Carbohydrate Research 457 (2018) 41–50

Contents lists available at ScienceDirect

Carbohydrate Research

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

Proper balance of solvent-solute and solute-solute interactions in the treatment of the diffusion of glucose using the Drude polarizable force field

Mingjun Yang [a, b], Asaminew H. Aytenfisu [a], Alexander D. MacKerell Jr. [a, *]

[a] Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, 20 Penn Street, Baltimore, MD 21201, USA
[b] XtalPi Inc., Shennan Road 6025, Futian District, Shenzhen, China

ARTICLE INFO

Article history:
Received 9 November 2017
Received in revised form
2 January 2018
Accepted 15 January 2018
Available online 31 January 2018

Keywords:
CHARMM
MD simulations
Carbohydrates
Monosaccharides
Molecular mechanics

ABSTRACT

Motivated by underestimation of the diffusion constant of glucose in aqueous solution at high glucose concentrations we performed additional optimization of the Drude polarizable hexopyranose monosaccharide force field. This indicated aggregation of the glucose at higher concentrations, which is a concern for studies of complex glycan systems such as the HIV Envelope where high effective concentrations of sugars are present. High-level quantum mechanical calculations were undertaken on water monohydrate-glucose interactions, on water cluster-glucose interactions and on glucose-glucose dimers in stacked (parallel) and perpendicular orientations. Optimization of the nonbond and dihedral parameters targeting these data yielded a revised model that showed improved agreement with experimental aqueous diffusion data. However, limitations in the diffusion constants were still present. These were due to the SWM4-NDP inherently overestimating the diffusion constant of water, a problem that was validated by calculation of the aqueous diffusion constants using the SWM6-NDP water model. In addition, results show the water diffusion constant to be significantly overestimated at high glucose concentrations though the glucose diffusion is in satisfactory agreement with experiment. These results indicate the subtle balance of water-sugar, water-water and sugar-sugar interactions that needs to be properly modeled to account for the full range of aqueous behavior of sugars in aqueous solution.

1. Introduction

Empirical force field based theoretical studies of biological macromolecules, including carbohydrates [1–4], is a growing field associated with increases in computer hardware as well as improvements on algorithms and in the force fields central to the accuracy of such methods. Towards this end work in our laboratory in collaboration with Roux and coworkers has focused on the development of a comprehensive biomolecular force field that explicitly treats electronic polarizability based on the classical Drude oscillator model [5,6]. The Drude force field is now available for proteins [7], DNA [8–10], lipids [11,12], monoatomic ions [5], and a range of small model compounds [13–17] as well as for acyclic carbohydrates, hexopyranose and furanose monosaccharides [18–21]. These models have now been utilized in a number application studies from which novel insights into the role of polarization in conformational transitions have been obtained [22–25]. In addition, these studies are also identifying limitations in the Drude model, motivating additional improvements in the model.

During development of the hexopyranose monosaccharides comparison of experimental and simulated self-diffusion constants for D-glucose showed good agreement with experiment at a glucose concentration of 1 M [20]. However, there was a trend for the self-diffusion constant to be underestimated at higher concentrations (2 and 5 M). Such a trend is indicative of aggregation of the monosaccharides, a problem that could be due to overestimation of monosaccharide-monosaccharide interactions, an underestimation of water-monosaccharide interactions or a combination of both. While many simulation studies of saccharides are on dilute solutions or glycans in the context of glycolipids and glycoproteins that are relatively isolated, there are many systems in which complex glycans are undergoing extensive glycan-glycan interactions such that they are essentially present at a high

* Corresponding author.
E-mail address: alex@outerbanks.umaryland.edu (A.D. MacKerell).

https://doi.org/10.1016/j.carres.2018.01.004
0008-6215/

concentration. Such systems are now accessible to simulations, such as the glycosylated HIV Envelope [26–28], where the proper balance of the aforementioned interactions is quite important. In addition, improvements in our optimization protocols and associated parameters from the development of the Drude furanose monosaccharide parameters required adjustment of the hexopyranose monosaccharides to maintain consistency [19]. Accordingly, we undertook additional optimization of the Drude hexopyranose model focused on maintaining the overall quality of the model while improving its properties in aqueous solution, including improving the balance of the monosaccharide-water (eg. solute-solvent) and monosaccharide-monosaccharide (eg. solute-solute) interactions.

2. Methods

Empirical force fields calculations were performed with the program CHARMM [29]. The initial parameters were the published Drude hexopyranose [20] and furanose monosaccharide parameters [19] along with the SWM4-NDP [30] and SWM6-NDP [31] water models. Quantum mechanical (QM) calculations used the Gaussian03 program [32] for energy minimizations to default tolerances with single point calculations performed using Psi4 and Q-Chem [33,34]. The specific QM model chemistries are described below.

Empirical energy minimization were performed using the steepest descent minimizer followed by the adapted-basis Newton-Raphson (ABNR) minimizer to a gradient of 10^{-5} kcal/mol/Å. Gas phase calculation did not use atom truncation. All Drude and QM interaction energies were calculated based on the total energy of the dimer (or multimers) minus the individual monomer energies as required to account for the electronic degrees of freedom in the model. This involved use of the gas phase optimized monomer geometries for both the dimers and monomers with the Drude energies calculated by restraining the real atoms with a harmonic force constant of 10^7 kcal/mol/Å and allowing the Drude oscillators to relax using the ABNR optimizer.

For molecular dynamics (MD) simulations the nonbonded Lennard-Jones interactions were computed within a cutoff of 12 Å, with a smoothing switch function over the range from 10 to 12 Å and a isotropic long-range correction term [35]. The electrostatic interactions were treated by the particle mesh Ewald method with a real space cutoff of 12 Å, a charge grid of 1 Å, and the 6-th order spline function for mesh interpolation [36]. The covalent bonds involving hydrogen atoms were constrained with the SHAKE algorithm and a time step of 1 fs was used to integrate the equations of motion [37]. The hard-wall constraint was applied to prevent the Drude particles moving more than 0.2 Å from their parent nuclei [11]. The Nose-Hoover thermostat was used to maintain the temperature of the real atoms at 298 K and the Drude particles at 1 K, respectively [38].

3. Results and discussion

The partial atomic charges for the Drude hexopyranose monosaccharides were transferred from the furanose force field to maintain the same charges on chemical groups with similar bonded environments. The atomic polarizabilities and Thole screening factors were then refitted against the QM dipole moments of 215 hexopyranose conformations using the Monte Carlo simulated annealing algorithm (MCSA) [39]. These include target QM dipole moments computed at the RIMP2/cc-pVQZ//MP2/6-31G(d) model chemistry for different conformations of the 16 hexopyranose isomers (Table 1). During MCSA optimization, the range that the parameters could vary was set to [-1.3 to −0.5] and [-2.1 to −1.0] for the atomic polarizabilities of oxygen atoms and the remaining non-hydrogen atoms, respectively, and [0.2 to 2.5] for all Thole scale factors. The MCSA was initialized at a temperature of 500 K and gradually reduced by a factor of 0.75 after every 200 steps for each parameter until it approached 0 K [19]. The final parameters were derived from averaging the last 20 MCSA steps of the optimization. Upon MCSA optimization, the root-mean-square difference (RMSD) between the MM and QM total dipole moments was reduced from 0.21 to 0.16 while the original Drude force field had a RMSD value of 0.19. Fig. 1 presents the comparison of the MM and QM results for both total and component dipole moments.

Table 1
Number of conformations used for the dipole moment calculation for each hexopyranose isomer. QM calculations were performed at the RIMP2/cc-pVQZ//MP2/6-31G(d) model chemistry unless noted.

Monosaccharide	no. of conformations
α-D-allose	11
β-D-allose	15
α-D-altrose	12
β-D-altrose	13
α-D-glucose	13
β-D-glucose	16
α-D-galactose	13
β-D-galactose	13
α-D-talose	16
β-D-talose	12
α-D-gulose	14
β-D-gulose	16
α-D-manose	14
β-D-manose	14
α-D-idose	14
β-D-idose	9
Total	215

To further validate the new electrostatic parameters, we examined the molecular polarizability of the glucose monosaccharide (Table 2). The result from the new Drude force field accounts for 70% of the QM polarizability, which is comparable to the original Drude alcohol force field [17] and the general trend with the Drude force field in which the gas phase molecular polarizabilities are scaled relative to the QM values [40].

Additional optimization focused on selected dihedral parameters targeting the conformational energies of the monosaccharides. For dihedral fitting, the QM target energies were computed at the RIMP2/cc-pVQZ//MP2/6-31G(d) model chemistry for different monosaccharide conformations (Table 3). For each monosaccharide, the relative energies of both anomers were all offset to the global minimum energy conformation. A relative energy cutoff of 20 kcal/mol was used to include more non-chair conformations that have a higher relative energy in the target data versus the previous cutoff value of 12 kcal/mol [20]. Fig. 2 shows the φ/θ distribution according to Cremer-Pople definition of the ring conformers [41]. The majority of the ring conformations are in the $^1C_4$ states though alternate puckers are included in the target data.

A total of 2004 conformational energies were used as target data in the fitting of all possible dihedrals including 4 non-hydrogen atoms or 3 non-hydrogen atoms and 1 hydroxyl hydrogen (Table 3). The multiplicity for intra-ring dihedrals is set to 1, 2, 3 and 4, and 1, 2 and 3 for the other dihedrals. The same energy offset was used for all 2004 conformations in the fitting, i.e. the previously reported non-group fitting was adopted in which all conformations were fit simultaneously. The *lsfitpar* program was used for dihedral fitting [42]. The RMSD for the updated parameters (Drude new), the original Drude force field and additive force field are shown in Table 3 [20,43]. Fig. 3 plots the relative conformational energies of the target data from QM, the new Drude force field and the original Drude force field. In addition to the 2004 conformations used in