Journal of Molecular Structure: THEOCHEM 864 (2008) 31–41

Contents lists available at ScienceDirect

Journal of Molecular Structure: THEOCHEM

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

New flexible boundary hybrid solvation models for efficient biomolecular simulations

Wusheng Zhu, Goran Krilov *

Department of Chemistry, Boston College, Merkert Center, 2609 Beacon Street, Chestnut Hill, MA 02467, USA

ARTICLE INFO

Article history:
Received 3 March 2008
Received in revised form 28 May 2008
Accepted 29 May 2008
Available online 7 June 2008

Keywords:
Molecular dynamics
Hybrid solvation
Polyalanine
Non-periodic boundary condition

ABSTRACT

Two simple models are developed to accurately account for short-range solvation effects in molecular dynamics simulations of biomolecular systems in water. The methods limit the amount of solvent that must be treated explicitly to two or three solvation shells, thereby eliminating the need for periodic boundary conditions, which significantly reduces the computational complexity. In the dynamic boundary model, a confining potential is imposed on the solvent that responds dynamically to fluctuations in the solvent distribution and conformations of the biomolecular solute. In the exchange boundary solvation model, the molecules at the surface of the explicit solvation shell are allowed to undergo pairwise exchanges with the bulk to maintain a uniform hydration of the solute. The models are used to study the solvation thermodynamics of a 11-residue polyalanine helix in water. A broad range of structural and dynamic properties characterizing both the protein and the surrounding solvent are computed. Comparison with the results obtained from conventional periodic boundary condition simulations shows excellent agreement, and the new solvation models are found to improve computational efficiency by up two orders of magnitude.

1. Introduction

The physiological solvent environment plays a critical role in modulating the thermodynamics and dynamics of biological macromolecules, which in turn controls their function. Hence, it is well known that the accurate treatment of solvation is crucial for the success of computer simulations of biomolecular processes [1,2]. A natural approach for incorporating solvation into molecular dynamics simulations is to embed the biomolecular solute in a box of solvent molecules with periodic boundary conditions [3]. However, in order to correctly capture the effects of solvation, the number of solvent molecules that need to be included in the simulation system is very large, with ratio of the number of solvent atoms to that of the solute frequently exceeding ten. For instance, a medium sized globular protein, such as myoglobin, with a diameter of 40 Å would require a 60 Å box filled with 6300 solvent molecules, resulting in the number of solvent atoms eight times larger than those of the solute. This increases the computational cost dramatically, consequently limiting the time scale of such simulations to nanoseconds. On the other hand, most biological processes of interest occur at much longer timescales, of milliseconds to seconds. Since by far the majority of the computational effort is spent on simulating the solvent, devising a more efficient means for treatment of solvation in large scale biomolecular simulations is highly desirable.

A common approach is to replace the expensive averaging over solvent configurations with an effective mean field approximation [4–10]. The polarizable solvent environment is generally treated as a dielectric continuum interacting with the solute via a reaction field. Two of the most widely used methods to estimate the reaction field are the generalized Born (GB) approximation [6] and the Poisson–Boltzmann equation (PB) [8]. Since the full solution of the continuum electrostatic problem is computationally demanding, a number of simple empirical distance-dependent screening functions have been developed [11–18] for the use in molecular dynamics simulations, where re-evaluation of solvent screening needs to be performed at every step. Although these 'implicit solvent' models have been used in multiple studies, their ability to reliably reproduce the energetics and dynamics of solvated biomolecules is still uncertain [19–21]. For example, through a series of extensive replica exchange simulations, Zhou et al. [22–24] have shown that the folding free energy landscape of the 16-residue beta hairpin from C terminus of protein G computed using the GB or PB continuum solvation models is qualitatively different from the one obtained from an explicit solvent simulation.

A principal drawback of continuum models is their inability to account for discrete nature of the solvent and the aspects of solvation inherent to this granularity, such as cavity dewetting, and specific solvent–solute interactions, including hydrogen bonding. Recognizing that these effects are most prominent in the first

* Corresponding author. Tel.: +1 617 552 8927; fax: +1 617 552 2705.
E-mail address: krilov@bc.edu (G. Krilov).

0166-1280/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2008.05.035

few solvation shells [25], significant effort has been invested over the years in methods that attempt to maintain a microscopically accurate description of solvation in this inner region, while realizing the computational savings by more approximate treatment of the outlying solvent [26]. Early work derived from statistical mechanical liquid theory includes the MTGLE method by Berkowitz et al. [27–29], the stochastic boundary model by Berkowitz and McCammon [30] and its extension by Brooks and Karplus in form of the mean field force approximation (MFFA) [31,32], and the RISM-HNC integral equation approach of Pettitt and Karplus [33]. First attempts to include the electrostatic polarization effects which are dominant in aqueous solutions of charged solutes are the surface constrained all-atom solvent model (SCAAS) by King and Warshel [34,35] and reaction field with exclusion (RFE) of Rullman and van Duijnen [36]. More recently, Beglov and Roux [37] introduced the spherical boundary potential, which allows for density fluctuations in the inner region, and further extended it to include a Poisson–Boltzmann treatment of polarization [38]. A number of studies were performed using this model, including electron transfer in proteins [39,40], solvation and thermal stability of biomolecules [41–44], and the folding kinetics of proteins [45–47].

A significant limitation of the above approaches is that the shape of the boundary is fixed and typically limited to the geometries which permit analytic solutions to the electrostatic polarization problem. A new generation of solvation models which attempt to alleviate this restriction has been introduced in the last decade. Primary hydration shell model of Beglov and Roux [48] relies on a sum-over-spheres approach to represent a non-spherical boundary, and uses a simple harmonic potential to keep the solvent molecules close to the solute surface, but lacks the reaction field treatment of electrostatics. Lounnas et al. introduced the shell approximation for protein hydration (SAPHYR) model [49], which uses a combination of Van der Waals interaction and average dipole–dipole electrostatic forces to keep the solvent molecules within a fixed distance of the solute surface. The model was successful in reproducing structural properties of several solvated proteins. In a different approach, Kimura et al. developed the surface of active polarons (SOAP) method [50], which strives to capture the electrostatic polarization through dynamic adjustment of oxygen partial charges on the solvent water molecules. The authors used this approach to estimate the solvation free energies of ions and amino-acid analogues.

Recently, Li et al. introduced the fluctuating elastic boundary (FEB) model [51], which is both suitable for molecular dynamics simulations, and removes some of the above restrictions. This method uses an elastic boundary comprising a mesh of quasi-particles interconnected by nearest neighbor bonds to confine the solute and its inner solvation shell. The boundary evolves dynamically in response to the shape fluctuations of the solute–solvent system inside. In addition to reducing the computational cost, since the boundary potential is constructed out of standard pairwise interactions, the model is easily incorporated into most major molecular dynamics simulation packages with little modification. The method was successful in reproducing static and dynamics properties of liquid water and conformational thermodynamics of solvated alanine dipeptide [52].

One of the disadvantages of the FEB model is the occasional solvent escape due to the inherently porous nature of the quasi-particle mesh boundary, sometimes accompanied by opening of large scale holes resulting in the loss of integrity of the confined explicit solvation shell and failure of the simulation. In addition, since the dynamics of the boundary are not directly coupled to the conformational changes of the solute, fluctuations in the confined solvent distribution due to local pressure variations were found to sometimes lead to non-uniform hydration and drying of parts of the solute surface.

In this paper, we present two new models which, by dispensing with the explicit representation of the quasi-particle mesh, maintain the advantages of the FEB approach, but eliminate many of these difficulties and hence improve both the accuracy and computational efficiency of simulations. In both cases, the position and shape of the 'implicit' solvent boundary are variable, and evolve dynamically with configurational fluctuations of the confined solute–solvent system, ensuring uniform hydration of the solute surface. Moreover, the nature of the boundary potential ensures that no solvent escape is possible.

The paper is organized as follows. In Section 2, we describe the dynamic boundary model and the exchange boundary models. In Section 3, we use these methods to compute the structural and dynamic properties of solvated polyalanine helices. We conclude in Section 4.

2. Flexible implicit boundary solvation models

For any non-periodic boundary condition solvation model, a properly designed confining potential which adequately captures the average influence of the external bulk solvent is crucial for reproducing the dominant solvation effects on the biomolecular solute [53]. Since the conformational changes of the solute influence the solvation profile, the potential must also be able to adapt its shape dynamically. In principle, the non-periodic boundary condition solvation model should include as few explicit solvent molecules as possible to reduce the computational cost, while still adequately describing the dominant effects of solvation. However, if only a single layer of solvent molecules is included, it would be very challenging to reproduce solvation correctly, since it is difficult to prevent some portions of the solvent-accessible solute surface from desolvating occasionally due to solvent motion during the simulation. With two or three layers of solvent molecules included, it should be feasible, although still not trivial, to design a solvation model that is both practical and accurate. Here we present two distinct approaches to non-periodic modeling of solvation which, requiring only a small number of solvent molecules to be treated explicitly, share the qualities of computational efficiency, simple implementation and relatively high accuracy.

2.1. Dynamic boundary solvation model

The aim of this model is to construct a boundary potential which will confine the solvent molecules of the explicitly represented solvation shell, and simultaneously, account for the variation in pressure on the solvent arising due to conformation changes of the solute. The origin of this confining interaction is primarily in the steric repulsion and packing forces exerted by the bulk solvent external to the boundary on the system inside. Such forces in liquids are generally represented via a Van der Waals potential of the form:

$$V(r) = \frac{Ae^{-Br}}{r} - \frac{C}{r^6}. \tag{1}$$

For computational efficiency, the latter is customarily approximated by a polynomial expression

$$V(r) \approx \frac{D_m}{r^m} - \frac{D_n}{r^n}, \tag{2}$$

with the $n = 6$, $m = 12$ Lennard-Jones function being the most widely used form. In light of this fact, we chose Eq. (2) as the functional form representing the interaction of the confined solvent with the dynamic boundary, such that: