Chemical Physics Letters 556 (2013) 272–277

Contents lists available at SciVerse ScienceDirect

Chemical Physics Letters

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

Statistical correction to effective interactions in the fragment molecular orbital method

Shigenori Tanaka [a,*], Chiduru Watanabe [b], Yoshio Okiyama [b]

[a] Graduate School of System Informatics, Department of Computational Science, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan
[b] Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

ARTICLE INFO

Article history:
Received 18 August 2012
In final form 27 November 2012
Available online 6 December 2012

ABSTRACT

A theoretical scheme to evaluate effective, screened interactions between fragments is proposed within the framework of the fragment molecular orbital (FMO) method. In this theory, the presence of implicit, dielectric continuum solvent is not assumed, but only the information on bare, inter-fragment interaction energies obtained through the FMO calculation for explicit, molecular system is employed. The effective interactions with inclusion of entropic effect are then described and optimized as a consequence of inter-fragment correlations on the basis of classical-mechanical many-body theories. Test calculations for a simple model system and a realistic protein system are performed, and their implications are discussed.

1. Introduction

Since its proposal in 1999 [1], the fragment molecular orbital (FMO) method [2–5] has provided a powerful and useful tool to perform *ab initio* electronic-state calculations for biomolecular and other related systems. One of very advantageous features in the FMO method for biomolecular analyses is its ability to evaluate "effective interactions" between fragments that are usually chosen to be an amino acid residue or a ligand molecule as a unit component in the protein–ligand complex system, for instance. This inter-fragment interaction is referred to as IFIE (Inter-Fragment Interaction Energy) or PIE (Pair Interaction Energy) in the literature [2–7], and plays a vital role in, e.g., docking analysis by specifying important interactions involved in the object system. In fact, the FMO-IFIE analysis has been extensively applied to the investigations of mechanisms of molecular recognition associated with protein–protein [8,9], protein–nucleic acid [10–12], protein–drug [13–17], and other [18–22] intermolecular interactions.

In many FMO studies performed previously, isolated biomolecular systems in vacuum have often been employed for the calculations. The inclusion of solvent effect has also been attempted by explicitly taking into account water molecules as parts of the calculated system [18,23–25]. Then, it has been recognized [26] as a serious difficulty in the FMO analysis that the calculated IFIEs often show unrealistically large values in magnitude from biochemical point of view; the magnitudes of IFIEs do not seem to be screened appropriately even between distant, charged fragments [8,9], and the situation is not improved significantly even by the presence of explicit water molecules surrounding the biomolecules, which indicates that the current FMO-IFIE analysis somehow overestimates the electrostatic interactions.

One possible solution to overcome this difficulty that the IFIE values represent a kind of "bare" interactions between fragments is to consider the dielectric screening effect by solvent or amino acids contained in the system in terms of dielectric constants for these screening materials [27,28]. For example, if we consider the screening effect by water solvent, the effective inter-fragment interactions would be substantially screened and reduced by about a factor of dielectric constant of water, i.e., 78. Thus, Fedorov and Kitaura [26] have recently proposed a computational scheme to evaluate the effective IFIEs in the FMO calculations, which could be applicable also to the description of screening effects due to amino acid residues. However, there are some drawbacks in this approach in common with other "implicit" solvent models: First, there are a lot of uncertainties concerning the modeling, approximation and parameters employed in actual applications. These ambiguities would further be enhanced when one performs some hybrid calculations employing both of explicit and implicit solvents simultaneously or dynamical calculations in which a number of time scales characterizing the screening may be mixed up in the simulations. In addition, it may be difficult for this implicit approach to quantitatively describe some of short-range interactions such as hydrogen-bonding one, which should play an important role in the context of molecular recognition.

In the present work, we study a way to partially correct the screening effects in the FMO-IFIE analysis in terms of classical-mechanical many-body correlations, where there are no "implicit" materials to screen the interactions between fragments. It is

* Corresponding author. Fax: +81 78 803 6621.
E-mail address: tanaka2@kobe-u.ac.jp (S. Tanaka).

0009-2614/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cplett.2012.11.085

remarked that the screening effects due to electronic polarization have already been included in the "bare" IFIE values calculated with the FMO method. The purpose of the present study is to explore a methodology to theoretically describe the effective interactions in the FMO calculations only with the use of fragments composed of explicit atoms. Starting from the IFIE values obtained through an FMO calculation for an isolated system that may or may not contain explicit solvent materials such as water molecules, we attempt to correct the effective interactions by considering the correlation effects between the fragments. We rely on some of theoretical techniques developed in the field of classical many-body problems [29], and thus propose a method to obtain the effective inter-fragment interactions relevant for a given configuration of molecular system. In many molecular simulations employing the FMO method, one often obtains a set of IFIE values for a number of configurations with and without inclusion of solvent molecules. The present approach could then provide a useful methodology to elicit the effective interactions with inclusion of entropic effect for each snapshot configuration, which may further be averaged statistically over multiple configurations to obtain the screened inter-fragment interactions realized in actual circumstances.

In the following section, we first develop a theoretical framework to calculate the effective IFIEs in the FMO analysis. Test calculations employing a simple model system and a realistic protein system are carried out in Section 3, and their implications are discussed. Concluding remarks are finally given in Section 4.

2. Theory

Let us consider a situation that the (symmetric) inter-fragment interaction energies (IFIEs) u_{ij} $(1 \leqslant i,j \leqslant N)$ for a system with N fragments have been given through an FMO calculation. Here, in the case of biopolymers, the IFIEs for the neighboring (bonding) fragments $(j = i \pm 1)$ as well as the diagonal components $(i = j)$ may be set to be zero by definition [2–7]. Our problem is then to evaluate effective interactions w_{ij} between fragments by considering the many-body effects associated with the inter-particle (fragment) correlations and various (density, conformational and energetic) fluctuations in the system. We assume in the following model that the "bare" interactions u_{ij} between the fragments are fixed as average values irrespective of possible conformational changes, and the model system may be mapped onto a classical system with a potential energy,

$$V = \sum_{i<j} u_{ij} n_i n_j + \sum_i \mu_i n_i, \tag{1}$$

where n_i refers to the site density (population) of fragment i and μ_i is the site energy or external ("confinement") potential for fragment i to keep the average density $\langle n_i \rangle$ fixed to be unity. We thus take account of the inter-fragment correlations not by the fluctuation of IFIE values u_{ij} but by the density fluctuation at site i associated with possible variations in conformations and energetics of the system.

Each fragment i in this system may be regarded as a classical particle embedded in the network of mutual interactions. The effective interaction or the potential of mean force w_{ij} is then related to the pair correlation function h_{ij} between the fragments via [29]

$$h_{ij} = e^{-\beta w_{ij}} - 1. \tag{2}$$

The parameter β in Eq. (2) is usually related to the absolute temperature T and the Boltzmann constant k_B via $\beta = (k_B T)^{-1}$, but, in the present formalism, may be regarded as an optimization parameter to control the degree of screening, which will be specified in the following.

The pair correlation function h_{ij} is related to the direct correlation function c_{ij} in terms of the Ornstein–Zernike relation [29]:

$$h_{ij} = c_{ij} + \sum_{k \neq i,j} c_{ik} h_{kj}. \tag{3}$$

The direct correlation function is conceptually introduced as a difference between the total and indirect parts of inter-particle correlation, and may be expressed, in the Percus–Yevick (PY) approximation [29] for classical many-body problem, as

$$c_{ij} = e^{-\beta w_{ij}} - e^{-\beta(w_{ij} - u_{ij})}, \tag{4}$$

providing a closure equation to determine w_{ij} for a given set of u_{ij}.

The procedure above may be formulated as follows. By introducing the Mayer function [29],

$$f_{ij} = e^{-\beta u_{ij}} - 1, \tag{5}$$

we can rewrite Eq. (4) as

$$c_{ij} = \frac{f_{ij}(h_{ij} + 1)}{f_{ij} + 1}. \tag{6}$$

Substituting this relation into Eq. (3), we find

$$h_{ij} = \frac{f_{ij}(h_{ij} + 1)}{f_{ij} + 1} + \sum_k \frac{f_{ik}(h_{ik} + 1)}{f_{ik} + 1} h_{kj}, \tag{7}$$

or

$$h_{ij} = f_{ij} + (f_{ij} + 1) \sum_k \frac{f_{ik}(h_{ik} + 1)}{f_{ik} + 1} h_{kj}. \tag{8}$$

For given u_{ij} and f_{ij}, we can obtain h_{ij} by Eq. (8), and thus find

$$w_{ij} = -\frac{1}{\beta} \ln(h_{ij} + 1). \tag{9}$$

We refer to this way of evaluating the effective interactions w_{ij} as "PY" scheme.

In the formulation above, the effective, screened interactions are obtained through the inter-fragment correlations in a system with given set of bare interactions. The origin of the screening may partially be attributed to the conformational fluctuations and associated entropic effects in the system, but the parameter β should not be chosen according to room temperature. The energy scale of IFIEs obtained by single-point FMO calculation is of electronic origin and hence much higher than that of thermal energy at room temperature. While the electronic state of the pertinent system is very sensitive to the conformational fluctuations due to thermal energy, the relevant value of β in the present screening model should be independent of the thermal energy associated with the system. As will be shown in the following sections, the value of β represents a high energy scale corresponding to the screening by "fast" modes in the system, and its optimal value can be determined by a "maximum screening" ansatz under constraint condition. The statistical averaging over the thermal motion at room temperature should be taken separately by generating many conformations for FMO calculations with, e.g., molecular dynamics simulations.

There is another well-known approximation often employed in classical many-body theory, that is, the hypernetted-chain (HNC) approximation [29]. In this approximation, the direct correlation function is expressed as

$$c_{ij} = e^{-\beta w_{ij}} - 1 + \beta(w_{ij} - u_{ij}) \tag{10}$$

instead of Eq. (4) for the PY approximation. We then find

$$h_{ij} = -1 + \exp\{-\beta u_{ij} + \sum_k [h_{ik} - \ln(h_{ik} + 1) - \beta u_{ik}] h_{kj}\} \tag{11}$$

instead of Eqs. (7) and (8). We refer to this scheme for obtaining h_{ij} and w_{ij} as "HNC" scheme.