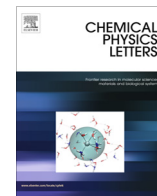




Contents lists available at ScienceDirect

Chemical Physics Letters

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

Research paper

Reactive molecular dynamics models from *ab initio* molecular dynamics data using relative entropy minimization

Christopher Arntsen^{a,1}, Chen Chen^{a,b,1}, Gregory A. Voth^{a,*}^a Department of Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA^b College of Chemistry and Molecular Sciences, Hubei Key Lab of Electrochemical Power Sources, Wuhan University, Wuhan 430072, China

ARTICLE INFO

Article history:

Received 24 January 2017

In final form 18 April 2017

Available online xxxxx

Keywords:

Reactive molecular dynamics

Hydrated excess proton

Relative entropy minimization

Acids

Multiscale modeling

ABSTRACT

We present two new multiscale molecular dynamics (MS-RMD) models for the hydrated excess proton in water developed directly from *ab initio* molecular dynamics (AIMD) simulation data of the same system. The potential of mean force along the proton transfer reaction coordinate and radial distribution functions for the MS-RMD models are shown to faithfully reproduce those of AIMD. The models are developed using an algorithm based on relative entropy minimization, thus demonstrating the ability of the method to rapidly generate accurate and highly efficient reactive MD force fields.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Proton transport is a fundamental process important to a variety of systems within the field of biology and materials science, relevant to systems ranging from membrane proteins [1,2] to fuel cells [3,4]. Proton transport is a unique process arising from both vehicular transport and proton Grotthuss hopping, which results from a rearrangement of covalent and hydrogen bonds [5–9]. Due to the importance of this process, it is necessary to have an accurate and physically realistic, yet efficient computational framework for simulating the process. While *ab initio* molecular dynamics (AIMD) would seem ideal, as it explicitly handles bond breaking and formation, the cost of such simulations prohibits their application to large biological or fuel cell systems. Similarly, classical molecular dynamics (MD) simulations can handle extremely large systems, but have fixed bonding topologies, precluding proton hopping and thereby ignoring the essential physics of the process.

Multiscale reactive molecular dynamics (MS-RMD) [10–12] and the earlier multistate empirical valence bond (MS-EVB) approach [8,13,14] have emerged as efficient computational methods for studying proton transport in complex systems. They have both been shown to capture the solvation structure of a hydrated excess proton in water and the physics involved in proton transfer. Addi-

tionally, MS-RMD and MS-EVB offer the computational efficiency necessary for running simulations too large for AIMD, e.g., MS-RMD has been applied to study the proton transport mechanism in membrane proteins [15,16], as well as to study and elucidate the proton transport mechanism in proton exchange membranes (PEMs) [4,17–19]. However, development of accurate MS-RMD models can be nontrivial [12].

Previous parameterization of MS-EVB models (as the precursor to MS-RMD) have been fit in part to the energetics as well as potential energy surfaces along the proton transfer reaction coordinate from *ab initio* calculations of small clusters of $[(\text{H}_2\text{O})_n\text{H}]^+$, where n is either 2 or 4 [8,13,14]. While demonstrably successful in several ways, the approach relies in part on gas-phase protonated water clusters to parameterize a model meant for bulk phase simulations. To overcome this, force-matching has been utilized to parameterize an MS-RMD model to Car-Parrinello MD (CPMD) [10,11] data directly. (We shall consider this model to be MS-RMD version 1 for the hydrated excess proton in water.) In this latter approach, all classical force field parameters, i.e. pair potentials, bonds, angles, are first force-matched to minimize the force-matching variational residual from nonreactive configurations of the CPMD data; while the reactive configurations are then force-matched to include proton hopping. The method successfully generated an MS-RMD force field which very closely matches both the radial distribution functions (RDFs) and potential of mean force (PMF) for proton transfer in the CPMD data. Despite the success of this method, all pair potentials utilize a numerical b-spline functional form, which precludes its simple application to MD simula-

* Corresponding author.

E-mail address: gavoth@uchicago.edu (G.A. Voth).¹ These authors contributed equally to this work.

tions of condensed phase systems as, e.g., simple mixing rules for the pair potentials cannot be used. (We note, however, that there is nothing that says force-matching cannot utilize predefined analytical forms, as has been done in Ref. [10]. Use of splines, however, improved the overlap of the model RDFs and PMF to those of the target.) The fitting MS-RMD models to CPMD data via force matching can be a challenging task, and the end result may not necessarily reproduce the desired pair distribution functions, although it will satisfy the Yvon-Born-Green equation connecting two- and three-body correlations [20]. As an alternative to force-matching, relative entropy minimization (REM) has been shown to effectively generate model coarse-grained force fields that can also reproduce the structure of a given target [21–24].

In this Letter, we develop a REM approach to parameterize MS-RMD force fields for the hydrated excess proton in water, fitting to state-of-the-art *ab initio* MD (AIMD) data. While we acknowledge that AIMD water does have its well-documented shortcomings [25,26], the fitting method described in this Letter is able to recapitulate the solvation structure and reactivity of hydrated excess protons without introducing the deficiencies seen in AIMD water. (This is discussed in more detail in the results section.) In that sense, the work is largely motivated by the fact that REM is agnostic to the target, and as more sophisticated quantum calculations become accessible for bulk phase calculations, REM would indeed be a powerful tool in reactive force field development. As such, the work herein is a demonstration of REM's ability to handle even highly dimensional, non-linear parameter space. We find that this approach can reproduce both the AIMD RDFs and proton transfer PMF. The MS-RMD models developed in this Letter can also utilize standard Lennard-Jones (LJ) pair potentials (so that, e.g., standard mixing rules can be used with novel materials) and specific functional forms for the repulsive and off-diagonal MS-RMD potentials.

The remaining content of this Letter is organized as follows: in the Methodology section, we discuss the details of MS-RMD and the REM fitting scheme. In the Results section, we then discuss the derived models, termed MS-RMD versions 2 and 3, and compare the two models to the most recent reactive MS-EVB model, MS-EVB 3.2 [13]. Finally, we conclude the Letter and discuss the outlook for the REM fitting scheme as applied to MS-RMD models, and outline desirable improvements in future models.

2. Methodology

Contrary to empirical classical MD, where the bonding topology is fixed, in MS-RMD (and MS-EVB before it), for a given nuclear configuration the ground state of a system $|\Psi\rangle$ is defined as a linear combination of unique bonding topologies (basis states): [8].

$$|\Psi\rangle = \sum_{i=1}^N c_i |i\rangle \quad (1)$$

where c_i is the coefficient of basis state i . The coefficients are determined by diagonalizing the Hamiltonian for each nuclear configuration such that:

$$\mathbf{H}\mathbf{c} = E_0\mathbf{c} \quad (2)$$

where E_0 is the ground state energy for that configuration. The diagonal matrix elements in the Hamiltonian are the diabatic energies of the basis states, defined by the classical force field energy of the basis state plus a repulsive term added to correct for an over-attraction between hydronium and water in the underlying classical force field. The off-diagonal elements, which introduce reactivity by allowing for "mixing" between basis states of different bonding topology, are a function of the distance between of hydronium and water. The functional forms of both the repulsive potentials and the off-diagonal coupling potential are shown in the [Supple-](#)

[mental Information](#), as well as all parameters used in the models described in this work. We note that variables fit using REM are shown in boldface, whereas those pertaining to the underlying classical force field or derived in previous models are shown regular typeface.

The MS-RMD models described below are parameterized using the REM scheme. The specific aim of REM is to find a model force field that generates a thermodynamic ensemble with maximal overlap to a target ensemble generated with a higher level of theory. In the present work, the target ensemble is AIMD and the model ensemble is the MS-RMD force field. The relative entropy of two ensembles is defined as:

$$S_{\text{rel}} = \sum_i p_T(i) \ln \frac{p_T(i)}{p_M(i)} \quad (3)$$

where $p(i)$ is the probability of configuration i in the target T and model M ensembles. In the constant NVT ensemble, the relative entropy is defined as

$$S_{\text{rel}} = \beta \langle U_M - U_T \rangle_T - \beta (A_M - A_T) + \langle S_{\text{map}} \rangle_T \quad (4)$$

where U is the potential energy, A is the configurational part of the Helmholtz free energy, and S_{map} is the entropy that results from degeneracies in the model.

The relative entropy is minimized on the condition that the average curvatures of the model potential energy with respect to the force field parameter λ_i over the model and target ensembles are equal: [21].

$$\left\langle \frac{\partial U_M}{\partial \lambda_i} \right\rangle_M = \left\langle \frac{\partial U_M}{\partial \lambda_i} \right\rangle_T \quad (5)$$

We therefore set out to minimize the difference between these two derivatives. The simplest numerical implementation for doing so uses the Newton-Raphson iteration scheme: [21].

$$\lambda_{i+1} = \lambda_i - \frac{\left(\frac{\partial S_{\text{rel}}}{\partial \lambda} \right)}{\left(\frac{\partial^2 S_{\text{rel}}}{\partial \lambda^2} \right)} = \lambda_i - \chi \left[\left\langle \frac{\partial U}{\partial \lambda} \right\rangle_T - \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_M \right] \left[\left\langle \frac{\partial^2 U}{\partial \lambda^2} \right\rangle_T - \left\langle \frac{\partial^2 U}{\partial \lambda^2} \right\rangle_M + \beta \left\langle \left(\frac{\partial U}{\partial \lambda} \right)^2 \right\rangle_M - \beta \left\langle \left(\frac{\partial U}{\partial \lambda} \right)^2 \right\rangle_M \right]^{-1} \quad (6)$$

Here, we introduce a mixing parameter χ . Notice that upon satisfaction of the condition in Eq. (6), the derivative of the relative entropy with respect to λ is zero, and the first bracketed term in Eq. (6) is zero, and λ is unchanged. It should also be noted that in regions where the second derivative of the relative entropy is negative, the optimization is far from the minimum, and the Newton-Raphson iteration scheme is inappropriate. The minimization procedure is performed iteratively, where the derivatives calculated in Eq. (6) are calculated over ensembles generated from updated parameters. As the minimization scheme utilized here requires the model to be somewhat close to the target ensemble, we began each minimization with the MS-EVB 3.2 parameters [13]. We should note that our implementation of the Newton-Raphson iteration scheme does not necessarily reach a global minimum, and may get trapped in a local minimum. Interestingly, we have found that all optimized parameter sets generate serviceable models, reproducing the target RDFs faithfully. This may be a result of the highly dimensional parameter space in which our minimization is performed (12 parameters are simultaneously optimized), which allows for greater flexibility – that is, certain parameters can compensate for deleterious effects of others.

The AIMD data to which the MS-RMD models in this paper were fit utilize the BLYP exchange-correlation functional with the D3 Grimme dispersion and the TZV2P basis set. The systems used in

Download English Version:

<https://daneshyari.com/en/article/5377754>

Download Persian Version:

<https://daneshyari.com/article/5377754>

[Daneshyari.com](https://daneshyari.com)