Chemical Physics Letters 651 (2016) 92–96

Contents lists available at ScienceDirect

Chemical Physics Letters

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

Molecular dynamics simulations to study the solvent influence on protein structure

Hector Dominguez [1]

Instituto de Investigaciones en Materiales, Universidad Nacional Autónoma de México, México D.F. 04510, Mexico

ARTICLE INFO

Article history:
Received 7 February 2016
In final form 10 March 2016
Available online 18 March 2016

ABSTRACT

Molecular simulations were carried out to study the influence of different water models in two protein systems. Most of the solvents used in protein simulations, e.g., SPC/E or TIP3P, fail to reproduce the bulk water static dielectric constant. Recently a new water model, TIP4P/ϵ, which reproduces the experimental dielectric constant was reported. Therefore, simulations for two different proteins, Lysozyme and Ubiquitin with SPC/E, TIP3P and TIP4P/ϵ solvents were carried out. Dielectric constants and structural properties were calculated and comparisons were conducted. The structural properties between the three models are very similar, however, the dielectric constants are different in each case.

1. Introduction

The study of electric properties in biological systems, such as the dielectric constant, have been conducted for several years [1–6]. In particular, electrostatic properties of biomolecules will help us to understand fundamental biochemical mechanisms, such as catalysis, redox reactions or ion homeostasis effects, among others [5,7,8]. From the theoretical point of view, it is important to make a good evaluation of the dielectric properties of those system, however, the calculated effective dielectric constants in proteins are in general larger than those observed in actual experiments. This discrepancy could show the significant influence of the solvent around the proteins. On the other hand, not only the electric properties are important but also the structure, the dynamics and the thermodynamics of the proteins are relevant quantities to study processes such as folding, binding, solvation, etc. [9–13].

In particular, computer simulations have played an important role in the study of such complex systems [14–17]. From the computational point of view, it is always desirable to use accurate models in order to reproduce experimental observations, i.e., it is important to choose the right force fields. People have used several force fields to simulate proteins (CHARMM, GROMOS, AMBER, OPLS) and most of the classical simulations are conducted with traditional water models (SPC, SPC/E, TIP3P). However, in simulations with explicit water contributions, it is expected to have a large solvent influence, in particular, for the study of dielectric properties [18]. Unfortunately, most of the atomistic classical water models have failed to reproduce the experimental static dielectric constant. Therefore, the problem is how to obtain good dielectric properties in biological systems if the solvent is not appropriate.

Nowadays, several non-polarizable water models have been proposed in the literature [19–21], however, none of them reproduces all bulk properties. Recently, a comparison between most of the classical non-polarizable models was conducted that showed that all of them fail to reproduce the experimental dielectric constant [22]. Some while ago, a new water model (TIP4P/ϵ) was reported which reproduced correctly the static dielectric constant [23]. Moreover, the new model also provided good agreement with thermodynamic and structural data. The TIP4P/ϵ model is a rigid molecule with three charges and one Lennard Jones site on the oxygen atom. Like the other empirical models, it was constructed with classical potentials whose parameters were obtained to fit specific experimental properties. Therefore, there are not vibrational or electronic contributions to describe the target properties. Nevertheless, the model should be good enough to describe the behaviour of bulk properties at the classical simulation level.

So, in the present work, we report studies of two proteins, Lysozyme and Ubiquitin, with TIP4P/ϵ water. Dielectric constants and structural properties were calculated and comparisons with other two water models, SPC/E and TIP3P, were conducted.

2. Model

Molecular dynamics simulations with an atomistic approach were conducted for the Lysozyme enzyme and the Ubiquitin

E-mail address: hectordc@unam.mx
[1] Present address: Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada. On sabbatical leave.

http://dx.doi.org/10.1016/j.cplett.2016.03.026
0009-2614/

protein. The initial configurations were taken from the research data bank, 1AKI1 and 1UBQ for the Lysozyme and Ubiquitin, respectively. All simulations were carried out in the NPT ensemble at a temperature $T = 300$ K and pressure $P = 1$ bar using the GROMOS43a1 force field. In addition, a velocity rescaling thermostat was used to maintain a constant temperature. This thermostat is essentially a Berendsen thermostat with an stochastic term to keep the correct kinetic energy, i.e., the thermostat produces a correct canonical ensemble [24]. Moreover, during the simulations two groups were created for temperature control, one for the protein and one for the solvent. In both cases, the temperature was monitored to ensure that both groups had the same value ($T = 300$ K). On the other hand, a Parrinello–Rahman barostat was used to maintain the pressure. Temperature and pressure relaxation time constants of $\tau_T = 0.1$ ps and $\tau_P = 2.0$ ps were used, respectively. The electrostatic interactions were handle with the particle mesh Ewald method. For the solvent, the SPC/E, the TIP3P and the TIP4P/ϵ water models were used. The first two solvents are 3-site models whereas the TIP4P/ϵ is a 4-site model. For the Lysozyme system 22 801 SPC/E or TIP3P water molecules were used whereas 23 299 water molecules were used for the TIP4P/ϵ system. For the Ubiquitin protein, 19 845 water molecules were used for the SPC/E and TIP3P models while 20 402 molecules were employed for the TIP4P/ϵ model. Simulations were performed for 200 ns after 50 ns to achieve equilibrium conditions with a timestep of dt = 0.002 ps. All simulations were ran with the GROMACS 4.5.6 software [25].

3. Results

In the next section, we present the results of two different proteins solvated with three different water models. In most of the biological simulations the SPC/E and TIP3P water models are used as common solvents. It is not common to use 4-site water models in those systems even when they might reproduce, sometimes, better bulk properties [22]. Here, we present simulations with a 4-site model, which better reproduces several bulk properties (even the dielectric constant) in contrast with the most common 3-site models. We show that the new model presents a different behaviour compare with the other two models.

3.1. Dielectric constant

The first analysis was conducted for the calculation of the static dielectric constant in the protein and in the whole system. In Figure 1 the results for the Lysozyme enzyme are shown. There, it is observed that the protein with TIP4P/ϵ has a slightly higher dielectric constant than that with SPC/E (top of Figure 1). On the other hand, the dielectric constant for the protein with the TIP3P water is higher than the other water models. Although, there are not unique protein dielectric constants reported in the literature, the general values ranged from 2 to 7 [1,3,18]. Higher dielectric constants have been reported, however, the reason of those values is still matter of several studies [1,3]. In our case, the value obtained for the TIP3P model is higher than the one expected.

The calculation of the dielectric constant from computer simulations is a demanding task. In fact, long runs are necessary to obtain a reasonable estimate of this property (in our case we ran the systems up to 200 ns). Here, the static dielectric constant was calculated from the time-average of the dipole moment fluctuations and using 'tinfoil' boundary conditions,

$$\epsilon = 1 + \frac{1}{3Vk_BT\epsilon_0}(\langle \mathbf{M}^2 \rangle - \langle \mathbf{M} \rangle^2) \quad (1)$$

where V is the volume of the system, T is the temperature, k_B is the Boltzmann constant, ϵ_0 is the vacuum permittivity, and $\mathbf{M}$ the dipole moment ($\mathbf{M} = \sum q_i \mathbf{r}$, q_i and $\mathbf{r}$ are the charge and the coordinate vector of atom i, respectively).

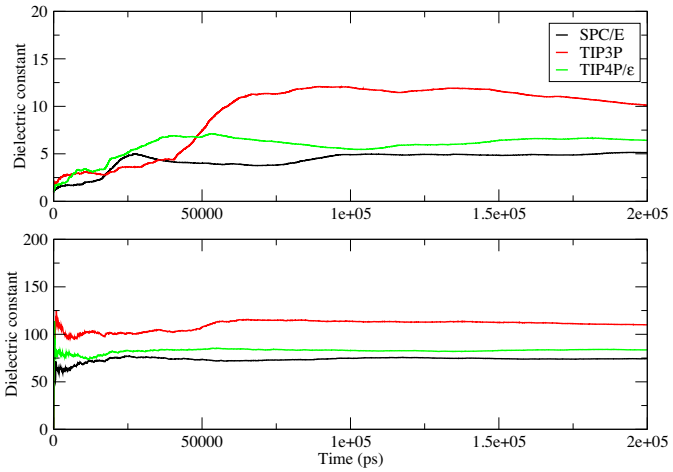

Figure 1. Static dielectric constant as a function of time for the Lysozyme system. Top figure for the protein and bottom figure for the whole system. Black lines are the simulations for the SPC/E model, red lines for the TIP3P model and the green lines for the TIP4P/ϵ model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

When the dielectric constant of the whole system is calculated, the same trends are observed, i.e., the TIP3P, the TIP4P/ϵ and the SPC/E water models have the highest, the intermediate and the lowest values, respectively (bottom of Figure 1). The dielectric constants obtained for the SPC/E, TIP3P and TIP4P/ϵ models were $\approx$74, 113 and 83, respectively. Previous works reported a dielectric constant of 62 for TIP3P [13], 68 for SPC [14] and 70.9 for SPC/E water [16]. Due to the small protein dielectric constant, those values are close to the bulk water of the corresponding model. Therefore, the TIP4P/ϵ model is close to the experimental value [23]. Average values and standard deviations of those quantities are given in Table 1. When the total dipole moment, $\langle M_0 \rangle$, was calculated, we found values of 165.6 D, 244.5 D and 190.0 D for the SPC/E, TIP3P and TIP4P/ϵ models, respectively. Previous works reported values of 184.2 D (using SPC water) [17] and 223 D (using TIP3P water) [13].

For the Ubiquitin protein system, the dielectric constant results are given in Figure 2. There, different trends are observed. For instance, the highest dielectric constant value of the protein is obtained for the SPC/E water model and the lowest for the TIP3P one.

Once again, the protein dielectric constant with the TIP4P/ϵ water model is nearly in the middle (top of Figure 2). For the static dielectric constant of the complete system, the SPC/E and TIP4P/ϵ models have nearly the same value ($\approx$94), which is close to the interval of the dielectric constants reported in an earlier experimental work (81–91) [5]. However, the TIP3P model shows a higher value. The calculated total dipole moments are 293.5 D, 229.3 D and 288.4 D for the SPC/E, TIP3P and TIP4P/ϵ models, respectively. Here, the value reported by other authors is 244 [13].

3.2. Structural properties

How the protein structure is modified by the different solvent models was also analyzed. In Figure 3, the radius of gyration is reported for the two proteins. For the Lysozyme, values of 1.36 nm, 1.41 nm and 1.40 nm are obtained for the SPC/E, TIP3P and TIP4P/ϵ water models, respectively (top of Figure 3 and Table 1). The values for the TIP3P and TIP4P/ϵ are in good agreement with experiments and other simulation works, $\approx$1.4 nm [9,10]. By using these

Table 1
Average ($\langle X \rangle$) data for the dielectric constant (ϵ), the radius of gyration (R_g), the solvent accessible surface (Area) and the number of hydrogen bonds (h_b). The respective standard deviation (σ^2) of calculated properties are given in the parentheses.

	Lysozyme			Ubiquitin		
	SPC/E	TIP3P	TIP4P/ϵ	SPC/E	TIP3P	TIP4P/ϵ
	$\langle X \rangle$ (σ^2)	$\langle X \rangle$ (σ^2)	$\langle X \rangle$ (σ^2)	$\langle X \rangle$ (σ^2)	$\langle X \rangle$ (σ^2)	$\langle X \rangle$ (σ^2)
ϵ (protein)	4.94 (0.76)	11.29 (0.29)	6.23 (0.15)	15.01 (0.09)	8.76 (0.38)	11.31 (0.98)
ϵ (system)	74.41 (0.44)	112.61 (1.4)	83.13 (0.37)	94.64 (0.27)	108.75 (0.26)	93.71 (3.45)
R_g (nm)	1.36 (5.0×10^{-5})	1.41 (5.5×10^{-4})	1.40 (1.1×10^{-4})	1.13 (2.5×10^{-5})	1.22 (4.3×10^{-4})	1.16 (1.5×10^{-4})
Area (nm^2)	65.5 (1.2)	83.7 (5.2)	80.3 (2.3)	45.8 (0.39)	56.9 (3.6)	50.6 (1.2)
h_b	228.9 (26)	386.9 (46)	360.4 (32)	168.8 (12)	258.9 (66)	219.5 (15)

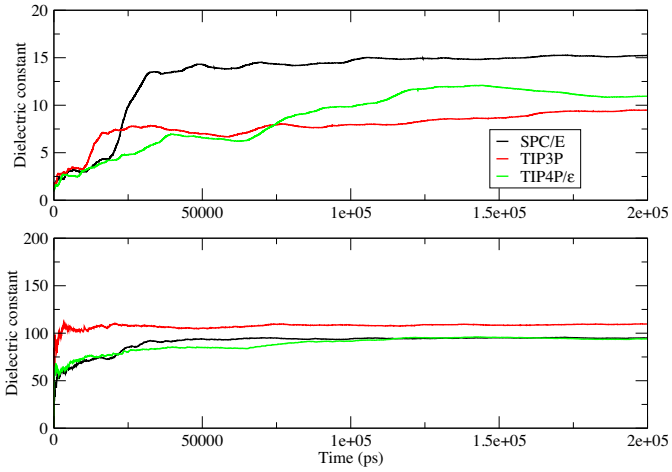

Figure 2. Static dielectric constant as a function of time for the Ubiquitin system. Top figure for the protein and bottom figure for the whole system. Black lines are the simulations for the SPC/E model, red lines for the TIP3P model and the green lines for the TIP4P/ϵ model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

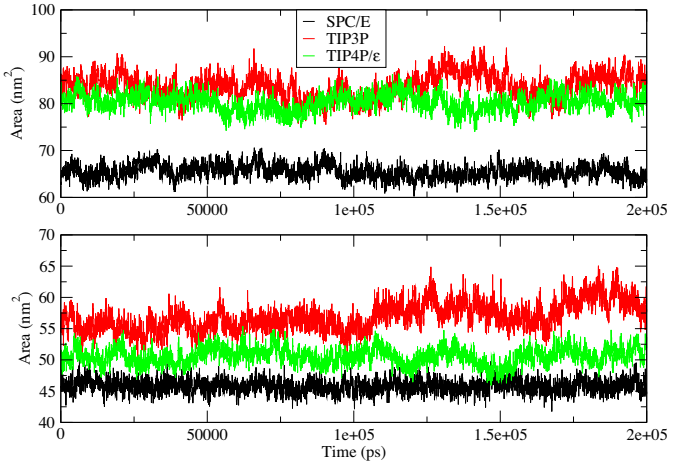

Figure 4. Total solvent accessible surface for the Lysozyme (top figure) and Ubiquitin (bottom figure) proteins as a function of time. Black lines are the simulations for the SPC/E model, red lines for the TIP3P model and the green lines for the TIP4P/ϵ model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

numbers, the average protein radii are 1.75 nm, 1.82 nm and 1.81 nm for the SPC/E, TIP3P and TIP4P/ϵ, respectively.

For the Ubiquitin protein the values are 1.13 nm, 1.22 nm and 1.16 nm for the SPC/E, TIP3P and TIP4P/ϵ models, respectively (Figure 3b). Then the average protein radii are 1.46 nm, 1.58 nm and 1.50 nm for the SPC/E, TIP3P and TIP4P/ϵ, respectively. There are not significant differences in these values, however, the value for the TIP4P/ϵ model is in the middle of the SPC/E and TIP3P models. The data, for the radius of gyration, are also in good agreement with previous works (≈1.1–1.2) [11,12].

The total solvent accessible surface was also calculated to investigate the influence of the different water models on the proteins. In Figure 4 and Table 1, the results for the Lysozyme protein

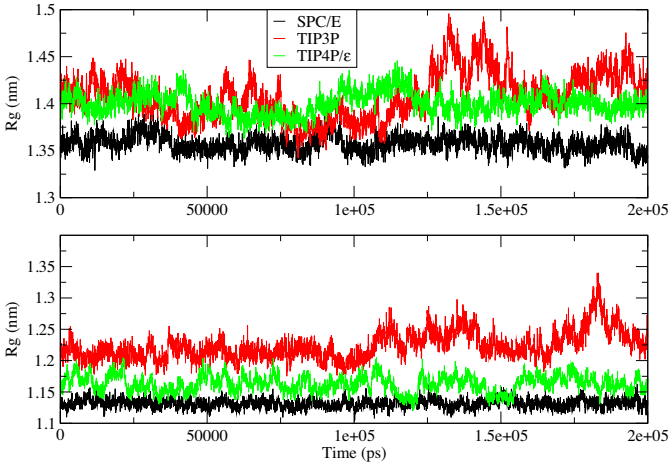

Figure 3. Radius of gyration for the Lysozyme (top figure) and Ubiquitin (bottom figure) proteins as a function of time. Black lines are the simulations for the SPC/E model, red lines for the TIP3P model and the green lines for the TIP4P/ϵ model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

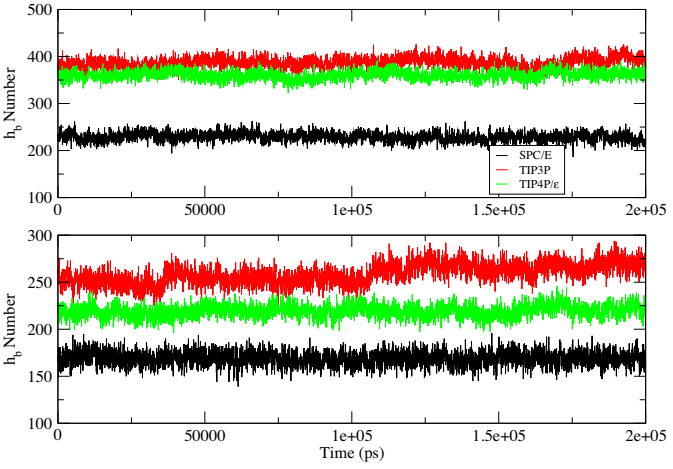

Figure 5. Total hydrogen bonds (protein-solvent) for the Lysozyme (top figure) and Ubiquitin (bottom figure) proteins as a function of time. Black lines are the simulations for the SPC/E model, red lines for the TIP3P model and the green lines for the TIP4P/ϵ model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

are shown. Whereas the TIP3P and TIP4P/ϵ models have similar surfaces ($\approx$83.7 nm^2 and 80.3 nm^2, respectively) the SPC/E model has a lower accessible surface area (65.5 nm^2). Here, the difference in area given by the SPC/E model is significantly lower than that of the TIP3P and TIP4P/ϵ models. For the Ubiquitin-solvent system, the surfaces are significant lower, 45.8 nm^2, 56.9 nm^2 and 50.6 nm^2 for the SPC/E, TIP3P and TIP4P/ϵ models, respectively (bottom of Figure 4). Once again, it is observed that the TIP4P/ϵ value is in the middle of the other two.

The number of hydrogen bonds was also calculated. The results for the Lysozyme protein are shown in Figure 5 and Table 1. It can be seen that nearly the same number of bonds is obtained for the TIP3P and TIP4P/ϵ solvents. However, in the case of the SPC/E, the number of hydrogen bonds decays significantly with respect to the other models. For the Ubiquitin protein, the same trends are observed in the number of hydrogen bonds, i.e., the TIP3P and TIP4P/ϵ solvents have similar number of bonds whereas the SPC/E model has a reduced number.

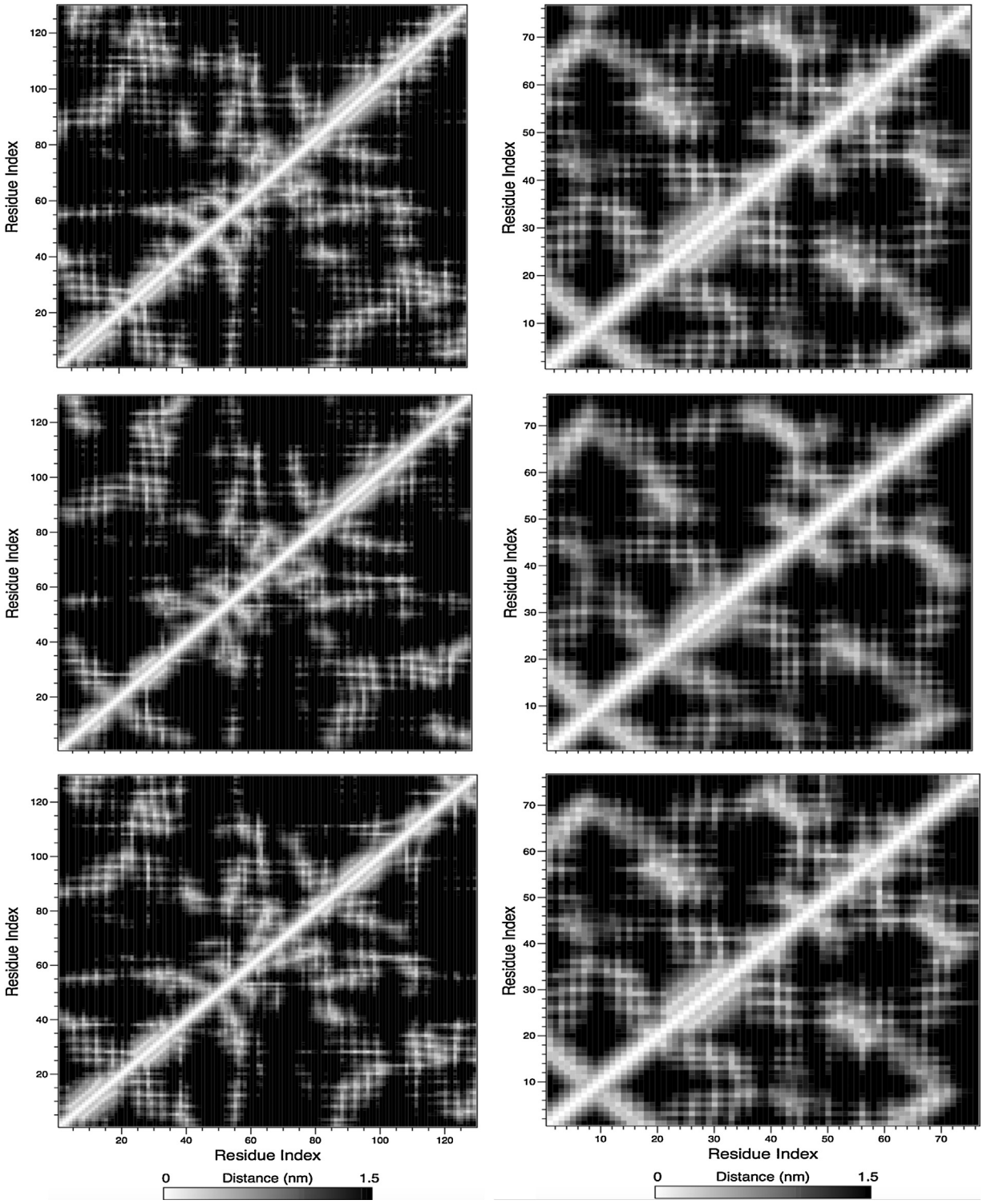

Figure 6. Mean smallest distance between protein residues. Left figures are for the Lysozyme protein with SPC/E (top), TIP3P (middle) and TIP4P/ϵ (bottom) models. Right figures are for the Ubiquitin protein with SPC/E (top), TIP3P (middle) and TIP4P/ϵ (bottom) models.

The mean residues distances in the proteins can also give us some information about the influence of the different solvents in the proteins structures. For the Lysozyme enzyme, the largest distance was found between residues 50–90 (left of Figure 6). Although the diagrams look alike, it is possible to observe more blank spots in the TIP4P/ϵ model suggesting that residues are closer than those in the other models. In the case of the Ubiquitin protein, the residues distances for the SPC/E and TIP4P/ϵ models look very similar to each other but slightly different compare with the TIP3P, in particular in regions of residues 25–35 and 55–70.

Diffusion coefficients were also evaluated for the proteins with the different solvents. For the Lysozyme we found values of 8.8×10^{-7} cm^2/s, 5.5×10^{-7} cm^2/s and 7.9×10^{-7} cm^2/s for the SPC/E, TIP3P and TIP4P/ϵ models, respectively. These numbers are smaller than the value reported in a previous simulation (10.9×10^{-7} cm^2/s) [23]. For the Ubiquitin protein the diffusions were 5.7×10^{-7} cm^2/s, 6.9×10^{-7} cm^2/s and 7.7×10^{-7} cm^2/s for the SPC/E, TIP3P and TIP4P/ϵ models, respectively. Here, previous simulations, using TIP3P water, reported diffusions between 2 and 5×10^{-7} cm^2/s [26].

4. Conclusions

Molecular dynamics simulations were conducted to study the differences of two proteins in contact with three different water models. Few works have been done to compare the influence of water models in protein systems and most of those works have used common solvents (SPC, SPC/E, TIP3P, TIP4P), which fail to reproduce the static dielectric constant. It is worth noticing that some models are worse than others, and they can even fail to reproduce other bulk properties. For instance, TIP4P is worse than TIP3P to calculate the dielectric constant [22]. Few years ago, a new rigid non-polarizable water model was developed, the TIP4P/ϵ. Although the TIP4P/ϵ model is also an empirical representation for simple systems, it describes most experimental observations. Since its parameters were obtained to reproduce target properties, we cannot expect that it could reproduce all experimental data. Moreover, with this classical force field it is not possible to include electronic contributions which might be involved in the calculations of complex properties such as the dielectric constant [27]. However, in order to include those contributions different approaches should be considered, such as density functional methods, which are beyond the purpose of the present work. For our simulations classical solvents will be enough to obtain the behaviour of biological systems. In fact, to the best of our knowledge there has not been works to test the recent TIP4P/ϵ water model with protein systems.

From the Lysozyme enzyme results, it is noted that the TIP3P water model overestimates the protein dielectric constant compared with the SPC/E model. However, in the case of the Ubiquitin protein we observe the opposite trend, i.e., the dielectric constant with the SPC/E model is higher than that with the TIP3P model. Nevertheless, in both cases the TIP4P/ϵ water model gives intermediate results and in some cases, it has a better agreement with experimental results.

In terms of structural properties, the three water models do not show significant differences among each other and in general they are in agreement with reported data. However, it is interesting to note that the TIP4P/ϵ water model once again gives values between the other two water models.

It is well known that the calculation of any protein property by using computer simulations depends on the particular force field model employed [11,12]. Some force fields can give good results in agreement with experiments however it is not possible to say in advance which model will provide the best results. In the present work, we have evaluated some protein properties by using, what it seems to be, a better water model. However, from the results, it is not completely evident that the new water model improves all the studied properties. For instance, the structural properties do not present many differences, whereas there is a slight improvement in terms of the dielectric constant value. So, if we are interested in structural properties any water model could give similar results, however, if we are interested in obtaining electric properties, the TIP4P/ϵ water model might give us better results.

Acknowledgments

This research work was supported by UNAM and Conacyt Mexico through sabbatical scholarships which I acknowledge. I also acknowledge DGTIC-UNAM for the supercomputer facilities through grant SC16-1-IR-21. I also thank all the reviewers for their valuable comments to improve the manuscript and Margarita Rivera for reading the manuscript.

References

[1] P. Kukic, D. Farrell, L.P. McIntosh, B. Garcia-Moreno, K.S. Jensen, Z. Toleikis, K. Teilum, J.E. Nielson, J. Am. Chem. Soc. 13 (2013) 16968.
[2] R. Affleck, C.A. Haynes, D.S. Clark, Proc. Natl. Acad. Sci. 89 (1992) 5167.
[3] D.A. Karp, A.P. Gittis, M.R. Stahley, C.A. Fitch, W. Stites, B. Garcia-Moreno, Biophys. J. 92 (2007) 2041.
[4] L.N. Rashkovich, V.A. Smirnov, E.V. Petrova, Phys. Solid State 50 (2008) 631.
[5] A. Knocks, H. Weingártner, J. Phys. Chem. B 105 (2001) 3635.
[6] D.C. Rees, J. Mol. Biol. 141 (1980) 323.
[7] S. Floros, M. Liakopoulou-Kyriakifes, K. Karatasos, G.E. Papadopoulos, Eur. Biophys. 44 (2015) 599.
[8] C. Cametti, S. Marchetti, C. Gambi, G. Onon, J. Phys. Chem. B 115 (2011) 7144.
[9] J. García de la Torre, M.L. Huertas, B. Carrasco, Biophys. J. 78 (2000) 719.
[10] W.R. Krigbaum, F.R. Kúgler, Biochemistry 9 (1970) 1216.
[11] D.B. Kony, P.H. Húnenberger, W.F. van Gunsteren, Protein Sci. 16 (2007) 1101.
[12] N.J. Marianayagam, S.E. Jackson, J. R. Soc. Interface 2 (2005) 47.
[13] D.V. Matyushov, J. Chem. Phys. 136 (2012) 085102.
[14] J.W. Pitera, M. Falta, W.F. van Gunsteren, Biophys. J. 80 (2001) 2546.
[15] T. Simonson, C.L. Brooks, J. Am. Chem. Soc. 118 (1996) 8452.
[16] R. Ghosh, S. Banerjee, M. Hazra, S. Roy, B. Bagchi, J. Phys. Chem. 141 (2014), 22D53.
[17] P.E. Smith, R.M. Brunne, A.E. Mark, W.F. van Gunsteren, J. Phys. Chem. 97 (1993) 2009.
[18] L. Li, C. Li, Z. Zhang, E. Alexov, J. Chem. Theory Comput. 9 (2013) 2126.
[19] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, J. Phys. Chem. 91 (1987) 6269.
[20] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926.
[21] J.L.F. Abascal, C. Vega, J. Chem. Phys. 123 (2005) 234505.
[22] C. Vega, J.L.F. Abascal, Phys. Chem. Chem. Phys. 13 (2011) 19663.
[23] R. Fuentes-Azcatl, J. Alejandre, J. Phys. Chem. B 118 (2014) 1263.
[24] P.H. Hünenberg, Adv. Polym. Sci. 173 (2005) 105.
[25] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, J. Chem. Theory Comput. 2008 (2008) 435.
[26] L.A. Abriata, M. Dal Peraro, Sci. Rep. 5 (2015) 1.
[27] J.B. Hasted, in: F. Franks (Ed.), Water: A Comprehensive Treatise, The Physics and Physical Chemistry of Water, vol. 1, Springer, New York, 1972 (Chapter 7).