

**MONTE CARLO THERMODYNAMIC AND STRUCTURAL PROPERTIES OF THE TIP4P WATER MODEL: DEPENDENCE ON THE COMPUTATIONAL CONDITIONS**

João Manuel Marques Cordeiro

Depto. de Física e Química - Faculdade de Engenharia de Ilha Solteira - UNESP - CP 31 - 15385-000 - Ilha Solteira - SP

Recebido em 7/7/97; aceito em 24/4/98

Classical Monte Carlo simulations were carried out on the NPT ensemble at 25°C and 1 atm, aiming to investigate the ability of the TIP4P water model [Jorgensen, Chandrasekhar, Madura, Impey and Klein; *J. Chem. Phys.*, 79 (1983) 926] to reproduce the newest structural picture of liquid water. The results were compared with recent neutron diffraction data [Soper; Bruni and Ricci; *J. Chem. Phys.*, 106 (1997) 247]. The influence of the computational conditions on the thermodynamic and structural results obtained with this model was also analyzed. The findings were compared with the original ones from Jorgensen *et al* [above-cited reference plus *Mol. Phys.*, 56 (1985) 1381]. It is notice that the thermodynamic results are dependent on the boundary conditions used, whereas the usual radial distribution functions  $g(O/O(r))$  and  $g(O/H(r))$  do not depend on them.

**Keywords:** Monte Carlo simulation; TIP4P water model; radial distribution function.

**INTRODUCTION**

It is difficult to over-emphasize the importance of water in chemical and biological processes. The study of its properties has attracted a good deal of attention in spite of the difficulties involved. Experimental results have provided a comprehensive picture of the way water molecules organize themselves in pure water and around particles dissolved in it and supplied a sensitive evaluation of existing theories of aqueous systems<sup>1</sup>. However there are a lot of obstacles in the theoretical study of the liquid state due to the complexity of the intermolecular forces. It is not feasible in general to calculate statistical mechanics averages by using standard mathematical approaches. With the dissemination of computers, methods such as Monte Carlo (MC) and molecular dynamics have become popular tools in this task<sup>2-5</sup>. Nevertheless, the success of the investigation is very much dependent on the availability of intermolecular potential functions that yield a reasonable model for the liquids. Consequently, there has been a great deal of interest to develop models that reproduce the water behavior. Among them one can find rigid models<sup>6-8</sup> and others including polarization effects<sup>9-11</sup>. In a recent work I have investigated the behavior of some models of water under various thermodynamic conditions and done some considerations about rigid and polarizable models<sup>12</sup>. As I have noticed, while polarizable potentials surely improve water models, in many ways they are less successful than non-polarization potentials which will continue to be used in water simulations for some time.

In recent years I have undertaken a theoretical investigation of the liquid properties emphasizing the liquid structure and the solute/solvent interaction in aqueous solutions<sup>13,32</sup>. For water I used the TIP4P model developed by Jorgensen *et al*<sup>7,14</sup>. This model is, together with the SPC and SPC/E models<sup>8</sup>, perhaps the most popular model in water simulations. In the quoted papers of Jorgensen *et al*, the authors have shown that the TIP4P model, besides reproducing the thermodynamic properties of water, represents very well the experimental  $g(O/O(r))$  radial distribution function (rdf) obtained from the

X-ray diffraction data of Narten and Levy<sup>15</sup>. The first peak in the  $g(O/O(r))$  function of the TIP4P model is slightly higher than the experimental one, which is a characteristic of several models reported by Jorgensen *et al*<sup>7</sup>. The  $g(O/H(r))$  and  $g(H/H(r))$  rdFs obtained with the TIP4P model have not been compared with experimental ones. Jorgensen *et al* pointed out that there would be many uncertainties in these radial distribution functions obtained from the X-ray diffraction experiments<sup>7</sup>. Experimental rdFs for water at 25°C and 1 atm were more recently obtained via neutron diffraction experiments by Soper and Philips<sup>16</sup>, in which the  $g(O/O(r))$  rdf was estimated by representing the total pair correlation as a Gaussian density distribution for the near-neighbor peak, plus a broader distribution for the rest of the pair correlation function. The peak obtained using this methodology is much higher than that reported by Jorgensen *et al* obtained from the Narten and Levy data<sup>7</sup>. Since then, some analytical models for water have been proposed by Blum *et al*<sup>17,18</sup> that represent very well the rdf published by Soper and Philips<sup>16</sup>. However, in 1994 Soper published a paper in which he used a new formalism to analyze the neutron diffraction data, that does not involve fitting Gaussians, but uses the minimum noise principle to estimate the entire  $g(r)$ <sup>19</sup>. The height of the first peak in the  $g(O/O(r))$  rdf, previously reported as 3.1, decreases in the new data to roughly 2.5. The author declares that in the new work the peak widths and heights have been determined more reliably<sup>19</sup>. Blum and Degreve do not refer to this work in a recently published paper<sup>20</sup>. In the present year, Soper and coworkers have made new improvements in the formalism and published site-site pair correlation functions for water from 25 to 400°C<sup>21</sup>. While in the work of 1994 the three  $g(r)$ s (OO, OH and HH) were calculated directly from the measured differential cross sections, each one independently of the others, in the later one, the three rdFs were estimated simultaneously to give the best fit to all three differential cross sections. In this way, systematic errors, which may otherwise cause a positive deviation in one rdf and a negative deviation in another, can be corrected for more accurately<sup>22</sup>. The  $g(O/O(r))$  plot reported agrees very well with the earlier X-ray results reported by Jorgensen and coworkers<sup>7</sup>. Thus, the TIP4P model has gained new strenght. By increasing, we have now an accurate experimental  $g(O/H(r))$  to compare with the theoretical one. On the other hand, in many situations it is necessary to investigate

e-mail: cordeiro@feis.unesp.br

systems under computational conditions different from those reported in the original papers of Jorgensen *et al.*<sup>7,14</sup>. Therefore, in the present work we investigated the behavior of the TIP4P model under new circumstances. The structural and thermodynamic results were compared with those reported by Jorgensen *et al.*<sup>7,14</sup> and with experimental ones described in the literature.

## DETAILS OF THE SIMULATION

**The Intermolecular Potential Function.** As the TIP4P is a rigid model, molecular relaxation effects are not considered. Following usual procedures in force field calculations, the energy  $E_{ab}$  between molecules **a** and **b** is represented by a sum of Coulomb and Lennard-Jones potentials centered on the sites, that is

$$E_{ab} = \sum_{ij} \left[ A_{ij} / r_{ij}^{12} - B_{ij} / r_{ij}^6 + q_i q_j / r_{ij} \right] \quad (1)$$

where  $r_{ij}$  is the distance between site  $i$  in **a** and site  $j$  in **b** and  $q_i$  and  $q_j$  are point charges located respectively on the  $i$  and  $j$  molecular sites. For each site  $k$ , the parameters  $A_{kk}$  and  $B_{kk}$  are given by  $A_{kk} = 4\epsilon_k \sigma_k^{12}$  and  $B_{kk} = 4\epsilon_k \sigma_k^6$ , where  $\epsilon_k$  and  $\sigma_k$  are the Lennard-Jones parameters for the  $k$ -th site.

**Monte Carlo Simulations** with Metropolis' importance sampling<sup>24</sup> and cubic box boundary conditions were carried out on the NPT ensemble at 298 K and 1.0 atm, using the DIADORIM code developed by L. C. G. Freitas<sup>25</sup>. To investigate size effects on the results two samples were simulated, one with 250 and another with 500 monomers in the reference box. Jorgensen and Madura have done the same for other sizes of samples and used a cutoff procedure to calculate the contribution of the Coulombic term to the configurational energy<sup>14</sup>. In the present work this procedure was not used, but instead the configurational energy was calculated considering all molecules in the box. Also the average was calculated over a significantly

larger number of configurations than by those authors. Details of the simulations of the present work and those ones with 216 and 512 molecules in the box, simulated by Jorgensen and Madura<sup>14</sup>, are recorded in table 1. It is interesting to notice that the cutoff distance used in both cases by Jorgensen and Madura is the same. Starting from an initial distribution of molecules in the central box, a new configuration was generated by randomly translating and rotating a randomly chosen molecule along Cartesian coordinates. In the NPT ensemble, new configurations are also generated by probing the density of the liquid with volume changes. After a volume change, the center of mass coordinates of all molecules in the reference box were scaled in a usual way<sup>23</sup>. Ranges for translations and rotations of monomers and also volume changes were adjusted to yield an acceptance/trial ratio between 0.4 and 0.45 for new configurations. Statistical uncertainties were calculated from separate averages over blocks of  $2 \times 10^5$  configurations. The calculations were carried out on IBM RISC 6000 workstations.

## RESULTS AND DISCUSSION

**Thermodynamic Results.** The thermodynamic results obtained for the simulations are presented in table 2 compared with those of Jorgensen and Madura<sup>14</sup> and with experimental data. The origin of the experimental intermolecular energy  $E_i$  will be discussed below. As one can observe the values obtained in the present work for density, heat of vaporization and intermolecular energy are about 5% greater than those reported by Jorgensen and Madura<sup>14</sup>. The values reported by these authors are in better agreement with the experimental ones than the results obtained in the present work. Simulations made using the DIADORIM program with cutoff procedure<sup>26</sup> have shown values in accordance with the results of Jorgensen and Madura<sup>14</sup>. Then the difference between the present results and those reported by Jorgensen and Madura must be due to the difference in the number of monomers that are interacting and, if this is true, the parameterization of the TIP4P model would

**Table 1.** Conditions of Monte Carlo simulations for TIP4P water at 25°C and 1 atm.

N	Equil.	Aver.	$r_c(\text{\AA})$	$\Delta r(\text{\AA})$	$\Delta q(\text{deg.})$	$\Delta V(\text{\AA}^3)$	edge( $\text{\AA}$ ) <sup>c</sup>
250 <sup>a</sup>	$8 \times 10^6$	$1.2 \times 10^7$	-	0.15	13.0	150	19.27
500 <sup>a</sup>	$8 \times 10^6$	$1.2 \times 10^7$	-	0.15	13.0	300	24.28
216 <sup>b</sup>	$4 \times 10^6$	$3 \times 10^6$	8.5	0.15	15.0	100	-
512 <sup>b</sup>	$1.2 \times 10^6$	$3 \times 10^6$	8.5	0.15	15.0	200	-

<sup>a</sup> number of molecules in the simulation box in this work

<sup>b</sup> number of molecules in the simulation box in reference 14

<sup>c</sup> edge of the box

$r_c$  is the cutoff ratio used by Jorgensen and Madura<sup>14</sup>. With this cutoff distance interactions between about 80 monomers are considered.

$\pm \Delta r$ ,  $\pm \Delta q$  and  $\pm \Delta V$  are the ranges for translations, rotations and volume, respectively.

**Table 2.** Calculated and experimental properties for water at 25°C and 1 atm.

	250 molec.	500 molec.	216 molec. <sup>a</sup>	512 molec. <sup>a</sup>	expert. <sup>c</sup>
$d(\text{g/cm}^3)$	$1.040 \pm 0.002$	$1.047 \pm 0.001$	$1.002 \pm 0.005$	$0.989 \pm 0.004$	0.997
$V_{mol}(\text{cm}^3/\text{mol})$	$17.31 \pm 0.03$	$17.18 \pm 0.02$	-	-	18.01
$\Delta H_{vap}(\text{kcal/mol})$	$11.16 \pm 0.01$	$11.12 \pm 0.01$	$10.64 \pm 0.02$	-	10.52
$-E_i(\text{kcal/mol})$	$10.57 \pm 0.01$	$10.53 \pm 0.01$	$10.11 \pm 0.02$	$10.00 \pm 0.01$	9.92
$C_\eta(\text{cal/mol.deg})$	$17.32 \pm 0.30$	$18.12 \pm 0.30$	$20.02 \pm 1.50$	-	17.99
$10^5 k(\text{atm}^{-1})$	$4.14 \pm 0.30$	$4.31 \pm 0.30$	$6.7 \pm 1.3$	-	4.58 <sup>d</sup>
$10^4 \alpha(\text{deg}^{-1})^b$	$2.79 \pm 0.30$	$6.62 \pm 0.30$	-	-	2.57 <sup>d</sup>

<sup>a</sup> reference 14

<sup>b</sup> Jorgensen and coworkers have reported a theoretical value of  $9.4 \times 10^{-4}$  for this property<sup>7</sup>.

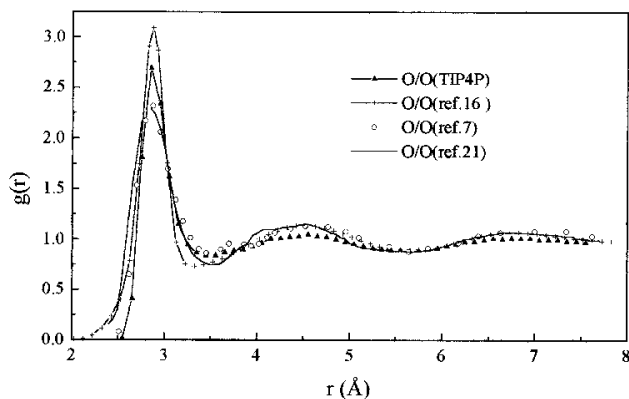
<sup>c</sup> reference 27

<sup>d</sup> reference 7

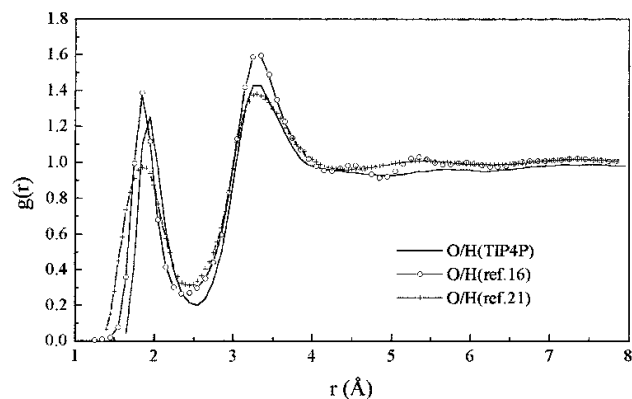
be dependent on the cutoff ratio. Previous results obtained by the author indicate that if a liquid is parameterized using a cutoff procedure to calculate the contribution of the Coulombic term to the interaction energy, these parameters do not reproduce the same results if one uses Ewald's sum. This is particularly expected for the TIP4P model for which parameterization was optimized to fit experimental properties using a cutoff procedure<sup>28</sup>. The experimental intermolecular energy was computed with the equation  $DH_{vap} = -E_i + P[V(g) - V(l)] = -E_i + RT$ , where one assumes that the gas is ideal and that the sum of kinetic and vibrational energies is the same for the liquid and the gas. As discussed by Jorgensen and Madura<sup>14</sup>, there is little error in making this approximation. The heat capacity  $C_p$ , the isothermal compressibility  $k$  and the thermal expansion  $\alpha$  have been computed from the standard fluctuation formulas<sup>23</sup>. The rates of convergence for these properties are slower than for the energy or density, especially  $k$  and  $\alpha$ <sup>29</sup>. The computed configurational contribution to the  $C_p$  has been summed to 3R (the classical kinetic energy contribution of the translation and rotation from the monomers) giving the value shown in table 2. This was the methodology also used by Jorgensen *et al*<sup>7</sup>. Considering the problems resulting from this approximation and from the way of calculating these properties, the obtained values are in good agreement with the experimental ones. The present results are in better agreement with the experiments than those of Jorgensen and Madura<sup>14</sup> and this can be attributed to the number of configurations averaged, which is considerably larger in the present work.

**Radial Distribution Functions.** Figs. 1 and 2 show the  $g(O/O(r))$  and  $g(O/H(r))$  rdf obtained in the present simulations compared with experimental results. The  $g(O/O(r))$  peak obtained by Soper and Philips<sup>16</sup> is presented for comparison. As one can observe, there is a good agreement between the  $g(O/O(r))$  obtained by Soper *et al*<sup>21</sup> and that obtained by Jorgensen *et al*<sup>7</sup> using the Narten and Levy experimental data<sup>14</sup>. The behavior verified for the TIP4P model in the present work is the same as that verified earlier by Jorgensen *et al*<sup>7,14</sup>. The height of  $g(O/O(r))$  peak is overestimated compared to the experimental one. The integration of the first peak up to the minimum point gives a value of 5.0 for the O-O coordination number, in agreement with the value reported by Jorgensen *et al*<sup>7</sup>.

As one can see in figure 2, the TIP4P model also produces an overestimation in the amplitude of the first peak of the O/H correlation compared with the experimental one. This seems to be in accord with the results of the simulations that show that the heat of vaporization obtained with the TIP4P model is greater than the experimental value. Therefore, it looks like the Lennard-Jones parameters are overestimated (at least the one of oxygen). For the second peak the agreement with the



**Figure 1.** O/O radial distribution functions computed for liquid TIP4P with 250 molecules in the box compared with experimental results. The data of ref. 13 are shown for comparison.

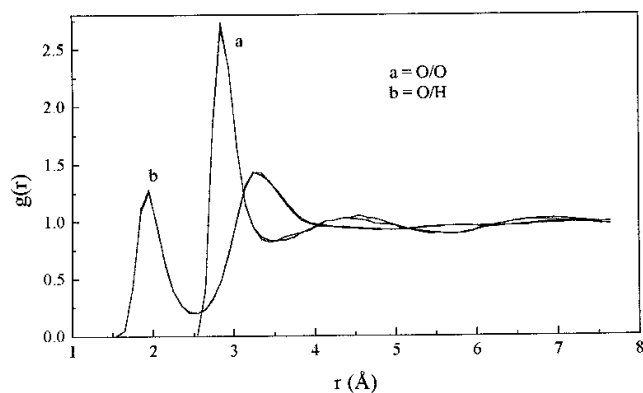


**Figure 2.** O/H radial distribution function computed for liquid TIP4P with 250 molecules in the box compared with experimental results.

experiment is very good. Nothing was found in the literature for comparing the  $g(O/H(r))$  obtained with the TIP4P model with experimental results. By integrating the first peak up to the minimum point a value of 4.0 for the O-H coordination number was obtained. This value agrees with that reported by Jorgensen *et al*<sup>7</sup> and is a little greater than the experimental one<sup>21</sup>. Some analyses suggest that the height of the OH peak is not a reliable measure of the degree of hydrogen bonding in water, so that geometric or energetic aspects, or both, should be taken into account<sup>30-31</sup>. Since the average O-O coordination number per monomer is about 1 unit greater than the average number of hydrogen bonds, one can conclude that one non-hydrogen-bonded molecule is trapped in the first hydration shell due to the intermolecular forces. It would be interesting to try to imagine the effect on the properties of water of this extra molecule bonded in the field of the water molecule. The distortion in the tetrahedral field arising from this molecule must result in a decrease in the melting point of water with clear effects on the development of life. In Fig. 3 we show the  $g(O/O(r))$  and  $g(O/H(r))$  rdfs for water obtained with the TIP4P model for boxes with 250 and 500 molecules. As one can observe, the plots are exactly equal meaning that the  $g(r)$  obtained are not dependent on box size.

## CONCLUSIONS

The TIP4P parameterization looks like being dependent on the number of molecules that interact in the simulation box. The parameterization of the TIP4P model was optimized to reproduce the experimental properties of water using a cutoff procedure to calculate the contribution of the Coulombic term



**Figure 3.** The O/O and O/H radial distribution function for water with 250 and 500 molecules in the simulation box.

to the configurational energy. The model super-estimates the experimental amplitude of the rdf for water as compared with the more recent results. The greater amplitude in the O/O rdf and of the first peak of the O/H rdf is in accord with the greater heat of vaporization yielded by the TIP4P model in the present simulations. Perhaps this is a consequence of attributing a too big  $\epsilon$  Lennard-Jones parameter to the oxygen atom. Considering the thermodynamic results and rdf plots, there is no difference between the simulation with 250 and 500 molecules in the simulation box. So one can do the simulation using the smaller number of molecules — one does not lose information and the simulation is faster. The number of configurations used in the present work yield an acceptable convergence of the isothermal compressibility and thermal expansion coefficients. Finally, the disagreement between the experimental value of the mean molar energy and the theoretical one obtained by Degrève and Blum<sup>18</sup> is in line with the fact that their water model fits very well the old data for the rdf of Soper and Phylips<sup>16</sup>, which according to more recent data must be overestimated<sup>21</sup>.

#### ACKNOWLEDGMENTS

The author is grateful to Dr. Alan K. Soper for the experimental  $g(O/O(r))$  and  $g(O/H(r))$  data and the profitable discussions.

#### REFERENCES

1. Soper, A. K.; *J. Phys.:Condens. Matter.*, **1997**, *50*, 2717.
2. Gibson, K. D.; Sheraga, H. A.; *J. Comp. Chem.* **1990**, *11*, 468.
3. Jorgensen, W. L.; *Chemtracts - Organic Chemistry* **1991**, *4*, 91.
4. Straatsma, T. P.; McCammon, J. A.; *Annu. Rev. Phys. Chem.* **1992**, *43*, 407.
5. Ladanyi, B. M.; Skaf, M. S.; *Annu. Rev. Phys. Chem.* **1993**, *44*, 335.
6. Stillinger, F. H.; Rahman, A.; *J. Chem. Phys.* **1974**, *60*, 1545
7. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.; *J. Chem. Phys.* **1983**, *79*, 926.
8. Berendsen, H. J. C.; Grigera J. R.; Straatsma, T. P.; *J. Phys. Chem.* **1987**, *91*, 6269.
9. Brodholt, J.; Sampoli, M.;Vallauri, R.; *Mol. Phys.* **1995**, *86*, 149.
10. Ahlstrom, P.; Wallqvist, A.; Engstrom, S.; Jonsson, B.; *Mol. Phys.* **1989**, *68*, 563.
11. Tellemann, O.; Jonsson, B.; *Mol. Phys.* **1987**, *60*, 193.
12. Cordeiro, J. M. M.; unpublished results.
13. Freitas, L. C. G.; Cordeiro, J. M. M.; *J. Mol. Struct. (Theochem)* **1995**, *335*, 189.
14. Jorgensen, W. L.; Madura, J. D.; *Mol. Phys.* **1985**, *56*, 1381.
15. Narten, A. H.; Levy, H. A.; *J. Chem. Phys.* **1971**, *55*, 2263.
16. Soper, A. K.; Phylips, M. G.; *Chem. Phys.* **1986**, *107*, 47.
17. Blum, L.; Vericat, F.; Bratko, D.; *J. Chem. Phys.* **1995**, *102*, 1461.
18. Degrève, L.; Blum, L.; *Physica A* **1996**, *224*, 550.
19. Soper, A. K.; *J. Chem. Phys.* **1994**, *101*, 6888.
20. Blum, L.; Degrève, L.; *Mol. Phys.* **1996**, *88*, 585.
21. Soper, A. K.; Bruni, F.; Ricci, M. A.; *J. Chem. Phys.* **1997**, *106*, 247.
22. Soper, A. K.; personal communication.
23. Allen, M. P.; Tildesley, D. J.; *Computer Simulations of Liquids* Oxford University Press, Oxford, 1987.
24. Metropolis, N. A.; Rosebluth, W.; Rosebluth, M. N.; Teller, A. H.; Teller, E.; *J. Chem. Phys.* **1953**, *21*, 108.
25. DIADORIM fortran code by L. C. G. Freitas, Theoretical Chemistry Lab., Chemistry Dept. - UFSCar - S. Carlos - SP - Brazil.
26. Freitas, L. C. G.; *J. Mol. Struct. (Theochem)* **1993**, *282*, 151.
27. Lide, D. R., editor-in-chief; *CRC Handbook of Physics and Chemistry*, 73RD Edition, CRC Press Inc, 1992.
28. Jorgensen, W. L.; personal communication.
29. Jorgensen, W. L.; *Chem. Phys. Letters* **1982**, *92*, 405.
30. Chialvo, A. A., Cummings, P. T.; *J. Chem. Phys.* **1994**, *101*, 4466.
31. Gorbaty, Yu. E. , Kalinishev, A. G.; *J. Phys. Chem.* **1995**, *99*, 5336.
32. note added in proof: Besides the work mentioned above, I have a paper accepted by the *J. Mol. Liq.* in collaboration with L. C. Gomide Freitas and F. Garbujo (Theoretical studies of liquids by computer simulations: the water-acetone mixture) and another one published in *Int. J. Quant. Chem.* **1997**, *65*, 709.