Published in final edited form as:

Chem Phys Lett. 2010 February 5; 486(1-3): 70–73. doi:10.1016/j.cplett.2009.12.089.

Solvation of Biomolecules by the Soft Sticky Dipole-Quadrupole-Octupole Water Model

Jerez A. Te, Ming-Liang Tan, and Toshiko Ichiye*

Department of Chemistry, Georgetown University, Washington, DC, 20057

Abstract

The soft sticky dipole-quadrupole-octupole (SSDQO) potential energy function represents a water molecule by a single site with a van der Waals sphere and point multipoles. Previously, SSDQO was shown to give good properties for liquid water and solvation of simple ions and is faster than three point models. Here, SSDQO is assessed for solvating biologically relevant molecules having a multi-site, partial charge description. Monte Carlo simulations of ethanol, benzene, and N-methylacetamide in SSDQO with SPC/E moments showed the water structure was as good as in SPC/E. Thus, SSDQO is potentially useful for simulations of biological macromolecules in aqueous solution.

INTRODUCTION

The structure and activity of biological macromolecules such as proteins are highly dependent on being solvated by water. Computer simulations of these systems must include solvent effects and the most accurate way of treating these effects is by using explicit water models. Moreover, simulations with explicit water are also a means of studying the underlying molecular basis of aqueous solvation. However, evaluating the water-water interaction is the most time consuming process in such simulations because of the large number of water molecules needed to solvate the solute. Since most water models use partial charges on fixed interaction sites to describe the electrostatics, a greater number of interaction sites exacerbates the problem because of the increased number of internuclear distances that must be calculated. On the other hand, a greater number of interaction sites generally makes it easier to fit more properties of liquid water. For instance, three-site models such as the SPC/E [1] and TIP3P [2] models give a reasonable description of water but have problems with dielectric and dynamical properties, respectively [3]. On the other hand, the five-site TIP5P [4] model has excellent properties for pure water but is computationally expensive.

Recently, we have developed the soft sticky dipole-quadrupole-octupole (SSDQO) model of water [5], which unlike typical models, has a single-site with a van der Waals sphere and point dipole, quadrupole, and octupole moments. Because the electrostatic interaction potential is the exact moment expansion up to order $1/r^4$ and contains an approximation for the $1/r^5$ term, it is both efficient and accurate. When the moments, geometry, and van der Waals parameters of SPC/E, TIP3P, and TIP5P are used, SSDQO reproduces the water

*corresponding author (ti9@georgetown.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

dimer potential energy and radial distribution function of the corresponding multipoint model [5]. Moreover, SSDQO using SPC/E moments, geometry, and van der Waals parameters has good thermodynamic, dielectric, and dynamic properties [6] as well as good structural properties around simple ions [7]. Although it also reproduces many of the flaws of SPC/E water, it is actually an improvement in being somewhat more structured than SPC/E. Moreover, SSDQO is computationally faster than even the three-site models because the nine distances required for three-site models are slower than the matrix multiplications and the single intermolecular distance needed in SSDQO for the moment expansion, since the approximation for the $1/r^5$ require only the lower order matrix multiplications and higher order terms are neglected. Previous studies have shown that SSDQO is about three times faster than three point models such as SPC/E and TIP3P in Monte Carlo simulations [5] and about two times faster in molecular dynamics simulations, where the dipole-dipole interaction was treated with Ewald summation and the higher order interactions by truncation [6]. We are also improving the properties of SSDQO by refining the moments against quantum chemical calculations to reproduce the experimental properties of liquid water.

However, despite the efficiency and accuracy of the SSDQO water model, it can be useful in improving the speed in typical simulations of biological macromolecules only if it describes the interaction with the solute accurately. This is not a trivial question since most force fields used for such simulations as CHARMM [8] and AMBER [9] describe the electrostatics of proteins, nucleic acids, carbohydrates, lipids, etc. by partial charges while the electrostatics of the water molecules in the SSDQO model are described by higher order multipoles. Thus it is crucial to assess the ability of SSDQO to interact with a multi-site, partial charge description of a solute.

Here, SSDQO is assessed for solvating biologically relevant molecules described by typical partial charge descriptions. The solutes studied are ethanol, benzene, and N-methylacetamide (NMA), which represent different molecular fragments found in proteins. Ethanol is important since it is amphiphilic, with a hydroxyl group and a hydrocarbon tail, and it mimics the side chain of serine. Benzene tests the ability to solvate a hydrophobic molecule and mimics the side chain of phenylalanine. Finally, NMA is the simplest compound with a peptide bond and thus serves as a test for solvating the polypeptide backbone. Since SSDQO is based on a multipole expansion, which is most accurate at long range, the focus here is on the accuracy of the short-ranged structure where the model is most uncertain. Thus, the radial distribution functions are examined, which are sensitive to short-ranged structure even when long-range correlations are poorly described [10]. Monte Carlo simulations of these solutes in SSDQO using moments and van der Waals parameters from SPC/E and in SPC/E water demonstrate that the solvation by SSDQO is as good as by SPC/E. Moreover, the slight differences indicate that SSDQO actually acts a better hydrogen bond donor than SPC/E. The efficiency and accuracy of the SSDQO model of water indicate that it is potentially very useful for computer simulation of macromolecules in aqueous solution.

METHODOLOGY

Detailed descriptions of the SSDQO water-water and water-ion potentials can be found elsewhere [5,7] so only a brief description is given here. The non-bonded interaction potential is given by

$$\begin{aligned}
U_{ij}(\mathbf{r}) = & 4\epsilon_{ij} \left\{ \left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right\} \\
& + \frac{1}{r} [q_i q_j] \\
& + \frac{1}{r^2} [-q_i (\boldsymbol{\mu}_j \cdot \mathbf{n}) + (\mathbf{n} \cdot \boldsymbol{\mu}_i) q_j] \\
& + \frac{1}{r^3} [q_i (\boldsymbol{\Theta}_j \cdot \mathbf{n}^{(2)}) + (\mathbf{n}^{(2)} \cdot \boldsymbol{\Theta}_i) q_j - 3(\mathbf{n} \cdot \boldsymbol{\mu}_i)(\boldsymbol{\mu}_j \cdot \mathbf{n}) + \boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j] \\
& + \frac{1}{r^4} [-q_i (\boldsymbol{\Omega}_j \cdot \mathbf{n}^{(3)}) + (\mathbf{n}^{(3)} \cdot \boldsymbol{\Omega}_i) q_j + 5(\mathbf{n} \cdot \boldsymbol{\mu}_i)(\boldsymbol{\Theta}_j \cdot \mathbf{n}^{(2)}) - 5(\mathbf{n}^{(2)} \cdot \boldsymbol{\Theta}_i)(\boldsymbol{\mu}_j \cdot \mathbf{n}) - 2\boldsymbol{\mu}_i \cdot \boldsymbol{\Theta}_j \cdot \mathbf{n} + 2\mathbf{n} \cdot \boldsymbol{\Theta}_i \cdot \boldsymbol{\mu}_j] \\
& + \frac{1}{r^5} [q_i \Phi_j (\boldsymbol{\mu}_j \cdot \mathbf{n})(\mathbf{o}_j \cdot \mathbf{n}^{(3)}) + (\mathbf{n}^{(3)} \cdot \mathbf{o}_i)(\mathbf{n} \cdot \boldsymbol{\mu}_i) \Phi_i q_j - c_{\text{DO}}(\mathbf{n} \cdot \boldsymbol{\mu}_i)(\boldsymbol{\Omega}_j \cdot \mathbf{n}^{(3)}) - c_{\text{DO}}(\mathbf{n}^{(3)} \cdot \boldsymbol{\Omega}_i)(\boldsymbol{\mu}_j \cdot \mathbf{n}) + c_{\text{QQ}}(\mathbf{n}^{(2)} \cdot \boldsymbol{\Theta}_i)(\boldsymbol{\Theta}_j \cdot \mathbf{n}^{(2)})]
\end{aligned}$$

where $\mathbf{r} = r\mathbf{n}$ is the internuclear vector from particle i to j . Also, the dyadic products are denoted by $[\mathbf{n}^{(2)}]_{ij} = n_i n_j$ and $[\mathbf{n}^{(3)}]_{ijk} = n_i n_j n_k$, and the matrix contractions are denoted by $\mathbf{A} \cdot \mathbf{B} = \sum_i A_i B_i$, $\mathbf{A} : \mathbf{B} = \sum_{ij} A_{ij} B_{ij}$, and $\mathbf{A} \cdot \mathbf{B} = \sum_{ijk} A_{ijk} B_{ijk}$. The electrostatic potential is an exact multipole expansion up to order $1/r^4$ and contains an approximate $1/r^5$ term. For water-water interactions, the water molecules i and j interact through the dipole $\boldsymbol{\mu}$, quadrupole $\boldsymbol{\Theta}$, and octupole $\boldsymbol{\Omega}$ moments of water, with the monopole $q=0$. For the solute-water interactions, the partial charges q_i of the solute molecule interact with the multipole moments of SSDQO water molecule j up to the hexadecapole Φ . In the approximate charge-hexadecapole interaction, $\mathbf{m}$ is a unit vector along the direction of $\boldsymbol{\mu}$, and $\mathbf{o}$ is a unit vector along the direction of $\boldsymbol{\Omega}$. The factor Φ for a tetrahedral molecule is defined as $7\Gamma b_{\text{OH}}/10\sqrt{3} = -H_{\text{ZZZZ}}/2$, where $\mathbf{H}$ is the hexadecapole moment matrix. For SPC/E, $\Gamma = 5/2\Omega$ so $\Phi = 7\Omega b_{\text{OH}}/4\sqrt{3}$. This potential allows straightforward combining rules for interaction with other molecules.

The Monte Carlo simulations used standard Metropolis sampling [11] in the NVT ensemble at 298 K for a cubic box (box length, $b = 24.835 \text{ \AA}$). In each case, one solute was solvated in box of water created at the experimental density of water ($0.03346 \text{ molecules/\AA}^3$). The simulations consisted of one ethanol in 504 water molecules, one benzene in 502 water molecules, and one NMA in 498 water molecules. Periodic boundary conditions and spherical switching functions between $(b/2 - 1) \text{ \AA}$ and $b/2 \text{ \AA}$ were applied.

The starting configurations of the ethanol and the NMA were taken from the Cambridge Structural Database while benzene was created from CHARMM parameters (C-C and C-H bond lengths are $1.375 \text{ \AA}$ and $1.080 \text{ \AA}$ respectively; C-C-C and H-C-C angles are 120°). The CHARMM22 potential energy function [12] was used for all solutes. The SSDQO potential energy function with SPC/E Lennard-Jones parameters and moments ($\mu = 2.3503 \text{ D}$, $\Theta = 2.0355 \times 10^{-26} \text{ esu-cm}^2$, $\Delta = 0$, $\Omega = 0.7834 \times 10^{-34} \text{ esu-cm}^3$, and $\Gamma = 1.9585 \times 10^{-34} \text{ esu-cm}^3$), which is referred to as SSDQO:SPC/E, was used for the water molecules [5]. In these calculations, the solute coordinates were fixed.

The initial configurations were equilibrated for 400,000 MC “passes” (one pass equals N attempted translational and rotational moves) and the radial distribution functions were calculated from the subsequent 400,000 MC passes. The acceptance ratio in all MC runs was approximately 40%. For comparison, the solutes were also simulated in SPC/E water using the same conditions.

RESULTS AND DISCUSSION

The radial distribution functions around each of the solutes were calculated for both SSDQO and SPC/E. For reference, the pure water radial distribution function is reproduced in Fig. 1. In the discussion below, g_{XO} and g_{XH} refer to the radial distribution function around the solute atom X of the water O and H, respectively.

The radial distribution functions of SSDQO and SPC/E water around ethanol demonstrate that the solvation of an hydrogen bonding amphiphilic solute by both models is very similar but with some minor structural differences (Fig. 1). The first solvation shell around the ethanol oxygen is at approximately the same distance for both models, with the first peak at ~ 2.8 Å in g_{OO} for both (Fig. 1a). However, the first peak of SSDQO was higher and wider than in SPC/E and integration of the first peak to 3.475 Å (the minimum for SSDQO) gave coordination numbers of 3.35 and 2.90 for SSDQO and SPC/E respectively, while integration of the first peak to 3.3 Å (the minimum for SPC/E) gave coordination numbers of 2.85 and 2.31 for SSDQO and SPC/E respectively (Table 1). This reflects the somewhat stronger hydrogen bonding capability of SSDQO with SPC/E moments. The more strongly pronounced peak in g_{OO} for SSDQO is consistent with an *ab initio* molecular dynamics (AIMD) simulation study [13] and the larger coordination number is consistent with neutron diffraction value of less than 3 waters within 3.0 Å [14,15] and the AIMD value of 3 waters within 3.3 Å [13]. The peaks in g_{OH} and g_{HO} had minima at similar positions for SSDQO and SPC/E and the slightly more pronounced peaks are again consistent with the AIMD study [13]. The number of hydrogen bond donating waters in the first shell found by integrating the first peak in g_{OH} to 2.55 Å (the minimum for SSDQO) gave coordination numbers of 1.86 and 1.33 for SSDQO and SPC/E, respectively, while the number of hydrogen bond accepting waters found by integrating the first peak in g_{HO} to 2.6 Å (the minimum for SSDQO) gave coordination numbers of 0.90 and 0.77 for SSDQO and SPC/E, respectively. Thus, both models are reasonable in showing that more waters act as hydrogen bond donors than acceptors [16]. Furthermore, this demonstrates that even though SSDQO does not have an explicit hydrogen, it can act as better hydrogen bond donor than SPC/E. Also, the solvation around the methyl carbon shows that the water is shows no hydrogen bonding to it for both models (Fig. 2b), as expected for a nonpolar atom.

The radial distribution functions of SSDQO and SPC/E water around benzene show that both models were almost identical in their solvation of this hydrophobic solute (Fig. 3). The distribution functions around the carbons, g_{CO} and g_{CH} , showed that water has relatively little structure around the carbon atoms for both models (Fig. 3a), as expected. In addition, the distribution functions around the center of mass (cm) showed that the packing was similar (Fig. 3b), with small peak in g_{cmO} at ~ 3.5 Å corresponding to water near the center of the ring and a larger peak at ~ 5 Å corresponding to water along the periphery of the ring. Moreover, coordination numbers found by integrating the small peak in g_{CO} at ~ 3.5 Å together with the peak in g_{cmH} at ~ 2.5 Å indicate that on average, one water molecule is found at face of the benzene near the center of the ring with the hydrogen pointing inward (Table 1) even though there are two equivalent positions, one on each face. To our knowledge, no experimental data was available for comparison.

The radial distribution functions of SSDQO and SPC/E water around NMA also showed similar structure of both with minor variations (Fig. 4). The first solvation shell around the carbonyl oxygen was very similar (Fig. 4a), with the first peak at 2.8 Å in g_{OO} for both. The peak was slightly wider for SSDQO, so the coordination numbers found by integrating to 3.35 Å (the minimum for SSDQO) were 2.55 and 2.35 for SSDQO and SPC/E, respectively, while coordination numbers found by integrating to 3.25 Å (the minimum for SPC/E) were 2.35 and 2.25 for SSDQO and SPC/E, respectively. The g_{OO} for both were consistent with

an AIMD simulation [17,18] and the coordination numbers were consistent with a neutron diffraction value of two hydrogen bond donor waters [19] and the AIMD value of two waters [17,18]. The first peak at 1.8 Å in g_{OH} for both showed that water is acting as a hydrogen bond donor to the NMA oxygen in both. The second peak in g_{OO} for both models was at ~5 Å, corresponding to the expected position for water hydrogen bonded to the trans NH group [20], although it was somewhat more defined for SSDQO. On the other hand, while the solvation around the amide nitrogen was relatively less structured than around the carbonyl oxygen for both SSDQO and SPC/E (Fig. 4b), SSDQO had a slight peak in g_{NO} at ~3 Å with a population of 0.77 whereas there was only a shoulder for SPC/E with a population of 0.5 (Table 1). The stronger peak in SSDQO is consistent with the AIMD study [17,18] and the larger coordination number is consistent with the neutron diffraction value of one hydrogen bond acceptor water [19] and the AIMD value of one water [17,18]. The peak in g_{HO} at ~2 Å shows that the water is acting as a hydrogen bond acceptor. Also, the density in g_{NH} does not become significant until ~3.5 Å, which indicates that the water in the first shell is in dipolar orientation with respect to the amide nitrogen. Finally, the solvation around the methyl carbons is very similar to the methyl carbon of ethanol for both models and so are not shown.

CONCLUSIONS

Here, the SSDQO model using moments and van der Waals parameters from SPC/E was shown to give reasonable solvation of molecules with varying polarities in which the electrostatics were described by partial charges, which were in good agreement with solvation by SPC/E. There was a tendency for stronger hydrogen bonds in SSDQO than SPC/E, which was also shown for the interactions in pure liquid water [5], and we are currently optimizing the moments and van der Waals parameters for SSDQO to reproduce thermodynamic, dielectric, dynamical, and solvation properties of water under different conditions. However, overall these studies demonstrate that the single point, multipole moment interaction potential of SSDQO can be used with the multiple point, partial charge interaction potential commonly used in force fields used to describe the aqueous solvation of biological macromolecules. The radial distribution functions, which are sensitive to short-range interactions [10], demonstrates that SSDQO can reproduce the local solvation structure around a solute, including solutes that have hydrogen-bonding interactions with water. This is an important test because while the multipole approximation is expected to good for long-range interactions, the ability to reproduce short-range interactions needed to be demonstrated. Furthermore, the computational efficiency of SSDQO makes it potentially valuable for computational simulations of large macromolecules in aqueous solution, where the number of water molecules needed is substantial.

Acknowledgments

The authors are grateful to the National Science Foundation for the support of this work through grant number MCB-0544629. The calculations were performed on facilities provided by Georgetown University and administered by the division of Advanced Research Computing (ARC). Support was also provided by the William G. McGowan Foundation. In addition, this research was supported in part by the Intramural Research Program of the NIH, National Heart, Lung, and Blood Institute (Laboratory of Computational Biology).

References

1. Berendsen HJC, Grigera JR, Straatsma TP. J Phys Chem 1987;91:6269.
2. Jorgensen WL. J Am Chem Soc 1981;103:335.
3. van der Spoel D, van Maaren PJ, Berendsen HJC. J Chem Phys 1998;108:10220.
4. Mahoney MW, Jorgensen WL. J Chem Phys 2000;112:8910.
5. Ichiye T, Tan ML. J Chem Phys 2006;124:134504. [PubMed: 16613458]

6. Chowdhuri S, Tan ML, Ichiye T. J Chem Phys 2006;125:144513. [PubMed: 17042615]
7. Tan ML, Lucan L, Ichiye T. J Chem Phys 2006;124:174505. [PubMed: 16689581]
8. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. J Comp Chem 1983;4:187.
9. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. J Comp Chem 2005;26:1668. [PubMed: 16200636]
10. Andrea TA, Swope WC, Andersen HC. J Chem Phys 1983;79:4576.
11. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. J Chem Phys 1953;21:1087.
12. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. J Phys Chem B 1998;102:3586.
13. van Erp TS, Meijer EJ. J Chem Phys 2003;118:8831.
14. Sidhu KS, Goodfellow JM, Turner JZ. J Chem Phys 1999;110:7943.
15. Takamuku T, Saisho K, Nozawa S, Yamaguchi T. J Mol Liq 2005;119:133.
16. Fileti EE, Chaudhuri P, Canuto S. Chem Phys Lett 2004;400:494.
17. Mantz YA, Gerard H, Iftimie R, Martyna GJ. J Am Chem Soc 2004;126:4080. [PubMed: 15053577]
18. Mantz YA, Gerard H, Iftimie R, Martyna GJ. J Phys Chem B 2006;110:13523. [PubMed: 16821879]
19. Trabelsi S, Nasr S, Bahri M, Bellissent-Funel MC. J Mol Struct 2007;842:93.
20. Jorgensen WL, Swenson CJ. J Am Chem Soc 1985;107:1489.

Figure 1. Radial distribution for liquid water: oxygen – oxygen, oxygen – hydrogen (shifted upward by 1), and hydrogen – hydrogen (shifted upward by 2) for SSDQO (black) and SPC/E (gray) water.

Figure 2. Radial distribution functions for ethanol in SSDQO (black) and SPC/E (gray) water: a) ethanol oxygen - water oxygen, ethanol oxygen - water hydrogen (shifted upward by 1), and ethanol hydrogen - water oxygen (shifted upwards by 2) and b) ethanol methyl carbon - water oxygen and ethanol methyl carbon - water hydrogen (shifted by 2).

Figure 3. Radial distribution functions for benzene in SSDQO (black) and SPC/E (gray) water: a) benzene carbon - water oxygen and benzene carbon - hydrogen (shifted upward by 2) and b) benzene center of mass - water oxygen and benzene center of mass - hydrogen (shifted upward by 2).

Figure 4. Radial distribution functions for N-methylacetamide in SSDQO (black) and SPC/E (gray) water: a) NMA oxygen - water oxygen and NMA oxygen - hydrogen (shifted upward by 2) and b) NMA nitrogen - water oxygen, NMA nitrogen - hydrogen (shifted upward by 1), and NMA hydrogen - oxygen (shifted upward by 2).

Table 1

Comparison of the number of water molecules surrounding various atoms of the solutes, integrated to the minimum in the SSDQO (SPC/E) radial distribution function.

	SSDQO:SPC/E	SPC/E
Ethanol O – water O	3.35 (2.85)	2.90 (2.31)
Ethanol O – water H	1.86	1.33
Ethanol H – water O	0.90	0.77
Benzene cm – water O	1.07	0.95
Benzene cm – water H	1.28	1.06
NMA O – water O	2.55 (2.35)	2.35 (2.25)
NMA O – water H	2.52	2.12
NMA N – water O	0.77	0.52
NMA H – water O	1.03	0.75