Unusual properties of aqueous solutions of L-proline: A molecular dynamics study

Monica Civera [a], Maurizio Sironi [a,c], Sandro L. Fornili [b,c,*]

[a] Department of Physical Chemistry and Electrochemistry, University of Milan, Via Golgi 19, I-20133 Milan, Italy
[b] CNR-INFM and Department of Information Technologies, University of Milan, Via Bramante 65, I-26013 Crema, Italy
[c] Centre for Biomolecular Interdisciplinary Studies and Industrial Applications, Via Fratelli Cervi 93, I-20090 Segrate, Italy

Received 25 July 2005

Abstract

Aqueous solutions of the bioprotectant proline are simulated for solute molar fractions ranging from 2.0×10^{-3} to 2.3×10^{-1}. Statistical analyses show that proline affects the water structure more strongly than glycine betaine and trimethylamine-N-oxide, two of the most effective bioprotectants widely diffuse in nature, and as strongly as *tert*-butyl alcohol, a protein denaturant which at high concentration self-aggregates. No evidence is found, however, that proline self-aggregates as it has been previously suggested to explain experimental findings on concentrated proline solutions. Nevertheless, the behavior of the diffusion coefficients of proline and water vs. solute concentration qualitatively agrees with such results.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

In adverse conditions, such as drought, low or high temperature, or high external salinity, many organisms accumulate small organic compounds (called compatible osmolytes), selected by evolution to protect the cellular machinery [1]. Osmolytes protect native proteins from denaturation and some of them behave as chemical chaperones by promoting the correct proteins refolding in vitro and in the cell [2–6]. This effect appears to be related to the osmolyte property of being displaced from the protein domain [7], due to their unfavorable interaction with the protein backbone [8], an 'osmophobic effect' which could be particularly relevant in a crowded environment, as within the cell.

Among the compatible osmolytes, the amino acid proline features unusual properties in aqueous solution: it is soluble up to ca. 6.5 M, its viscosity behavior is similar to that of larger hydrophilic colloids [9], it improves the solubility of hydrophobic compounds [9–12] and inhibits aggregation during protein refolding [12,13]. Such experimental findings have been interpreted as suggesting that proline self-aggregates in concentrated aqueous solutions, with pyrolidine rings stacking one over the other [9]. This kind of arrangement has been observed in L-proline monohydrate crystals by X-ray analysis at 100 K [14].

The present work has been undertaken to shed some light at molecular level of details on the properties of proline aqueous solutions, which are investigated using molecular dynamics (MD) simulation up to its solubility limit in water.

2. Computational methods

Simulations were performed using the Amber ff99 force field [15] and the SANDER module of AMBER7 package [16]. The temperature was controlled according to the Berendsen coupling algorithm [17] with $\tau_T = 1$ ps; $\tau_P = 1$ ps was used for the pressure control. A 9-Å cutoff and Ewald sums were used for the non-bonded interactions. SHAKE [17] kept constrained bonds involving hydrogen atoms. NPT ($T = 298$ K, $P = 1$ bar) simulations were performed

with 0.5-fs time-step for at least 500 ps. Systems at different solute concentrations were obtained by solvating a given number of proline molecules with 503 TIP3P [18] waters (see Table 1). System configurations were stored on disk every 0.01 ps for statistical analyses based on the PTRAJ and CARNAL modules of the AMBER7 package. Graphical visualization was performed with gOpenMol [19].

3. Results and discussion

As indicated in Table 1, the simulated systems correspond to an ample range of solute concentrations up to proline solubility limit in water.

To provide an overall view about the hydration characteristics of proline, a 3D visualization is presented in Fig. 1

in terms of spatial distribution function (SDF) of oxygen atoms of the water molecules residing within the first hydration shell of proline. This distribution has been evaluated over a 2.5-ns MD trajectory of the system consisting of one proline and 503 TIP3P [18] water molecules, using the following expression [20]:

$$g(\mathbf{r}) = \rho(\mathbf{r})/\rho_0,$$

where $\rho(\mathbf{r})$ indicates the number density of the water oxygen atoms at the position $\mathbf{r}$ and ρ_0 its average within the simulation box. The position vector $\mathbf{r}$ is defined in a coordinate system attached to the solute molecule.

The isosurfaces shown in Fig. 1 correspond to $g(\mathbf{r}) = 2.3$. In the insets of the upper row we report front-, lateral- and rear-views of the water oxygen SDF around proline. For

Table 1
Average number per proline of water molecules within 3.5 Å from the atoms of the solutes (first-shell waters). Some of them are hydrogen-bonded to solute (H-B waters). D_w and D_s (in 10^{-3} nm^2 ps^{-1}) represent the water and solute diffusion coefficients, respectively

Solutes	Solute molar fraction	First-shell waters	H-B[a] waters	D_w	D_s
1	1.98×10^{-3}	19.5	4.2	2.8	1.3
4	7.89×10^{-3}	19.2	4.2	2.7	1.2
8	1.57×10^{-2}	18.7	3.9	2.6	1.0
14	2.71×10^{-2}	18.3	3.9	2.5	1.2
16	3.08×10^{-2}	17.9	3.8	2.4	1.3
22	4.19×10^{-2}	17.8	3.8	2.2	1.2
27	5.09×10^{-2}	17.2	3.8	2.1	1.0
36	6.68×10^{-2}	17.0	3.6	1.9	1.1
43	7.88×10^{-2}	15.9	3.6	1.8	0.90
51	9.21×10^{-2}	15.7	3.3	1.6	0.90
64	1.13×10^{-1}	14.8	3.3	1.5	0.70
72	1.25×10^{-1}	14.5	3.1	1.3	0.60
85	1.45×10^{-1}	13.9	3.1	1.1	0.50
100	1.66×10^{-1}	13.1	2.9	1.0	0.50
125	1.99×10^{-1}	12.1	2.8	0.80	0.40
150	2.30×10^{-1}	11.3	2.7	0.63	0.30

[a] Hydrogen bond definition: distance $r_{OO} \leqslant 3.5$ Å; angle $\theta_{OHO} \geqslant 150°$ [25].
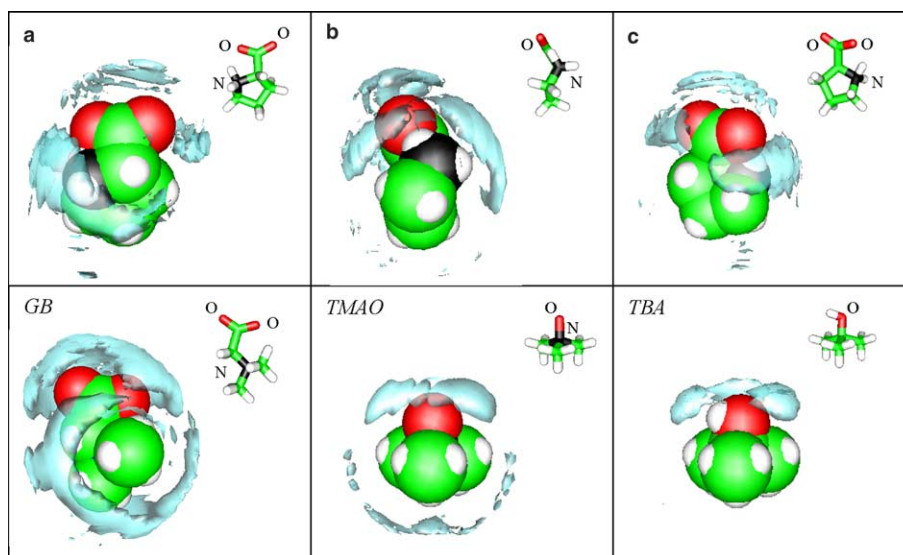

Fig. 1. Spatial distribution functions (SDFs), $g(\mathbf{r})$, of the water oxygen atoms surrounding different solutes. (Upper row) Front (a), lateral (b) and rear (c) views for proline. (Lower row) SDFs for glycine betaine (GB), trimethylamine-*N*-oxide (TMAO) and *tert*-butyl alcohol (TBA). Reference molecular structures are shown on the top right part of each panel. The isosurfaces correspond to $g(\mathbf{r}) = 2.3$.
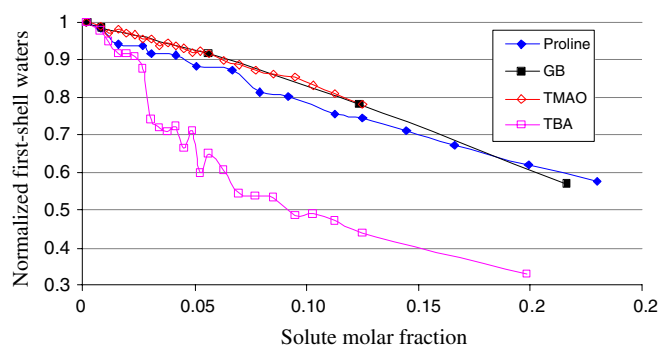
Fig. 2. Normalized number of the first shell water molecules vs. solute molar fraction of proline, GB, TMAO and TBA. Lines are eye-guides.

comparison, in the lower row of Fig. 1, water SDFs are shown around the bioprotectans glycine betaine (GB) and trimethylamine-*N*-oxide (TMAO), and the protein denaturant *tert*-butyl alcohol (TBA), which have been evaluated from previous simulation data [21,22]. As expected, the water oxygen distribution is more relevant in the space regions surrounding the hydrophilic carboxyl, hydroxyl and amino groups. However, while at this SDF level the water distribution extends somewhat towards the hydrophobic moiety in the methylamines GB and TMAO, due to the presence of the nitrogen atom buried among the methyl groups of these molecules, the methyl groups of TBA and the hydrophobic moiety of the pyrolidine ring of proline are left totally 'naked'. For the latter the effect appears as partially compensated by the well hydrated amino group, so that the overall water distribution around proline seems closer to that around GB and TMAO than that surrounding TBA. This interpretation agrees with the average numbers of first-shell waters for one-solute systems (FSWs, i.e.,

waters within 3.5 Å from solute atoms). Indeed, from Table 1 one sees that proline has 19.5 FSWs, while the corresponding values for GB, TMAO and TBA are 20.7, 17.2 and 16.5, respectively [21,22].

The proline hydration as a function of solute molar fraction is shown in Fig. 2, where we report the normalized number of the first-shell waters (i.e., the number of the FSWs for a given solute concentration divided by the corresponding number for the one-solute systems). For comparison, analogous values are presented for aqueous solutions of GB, TMAO and TBA [21,22]. One can see that the general behavior of proline is much more similar to those of GB and TMAO than that of TBA, although the proline hydration seems to decay more slowly with increasing solute concentration.

The proline similarity with GB and TMAO is also pointed out by the analyses of the solute distribution around water at solute molar fraction close to the solubility limit in water. Indeed, under this condition a water molecule is on the average shared by ca. 3 proline molecules, as in the GB case, while the corresponding value for TMAO is 2 and no water sharing has been observed for TBA [21,22]. These data indicate that also for proline a diffuse solute–solvent intermixing persists up to the highest solute concentration, as for GB and TMAO.

A rather different view seems to emerge from Fig. 3, where we report the water oxygen–oxygen radial distribution functions (RDFs), $g_{OO}(r)$, evaluated for one-solute solutions (thin lines) and for the most concentrated solutions (thick lines) of proline, GB, TMAO and TBA. Indeed, the heights of the RDFs first peaks for proline and TBA are similar, being larger than those for GB and TMAO. This indicates that the water structuring induced
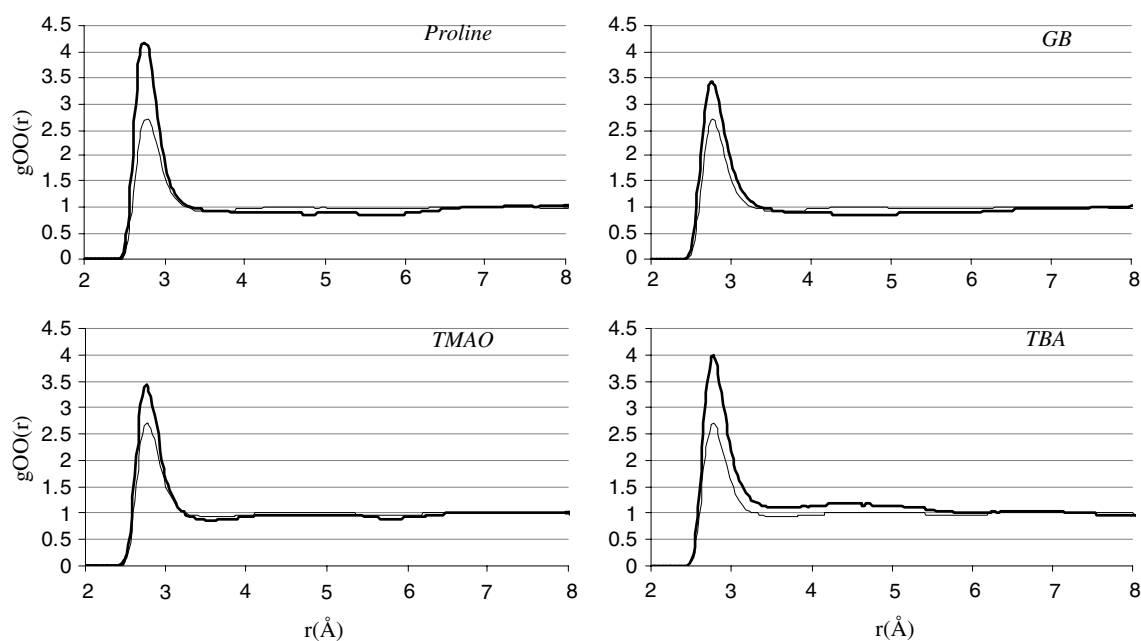

Fig. 3. Radial distribution functions, $g_{OO}(r)$, of the water oxygen atoms for the lowest (thin lines) and highest (thick lines) solute concentrations, which correspond to solute molar fraction of 2.3×10^{-1}, 2.2×10^{-1}, 1.3×10^{-1} and 2.0×10^{-1} for proline, GB, TMAO and TBA, respectively.

at high concentration by proline and by TBA is comparable and more relevant than those caused by the presence of GB and TMAO. Since TBA self-aggregates at such concentrations [23], these results could be interpreted as suggesting that also for proline the water structure enhancement comes from self-aggregation, a mechanism previously hinted to explain experimental findings [9–12], such as anomalously high viscosity of concentrated aqueous solutions of proline. However, this view contrasts with the previous results about the proline hydration vs. solute concentration (Fig. 2) and its propensity to be intermixed with water even at high solute concentrations. This is clearly evidenced by the molecular visualization shown in Fig. 4, where we report a typical configuration of the solution corresponding to the highest proline concentration examined together with analogous data for GB, TMAO and TBA [21,22]. Solute segregation from the solvent is well evident only for TBA, confirming that self-aggregation is not relevant for proline up to solute concentrations close to its solubility limit. Further, direct inspection shows that solute and solvent molecules appear as extensively connected via hydrogen bonds (HBs) for proline, GB and TMAO. From Table 1 one can see that, at the highest concentration examined, the average number of water molecules hydrogen-bonded to one molecule of proline is 2.7. The corresponding values for GB, TMAO and TBA are 2.5, 2.2 and 1.3, respectively [21,22]. Furthermore, solute–solute HBs can occur for proline and TBA, due to the presence on these molecules of acceptor and donor groups. The average numbers of solute–solute HBs per solute at the highest concentrations examined are 0.4 and 0.2 for proline and TBA, respectively. This extra hydrogen bonding capability can cause the apparent higher HB connectivity of proline solutions with respect to the GB and TMAO cases. It could also bring about the tighter water localization evidenced by the RDFs shown in Fig. 3.

In order to examine the translational dynamics of the solute and water molecules as a function of the solute concentration we evaluated their diffusion coefficient, D_s and D_w, respectively, according to the Einstein formula [17]. In Fig. 5, we report the ratio of the diffusion coefficient values for one-proline solution, D_{s0} and D_{w0}, to the corresponding values at a given solute molar fraction. Indeed,

Fig. 4. Statistically representative snapshots of the most concentrated solutions. Lighter molecules represent water.

one would expect that this quantity is related to the viscosity properties [24] of proline solutions. For comparison, analogous values are reported for GB, TMAO and TBA, based on previous simulation data [21,22]. Data of Fig. 5 show that as the solute concentration increases, the translational dynamics of proline solutes increasingly diverges from that of the other solutes examined. Further, the diffusion coefficients of both the solvent and solutes molecules show the same behavior although the statistical error is larger for the latter since the number of the solutes molecules is lower than that of the water molecules.

4. Conclusions

Statistical analyses of MD simulations of proline aqueous solutions at solute molar fraction ranging from
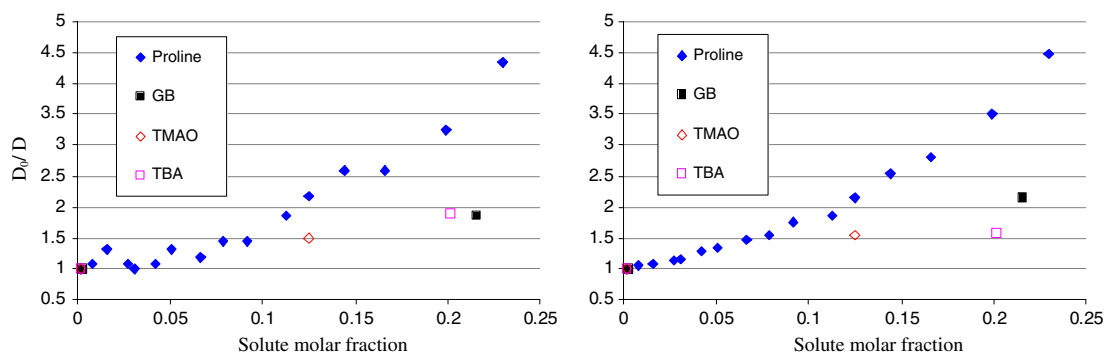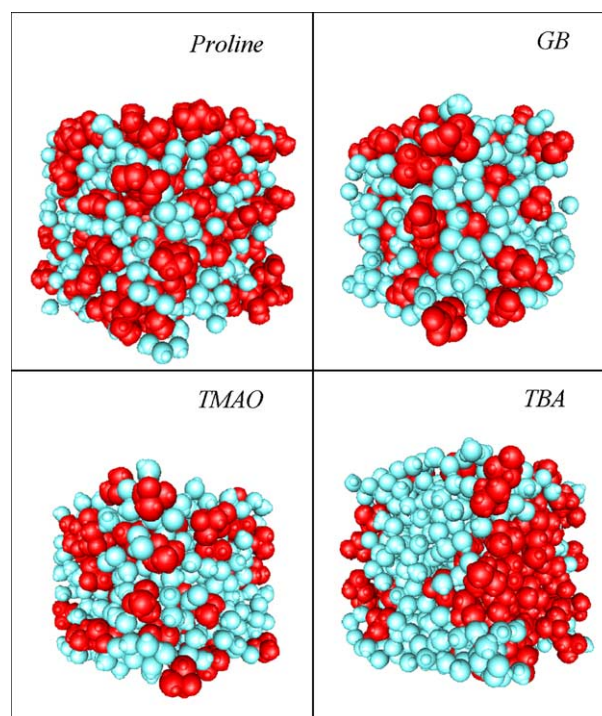

Fig. 5. Diffusion coefficient ratio between the value for one-solute solution (D_0) and the value at a given solute concentration (D) vs. solute molar fraction. (Left) D_{s0}/D_s for solute molecules; (right) D_{w0}/D_w for water molecules.

2.0×10^{-3} to 2.3×10^{-1} indicate that proline induces a strong localization of the surrounding water, comparable to that caused by the self-aggregating protein denaturant *tert*-butyl alcohol (TBA), and more pronounced than that due to the related osmoprotectants glycine betaine (GB) and trimethylamine-*N*-oxide (TMAO). Analyses show that proline, like GB and TMAO, is well intermixed with water up to the highest concentration examined forming extensive HB network whose connectivity is enhanced in the proline case by its capability to make solute–solute HBs. This view can explain the strong water structuring induced by proline and the behavior of the diffusion coefficients of both solute and water molecules vs. proline concentration which is in qualitative agreement with the anomalous viscosity properties of proline solutions evidenced experimentally [9]. At no concentration, however, molecular visualization of system snapshots has shown any significant presence of the proline self-aggregates hinted to explain the experimental findings [9–12].

Acknowledgments

References

[1] P.H. Yancey, M.E. Clark, S.C. Hand, R.D. Bowlus, G.N. Somero, Science 217 (1982) 1214.

[2] D.W. Bolen, Methods 34 (2004) 312.

[3] A.J. Wang, D.W. Bolen, Biochemistry 36 (1997) 9101.

[4] K. Goeller, E.A. Galinski, J. Mol. Catal. B 7 (1999) 37.

[5] S.A. Celinski, J.M. Scholtz, Protein Sci. 11 (2002) 2048.

[6] M.K. Chattopadhyay, R. Kern, M.Y. Mistou, A.M. Dandekar, S.L. Urastu, G. Richarme, J. Bacteriol. 186 (2004) 8149.

[7] T. Arakawa, S.N. Timasheff, Biophys. J. 47 (1985) 411.

[8] D.W. Bolen, I.V. Baskakov, J. Mol. Biol. 310 (2001) 955.

[9] B. Schobert, H. Tschesche, Biochim. Biophys. Acta 541 (1978) 270.

[10] V. Srinivas, D. Balasubramanian, Langmuir 11 (1995) 2830.

[11] D. Samuel, T.K.S. Kumar, G. Jayaraman, P.W. Yang, C. Yu, Biochem. Mol. Biol. Int. 41 (1997) 235.

[12] D. Samuel, T.K.S. Kumar, G. Ganesh, G. Jayaraman, P.W. Yang, M.M. Chang, V.D. Trivedi, S.L. Wang, K.C. Hwang, D.K. Chang, C. Yu, Protein Sci. 9 (2000) 344.

[13] O.P. Chilson, A.E. Chilson, Eur. J. Biochem. 270 (2003) 4823.

[14] J. Janczak, P. Luger, Acta Crystallogr. C 53 (1997) 1954.

[15] J. Wang, P. Cieplak, P.A. Kollman, J. Comput. Chem. 21 (2000) 1049.

[16] D.A. Case et al., Amber7, University of California, San Francisco, 2002.

[17] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon Press, Oxford, 1987.

[18] W.L. Jorgensen, J. Chandrasekhar, D. Madura, R.W. Impey, M.L. Kein, J. Chem. Phys. 79 (1983) 926.

[19] D.L. Bergman, L. Laaksonen, A. Laaksonen, J. Mol. Graph. Model. 15 (1997) 301.

[20] I.M. Svishchev, P.G. Kusalik, J. Chem. Phys. 99 (1993) 3049.

[21] M. Civera, A. Fornili, M. Sironi, S.L. Fornili, Chem. Phys. Lett. 367 (2003) 238.

[22] A. Fornili, M. Civera, M. Sironi, S.L. Fornili, Phys. Chem. Chem. Phys. 5 (2003) 4905.

[23] A. Di Michele, M. Freda, G. Onori, A. Santucci, J. Phys. Chem. A 108 (2004) 6145.

[24] K. Rah, B.C. Eu, Phys. Rev. E 60 (1999) 4105.

[25] H. Xu, B.J. Berne, J. Phys. Chem. B 105 (2001) 386.