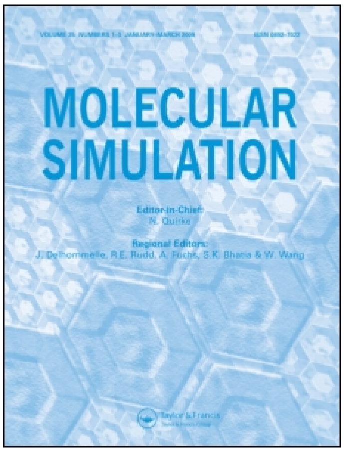


Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Proton Transfer in Liquid Water II; A Semiempirical Method to Describe Chemical Reactions

Tamiki Komatsuzaki[a]; Iwao Ohmine[b]
[a] Institute for Fundamental Chemistry, Kyoto, Japan [b] Department of Chemistry, Faculty of Science, Nagoya University, Nagoya, Japan

To cite this Article Komatsuzaki, Tamiki and Ohmine, Iwao(1996) 'Proton Transfer in Liquid Water II; A Semiempirical Method to Describe Chemical Reactions', Molecular Simulation, 16: 4, 321 — 344

To link to this Article: DOI: 10.1080/08927029608024083

URL: http://dx.doi.org/10.1080/08927029608024083

Molecular Simulation, 1996, Vol. 16, pp. 321–344
Reprints available directly from the publisher
Photocopying permitted by license only

PROTON TRANSFER IN LIQUID WATER II; A SEMIEMPIRICAL METHOD TO DESCRIBE CHEMICAL REACTIONS

TAMIKI KOMATSUZAKI*

Institute for Fundamental Chemistry,
34-4, Takano-Nishihiraki-cho, Sakyo-ku, Kyoto 606, Japan

IWAO OHMINE

Department of Chemistry, Faculty of Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya, 464-01, Japan

(Received March 1995, accepted June 1995)

A new semiempirical method is developed to deal with the proton transfer in liquid water. In the previous work, we have shown that two- and three-body charge transfer interactions and electrostatic interactions are the most important factors to describe the potential energy surfaces (PES) of the proton transfer in liquid water [*Chemical Physics* **180**, 239–269, 1994]. In order to take account of these factors, we develop a semiempirical method imposing the principle of electronegativity equalization to the Atoms in Molecule (AIM) method. The method is free from the well-known discrepancy of the traditional AIM methods, that is, the fractional molecular charges at large molecular separation, and thus can be applied to the charge transfer reactions. Intra- and intermolecular physical quantities, such as total energies, force vectors, dipole moment vectors and intermolecular charge transfer, obtained by the present method are found to be in good agreement with those by *ab initio* calculation.

KEY WORDS: Proton transfer, electronegativity equalization principle, atoms in molecule method.

1 INTRODUCTION

Proton transport in liquid water is one of the most fundamental processes in chemical reactions [1]. Despite its importance, there have been only few theoretical investigations on this reaction dynamics. This is because the reaction is essentially multidimensional in nature; the solvent water molecules directly participate in the proton transfer. Many degrees of freedom are to be explicitly included in the analysis. On the other hand the theoretical studies for other chemical reactions in solutions have assumed that reaction coordinates are well separated from solvent coordinates, and thus the generalized Langevin equation has been widely applied within the framework of the linear response theory [2].

---

* To whom all correspondences should be addressed.

A molecular dynamics simulation will be a powerful technique to investigate this proton transfer dynamics and to find how the hydrogen-bond network rearrangement of liquid water is related to this dynamics [3]. In the past, only the work that elucidates the proton transfer in $H_5O_2^+ + H_3O_2^-$ by the MD simulation was by Weber and Stillinger [4]. Their polarization model was, however, not applied to the proton transport in a larger system. It is recently that the proton transfer dynamics in water has been investigated by a reliable calculation, Car-Parrinello-Vanderbilt method, but the investigation is still for the short time dynamics of a small water cluster [5].

In this paper, we propose a semiempirical method usable in an MD calculation for large systems, which explicitly takes account of the intra- and intermolecular charge reorganization associated with the proton transfer. The method is based on the electronegativity equalization principle [6]. Many variations of Atoms in Molecule (AIM) method, using the principle of electronegativity equalization, have been proposed in order to deal with the charge reorganization of molecules [7–10]. The AIM model was recently extended to be combined with the extended Lagrangian method by Rick, Stuart and Berne [11], and applied to a molecular dynamics calculation of liquid water. Their method was found to be very useful to describe the polarization effects and well reproduce many physical properties of water. It can not, however, be applied straightforwardly to chemical reactions involving the bond-forming and –breaking such as a proton transfer process. In the present work, we propose a new method based on the AIM model which can be applied to reactions. We introduce 'clustering of charge transferable space' in order to avoid the well-known discrepancy of fractional molecular charge distribution at an infinite separation, intrinsic in traditional AIM methods [12,13]. We also derive a simple formula evaluating forces, which can be easily implemented in MD calculations.

2 THEORY AND METHOD

In the AIM method, the total electronic energy is assumed to be decomposed into the intra- and interatomic electronic energies. The intra-atomic electronic energy of atom X, relative to the energy of the neutral atom X, is assumed to be described by the second-order Taylor expansion with respect to an electron density on atom X, ΔN_X. N_X is number of electrons on atom X, N_X^0 is that on X^0 and ΔN_X is $N_X - N_X^0$. Here, we use the notation X^0 to indicate atom X with the isolated atomic electron density. The interatomic electronic energy is approximated by the sum of the Coulomb type interactions between the atoms. The $\{\Delta N_X\}$, a set of ΔN_X, are variational variables and determined to equalize the chemical potential of all atoms. The total energy of a system at a given configuration, E, is given by a sum of two terms, E_0 and ΔE. E_0 is the total energy of the system without the charge reorganization, (*i.e.*, the total system energy given by assuming that the atoms having the isolated atomic electron densities, X^0s, mutually interact). E_0 is thus independent on the variational variable ΔN_X [9]. ΔE is the charge reorganization energy.

$$E = E_0 + \Delta E. \qquad (1)$$

ΔE depends on ΔN_X and corresponds to the stabilization energy due to the charge reorganization. ΔE can be described as follows.

$$\Delta E = \sum_X \mu_{X^+}(\mathbf{R}_X)\Delta N_X + \frac{1}{2}\sum_{X,Y} \eta_{XY}(\mathbf{R}_X,\mathbf{R}_Y)\Delta N_X \Delta N_Y, \tag{2}$$

where

$$\mu_{X^+}(\mathbf{R}_X) = \mu_{X^0} + \sum_Y{}' [-Z_Y U_{en}(|\mathbf{R}_X - \mathbf{R}_Y|) + N_{Y^0} U_{ee}(|\mathbf{R}_X - \mathbf{R}_Y|)], \tag{3}$$

$$\eta_{XX}(\mathbf{R}_X, \mathbf{R}_X) = \eta_{X^0}, \tag{4}$$

$$\eta_{XY}(\mathbf{R}_X, \mathbf{R}_Y) = U_{ee}(|\mathbf{R}_X - \mathbf{R}_Y|), \tag{5}$$

$$\sum_X N_X = \sum_X [N_{X^0} + \Delta N_X] = N. \tag{6}$$

Here, μ_{X^0} and η_{X^0} denotes the chemical potential and the hardness of atom X in the isolated case [14], respectively and N is the total number of electrons of the system. $\mathbf{R}_X$ is a position vector of atom X. The prime in equation (3) means that the summation over Y is for Y $\neq$ X. $\mu_{X^+}(\mathbf{R}_X)$[equation (3)] depends on all atomic coordinates of the system, $\{\mathbf{R}_Z\} = (\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \ldots)$. $U_{\xi\zeta}(R)$ denotes the Coulomb type potential where ξ and ζ denotes electron (e) or nucleus (n).

E_0 is the sum of the total energy of the X^0 and the Coulomb type interaction energies between the unperturbed atoms located at the system configuration.

$$\begin{aligned} E_0 = {} & \sum_X E_{X^0} \\ & + \sum_{X>Y} [Z_X Z_Y U_{nn}(|\mathbf{R}_X - \mathbf{R}_Y|) + N_{X^0} N_{Y^0} U_{ee}(|\mathbf{R}_X - \mathbf{R}_Y|) - N_{X^0} Z_Y U_{en}(|\mathbf{R}_X - \mathbf{R}_Y|)] \\ & + f_\theta + f_{\text{rep}}. \end{aligned} \tag{7}$$

Here, E_0 also includes the effective empirical potentials f_θ and f_{rep} representing the atomic dipole force and the exchange repulsion effects, respectively, as described later.

In AIM models truncated at the second-order Taylor expansion with respect to charge densities, imposing that the chemical potential is equal over all atoms results in the non-zero intermolecular charge transfers among atoms even at the infinite separation. This obviously contradicts with the real situation. One way of resolving this discrepancy is to introduce the concept of 'temperature' in the electronic structure theory, as shown by Perdew *et al.* [12]. They showed that the quadratic expansion of the electronic energies of fragments in terms of their electron density is invalid at the asymptotic regime. The infinite expansion is required to describe the discontinuity of the chemical potential against the charge density, needed to have proper charge distribution at infinite separation. The presence of this discontinuity is confirmed by the charge-constrained *ab initio* calculation by Cioslowski and Stefanov [13].

The second-order Taylor expansion of the interaction energy [equation (2)] is a good approximation when molecules (atoms) are in the interaction regime where molecules are close to each other and making strong interactions [13]. Furthermore, the expression for the interaction energy is simple. So we have tried to retain this functional formula and yet to avoid the fractional charge distribution at near infinite separation (called asymptotic regime). This is achieved by restricting the space of the interatomic charge transfers in the following way. The first step is to group atoms into clusters by a certain criterion depending on the distances among atoms; the distant atoms belong to different clusters. For example, a cluster can be H_2O and another be H_3O^+ in a ($H_2O + H_3O^+$) system with large H_2O-H_3O^+ separation (see Fig. 1). The chemical potentials are then equalized among atoms within each cluster by allowing the charge transfers to occur among these atoms under the influence of Coulomb forces from atoms of other clusters. No charge transfer is allowed among atoms belonging to different clusters and thus the total atomic charges of individual clusters are integer. In this approximation, the chemical potentials of atoms belonging to different clusters thus can not be mutually equal. We term this clustering of the atoms as CCTS (*C*lustering of *C*harge *T*ransferable *S*pace).

In the case of close H_2O-H_3O^+ distance, for example, all atoms belong to a cluster (interacting regime), and the atomic charges are obtained by allowing charge transfers to occur among all atoms. The atomic charges of this interacting regime (*e.g.*, with CCTS2) must be smoothly connected with those at asymptotic regime (*e.g.*, with CCTS1) by an interpolation. We use the following interpolation scheme. When the system is in between the interacting regime and the asymptotic regime, the interaction energy is evaluated by substituting the average electron density $\bar{N}_X$ into equation (2) [15]. The average electron density $\bar{N}_X$ in this intermediate region is calculated by employing a weighting function $\omega_i(\{\mathbf{R}_X\})$ smoothly connecting these two regimes. In general, $\bar{N}_X$ is defined by

$$\bar{N}_X = \sum_i \omega_i(\{\mathbf{R}_X\}) N_X^i, \tag{8}$$

and

$$\sum_i \omega_i(\{\mathbf{R}_X\}) = 1, \tag{9}$$

where N_X^i is the number of electrons on atom X calculated in the i-th CCTS and $\omega_i(\{\mathbf{R}_X\})$ is the normalized weighting function of the i-th CCTS at a given system configuration, $\{\mathbf{R}_X\}$.

The reorganization electronic energy ΔE is then given by

$$\Delta E = \Delta E[\{\overline{\Delta N}_X\}, \{\mathbf{R}_X\}] \tag{10}$$

$$= \sum_X \mu_{X^+}(\mathbf{R}_X)\overline{\Delta N}_X + \frac{1}{2}\sum_{X,Y} \eta_{XY}(\mathbf{R}_X, \mathbf{R}_Y)\overline{\Delta N}_X\overline{\Delta N}_Y. \tag{11}$$

$$\Delta E^i = \Delta E[\{\Delta N_X^i\}, \{\mathbf{R}_X\}], \tag{12}$$

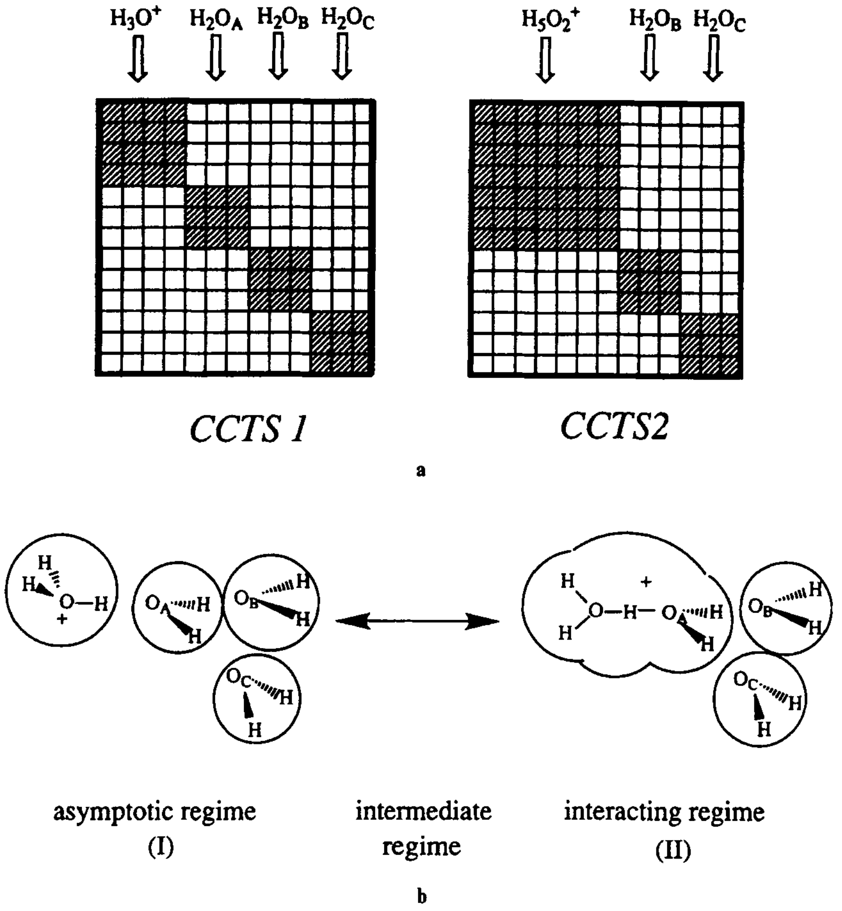


Figure 1 a) The clustering of charge transferable space (CCTS) in hardness matrix. b) The molecular configurations in an asymptotic regime (I) and in an interacting regime (II). CCTS1 for the configuration I and CCTS2 for the configuration II.

where $\{\overline{\Delta N_X}\}$ and $\{\Delta N_X^i\}$ stands for $(\overline{\Delta N_1}, \overline{\Delta N_2}, \overline{\Delta N_3}, \ldots.)$ and $(\Delta N_1^i, \Delta N_2^i, \Delta N_3^i, \ldots.)$, respectively. This procedure is readily applied to a case with many CCTSs.

An example of CCTS for a proton transferring in a water cluster is shown in Figure 1.

In CCTS1, the charge reorganization is carried out within a water molecule and within H_3O^+ corresponding to each diagonal-block of hardness matrix, η. In CCTS2, a water molecule strongly interacts with H_3O^+ and the intermolecular charge transfer is allowed among all atoms of the segment $H_2O + H_3O^+$. In an asymptotic regime I, the average electron density, $\bar{N}_X$, is to be the electron density of CCTS1 ($\omega_1 = 1$ and $\omega_2 = 0$) and ΔE is equal to ΔE^1. In the interacting regime II, ΔE is equal to $\Delta E^2(\omega_1 = 0$ and $\omega_2 = 1)$. At each system configuration $\{\mathbf{R}_X\}$ in the intermediate regime between regime I and II, we calculate the electron densities for both CCTS1 and CCTS2 and evaluate $\bar{N}_X$ as their average weighted with $\omega_i(\{\mathbf{R}_X\})$.

The electron densities $\{N^i_{\mathrm{X}}\}$ of the atoms X at the $\{\mathbf{R}_{\mathrm{X}}\}$ configuration in the i-th CCTS are obtained by minimizing the reorganization energy under the constraint that the total number of the charge in the κ-th cluster should be integer. The Lagrangian for the i-th CCTS is then given by

$$L_i = \Delta E[\{\Delta N^i_{\mathrm{X}}\},\{\mathbf{R}_{\mathrm{X}}\}] - \sum_{\kappa=1}^{l_i} \mu^i_\kappa \left[\sum_{X\in\kappa}^{m_{i\kappa}} \Delta N^i_{\mathrm{X}} - \Delta N^i_\kappa \right], \tag{13}$$

where μ^i_κ and ΔN^i_κdenotes the Lagrange multiplier and the total charge of the κ-th cluster in the i-th CCTS, respectively. l_i and $m_{i\kappa}$ are the total number of the clusters in the i-th CCTS and the total number of atoms in the κ-th cluster of the i-th CCTS, respectively. The variational condition $\delta L_i = 0$ yields a set of $l_i(m_{i\kappa}+1)$ inhomogeneous linear equations for $l_i(m_{i\kappa}+1)$ unknown variables. For example, the variations with respect to the charges of atoms X belonging to the κ-th cluster, ΔN^i_{X}, yield $m_{i\kappa}$ equations

$$\mu_{\mathrm{X}^+} + \sum_{Y\notin\kappa} \eta_{\mathrm{XY}} \Delta N^i_{\mathrm{Y}} + \sum_{Y\in\kappa} \eta_{\mathrm{XY}} \Delta N^i_{\mathrm{Y}} = \mu^i_\kappa \tag{14}$$

and the variation with respect to the Lagrange multiplier, μ^i_κ, [the chemical potential of the κ-th cluster when equation (14) is satisfied] gives an equation

$$\sum_{X\in\kappa} \Delta N^i_{\mathrm{X}} - \Delta N^i_\kappa = 0. \tag{15}$$

We have l_i sets of them. For a small or an intermediate size system, these equations can be straightforwardly solved for $n(=l_i m_{i\kappa}$; total number of atoms in the system) unknown variables ΔN^i_{X} and l_i unknown variables μ^i_κ by using a matrix inversion method.

For a large system, an equation of motion method might be used instead of a matrix inversion method. The Lagrangian including the kinetic energy term with the fictitious charges 'mass', M_Q, for the i-th CCTS, L_i is given by

$$L_i = \frac{1}{2} M_Q \sum_{\kappa=1}^{l_i} \sum_{X\in\kappa}^{m_{i\kappa}} (\Delta \dot{N}^i_{\mathrm{X}})^2 - \Delta E[\{\Delta N^i_{\mathrm{X}}\},\{\mathbf{R}_{\mathrm{X}}\}] - \sum_{\kappa=1}^{l_i} \mu^i_\kappa \left[\sum_{X\in\kappa}^{m_{i\kappa}} \Delta N^i_{\mathrm{X}} - \Delta N^i_\kappa \right]. \tag{16}$$

Then, the equation of motion for ΔN^i_{X} is

$$M_Q \Delta \ddot{N}^i_{\mathrm{X}} = -\frac{1}{m_{i\kappa}} \sum_{Y}^{m_{i\kappa}} (\tilde{\mu}^i_{\mathrm{X}} - \tilde{\mu}^i_{\mathrm{Y}}), \tag{17}$$

$$\tilde{\mu}^i_{\mathrm{X}} = \frac{\partial \Delta E[\{\Delta N^i_{\mathrm{X}}\},\{\mathbf{R}_{\mathrm{X}}\}]}{\partial \Delta N^i_{\mathrm{X}}}, \tag{18}$$

for atoms X belonging to the κ-th cluster of the i-th CCTS. These equations of motion are solved by using a standard MD calculation. This method does not need a matrix inversion and may be more efficient [11].

2.1 *Force Calculation*

The force arising from E_0 can be straightforwardly calculated because it does not include the variables, $\Delta N_{\mathrm{X}}^{i}$. The force arising from ΔE is given by

$$\frac{d\Delta E}{d\mathbf{R}_{\mathrm{X}}}=\frac{\partial\Delta E}{\partial\mathbf{R}_{\mathrm{X}}}+\sum_{i,Y}\frac{\partial\Delta E}{\partial\Delta N_{\mathrm{Y}}^{i}}\frac{\partial\Delta N_{\mathrm{Y}}^{i}}{\partial\mathbf{R}_{\mathrm{X}}}+\sum_{i}\frac{\partial\Delta E}{\partial\omega_i}\frac{\partial\omega_i}{\partial\mathbf{R}_{\mathrm{X}}}. \tag{19}$$

Using

$$\frac{\partial\Delta E}{\partial\Delta N_{\mathrm{Y}}^{i}}=\frac{\partial\Delta E}{\partial\overline{\Delta N}_{\mathrm{Y}}}\frac{\partial\overline{\Delta N}_{\mathrm{Y}}}{\partial\Delta N_{\mathrm{Y}}^{i}}=\bar{\mu}_{\mathrm{Y}}\omega_i, \tag{20}$$

$$\frac{\partial\Delta E}{\partial\omega_i}=\sum_{Y}\frac{\partial\Delta E}{\partial\overline{\Delta N}_{\mathrm{Y}}}\frac{\partial\overline{\Delta N}_{\mathrm{Y}}}{\partial\omega_i}=\sum_{Y}\bar{\mu}_{\mathrm{Y}}\Delta N_{\mathrm{Y}}^{i}, \tag{21}$$

equation (19) becomes

$$\frac{d\Delta E}{d\mathbf{R}_{\mathrm{X}}}=\frac{\partial\Delta E}{\partial\mathbf{R}_{\mathrm{X}}}+\sum_{i,Y}\left[\bar{\mu}_{\mathrm{Y}}\omega_i\frac{\partial\Delta N_{\mathrm{Y}}^{i}}{\partial\mathbf{R}_{\mathrm{X}}}+\bar{\mu}_{\mathrm{Y}}\Delta N_{\mathrm{Y}}^{i}\frac{\partial\omega_i}{\partial\mathbf{R}_{\mathrm{X}}}\right] \tag{22}$$

$$=\frac{\partial\Delta E}{\partial\mathbf{R}_{\mathrm{X}}}+\sum_{Y}\bar{\mu}_{\mathrm{Y}}\frac{\partial\overline{\Delta N}_{\mathrm{Y}}}{\partial\mathbf{R}_{\mathrm{X}}} \tag{23}$$

In a single CCTS case (*e.g.*, *i*-CCTS), the second term of equation (23) vanishes since

$$\sum_{Y}\mu_{\mathrm{Y}}^{i}\frac{\partial\Delta N_{\mathrm{Y}}^{i}}{\partial\mathbf{R}_{\mathrm{X}}}=\sum_{\kappa=1}^{l_i}\mu_{\kappa}^{i}\frac{\partial\sum\limits_{Y\in\kappa}^{m_{i\kappa}}\Delta N_{\mathrm{Y}}^{i}}{\partial\mathbf{R}_{\mathrm{X}}}=0, \tag{24}$$

for

$$\sum_{Y\in\kappa}^{m_{i\kappa}}\Delta N_{\mathrm{Y}}^{i}=\Delta N_{\kappa}^{i}=constant. \tag{25}$$

$d\Delta E/d\mathbf{R}_{\mathrm{X}}$ is then given by

$$\frac{d\Delta E}{d\mathbf{R}_{\mathrm{X}}}=\frac{\partial\Delta E}{\partial\mathbf{R}_{\mathrm{X}}}=\sum_{Y}\frac{\partial\mu_{\mathrm{Y}^{+}}}{\partial\mathbf{R}_{\mathrm{X}}}\Delta N_{\mathrm{Y}}+\frac{1}{2}\sum_{Y,Z}\frac{\partial\eta_{\mathrm{YZ}}}{\partial\mathbf{R}_{\mathrm{X}}}\Delta N_{\mathrm{Y}}\Delta N_{\mathrm{Z}}. \tag{26}$$

This is analogous to the energy gradient calculation in Hartree-Fock theory, in which derivatives of molecular orbital coefficients with respect to $\mathbf{R}_{\mathrm{X}}$ are not needed because a wave function is normalized.

We need, however, to use multi-CCTS to describe atomic transfer reactions such as a proton transfer process in liquid water. In an intermediate regime where more than two CCTS calculations are needed, the force calculation then becomes rather complicated. This is because ΔE is not minimized with respect to the *average* electron density (although it is minimized to $\{\Delta N_{\mathrm{X}}^{i}\}$ of each *i*-CCTS) and the second term of equation (23) does not vanish. A set of derivatives $\{\partial\overline{\Delta N}_{\mathrm{Y}}/\partial\mathbf{R}_{\mathrm{X}}\}$ should be explicitly estimated. As seen in equation (22), each of these derivatives consists of two terms; a derivative of the weighting function and that of the optimal electron density. The former derivative can be easily estimated for ω_i known as a function of $\mathbf{R}_\mathrm{X}$. The latter derivative can be obtained as follows, as first shown by Wales [16].

Differentiating the Lagrangian gradients with the optimal electron density [equations (14) and (15)] by $\mathbf{R}_\mathrm{X}$, we obtain

$$\frac{\partial^2 L_i}{\partial \Delta N^i_\mathrm{X} \partial \mathbf{R}_\mathrm{X}} = \frac{\partial \mu_{\mathrm{X}^+}}{\partial \mathbf{R}_\mathrm{X}} + \sum_Y \frac{\partial \eta_\mathrm{XY}}{\partial \mathbf{R}_\mathrm{X}} \Delta N^i_\mathrm{Y} + \sum_Y \eta_\mathrm{XY} \frac{\partial \Delta N^i_\mathrm{Y}}{\partial \mathbf{R}_\mathrm{X}} - \frac{\partial \mu^i_\kappa}{\partial \mathbf{R}_\mathrm{X}} = 0, \tag{27}$$

$$\frac{\partial^2 L_i}{\partial \mu^i_\kappa \partial \mathbf{R}_\mathrm{X}} = -\sum_{Y \in \kappa}^{m_{i\kappa}} \frac{\partial \Delta N^i_\mathrm{Y}}{\partial \mathbf{R}_\mathrm{X}} = 0, \tag{28}$$

which are the linear inhomogeneous equations for $n(=l_i m_{i\kappa})$ variables $\partial \Delta N^i_\mathrm{Y}/\partial \mathbf{R}_\mathrm{X}$ and l_i variables $\partial \mu^i_k/\partial \mathbf{R}_\mathrm{X}$. As we know the functional forms of μ_{X^+} and η_XY, we can easily calculate $\partial \mu_{\mathrm{X}^+}/\partial \mathbf{R}_\mathrm{X}$ and $\partial \eta_\mathrm{XY}/\partial \mathbf{R}_\mathrm{X}$. $\partial \Delta N^i_\mathrm{Y}/\partial \mathbf{R}_\mathrm{X}$ are obtained by solving these linear equations with a matrix inversion method or by using an equation of motion method, similar as in equations (14) and (15). This is performed for each $\mathbf{R}_\mathrm{X}$ and each CCTS.

It is mentioned here that in a similar fashion we can derive a simple formula of a Hessian matrix with respect to the nuclear coordinates for AIM models with single CCTS [16].

2.2 *CCTS and Weighting function* ω_i

We must carefully select a set of CCTSs and choose the weighting function $\omega_i(\{\mathbf{R}_\mathrm{x}\})$ in order to obtain the reliable PES and forces of a reaction.

A procedure to specify the network connection in a percolation problem, called Hoshen-Kopelman algorithm [17], can be applied to efficiently construct CCTSs. We employ the hardness η_{ij} among i and j atoms (*i.e.*, interaction strength between atoms) as a criterion to construct CCTSs. There are, of course, many other criteria to construct CCTS; for example, a distances between atoms or an overlaps of atomic wave functions. Two threshold parameters, η_1 and η_2 $(0 \leqslant \eta_1 < \eta_2)$, are introduced to build a discretized interaction matrix, H_{ij}, as

$$\begin{aligned} &\text{(a)}\ H_{ij} = 0 \quad \text{for } \eta_1 > \eta_{ij} \\ &\text{(b)}\ H_{ij} = 1 \quad \text{for } \eta_2 > \eta_{ij} \geqslant \eta_1 \\ &\text{(c)}\ H_{ij} = 2 \quad \text{for } \eta_{ij} \geqslant \eta_2 \end{aligned} \tag{29}$$

Applying the clustering algorithm of Hoshen and Kopelman [17] to this interaction matrix, we obtain a 'clustered interaction matrix' C_{ij} which consists of several block diagonal segments, called clusters; an element C_{ij} for a pair of atoms (i, j) belonging to different clusters is 0, and that for atoms belonging to a same cluster is 1 or 2, depending on the strength of their 'interconnected' interaction. Here we construct C_{ij} from H_{ij} in the following procedure. For an atomic pair (i, j) with $H_{ij} = 2$, $C_{ij} = 2$. For an atomic pair (i, j) with $H_{ij} = 0$ or 1 but whose 'interconnected interactions' satisfy, for example, both $H_{ik} = 2$ and $H_{kj} = 2$, then $C_{ij} = 2$. For an atomic pair (i, j) with $H_{ij} = 0$ or 1 but whose all 'interconnected interactions' satisfy $H_{lk} = 1$, then $C_{ij} = 1$. If $H_{ij} = 0$ and all its 'interconnected interactions' are also 0, then $C_{ij} = 0$. All diagonal elements are 2. This is an extension of the original Hoshen-Kopelman procedure.

The reorganization energy of the system whose clustered interaction matrix includes $C_{ij}=1$ should be interpolated from a set of the reorganization energies calculated for several CCTSs. We can easily construct a set of CCTSs from the 'clustered interaction' matrix. For example, suppose that we have following matrices for a system with three atoms.

$$C=\underset{\mathrm{I}}{\begin{pmatrix}2 & 2 & 2\\ 2 & 2 & 2\\ 2 & 2 & 2\end{pmatrix}},\underset{\mathrm{II}}{\begin{pmatrix}2 & 2 & 0\\ 2 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}},\underset{\mathrm{III}}{\begin{pmatrix}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{pmatrix}},\underset{\mathrm{IV}}{\begin{pmatrix}2 & 1 & 0\\ 1 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}}, \tag{30}$$

where the columns (rows) denote three atoms, say, A, B and C. Those cases correspond to the following situations, respectively: (I) all three atoms interact strongly, (II) A and B interact strongly but C is far apart from these two atoms, (III) all three atoms interact with intermediate strength and (IV) A and B interact with intermediate strength and C is far apart from A and B.

Cases I and II are represented by single CCTS, respectively. In Case (I) all atoms are included in one cluster and the charge transfer is allowed among all atoms. In case (II), atoms A and B make a block (cluster) and C makes another block. The charge transfer is allowed between A and B but not from/to C. Cases III and IV correspond to the intermediate regime with two CCTSs. in Case III the system energy can be calculated from the average electron density $\bar{N}_X(X=A,B,C)$ which are interpolated from those of two CCTSs. These two CCTSs (*e.g.*, CCTS1 and CCTS2) are represented by following clustered interaction matrices;

$$\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix},\begin{pmatrix}2 & 2 & 2\\ 2 & 2 & 2\\ 2 & 2 & 2\end{pmatrix} \tag{31}$$

That is, each atom is belonging to different cluster (the charge transfer is not allowed between any atoms), and all atoms, A, B and C, belong to a cluster (the charge transfer is allowed among all atoms), respectively. Similarly, the system energy of Case IV can be interpolated from two CCTSs of

$$\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix},\begin{pmatrix}2 & 2 & 0\\ 2 & 2 & 0\\ 0 & 0 & 2\end{pmatrix} \tag{32}$$

Generalization of the present procedure to a multi-CCTS case can be readily achieved.

The average electron density of the atom X, $\overline{\Delta N}_X$ becomes, for example,

$$\overline{\Delta N}_X=\omega_1\Delta N_X^1+\omega_2\Delta N_X^2,\ \omega_2=1-\omega_1, \tag{33}$$

where $\Delta N_X^1(\Delta N_X^2)$ is evaluated at the CCTS1(CCTS2). $\omega_1(0\leqslant\omega_1\leqslant1)$ must be a smooth function interpolating between two regimes thus satisfy the following condition,

$$\omega_1 = 0 \text{ for } \eta_{\max\{1\}} \geqslant \eta_2 \text{ and } 1 \text{ for } \eta_{\max\{1\}} \leqslant \eta_1, \tag{34}$$

where $\eta_{\max\{1\}}$ is the maximum value in the hardness matrix elements of atomic pairs (i,j) whose C_{ij} is equal to 1. In our calculation, we use a following formula for ω_1,

$$\omega_1 = \begin{cases} 1 \ (x < \eta_1) \\ \sum\limits_{n=0}^{n_1} a_n x^{2n} (\eta_1 \leqslant x < \eta_B) \\ \dfrac{-\eta_D}{\eta_A - \eta_D} + \dfrac{x}{\eta_A - \eta_D} (\eta_B \leqslant x \leqslant \eta_C) \\ \sum\limits_{n=0}^{n_2} b_n x^{2n} (\eta_c < x < \eta_2) \\ 0 \ (\eta_2 < x), \end{cases} \tag{35}$$

where

$$x = \eta_{\max\{1\}}, \tag{36}$$

$$\eta_1 < \eta_A < \eta_B < \eta_C < \eta_D < \eta_2. \tag{37}$$

a_n and b_n are coeffieient of n-power of x to connect smoothly at η_1 and η_B and at η_C and η_2, respectively. This function ω_1 is plotted against x in Figure 2. $\eta_I (\mathrm{I} = \mathrm{A}, \mathrm{B}, \mathrm{C}$ and D) are adjustable parameters to reproduce the reference data. We use *ab initio* data of various quantities as the reference data (see Section 3).

To assign the number of electrons in the κ-th cluster of the i-th CCTS, N^i_κ, we decompose molecular species ($H_2O, H_3O^+, H_5O_2{}^+$, etc.) into H^+ and O^{2-} atoms and then count the number of oxygen atoms included in the κ-th cluster, N^κ_0. N^i_κ is then given by

$$N^i_\kappa = N_e(O^{2-}) \times N^\kappa_0. \tag{38}$$

where $N_e(O^{2-})$ is the number of electrons of O^{2-}. Since we deal with only valence electrons of a system, counting a charge of core electrons as a part of a nuclear charge, $N_e(O^{2-})$ is eight.

2.3 Coulomb type potential $U_{\xi\zeta}$

Coulomb type potential $U_{\xi\zeta}$ We postulate that the Coulomb type potential $U_{\xi\zeta}$ used in equations (3), (5) and (7) can be described by

$$U_{\xi\zeta}(R_{XY}) = U_{\xi\zeta}(|\mathbf{R}_X - \mathbf{R}_Y|) = \begin{cases} \dfrac{1}{2v_1}(\eta_X{}^0 + \eta_Y{}^0) + AR_{XY} + BR_{XY}{}^2 (R_{XY} < R_0) \\ \dfrac{1}{\left[R_{XY}{}^2 + v_2\left(\dfrac{2}{\eta_X{}^0 + \eta_Y{}^0}\right)^2\right]^{1/2}} (R_{XY} \geqslant R_0), \end{cases} \tag{39}$$

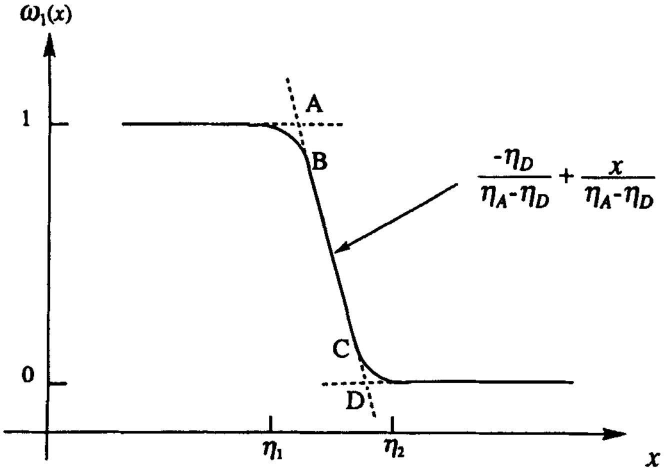


Figure 2 A function $\omega_1(x)$ [equation (35)].

where the constants A and B are chosen to smoothly connect these two expressions at R_0. The second expression of equation (39) is the Ohno formula [18, 19]. A similarly expression was used in the treatment of Nalewajski *et al.* [20]. The same formula is employed for U_{en} on the analogy of a semiempirical MO theory [21]. The following formula [22] is used for U_{nn}.

$$U_{nn}(R_{XY}) = \gamma_{nn}(R_{XY}) + \left(\frac{1}{R_{XY}} - \gamma_{nn}(R_{XY})\right)\exp(-\beta R_{XY}) \tag{40}$$

where $\gamma_{nn}(R_{XY})$ is the same formula to equation (39) and β is an adjustable parameter. The adjustable parameters in these equations, R_0, v_1, v_2 and β, are to be determined to reproduce reference *ab initio* data (see Section 3) for atomic pairs and pair of electron/nucleus.

2.4 *Effective empirical potential f_θ and f_{rep}*

The shape change of an atomic electron density distribution caused by interactions between atoms is one of main factors that determine the (intra-) molecular structure, being known as the atomic dipole force effect [23]. In order to take into account the atomic dipole force effect, we have introduced an empirical three-body potential function f_θ in equation (7) to reproduce the intramolecular bending potential of water and hydronium molecules. f_θ does not depend on the electron density.

The potential function, f_θ is defined as follows:

$$f_\theta = \sum_{X}^{\text{oxygen}} \sum_{Y\neq X}^{\text{all}} \sum_{Z\neq X,Y}^{\text{all}} T(R_{XY})T(R_{XZ}) \sum_{n=0}^{n_\theta} \vartheta_n(\mathbf{e}_{XY}\cdot\mathbf{e}_{XZ})^n, \tag{41}$$

$$T(R)=\frac{1}{1+\exp\{\alpha(R-R_C)\}}\sum_{n=0}^{n_T}c_nR^n, \tag{42}$$

where

$$\mathbf{e}_{XY}=\frac{\mathbf{R}_{XY}}{|\mathbf{R}_{XY}|}.$$

$\mathbf{e}_{XY}\cdot\mathbf{e}_{XZ}$ is a scalar product of $\mathbf{e}_{XY}$ and $\mathbf{e}_{XZ}$. The first summation in equation (41) is over oxygen atoms in the system. The second and third summations are over all atoms in the system. $T(R)$ is a polynomial of n_T-th order, multiplied by an exponentially decaying function. It contains adjustable parameters α, R_C and $c_n(n=0,1,2,3,..,n_T)$. ϑ_n $(n=0,1,2,\ldots,n_\theta)$ is a coefficient for the n-th power of $\mathbf{e}_{XY}\cdot\mathbf{e}_{XZ}$ and is an adjustable parameter.

In order to take account of the exchange repulsion effect, we have introduced an empirical two-body potential function f_{rep} in equation (7),

$$f_{\mathrm{rep}}\approx\sum_{X>Y}^{\mathrm{all}}T[\log(R_{XY})], \tag{43}$$

where $\log(R_{XY})$ is used instead of R_{XY} in order to reproduce the repulsive region properly. The adjustable parameters in equations (41)–(43) are optimized to reproduce reference *ab initio* data (see Section 3).

3 MODEL CALCULATIONS AND DISCUSSION

All parameters in the potential functions [equations (3)–(5), (35) and (39)–(43)] are fitted to accurately reproduce *various physical quantities* calculated at *various geometries* of *reference model systems* by an *ab initio* SCF method with MIDI-4** basis set. The physical quantities used in the fitting procedure are (1) force, (2) total energy, (3) dipole moment, and (4) gross atomic Mulliken charge. An intermolecular quantity Q_{INT} here is defined as

$$Q_{\mathrm{INT}}=Q(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3,\ldots,\boldsymbol{r}_N)-\sum_{i=1}^{N}Q^{(1)}(\boldsymbol{r}_i), \tag{44}$$

where $\boldsymbol{r}_i$ denotes the atomic nuclear position vectors of the i-th molecule and $Q^{(1)}(\boldsymbol{r}_i)$ denotes the corresponding intramolecular quantity of the i-th molecule with the intra-molecular geometry in the aggregate.

As reference model systems, we selected H_2O and H_3O^+ to calculate intramolecular physical quantities, and H_4O_2 and $H_5O_2^+$ to calculate intermolecular physical quantities. In order to avoid generating unrealistically high energy species in our model, we also added the tetrahedral H_4O^{2+} and H_4O^+ to the references. For the fitting, various geometries of these reference systems are generated. For H_2O, H_3O^+, H_4O_2 and $H_5O_2^+$, we generated many structures along displacement vectors of selected normal modes; the selected modes are all modes in H_2O, H_3O^+ and eleven lowest frequency modes in H_4O_2 and $H_5O_2^+$. For H_2O, we included large bending

($\angle$HOH = 104.5° → 180°) motion, and for H_3O^+, we added the inversion motion. For $H_5O_2^+$, the distance between the two oxygen atoms of H_2O and H_3O^+, R_{OO}, is arbitrarily chosen as 3.0, 3.5, 4.0, 5.0, 6.0, 7.0 and 8.0 Å, and their orientation angle is randomly generated for the system configurations whose interaction energy variations are within 9 kcal/mol. For H_4O_2, R_{OO} of two H_2O is arbitrarily chosen as 2.5, 3.0, 3.5, 4.0, 5.0, 6.0 and 7.0 Å and their orientation angle is randomly generated for the system configurations whose interaction energy variations are within 13 kcal/mol. In these configurations, the intramolecular structure of H_2O and H_3O^+ is fixed as those optimized in their isolated form, respectively. The energy minimum structures found by the *ab initio* calculation [24] are included for H_4O_2 and $H_5O_2{}^+$. The proton transfer PESs for $H_2O + H_3O^+ \rightarrow H_3O^+ + H_2O$ at several frozen oxygen-oxygen distances (R_{OO} = 2.38, 2.55, 2.75 and 2.95 Å) and the bond-breaking reaction, $H_3O^+ \rightarrow H_2O + H^+$ is also included.

The magnitudes of the intra-molecular physical quantities are quite different from those of the inter-molecular quantities. For example, the total intramolecular energy variation of H_2O reaches about 40 kcal/mol while the two-body interaction energy between H_2O molecules alters within $-5 \sim 12$ kcal/mol. We thus have fitted the potential parameters for the intra- and for inter-molecular quantities independently by using the following 'step by step' procedure:

(*i*) The chemical potential and the hardness of the isolated oxygen and hydrogen atoms [equations (3) and (4)] and the parameters in $U_{\xi\zeta}$ [equation (39)] are determined by fitting *inter*molecular quantities at the asymptotic region where the intermolecular electrostatic interactions are dominant [25].

(*ii*) The parameters in ω_1 [equation (35)] are determined by fitting *inter*molecular quantities, which depend on electron densities, for all configurations except proton transfer reaction of $H_5O_2^+$. Other *inter*molecular quantities, which are independent of electron densities, are used to determine the parameters of f_{rep} [equation (43)].

(*iii*) The parameters of f_θ [equation (41)] are then determined from fitting of *intra*- and *inter*molecular quantities.

In each step (*i*)–(*iii*), the other adjustable parameters are fixed. These steps are repeated until the reference *ab initio* values are reproduced in a certain accuracy.

We used the nonlinear fitting of the 'modified Marquardt method' [26]. In this nonlinear fitting process, we weighted the factors on individual quantities, depending on their importance and magnitude; we put large factor for an important quantity and for a quantity of small magnitude. For example, we put large factor for the intermolecular charge transfer term in the proton transfer of $H_5O_2^+$, since this term was found to be one of the most important key elements to determine the proton transfer PES [27]. We also multiplied larger factors to the intermolecular quantities of H_4O_2 than to those of $H_5O_2^+$, since their magnitudes of H_4O_2 are much less than those of $H_5O_2^+$. The parameters obtained by this nonlinear fitting procedure are listed in Tables A.I–A.IV in Appendix.

By using these obtained parameters, we have calculated the physical quantities by the present AIM model and compared them with *ab initio* data. The correlation between an AIM data and a corresponding *ab initio* data is summarized in Tables 1

Table 1 The correlation between AIM data and *Ab-initio* data of intra- and intermolecular quantities for H_4O_2 and $H_5O_2^+$.

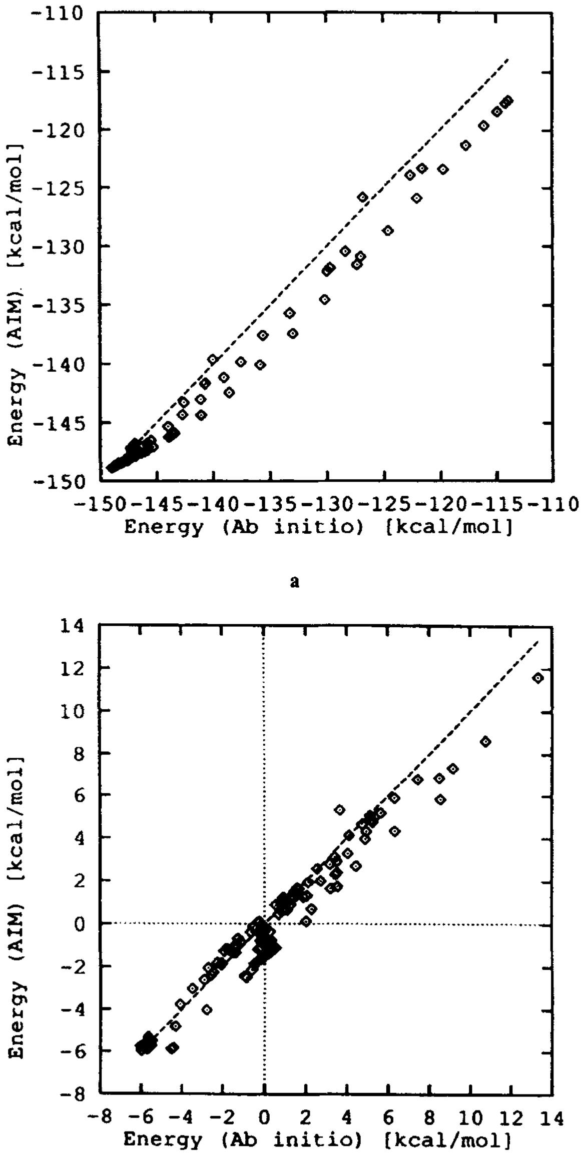
reference quantities	H_4O_2			$H_5O_2{}^+$		
	$n^{a)}$ (*range*$^{b)}$)	$s^{c)}$	$r^{d)}$	*n (range)*	*s*	*r*
intermolecular quantities						
Interaction energy [kcal/mol]	206(−6.0 ~ 13.3)	0.9	0.983	231(−79.7 ~ 8.7)	1.6	0.998
Force acting on bond length [kcal/mol. Bohr]	1030(−21.5 ~ 25.2)	0.7	0.979	1386(−42.5 ~ 85.4)	3.6	0.973
Force acting on bending and torsional angles [kcal/mol. rad]	1442(−11.5 ~ 22.0)	0.8	0.950	2079(−22.5 ~ 22.5)	1.6	0.907
Dipole moment vector [Debye]	618(−3.9 ~ 1.6)	0.3	0.994	924(−4.4 ~ 3.7)	0.7	0.975
Charge transfer$^{e)}$ [au]				462(−0.43 ~ 0.43)	0.04	0.972
intermolecular quantities						
Energy$^{f)}$ [kcal/mol]	206(−304 ~ 285)	0.8	0.979	231(−197 ~ −155)	1.4	0.996
Force acting bond length [kcal/mol. Bohr]	1030(−28.4 ~ 39.1)	0.7	0.986	1386(−42.5 ~ 72.1)	3.4	0.904
Force acting on bending and torsional angles [kcal/mol. rad]	1442(−33.1 ~ 30.8)	1.1	0.950	2079(−22.9 ~ 22.8)	1.3	0.866
Dipole moment vector [Debye]	618(−1.9 ~ 3.4)	0.2	0.999	693(−4.2 ~ 7.0)	0.4	0.990
Atomic Mulliken charge [au]	1236(−0.82 ~ 0.41)	0.06	0.998	1617(−0.88 ~ 0.74)	0.08	0.995

[a)] the number of sample points. [b)] the range of value to be fitted. [c)] the standard error. [d)] the correlation coefficient. [e)] the amount of charge transfer between H_2O molecules in H_4O_2 is found to be in between −0.04 and 0.04 (*ab initio*) and in between −0.02 and 0.02 (AIM). Therefore, this reference quantity of H_4O_2 was not weighted in the parameter fitting procedure. [f)] the interatomic energies $E - \Sigma E_{X^0}$, where E and E_{X^0} denotes the total energy of the system and of the isolated atom X, respectively. Here, the total energy of the isolated oxygen and hydrogen atom is −74.707703 hartree (8 electrons, triplet) and −0.499278 hartree (1 electron, doublet) at UHF/MIDI-4**, respectively.

Table 2 The correlation between AIM data and *Ab-initio* data H_2O and H_3O^+.

reference quantities	*H_2O*			*H_3O^+*		
	n[a](*range*[b])	*s*[c]	*r*[d]	*n(range)*	*s*	*r*
Energy[e] [kcal/mol]	700(− 149.0 ~ − 113.9)	1.5	0.984	318(− 14.7 ~ 61.4)	0.9	0.998
Force acting on bond length [kcal/mol.Bohr]	1400(− 67.1 ~ 129.0)	3.0	0.990	954(− 76.6 ~ 76.0)	2.3	0.994
Force acting on bending and torsional angles [kcal/mol.rad]	700(− 41.1 ~ 93.2)	4.5	0.916	954(− 41.2 ~ 42.4)	2.6	0.954
Dipole moment vector [Debye]	1400(0.0 ~ 2.1)	0.2	0.987	954(− 2.1 ~ 3.1)	0.2	0.990
Atomic Mulliken charge	2100(− 0.88 ~ 0.44)	0.08	0.999	1272(− 0.80 ~ 0.84)	0.13	0.989

[a] the number of sample points. [b] the range of value to be fitted. [c] the standard error. [d] the correlation coefficient. [e] see the footnote [f] of Table 1.



a

Figure 3

b

and 2 for H_2O, H_4O_2, H_3O^+ and $H_5O_2^+$. Their correlation for intra- and intermolecular energies is plotted in Figure 3.

We can see from these Tables and Figure that the present AIM model gives the good correlation with the *ab initio* data for both intra- and intermolecular quantities, energies, forces and dipole moment vectors, despite that their magnitudes are mutually quite different. The deviations in intra- and inter-molecular energies are at

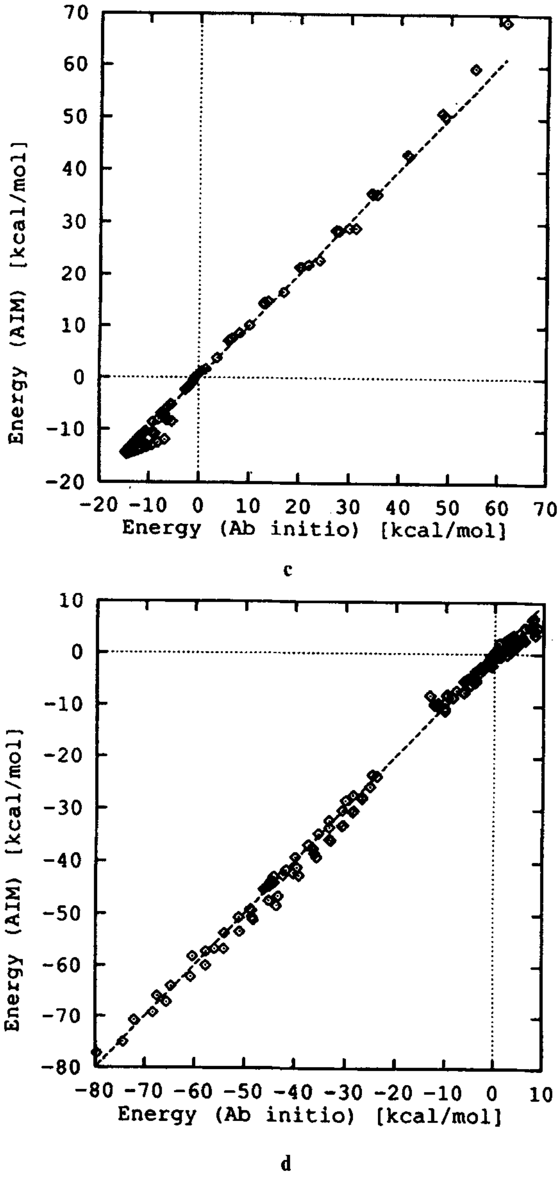


Figure 3 The correlation of the energies between *ab initio* RMF/MIDI-4** calculation and the present AIM model calculation. (a) for H_2O, (b) for $(H_2O)_2$, (c) for H_3O^+ and (d) $H_5O_2^+$. The energies [kcal/mol] in (a) and (c) are the interatomic energies (see the footnote in Table 1).

most 2–3 kcal/mol. The inter-molecular charge transfer is also well reproduced by the present AIM model (see Tab. 1).

The proton transfer PESs of an $(H_2O + H_3O^+)$ system (see Fig. 4) obtained by the present AIM model and *ab initio* calculation are shown in Figs. 5 and 6. The present AIM model gives the proton transfer PES in good agreement with the *ab initio* PES.

The absolute value of the intermolecular potential energy is off about 2 kcal but the overall shape of the potential surface is well reproduced in the present AIM calculation; there is a large energy barrier along the proton transfer for a large O—O distance (Fig. 6) and almost no barrier for a small O—O distance (Fig. 5). The PES by the AIM model may be even in better agreement with the *ab initio* PES if the weighting factor on the proton transfer PES is increased in the fitting procedure.

The interaction matrices [equation (29)] and the corresponding clustered interaction matrices at the points indicated by a, b, b′ and c in Figures 5 and 6 are shown in Fig. 7. When the proton transfers between the oxygen atoms of two water molecules whose mutual distance is small (Fig. 5), the requisite CCTS near the transition

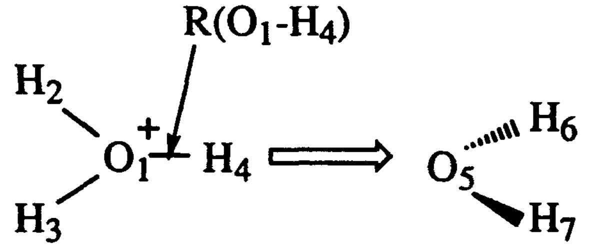


Figure 4 Proton transfer reaction in a ($H_2O + H_3O^+$) system. H_4 is the hydrogen transferring from one oxygen O_1 to the other oxygen O_2.

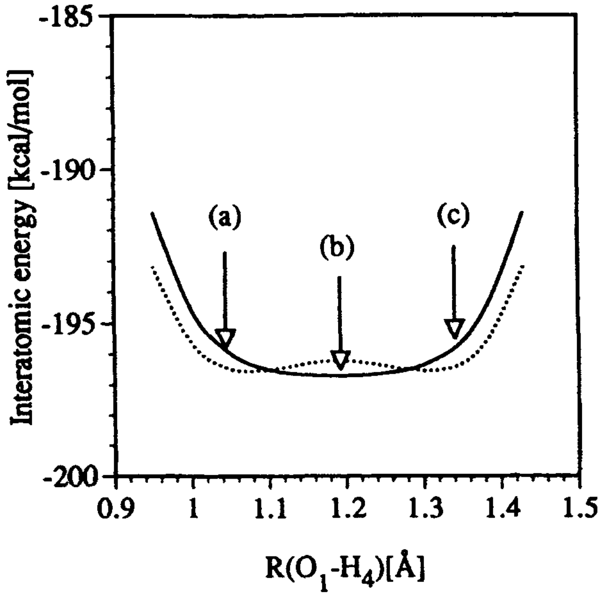


Figure 5 The proton transfer PES in the ($H_2O + H_3O^+$) system with a short distance O–O (2.379 Å). The geometries are taken to be those obtained by *ab initio* calculation; all degrees of freedom are optimized along the proton movement except that $R(O_1-O_5)$ is fixed to be 2.379 Å. The solid and dash lines refers the energy obtained by the present AIM model and that obtained by *ab initio* calculation, respectively. The energies are in the interatomic energy (see the footnote in Table 1).

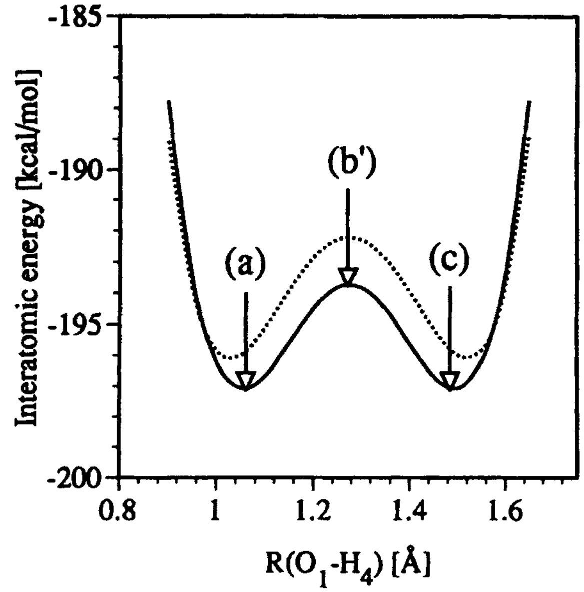


Figure 6 The proton transfer PES in the ($H_2O + H_3O^+$) system with an intermediate O–O distance (2.55 Å). The geometries are taken to be those obtained by *ab initio* calculation; all degrees of freedom are optimized along the proton movement except that $R(O_1–O_5)$ is fixed to be 2.55 Å. The solid and dash lines refers the energy obtained by the present AIM model and that obtained by *ab initio* calculation, respectively. The energies are in the interatomic energy (see the footnote in Table 1).

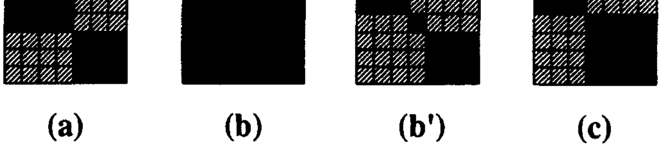


Figure 7 The interaction and the clustered interaction matrixes at the points, (a), (b), (b') and (c), indicated on the PESs of the present AIM model in Figures 5 and 6. Each row (column) corresponds to an atom numbered in Figure 4; for example, 4-th row (column) corresponds to the transferring hydrogen H_4 (atom 4) in Figure 4. The dark-, light- and non-shaded boxes indicate the (clustered) interaction matrix elements with the values, 2, 1 and 0, respectively [equation (29) and C_{ij}].

state is single, a cluster with one block involving the charge reorganization over all atoms (Figure 7b). When a proton transfers between two oxygen atoms intermediately apart (Figure 6), the requisite CCTS near the transition state consists of two type of clusters, one with three blocks of the two individual water molecules and a bare proton, $[H_2O]—[H]^+—[H_2O]$ (Fig. 7b′) and the other with one block involving the charge reorganization over all atoms, $[H_2O—H—H_2O]^+$ (Fig. 7b). The requisite CCTS of the product (Fig. 7c) is anti-symmetric to that of the reactant (Fig. 7a).

In the case of water dimer, H_4O_2, the individual water molecule has almost identical molecular chemical potential, which might slightly alter with the intramolecular geometrical change caused by the intermolecular interaction. The resultant molecular charge transfer between water molecules is thus found to be negligible (see footnote in Tab. 1). Unless the bond-breaking and-forming in H_4O_2 is not considered, a multi-CCTS calculation is almost equivalent to the single CCTS calculation even for the strongly interacting two water molecules, because only very small intermolecular charge transfer is involved.

4 CONCLUSION

We have proposed a new semiempirical AIM method to generate the PES of the proton transport in water, which can be easily implemented in a molecular dynamics calculation. The method can take account of the effects beyond classical polarization, *i.e.*, intra- and intermolecular charge reorganization. The method is based on the extended AIM method in the density functional theory. We have introduced the 'clustering of charge transferable space (CCTS)' in our AIM model in order to deal with a charge transfer reaction. This is because the traditional AIM methods involve the well-known discrepancy, the 'fractional molecular charges' in a large molecular separation. The present multi-CCTS AIM model is free from this discrepancy and can even describe charge transfer reactions or ionic molecular interactions. It was found that the present AIM method can well reproduce many physical quantities and provide a reliable PES for the proton transfer in at least small systems. We must extend our calculation to see how well the present model can be applied to larger systems such as a proton transfer in liquid water and analyze how well the present model can reproduce three-body charge transfer effect, which was found to be one of most important interaction in the proton transfer in water [27], by choosing proper CCTSs.

We have proposed a method calculating the force in the present AIM scheme. In an MD calculation, the forces must be calculated along dynamics. In a single CCTS case, the calculation of forces is relatively simple; it is obtained without evaluating the derivative of electron density with respect to nuclear coordinates. We have found the procedure, equations (19)–(28), to calculate the forces in multi-CCTS, which is essential to describe the reaction involving bond-breaking and -forming.

It is important to select proper CCTS to describe a reaction. We have developed a method to select CCTS based on the physical intuition. A new method to select CCTS based on more sound physical basis should be developed. For certain reactions, we may need to expand a functional form of the atomic interaction beyond the quadratic form with respect to the electron densities used in the present approximation [28].

Acknowledgements

The major part of the present work was carried out during the doctor course of T. K. (1991–1994) at Institute for Molecular Science. T. K. thanks to Dr. S. Takada, Mr. M. Matsumoto, Dr. S. W. Rick and Dr. M. Nagaoka for stimulating discussions. We also thank Mr. M. Matsumoto for providing us his source code of a fast clustering algorithm. This work was partially supported by the grants in aid for Scientific Research on Priority Area of "Theories of Chemical Reactions", "Photochemical Reactions", "Manybody Chemical Systems" and "Complex Fluid" from the Ministry of Education in Japan for I. O. The calculations were carried out with HITAC S820 and M680 at the computer center in IMS and the workstations in I. O. laboratory and in the Institute for Fundamental Chemistry.

APPENDIX

In general, basis functions x^n ($n = 1, 2, 3, \ldots.$) in a simple polynomial expansion are mutually strongly correlated in the optimization of f_θ [equations (41) and (42)] and it is, thus, rather difficult to find the optimal expansion coefficients c_n. We therefore use, instead, Chebyshev orthonormal polynomials [29],

$$T(x) = \sum_{n=0}^{n_T} \bar{c}_n \cos(n \cos^{-1} y), \; (x_{\text{min}} \leqslant x \leqslant x_{\text{max}}),$$

where

$$y = \frac{x - \frac{1}{2}(x_{\text{max}} + x_{\text{min}})}{\frac{1}{2}(x_{\text{max}} - x_{\text{min}})},$$

for the intramolecular x and the intermolecular x. We define the maximum and minimum values of valuables x, x_{min} and x_{max}, to separate the intra- and intermolecular regions; x_{max} of the intramolecular region is equal to the x_{min} of the intermolecular region. The expansion coefficients $\bar{c}_n$ chosen to smoothly connect the function f_θ of the intermolecular x to that of the intramolecular x. ϑ_n are also determined by using Chebyshev orthonormal polynomials, where x_{min} and x_{max} is -1 and $+1$, respectively.

For f_{rep} [equation (43)], x, x_{max} and x_{min} is replaced by x', x'_{max} and x'_{min}, respectively, where

$$x' = \log(x - x_0), \quad (x'_{\text{min}} \leqslant x' \leqslant x'_{\text{max}}),$$

where

$$x'_{\text{max}} = \log(x_{\text{max}} - x_0), \quad x'_{\text{min}} = \log(x_{\text{min}} - x_0),$$

$$x_0 = x_{\text{min}} - d.$$

The Chebyshev coefficients (*i.e.*, $\bar{c}_n$) for the transformed variable can be easily converted to the expansion coefficients (*i.e.*, c_n) for the original variable [29].

The optimized adjustable parameters are summarized in Tables A.I–A.IV.

Table A. I The optimized potential parameters.

	μ_{X0} [au]	η_{X0} [au]
H	-0.200	0.852
O	-0.278	0.460
	$U_{\xi\zeta}$	
	v_1	v_2
core-core		
$O_n - H_n$	0.9843	0.1820
$H_n - H_n$	3.2765	0.1173
$O_n - O_n$	0.0284	0.1218
electron-electron		
$O_e - O_e$	1.7103	0.1317
$O_e - H_e$	1.3319	0.0458
$H_e - H_e$	9.7847	0.2990
core-electron		
$O_e - O_n$	1.0788	0.1257
$O_e - H_n$	1.1442	0.1204
$O_n - H_e$	1.1910	0.0785
$H_e - H_n$	5.2521	0.2695
	R_0[au]	β[au^{-1}]
O—H	2.835	1×10^4
H—H	2.835	1×10^4
O—O	5.597	1×10^4

Table A. II The optimized optential parameters of f_{rep}.

	H—H	O—H	O—O
$\bar{c}_n$	$[\times 10^{-3}]$	$[\times 10^{-1}]$	$[\times 10^{2}]$
$\bar{c}_0$	8.981	-8.968	3.299
$\bar{c}_1$	-7.916	-8.809	3.008
$\bar{c}_2$	-9.135	-6.321	2.271
$\bar{c}_3$	-1.961×10^1	-3.971	1.402
$\bar{c}_4$	-1.340×10^1	-1.837	6.896×10^{-1}
$\bar{c}_5$	-5.132	-6.673×10^{-1}	2.561×10^{-1}
$\bar{c}_6$	3.497	5.685×10^{-3}	6.368×10^{-2}
$\bar{c}_7$	5.931	2.414×10^{-2}	6.649×10^{-3}
$\bar{c}_8$	3.628	-1.179×10^{-2}	-1.345×10^{-3}
$\bar{c}_9$	1.546	1.111×10^{-2}	-3.890×10^{-4}
n_T	9	9	9
d[au]	0.549	0.549	0.549
α[au^{-1}]	2.201	2.201	2.201
R_c[au]	5.670	5.670	5.670
x_{min} [au]	1.952	1.483	4.406
x_{max} [au]	26.846	25.811	24.614

Table A.III The optimized potential parameters of f_θ.

	O—H	O—O		
$\bar{c}_n$*(intra)*	$[\times 10^{-1}]$	$[\times 10^{1}]$	$\bar{\theta}_n^{a)}$	$[\times 10^{-2}]$
$\bar{c}_0$	4.769×10^{1}	-1.627×10^{1}	$\bar{\theta}_0$	-2.477×10^{1}
$\bar{c}_1$	1.045×10^{1}	-3.783	$\bar{\theta}_1$	-6.357
$\bar{c}_2$	-4.497	-5.683	$\bar{\theta}_2$	2.351
$\bar{c}_3$	-1.220	1.911	$\bar{\theta}_3$	-2.649
$\bar{c}_4$	4.511×10^{-1}	7.684	$\bar{\theta}_4$	-5.986×10^{-1}
$\bar{c}_5$	7.164×10^{-1}	2.483	$\bar{\theta}_5$	-5.143×10^{-1}
$\bar{c}_6$	1.074		$\bar{\theta}_6$	3.116×10^{-2}
			$\bar{\theta}_7$	1.779×10^{-1}
n_T*(intra)*	6	5	n_θ	7
n_T*(intra)*$^{b)}$	1	1		
$\alpha[\mathrm{au}^{-1}]$	2.201	2.201		
$R_c[\mathrm{au}]$	1.890	1.890		
intra-region				
$x_{\min}[\mathrm{au}]$	1.483	4.406		
$x_{\max}[\mathrm{au}]$	5.000	8.000		
intra-region				
$x_{\min}[\mathrm{au}]$	5.000	8.000		
$x_{\max}[\mathrm{au}]$	25.811	24.614		

intra and *inter* denotes the *intra*- and *inter*-molecular region, respectively. *a)* $\bar{\theta}_n$ are the *n*-th Chebyshev coefficient for *n*-th Chebyshev polynomial. *b)* we used the value of n_T(*inter*) to require the number of the corresponding adjustable variables, $\{\bar{c}_n(\mathit{inter})\}$ to connect in energies and in their derivatives at the boundary to be minimum. Therefore, $\{\bar{c}_n(\mathit{inter})\}$ are definitely determined for the given $\{\bar{c}_n(\mathit{intra})\}$.

Table A.IV The optimized potential parameters of ω_1.

η_1 [au]	0.200
η_A [au]	0.250
η_B [au]	0.270
η_C [au]	0.380
η_D [au]	0.385
η_2 [au]	0.390
n_1	3
n_2	3

we used the values of n_1 and n_2 to require each number of the corresponding adjustable variables, $\{a_n\}$ and $\{b_n\}$ to connect in functions and in their derivatives at the two boundaries to be minimum. Therefore, $\{a_n\}$ and $\{b_n\}$ are definitely determined for the given η_ω ($\omega = 1, 2, A, B, C$ and D).

References

[1] N. Agmon, "The Grotthuss Mechanism", *Chem. Phys. Lett.*, **244**, 456 (1995); K. Ando and J. T. Hynes, "Molecular Mechanism of HCl Acid Ionization in Water", private communication; R. Pomès and B. Roux, "Quantum effects on the structure and energy of a protonated linear chain of hydrogen-bonded water molecules", *Chem. Phys. Lett.*, **234**, 416 (1995); F. H. Stillinger in *Theoretical Chemistry. Advances and Perspectives* Vol. 3, edited by H. Eyring and D. Henderson, (Academic Press, New York 1978).

[2] J. T. Hynes, *Theory of Chemical Reaction Dynamics* Vol. 4, edited by Baer (Chemical Rubber, Florida 1985).

[3] I. Ohmine and H. Tanaka, "Fluctuation, Relaxations, and Hydration in Liquid Water", *Chem. Rev.*, **93**, 2545 (1993).

[4] T. A. Weber and F. H. Stillinger, "Dynamical study of the $H_5O_2^+ + H_3O_2^-$ neutralization reaction", *J. Chem. Phys.*, **77**, 4150 (1982).

[5] M. Tuckerman, K. Laasonen, M. Sprik and M. Parrinello, "*Ab initio* molecular dynamics simulation of the solvation and transport of H_3O^+ and OH^- ions in water", *J. Phys. Chem.*, **99**, 5749 (1995).

[6] R. T. Sanderson, "Partial charges on atoms in organic compounds", *Science*, **121**, 207 (1955); R. G. Parr, R. A. Donnelley, M. Levy and W. E. Palk, "Electronegativity: The density functional viewpoint", *J. Chem. Phys.*, **68**, 3801 (1978).

[7] R. F. Nalewajski, "A study of electronegativity equalization", *J. Phys. Chem.*, **89**, 2831 (1985).

[8] W. J. Mortier, S.K. Ghosh and S. Shankar, "Electronegativity equalization method for the calculation of atomic charges in molecules", *J. Am. Chem. Soc.*, **108**, 4315 (1986).

[9] S. K. Ghosh and R. G. Parr, "Toward a semiempirical density functional theory of chemical binding", *Theor. Chim. Acta.*, **72**, 379 (1987).

[10] A. K. Rappe and W. A. Goddard III, "Charge equilibration for molecular dynamics simulations", *J. Phys. Chem.*, **95**, 3358 (1991).

[11] S. W. Rick, S. J. Stuart and B. J. Berne, "Dynamical fluctuating charge force fields: Application to liquid water", *J. Chem. Phys.*, **101**, 6141 (1994).

[12] J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, "Density-functional theory for fractional particle number: derivative discontinuities of the energy", *Phys. Rev. Lett.*, **49**, 1691 (1982).

[13] J. Cioslowski and B. B. Stefanov, "Electron flow and electronegativity equalization in the process of bond formation", *J. Chem. Phys.*, **99**, 5151 (1993).

[14] R. G. Parr and R. G. Pearson, "Absolute hardness: companion parameter to absolute electronegativity", *J. Am. Chem. Soc.*, **105**, 7512 (1983).

[15] E. K. U. Gross, L. N. Oliveira and W. Kohn, "Density functional theory for ensembles of fractionally occupied state II: Basic formalism", *Phys. Rev.*, **A37**, 2809 (1988).

[16] D. J. Wales, "*Ab initio* calculation of molecular structure by expansion of the electron density", *Chem. Phys. Lett.*, **217**, 302 (1994).

[17] J. Hoshen and R. Kopelman, "Percolation and cluster distribution. I. cluster multiple labeling technique and critical concentration algorithm", *Phys. Rev.*, **B14**, 3438 (1976); The efficient fast clustering algorithm in an MD calculation can be developed and will be given elsewhere.

[18] R. Pariser and R. G. Parr, "A semi-empirical theory of the electronic spectra and electronic structure", *J. Chem. Phys.*, **21**, 767(1953).

[19] K. Ohno, "Some remarks on the Pariser-Parr-Pople method", *Theor. Chim. Acta.*, **2**, 219 (1964).

[20] R. F. Nalewajski, J. Korchowiec and Z. Zhou, "Molecular hardness and softness parameters", *Int. J. Quantum. Chem. Symp.*, **22**, 349 (1988).

[21] J. A. Pople and G. A. Segal, "Approximate self-consistent molecular orbital theory. III.", *J. Chem. Phys.*, **44**, 3289 (1966).

[22] N. C. Baird and M. J. S. Dewar, "Ground states of sigma-bonded molecules. IV.", *J. Chem. Phys.*, **50**, 1262 (1969).

[23] H. Nakatsuji, "Electrostatic force theory for a molecule and interacting molecules. I.", *J. Am. Chem. Soc.*, **95**, 345 (1973).

[24] M. D. Newton and S. Ehrenson, "*Ab initio* studies on the structures and energetics of inner- and outer-shell hydrates of the proton and the hydroxide ion", *J. Am. Chem. Soc.*, **93**, 4971 (1971).

[25] K. Kitaura and K. Morokuma, "A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation", *Int. J. Quantum. Chem.*, **10**, 325 (1976).

[26] T. Nakagawa and Y. Koyanagi in '*Recent Developments in Statistical Inference and Data Analysis*', edited by K. Matsusita (North Holland, Amsterdam, 1980).

[27] T. Komatsuzaki and I. Ohmine, "Energetics of proton transfer in liquid water. I.", *Chemical Physics*, **180**, 239 (1994).

[28] D. J. Wales, A. M. Lee and N. C. Handy, "Can Thomas-Fermi theory describe chemical bonding?", submitted to *Phys. Rev. Lett.*

[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, "*NUMERICALRECIPES: The art of scientific computing*", 2nd edition, (Cambridge University Press, New York, 1992).