THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

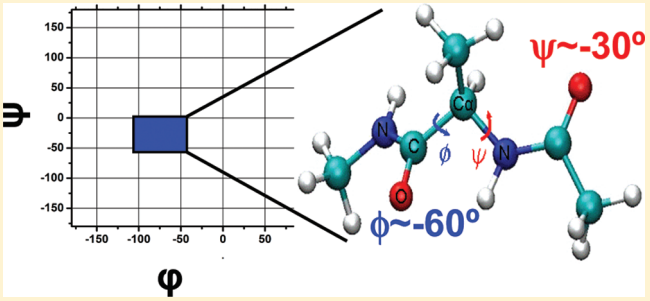
Assessment of the Intrinsic Conformational Preferences of Dipeptide Amino Acids in Aqueous Solution by Combined Umbrella Sampling/MBAR Statistics. A Comparison with Experimental Results

Victor L. Cruz,* Javier Ramos, and Javier Martinez-Salazar

BIOPHYM, Department of Macromolecular Physics, Instituto de Estructura de la Materia, CSIC Serrano 113-bis, 28006 Madrid, Spain

Supporting Information

ABSTRACT: The propensities of 19 amino acid dipeptides have been calculated by a distributed umbrella sampling molecular dynamics simulation procedure using the OPLS-AA force field. The potential of mean force maps was estimated with the multiple Bennett acceptance ratio statistics. The resulting propensities compare satisfactorily well with very recently published experimental data on equivalent systems. In particular, α conformation-probabilities for all of the dipeptides remain much lower than either β or P_{II} propensities. This result is in agreement with most experimental data for dipeptides. However, it is also in contrast with most simulation studies performed so far with other force fields, where α conformations result even more probable than P_{II} or β ones. We discuss the behavior of the OPLS-AA force field, which can be useful for the improvement of this model in reproducing the recent experimental observations on amino acid dipeptides.



I. INTRODUCTION

Dipeptides have been extensively studied from both theoretical and experimental points of view as models for the unfolded protein state.[1−15] For example, a high propensity for the extended P_{II} conformation is generally observed for dipeptides in contrast to a preference toward the helical α_R conformation reported for the same amino acid in the context of protein coil databases.[16−18] Thus, it has been proposed that coil regions in proteins are biased by the secondary structure adopted by the contiguous amino acids in the protein backbone.[19]

Recent IR and Raman experiments on 19 amino acid dipeptides in aqueous solution reveal that P_{II} and β are the predominant conformations and that the α_R structure is only noticeable in small amounts.[8,20] This observation is in line with other results such as those obtained by Schweitzer-Stenner et al. on several GXG tripeptides combining vibrational and NMR spectroscopy.[21] On the other hand, GGXGG peptides present similar propensity tendencies to the above-mentioned, but these authors recognize also that probabilities can be sequence and context dependent.[22] That seems to be the case in view of the propensities that were obtained by Dunbrack et al. with a different selection made from a set of 3038 proteins from the Uppsala Electron Density Server.[23] These authors showed that P_{II} (42.1%) and β (32.7%) conformations are more populated than α_R (17.1%) ones in the case of amino acids located in random coil motifs. However, they observed an increase of α_R population (from 17.1 to 42.4%) when they analyzed residues belonging to turn structural motifs. Also, from analysis of the Protein Coil Library,[16] Rose and co-workers reported a 32.9% abundance of the α-helix motif[18] on a selected subset of their coil library compilation including turns, bridges, and coils. Other compilations from the Protein Coil Library show also a clear preference of the α-helix conformation.[24] However, Jha[10] obtained P_{II} (35.5%) and β (32.9%) conformer distributions, larger than α_R (27.4%) in a coil library compiled without residues in helices, sheets, turns, and flanking residues belonging to any of these substructures. Thus, the amino acid conformation populations in dipeptides and proteins seem to be different, at least, when the amino acid belongs to or is close to a secondary structural motif.

On the other hand, a huge number of simulation articles concerning individual dipeptides, such as alanine, have been released.[25−35] However, theoretical systematic studies are less common. For example, Feig[7] calculated the potential of mean force (PMF) of the 20 amino acid dipeptides (Ac-X-NH_2) by performing 150 ns unrestricted MD simulations using the CHARMM22/CMAP force field. His results show that extended (β/P_{II}) and helical (α_R/α') basins are almost equally populated in all nonglycine and nonproline amino acids. As mentioned above, it has been experimentally found that P_{II} and β are the dominant conformation for dipeptides in water. Thus, the author suggested that the overpopulation of α basin in its simulations is due to a bias toward α-helical conformations produced by the

Received: July 15, 2011
Revised: November 11, 2011
Published: December 04, 2011

CHARMM22/CMAP force field. A similar conclusion can be drawn for alanine dipeptide simulations.[28] Therefore, a central question in the simulation of small peptides or unfolded states is the suitability of the force field to study these systems. Several force fields and multiple variants have been used in this context with different results.[6,10,23,25,28−33,36−43] An exhaustive review of these results is beyond the scope of this paper. However, we can summarize some important observations for the interpretation of our results. Among the most used force fields in atomistic simulations of biosystems, OPLS-AA[44−46] seems to give higher P_{II} or β content in the conformational analysis of dipeptides in comparison with other commonly used force fields.[28] This observation is in agreement with experimental results showing a higher preponderance of extended conformations over α-helical ones in dipeptides. This β preference in dipeptides obtained with OPLS-AA has been observed in the particular case of the alanine dipeptide, which is used as a reference case for this kind of study. We are not aware of any study on simulations carried out with OPLS-AA over other dipeptides than ALA, GLY, or PRO.

In this work we have performed a distributed umbrella sampling simulation to study the conformational space of 19 out of 20 amino acid dipeptides capped with the acetyl (ACE) and *N*-methyl amide (NME) groups.

The umbrella sampling can cover the available conformational variability by performing a big number of simulations in a distributed architecture, so that each φ/ψ biased simulation can be performed independently in a different CPU. The citizen volunteer computing project Ibercivis[47] is an adequate platform to perform such a distributed work. This platform operates in the same way as other popular citizen distributed computing platforms such SETI@HOME or FOLDING@HOME. Each volunteer participating in the project allows the utilization of his/her personal computer during those periods of time when the computer is idle. At these moments the central Ibercivis server sends a work to that computer and receives the calculation output after completion.

For each dipeptide, MD independent simulations are performed one for each pair of restrained φ/ψ dihedral angles. Each simulation is run in turn at each available CPU.

It is expected that an extensive coverage of the conformational space will result from this strategy.

We analyze the different potential of mean force (PMF) maps for each dipeptide calculated using the multiple Bennet acceptance ratio (MBAR) estimator.[48] This methodology presents several advantages over the weighted histogram analysis method (WHAM),[49] among others, the direct calculation of estimation errors. From these maps we calculate the propensities for each of the 19 amino acid dipeptides and compare them with recent experimental data on similar systems.

II. COMPUTATIONAL METHODS

II.1. Initial Structures. The amino acid dipeptides were built in the usual way with three residues, namely, the monomer 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 as collective variables. Each torsion angle was systematically varied from −180.0° to +180.0° centered at 10° intervals. Each conformation is embedded in a simulation box of water molecules using the genbox utility of the GROMACS software suite.[50] Box dimensions were chosen to have a system with a density of about 1 g/cm^3 and large enough to avoid periodic image artifacts during the MD simulations. All systems contain the same number of water molecules (550 water molecules).

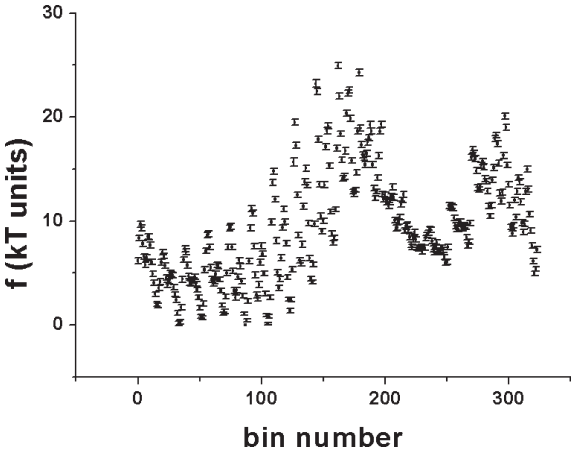


Figure 1. PMF values and the corresponding error bars estimated by the MBAR statistical analysis for the ASP dipeptide. The bin number in the abscissa axis represents each pair of φ,ψ backbone dihedral angles used to represent the PMF Ramachandran maps. For example, bin = 1 stands for the (−170,−170) pair in accordance to the numbering in Table S4 (Supporting Information).

II.2. Molecular Dynamics Simulation and Umbrella Sampling. The OPLS-AA force field for the dipeptide in conjunction with the TIP4P model for the water solvent has been used along this work.[44,45,51] Some authors reported that the OPLS-AA/TIP4P combination was the most successful among different force fields to reproduce the conformational space of the alanine dipeptide.[31]

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 an equilibration MD simulations in the NPT ensemble with position restraints in every peptide atom (LINCS algorithm[52]).

Subsequently, for each umbrella structure, 2-ns production MD simulations were run at 300 K in the NVT ensemble with bond restraints through the LINCS algorithm, so that the time step was 2 fs. We estimated that this time is enough to adequately sample each conformation. To ensure this, we carried out simulations for selected dipeptides extending to 4 ns the collection of dynamics trajectories. The PMF map obtained with the 4 ns simulations is indistinguishable to that obtained with the shorter ones.

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. The temperature (300 K) was kept constant by coupling the system to an external bath following the Berendsen algorithm.[53] The coupling constant was 0.1 ps for temperature. The particle mesh Ewald (PME) method[54,55] was used to calculate the long-range electrostatic interactions. A total of 1260 (36 × 35) molecular dynamics simulations were ran for each dipeptide and were collected to subsequent analysis.

II.3. Potential of Mean Force. We used the MBAR estimator[48] 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 48) until

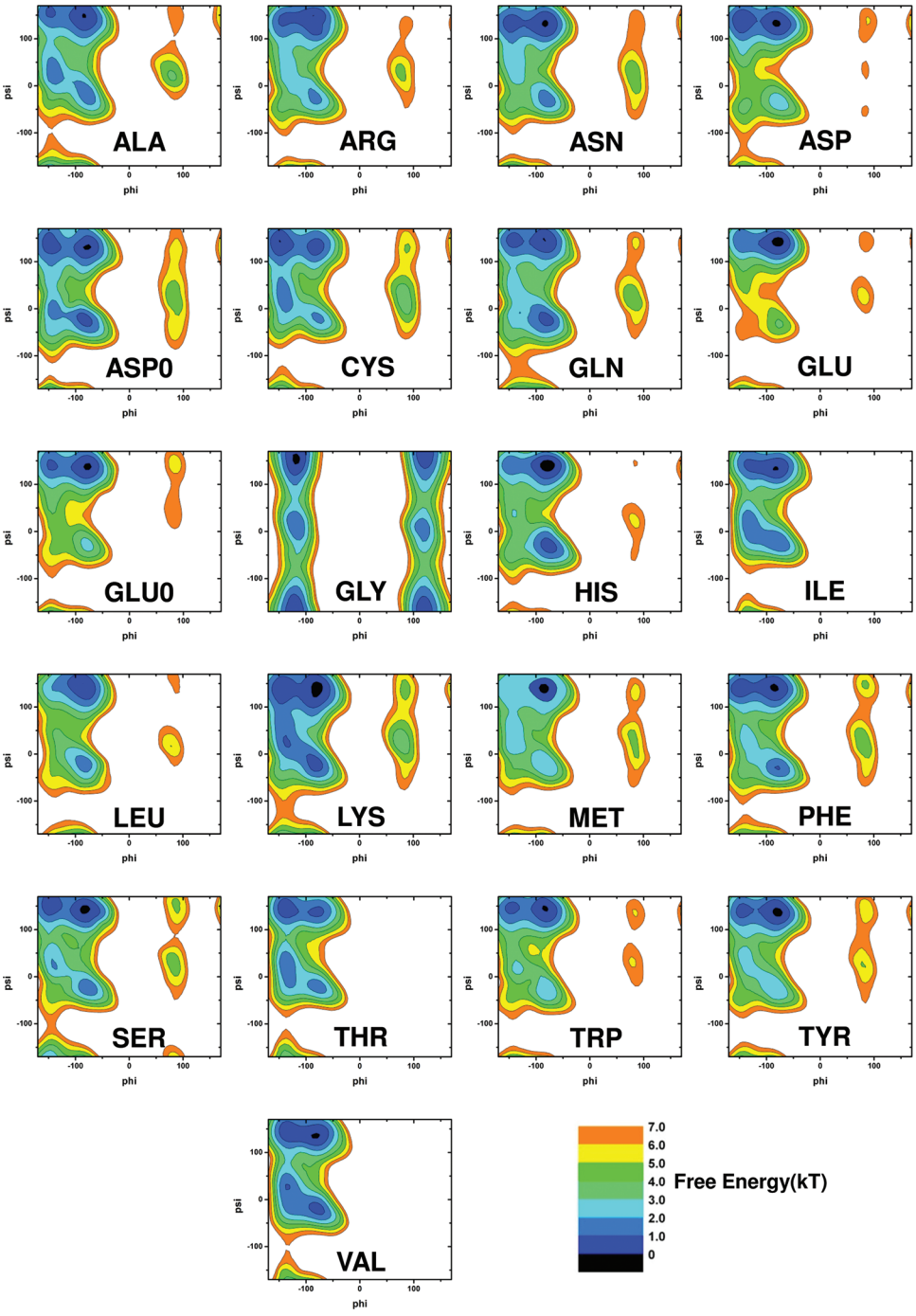


Figure 2. PMF Ramachandran maps for the all amino acid dipeptides studied in this work. The color legend is at the bottom of the figure. Energies are expressed in kT units.

the maximum relative change between successive PMF values is below 10^{-7} (relative tolerance, see Appendix C.1.a., in ref 48). 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 *g*. Further details about the calculation of statistical inefficiency can be found in Appendix A of ref 48.

III. RESULTS AND DISCUSSION

The dihedral restraint force constant used for each umbrella allows a Gaussian dihedral angle distribution with a 10° fluctuation around each restrained value, as can be observed in Figure S.1 in the Supporting Information, where we show the dihedral angle distribution obtained for a selected umbrella of the α_R configuration of the ASP dipeptide.

During the equilibration period, the energy terms concerning the solvent−solute and solvent−solvent interactions converge readily in the first hundred picoseconds. The simulation box

dx.doi.org/10.1021/jp206757j |*J. Phys. Chem. B* 2012, 116, 469–475

Table 1. Basin Definition[a]

basin	Φ_{min}	Φ_{max}	ψ_{min}	ψ_{max}
P_{II}	−100.0	0.0	70.0	180.0
β	−180.0	−100.0	70.0	180.0
α_R	−100.0	0.0	−80.0	40.0
α'	−160.0	−100.0	−80.0	90.0

[a] Angles are given in degrees.

dimensions and the associated density also converge within the same time period. The trajectories and energies are collected every picosecond and subsequently submitted to the MBAR analysis. The evaluation of the PMF through the MBAR estimator requires that samples should be uncorrelated. In the time series of each umbrella simulation, the energy values are, in principle, correlated along the molecular dynamics trajectories. We estimated the time correlation function from each time series by evaluating the statistical inefficiency parameter according to Chodera.[48] Thus, the correlation time was found to be between 20 and 30 ps for all dipeptides.

The resulting PMF values estimated along with the corresponding statistical uncertainties are presented in Tables S1−S21 of the Supporting Information. The mean relative error averaged over the different calculations is less than 5%. Figure 1 shows as an illustrative example the PMF values and corresponding error bars obtained for the ASP dipeptide. As can be observed the errors are very similar for all the PMF values and slightly decreased for the most populated bins, which, in general, correspond to the most stable conformations.

In Figure 2, we present contour representations of the different PMF maps obtained for each dipeptide. These maps show the expected landscape with four main basins corresponding to P_{II}, β, α_R, and α' conformations. Other minor conformations appear also at positive φ values, namely, α_L and Y conformations. It is worth to say that the four basins appear clearly defined in the PMF maps. The P_{II} and β basins present the lower energy values, and in the case of ARG, ASN, LYS, PHE, and VAL, a small overlap between these regions is visible. The GLY dipeptide represents a special case due to the lack of a lateral chain. The PMF map is almost symmetrical with basins placed in the four quadrants. In this case the P_{II} basin is again the lowest PMF value in agreement with the overall tendency showed for the other amino acids and in clear discrepancy with the results reported by Baldwin.[20]

We calculated the normalized probability of each different basin using the Boltzmann transformation between free energy and probability for each point i in the map

$$P(i)\alpha \exp(-F(i)/kT) \qquad \text{(eq.1)}$$

where $P(i)$ is the unnormalized probability and $F(i)$ is the calculated PMF at each umbrella point in kT units. The normalized values were obtained so that $\Sigma_i P(i) = 1$

Table 1 shows the ranges used to delimitate the main different basins.

Table 2 shows the propensities obtained for each amino acid dipeptide. In the same table, we include the experimental results obtained recently by Baldwing's group on the same dipeptides[20] and by Schweitzer-Stenner's group on some amino acids flanked by two GLY residues (GXG),[21] both in aqueous solution.

In general, the calculated probabilities are qualitatively in agreement with the experimental data in the sense that the β and P_{II} conformations are substantially more populated than the α_R basins.

Table 2. Propensities Obtained for the Different Amino Acid Dipeptides[a]

	this work				ref 20			ref 21	
dipeptide	P_{II}	β	α_R	α'	P_{II}	β	α_R	P_{II}	β
ARG	0.45	0.38	0.08	0.09	0.54	0.39	0.07		
ASN	0.40	0.42	0.08	0.10	0.40	0.58	0.02		
ASP	0.47	0.44	0.05	0.04	0.49	0.46	0.05		
ASP-0	0.44	0.37	0.11	0.08	0.43	0.55	0.02		
CYS	0.27	0.52	0.06	0.15	0.43	0.54	0.03		
GLN	0.38	0.35	0.19	0.08	0.44	0.48	0.08		
GLU	0.63	0.34	0.01	0.02	0.59	0.36	0.05	0.54	0.26
GLU-0	0.52	0.35	0.08	0.05	0.47	0.48	0.05		
GLY	0.40	0.40	0.01	0.19	0.22	0.12	0.66		
HIS	0.52	0.22	0.22	0.04	0.38	0.58	0.04		
ILE	0.44	0.32	0.13	0.11	0.46	0.52	0.02		
LEU	0.58	0.27	0.11	0.04	0.55	0.35	0.10	0.58	0.24
LYS	0.47	0.35	0.12	0.06	0.55	0.41	0.04	0.50	0.41
MET	0.51	0.31	0.07	0.11	0.50	0.47	0.03	0.64	0.36
PHE	0.40	0.42	0.12	0.06	0.45	0.49	0.06	0.42	0.40
SER	0.43	0.40	0.09	0.08	0.49	0.47	0.04	0.45	0.30
THR	0.35	0.41	0.13	0.11	0.39	0.58	0.03		
TRP	0.50	0.39	0.06	0.05	0.54	0.44	0.02		
TYR	0.49	0.38	0.06	0.07	0.46	0.47	0.07		
VAL	0.39	0.34	0.15	0.12	0.47	0.51	0.02	0.38	0.40

[a] Our results in the first column set are compared with those obtained by Baldwin et al. (ref 20) and Schweitzer-Stenner et al. (ref 21), whose data are reported in column sets 2 and 3, respectively.

The agreement between the calculated and experimental P_{II} propensities is very satisfactory with the exception of CYS and HIS, which give too low and too high P_{II} propensities, respectively. The remaining amino acids give a propensity value which differs from the experimental ones in less than 0.1.

Regarding the β conformations, the differences between calculated and experimental propensities is more noticeable than for the P_{II} ones. All of the amino acids show calculated β propensities lower than the experimental values. The largest differences correspond to HIS and ILE. The calculated β propensities for the other dipeptides present differences below 0.2 with the experimentally determined data.

The α_R conformations are, on the contrary, significantly higher than those reported by Baldwing or the experimental results available on GXG tripeptides. HIS, GLN, and the three amino acids with secondary C_β atoms, namely, ILE, THR, and VAL, present the largest α content, between 0.24 and 0.27, with a notable contribution of the α_R basin. It is also noticeable than the smallest α content correspond to the two negatively charged side chain amino acids, ASP and GLU.

The HIS and GLY amino acids seem to have the worse comparison with the experimental propensity values. It should be mentioned here that this HIS amino acid was parametrized in a different way, as was reported by Jorgensen in the original OPLS paper.[44] This amino acid was parametrized by fitting to ab initio 6-31G(d) results on the charged imidazole moiety instead of experimental data as was the case with other functional groups.

dx.doi.org/10.1021/jp206757j | J. Phys. Chem. B 2012, 116, 469–475

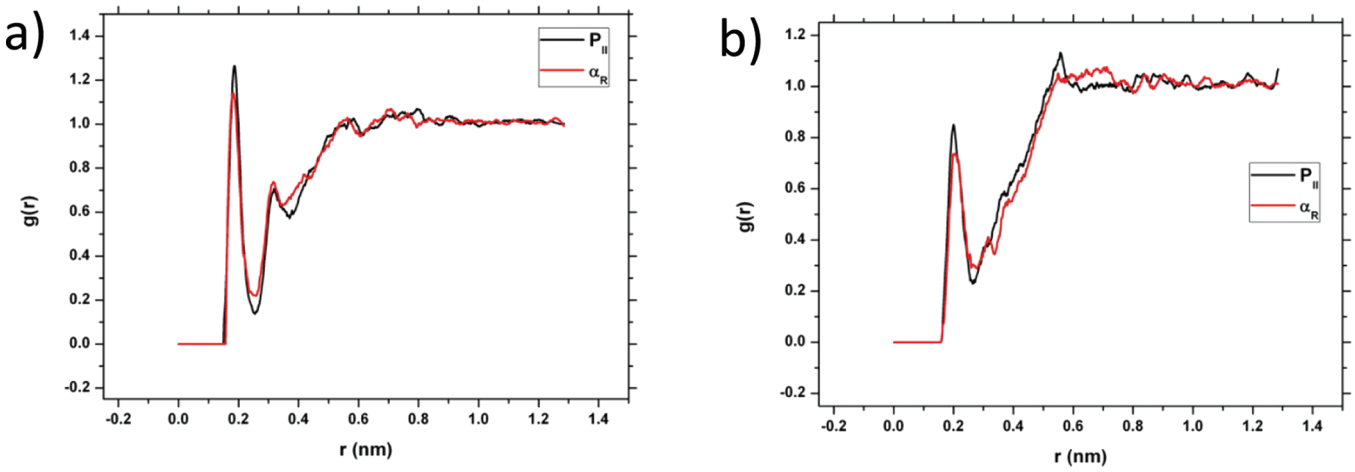


Figure 3. Radial distribution functions of water atoms around backbone atoms of the ASP dipeptide in the P_{II} and α_R conformations. (a) Distribution of H water around backbone carbonyl O atoms and (b) distribution of O water around backbone amino H atoms.

Thus, we might speculate that this different parametrization was the source of the above-mentioned discrepancies.

The ability of OPLS-AA to compare satisfactorily well with experimental probabilities on amino acid dipeptides may be due either to the force field, the distributed umbrella sampling procedure, or the PMF estimation through MBAR statistical analysis.

Regarding the behavior of the commonly used force fields in biomolecular simulations, the general trend gives larger propensities for the α conformations as was mentioned in the Introduction section. This tendency is somewhat changed toward lower α structure probabilities when additional corrections are incorporated in the force field such as polarization effects,[31] which also introduces an important increase in computational cost.

Here the question is to know what peculiarities of OPLS-AA might be responsible of the lower α content in amino acid dipeptides.

Other simulations reported in the literature show the influence of the water network structure around the protein molecule in terms of system stabilization.[25,27,29,32,34,56−59] Most of these studies conclude that it is necessary to take into account explicit water molecules to adequately simulate these systems.

We have tried to search differences in the water structure around P_{II} and α_R dipeptide conformations in order to give further data to explain the overstabilization of the P_{II} structures in water. In Figure 3 we present the radial distribution function $g(r)$ of the ASP dipeptide backbone atoms involved in H bonds with the surrounding water molecules. In particular, we present in Figure 3a the radial distribution function of water H atoms around the backbone carbonyl O atoms and in Figure 3b the distribution of O water atoms around the backbone amino H atoms. This dipeptide was selected as an example because it has one of the largest differences between P_{II} and α_R propensities. The radial distribution functions were calculated for the φ,ψ umbrella pairs corresponding to the lowest PMF values in the P_{II} (−90,130) and α_R (−70,−30) basins, respectively. The two plots point in the same direction; that is, the P_{II} conformation shows more probability to form H bonds with the solvent than the α_R one. The largest probabilities of H-bonding between the backbone atoms and water are associated with carbonyl O atoms.

On the other hand, we also calculate the radial distribution function of water H atoms around the carboxylic O atoms of the ASP side chain and found no differences between the P_{II} and α_R conformations in this case (see Figure S2 in the Supporting Information).

The main difference among the usual force fields is the parametrization method followed in each case. For example, the OPLS-AA force field shares with AMBER most of the parameters associated to bonding interactions, such as bonds, bond angles, and dihedrals. These parameters have been obtained by fitting to ab initio calculations, with continuous improvements as more accurate ab initio models were within the reach of available computational resources.[60]

A distinctive feature of the OPLS-AA parametrization is that experimental data were used to get parameters for the nonbonded interaction energy terms, namely, Coulombic and van der Waals energy terms. These terms have an obvious influence in the potential energy when the interaction with the solvent is taken into account and even in the internal nonbonded interaction between the amino acid atoms.

IV. CONCLUSIONS

We have performed an exhaustive exploration of the conformational space of some dipeptides using a distributed umbrella sampling over the peptide torsion angles φ and ψ. We selected the OPLS-AA force field which gives, in general, the lower α conformation content among the commonly used force fields in biomolecular simulations. Our results confirm the above observation in all dipeptides and show a good agreement with recently published propensities obtained by IR and Raman spectroscopy on the same amino acid dipeptides.

The reason for the behavior of the OPLS-AA force field may reside in the parametrization method used for the nonbonded interaction energy terms which is the main difference with other commonly used force fields. OPLS-AA parameters for these terms were obtained by fitting experimental data of some organic liquids which contain groups commonly found in proteic molecules. The OPLS-AA parameters for bonds, bond angles and torsion angles were obtained by fitting to quantum ab initio calculations. This is the method also followed by other force fields.

In principle, it can be assumed that the stability order of amino acid dipeptide conformations that can be obtained with force

dx.doi.org/10.1021/jp206757j |J. Phys. Chem. B 2012, 116, 469–475

fields parametrized according to ab initio models will reproduce those obtained with the QM calculations. That should be the case with force fields such as AMBER and CHARMM, in view of the results reported in the literature. Our results suggest that in the OPLS-AA force field case, the particular origin of the nonbonded interaction energy term parameters makes a difference in the simulation of amino acid dipeptide systems. A new reparameterization of these nonbonded interactions against either high level QM models including explicit water molecules or against new more extensive sets of experimental data will improve the match to experimental propensities in amino acid dipeptides.

■ ASSOCIATED CONTENT

Supporting Information. Twenty-one tables containing the PMF values and the corresponding statistical uncertainties for each amino acid. Figure S1 showing the φ,ψ dihedral angle distribution for an α_R umbrella of the ASP dipeptide. Figure S2 showing the radial distribution function of water H atoms around the carbonyl O atoms for the ASP dipeptide case. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: victor.cruz@iem.cfmac.csic.es.

■ ACKNOWLEDGMENT

Thanks are due to the CICYT (MAT2009-12364 Project) for financial support. Computational support in the use of the distributed computing to the Ibercivis team and to the IBERGRID project 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) Anderson, A. G.; Hermans, J. *Proteins: Struct., Funct., Genet.* **1988**, *3*, 262–265.
(2) Avbelj, F.; Grdadolnik, S. G.; Grdadolnik, J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1272–1277.
(3) Beck, D. A. C.; Alonso, D. O. V.; Inoyama, D.; Daggett, V. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 12259–12264.
(4) Betancourt, M. R. *J. Phys. Chem. B* **2008**, *112*, 5058–5069.
(5) Chin, W.; Piuzzi, F.; Dimicoli, I.; Mons, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1033–1048.
(6) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2001**, *263*, 203–219.
(7) Feig, M. *J. Chem. Theory Comput.* **2008**, *4*, 1555–1564.
(8) Grdadolnik, J.; Grdadolnik, S. G.; Avbelj, F. *J. Phys. Chem. B* **2008**, *112*, 2712–2718.
(9) Jha, A. K.; Colubri, A.; Freed, K. F.; Sosnick, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13099–13104.
(10) Jha, A. K.; Colubri, A.; Zaman, M. H.; Koide, S.; Sosnick, T. R.; Freed, K. F. *Biochemistry (Mosc)* **2005**, *44*, 9691–9702.
(11) Kameda, T.; Takada, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17765–17770.
(12) Kang, Y. K. *J. Phys. Chem. B* **2006**, *110*, 21338–21348.
(13) Ramakrishnan, V.; Ranbhor, R.; Kumar, A.; Durani, S. *J. Phys. Chem. B* **2006**, *110*, 9314–9323.
(14) Ranbhor, R.; Ramakrishnan, V.; Kumar, A.; Durani, S. *Biopolymers* **2006**, *83*, 537–545.
(15) Shi, Z. S.; Chen, K.; Liu, Z. G.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877–1897.
(16) Fitzkee, N. C.; Fleming, P. J.; Rose, G. D. *Proteins: Struct., Funct., Bioinf.* **2005**, *58*, 852–854.
(17) Perskie, L. L.; Rose, G. D. *Protein Sci.* **2010**, *19*, 1127–1136.
(18) Perskie, L. L.; Street, T. O.; Rose, G. D. *Protein Sci.* **2008**, *17*, 1151–1161.
(19) Pizzanelli, S.; Forte, C.; Monti, S.; Zandomeneghi, G.; Hagarman, A.; Measey, T. J.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2010**, *114*, 3965–3978.
(20) Grdadolnik, J.; Mohacek-Grosev, V.; Baldwin, R. L.; Avbelj, F. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 1794–1798.
(21) Hagarman, A.; Measey, T. J.; Mathieu, D.; Schwalbe, H.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2010**, *132*, 540–551.
(22) 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*, 17964–17968.
(23) Ting, D.; Wang, G.; Shapovalov, M.; Mitra, R.; Jordan, M. I.; Dunbrack, R. L., Jr. *PLoS Comp. Biol.* **2010**, *6*, e1000763.
(24) Rata, I. A.; Li, Y. H.; Jakobsson, E. *J. Phys. Chem. B* **2010**, *114*, 1859–1869.
(25) Apostolakis, J.; Ferrara, P.; Caflisch, A. *J. Chem. Phys.* **1999**, *110*, 2099–2108.
(26) Cruz, V.; Ramos, J.; Martinez-Salazar, J. *J. Phys. Chem. B* **2011**, *115*, 4880–4886.
(27) Degtyarenko, I. M.; Jalkanen, K. J.; Gurtovenko, A. A.; Nieminen, R. M. *J. Phys. Chem. B* **2007**, *111*, 4227–4234.
(28) Hu, H.; Elstner, M.; Hermans, J. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 451–463.
(29) Jono, R.; Watanabe, Y.; Shimizu, K.; Terada, T. *J. Comput. Chem.* **2010**, *31*, 1168–1175.
(30) Kim, M.; Choi, S. H.; Kim, J.; Choi, K.; Shin, J. M.; Kang, S. K.; Choi, Y. J.; Jung, D. H. *J. Chem. Inf. Model.* **2009**, *49*, 2528–2536.
(31) Kwac, K.; Lee, K. K.; Han, J. B.; Oh, K. I.; Cho, M. *J. Chem. Phys.* **2008**, *128*, 105106.
(32) Law, P. B.; Daggett, V. *Protein Eng., Des. Sel.* **2010**, *23*, 27–33.
(33) Liu, Z.; Ensing, B.; Moore, P. B. *J. Chem. Theory Comput.* **2011**, *7*, 402–419.
(34) Poon, C. D.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 5642–5643.
(35) Wang, J.; Gu, Y.; Liu, H. *J. Chem. Phys.* **2006**, *125*, 094907.
(36) Gnanakaran, S.; Garcia, A. E. *Proteins: Struct., Funct., Bioinf.* **2005**, *59*, 773–782.
(37) Lee, H. J.; Park, H. M.; Lee, K. B. *J. Theor. Comput. Chem.* **2009**, *8*, 799–811.
(38) Levitt, M.; Hirshberg, M.; Sharon, R.; Daggett, V. *Comput. Phys. Commun.* **1995**, *91*, 215–231.
(39) Seabra, G. D.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. *J. Phys. Chem. A* **2007**, *111*, 5655–5664.
(40) Sun, C. L.; Jiang, X. N.; Wang, C. S. *J. Comput. Chem.* **2009**, *30*, 2567–2575.
(41) Vymetal, J.; Vondrasek, J. *J. Phys. Chem. B* **2010**, *114*, 5632–5642.
(42) Vymetal, J.; Vondrasek, J. *Chem. Phys. Lett.* **2011**, *503*, 301–304.
(43) Verbaro, D.; Ghosh, I.; Nau, W. M.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2010**, *114*, 17201–17208.
(44) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
(45) Jorgensen, W. L.; Tirado-Rives, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6665–6670.
(46) Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J. *J. Phys. Chem. B* **2004**, *108*, 16264–16270.
(47) http://www.ibercivis.es, 2010.
(48) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(49) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(50) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
(51) Lawrence, C. P.; Skinner, J. L. *Chem. Phys. Lett.* **2003**, *372*, 842–847.
(52) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(53) Berendsen, H. J. C.; Postma, J. P. M; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(54) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(55) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(56) Jalkanen, K. J.; Degtyarenko, I. M.; Nieminen, R. M.; Cao, X.; Nafie, L. A.; Zhu, F.; Barron, L. D. *Theor. Chem. Acc.* **2008**, *119*, 191–210.
(57) Lee, M. E.; Lee, S. Y.; Joo, S. W.; Cho, K. H. *J. Phys. Chem. B* **2009**, *113*, 6894–6897.
(58) Mirkin, N. G.; Krimm, S. *Biopolymers* **2009**, *91*, 791–800.
(59) Mullin, J. M.; Gordon, M. S. *J. Phys. Chem. B* **2009**, *113*, 14413–14420.
(60) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.