A Simplified Model of Local Structure in Aqueous Proline Amino Acid Revealed by First-Principles Molecular Dynamics Simulations

Raphael Z. Troitzsch,* Paul R. Tulip,* Jason Crain,*† and Glenn J. Martyna‡
*School of Physics, The University of Edinburgh, Edinburgh, United Kingdom; †National Physical Laboratory, Teddington, United Kingdom; and ‡IBM T. J. Watson Research Center, Yorktown Heights, New York

ABSTRACT Aqueous proline solutions are deceptively simple as they can take on complex roles such as protein chaperones, cryoprotectants, and hydrotropic agents in biological processes. Here, a molecular level picture of proline/water mixtures is developed. Car-Parrinello ab initio molecular dynamics (CPAIMD) simulations of aqueous proline amino acid at the B-LYP level of theory, performed using IBM's Blue Gene/L supercomputer and massively parallel software, reveal hydrogen-bonding propensities that are at odds with the predictions of the CHARMM22 empirical force field but are in better agreement with results of recent neutron diffraction experiments. In general, the CPAIMD (B-LYP) simulations predict a simplified structural model of proline/water mixtures consisting of fewer distinct local motifs. Comparisons of simulation results to experiment are made by direct evaluation of the neutron static structure factor $S(Q)$ from CPAIMD (B-LYP) trajectories as well as to the results of the empirical potential structure refinement reverse Monte Carlo procedure applied to the neutron data.

INTRODUCTION

Aqueous proline amino acid has long been known to possess unusual physical, colligative, and transport properties such as viscosities that are anomalously high for such a low molecular weight solute (1–3). More importantly, due to aqueous proline's unique properties, proline-water mixtures, which show surprisingly large proline solubility, demonstrate rudimentary biochemical functionality. For example, aqueous proline exhibits hydrotropism—the property whereby increased proline content increases the solubility in water of hydrophobic compounds (4). Due to this behavior, it has been suggested that aqueous proline may act as a simple chemical chaperone in protein folding (5). Aqueous proline is also considered to be a natural bioprotectant expressed by plants and other organisms (such as bacteria and protozoa) under adverse conditions such as low temperature stress (6–11).

There has been extensive speculation in the literature as to the molecular origin of the unusually diverse properties of aqueous proline and, in particular, how they are connected to proline's hydration structure and hydrogen-bonding propensities. One prevailing view has been that ordered intermediate range aggregates are formed in solution and that these lead to the peculiar properties of the proline-water mixtures (1). While there has been some indirect evidence for the presence of aggregates from spectroscopic, light scattering, and calorimetric measurements (3,4,12), very recent neutron diffraction obtained at intermediate and low Q appear to rule out persistent aggregates having a well-defined characteristic length scale, implying that only local correlations are relevant for an understanding of the physical properties (13). Extensive computer simulations based on empirical potential models such as CHARMM22 (14) with sampling efficiency improved by parallel tempering have also ruled out the presence of well-ordered mesoscale structures, but do provide evidence for more diffuse heterogeneities arising from concentration fluctuations which nicely explains the light scattering data. The empirical model simulations also suggest that proline suppresses the normal temperature dependence of structure of water and preserves ambient correlations even in very cold solutions (15,16) by frustrating the formation of ice crystals via competitive hydrogen-bonding of proline to water. This latter insight provides an explanation of the prolines' colligative and cryoprotective properties.

Although the empirical model simulation studies have contributed to the understanding of proline solutions, significant open questions remain. The structure of the mixtures predicted by various rigid, nonpolarizable (fixed charge) water potentials (SPC/E, TIP3P, and TIP4P) in conjunction with CHARMM22 was compared to neutron diffraction data through direct evaluation of the static structure factor $S(Q)$ from the MD trajectories. This approach deals directly with the experimental observable(s) rather than the site-site distributions $g_{ij}(r)$, which are not directly measured and can only be inferred through reverse Monte Carlo procedures such as the empirical potential structure refinement method (17). Some features of $S(Q)$ were found to be well described by all three empirical potential models whereas others were poorly represented by all. Ultimately, no empirical model has provided a clearly superior description of the experimental data and none could be ruled out (16). It is clear that all three empirical models are limited in the sense that all are rigid and neglect polarization effects that may be significant in this system. Therefore, a molecular level understanding of proline-water solutions is yet lacking; a rather surprising conclusion.

Submitted April 8, 2008, and accepted for publication July 21, 2008.

Address reprint requests to Dr. Paul Robert Tulip, Tel.: 44-01-31-650-5273; E-mail: ptulip@ph.ed.ac.uk.

Editor: Steven D. Schwartz.

0006-3495/08/12/5014/07 $2.00

doi: 10.1529/biophysj.108.134916

To obtain a more complete understanding of the molecular structure of aqueous proline solutions, we have performed first-principles molecular dynamics simulations (18) on IBM's Blue Gene/L supercomputer at the B-LYP level of theory enabled by massively parallel software (19,20). Comparing these new results to the previously reported empirical models and to experimental diffraction data, it is possible to determine the structural motifs that underlie proline's remarkable properties and to assess the importance of the physics that drive them. In particular, the influence of flexibility and polarizability on the structural properties of this important system can be discerned and are described herein.

METHODS

Aqueous proline was modeled from first-principles within the following approximations. The gradient-corrected Becke-Lee-Yang-Parr (21,22) (B-LYP) approximate density functional and a plane-wave basis set (B-LYP/pw) were used in conjunction with a 70 Ry energy cutoff and norm-conserving pseudopotentials (23) at the Γ-point. Car-Parrinello ab initio molecular dynamics (18) (CPAIMD) were performed in the canonical ensemble using Nosé-Hoover chain thermostats (24–28) and a 0.125 fs time-step on IBM's Blue Gene/L supercomputer using new massively parallel software (19,20).

A 38:2 water/proline mixture corresponding to a 2.75 M solution was examined. The system was first equilibrated under the CHARMM22 force field (C22) (14) in conjunction with the rigid, nonpolarizable TIP4P (29) water model via a 3-ns parallel tempering simulation. Furthermore, the empirical model results were used to allow a three-way comparison between previously published empirical model results (16), which consisted of considerably larger systems (6860:343 water/proline molecules), the finite size empirical model results, and the CPAIMD (B-LYP) ab initio calculations. Four starting configurations for a series of four CPAIMD (B-LYP) simulations were picked at random from the $T = 300$ K walker of the empirical model parallel tempering computation. In each case, the systems was equilibrated under CPAIMD (B-LYP) dynamics for 10 ps. Data were subsequently collected for all initial conditions for 200 ps. In this way, a total of 800 ps of sampling was performed for the 38:2 system under B-LYP.

RESULTS

In Fig. 1, snapshots from CPAIMD (B-LYP) trajectories, which depict selected hydrogen-bonded motifs as well as the arrangement of water around the apolar groups, are presented. Also shown is the labeling scheme for the proline atoms that we will refer to in later discussion of the structure.

Direct comparison to experiment: Validation of CPAIMD (B-LYP) approach

In Fig. 2, *a–c*, we give the static structure factor $S(Q)$ obtained by CPAIMD-B-LYP simulation, by classical MD simulation using two empirical water potentials (TIP3P and TIP4P) and experimental neutron diffraction on isotopically substituted samples. We emphasize that the experimental structure factor is obtained directly from the measured diffraction intensity once corrections for background, multiple scattering, and nuclear recoil are applied, as described in McLain et al. (13). For all simulated systems, $S(Q)$ is generated by direct evaluation from the simulation trajectories and not via Fourier transform of the corresponding total pair correlation function. We thereby avoid truncation errors in $S(Q)$ as discussed in detail in Troitzsch et al. (16). The method used is summarized in the Supplementary Material in Data S1. We have incorporated nuclear quantum effects using the same procedure as described in Troitzsch et al. (16); as discussed therein, compounds containing light atoms exhibit nuclear quantum effects that impact upon structural properties, and hence upon the structure factor.

Three isotopically distinct systems are considered: fully deuterated proline in D_2O, partially deuterated proline (*D* substitution only on amide sites in D_2O), and fully hydrogenous proline in H_2O. For fully hydrogenous proline, we present results for $Q > 2$ Å^{-1} because, as discussed in McLain et al. (13), in the experimental results, the background and inelasticity corrections are most difficult to remove in the low

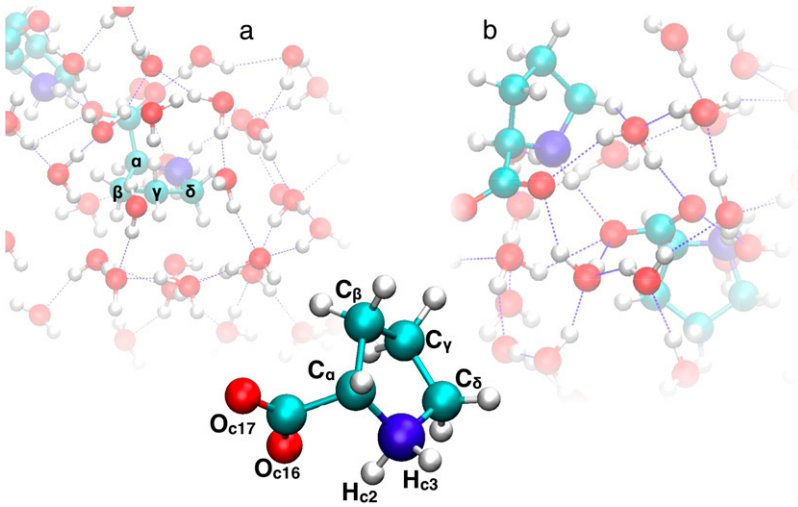

FIGURE 1 Snapshots of the molecular trajectory showing selected hydrogen-bonding motifs, as well as the structure of water around the apolar groups. Further, the labeling scheme for the proline atoms is shown.

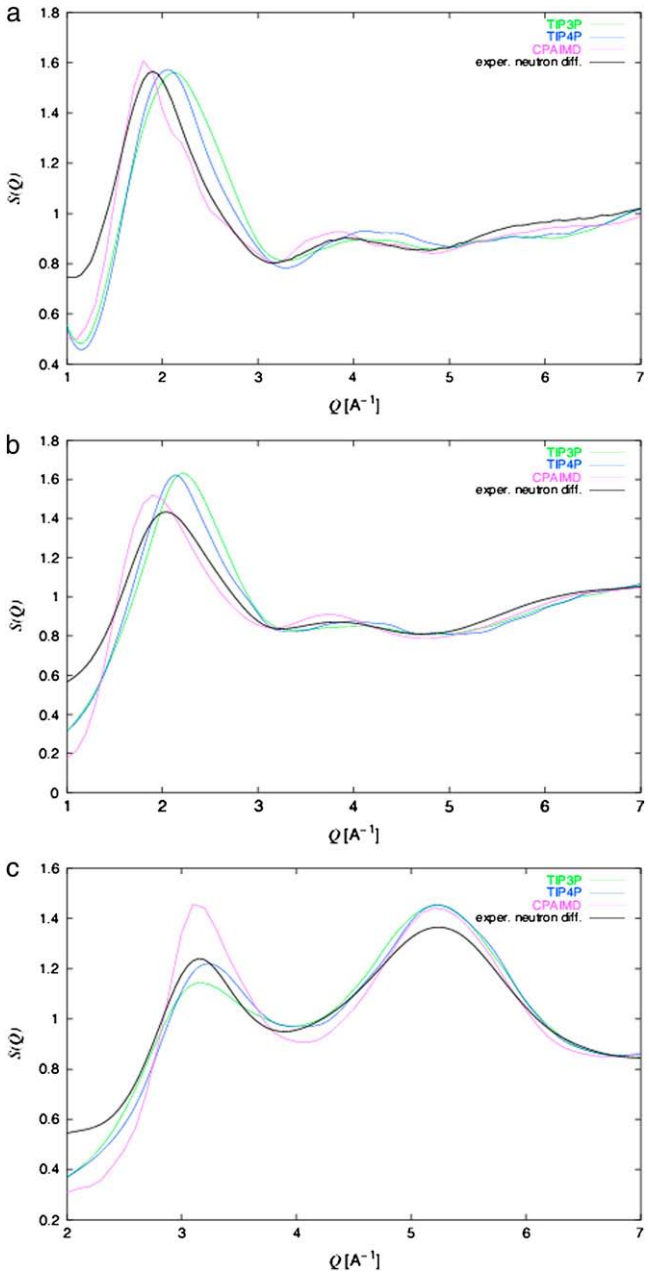

FIGURE 2 Static structure factor, $S(Q)$, as obtained from the trajectories for classical and ab initio calculations, as well as from neutron experiments (13). Panels show (*a*) the system fully deuterated, apart from the proline backbone, (*b*) fully deuterated and (*c*) fully hydrogenous system.

Q (i.e., $Q < 2$ Å^{-1}) region for compounds containing hydrogen. Since backbone hydrogens are nonexchangeable, we can neglect H/D exchange in these systems. In Figs. S1 and S2 in Data S1, we demonstrate, through explicit calculation, that finite size effects are not significant, and therefore do not alter the conclusions drawn from the $S(Q)$ calculations.

Discrepancies between the empirical potential model and experiment are largest for the fully deuterated system (Fig. 2 *b*) where the first peak height and position are both significantly overestimated. By contrast, $S(Q)$ obtained from the CPAIMD trajectories is in better agreement with the measured data for this isotopic composition. Both the peak position and intensity are much improved relative to the empirical models. We will explore the structural dif-

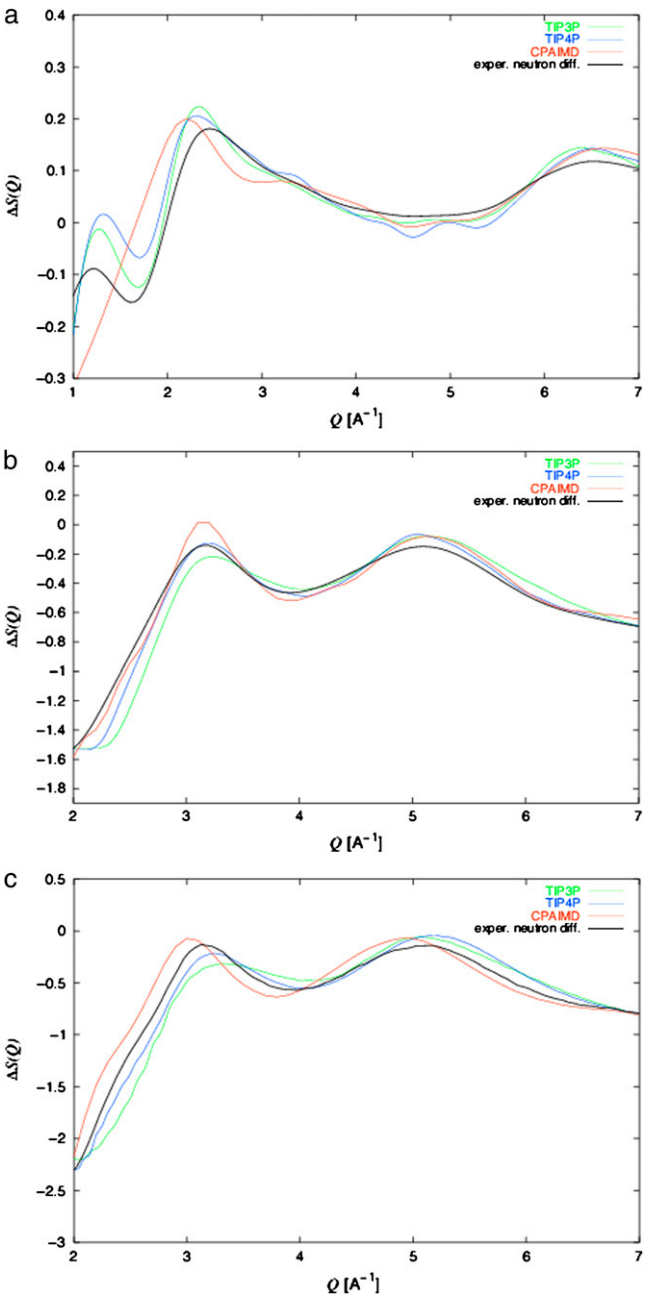

FIGURE 3 Comparison of differences in the static structure factors given in Fig. 2, as obtained from classical and ab initio calculations, as well as from neutron experiments (13). (*a*) Difference between the fully deuterated and the partially deuterated cases (i.e., between panels *b* and *a* in Fig. 2). (*b*) Difference between the fully hydrogenous and the partially deuterated cases (i.e., between panels *c* and *a* in Fig. 2). (*c*) Difference in static structure factor for water and heavy water.

ferences between the empirical models that lead to these results after consideration of the other two isotopic compositions. In the case of partially deuterated proline in D_2O (Fig. 2 *a*), agreement between empirical potentials and experiment is better, though the first peak position is still overestimated. The CPAIMD-B-LYP model, although slightly overestimating the peak intensity, provides a better description of the peak position and shape in comparison with the experimental data. We should note here that the B-LYP model provides an inadequate description of the dispersion interaction and hence treats hydrophobic interactions less accurately; this particularly manifests itself in the partially deuterated system. A correction scheme to this problem is outlined in Data S1. Finally, in the fully hydrogenous case (Fig. 2 *c*), none of the simulation models accounts well for the measured intensity near $Q \approx 5.2$ and the TIP4P model system appears to reproduce the first peak better than the CPAIMD-B-LYP model. It is likely that the failure of the CPAIMD model owes its origin to a less accurate description of the water structure.

To illustrate the contributions from the backbone hydrogens to the structure factor, Fig. 3 *a* displays the difference between the structure factors for the fully deuterated and the partially deuterated systems (i.e., between panels *b* and *a* in Fig. 2), while in Fig. 3 *b*, we do likewise for the amine and water hydrogens by illustrating the difference between the fully hydrogenous and the partially deuterated system (i.e., between panels *c* and *a* in Fig. 2). Fig. 3 *a* indicates that the classical models appear to provide a better description than the CPAIMD model. This is probably because this quantity emphasizes the contributions of the hydrophobic interactions to the static structure factor, and as discussed above, the CPAIMD model describes the dispersion interactions underlying the hydrophobic interactions inadequately, a point that manifests itself again when we discuss the structural motifs present. In Fig. 3 *b*, we find that the classical models appear to perform better than the CPAIMD model. This is

TABLE 1 Coordination numbers for bifurcated and conventional H-bonds in proline-proline and proline-water contacts at the carboxyl group

	TIP4P	SPC/E	CPAIMD	EPSR (13)	*a*	*b*
H_{C2} bifurcated	0.01	0.01	0.00		2.4	2.1
H_{C2}	0.03	0.03	0.06		2.4	2.1
H_{C3} bifurcated	0.06	0.01	0.01		2.4	2.1
H_{C3}	0.05	0.04	0.11		2.4	2.1
H_W bifurcated	0.42	0.50	0.03		2.4	2.1
H_W	2.29	2.28	1.53		2.4	2.1
Total H_W	2.71	2.78	1.56	1.67	2.4	
Total bifurcated	0.49	0.52	0.04			
Total nonbifurcated	2.37	2.35	1.80			
Total	2.86	2.86	1.84	1.96		

Figs. 4 and 5 show the corresponding distributions. The column headers *a* and *b* (in Å) delimit the symmetric areas of integration as defined in Troitzsch et al. (16).

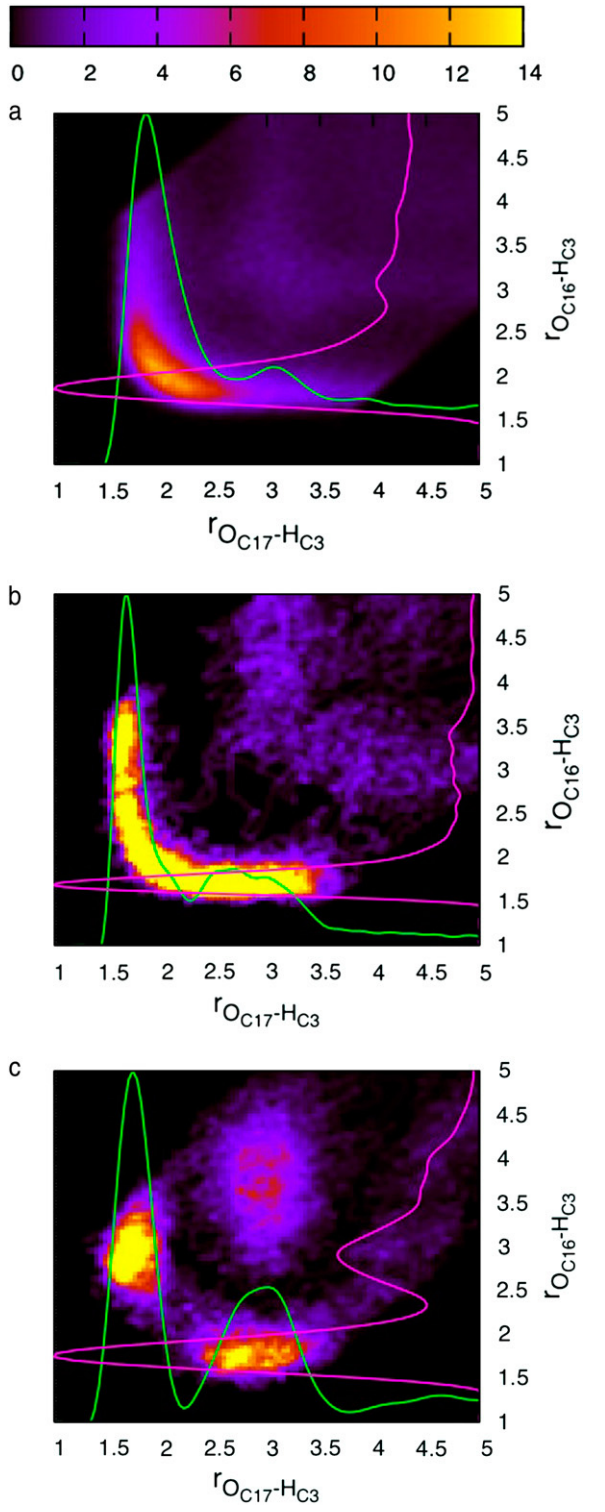

FIGURE 4 Two-dimensional radial distribution indicating the distances of the carboxyl group oxygens to one (distinct) hydrogen on the amide group. The classical simulation in panel *a* shows a strong propensity for bifurcation that is present in the finite time and size run in panel *b*, but wholly absent in the CPAIMD result shown in panel *c*.

due to the less accurate description of the fully hydrogenous system, as noted above. It should be noted that it is not possible to correct for the effects of water structure upon the CPAIMD results, as the small size and short timescale of CPAIMD leads to poor statistics, and hence large errors when undertaking such a calculation. In Fig. 3 *c*, we present the difference in water and heavy water $S(Q)$s with appropriate weighting. It can be seen that the TIP4P model is in best agreement with the experimental results; the CPAIMD model underestimates the location of the first peak, while also overestimating its intensity.

Structural motifs

We now consider the structural differences contributing to the $S(Q)$ variations. First we address the question of hydration of the proline molecules by reporting the coordination number of water around the carboxyl oxygen atoms. For the CPAIMD (B-LYP) system the average carboxyl oxygen-to-water hydrogen coordination number is $\sim$1.6 whereas it is much larger ($\approx$2.8) in the empirical model with little variation observed between SPC/E and TIP4P potentials. Empirical potential structure refinement of the neutron diffraction data implies a coordination number of $\approx$1.7. The CPAIMD (B-LYP) result is in better agreement with this value and is in fact marginally underhydrated relative to the empirical potential structure refinement (EPSR) result (see Table 1). We will return to the structural origin of the discrepancy between the empirical model and CPAIMD (B-LYP) results below.

The empirical models show two prominent structural features that appear to be absent from the CPAIMD(B-LYP) simulation. The first motif predicted by the empirical models is the bifurcated hydrogen bond in which a water or amide hydrogen acts as a proton donor to two carboxyl oxygens simultaneously (16). These bifurcated H-bonds are evident in the joint distribution function, $g(r_{O_{C17}-H_{C3}}; r_{O_{C16}-H_{C3}})$, which measures the joint probability of finding the amide proton H_{C3} at a given distance from the carboxyl oxygens (O_{C17} and O_{C16}) where the labeling is as defined in Fig. 1. For the empirical model system, Fig. 4 *a*, we see that there is only one preferred inter-proline dimer contact where the amide proton is equidistant from the carboxyl oxygens. The H-bond therefore shows a strong tendency to bifurcate. The results for a smaller system size and shorter time empirical model simulation are shown in Fig. 4 *b*. It is clear that finite size and finite sampling time do not lead to qualitative differences in the empirical model simulation results and the choice of system size and timescale for the CPAIMD (B-LYP) study can be justified. We discuss this point in more detail in Data S1. The corresponding joint distribution for the CPAIMD (B-LYP) system (which can be compared directly to the empirical model in Fig. 4 *b*) is shown in Fig. 4 *c*. Here the signature for H-bond bifurcation is absent and there is a distinct preference for a single, somewhat shorter H-bond; two distinct contact distances are present.

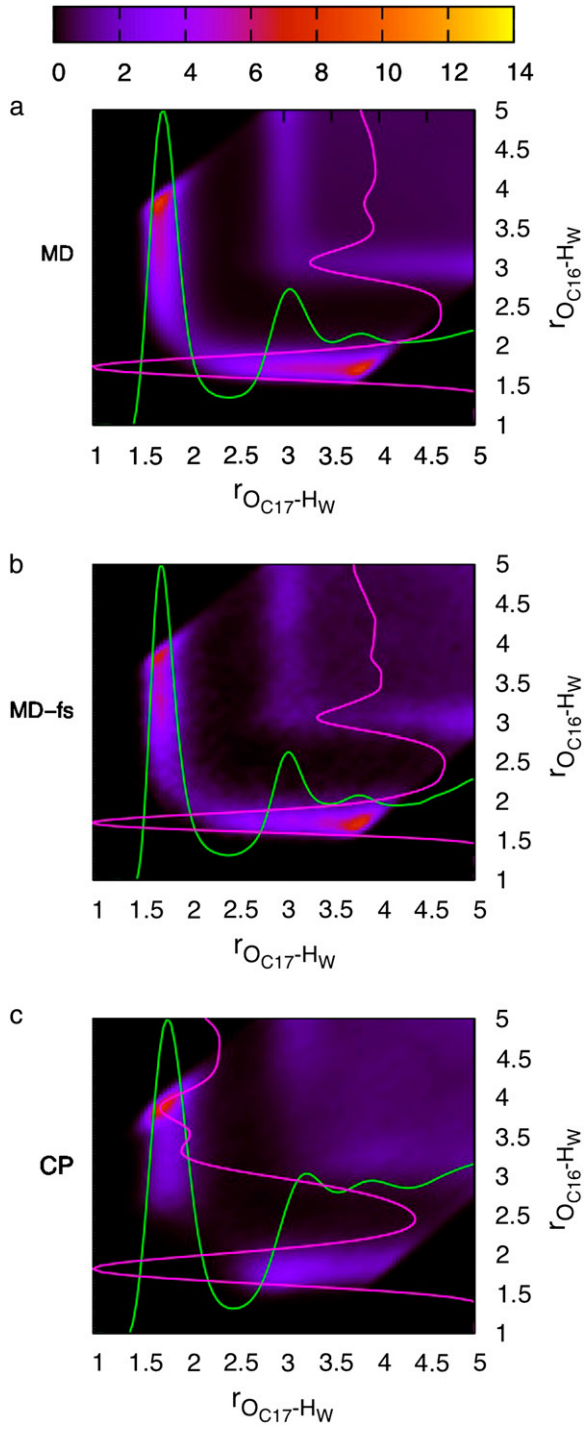

FIGURE 5 Analogous to Fig. 4: two-dimensional radial distribution indicating the distances of the carboxyl group oxygens to water hydrogen. Again, classical results (*a*) shows bifurcation occurs, even in the finite time and size run (*b*), but is absent in the CPAIMD result (*c*).

In the case of proline-water hydrogen bonding, we also see evidence of similar H-bond bifurcation in water-proline hydration structure of the empirical model. Two-dimensional distributions analogous to those for proline-proline dimer

contacts are shown in Fig. 5 *a* for the empirical model in which the amide proton position is replaced by the water proton. Diffuse intensity is observed over a wide range of distances with appreciable contribution arising at the point where the water proton is equidistant from the carboxyl oxygen and thus participating in a bifurcated H-bond. The presence of this bifurcated H-bond may be responsible for the overhydration of proline observed in the empirical model relative to CPAIMD (B-LYP) and EPSR-generated ensembles based on neutron diffraction data, and is likely due to the neglect of many-body polarization in the CHARM22 empirical model calculations. The suppression of this bifurcated motif in the CPAIMD (B-LYP) system is clearly seen in Fig. 5 *c*, where no intensity is observed at the equidistant position corresponding to an H-bond contact.

The second motif of interest predicted by empirical model simulations is proline dimers formed by apolar (hydrophobic) associations that contribute a distinct close contact between carbon atoms of the ring. The latter is illustrated in the two-dimensional joint distribution functions between ring carbons shown in Fig. 6 *a* for the empirical model. Specifically, we depict the distribution function, $g(r_{C_\gamma-C_\beta}; r_{C_\gamma-C_\delta})$. Under the empirical model, it is clear that contacts between these atoms exhibit very close approaches implying that proline associates via apolar contacts as expected for hydrophobic systems. In Fig. 6 *b*, the results of an empirical model simulation performed in a small system for short times is given to explore finite size and finite timescale effects on the results as above. Again, the choice of system size and timescale for the CPAIMD (B-LYP) study is thereby justified, as discussed in more detail in the Data S1. The same two-dimensional joint distribution for the CPAIMD (B-LYP) system is shown in Fig. 6 *c*. Under CPAIMD (B-LYP), we find that the only significant features arise from hydrogen-bonded dimers, a result that cannot be explained by finite system size and timescale effects. Note, however, that the B-LYP functional does not provide an adequate description of dispersion (30,31), which may be the overriding factor in the absence of hydrophobic association in the CPAIMD (B-LYP) study. Comparison of the radial distribution of the backbone carbons between CPAIMD (B-LYP) and EPSR (13) (not shown) supports the suggestion that the ab initio result is undercoordinated in the backbone association.

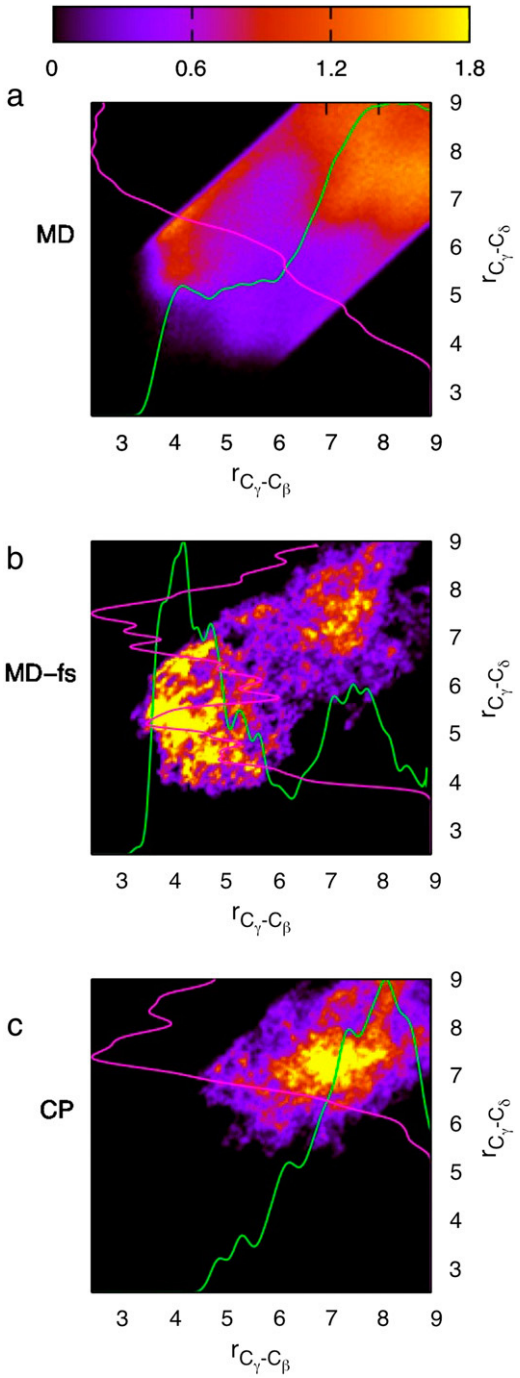

FIGURE 6 Two-dimensional radial distribution functions of the backbone (hydrophobic) contact for (*a*) classical, (*b*) classical finite size and time, and (*c*) CPAIMD. It is clear that the close contact between the backbones exists in the classical system despite short runtime and finite box size.

SUMMARY AND CONCLUSION

In summary, first-principles molecular dynamics simulations of aqueous proline amino acid reveal considerable differences in local structure, which effect the description of hydration and dimer formation via hydrogen bonding. Specifically, in the CPAIMD system, carboxyl group hydration is diminished; dimerization via hydrophobic contacts is suppressed; and hydrogen-bond bifurcation is virtually eliminated relative to the empirical model. Overhydration of the carboxyl groups by nonpolarizable force fields is likely an important source of error in protein simulations. Similarly, the overstabilization of bifurcated H-bonds seems likely to bias empirical model simulations in unphysical ways. Direct evaluation of the structure factor from classical and CPAIMD trajectories generally

reveals that the CPAIMD structures are in better agreement with experimental neutron diffraction data on isotopically labeled samples and that some significant discrepancies between classical MD and measured $S(Q)$ are resolved in the CPAIMD simulations. The local structural motifs that emerge from the B-LYP CPAIMD simulations therefore appear not only to be simpler than those found observed in the empirical models computations, but more importantly, also represent a better comparison to the experimental data.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

J.C. acknowledges support from Scottish Enterprise and the Edinburgh Parallel Computing Center. G.J.M., R.Z.T., and P.R.T. thank IBM research for support and computer time, as well as E. Bohm and L. Kalé (both at University of Illinois at Urbana-Champaign) and S. Kumar (32) and R. Walkup (both at IBM).

REFERENCES

1. Schobert, B., and H. Tschesche. 1978. Unusual solution properties of proline and its interaction with proteins. *Biochim. Biophys. Acta.* 541:270–277.
2. Chilson, O. P., and A. E. Chilson. 2003. Perturbation of folding and reassociation of lactate dehydrogenase by proline and trimethylamine oxide. *Eur. J. Biochem.* 270:4823–4834.
3. Schobert, B. 1977. Anomalous colligative properties of proline. *Naturwissenschaften.* 64:386–387.
4. Rudolph, A. S., and J. H. Crowe. 1986. A calorimetric and infrared spectroscopic study of the stabilizing solute proline. *Biophys. J.* 50: 423–430.
5. Chattopadhyay, M. K. 2004. The chemical chaperone proline relieves the thermosensitivity of a DnaK deletion mutant at 42°C. *J. Bacteriol.* 186:8149–8152.
6. Morita, Y., S. Nakamori, and H. Takagi. 2003. L-proline accumulation and freeze tolerance in *Saccharomyces cerevisiae* are caused by a mutation in the PRO1 gene encoding γ-glutamyl kinase. *Appl. Environ. Microbiol.* 69:212–219.
7. Paquin, R., and G. Pelletier. 1981. Cold-acclimation of alfalfa (*Medicago-media* Pers.) under field conditions. 1. Variations of the free proline content of leaves and crowns. *Physiol. Veg.* 19:103–117.
8. Koster, K. L., and D. V. Lynch. 1992. Solute accumulation and compartmentation during the cold-acclimation of Puma rye. *Plant Physiol.* 98:108–113.
9. Hare, P., W. Cress, and J. van Staden. 1999. Proline synthesis and degradation: a model system for elucidating stress-related signal transduction. *J. Exp. Bot.* 50:413–434.
10. Nanjo, T., Y. Kobayashi, Y. Yoshiba, K. Kakubari, K. Yamaguchi-Shinozaki, and K. Shinozaki. 1999. Antisense suppression of proline degradation improves tolerance to freezing and salinity in *Arabidopsis thaliana*. *FEBS Lett.* 461:205–210.
11. Yoshiba, Y., T. Kiyosue, K. Nakashima, K. Yamaguchi-Shinozaki, and K. Shinozaki. 1997. Regulation of levels of proline as an osmolyte in plants under water stress. *Plant Cell Physiol.* 38:1095–1102.
12. Samuel, D., 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, and C. Yu. 2000. Proline inhibits aggregation during protein refolding. *Protein Sci.* 9:1604.
13. McLain,S. E., A. K. Soper, A. E. Terry, and A. Watts. 2007. Structure and hydration of L-proline in aqueous solutions. *J. Phys. Chem. B.* 111:4568–4580.
14. MacKerell, A., D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
15. Troitzsch R. Z., H. Vass, W. Hossack, G. J. Martyna, and J. Crain. 2008. Molecular mechanisms of cryoprotection in aqueous proline: light scattering and molecular dynamics simulations. *J. Phys. Chem. B.* 112:4290–4297.
16. Troitzsch, R., G. Martyna, S. E. McLain, A. K. Soper, and J. Crain. 2007. Structure of aqueous proline via parallel tempering molecular dynamics and neutron diffraction. *J. Phys. Chem. B.* 111:8210–8222.
17. Soper, A. K. 2001. Tests of the empirical potential structure refinement method and a new method of application to neutron diffraction data on water. *Mol. Phys.* 99:1503–1516.
18. Car, R., and M. Parrinello. 1985. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* 55:2471–2474.
19. Bohm, E., A. Bhatele, L. V. Kale, M. E. Tuckerman, S. Kumar, J. A. Gunnells, and G. J. Martyna. 2007. Fine-grained parallelization of the Car-Parrinello ab initio molecular dynamics method on the IBM Blue Gene/L supercomputer. *IBM J. Res. Dev.* 52:159–175.
20. Vadali, R., Y. Shi, S. Kumar, L. V. Kale, M. E. Tuckerman, and G. J. Martyna. 2004. Scalable fine-grained parallelization of plane-wave-based ab initio molecular dynamics for large supercomputers. *J. Comp. Chem.* 25:2006.
21. Becke, A. D. 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A.* 38:3098–3100.
22. Lee, C. T., W. T. Yang, and R. G. Parr. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional the electron-density. *Phys. Rev. B.* 37:785–789.
23. Troullier, N., and J. L. Martins. 1991. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B.* 43:1993.
24. Martyna, G. J., M. Tuckerman, D. J. Tobias, and M. L. Klein. 1996. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* 87:1117–1157.
25. Martyna, G. J., M. L. Klein, and M. Tuckerman. 1992. Nosé-Hoover chains—the canonical ensemble via continuous dynamics. *J. Chem. Phys.* 97:2635–2643.
26. Tuckerman, M. E., and M. Parrinello. 1994. Integrating the Car-Parrinello equations. 1. Basic integration techniques. *J. Chem. Phys.* 101:1302–1315.
27. Tuckerman, M. E., and M. Parrinello. 1994. Integrating the Car-Parrinello equations. 2. Multiple time-scale techniques. *J. Chem. Phys.* 101:1316–1329.
28. Hutter, J., M. Tuckerman, and M. Parrinello. 1995. Integrating the Car-Parrinello equations. 3. Techniques for ultrasoft pseudopotentials. *J. Chem. Phys.* 102:859–871.
29. Hernandes, M. Z., J. B. P. da Silva, and R. L. Longo. 2003. Chemometric study of liquid water simulations. I. The parameters of the TIP4P model potential. *J. Comp. Chem.* 24:973–981.
30. Kohn, W., Y. Meir, and D. Makarov. 1998. van der Waals energies in density functional theory. *Phys. Rev. Lett.* 80:4153–4156.
31. Zimmerli, U., M. Parrinello, and P. Koumoutsakos. 2004. Dispersion corrections to density functionals for water aromatic interactions. *J. Chem. Phys.* 120:2693–2699.
32. Kumar, S., C. Huang, G. Zheng, E. Bohm, A. Bhatele, J. C. Phillips, H. Yu, and L. V. Kalé. 2007. Scalable molecular dynamics with NAMD on the IBM Blue Gene/L system. *IBM J. Res. Dev.* 52:177–188.