THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

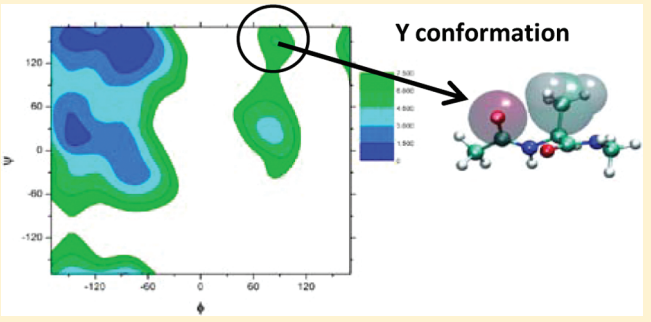
Water-Mediated Conformations of the Alanine Dipeptide as Revealed by Distributed Umbrella Sampling Simulations, Quantum Mechanics Based Calculations, and Experimental Data

Víctor Cruz,* Javier Ramos, and Javier Martínez-Salazar

BIOPHYM, Instituto de Estructura de la Materia, CSIC, Serrano 113bis, 28006, Madrid, Spain

Supporting Information

ABSTRACT: An exhaustive umbrella sampling simulation of the alanine dipeptide in solution is reported. The presence of stable Y conformations in solution is assessed by both quantum calculations and experimental data from X-ray and NMR protein-solved structures available in the protein coil library. The agreement between experimental and simulation Ramachandran plots of the dipeptide in solution is excellent. A suitable explanation of the stabilization of the Y conformation mediated by water for the alanine dipeptide is discussed on the basis of Car−Parrinello MD calculations of the dipeptide in water.



INTRODUCTION

The alanine dipeptide has been extensively used to study the structure of aminoacids. The main reason resides in the fact that it is able to model almost all known aminoacid conformations in proteins.[1−3] Experimental data as well as simulation have identified several regions of the conformational space more plausible for the alanine dipeptide. Among these regions, the polyproline-II one (Table 1) has been proved to be the most favorable.[4−8] Other regions are also important such as the β-like and the α_R conformations, associated with secondary structure motifs in proteins (β-strands and α-helix, respectively). Further conformations have been identified on the conformational map such as the α_L, δ, γ, and τ structures but in small proportions.[9] The application of empirical force field based simulations has been proved to be robust enough to study protein structure[10−40] (for details, see the work by Rueda et al.).[41] There are different protocols to explore the molecular conformational space. The most commonly used in the analysis of small peptides is replica exchange molecular dynamics (REMD).[42] Some drawbacks are, however, reported in the literature concerning the free energies of some interesting states, which are too high compared to the most stable system configuration[36] and therefore may not be sufficiently sampled. This can be alleviated by using umbrella sampling strategies.[29,36] In this methodology, biased potentials are added to the Hamiltonian to explore the system energy changes along some particular coordinate reaction. In the present work, we take the φ, ψ torsion angles as the geometric parameters over which we applied the biasing potential. For each umbrella, we applied a multinanosecond MD simulation protocol to sample the configurational space. All the information obtained from the umbrella sampling procedure can be submitted to the weighted histogram analysis method (WHAM)[43] to obtain an unbiased potential of mean force (PMF) with respect to the chosen reaction coordinate during the sampling procedure.

However, an alternative to evaluate the PMF profile, known as the multistate Bennett acceptance ratio (MBAR) estimator,[44] has been recently developed, which has significant advantages over histogram reweighting methods such as WHAM. On one hand, it does not require the sampled energy range to be discretized to produce histograms, eliminating bias due to energy binning and significantly reducing the time complexity of computing a solution to the estimating equations in many cases. Additionally, a direct estimate of the statistical uncertainty is provided for all calculated quantities.

The umbrella sampling strategy produces a conformational grid in which each cell is sufficiently small and independent to be suitable for distributed computations in ordinary desktop computers (for details, see the Computational Methods section). In this sense, we take advantage of the resources available in a local Spanish initiative for distributed computing known as Ibercivis,[45] which is similar in spirit to initiatives such as SETI@HOME. This is a volunteer computing platform designed to run scientific problems in a distributed computing paradigm using available CPU time allocated in volunteer citizens' desktop computers.

COMPUTATIONAL METHODS

Umbrella Sampling. The alanine dipeptide was built in the usual way with three residues, namely, the ALA monomer

Received: March 10, 2011
Published: April 06, 2011

Table 1. Definition of the Basin Boundaries in the Ramachandran Plots (Angles Are Given in Degrees)

	ϕ_{min}	ϕ_{max}	ψ_{min}	ψ_{max}
α_R	−160	−40	−70	30
β	−180	−105	100	−130
PP_{II}	−105	−50	100	180
α_L	40	120	−60	60
Y	60	120	70	−140

flanked by the acetyl (ACE) and N-methyl amide (NME) terminal blocks. The φ and ψ backbone dihedral angles were selected for the umbrella sampling procedure. Each torsion angle was varied from −180.0 to 180.0° in 10° intervals. Each conformation is embedded in a simulation box of water molecules using the genbox utility of the GROMACS software suite.[46] Cubic box dimensions were large enough to avoid periodic image artifacts during the molecular dynamics simulation (2.6 nm of box edge). The generated trajectories were analyzed using the MBAR estimator[44] to calculate the potential of mean force (PMF). The final PMF values were obtained by self-consistently solving the corresponding set of coupled equations (eq 11 in ref 44) until the maximum relative change between successive PMF values is below 10^{-7} (relative tolerance, see Appendix C.1.a. in ref 44). In order to estimate the associated statistical uncertainties corresponding to each PMF value following the MBAR formalism, it is desirable to have a set of uncorrelated samples from each umbrella simulation. These samples can be obtained from the corresponding correlated time series by subsampling with a time interval greater than the maximum of the statistical inefficiency. The interested reader can find further details about the calculation of statistical inefficiency in Appendix A of ref 44.

Molecular Dynamics Simulation. The OPLS-AA force field for the dipeptide in conjunction with the TIP4P model for the water solvent has been used along this work.[47−49] Some authors reported that the OPLSAA/TIP4P combination was the most successful among different force fields to reproduce the conformational space of the alanine dipeptide.[40] Prior to the molecular dynamics simulation, the energy of the system is minimized to release unphysical atomic contacts. After minimization, the system is subjected to MD simulations in the NPT ensemble with position restraints in every protein atom (LINCS algorithm). The final simulation cell contained 600 water molecules in a cubic box of 26.476 Å axis length, giving a water number density of 0.03233 $Å^{-3}$.

Subsequently, for each umbrella structure, a 2 ns molecular dynamics simulation was run at 300 K in the NVT ensemble with bond restraints through the LINCS algorithm,[50] so that the time step was 2 fs. Dihedral angle restraints were introduced by a quadratic harmonic potential with a force constant of 350 kJ mol^{-1} rad^{-1} which allowed the torsion angles to sample ±10° around the reference φ and ψ values. Temperature (300 K) was kept constant by coupling the system to an external bath following the Berendsen algorithm.[51] The coupling constant was 0.1 for temperature. The particle mesh Ewald (PME) method[52,53] was used to calculate the long-range electrostatic interactions. A total of 1260 (36 × 35) molecular dynamics simulations were ran and were collected for subsequent analysis.

In order to illustrate the backbone dihedral angle variability produced in each biased simulation, we show in Figure S1 of the Supporting Information the φ and ψ dihedral angle distributions for umbrellas in the α_R and PPII basins. In particular, those corresponding to −70°, −30° and 80°, 150° (φ, ψ) pairs, respectively. As can be observed in this figure, the dihedral restraint force constant selected allows a 10° dispersion around each restricted φ, ψ pair, which, in principle, allows one to sample the whole conformational space spanned by the two backbone dihedral angles.

Quantum Mechanics Calculations. Quantum mechanics calculations at the MP2/6-311G(d,p) level were performed for selected alanine dipeptide conformations. Solvation effects were taken into account through the polarizable continuum model (PCM) of Tomasi.[54] All the geometry optimizations with this MP2 Hamiltonian were performed with the Gaussian 09 package.[55]

Additionally, we performed ab initio molecular dynamics (AIMD) calculations on α_R and Y conformations with explicit water molecules. These DFT Car−Parrinello simulations (DFT-CP) were performed using CPMD v.3.11. The Becke, Lee, Yang, and Parr (BLYP) gradient-corrected functional has been used. The one-electron orbitals are expanded in a plane-wave basis set with a kinetic energy cutoff of 70 Ry. The norm-conserving pseudopotentials generated according to the Goedecker procedure have been used.

Periodic boundary conditions have been imposed using cubic boxes with 1 alanine dipeptide molecule and 103 water molecules with a box length of 14.9 Å. This corresponds to a macroscopic density of 1.004 g/cm^3. The starting configurations have been obtained from the umbrella sampling data set. Canonical ensemble simulations (NVT ensemble) were performed using a fictitious electron mass of 500 au and a time step of 5 au (0.12 fs). The temperature was controlled using Nose−Hoover chain thermostats with a target temperature of 300 K and a characteristic frequency of 3000 cm^{-1}. The runs extended up to 40000 steps (4.8 ps) and were sampled every 100 steps. To conserve the hydrogen bond network, the hydrogen atom keeps its real atomic mass. This size represents a compromise between spatial extent and computational cost.

The criterion to identify hydrogen bonds between alanine dipeptides and water molecules is based on the following geometrical parameters: donor−acceptor distance and angle bond less than 3.0 Å and 20°, respectively.

RESULTS AND DISCUSSION

All the restricted MD production runs in the NVT ensemble remain well equilibrated along the 2 ns simulation period. The absence of any drift in the various energy components, in particular dipeptide−solvent interactions, confirms that statement.

The PMF maps were calculated applying the MBAR estimation model as was specified in the Computational Methods section. The convergence criteria was to accept the results when the maximum relative change in PMF values calculated between successive iterations was less than 10^{-7}

$$\max_{i=2,\ldots,K} |f_i^{(n+1)} - f_i^{(n)}| / |f_i^{(n)}| < 10^{-7}$$

where f_i is the free energy value that runs over the K states and n is the iteration number.

The corresponding PMF values for each φ, ψ bin, the number of samples in each bin, and the statistical uncertainty associated with each PMF value, directly calculated by the MBAR model,

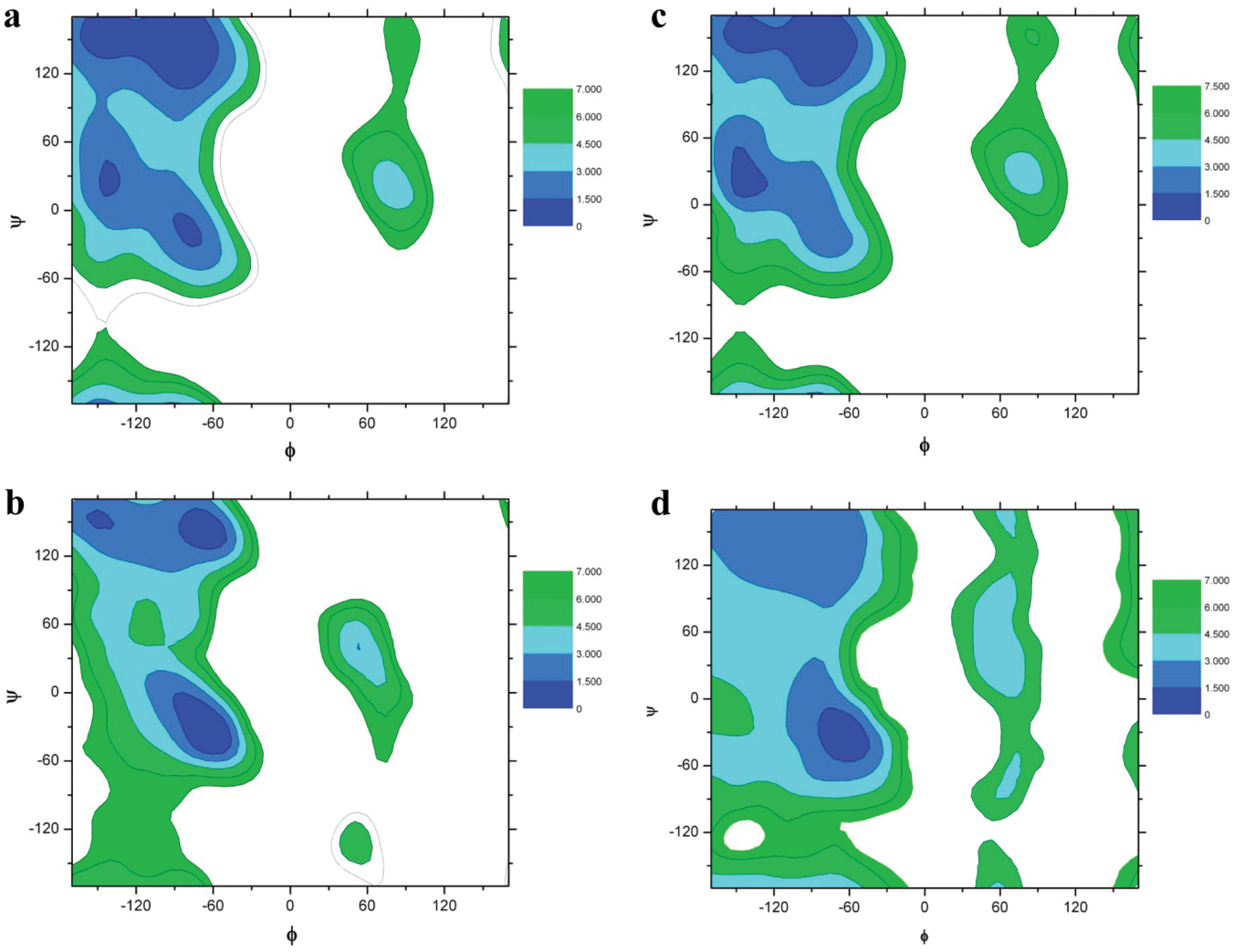


Figure 1. (a) PMF for the alanine dipeptide, taking into account those energy terms associated only with the dipeptide. (b) Ramachandran plot for alanine residues extracted from the protein coil library from X-ray solved proteins. (c) PMF for the alanine dipeptide, including energy terms associated with dipeptide−water interaction. (d) Ramachandran plot for alanine residues extracted from the protein coil library selecting structures elucidated by solution NMR. The color legend stands for energy contours giving in kT units. The dihedral angles are in degrees.

can be found in Table S1 in the Supporting Information. Here, just to mention, the mean relative uncertainty estimated for the calculated PMF values is less than 6%.

Figure 1a shows the PMF for the alanine dipeptide, taking into account those energy terms associated with the dipeptide itself, namely, bending, torsion, and nonbonding interactions. The energy corresponding to the φ, ψ dihedral restraints biasing the trajectories was subtracted. Several basins can be observed (basin boundaries are available in Table 1), with PPII being slightly more stable than that associated with the β conformation. Also noticeable is the α_R basin with an energy level somewhat above those conformations. Also mapped is the α_L basin but at higher energy levels than the other conformations. This map compares reasonably well with the PMF plot (Figure 1b) obtained for alanine residues in a list extracted from the Protein Coil Library (PCL).[56] This library contains peptide fragments extracted from the protein databank that are not associated with secondary structure motifs. The particular list we used at this point contains fragments of X-ray-elucidated proteins from the PISCES database server[57] according to the following criteria: less than 20% sequence identity, better than 1.6 Å resolution, and a refinement factor of 0.25. We have calculated the PMF of the different alanine conformations following these steps:

(1) The Ramachandran plot is divided in cells of 20° grid size and the number N_i of φ, ψ pairs is evaluated in each cell.
(2) The raw normalized probability associated with each cell, P_i, is calculated as the ratio $P_i = N_i/\Sigma N_i$.
(3) Finally, the PMF (f), in kT units to compare with the MBAR maps, is evaluated at each cell as $f_i = -\ln P_i$.

In Figure 1c, we present the PMF map calculated taking into account the dipeptide−water intermolecular interaction terms. In this case, we obtain a clear preponderance of the PPII conformation in detriment of the β and α_R structures. This is in agreement with the general result, experimentally obtained by different techniques, which shows PPII as the main conformation for the alanine dipeptide in solution. Our results confirm that inclusion of energy terms concerning intermolecular water−dipeptide interactions stabilizes the PPII conformation with

dx.doi.org/10.1021/jp2022727 |J. Phys. Chem. B 2011, 115, 4880–4886

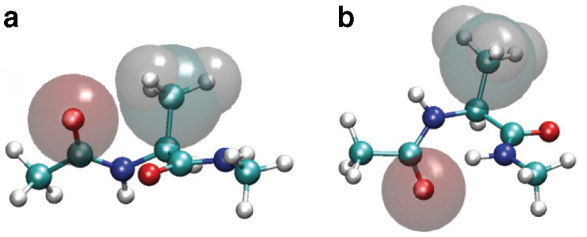


Figure 2. (a) Ball and Stick representation of the alanine dipeptide Y conformation combined with a CPK representation of the methyl side chain and the carbonyl O atoms. (b) Ball and Stick representation of the alanine dipeptide α_R conformation.

respect to the β and α_R ones in close agreement with other authors.[4,6,22,58]

In addition to the PPII region, a new basin appears enhanced in the calculated PMF map which can be assigned to the Y conformation.[9,59] This result is particularly striking because the Y structure is catalogued as a disfavored region for all aminoacids except glycine.[9] This conformation has been observed by other authors on calculated PMF maps of alanine,[60] but to our knowledge, there is no discussion concerning the stabilization of this structure.

At this point, we wondered if the overstabilization around the Y basin in the Ramachandran representation we obtained for the alanine residue could be an artifact of the force field or could have any physical sense.

On one hand, we examined again the data compiled in the PCL database looking for solution NMR resolved structures with less than 20% sequence identity. It is agreed that, in general, crystallographic derived structures have more resolution than solution NMR ones. However, the experimental procedures to obtain a good enough crystal could significantly "dry" the peptide samples. In this sense, solution NMR has the advantage of dealing with samples near physiological conditions, hydration included, but at the expense of lower resolution structures. With this aim, we searched in the PDB those structures resolved by solution NMR and extracted the corresponding torsion files by a batch searching of the PDB codes list in the PCL server. In Figure 1d, we present the PMF plot of the (φ, ψ) pairs corresponding to alanine residues obtained from the torsion files. As it can be observed, there is a noticeable Y-basin.

In order to explain these results, we analyzed the steric hindrance between the CH_3 substituent on the Cα atom and the carbonyl O of the ACE terminus. This steric factor is argued to be the reason why the Y-basin is classified as a disfavored region for all aminoacids but glycine. That is the case if we observe Figure 2a, where we have represented the alanine dipeptide with φ, ψ torsion angle values of 80.0 and 160.0°, respectively. As it can be observed, the O atom partially eclipses the CH_3 group. This eclipsed conformation has an additional energetic cost with respect to an unstrained structure (Figure 2b, $\varphi \sim -70.0°$, $\psi \sim -20.0°$) corresponding to an α_R-like conformation. This energetic cost of the dipeptide in the gas phase is calculated at the MP2 level, with the α_R conformation being more stable by 1.8 kcal/mol.

However, the solvation effect, calculated as the difference between the energies of the PCM implicit water model and the gas phase structure at the same MP2 level, is bigger for the α_R conformation (−9.5 kcal/mol) than for the Y conformation (−8.1 kcal/mol). We include as Supporting Information the

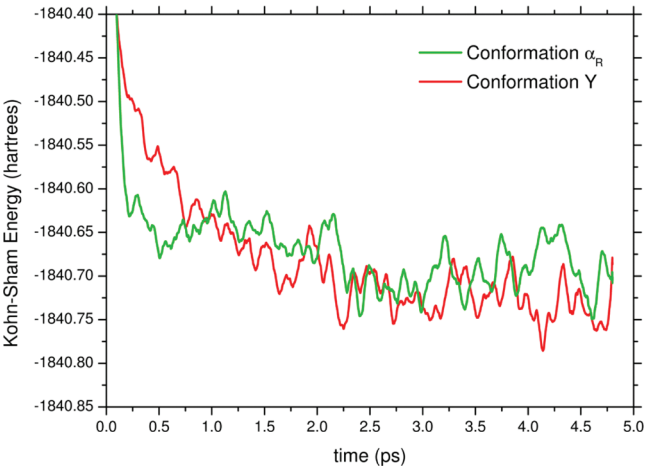


Figure 3. Evolution of the Kohn−Sham energy along the CPMD simulation trajectory.

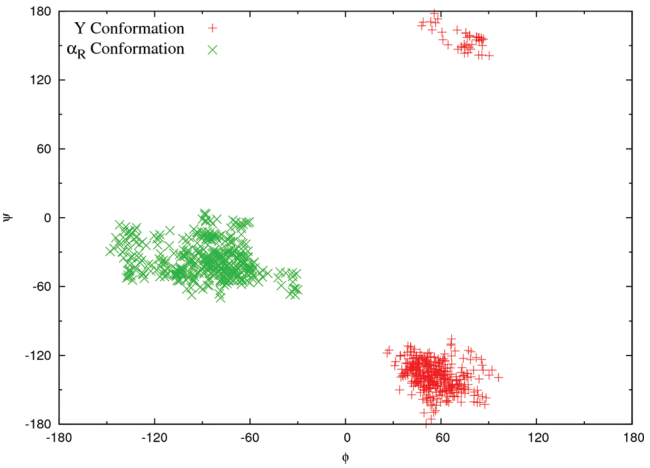


Figure 4. Ramachandran plot corresponding to structures calculated along the CPMD trajectory.

Gaussian 09 output files corresponding to the implicit solvent calculations with the α_R and Y conformations.

The methyl side chain in alanine is a hydrophobic group which tends to be naturally unstable in water. This observation has been clearly confirmed by the ab initio MD calculation on alanine in a water droplet made by Degtyarenko and co-workers.[14] These authors showed that the methyl side chain group of the alanine molecule tends to move to the droplet surface during the MD simulation in order to minimize the free energy.[14] Thus, the partial shielding of the methyl group exerted by the carbonyl O in the Y conformation could release some of the hydrophobic instability of the alanine side chain in water. In fact, the accessible surface area to solvent calculated from the MD trajectories for the methyl group in the Y structure is only 0.49 ± 0.05 nm^2, whereas for the α_R geometry it is 0.55 ± 0.05 nm^2.

However, it has been shown that implicit solvent models are crude approximations to the real solvent structure and do not take into account some subtleties associated with the H-bond network of water molecules around a solute.[61,62] Therefore, we decided to carry out CPMD calculations taking into account a set of explicit water molecules. We take a snapshot from the MD simulation box corresponding to the alanine dipeptide structures

dx.doi.org/10.1021/jp2022727 |J. Phys. Chem. B 2011, 115, 4880–4886

already used for the in vacuum and implicit solvent calculations for both conformations. Then, we selected a subset of the simulation box and carried out a 4.8 ps Car−Parrinello MD simulation as specified in the Computational Methods section.

The simulated systems seem to be acceptably equilibrated regarding system temperature and various energy components.

In Figure 3, we present the evolution of the Kohn−Sham energy along the simulation period. This energy is comparable to the energy that can be obtained from classical MD simulations. As can be observed, the energy values of the Y system tend to be of the same magnitude as those obtained for the α_R conformation. The Ramachandran representation of the φ, ψ backbone angles obtained along the CPMD simulation trajectory for both systems is presented in Figure 4. As it can be seen, the backbone dihedral angles remain in a wide value range around their respective conformations.

In order to gain new insights concerning the alanine−water interaction, we evaluated different radial distribution functions, $g(r)$, corresponding to selected atomic pairs, from the CPMD simulations. The $g(r)$ of water O atoms around the methyl side-chain center-of-mass for α_R and Y conformations is presented in Figure 5. The water molecules are more uniformly distributed around the Y structure, as can be deduced from the softer $g(r)$ profile for this conformation. This is in agreement with a slight increase in entropy with respect to the α_R case. Then, this shielding effect is expected to compensate the steric hindrance in the Y conformation. Regarding the H water distribution around the dipeptide carbonyl O atoms, in Figure 6, we show the $g(r)$ corresponding to interaction between the two dipeptide carbonyl O atoms and H water atoms. A similar alanine−water interaction can be observed for both conformations around the carbonyl O atom belonging to the alanine residue (Figure 6a). The $g(r)$ profile for the carbonyl O belonging to the ACE residue, which is next to the methyl group, is slightly weaker in the Y configuration (Figure 6b). In spite of the somewhat different atomic environment around these carbonyl oxygens, the hydrogen bonds between the carbonyl groups and water molecules are quite similar for both conformations fluctuating between one and four hydrogen bonds along the CPMD trajectories.

The bulk water structure in the simulation box remains very similar for both systems, as can be observed in Figure 7 where we represent the radial distribution function of the O water−O water interatomic interaction. Furthermore, the $g(r)$ functions calculated for the different water interatomic interactions compare well with the experimental $g(r)$ obtained for pure water[63] (see Figure 2 in ref 63). The corresponding $g(r)$ plots are included as Supporting Information (Figures S2 and S3).

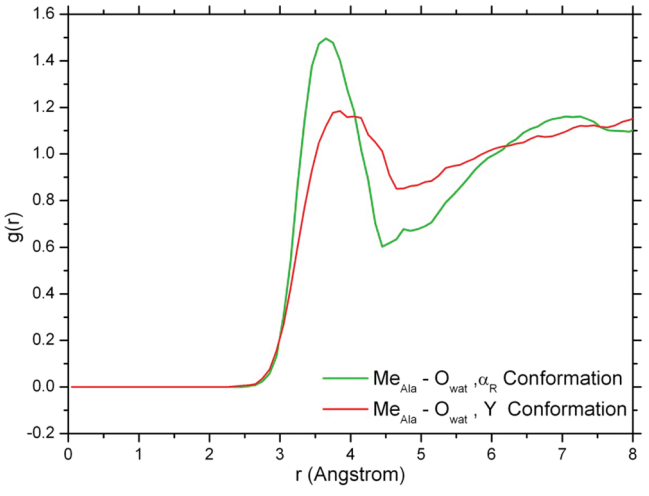


Figure 5. Radial distribution function of O water atoms around the alanine methyl center of mass.

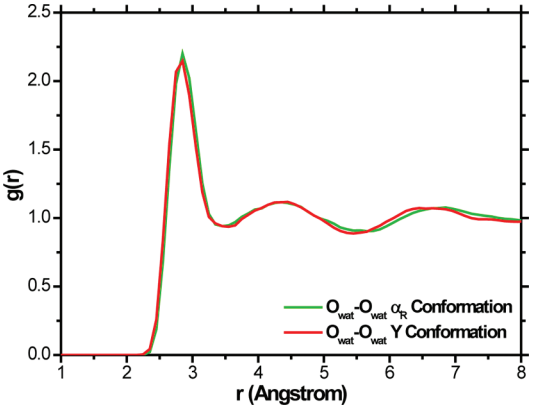


Figure 7. Radial distribution function corresponding to water O−water O interaction.

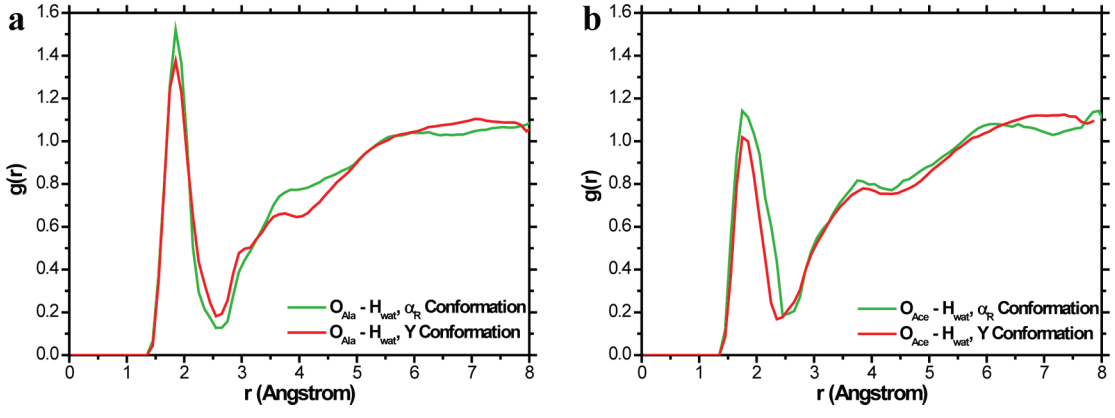


Figure 6. (a) Radial distribution function of H water atoms around the alanine O carbonyl atom. (b) Radial distribution function of H water atoms around the ACE O carbonyl atom.

dx.doi.org/10.1021/jp2022727 |J. Phys. Chem. B 2011, 115, 4880–4886

CONCLUSIONS

Taking advantage of the large amount of distributed computational resources available through the Ibercivis platform, we have performed an exhaustive sampling of the conformational space of the alanine dipeptide in solution. When we consider protein structures resolved by solution NMR, the Ramachandran plot corresponding to alanine residues presents a significative number of points in the Y basin conformation, in contrast with the absence of this region in plots derived from X-ray crystallographic data. Our PMF maps catch this experimental observation.

The PMF maps obtained after the weighted histogram treatment of the umbrella sampling showed that the Y conformation can be favored when we take into account explicitly solute−solvent interactions. We have supported this observation by means of QM calculations showing that, in explicit water, the Y conformation is stable and can be competitive with the α_R structure. This Y conformation is supposed to be found only in the case of glycine dipeptide, due to the absence of steric hindrance between the backbone and the side chain. However, in the case of the alanine dipeptide, we have shown that partial shielding of the methyl side chain by backbone carbonyl oxygen can overcome the steric hindrance by reduction of the hydrophobic effect of the alkyl group in a water environment.

These results might provide new insights for understanding the conformation of unfolded peptides and proteins as well as the unfolded state during the protein folding.

ASSOCIATED CONTENT

Supporting Information. Table S1 includes the corresponding PMF values for each φ, ψ bin, the number of samples in each bin, and the statistical uncertainty associated with each PMF value. Figure S1 shows φ and ψ dihedral angle distributions corresponding to simulations of α_R and Y conformations. Figures S2 and S3 show different radial distribution functions corresponding to interactions between dipeptide and water or between water molecules. The Gaussian 09 output files for the implicit solvent MP2 calculations are given in two log files, namely, "jp2022727_si_002.txt" and "jp2022727_si_003.txt" for the α_R and Y conformations, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +34915616800. E-mail: victor.cruz@iem.cfmac.csic.es.

ACKNOWLEDGMENT

Thanks are due to the CICYT (MAT2009-12364 project) for financial support. This research was also funded by the CSIC (Consejo Superior de Investigaciones Científicas, Spain) under the Grant PIE-200850I072. Computational support in the use of the distributed computing to the Ibercivis team is acknowledged (www.ibercivis.es). We are also very grateful to the anonymous citizens who have made available their desktop computers in an altruistic way.

REFERENCES

(1) Avbelj, F.; Grdadolnik, S. G.; Grdadolnik, J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1272.
(2) Smith, P. E.; Pettitt, B. M. *J. Phys. Chem.* **1993**, *97*, 6907.
(3) Apostolakis, J.; Ferrara, P.; Caflisch, A. *J. Chem. Phys.* **1999**, *110*, 2099.
(4) Lee, M. E.; Lee, S. Y.; Joo, S. W.; Cho, K. H. *J. Phys. Chem. B* **2009**, *113*, 6894.
(5) Shi, Z. S.; Chen, K.; Liu, Z. G.; Ng, A.; Bracken, W. C.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17964.
(6) Chellgren, B. W.; Miller, A. F.; Creamer, T. P. *J. Mol. Biol.* **2006**, *361*, 362.
(7) Weise, C. F.; Weisshaar, J. C. *J. Phys. Chem. B* **2003**, *107*, 3265.
(8) Hagarman, A.; Measey, T. J.; Mathieu, D.; Schwalbe, H.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2010**, *132*, 540.
(9) Perskie, L. L.; Rose, G. D. *Protein Sci.* **2010**, *19*, 1127.
(10) Anderson, A. G.; Hermans, J. *Proteins: Struct., Funct., Genet.* **1988**, *3*, 262.
(11) Baldwin, R. L. *J. Mol. Biol.* **2007**, *371*, 283.
(12) Berezhkovskii, A.; Hummer, G.; Szabo, A. *J. Chem. Phys.* **2009**, *130*, 205102.
(13) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Dill, K. A. *Multiscale Model. Simul.* **2006**, *5*, 1214.
(14) Degtyarenko, I. M.; Jalkanen, K. J.; Gurtovenko, A. A.; Nieminen, R. M. *J. Phys. Chem. B* **2007**, *111*, 4227.
(15) Deplazes, E.; van Bronswijk, W.; Zhu, F.; Barron, L. D.; Ma, S.; Nafie, L. A.; Jalkanen, K. J. *Theor. Chem. Acc.* **2008**, *119*, 155.
(16) DePristo, M. A.; de Bakker, P. I. W.; Lovell, S. C.; Blundell, T. L. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 41.
(17) Escobedo, F. A.; Borrero, E. E.; Araque, J. C. *J. Phys.: Condens. Matter* **2009**, *21*, 333101.
(18) Gallicchio, E.; Paris, K.; Levy, R. M. *J. Chem. Theory Comput.* **2009**, *5*, 2544.
(19) Henin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. *J. Chem. Theory Comput.* **2010**, *6*, 35.
(20) Hu, H.; Elstner, M.; Hermans, J. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 451.
(21) Ireta, J.; Scheffler, M. *J. Chem. Phys.* **2009**, *131*, 085104.
(22) Jono, R.; Watanabe, Y.; Shimizu, K.; Terada, T. *J. Comput. Chem.* **2010**, *31*, 1168.
(23) Liu, C.; Zhao, D. X.; Yang, Z. Z. *J. Theor. Comput. Chem.* **2006**, *9*, 77.
(24) Mijajlovic, M.; Biggs, M. J. *J. Phys. Chem. B* **2007**, *111*, 7591.
(25) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902.
(26) Mulder, F. A. A.; Filatov, M. *Chem. Soc. Rev.* **2009**, *39*, 578.
(27) Penev, E. S.; Lampoudi, S.; Shea, J. E. *Comput. Phys. Commun.* **2009**, *180*, 2013.
(28) Prada-Gracia, D.; Gomez-Gardenes, J.; Echenique, P.; Falo, F. *PLoS Comput. Biol.* **2009**, *5*, e1600415.
(29) Ravindranathan, K. P.; Gallicchio, E.; Levy, R. M. *J. Mol. Biol.* **2005**, *353*, 196.
(30) Seabra, G. D.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. *J. Phys. Chem. A* **2007**, *111*, 5655.
(31) Seabra, G. D.; Walker, R. C.; Roitberg, A. E. *J. Phys. Chem. A* **2009**, *113*, 11938.
(32) Strodel, B.; Wales, D. J. *Chem. Phys. Lett.* **2008**, *466*, 105.
(33) Sun, C. L.; Jiang, X. N.; Wang, C. S. *J. Comput. Chem.* **2009**, *30*, 2567.
(34) Unal, E. B.; Gursoy, A.; Erman, B. *Phys. Biol.* **2009**, *6*.
(35) Velez-Vega, C.; Borrero, E. E.; Escobedob, F. A. *J. Chem. Phys.* **2009**, *130*, 225101.
(36) Wang, J.; Gu, Y.; Liu, H. Y. *J. Chem. Phys.* **2006**, *125*, 9.
(37) Xu, C.; Wang, J.; Liu, H. Y. *J. Chem. Theory Comput.* **2008**, *4*, 1348.
(38) Yang, S.; Cho, M. *J. Chem. Phys.* **2009**, *131*, 135102.
(39) Zheng, W. H.; Andrec, M.; Gallicchio, E.; Levy, R. M. *J. Phys. Chem. B* **2009**, *113*, 11702.
(40) Kwac, K.; Lee, K. K.; Han, J. B.; Oh, K. I.; Cho, M. *J. Chem. Phys.* **2008**, *128*, 105106.
(41) Rueda, M.; Ferrer-Costa, C.; Meyer, T.; Perez, A.; Camps, J.; Hospital, A.; Gelpi, J. L.; Orozco, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 796.

(42) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.
(43) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.
(44) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, 129.
(45) http://www.ibercivis.es, 2010.
(46) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.
(47) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(48) Jorgensen, W. L.; Tirado-Rives, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665.
(49) Lawrence, C. P.; Skinner, J. L. *Chem. Phys. Lett.* **2003**, *372*, 842.
(50) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463.
(51) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(52) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(53) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(54) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
(55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
(56) Fitzkee, N. C.; Fleming, P. J.; Rose, G. D. *Proteins: Struct., Funct., Bioinf.* **2005**, *58*, 852.
(57) Wang, G. L.; Dunbrack, R. L. *Bioinformatics* **2003**, *19*, 1589.
(58) Law, P. B.; Daggett, V. *Protein Eng., Des. Sel.* **2010**, *23*, 27.
(59) Perskie, L. L.; Street, T. O.; Rose, G. D. *Protein Sci.* **2008**, *17*, 1151.
(60) Feig, M. *J. Chem. Theory Comput.* **2008**, *4*, 1555.
(61) Mirkin, N. G.; Krimm, S. *Biopolymers* **2009**, *91*, 791.
(62) Mukhopadhyay, P.; Zuber, G.; Beratan, D. N. *Biophys. J.* **2008**, *95*, 5574.
(63) Soper, A. K. *Chem. Phys.* **2000**, *258*, 121.