Chemical Physics 472 (2016) 24–35

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Fast calculation of molecular total energy with ABEEMσπ/MM method – For some series of organic molecules and peptides

Zhong-Zhi Yang *, Xiao-Ting Lin, Dong-Xia Zhao *

School of Chemistry and Chemical Engineering, Liaoning Normal University, Dalian 116029, China

ARTICLE INFO

Article history:
Received 15 October 2015
In final form 6 March 2016
Available online 14 March 2016

Keywords:
Molecular total energy
ABEEMσπ/MM method
Organic molecules and peptides
Fast calculation
High accuracy

ABSTRACT

A new ABEEMσπ/MM method for fast calculation of molecular total energy is established by combining ABEEMσπ model with force field representation, where ABEEMσπ is the atom-bond electronegativity equalization model at the σπ level. The calibrated parameters are suitable and transferable. This paper demonstrates that the total molecular energies for series of alcohols, aldehydes, carboxylic acids and peptides calculated by ABEEMσπ/MM method are in fair agreement with those obtained from calculations of ab initio MP2/6-311++G(d, p) method with mean absolute deviation (MAD) being 1.45 kcal/mol and their linear correlation coefficients being 1.0000. Thus it opens good prospects for wide applications to chemical and biological systems.

1. Introduction

Energy is basic for describing various static and dynamic properties of a molecular system. Total energy of a molecule is an important physical quantity in science. Generally, it is calculated by solving Schrödinger equation under Born–Oppenheimer approximation, which can be realized by using ab initio methods at various levels. However, as well known, this solving procedure with high accuracy is only feasible for small molecules due to computational limitation. Even using well-developed density functional method, this is also limited to mediate molecules that contain a few hundreds of atoms. But, we often need to treat macromolecules, such as a bio-molecule, or a condensed system that contains hundreds of thousands atoms. Thus one wants to develop fast methods to treat those large molecular systems yet with proper accuracy for calculating total energies or potential functions. It is particularly noted that in recent years great efforts have been made to develop approximate methods to calculate the ground-state energy of a large molecular systems with high accuracy and speed, such as generalized energy-based fragmentation approach (GEBF) by Li and his coworkers [1], quantum mechanical fragment methods based on partitioning atoms or partitioning coordinates by Truhlar and his coworkers [2,3], and all the papers in the special issue of Accounts of Chemical Research, Beyond QM/MM: Fragment Quantum Mechanical Methods [4].

Well-developed various force fields offer simple potential energy functions for large complicated molecular systems and have been widely used in molecular dynamic (MD) simulations for understanding and predicting their structures and properties [5,6]. In order to improve accuracy and reliability of MD simulations, a special issue of Journal of Chemical Theory and Computation in 2007 edited by Jorgensen has called upon to develop the polarizable force fields [7]. Actually, a great deal of efforts and designs have been made to pursue this. In this respect, CHARMM [8–11], OPLS [12,13], NEMO [14,15], AMBER [16–19], AMOEBA [20,21], ABEEM [22–25] etc. force fields have brought about fair progresses in terms of various approaches of adequately treating the polarization effect in the molecular systems. In some polarizable force fields, such as in AMOEBA force field, the potential energy function is expressed by containing some orders of induced multipoles, such as dipole moments, quadrupole moments, and even higher order multipoles at atomic sites. In addition, it is noticed that Xie and Gao have designed and developed a new generation force field, the X-POL potential, which is based on a quantum mechanical model in order to avoid deficiencies in the usual force fields [26]. Besides, Mackerell and his coworkers have developed CHARMM Drude polarizable force field and investigated nucleic acid bases, peptides and proteins as well as accurate calculation of the hydration free energies [27–31]. The van Gunsteren group are developing polarizable force fields using Drude-type approaches for molecular dynamics simulation of liquid hydrocarbons [32]. The polarizable force fields also include inducible point dipole models, such as the AMBER FF02 polarizable force field, Duan et al. have developed it for simulations of proteins and peptides [19]. Although the polarizable force fields have been developed since 1970s [33], broadly applicable polarizable force fields

---

* Corresponding authors.

http://dx.doi.org/10.1016/j.chemphys.2016.03.004
0301-0104/© 2016 Elsevier B.V. All rights reserved.

have not emerged and simulations of biomolecular systems with polarizable force fields are still uncommon, as pointed out by Jorgensen in the special issue on polarization [7], in which we may know more details.

At the same period of time, a few fluctuating charge models for polarizable force fields have also been developed and applied by some groups, such as Rick [34], Patel and Brooks III [35], van der Graaf [36,37], Rappé and Goddard [38,39], Chelli and Procacci [40,41], Stern and Berne [42,43], and Yang and his coworkers [22–25,44–48], and so on. The fluctuating charge model for calculating the charge distribution is based on the electronegativity equalization method (EEM) [49,50], the charge equilibration (QEq) [51], chemical potential equalization principle (CPE) [52], or the atom-bond electronegativity equalization method (ABEEM) [53,54], which fall in category of conceptual density functional theory (DFT) [55–57]. The advantage of these methods is its fast computation by only solving a set of linear equations for charge distribution without any iteration procedure in the minimization or simulation process.

Here, a new method of fast calculation of molecular total energy is proposed by combining ABEEM$\sigma\pi$ method (atom-bond electronegativity equalization model at $\sigma\pi$ level) with a polarizable force field scheme, so that it has advantages of the two methods. It is exemplified by calculating total energies of series of alcohols, aldehydes, carboxylic acids as well as peptides. The calculation results obtained by ABEEM$\sigma\pi$/MM method are in fair agreement with those obtained by ab initio method at MP2/6-311++G(d,p) level with high accuracy.

The paper is organized as follows: Section 2 describes the methods including the formalism of ABEEM$\sigma\pi$/MM and the ab initio methods employed; Section 3 presents the results and discussion; Section 4 provides a brief conclusion.

2. Methods

2.1. ABEEM$\sigma\pi$/MM method

Historically, ABEEM$\sigma\pi$ model or method has experienced the following three stages: (1) first, it partitions a molecule into atomic and bond regions; [53] (2) second, the lone pair electron regions are added to those of the first [54]; (3) based on that of the second stage, the regions of σ bond and π bond are distinguished and added, which is named ABEEM$\sigma\pi$ model or method [22–24,44,58]. ABEEM$\sigma\pi$ model is able to give systematically fair charge distributions for large chemical and biological molecular systems, which can be used to investigate various properties. Here, we will develop it to calculate the total energy of a molecule with very high speed and fair accuracy comparable to ab initio results at MP2/6-311++G(d,p) level as benchmark by using the Gaussian 09 [59].

In Born–Oppenheimer approximation, based on the density functional theory (DFT) [56], the total energy of a molecule in the ABEEM$\sigma\pi$ model [24,53,54,58], E_{ABEEM}, can be expressed as:

$$\begin{aligned} E_{\text{ABEEM}} &= \sum E_i^{\text{intra}} + \sum E_{i,j\neq i}^{\text{inter}} \\ &= \sum_i (E_i^* + \chi_i^* q_i + \eta_i^* q_i^2) + k \sum_{i<j} q_i q_j / R_{ij} \end{aligned} \tag{1}$$

where i and j denote the regions partitioned a molecule in some way, E_i^{intra} is the intra-region i contribution to the total molecular energy, and $E_{i,j\neq i}^{\text{inter}}$ is the inter-region contribution to the total molecular energy. $E_i^*, \chi_i^*, \eta_i^*$, are valence state energy, electronegativity, and hardness of region i of the molecule, respectively. q_i and q_j are the partial charges of regions or sites i and j. k is a adjusted parameter. R_{ij} denotes the distance between sites i and j. Here, E_{ABEEM} is explicitly expressed through partial charges of all regions while all the rest are attributed to the $\sum_i E_i^*$ that are implicitly related to the partial charges in order to simply derive and to express the electronegativity equalization equation [53,54]. In ABEEM$\sigma\pi$ method, the sites include the atoms, σ and π bonds, as well as lone pairs [22,24,58].

In a usual force filed, the potential energy can be written as the sum of several terms which explicitly describe various interaction energies between different pairs in a molecule [5,6]. The energy of the molecule is expressed as:

$$E_{\text{FF}} = \sum_{\text{bonds}} E_r + \sum_{\text{angles}} E_\theta + \sum_{\text{torsions}} E_{\text{torsions}} + \sum_{\text{imptors}} E_{\text{imptors}} + \sum_{\text{vdW}} E_{\text{vdW}} + \sum_{\text{elec}} E_{\text{elec}} \tag{2}$$

where E_r is the stretching vibration energy of a bond between two atoms, E_θ is an angle bending energy between two bonds involved three atoms, E_{torsions} and E_{imptors} are torsion and improper dihedral angle energies involved four atoms, respectively. These terms mentioned above represent the energies between the bonded atoms. For the non-bonded interaction energies, E_{vdW} and E_{elec} are van der Waals and electrostatic interaction energies, respectively. All the terms are explicitly given in the following discussion.

One evaluates the E_r value of the energy at an arbitrary bond length by taking a Taylor expansion about the equilibrium bond length, r_{eq},

$$\begin{aligned} E_r = E(r_{\text{eq}}) + \left.\frac{dE}{dr}\right|_{r=r_{\text{eq}}} (r - r_{\text{eq}}) + \left.\frac{d^2E}{dr^2}\right|_{r=r_{\text{eq}}} (r - r_{\text{eq}})^2 \\ + \left.\frac{d^3E}{dr^3}\right|_{r=r_{\text{eq}}} (r - r_{\text{eq}})^3 + \cdots \end{aligned} \tag{3}$$

Note that the first term of Eq. (3), $E(r_{\text{eq}})$, denotes the energy at equilibrium bond length and is a constant, and the second by virtue of r_{eq} being the minimum is equal to zero. If we truncate those terms after the second non-zero term, we have the simplest possible expression for the stretching vibration energy of a bond between two atoms.

$$\begin{aligned} E_r &= E(r_{\text{eq}}) + \left.\frac{d^2E}{dr^2}\right|_{r=r_{\text{eq}}} (r - r_{\text{eq}})^2 \\ &= E(r_{\text{eq}}) + k_r (r - r_{\text{eq}})^2 \end{aligned} \tag{4}$$

where k_r represents force constant of the stretching, r is the actual bond length, and r_{eq} denotes the equilibrium bond length. Eq. (4) is the simplest harmonic expression, and one also uses other forms of empirical functions, such as the Morse function.

By the similar way, the bending energy, E_θ, can be expressed as:

$$E_\theta = E(\theta_{\text{eq}}) + k_\theta (\theta - \theta_{\text{eq}})^2 \tag{5}$$

where $E(\theta_{\text{eq}})$, denotes the energy at equilibrium bond angle, and is a constant, k_θ represents force constant of the bending, θ is the actual bond angle, and θ_{eq} denotes the equilibrium bond angle.

The torsion and improper dihedral angle energies, E_{torsions} and E_{imptors} can be written,

$$\begin{aligned} E_{\text{torsions}} + E_{\text{imptors}} = \left[\frac{V_1}{2}(1 + \cos\phi) + \frac{V_2}{2}(1 - \cos 2\phi) \right. \\ \left. + \frac{V_3}{2}(1 + \cos 3\phi)\right] + \nu(1 - \cos 2\omega) \end{aligned} \tag{6}$$

in which V_1, V_2, V_3 and ν are force constants of the dihedral angle (ϕ) and improper dihedral angle (ω), respectively.

The non-bonded energy can be written as the sum of the van der Waals interaction energy and electrostatic energy.