9 November 2001

Chemical Physics Letters 348 (2001) 350–356

CHEMICAL PHYSICS LETTERS

www.elsevier.com/locate/cplett

Transition-state optimization by the free energy gradient method: Application to aqueous-phase Menshutkin reaction between ammonia and methyl chloride

Hajime Hirao, Yukihiko Nagae, Masataka Nagaoka *

Graduate School of Human Informatics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

Received 4 June 2001; in final form 27 August 2001

Abstract

The transition state (TS) for the Menshutkin reaction $H_3N + CH_3Cl \rightarrow H_3NCH_3^+ + Cl^-$ in aqueous solution was located on the free energy surface (FES) by the free energy gradient (FEG) method. The solute–solvent system was described by a hybrid quantum mechanical and molecular mechanical (QM/MM) method. The reaction path in water was found to deviate largely from that in the gas phase. It was concluded that, in such a reaction including charge separation, TS structure optimization on an FES is inevitable for obtaining valid information about a TS in solution. © 2001 Published by Elsevier Science B.V.

1. Introduction

Transition state (TS) is an important concept in chemical reactions [1,2]. It is closely related to reaction rates, providing valuable insights into molecular reactivities. This has prompted quantum chemical TS structure optimizations for isolated molecular systems including complex organometallic reactions for these decades [2]. However, for reactions in solution, since the bond breaking and forming processes have to be dealt with, and moreover, the Avogadro number of molecules are essentially taken into consideration, it becomes a very difficult task to execute geometry optimizations with the chemical accuracy for not only TSs but also stable states (SSs). Although several gradient methods for treating solvated molecules have been developed [3,4], they are based on mean field approximations and the structures of solvent molecules are described only effectively. Since:

1. local solute–solvent interactions,
2. microscopic entropy contributions,
3. non-equilibrium solvent dynamics

are of crucial importance in solution, it is urgently desirable to establish such a method that treats explicit solvent structures. Monte Carlo (MC) and molecular dynamics (MD) simulations [5], which can reproduce the liquid behavior on computers, are powerful methods satisfying the above requirements. In addition, an empirical valence bond (EVB) [6] and a hybrid quantum mechanical and

* Corresponding author. Fax: +81-52-789-5623.
E-mail address: mnagaoka@info.human.nagoya-u.ac.jp (M. Nagaoka).

0009-2614/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S0009-2614(01)01131-9

molecular mechanical (QM/MM) potential [7,8] have allowed the simulations to model solution chemical reactions theoretically. For example, Gao and Xia [9] constructed a *two-dimensional* free -energy surface (FES) for the Menshutkin reaction $H_3N + CH_3Cl \rightarrow H_3NCH_3^+ + Cl^-$ in aqueous solution by MC–QM/MM simulations, thereby predicting an early TS in accord with the Hammond postulate [10]. However, when a reaction is asymmetric, an FES with more than two dimensions is necessary, and accordingly, application of such an MC strategy seems to be computationally impractical.

On the other hand, Okuyama-Yoshida et al. [11] proposed and have recently developed the free energy gradient (FEG) method [12–14] for locating stationary optimal structures of molecules on a *multi-dimensional* FES. Although the method necessitates the gradient on a multi-dimensional FES, its negative is nothing but the force vector and can be calculated as such an average quantity that a solute molecule with a fixed structure might feel in solution. Numerically, such a gradient can be generated through an equilibrium MD simulation assuming solvent molecules would be equilibrated prior to the structural change of a solute molecule. The assumption is a kind of adiabatic approximation and enables us to locate SS and TS structures efficiently on an FES. It is emphasized that the manner of optimization is different from a grid search method [9] that requires calculations for unimportant regions as well as for reactive ones. Recently, the stable zwitterionic structure of glycine in water has been optimized using the FEG method with the EVB potential [14]. However, it should be noted that the EVB method requires a much more laborious procedure than the QM/MM method in the parameter fitting. Furthermore, in the latter, the effect of solvent upon quantum mechanical charge redistribution of solute can be taken into account in a self-consistent way [7,8].

Under these circumstances, we have been developing the FEG method combined with a QM/MM potential, which should become, in the next decade, an important and essential step to understand not only chemical reactivities in solution but also some hidden roles which might be microscopic or dynamic. In this Letter, we introduced the FEG method combined with the AM1/TIP3P potential and applied it, for the first time, to the TS structure optimization for the title Menshutkin reaction [9,15–20].

In Section 2, the computational method is explained concisely, followed by the results and discussion in Section 3.

2. Computational method

We used a QM/MM method to describe the solution system for the purpose of including the solvent structures explicitly. The reactant molecules, i.e., H_3N and CH_3Cl were considered the QM portion and described by the AM1 Hamiltonian [21], while the solvent water molecules were treated molecular mechanically by the TIP3P model [22]. For the Lennard–Jones-type interaction between QM and MM atoms, we used the parameters developed by Gao and Xia specifically for the reaction [9]. The core–core interaction energy between QM and MM atoms was evaluated by the method of Cummins and Gready [23], in which some additional parameters [7] such as the Ohno–Kloppman factors are set to zero.

MD calculations were carried out for a system containing a couple of reactant molecules and 445 water molecules in a rectangular box ($22.3 \times 22.3 \times 27.0$ Å^3) with the periodic boundary condition at 300 K. The temperature was kept constant with the Nosé-Hoover chain algorithm [24] and the system was maintained to be a canonical (NVT) ensemble. As a result, the mass density in the box was 1.0 g/cm^3. The SHAKE/RATTLE algorithms were used for the geometry constraints [25,26], and the non-bonded cutoff distance of 10.0 Å was employed. After equilibration MD run for 10 ps, sampling run was executed for 10 ps with a time step of 0.5 fs. All the MD calculations were performed with the AM1/TIP3P potential using ROAR 2.0 [27–29] partly modified for the present purpose.

According to the FEG method [11], a force vector $\mathbf{F}^{\mathrm{FE}}$ on an FES was obtained by

$$\mathbf{F}^{\mathrm{FE}}(\mathbf{q}^{\mathrm{s}}) = -\frac{\partial G(\mathbf{q}^{\mathrm{s}})}{\partial \mathbf{q}^{\mathrm{s}}} = -\left\langle \frac{\partial V_{\mathrm{RS}}(\mathbf{q}^{\mathrm{s}})}{\partial \mathbf{q}^{\mathrm{s}}} \right\rangle, \tag{1}$$

where $\mathbf{q}^s$ denotes the solute coordinates, $G(\mathbf{q}^s)$ is the free energy function, $V_{RS}(\mathbf{q}^s)$ is the sum of the solute potential energy and the solute–solvent interaction energies, and $\langle\cdots\rangle$ represents an equilibrium ensemble average. In the QM/MM formalism, V_{RS} can be represented as

$$V_{RS} = \langle \Psi | \hat{H}_{QM} + \hat{H}_{QM/MM} | \Psi \rangle, \quad (2)$$

where Ψ denotes the SCF wavefunction in solution and $\hat{H}_{QM}$ and $\hat{H}_{QM/MM}$ are the Hamiltonians for the QM part and the solute–solvent interaction energy, respectively.

3. Results and discussion

At first, we calculated the gas-phase reaction path at the AM1 level with GAUSSIAN 98 [30] through the following procedure:

1. the TS geometry optimization in the gas phase,
2. on the reactant side, the optimization of the structural parameters except N–C distance R_{N-C}, ranging in 0.05 Å increments from 1.70 to 3.20 Å,
3. on the product side, the optimization of the structural parameters except C–Cl distance R_{C-Cl}, ranging in 0.05 Å increments from 2.30 to 3.40 Å.

The reaction coordinate was defined by $R_{C-Cl} - R_{N-C}$. Here the free energy profile in the gas phase is assumed to be the same as the enthalpy of formation profile. The free energy profile for the reaction in water along the gas-phase structural change obtained above was calculated by statistical perturbation theory (SPT) [31]. The resultant free energy profiles in the gas phase and in water are shown in Fig. 1a. Fig. 1b represents the free energy profile in the vicinity of the top of free energy barrier in aqueous solution in Fig. 1a, where the intervals of N–C in partial geometry optimizations were reduced to 0.01 Å.

The free energy of activation and reaction in aqueous solution were calculated to be 31.7 and −18.7 kcal/mol, respectively. Although the latter is in good agreement with the value of Gao and Xia (−18 kcal/mol), the former is higher than their value (26.3 kcal/mol) by ca. 5.4 kcal/mol. This indicates that the reaction path in the gas phase should not be substituted directly for that in aqueous solution.

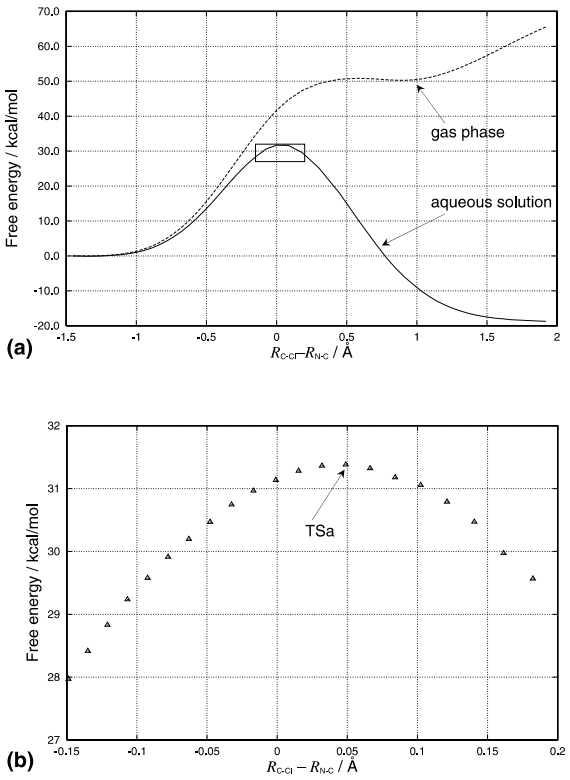

Fig. 1. (a) Free energy profiles in the gas phase and in aqueous solution along the gas-phase reaction coordinate $R_{C-Cl} - R_{N-C}$. (b) Detailed free energy change around the top of the free energy curve in water (rectangular region in Fig. 1a).

Further, for comparison, we performed the IRC calculation [32] to obtain the steepest-descent path from the TS in the gas phase and obtained the free energy changes in the gas phase and in aqueous solution along the whole IRC (Fig. 2a) and around the top of the barrier in solution (Fig. 2b). The intervals of the IRC were taken to be $\sim$0.12 amu$^{1/2}$bohr except the region around the TS in solution, in which the intervals were as short as $\sim$0.03 amu$^{1/2}$bohr.

The free energy of activation obtained by this procedure was 27.6 kcal/mol, whereas the reaction free energy could not be calculated because of the limited range that the IRC calculation can trace. Those structures that correspond to the maxima of the free energy curves in Figs. 1b and 2b are hereafter referred to as **TSa** and **TSb**, respectively.

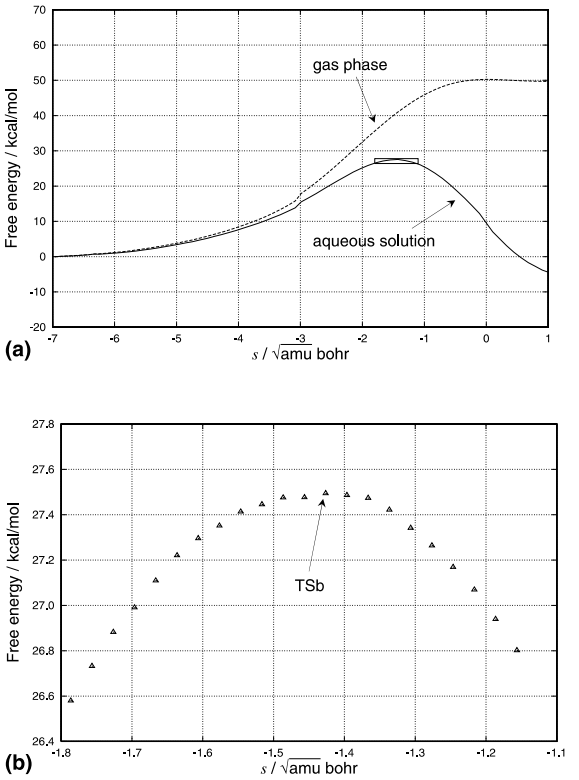

Fig. 2. (a) Free energy profiles in the gas phase and in aqueous solution along the gas-phase IRC. (b) Detailed free energy change around the top of the free energy curve in water (rectangular region in Fig. 2a).

Next, for the TS structure optimization in solution, we chose **TSb** as the initial geometry and then started the optimization in a similar way to the steepest-descent-path following procedure. At each optimization step, the atomic positions were updated along the direction of the average force vector $\mathbf{F}^{\mathrm{FE}}(\mathbf{q}^{s})$ by the displacement vector that was calculated by multiplying $\mathbf{F}^{\mathrm{FE}}(\mathbf{q}^{s})$ by a scaling constant. During the optimization, if the componential force acting on N, C and Cl atoms remain, their signs were changed. Despite the simple scheme adopted here, the minimum root mean square (RMS) force resulted in 0.010 hartree/bohr, which is smaller than 0.025 hartree/bohr in [14]. Thus, the TS structure (**TSc**) was fully optimized. The geometrical parameters for the TS in the gas phase, **TSa**, **TSb**, and **TSc** are summarized in Fig. 3.

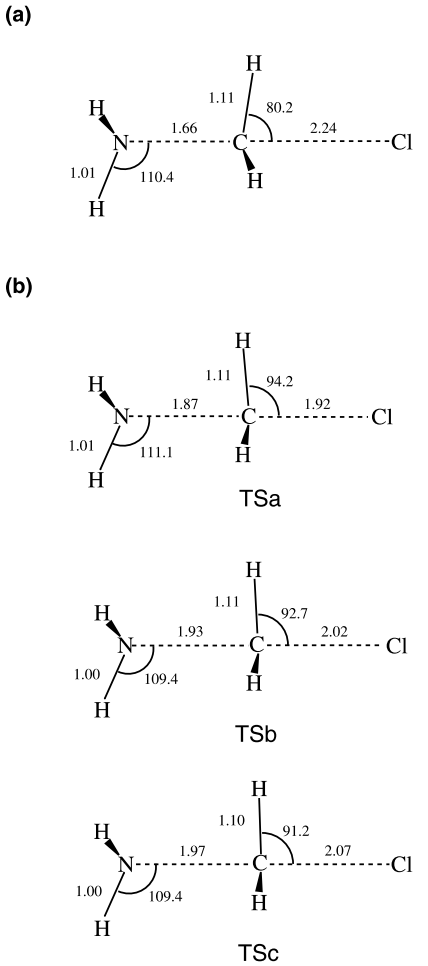

Fig. 3. Transition-state structures (a) in the gas phase and (b) in aqueous solution. See text for details. Bond lengths are in angstroms and bond angles in degrees.

All TSs obtained are earlier than the TS in the gas phase, which is in accord with the Hammond postulate. As is the case in **TSa** or **TSb**, **TSc** shows the C_{3v} symmetry with the optimized values of $R_{\mathrm{N-C}}$ and $R_{\mathrm{C-Cl}}$ are 1.97 and 2.07 Å, respectively. They are almost the same values as those obtained by Gao and Xia [9], which were 1.96 and 2.09 Å, respectively. However, it can be recognized in Fig. 3 that the N–C bond lengths for **TSa** and **TSb** are shorter than $R_{\mathrm{N-C}}$ for **TSc** by 0.10 and 0.04 Å, respectively, while the C–Cl distances are shorter by 0.15 and 0.05 Å, respectively. **TSc** is the earliest among the three **TSs** in solution, while **TSa** and

TSb, which reflect the gas-phase characters, are closer to the TS in the gas phase than **TSc**. One can understand therefore that, in aqueous solution, the reaction should proceed accompanied with such a structural deformation to enhance the charge separation and, inversely in the gas phase, it should do keeping itself tight to inhibit the charge separation. As for the free energy differences that were calculated by SPT, it was found that **TSa** and **TSb** are less stable than **TSc** by 4.6 and 0.8 kcal/mol, respectively. Taking account of these errors in the free energy in Figs. 1 and 2, the free energy of activation become 27.1 and 26.8 kcal/mol, respectively. Thus, under the additional influence of the TS shift toward the reactant side, there should have emerged large errors in the structure and energy at TS, especially when such bond lengths as R_{N-C} and R_{C-Cl} were used as a reaction coordinate to plot the free energy profile. It should be additionally noted that, in contrast to the case of asymmetric reactions discussed here, the effect of hydration upon the TS structure can be small in the symmetric S_N2 reaction $Cl^- + CH_3Cl \rightarrow CH_3Cl + Cl^-$ [33]. Although the large deviation in the asymmetric reaction is mainly because the partial optimization at a fixed structural parameter does not give a steepest-descent path for a reaction in the gas phase, the IRC calculation, which gives the steepest-descent one in the gas phase, cannot necessarily work properly for a simulation in solution; the area the IRC can trace is not so wide. Therefore, a free energy calculation with respect to a bond length is unavoidable in many cases, but it should be followed by the TS optimization using, e.g. the FEG method, to obtain a reasonable result.

Finally, the solvation structures for TSs are discussed. The radial distribution functions (RDFs) with respect to the Cl–OW and Cl–HW distances are shown in Figs. 4a and b, respectively.

For the better statistics, the RDFs were obtained by taking an average over three RDFs that were calculated from three independent trajectories each containing 30 ps run after 10 ps equilibration. One observes that the strength of the interaction between Cl and water molecules increases in the order of **TSa**, **TSb**, and **TSc**. The

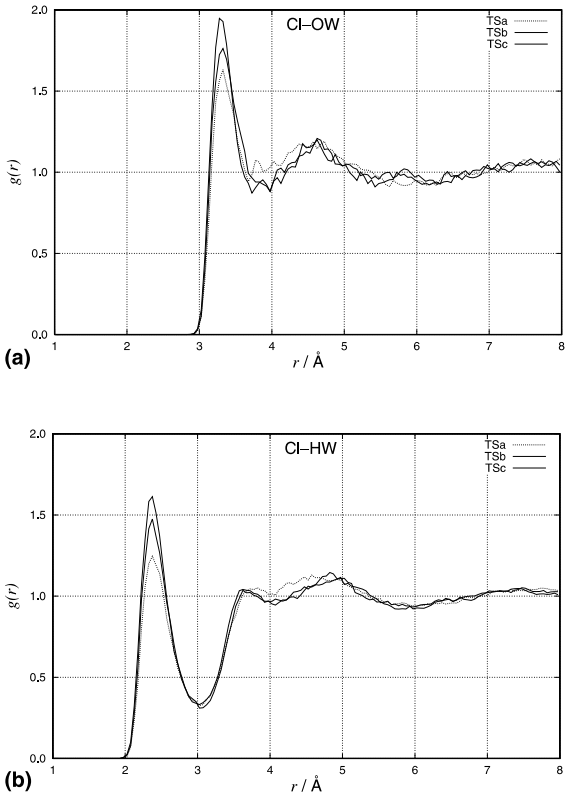

Fig. 4. (a) Cl–OW and (b) Cl–HW radial distribution functions for TSs in aqueous solution.

calculated Mulliken charges of Cl for **TSa**, **TSb**, and **TSc** were −0.58, −0.65, and −0.67, respectively (Table 1), accounting for the order of the strength of solute–solvent interaction.

Also noted is that the first peaks for **TSc** are both sharp compared to those broad ones obtained by Gao and Xia [9]. This is probably due to the difference in the formulas employed for calculating QM-MM interaction energies. Taking into account that the method of Cummins and Gready adopted presently gives good interaction energy curves as found by Luque et al. [34] and that the peaks for the TS for the symmetric S_N2 reaction are not broad but relatively sharp [33], the present RDFs are considered to be more appropriate and reliable.

In this Letter, we studied the TS for the Menshutkin reaction between ammonia and methyl chloride in aqueous solution by the combined

Table 1
Charge distributions for **TSa**, **TSb**, and **TSc**[a]

Species	Gas phase			Aqueous solution		
	H_3N	CH_3	Cl	H_3N	CH_3	Cl
TSa	0.11	0.33	−0.44	0.19	0.39	−0.58
TSb	0.11	0.37	−0.48	0.21	0.44	−0.65
TSc	0.11	0.39	−0.50	0.20	0.47	−0.67

[a] Calculated by the Mulliken population analysis.

FEG and QM/MM method. It has been found, through the comparison with other theoretical results for the reaction, that the FEG method provides quite reasonable information about the structure and energy for reactions in solution within the theoretical level employed. Its applications to other chemically reacting systems are also possible, even if some more sophisticated optimization algorithms [2] might be needed. In addition, because of its inherent properties to use the MD simulation and to describe solvent molecules explicitly, it could be extended so as to investigate the dynamic aspects of solution chemical reactions. Further development of the method along this line is now in progress.

Acknowledgements

We are grateful to Prof. M.J. Field and Prof. K.N. Houk for illuminating comments and constructive discussions. This work was supported partly by a Grant-in-Aid for Science Research from the Ministry of Education, Science and Culture in Japan and also by the Research and Development Applying Advanced Computational Science and Technology from the Japan Science and Technology Corporation (Project ACT-JST-98-A5-1).

References

[1] S. Glasstone, K.J. Laidler, H. Eyring, The Theory of Rate Processes, McGraw-Hill, New York, 1941.
[2] T. Fueno (Ed.), The Transition State – A Theoretical Approach, Kodansha, Tokyo, 1998.
[3] R. Cammi, J. Tomasi, J. Chem. Phys. 100 (1994) 7495.
[4] H. Sato, F. Hirata, S. Kato, J. Chem. Phys. 105 (1996) 1546.
[5] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Oxford, New York, 1987.
[6] A. Warshel, Computer Modeling of Chemical Reactions in Enzymes and Solutions, Wiley, New York, 1991.
[7] M.J. Field, P.A. Bash, M. Karplus, J. Comp. Chem. 11 (1990) 700.
[8] J. Gao, Acc. Chem. Res. 29 (1996) 298.
[9] J. Gao, X. Xia, J. Am. Chem. Soc. 115 (1993) 9667.
[10] N. Isaacs, Physical Organic Chemistry, Longman, Harlow, 1995.
[11] N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, Int. J. Quantum Chem. 70 (1998) 95.
[12] M. Nagaoka, N. Okuyama-Yoshida, T. Yamabe, J. Phys. Chem. A 102 (1998) 8202.
[13] N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, J. Phys. Chem. A 102 (1998) 285.
[14] N. Okuyama-Yoshida, K. Kataoka, M. Nagaoka, T. Yamabe, J. Chem. Phys. 113 (2000) 3519.
[15] N.Z. Menshutkin, Phys. Chem. 5 (1890) 589.
[16] M. Solà, A. Lledòs, M. Duran, J. Bertrán, J.-L.M. Abboud, J. Am. Chem. Soc. 113 (1991) 2873.
[17] V. Dillet, D. Rinaldi, J. Bertrán, J.-L. Rivail, J. Chem. Phys. 104 (1996) 9437.
[18] T.N. Truong, T.-T.T. Truong, E.V. Stefanovich, J. Chem. Phys. 107 (1997) 1881.
[19] C. Amovilli, B. Nennucci, F.M. Floris, J. Phys. Chem. B 102 (1998) 3023.
[20] K. Naka, H. Sato, A. Morita, F. Hirata, S. Kato, Theor. Chem. Acc. 102 (1999) 165.
[21] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.
[22] W.L. Jorgensen, J. Chandrasekhar, J.D. Madula, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926.
[23] P.L. Cummins, J.E. Gready, J. Comp. Chem. 18 (1997) 1496.
[24] G.J. Martyna, M.L. Klein, J. Chem. Phys. 97 (1992) 2635.
[25] W.F. van Gunsteren, H.J.C. Berendsen, Mol. Phys. 34 (1977) 1311.
[26] H.C. Andersen, J. Comp. Phys. 52 (1983) 24.
[27] A. Cheng, R.S. Stanton, J.J. Vincent, A. van der Vaart, K.V. Damodaran, S.L. Dixon, D.L. Hartsough, M. Mori, S.A. Best, G. Monard, M. Garcia, L.C. Van Zant, K.M. Merz Jr., ROAR 2.0, The Pennsylvania State University, 1999.
[28] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham III, D.M. Ferguson, G.L. Seibel, U.C. Singh,

P.K. Weiner, P.A. Kollman, AMBER 4.1, Unversity of California, San Francisco, CA, 1995.
[29] J.J.P. Stewart, MOPAC 7.0, QCPE No. 455 13 (1993) 42.
[30] M.J. Frisch et al., GAUSSIAN 98, Revision A.7, Gaussian Inc., Pittsburgh, PA, 1998.
[31] P. Kollman, Chem. Rev. 93 (1993) 2395.
[32] K. Fukui, J. Phys. Chem. 74 (1970) 461.
[33] W.L. Jorgensen, J.K. Buckner, J. Phys. Chem. 90 (1986) 4651.
[34] F.J. Luque, N. Reuter, A. Cartier, M.F. Ruiz-López, J. Phys. Chem. A 104 (2000) 10923.