme for Flexible Chains. *Mol. Phys.* **1992**, *75*, 59–70.
- (19) Dodd, L. R.; Boone, T. D.; Theodorou, D. N. A Concerted Rotation Algorithm for Atomistic Monte Carlo Simulation of Polymer Melts. *Mol. Phys.* **1993**, *78*, 961–996.
- (20) Martin, M. G.; Siepmann, J. I. Predicting Multicomponent Phase Equilibria and Free Energies of Transfer for Alkanes by Molecular Simulation. *J. Am. Chem. Soc.* **1997**, *119*, 8921–8924.
- (21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7.; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (22) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*(6), 3098–3100.
- (23) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*(2), 785–789.
- (24) Csaász, P.; Pulay, P. Geometry Optimization by Direct Inversion in the Iterative Subspace. *J. Mol. Struct. (THEOCHEM)* **1984**, *114*, 31–34.
- (25) Farkas, O.; Schlegel, H. B. Methods for Optimizing Large Molecules. I. An  $O(N^2)$  Algorithm for Solving Systems of Linear Equations for Coordinate Transformations in Quantum Chemical Geometry Optimization. *J. Chem. Phys.* **1998**, *109*, 1700–1704.
- (26) Kirkpatrick, S.; Gelatt, C. D., Jr.; Vecchi, M. P. Optimization by Simulated Annealing. *Science* **1983**, *220*(4598), 671–680.
- (27) Phillips, J. A.; Canagaratna, M.; Goodfriend, H.; Grushow, A.; Almlöf, J.; Leopold, K. R. Microwave and Ab Initio Investigation of  $HF-BF_3$ . *J. Am. Chem. Soc.* **1995**, *117*(50), 12549–12556.
- (28) Kim, D.; Klein, M. L. Ab Initio Study of  $BF_3 + (HF)_{1-7}$  Clusters. *Chem. Phys. Lett.* **1999**, *308*, 235–241.
- (29) Kazansky, V. B.; Senchenya, I. N.; Pankov, A. A. On the Nature of Hydrated Aliphatic Carbenium Ions and on Their Role in Homogeneous Acid Catalysis. *J. Mol. Catal.* **1991**, *70*(2), 189–195.

CI0202887