

## Quantitative Assessment of Force Fields on Both Low-Energy Conformational Basins and Transition-State Regions of the ( $\phi-\psi$ ) Space

Zhiwei Liu,<sup>†</sup> Bernd Ensing,<sup>‡</sup> and Preston B. Moore\*,<sup>†</sup>

*West Center for Computational Chemistry and Drug Design, Department of Chemistry & Biochemistry, University of the Sciences in Philadelphia, 600 South 43rd Street, Philadelphia, Pennsylvania 19104, United States and Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands*

Received July 16, 2010

**Abstract:** The free energy surfaces (FESs) of alanine dipeptide are studied to illustrate a new strategy to assess the performance of classical molecular mechanics force field on the full range of the ( $\phi-\psi$ ) conformational space. The FES is obtained from metadynamics simulations with five commonly used force fields and from ab initio density functional theory calculations in both gas phase and aqueous solution. The FESs obtained at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) level of theory are validated by comparison with previously reported MP2 and LMP2 results as well as with experimentally obtained probability distribution between the C<sub>5</sub>- $\beta$  (or  $\beta$ -PPII) and  $\alpha_R$  states. A quantitative assessment is made for each force field in three conformational basins, LeRI (C<sub>5</sub>- $\beta$ -C<sub>7eq</sub>), LeRII ( $\beta_2$ - $\alpha_R$ ), and LeRIII( $\alpha_L$ -C<sub>7ax</sub>- $\alpha_D$ ) as well as three transition-state regions linking the above conformational basins. The performance of each force field is evaluated in terms of the average free energy of each region in comparison with that of the ab initio results. We quantify how well a force field FES matches the ab initio FES through the calculation of the standard deviation of a free energy difference map between the two FESs. The results indicate that the performance varies largely from region to region or from force field to force field. Although not one force field is able to outperform all others in all conformational areas, the OPLSAA/L force field gives the best performance overall, followed by OPLSAA and AMBER03. For the three top performers, the average free energies differ from the corresponding ab initio values from within the error range (<0.4 kcal/mol) to ~1.5 kcal/mol for the low-energy regions and up to ~2.0 kcal/mol for the transition-state regions. The strategy presented and the results obtained here should be useful for improving the parametrization of force fields targeting both accuracy in the energies of conformers and the transition-state barriers.

### Introduction

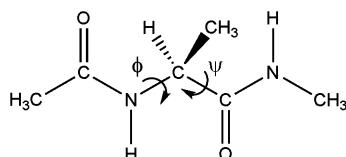
Classical molecular dynamics (MD) simulations have become an indispensable tool to study the structure and dynamics of biological macromolecules, such as proteins, nuclear acids, and lipid assemblies.<sup>1–7</sup> Assuming the validity of the

underlying classical approximation and sufficient sampling of the phase space, the accuracy of a MD calculation primarily depends on the quality of the molecular mechanics (MM) force field employed. Relentless effort to develop and parametrize MM force fields during the past three decades has led to the development of standardized force field libraries with brand names, such as the AMBER,<sup>8</sup> CHARMM,<sup>9</sup> GROMOS,<sup>10</sup> and OPLS models.<sup>11</sup> Considering the somewhat simple standard functional form of additive pair-potentials, (which include many-body effects, such as

\* Corresponding author. E-mail: p.moore@usp.edu. Telephone: 215-596-7537.

<sup>†</sup> University of the Sciences in Philadelphia.

<sup>‡</sup> University of Amsterdam.



**Figure 1.** Structure of Ace-Ala-Nme (alanine dipeptide, AD).

polarization only in an effective manner), these force fields have been shown to reproduce experimentally determined equilibrium structures of biological macromolecules surprisingly well. However, with the constantly advancing computer power, the emphasis of MD simulations has gradually moved from structural prediction to dynamical and kinetic processes, such as protein folding.<sup>12–19</sup> The performance of the standard force fields in these areas of study has not yet been evaluated thoroughly. This is partially because the accuracy of a thermodynamic and/or kinetic process depends not only on a correct force field description of the equilibrium structure but also that of the conformational propensity between different conformers and transition states.

To investigate the conformational dynamics and kinetics of proteins, the alanine dipeptide (Ace-Ala-Nme, AD, Figure 1) molecule has been a standard model system for theoretical<sup>16,20–29</sup> and experimental studies for the past 20 years.<sup>30–36</sup> Experimentally, two-dimensional (2D) infrared (IR),<sup>37</sup> vibrational,<sup>38</sup> Raman,<sup>39</sup> vibrational circular dichroism,<sup>40</sup> and NMR spectroscopy<sup>40</sup> have been used to study the conformational preference of AD and other alanine peptides in aqueous solution. Theoretically, high-level ab initio methods, such as the second-order Møller–Plesset perturbation theory (MP2), have been used to calculate the potential energies of AD across the full range of the ( $\phi$ – $\psi$ ) conformational space, either in gas phase or using the implicit continuum solvent models.<sup>25,41</sup> The energy profile as a function of the backbone dihedral angles  $\phi$  and  $\psi$ , known as the Ramachandran plot, has been used to parametrize and examine MM force fields<sup>9,24,41–43</sup> and in the development and evaluation of new modeling techniques.<sup>44–46</sup> Various enhanced sampling techniques, such as umbrella sampling,<sup>47,48</sup> adiabatic free energy dynamics (AFED)<sup>49</sup> method, and metadynamics,<sup>50–52</sup> have been developed and applied on mapping the ( $\phi$ – $\psi$ ) free energy surface (FES) of AD in both gas phase and aqueous solution. Most of these studies, with a focus on methodology development, have either compared FESs of AD among different force fields<sup>52</sup> or evaluated the force field FESs by high-level ab initio calculations in terms of positions ( $\phi$ ,  $\psi$  values) of energy minima, such as C<sub>7eq</sub>, C<sub>5</sub>,  $\alpha_R$ ,  $\beta$ , etc. and their relative stabilities. Both types of evaluation have limited abilities to give a full spectrum of assessment of the force field FESs. This is especially true for free energy in aqueous solution phase, even though experimental<sup>27,37</sup> and high-level ab initio data have become available.<sup>25</sup> Critically, no emphasis has been made on assessing the accuracy of transition states, which are essential in describing kinetics.

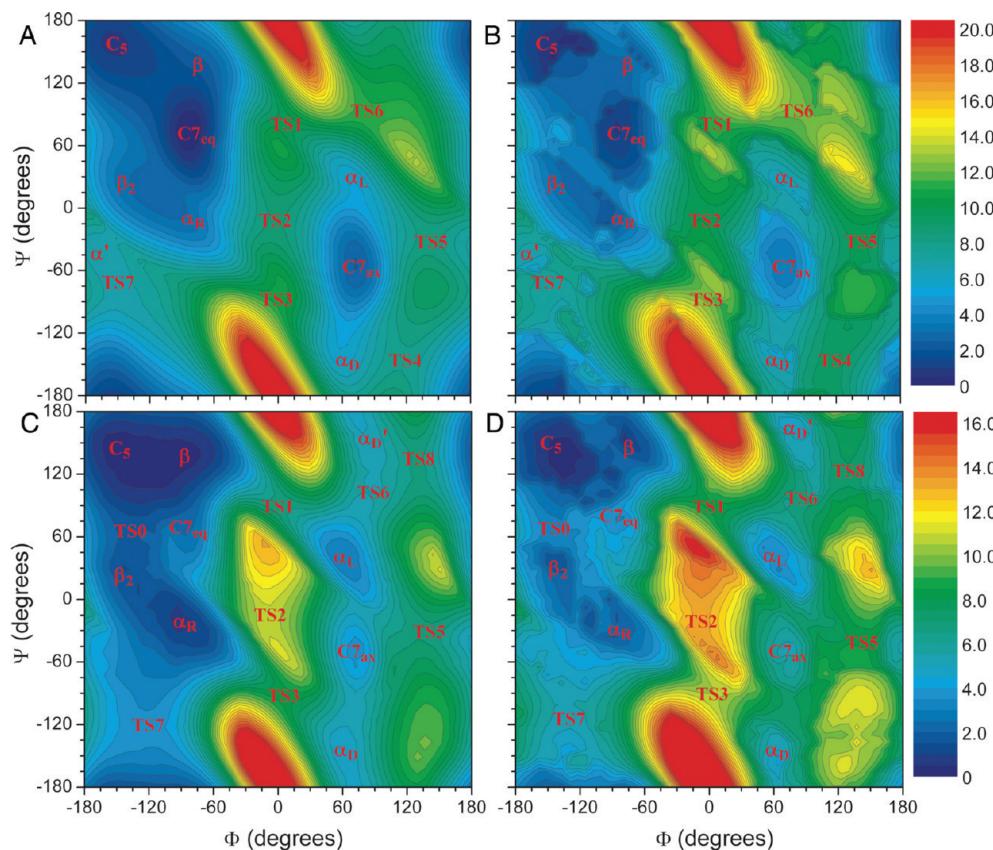
With respect to the dynamics, MD simulations of AD (and other small alanine peptides) up to 20 ns have been carried out using various MM force fields and semiempirical [such as self-consistent charge density functional tight binding

(SCC-DFTB), PM3, and AM1] methods.<sup>24,27,53,54</sup> The performance of the force field has been evaluated by comparing the conformational propensity obtained from the force field and semiempirical MD simulations. However, with 20 ns simulations, only a part of the negative  $\phi$  side of the ( $\phi$ – $\psi$ ) conformational space has been sufficiently sampled. Therefore, assessment of the force fields is made mostly on whether or not a proper ratio of population is achieved between the  $\beta$  and  $\alpha_R$  conformers, not the full ( $\phi$ – $\psi$ ) space.

In the present work, we have applied the metadynamics method,<sup>50,55</sup> in combination with five different force fields, to explore the 2D ( $\phi$ – $\psi$ ) FESs of AD in both gas and aqueous solution phases. Further, we have mapped out the potential energy profiles of AD using ab initio method, specifically the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) combination, on a 50 × 50 grid for both gas and implicit solvent models. The quality of the B3LYP energy surfaces is carefully examined by comparing them with both experimental and high-level ab initio results, such as those using MP2 and LMP2 with fairly large basis sets.<sup>25,56</sup> The free energy corrections to the potential energy profiles are obtained from frequency calculations at each grid point. The assessment of the force field FESs was not done by simple comparison of geometry and energy of each individual minimum- or transition-state structure against the ab initio result. Instead, the 2D ( $\phi$ – $\psi$ ) map was divided into three low-energy conformational basins and three transition-states regions. Quantitative assessment of each force field was carried out by two comparisons between force field and ab initio results: (1) the average free energy of each region; (2) how well a force field FES would match the ab initio FES in each region by calculating the standard deviation of the free energy difference between the two FESs.

## Methods

**Ab Initio Calculations.** Ab initio energy surfaces of AD were computed on a 50 × 50 grid in the ( $\phi$ – $\psi$ ) conformational space in which each dihedral angle had a range of  $-180^\circ$  to  $180^\circ$  with a  $7.2^\circ$  interval. At each grid point, geometry parameters, except the constrained ( $\phi$ – $\psi$ ) dihedral angles, were fully optimized at the B3LYP/6-31G(d,p) level of theory.<sup>57–59</sup> For AD in water, the optimization was performed in a polarized reaction field. The dielectric constant of water and the polarized continuum model (PCM)<sup>60–62</sup> implemented in Gaussian 03<sup>63</sup> were used. All energy minima and transition-states found on the constructed energy profiles were further fully optimized by removing the  $\phi$ – $\psi$  constraints. The potential energy of AD at each grid point, minimum or transition state, was then refined by single point calculation at the B3LYP/6-311+G(2d,p) level of theory. Frequency calculation at the B3LYP/6-31G(d,p) level of theory was carried out on each partially (grid point) or fully optimized structure to characterize the minimum (no imaginary frequency) or the transition state (one imaginary frequency) and to obtain the free energy corrections, including zero point energies (ZPE) (scaling factor of 0.9804)<sup>64</sup> and thermal energy (enthalpy and entropy) corrections at 298.15 K and 1 atm.



**Figure 2.** B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) energy contour maps of AD. (A) Potential energy (relative to  $C_7_{eq}$ ) in gas phase. (B) Free energy (relative to  $C_7_{eq}$ ) in gas phase. (C) Potential energy (relative to  $C_5$ ) in aqueous solution phase. (D) Free energy (relative to  $C_5$ ) in aqueous solution phase.

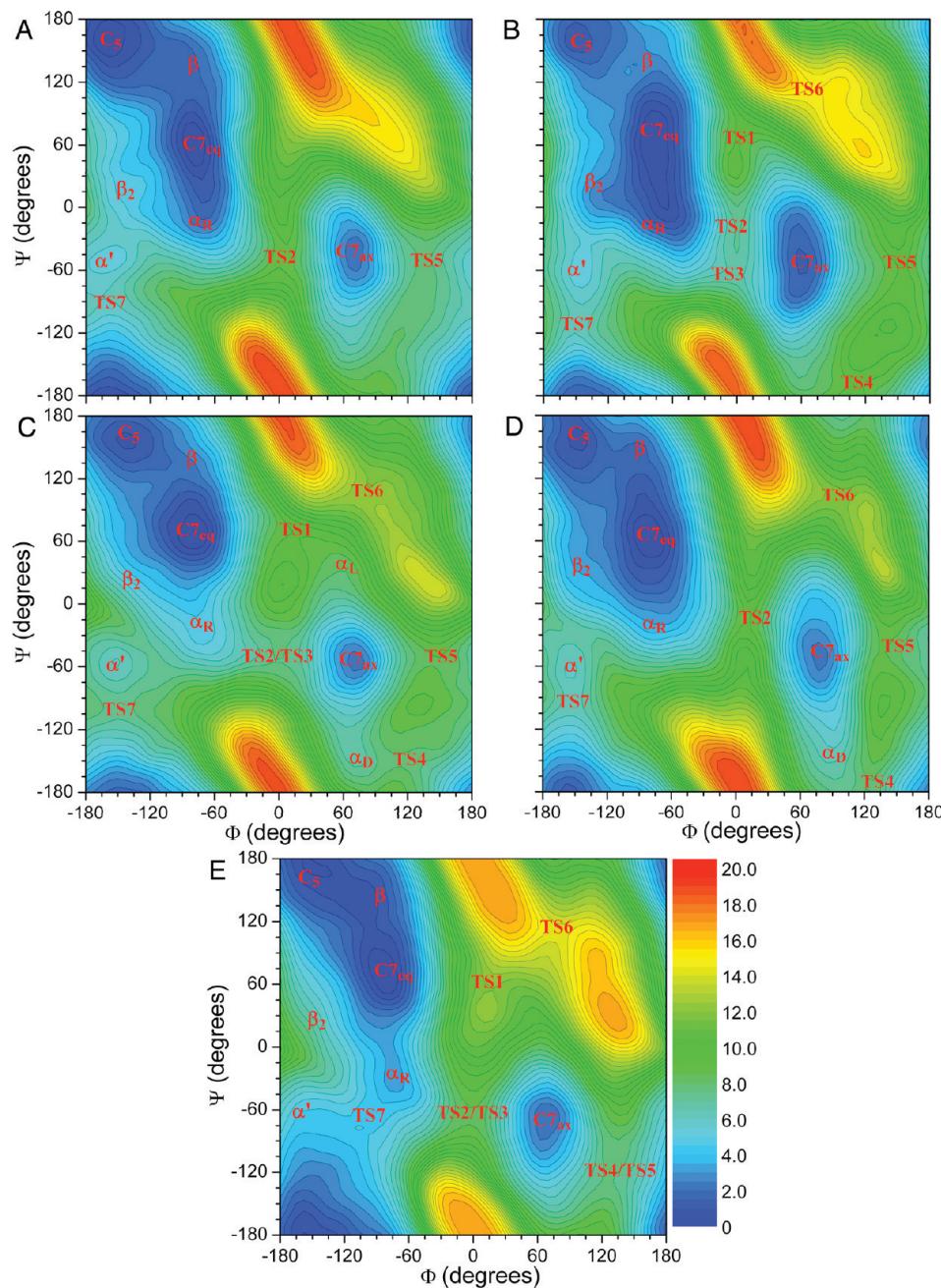
**Metadynamics Simulations.** All metadynamics simulations were carried out using the CM3D program.<sup>65</sup> Five different force fields, CHARMM27,<sup>9,41</sup> AMBER94,<sup>8</sup> AMBER03,<sup>42</sup> OPLSAA,<sup>11</sup> and OPLSAA/L (modified torsion parameters),<sup>43</sup> were used to describe AD and the water molecules, with the exception of using SPC/f for the water model.<sup>66,67</sup> All of these force fields use pair potentials that have been parametrized using a combination of quantum mechanical and experimental data to reproduce the minimum structures of macromolecular molecules, such as proteins, RNA, and DNA. These force fields also have similar functional forms (bonded and nonbonded terms). While there are no gross differences, they do differ in the details of parametrization. For example, AMBER03 is reparameterized using new point charges and backbone torsion parameters from AMBER94 to improve the performance of condensed-phase simulations. Similarly, OPLSAA/L is an improved version of OPLSAA by refitting the backbone torsion parameters. As a result, specific molecular systems will show differences, such as in simulations of membranes<sup>68</sup> or DNA.<sup>69</sup> It should be noted that the cross terms (CMAP) in the CHARMM27 force field were not used to evaluate the force fields within the boundary of the standard functional forms. For the simulation of aqueous solution, an AD molecule was placed in a periodic cubic box ( $L = 18.8 \text{ \AA}$ ) with 216 water molecules. The electrostatic interactions were calculated using Ewald summation,<sup>1,70,71</sup> and the real space cut-off was half of the cell dimension (9.4  $\text{\AA}$ ). Prior to the metadynamics runs, NPT simulation at 1 atm and 298 K for

at least 100 ps was carried out to equilibrate the cell volume. The metadynamics simulations were carried out using the NVT ensemble and a time step of 0.5 fs.

We used two sets of parameters related to the Gaussian potentials (or “hills”):<sup>50,51</sup> (1)  $w = 0.2 \text{ rad}$ ,  $h = 0.02 \text{ kcal/mol}$ , and  $\Delta = 0.02 \text{ rad}$  and (2)  $w = 0.1 \text{ rad}$ ,  $h = 0.02 \text{ kcal/mol}$ , and  $\Delta = 0.075 \text{ rad}$ . Here,  $\Delta$  is the minimum distance that the systems must move in  $(\phi-\psi)$  space before the next Gaussian potential is added. Hills setting (2) thus employs smaller potentials that are also further away from each other than hills setting (1). It should increase the accuracy of the FES, however, it takes longer simulation time to sample the entire conformational space. For AD in gas phase, both settings were employed. For hills setting (1), a 5 ns productive run was carried out to achieve sufficient sampling of the whole  $(\phi-\psi)$  conformational space. For hills setting (2), 20 ns are needed to achieve reasonably sufficient sampling for most of the force fields. The reconstruction of the  $(\phi-\psi)$  FES was done on a  $50 \times 50$  grid, in which the grid interval of  $7.2^\circ$  was comparable with the width of the hills in both settings. The same grid resolution was used throughout the paper so that the energy difference map between any two FESs, which is essential for our quantitative analysis, can be easily calculated.

## Results and Discussions

**Highlights.** Results obtained from this study fall into three subsections: (1) ab initio FESs of AD in gas and aqueous



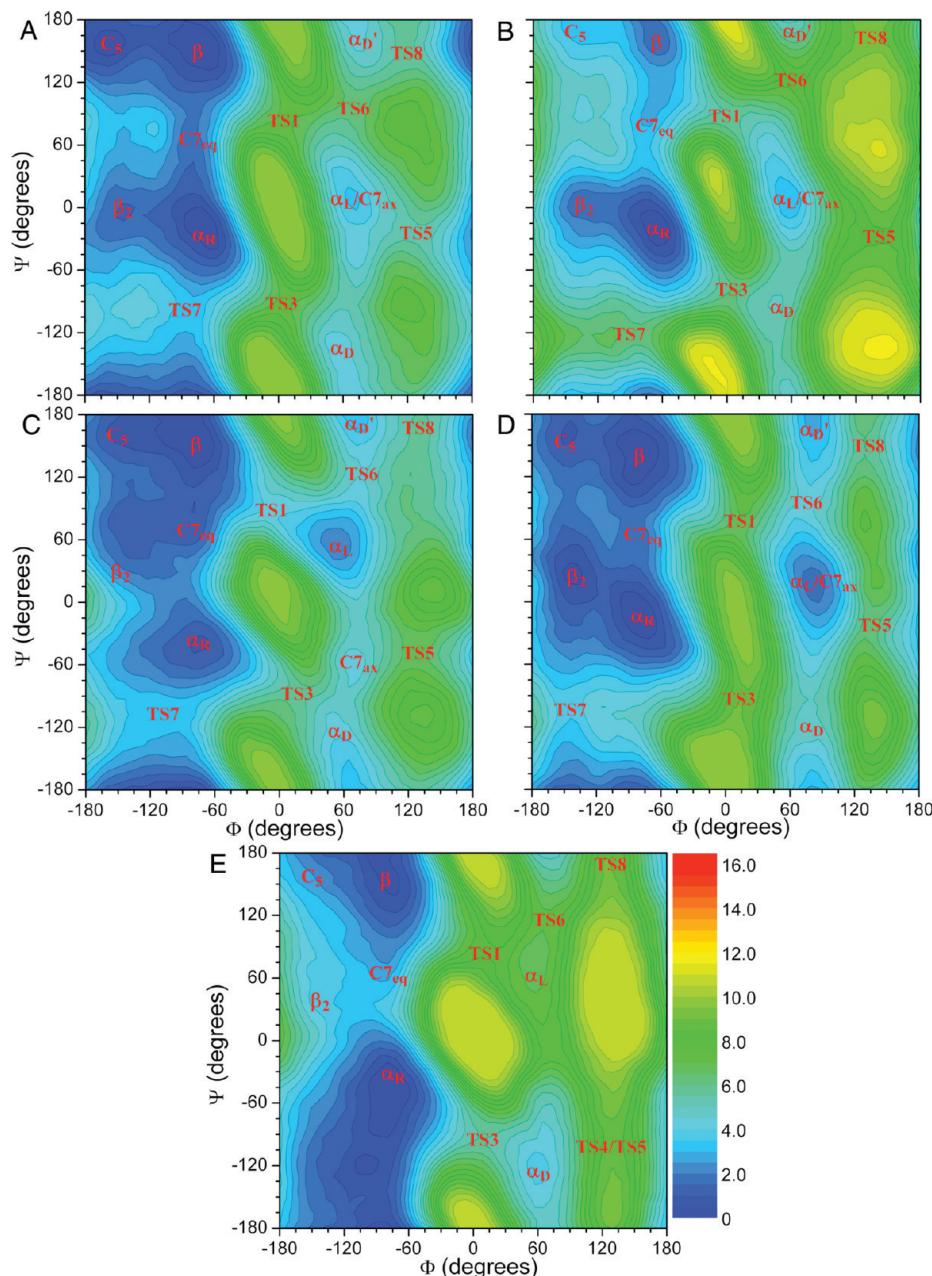
**Figure 3.** Free energy contour maps of AD in gas phase: (A) AMBER03; (B) AMBER94; (C) OPLSAA; (D) OPLSAA/L; and (E) CHARMM27.

phases; (2) FESs of AD obtained by metadynamics simulations using the five force fields; (3) quantitative assessment of the force fields according to the ab initio FESs. Part (3), though comes at the end, is the most important part of this study. Therefore, to give the readers a synopsis of what this study concludes, here we are providing the highlights of our results, which mainly focus on part (3), followed by more detailed discussion of each subsection.

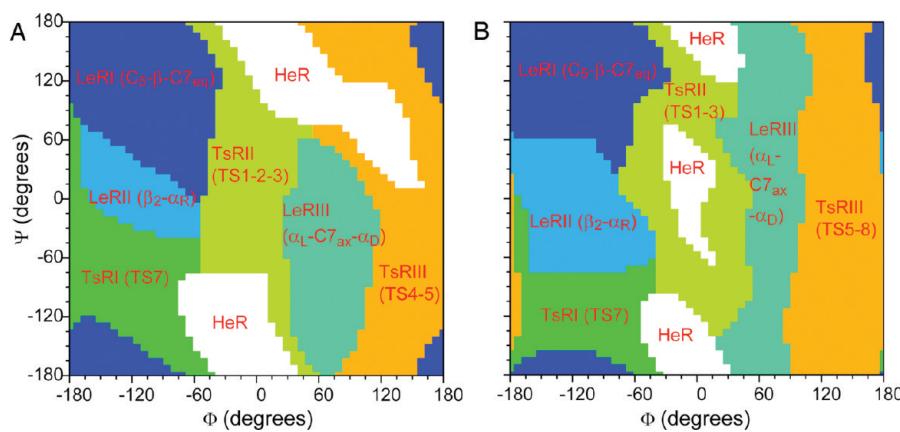
Our study shows that the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) method (with PCM for aqueous phase) is able to generate energy contour maps of AD in gas (Figure 2A) and aqueous (Figure 2C) phases that are as accurate as the previous MP2 approaches.<sup>25,41</sup> The computationally cost-efficient hybrid density functional theory (DFT) method also allowed us to perform frequency calculation and obtain free

energy corrections at each grid point so that the ab initio FESs (Figure 2B and D) can be constructed. Second, metadynamics simulations in combination with several commonly used force fields, namely AMBER94, AMBER03, CHARMM27, OPLSAA, and OPLSAA/L, are used to generate FESs of AD in gas (Figure 3) and aqueous (Figure 4) phases (explicit solvents). The estimated errors of the FESs are 0.2–0.4 kcal/mol for the ab initio FESs and 0.3–0.4 kcal/mol or slightly larger (up to 0.7 kcal/mol) in some cases due to minor insufficient sampling for the force field FESs.

The ab initio FESs are used as standards to assess the accuracy of the MM force field FESs in the low-energy conformational basins and critically for transition states linking the minima. Specifically, the ( $\phi$ – $\psi$ ) conformational space has been divided into six regions (Figure 5). The



**Figure 4.** Free energy contour maps of AD in aqueous solution. (A) AMBER03; (B) AMBER94; (C) OPLSAA; (D) OPLSAA/L; and (E) CHARMM27.



**Figure 5.** Definitions of low-energy (LeR) and transition-state (TsR) regions: (A) gas and (B) aqueous solution phases.

partitions are based on common practice of protein backbone conformational subsets as well as the ab initio FESs. There are three low-energy conformational basins (denoted as LeR), in which LeRI encloses minima  $C_5$  (corresponding to  $\beta$  in some literatures),  $\beta$  (corresponding to PPII in some literatures), and  $C7_{eq}$ . The LeRII basin includes minima  $\beta_2$  and  $\alpha_R$ , and both the LeRI and LeRII regions are on the negative  $\phi$  side. LeRIII is on the positive  $\phi$  side and covers minima  $\alpha_L$ ,  $C7_{ax}$ , and  $\alpha_D$ . There are also three transition-state regions, in which TsRI is the low-lying transition region between LeRII and LeRI. Both TsRII and TsRIII link the negative (LeRI and LeRII) and positive (LeRIII)  $\phi$  sides through rotation of the  $\phi$  dihedral angle clockwise (through 0°) or counterclockwise (through 180°), respectively. In this study, we have measured two quantities (Tables 3 and 4) per region for any FES. The first one is the mean free energy (average free energy of all grid points included in one basin), and the other is the standard deviation of the free energy difference map (FEDM) (obtained by one to one subtraction at each grid point between two FESs) of the designated region between the force field and the ab initio FESs. Comparison of the mean free energy of each basin (relative to the overall mean of all six regions on a FES) between ab initio and force field roughly indicates the accuracy of free energies of a particular group of conformers or transition states. The standard deviation of the FEDM measures how well the two FESs match each other and therefore provides a good indication of whether or not the force field predicts the positions of minima or transition states accurately. The following paragraphs give a brief summary of the assessment for each force field employed. It is worth mentioning that the average free energy of LeRII ( $\beta_2-\alpha_R$ ) is 0.9 kcal/mol higher than that of LeRI ( $C_5-\beta-C7_{eq}$ ) in aqueous solution based on our ab initio FES. This free energy gap is in agreement with a probability partition of 80–20 between the  $C_5-\beta$  and the  $\alpha_R$  states, as reported in recent experimental/computational investigations.<sup>27,38</sup>

For AMBER03, for the gas phase, the mean free energies of most regions agree well with the ab initio values. However, the LeRII ( $\beta_2-\alpha_R$ ) stability is slightly underestimated (0.9–1.2 kcal/mol higher). The energy surfaces match better with the ab initio FES on the negative  $\phi$  side (deviations ~1 kcal/mol) than on the positive  $\phi$  side (deviations ~1.5 kcal/mol). For aqueous solution phase, all three low-energy regions have higher, while all three transition-state regions have lower, mean free energies than ab initio. This means that the energy barriers for transition between conformational basins are 1.2–2.4 kcal/mol underestimated. In terms of the standard deviations, the AMBER03 aqueous FES matches well in all regions with the ab initio FES except for TSRI and TsRII, which have slightly larger deviations (1.5–1.7 kcal/mol). In addition, the aqueous phase FES matches with ab initio slightly better than the gas-phase FES, which reflects the parametrization emphasis of AMBER03 for condensed phase.

For AMBER94, for both gas and solution phases, the LeRI ( $C_5-\beta-C7_{eq}$ ) stability is underestimated. The effect of this underestimation is severe for the aqueous solution phase FES since it has caused a reversed order of stability between LeRI

( $C_5-\beta-C7_{eq}$ ) and LeRII ( $\beta_2-\alpha_R$ ). In addition, the mean free energy is lower for TSRII and higher for TSRIII than ab initio for both gas and aqueous phases, but the differences in energy are bigger in aqueous solution. The energy barriers decrease up to 5.5 kcal/mol for TsRII and increase 1.3 kcal/mol for TsRIII, which will affect any kinetic model built based on the AMBER94 force field.

CHARMM27 underestimates the stability of LeRII ( $\beta_2-\alpha_R$ ) in gas phase but not in aqueous solution and vice versa for LeRIII ( $\alpha_L-C7_{ax}-\alpha_D$ ). It also gives a higher mean free energy for LeRI ( $C_5-\beta-C7_{eq}$ ) and a lower mean for TsRII in aqueous solution, thus lowering the barrier 1.8–2.8 kcal/mol. In addition, CHARMM27 also heavily overestimates the stability of TsRI, i.e., the lower left quadrant of the ( $\phi-\psi$ ) conformational space in both gas and aqueous phases. The standard deviations in all regions are generally the first or second highest among all force fields. These larger deviations from ab initio are probably the result of the CHARMM parametrization procedure, which includes more empirical adjustments to fit with experimental crystallographic data.<sup>9</sup>

The performances of OPLSAA and OPLSAA/L are generally similar, and both have lower standard deviations (0.7–1.5 kcal/mol) in all six regions comparing with other force fields. For both gas and aqueous phases, the OPLSAA mean free energies of LeRII ( $\beta_2-\alpha_R$ ) are higher than the respective ab initio values. OPLSAA/L improves the average free energies of LeRII significantly. However, OPLSAA gives a much better relative stability of LeRI ( $C_5-\beta-C7_{eq}$ ) versus LeRII ( $\beta_2-\alpha_R$ ) due to similar underestimation of stabilities of both regions for the aqueous phase. Conversely, the improvement of OPLSAA/L in LeRII ( $\beta_2-\alpha_R$ ), in combination with the underestimation of stability of LeRI, leads to too small of an energy difference between LeRI and LeRII for aqueous phase. Both force fields also have lower mean free energies for TsRII, which leads to a 3.0 kcal/mol decrease in barrier for OPLSAA and a 2.4 kcal/mol for OPLSAA/L.

In summary, in our opinion, OPLSAA/L gives the best performance overall, followed by OPLSAA and AMBER03. As recently pointed out by Feig,<sup>72</sup> force fields parametrized based on AD can accurately reflect amino acid backbone torsional preferences when used in MD simulations of proteins. Therefore, the quantitative assessment and strategy presented here can be used for improving force field parametrization targeting not only the accuracy of energies of conformers but also transition-state barriers. It should be noted that the torsional profiles of AD obtained should not be directly used to represent the torsional preference of amino acids in protein structures.<sup>72</sup>

**Free Energy Surfaces of AD by Ab Initio Calculations.** Previously, various ab initio calculations at different levels of theory have been used to study AD in gas phase, water, and/or other medium, in order to gain insights into the conformational preferences of protein backbones and/or in assisting parametrization of MM force fields.<sup>41,56,73,74</sup> Particularly, Wang et al. have used the MP2/cc-pVTZ//MP2/6-31G(d,p) level of theory, in combination with the PCM model, for solvation effects to obtain fully relaxed ( $\phi-\psi$ ) potential energy surfaces of AD.<sup>25</sup> In addition to fully relaxed

**Table 1.** Optimized Geometries ( $\phi$  and  $\psi$  in  $^\circ$ ) and Relative and Free Energies ( $\Delta E$  and  $\Delta G$  in kcal/mol) of Stationary Points of AD in a Gas Phase, Obtained at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) (this work) and MP2/cc-pVTZ//MP2/6-31G(d,p)<sup>25</sup> Levels of Theory

B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p)				MP2/cc-pVTZ//MP2/6-31G(d,p)					
	$\phi$	$\psi$	$\Delta E$	$\Delta G$		$\phi$	$\psi$	$\Delta E^c$	$\Delta G$
C7 <sub>eq</sub>	-83.1	72.6	0.00	0.00		-82.0	80.6	0.0	0.00
C <sub>5</sub>	-158.4	164.6	0.92	0.06		-159.7	159.3	1.47 (0.91–1.11)	0.67
$\alpha_R$	-80.0	-20.0	3.37 <sup>a</sup>	2.28		-80.0	-20.0	3.27	3.74
$\alpha_L$	68.4	26.5	5.48	5.19		63.2	35.4	4.52 (4.36–5.19)	4.99
C7 <sub>ax</sub>	73.6	-57.7	2.48	2.63		75.8	-62.8	2.50 (2.06–2.48)	2.32
$\beta_2$	-125.7	21.6	2.72	1.87		-141.6	23.8	3.25 (2.51–2.84)	2.63
$\alpha'$	-169.9	-39.2	6.59	5.88		-166.1	-36.7	6.07 (5.49)	6.15
$\alpha_D$	59.8	-136.2	5.53 <sup>b</sup>	4.55		53.0	-133.4	4.75	5.24
TS1	5.6	81.4	9.72	10.31		4.4	84.3	8.94	
TS2	-1.4	-8.9	8.64	8.94		-0.2	-26.2	9.37	
TS3	2.8	-77.3	10.68 <sup>b</sup>	10.39		4.4	-85.5	10.20	
TS4	112.8	-146.7	7.93	8.24		115.4	-151.9	8.44	
TS5	135.9	-26.2	8.17	8.68		135.0	-25.3	8.12	
TS6	79.0	86.4	11.23	11.04		75.2	90.7	11.20	
TS7	-149.8	-87.3	6.80	6.27					

<sup>a</sup> Partially optimized by constraining ( $\phi, \psi$ ) at (-80.0°, -20.0°). <sup>b</sup> Located by a series of partial optimization. <sup>c</sup> Values in parentheses are from ref 56 at levels ranging from LMP2/cc-pVTZ(-f)//MP2/6-31G\* level to LMP2/cc-VQZ(-g)//MP2/6-311++G\*\*.

energy surfaces in the gas phase, ether and water, they have also optimized and characterized all energy minima and transition states as well as calculated free energy corrections of energy minima. Mackerell et al. have constructed the potential energy surface of AD in gas phase at a similar level of theory.<sup>41</sup> In addition, they and others have also optimized a few energy minima in the gas phase using levels up to MP2/6-311++G(d,p). In this paper, we have utilized a more time-efficient hybrid DFT method, specifically, the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) method, in combination with the PCM model for solvation effects. Our energy contour maps were obtained in a similar fashion with a finer grid. Furthermore, we have also estimated the free energy corrections to the energy contour maps. Finally, we optimized and characterized all energy minima and transition states. The comparison with the previous results will not only give us an estimate on the accuracy of the more time-efficient DFT methods but also test the robustness of the approach.

Figure 2A and B shows the potential energy surface and the FES of AD in the gas phase. The general features of the potential energy map are very similar to an energy map calculated at the MP2/cc-pVTZ//MP2/6-31G(d,p) level of theory.<sup>25</sup> The positions and numbers of energy minima and transition states on the two energy maps coincide with each other. Particularly, both the MP2 and DFT energy map indicate a transition state (TS2) at around (0°, 0°), which is different from the HF map where only the energy maximum is observed between TS1 and TS3. This has indicated that the DFT method is able to describe the electrostatic interactions between the C=O and N–H groups in AD, when compared to the HF method. Full geometry optimization at the B3LYP/6-31G(d,p) level has located energy minima C7<sub>eq</sub>, C<sub>5</sub>, C7<sub>ax</sub>,  $\beta_2$ ,  $\alpha_L$ , and  $\alpha'$  as well as transition-states TS1, TS2, and TS4–TS7. Similar to all previous theoretical studies, minimum  $\alpha_R$  cannot be located in gas phases, therefore optimization with ( $\phi$ – $\psi$ ) constrained at (-80°, -20°) is carried out. For  $\alpha_D$  and TS3, no stationary points can be definitely located at the B3LYP/6-31G(d,p) level because the energy derivative with respect to the  $\psi$  dihedral angle

does not converge to zero. However, we have located the best approximation of  $\alpha_D$  and TS3 by a series of partial optimizations (detailed in Supporting Information) as (59.8°, -136.2°) and (2.8°, -77.3°), respectively. The energy gradients with respect to  $\psi$  are estimated at less than 0.008 kcal/mol deg.

As shown in Table 1, the optimized structures, potential and free energies of minima, and transition states agree fairly well with previous theoretical results. The relative potential energy order of all energy minima C7<sub>eq</sub>, C<sub>5</sub>, C7<sub>ax</sub>,  $\beta_2$ ,  $\alpha_R$ ,  $\alpha_L$ ,  $\alpha_D$ , and  $\alpha'$  is the same between the DFT and MP2 methods. Quantitatively, the differences in optimized dihedral angles  $\phi$  and  $\psi$  between the two methods are mostly about 0–9° with two exceptions ( $\beta_2$  and TS2) at around 15–17°. This is understandable since both minimum  $\beta_2$  and transition state TS2 lie on the flat regions of the energy landscape. In fact, theoretical studies from different sources have shown significant discrepancy between optimizations using the MP2 method with different basis sets. Specifically, the  $\beta_2$  conformer was optimized at (-141.6°, 23.8°) with MP2/6-31G(d,p),<sup>25</sup> (-125.7°, 21.6°) with B3LYP/6-31G(d,p), and (-90.7°, -7.8°) with MP2/6-311++G(d,p).<sup>41</sup> Interestingly, the optimized  $\phi$  angle value determined by the DFT method is closer to that from the MP2/6-311++G(d,p) level with a much larger basis set than the MP2/6-31G(d,p) value with the same basis sets. The results indicate that the sizes of the basis sets (perhaps the addition of the diffusive functions) affect the optimized geometry more than the difference in methodology. Therefore, we are confident that the DFT method performance is at least as good as the MP2 method in terms of geometry optimization.

The differences in relative potential energies (C7<sub>eq</sub> as reference zero) of  $\alpha_R$ , C7<sub>ax</sub>, TS5, and TS6 are within 0.1 kcal/mol between the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) and MP2/cc-pVTZ//MP2/6-31G(d,p) methods. The relative energies of C<sub>5</sub>,  $\beta_2$ , TS2, and TS4 from the DFT calculations are ~0.5–0.7 kcal/mol less than that of the MP2 results. While the relative energies of  $\alpha_L$ ,  $\alpha'$ ,  $\alpha_D$ , TS1, and TS3 from the DFT calculations are ~0.5–0.9 kcal/mol more

than that of the MP2 results. Generally, the DFT method predicts more stable extended structures ( $C_5$ ,  $\beta_2$ ) and less stable compact conformers ( $\alpha_L$ ,  $\alpha_D$ ) than the MP2 method. It has been recognized that the MP2/cc-pVTZ//MP2/6-31G\*\* may artificially stabilize the compact conformers with respect to the extended ones due to intramolecular basis set superposition errors (BSSE).<sup>56</sup> The current energies calculated at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) could be more accurate than that of the MP2/cc-pVTZ//MP2/6-31G(d,p) level. Furthermore, Table 1 has also listed the relative energies of  $C_5$ ,  $\alpha_L$ ,  $\beta_2$ ,  $C_{7\text{ax}}$ , and  $\alpha'$  with respect to  $C_{7\text{eq}}$  at the levels up to LMP2/cc-pVQZ(-g)//MP2/6-311++G(d,p).<sup>56</sup> Since the local MP2 (LMP2) method can circumvent the BSSE issue, the fact that the LMP2 energies agree better with the DFT method (on  $C_5$ ,  $\beta_2$ , and in one case  $\alpha_L$ ) further confirms the above conclusion.

The free energies of all energy minima also agree fairly well, with a difference generally smaller than 0.7 kcal/mol, between the DFT and MP2 methods, with an exception of conformer  $\alpha_R$  which has a free energies difference of 1.5 kcal/mol between the two methods. This difference, however, has no indication on the accuracy of the employed DFT method based on the fact that Wang et al. have estimated the  $\alpha_R$  free energy corrections using the  $\alpha_L$  conformer, while we have calculated them using the  $\alpha_R$  conformer. Apart from the  $\alpha_R$  conformer, the DFT method also tends to produce more stable extended structures ( $C_5$ ,  $\beta_2$ ,  $\alpha'$ ,  $\alpha_D$ ) and less stable compact conformers ( $\alpha_L$ ,  $C_{7\text{ax}}$ ) in terms of free energies relative to  $C_{7\text{eq}}$ , when compared to the MP2 method. The contributions to this trend come from either the potential energies ( $C_5$ ,  $\beta_2$ , and  $\alpha_L$ ) or the free energy corrections ( $\alpha'$ ,  $\alpha_D$ , and  $C_{7\text{ax}}$ ). The free energy of the  $C_5$  conformer is 0.06 kcal/mol higher than the  $C_{7\text{eq}}$  conformer at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) level and 0.67 kcal/mol at the MP2/cc-PVTZ//MP2/6-31G(d,p) level. Both results agree well with the experimental observation that  $C_5$  and  $C_{7\text{eq}}$  were the dominant species at room temperature in the  $\text{CCl}_4$  solution and at low temperature in Ar matrices.<sup>75</sup>

Although the discrepancies between the two approaches are not large enough to raise serious concerns, we do want to note two facts that favor the approach in the current paper. One is that the frequency calculations and the geometry optimizations in this paper are done at the same B3LYP/6-31G(d,p) level of theory, while in the previous case, HF/6-31G(d) was used to optimize the structures and calculate the frequencies, and the free energy corrections obtained were then applied on optimized structures at the MP2/6-31G(d,p) level of theory. The second is that the current calculations are performed with tight optimization criteria and an ultra fine grid for integration.

The FES of AD in gas phase, as shown in Figure 2B, is constructed on the basis of potential energy map plus free energy corrections, including zero point energies and thermal energy corrections based on frequency calculations. The accuracy of the free energy corrections is subject to the magnitudes of the remaining nonzero first derivatives (force) of the optimized structures at each grid point. Generally, there are two types of errors. One behaves more like noises since it comes from the uneven qualities of the optimized structures

at each grid point, i.e., differences in remaining forces below the convergence criteria. Other sources of inaccuracy, such as vibrational–rotational coupling and hindered rotors, may also contribute to the noise of the FES. We have applied a linear smoothing function to even out noises. Specifically each grid point has only 50% contribution from the frequency calculation at this grid point, the other 50% comes evenly from its four neighboring grid points ( $\phi - 7.2^\circ$ ,  $\phi + 7.2^\circ$ ,  $\psi - 7.2^\circ$ , and  $\psi + 7.2^\circ$ ). This technique works quite well based on the comparison of the FESs before and after the smoothing. The second type of error is more systematic since it comes from the remaining forces (nonharmonic character outside the minima or transition states) caused by the constraints of the two dihedral angles. We have also analyzed the behaviors and the magnitudes of the second type of systematic errors. Naturally, the free energy corrections are most accurate around the true stationary points, i.e., minima and transition states. The frequency calculations lead to underestimation or overestimation of the free energy corrections in the area that extends out from a minimum or a transition state, respectively. By analyzing the trends of free energy corrections along the paths from minimum to transition state, we can estimate that the maximum error would be around 0.6–0.7 kcal/mol, while the average error would be around 0.2–0.4 kcal/mol.

The overall features of the FES are similar to those of the potential energy surface as expected. Although the free energy corrections do not alter the positions of energy minima and transition states, they do slightly increase the slopes of the energy profiles, which results in more distinguished local energy minima, such as the separation of the  $\beta_2$ – $\alpha_R$  area from the  $C_5$ – $\beta$ – $C_{7\text{eq}}$  area or  $\alpha_L$  and  $\alpha_D$  from the  $C_{7\text{ax}}$  center. The average free energy correction to all the transition states is 0.7 kcal/mol higher than the average correction to all the minima, as also shown in Table 1.

The fully relaxed energy maps without and with free energy corrections of AD in water, obtained at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) level with the PCM model, are shown in Figure 2C and D, respectively. The energy map in Figure 2C includes the potential energy of solute and polarized solute–solvent (PS-S) interaction energy, while the free energy corrections include zero point energies and thermal (enthalpy and entropy) corrections to the solute from the frequency calculation as well as the nonelectrostatic energy term, i.e., cavitation, dispersion and repulsion energies, calculated by the PCM model. The terms, energy maps without and with free energy corrections, are used in the following discussion with regard to Figure 2C and D, respectively.

The topological features of the energy map and changes from the gas-phase energy map to the aqueous map are almost identical to that obtained at the MP2/cc-pVTZ//MP2/6-31G(d,p) level with PCM model for solvation effects.<sup>25</sup> The changes include: (1) expansion of low-energy regions, more accessible and flatter energy surface (note the different contour color scale for gas (0–20 kcal/mol) and solution (0–16 kcal/mol) phases in Figure 2); (2) diminishing of the gas phase global minimum  $C_{7\text{eq}}$  and emerging of the new minima  $\alpha_R$  and  $\beta$ ; (3) energy barrier (TS0) between  $C_5$ – $\beta$

**Table 2.** Optimized Geometries ( $\phi$  and  $\psi$  in  $^\circ$ ) and Relative and Free Energies ( $\Delta E$  and  $\Delta G$  in kcal/mol) of Stationary Points of AD in Water, Obtained at the B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) (this work) and MP2/cc-pVTZ//MP2/6-31G(d,p)<sup>25</sup> Levels of Theory

	B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p)				MP2/cc-pVTZ//MP2/6-31G(d,p)			
	$\phi$	$\psi$	$\Delta E$	$\Delta G$	$\phi$	$\psi$	$\Delta E$	$\Delta G^b$
C7 <sub>eq</sub>	-85.4	73.4	2.06	3.09	-86.3	90.1	0.92	2.42
C <sub>5</sub>	-151.6	147.6	0.00	0.00	-156.4	143.8	0.00	0.00
$\alpha_R$	-78.1	-27.2	0.84	1.60	-70.5	-32.1	0.08	0.94
$\beta$	-75.1	143.3	0.26 <sup>a</sup>	0.06	-64.0	142.1	0.17	0.39
$\alpha_L$	61.3	40.9	2.84	3.30	59.4	41.4	1.27	1.67
C7 <sub>ax</sub>	73.4	-53.0	3.48	4.32	74.9	-54.3	2.69	4.00
$\beta_2$	-138.5	27.3	1.36	1.37	-145.6	27.2	1.27	1.57
$\alpha_D$	60.1	-147.7	4.43	4.30	55.6	-144.9	3.59	3.57
$\alpha_{D'}$	72.8	164.8	4.79	4.01	60.5	-170.9	4.08	3.68
TS0	-129.8	62.6	1.74	2.63	-143.1	70.9	1.80	
TS1	0.3	91.6	6.41	7.88	0.1	91.4	5.83	
TS2	-11.0	-11.7	10.55	12.52	-6.6	-29.6	10.70	
TS3	7.3	-92.1	8.08	9.99	7.5	-92.2	7.50	
TS5	132.2	-28.1	6.58	7.78	130.4	-28.7	6.51	
TS6	81.3	104.2	5.28	6.45	76.4	104.5	6.30	
TS7	127.9	133.7	6.45	7.09				
TS8	-114.0	-115.6	3.44	4.99				

<sup>a</sup> Located by a series of partial optimization. <sup>b</sup> Free energies corrections were made at the HF/6-31G\* level.

and  $\alpha_R$ - $\beta_2$  regions; and (4) the shift of dominant region from C7<sub>ax</sub> to  $\alpha_L$ . The optimized geometries and relative energies of all energy minima and transition states of AD in water are also shown in Table 2. The agreement on both the optimized ( $\phi$ ,  $\psi$ ) angles and energies is again reasonable, indicating the robustness of both ab initio treatments. The topological feature changes from the gas phase to the aqueous map bring the aqueous map into closer agreement with the experimental Ramachandran plot derived from experimental protein structures.<sup>76</sup> The results indicate that the PCM approach gives good description of the bulk solvent polarization effects despite the lack of specific interactions between the solute and solvent molecules.

The discrepancies in optimized geometries between DFT and MP2, however, are slightly larger than that from the gas-phase calculations. This directly results from more extended and flatter low-energy regions in the aqueous energy surface. There are five stationary points (C7<sub>eq</sub>,  $\beta$ ,  $\alpha_{D'}$ , TS0, and TS2), instead of the two in the gas phase, whose optimized  $\phi$  or  $\psi$  angles differ by more than  $10^\circ$  but no more than  $25^\circ$  between the DFT and MP2 results. For relative energies (C<sub>5</sub> as reference) without the corrections, the two methods agree very well (within 0.1 kcal/mol) on extended structures like  $\beta$  and  $\beta_2$ . However, for more compact structures, such as C7<sub>eq</sub>,  $\alpha_L$ ,  $\alpha_R$ , and C7<sub>ax</sub>, the DFT method gives higher energies (relative to C<sub>5</sub>), roughly 0.8–1.5 kcal/mol higher than that from the MP2 method. The discrepancies are systematic since relative stabilities (e.g., relative energies to C7<sub>eq</sub>) among the more compact structures are similar (difference 0.3–0.4 kcal/mol) between the two methods. This is consistent with what we have observed in the gas-phase calculations, which states that BSSE errors at the MP2/cc-pVTZ//MP2/6-31G(d,p) level might have artificially stabilized more compact structures. For energy without corrections, both methods predict C<sub>5</sub> as the global minimum. The orders of energies for the minima are not exactly the same but very similar. Due to the stabilization on the extended structures by the DFT method, the  $\beta$  and  $\beta_2$

conformers switch their places with their respective neighbors, the  $\alpha_R$  and C7<sub>eq</sub> conformers, in the energy order (low to high) C<sub>5</sub>- $\alpha_R$ - $\beta$ -C7<sub>eq</sub>- $\beta_2$ - $\alpha_L$ -C7<sub>ax</sub>- $\alpha_D$ - $\alpha_{D'}$  from the MP2 method. As a result, the energy order of C<sub>5</sub>- $\beta$ - $\alpha_R$ - $\beta_2$ -C7<sub>eq</sub>- $\alpha_L$ -C7<sub>ax</sub>- $\alpha_D$ - $\alpha_{D'}$  is obtained at the level of B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p).

Similar to the MP2 study,<sup>25</sup> the free energy corrections have also changed the order of energy minima slightly in the current DFT study. The overall agreement between the DFT and MP2 free energies is reasonable. Both methods give the C<sub>5</sub> conformer as the global minimum, with the  $\beta$  conformer with slightly higher free energy (0.39 for MP2 and 0.06 kcal/mol for DFT). Two other conformers,  $\alpha_R$  and  $\beta_2$ , follow C<sub>5</sub>- $\beta$  with relative free energies roughly 1–1.6 kcal/mol higher than that of the C<sub>5</sub>. Both methods partially agree with the experimental observations that the  $\beta$  and  $\alpha_R$  conformers are the dominant species in water.<sup>27,32,37,38,40</sup> However, the lack of experimental evidence on the existence of conformers C<sub>5</sub> and  $\beta_2$  does not explain the theoretical results in which they have very similar energies with  $\beta$  and  $\alpha_R$ , respectively. As indicated by Wang et al., the lack of explicit interactions of AD with water molecules as a result of using the PCM model may attribute to this disagreement. They have also speculated two intermediate conformers with explicit hydrogen-bonded water molecules whose geometries lie between C<sub>5</sub> and  $\beta$  and  $\alpha_R$  and  $\beta_2$ , respectively. These intermediate conformers are believed to be the dominant species  $\beta$  and  $\alpha_R$  observed in experiments.

The differences in the relative free energies between MP2 and DFT are similar to the differences in the relative energies before the free energy corrections, indicating inheritance of the problem. That is, there is a general trend of more stable extended structures and less stable compact structures from the DFT method than the MP2 methods. Another significant discrepancy is that the free energy (relative to C<sub>5</sub>) of the  $\alpha_L$  conformer in the DFT results is about 1.63 kcal/mol higher than that from the MP2 results. This energy difference comes from energy without corrections (1.57). There is no direct

experimental or theoretical evidence to determine which set of results is closer to the real value. However, from the MP2 results, the free energy of  $\alpha_L$  is just 0.7 kcal/mol higher than the  $\alpha_R$  conformer. This small difference should have resulted in a population of the  $\alpha_L$  conformer being observed experimentally. Therefore, a much larger energy gap of 1.7 kcal/mol between the  $\alpha_L$  and  $\alpha_R$  conformers from the DFT calculations fits better with the experimental results where only the population of  $\alpha_R$  and no  $\alpha_L$  is observed.<sup>27,37</sup>

The free energy corrections for the aqueous energy map have been subjected to the aforementioned linear smoothing technique. Analysis also has shown that the systematic errors caused by constraints of dihedral angles in optimization are similar to those of the gas-phase free energy corrections in terms of behaviors and magnitudes. As shown in Figure 2D, the aqueous FES is slightly nosier than that of the gas-phase FES. This is most likely because the convergence criteria used for partial geometry optimization of AD in aqueous phase at each grid point is slightly larger than that used for gas-phase optimization. The remaining forces will lead to larger noises in frequency calculations and therefore thermal corrections. Similar to the gas-phase behavior, the free energy corrections tend to raise the slope of the energy profiles. The increase is even greater in the aqueous phase than in the gas phase. The average free energy correction to the transition state is 1.1 kcal/mol higher than the average to the minima, as comparing to 0.7 kcal/mol in gas phase.

In a brief summary, the current B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) with PCM model approach was able to generate energy contour maps of AD in the gas phase and in water that is at least as accurate as the previous MP2 approaches. The cost-efficient hybrid DFT method has also allowed us to generate the energy contour maps using a finer grid as well as performing frequency calculation at each grid point to generate free energy corrections to the map. The validated ab initio FESs will be used as standards to access the accuracy of FESs generated by metadynamics simulations in combination with several commonly used force fields, namely AMBER94, AMBER03, CHARMM27, and OPLSAA.

**Free Energy Surfaces of AD by Metadynamics Simulations using MM Force Fields.** Figure 3 displays FESs of AD in a gas phase, and Figure 4 exhibits FESs of AD in an aqueous solution. In this section, we will visually examine and compare the general features of the force field FESs against that of the ab initio FESs to provide a qualitatively assessment of different force fields and the metadynamics method.

The gas-phase FESs are obtained using both hills setting (1) (Figure 3A) and (2) (Figure S1, Supporting Information). Comparison of the gas-phase FESs from two hills setting shows that using the same force field but different hill sizes produces almost identical FESs for each force field. This indicates the robustness of the metadynamics method to hills settings. The efficiency of the metadynamics simulations in sampling conformational space of the selected variables is shown in Figure 3, in which the contour lines cover up almost all conformational space except a very small high-energy area at around  $\phi = 0^\circ$  and  $\psi = 180^\circ$ . The excellent coverage indicates a good conformational sampling in just 5 ns using

hills setting (1). The dependency of the sampling efficiency on the metadynamics (hills) parameters can be analyzed based on the difference in conformational coverage. The metadynamics simulations using hills setting (2) add Gaussian potentials (hills) with half of the width and less frequently when compared with hills setting (1). Therefore, even with a 20 ns of simulation, the uncovered area is still larger than that of the FES from hills setting (1) with 5 ns. However, the blank areas are high-energy hills and generally do not affect the minima and important transition states. The unsampled areas in the aqueous solution phase FESs (Figure 4) are much less than that of the gas phase (using hills setting (2)). The result indicates that solvation effect has made the free energy surface much more accessible, which is consistent with what we observed in the ab initio calculations.

The general shapes of all force field gas-phase FESs are similar to that of the ab initio FES (Figure 2B). All force field FESs capture the major minima  $C_5$ ,  $C_{7_{eq}}$ , and  $C_{7_{ax}}$ . Their positions in terms of the values of  $(\phi, \psi)$  are also in good agreement with that of the ab initio FES. However, further inspection of Figure 3 indicates that there are characteristic differences in details of the gas-phase FESs of different force fields and the ab initio FES. For example, most force field FESs do not have local minima for conformers  $\alpha_L$  and  $\alpha_D$ , whose positions are clearly defined in the ab initio FES. On the OPLSAA, OPLSAA/L, or the AMBER94 FES, there is a small relatively flat region below the  $C_{7_{ax}}$  minimum that can be classified as the  $\alpha_D$  conformer. For  $\alpha_L$ , a similar flat region above  $C_{7_{ax}}$  is seen only on the OPLSAA FES. Second, on the AMBER03, OPLSAA, and CHARMM27 force field FESs, the stability of the  $\beta_2$  conformer area is underestimated. It should be noted that Mackerell et al. have noticed this underestimation and have added  $\phi, \psi$  dihedral cross terms (CMAP) to the CHARMM force field to improve the energetics of the  $\beta_2$  conformer.<sup>41</sup> However, the current FES is calculated without the cross terms to evaluate the force fields within the boundary of the standard functional forms. Another significant difference between the force field and ab initio FES is that the CHARMM27 force field has overestimated the stabilities of the conformational space at the lower left quadrant of the FES, i.e., the area around  $(-120^\circ, -120^\circ)$ .

The rest of the transition states fall into two major groups that connect the low-energy regions ( $LeRI$  and  $LeRII$ ,  $C_5-C_{7_{eq}}-\beta_2-\alpha_R$ ) on the negative  $\phi$  side, to another region on the positive  $\phi$  side ( $LeRIII$ ,  $\alpha_L-C_{7_{ax}}-\alpha_D$ ) through rotation of the dihedral angle  $\phi$  clockwise and counterclockwise, respectively. One group includes TS1–TS3 which have  $\phi$  roughly at  $0^\circ$ . As shown in Figure 3, only the AMBER94 force field FES features all three transition states with TS3 shifted slightly upward (larger  $\psi$ ). The OPLSAA and CHARMM27 FESs feature TS1 and another transition state (TS2/TS3) which positions somewhere between the ab initio TS2 and TS3. The AMBER03 and OPLSAA/L force fields, on the other hand, only show TS2, which also shifted slightly downward (smaller  $\psi$ ) in the AMBER03 FES. The other group includes TS4 and TS5 whose  $\phi$  angle values are roughly at  $120^\circ$ – $150^\circ$ . Both transition states are featured in the FESs of three force fields, AMBER94, OPLSAA, and

OPLSAA/L, while only TS5 is shown by AMBER03, and a combined TS4/TS5 is shown by CHARMM27. The last transition state TS6 is of less importance due to its higher free energy ( $\sim 11$  kcal/mol relative to  $C_{7\text{eq}}$  by DFT). Generally, all force fields have predicted its position reasonably well but overestimated the energy barrier with respect to the global minimum  $C_{7\text{eq}}$ , in which OPLSAA and OPLSAA/L did slightly better than the other three force fields.

The comparison of Figures 3 and 4 shows that the changes from the gas phase to the solution phase FESs are common among all force fields and also very similar to those observed in the ab initio FESs (Figure 2). These changes include more extended low-energy regions, flatter surface, diminishing of gas-phase global minimum  $C_{7\text{eq}}$  and emerging of  $\beta$  and  $\alpha_R$ , and shifting from  $C_{7\text{ax}}$  to  $\alpha_L$ . Considering the force field aqueous FESs are obtained from simulations with explicit solvent molecules, the results indicate that the PCM model used in the ab initio calculations, although lacking specific AD-water interactions, does capture the major solvation effect.

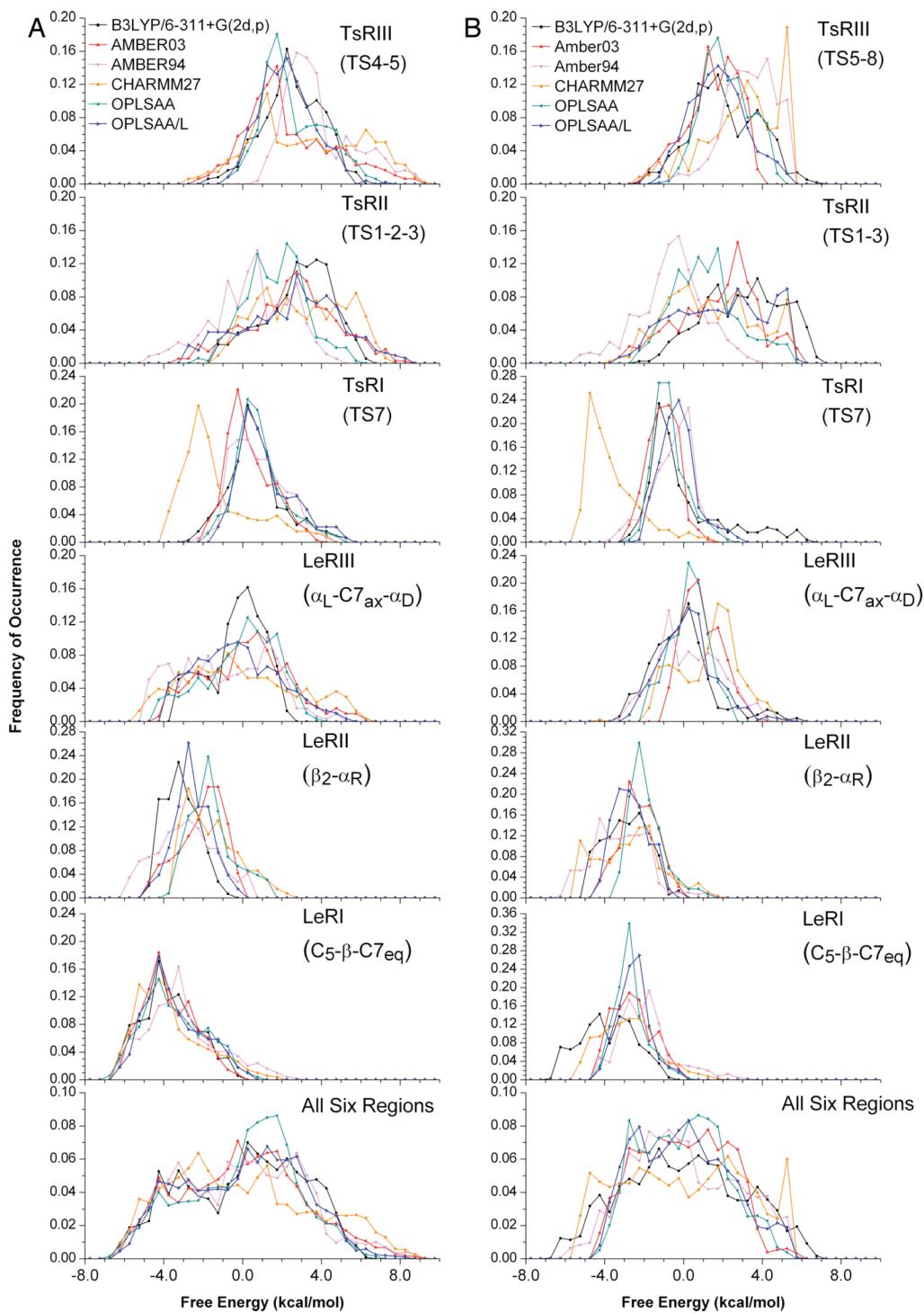
On the other hand, the difference in the details of the FESs between force field and ab initio is also more evident in the aqueous solution phase. Although all force fields have captured the low-lying minima  $\beta$  and  $\alpha_R$  accurately in terms of both their energies and positions. When compared with the ab initio FES, AMBER94, and CHARMM27 fail to capture minimum  $C_5$ . For the  $\beta_2$  conformer, OPLSAA/L has done a good job, followed by the two AMBER force fields which predict the position of  $\beta_2$  to be slightly higher than that of the ab initio calculations. The OPLSAA gives a local minimum roughly  $60^\circ$  (in  $\psi$ ) above the position of the ab initio  $\beta_2$  conformer. The CHARMM27 force field has failed to produce a  $\beta_2$  minimum and underestimated the stability of the  $\beta_2$  region significantly. The performance of the force fields on the second group of minima  $\alpha_L$ - $C_{7\text{ax}}$ - $\alpha_D$ - $\alpha_D'$  on the positive (in  $\phi$ ) side of the contour map also varies. The OPLSAA force field seems to produce the best fit to the ab initio result, featuring well-separated  $\alpha_L$  and  $C_{7\text{ax}}$  minima with  $\alpha_L$  being the lowest in free energy among the four minima. The two AMBER and the OPLSAA force fields combine  $\alpha_L$  and  $C_{7\text{ax}}$  into one minimum, which positions in the middle. The CHARMM27 force field predicts the position of  $\alpha_L$  quite well but underestimates its stability significantly. In terms of transition states, all force fields except CHARMM27 predict the position of TS7 fairly well, and its stability varies from 2 to 6 kcal/mol with respect to the  $\alpha_R$  or  $\beta$  conformer. Similar to the gas-phase CHARMM27 FES, the stability of the lower left quadrant of the aqueous CHARMM27 FES has been overestimated, therefore, no TS7 is observed. Instead, a transition state above the  $\alpha_R$  and below the  $C_{7\text{eq}}$  has been identified. For the transition-state group TS1–TS3, all force field FESs feature TS1 and TS3 whose positions and energetics also agree well with the ab initio FES. TS2 is missing from all force field energy maps, however, it is not of serious concern since TS2 is much higher in energy than the other two transition states. All force field FESs feature both TS5 and TS8, and their positions similar to that have been observed in the ab initio FES.

In addition to the force field-specific discrepancies discussed above, there is also one common difference between all MM force field FESs and the ab initio FES of AD in aqueous phase, i.e., the free energy span of the ab initio FES (0 to  $\sim 14$  kcal/mol), which is wider than that of the force field FESs (0 to  $\sim 10$ –12 kcal/mol). Since we have excluded the high-energy areas of the FES when calculating the above free energy range, we can safely say that the insufficient conformational sampling in these high-energy areas by the force fields does not influence our results. The lack of specific interactions (hydrogen bonds) between solvent and AD in the ab initio PCM calculations may have contributed to this discrepancy. We could not determine to what extent the lack of specific interactions has on the FES until ab initio MD simulation of AD with explicit solvent can be afforded. In addition, the free energy corrections may also have a small contribution (0.2–0.4 kcal/mol) to this common discrepancy.

The qualitative visual inspection of both ab initio and force field FESs reveals that although they all bear similar general features, they are distinctly different in characteristic details. In the following section, we will present quantitative assessment of the performance of each force field.

**Quantitative Assessment of the Force Field FESs.** Before a detailed quantitative comparison between force field and ab initio FESs can be made, it is necessary to determine the error (noise) of each individual FES so that the difference in energy measured later can be put into context. The error of the ab initio FES, ignoring errors from potential energy which is not the scope of this paper, is 0.2–0.4 kcal/mol based on analysis on the free energy corrections in the previous section. For the force field FESs, the method we use to estimate their resolutions was briefly mentioned in a previous article.<sup>51</sup> Basically, the errors are calculated as the standard deviation of the FEDM between two FESs that should be the same, e.g., two FESs using the same force field but different hills settings. The errors are estimated to be around 0.3–0.4 kcal/mol, or slightly larger (up to 0.7 kcal/mol) in some areas due to insufficient sampling, for both gas-phase and aqueous phase FESs. More details of the error analysis can be found in the Supporting Information.

As we have mentioned in the Introduction, most force field parametrization processes that integrate ab initio calculations have focused on getting the position and the relative energies of a few low-lying minima right. Since such parametrization does not pay much attention to the transition states, the resulting force field could introduce large errors when studying protein dynamics and kinetics. In this paper, we present a strategy to assess the performance of any force field not only in the low-energy regions but also transition-state areas. As shown in Figure 5, the  $\phi$ – $\psi$  space has been partitioned into six regions, which in total account for more than 80% of the conformational space. The rest of the space is of high energy and excluded from our analysis. The partitions are different from gas to aqueous solution phase and are based on common practice of protein backbone conformational partitions as well as the ab initio FESs. There are three low-energy regions in which LeRI encloses minima  $C_5$ ,  $\beta$ , and  $C_{7\text{eq}}$ , and LeRII includes minima  $\beta_2$  and  $\alpha_R$ , and both of these regions are on the negative  $\phi$  side. LeRIII is on the



**Figure 6.** Free energy distributions: (A) gas and (B) aqueous solution phases.

positive  $\phi$  side and covers minima  $\alpha_L$ ,  $C7_{ax}$ , and  $\alpha_D$ . There are also three transition regions in which TsRI is the low-lying transition region between LeRII and LeRI. Both TsRII and TsRIII link the negative and positive  $\phi$  sides through rotation of the  $\phi$  dihedral angle clockwise (through  $0^\circ$ ) or counterclockwise (through  $180^\circ$ ).

Based on the  $50 \times 50$  grid map, the distribution of free energies of each region has been plotted in Figure 6 for each force field or ab initio FES in gas or aqueous solution phase. In addition, we have also measured two quantities per region for each force field FES. As shown in Tables 3 and 4, the first one is the mean free energy, and the other is the standard

deviation of the FEDM between the force field and ab initio FESs for each region. Unlike the FESs (Figures 2–4), where free energy is relative to the global minimum, the mean free energies in Tables 3 and 4 or the free energies in the distribution plots of Figure 6 are relative to the overall mean of all six regions for each force field. This overall mean, as a reference, is unbiased and not subject to the error in energy of a single point (the global minimum) on the FES. Comparison of the mean free energy of each region between the ab initio and any force field FESs roughly indicates the average accuracy of the force field on energetics of a particular group of conformers or transition states. The

**Table 3.** Mean Free Energies of the Six Regions of the Ab Initio and MM Force Field FESs

	DFT	AMBER03	AMBER94	CHARMM27	OPLSAA	OPLSAA/L
			Gas Phase <sup>a</sup>			
all six regions	0.00	0.00	0.00	0.00	0.00	0.00
LeRI ( $C_5\beta\text{-}C7_{\text{eq}}$ )	-3.72	-3.82/-3.74	-3.16/-3.07	-4.14/-3.69	-3.69/-3.50	-3.63/-3.51
LeRII ( $\beta_2\text{-}\alpha_R$ )	-3.11	-1.86/-2.09	-2.60/-2.89	-1.27/-1.55	-1.22/-1.60	-2.66/-2.57
LeRIII ( $\alpha_L\text{-}C7_{\text{ax}}\text{-}\alpha_D$ )	-0.32	0.42/0.18	-0.35/-0.88	0.90/-0.11	0.32/-0.07	-0.07/-0.31
TSRI (TS7)	0.67	0.45/0.35	0.64/0.76	-1.51/-1.29	0.81/0.85	1.08/1.15
TSRII (TS1–TS3)	2.82	2.82/2.50	0.74/0.48	3.02/3.01	2.10/1.70	2.27/2.49
TSRIII (TS4–TS5)	2.48	1.87/2.34	3.38/3.86	2.87/3.28	2.11/2.54	2.26/2.23
			Aqueous Solution			
all six regions	0.00	0.00	0.00	0.00	0.00	0.00
LeRI ( $C_5\beta\text{-}C7_{\text{eq}}$ )	-3.85	-2.57	-1.91	-2.79	-2.68	-2.52
LeRII ( $\beta_2\text{-}\alpha_R$ )	-2.90	-2.29	-3.21	-2.74	-1.90	-2.66
LeRIII ( $\alpha_L\text{-}C7_{\text{ax}}\text{-}\alpha_D$ )	-0.07	0.94	0.30	1.27	0.26	0.21
TSRI (TS7)	0.04	-1.03	-0.50	-3.60	-0.71	-0.27
TSRII (TS1–TS3)	3.02	1.74	-0.60	1.37	1.12	1.87
TSRIII (TS5–TS8)	2.07	1.48	3.34	2.97	2.15	1.75

<sup>a</sup> First number is from a 20 ns gas-phase simulation using hills setting (2), while second number after the slash is from a 5 ns gas-phase simulation using hills setting (1).

**Table 4.** Standard Deviation (kcal/mol) of the FEDMs between the Ab Initio and Force Field FESs

	AMBER03	AMBER94	CHARMM27	OPLSAA	OPLSAA/L
		Gas Phase <sup>a</sup>			
all six regions	1.50/1.39	1.78/1.96	2.13/2.17	1.31/1.30	1.25/1.25
LeRI ( $C_5\beta\text{-}C7_{\text{eq}}$ )	0.89/0.75	1.39/1.53	1.03/1.06	0.76/0.89	0.89/0.89
LeRII ( $\beta_2\text{-}\alpha_R$ )	1.09/1.24	1.25/1.45	1.03/1.21	1.04/1.07	1.03/0.99
LeRIII ( $\alpha_L\text{-}C7_{\text{ax}}\text{-}\alpha_D$ )	1.58/1.53	1.92/1.83	2.21/2.32	1.03/1.09	1.37/1.51
TSRI (TS7)	1.01/0.90	1.04/1.15	2.25/2.22	1.04/0.98	0.76/0.73
TSRII (TS1–TS3)	1.55/1.55	1.64/1.62	1.65/2.02	1.47/1.60	1.96/1.90
TSRIII (TS4–TS5)	1.82/1.70	1.34/1.45	2.14/2.28	1.41/1.24	0.87/1.00
		Aqueous Solution			
all six regions	1.67	2.50	2.52	1.62	1.51
LeRI ( $C_5\beta\text{-}C7_{\text{eq}}$ )	1.23	1.84	1.69	1.26	1.17
LeRII ( $\beta_2\text{-}\alpha_R$ )	1.12	1.45	1.87	1.15	0.79
LeRIII ( $\alpha_L\text{-}C7_{\text{ax}}\text{-}\alpha_D$ )	1.23	1.96	2.11	1.27	1.19
TSRI (TS7)	1.74	2.17	2.14	1.55	1.32
TSRII (TS1–TS3)	1.50	1.52	1.92	1.31	1.82
TSRIII (TS5–TS8)	1.30	1.39	1.94	1.05	1.22

<sup>a</sup> First number is from a 20 ns gas-phase simulation using hills setting (2), while second number after the slash is from a 5 ns gas-phase simulation using hills setting (1).

standard deviation of the FEDM indicates how well the two energy surfaces match, therefore is a good measure of whether or not the force field predicts the positions of minima or transition states as accurately as the ab initio method. We will discuss, region by region, first the results of the gas-phase FESs, then the aqueous phase FESs.

**Gas-Phase Region by Region.** As shown in the bottom panel of Figure 6A, the overall free energy distributions, obtained by metadynamics simulations using the five different force fields, fit reasonably well with the ab initio calculations. The width and the position of the peaks all match closely with the ab initio distribution. Table 4 shows that the standard deviation of the difference between any force field and the ab initio FESs ranges from 1.25 to 2.17 kcal/mol, which is much larger than the estimated errors (0.2–0.7 kcal/mol), indicating true difference. More specifically, the two OPLSAA and the AMBER03 force fields differ from the ab initio free energy map about 1.2–1.4 kcal/mol, while AMBER94 and CHARMM27 have larger deviations from 1.8 to 2.2 kcal/mol.

The performance of all force fields in the first low-energy region, LeRI, is also good. As shown in Table 3, the average

free energies (−3.07 and −3.50 to −3.74) of LeRI are mostly within the error range with that of the ab initio FES (−3.72). The only exception is AMBER94 (−3.07), which underestimates the stability of LeRII for about 0.65 kcal/mol, just outside the error range of 0.3–0.4 kcal/mol. The standard deviation of the difference energy map between any force field and ab initio FES in this region mostly ranges from 0.8–1.0 kcal/mol. The AMBER94, again, has a slightly larger deviation of 1.5 kcal/mol from the ab initio energy map. In gas phase, all five force fields seem to underestimate the stability of the second low-energy region LeRII, i.e., the  $\beta_2\text{-}\alpha_R$  conformers, especially the CHARMM27, OPLSAA, and AMBER03 force fields, whose average free energies are 1.0–1.5 kcal/mol higher than the ab initio mean of LeRII. Both AMBER94 and OPLSAA/L give reasonable average free energies. However, the distribution of free energies (as shown in Figure 6A) of AMBER94 is much wider, and not surprisingly the standard deviation from the ab initio map is the largest (1.5 kcal/mol). That leaves the OPLSAA/L force field that performs well in both average free energy (~0.5 kcal/mol difference from ab initio) and standard deviation (~1 kcal/mol). For the third low-energy region LeRIII, the

average free energies ( $-0.88$  to  $0.28$ ) of all five force fields are all within a reasonable range, comparing with the ab initio mean ( $-0.32$ ) of LeRIII. Figure 6A shows that the force field free energies distributions are usually wider than that of the ab initio. The standard deviations of the difference between any force field and the ab initio FESs rank as the following: OPLSAA has the smallest deviation ( $1.09$  kcal/mol), followed by OPLSAA/L and AMBER03 ( $\sim 1.5$  kcal/mol), then followed by AMBER94 and CHARMM27 ( $1.8$ – $2.3$  kcal/mol). This is consistent with the observations based on the FESs, in which most force field FESs miss the  $\alpha_L$  and  $\alpha_D$  conformers, except OPLSAA which shows some signs of the two minima.

The next region, TsRI, features transition state TS7 and minimum  $\alpha'$ . The performance of all force fields except CHARMM27 are good, with their mean free energies within  $\pm 0.4$  kcal/mol, standard deviation  $\sim 1.0$  kcal/mol, and distribution fits well with the ab initio calculations. The exception, CHARMM27, heavily overestimates the stability of this region. The mean free energy calculated by CHARMM27 is nearly  $2$  kcal/mol lower than the ab initio mean. In addition, the standard deviation of the difference map between the CHARMM27 and ab initio FESs in this region is  $2.2$  kcal/mol, which is two times of any other force field.

The second transition-state region, TsRII, is one of the two paths between the  $\phi < 0$  (LeRI and LeRII) and the  $\phi > 0$  (LeRIII) regions. We have observed a general increase in standard deviation from the minima regions to this transition region. The results indicate that match between any force field FES with the ab initio FES in this region is less satisfactory than that of the low-energy minima regions. This is probably due to the fact that all parametrization focuses only on matching the positions and energetics of minima. The best standard deviation is  $\sim 1.5$  kcal/mol from the two AMBER and the OPLSAA force fields, while the other two are at  $\sim 2.0$  kcal/mol. The average free energies of TsRII of AMBER03, OPLSAA/L, and CHARMM27 are reasonable, while OPLSAA and AMBER94 underestimate the average free energy by  $1.1$ – $2.3$  kcal/mol, respectively. Overall, the AMBER03 force field gives the best fit to the ab initio FES in this region (TsRII).

The last transition-state region, TsRIII, also links the  $\phi < 0$  (LeRI and LeRII) and the  $\phi > 0$  (LeRIII) regions but from an opposite direction. AMBER94 and CHARMM27 overestimate the energy of this region for  $1.4$  and  $0.8$  kcal/mol, respectively. The average free energy of the other three force fields is good, however, the standard deviation between AMBER03 ( $1.7$  kcal/mol) and ab initio is larger than the two OPLSAA ( $1.0$  to  $1.2$  kcal/mol) force fields. Overall, the OPLSAA/L FES fits the best with ab initio for TsRIII, in terms of both mean and standard deviation.

**Aqueous Phase Region by Region.** Consistent with what we have observed in the free energy maps, the distribution of free energies among all six regions is generally wider for ab initio than the force fields. It should be noted that the sharp peaks we saw at the high energy end of the CHARMM27 distribution are due to insufficient conformational sampling, mostly in TsRIII region, as shown in Figure 4. Standard deviations between the force field and ab initio

FESs are generally larger than that in the gas phase, indicating less satisfactory matches. As shown in Table 4, the standard deviation of individual region or among all six regions falls into two groups. The OPLSAA, OPLSAA/L, and AMBER03 force fields are in one group ( $1.5$ – $1.7$  kcal/mol) with smaller deviations than AMBER94 and CHARMM27 ( $2.5$  kcal/mol). This trend is similar to the performance of these force fields in gas phase.

For the first low-energy region, LeRI ( $C_5\text{-}\beta\text{-}C7_{eq}$ ), the average free energies of all force field FESs are higher than the ab initio mean. This is directly related to the fact that ab initio has a wider distribution, therefore the average free energy of the lowest energy region, relative to the average of the overall distribution, is lower than that of the force fields. Since energy is only relevant in terms of relative stability, we will discuss the effects of this discrepancy later together with the average energies of other regions. However, we would like to mention that the average free energy of LeRI of AMBER94 is  $2.0$  kcal/mol higher than the ab initio mean. This difference is slightly larger than the  $1.0$ – $1.3$  kcal/mol differences between the other force fields and ab initio FESs. This is consistent with the gas-phase results in which AMBER94 underestimates the stability of LeRI. In terms of standard deviation, i.e., how well a force field FES matches with the ab initio FES, OPLSAA/L, OPLSAA, and AMBER03 give  $\sim 1.2$  kcal/mol deviation, while AMBER94 and CHARMM27 go up to  $1.7$ – $1.8$  kcal/mol.

For the second low-energy region, LeRII ( $\beta_2\text{-}\alpha_R$ ), the average free energy is  $0.9$  kcal/mol higher than that of LeRI ( $C_5\text{-}\beta\text{-}C7_{eq}$ ) based on the ab initio FES. This difference is very much in agreement with a probability partition of  $80$ – $20$  between the  $C_5\text{-}\beta$  (or  $\beta\text{-}P_{II}$ ) and the  $\alpha_R$  states, according to an experimental/computational investigation.<sup>27</sup> For AMBER03, CHARMM27, and OPLSAA/L, the mean free energy of LeRII agrees well with that of the ab initio. However, the free energy differences between the LeRI and LeRII regions for these force fields are too small ( $-0.3$ ,  $-0.05$ , and  $0.1$  kcal/mol, respectively) because of the underestimation of the stability of the LeRI regions. The near to zero ( $-0.05$ ) energy difference obtained by CHARMM27 is consistent with previous MD simulation which gives a  $50$ – $50$  probability distribution<sup>24</sup> between the two regions. AMBER03 is slightly better than CHARMM27, while OPLSAA/L is slightly worse. For AMBER94, the mean free energy of LeRII is in line with ab initio as well. However, the order of mean free energy between LeRI and LeRII is reversed due to the underestimation of stability of LeRI. Again, the reversed order of stability is consistent with previous MD simulation, which gives a reversed  $20$ – $80$  probability distribution between LeRI and LeRII.<sup>24</sup> For OPLSAA, the mean free energy of LeRII is  $1.0$  kcal/mol higher than the ab initio mean. However, since both LeRI and LeRII's stabilities have been underestimated by  $1.0$ – $1.1$  kcal/mol, their relative stability is similar to ab initio. This is again in line with previous MD simulation that gives an  $85$ – $15$  probability distribution between the two regions.<sup>24</sup> The consistency observed also validates the current approach of using the mean free energy and the definition of regions. The standard deviations for LeRII between the force field

and ab initio FESs are 0.8 kcal/mol for OPLSAA/L, ~1.1 kcal/mol for OPLSAA and AMBER03, and 1.5–1.9 kcal/mol for AMBER94 and CHARMM27.

For the low-energy region on the negative  $\phi$  side, LeRIII, the mean free energies are 1.0–1.3 kcal/mol higher than the ab initio mean for AMBER03 and CHARMM27. This results in relative stability between LeRI and LeRIII, in good agreement with ab initio, but not for relative stability between LeRII and LeRIII. For the other three force fields, the mean free energies of LeRIII are close to the ab initio value. Therefore, the relative stabilities between LeRII and LeRIII are good for AMBER94 and OPLSAA/L but not so good for OPLSAA due to the higher mean free energy of LeRII. Furthermore, none of these three produces a good relative stability between LeRI and LeRIII. The standard deviations of LeRIII between force field and ab initio are very similar to LeRI, in which the first group of force fields OPLSAA/L, OPLSAA, and AMBER03 is at ~1.2 kcal/mol, while the other group including AMBER94 and CHARMM27 is at ~2.0 kcal/mol.

For the first transition-state region, TSRI, the results are very similar to the gas phase, in which CHARMM27 heavily and AMBER03 and OPLSAA slightly overestimate the stability of this area. Combining performance in both mean and standard deviation, the OPLSAA/L is the best as comparing with ab initio. For the second transition-state region, TSRII, the average free energy from all force fields is lower than that of the ab initio. This directly contributes to the fact that ab initio FES has a wider free energy distribution than all the force field FESs. More specifically, AMBER94 underestimates the free energy of this region for 3.6 kcal/mol. The other force fields perform better than AMBER94 and only underestimate the average free energy for 1.1 to 1.9 kcal/mol. Similar to the gas-phase result, AMBER03 is the best match to the ab initio values combining the performance of mean and standard deviation. OPLSAA/L has the best mean but slightly larger deviation, while OPLSAA has the best deviation but slightly lower mean. For the third transition-state region, TsRIII, the mean free energies obtained by CHARMM27 and AMBER94 are 0.9 and 1.3 kcal/mol higher than the ab initio mean. Both force fields also have higher standard deviations from the ab initio FES. The other three force fields match with the ab initio values in both mean free energy (difference within 0.6 kcal/mol) and standard deviation (1.0–1.3 kcal/mol). The two OPLSAA force fields perform slightly better than AMBER03 in both mean and standard deviation.

## Conclusions

The current B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d,p) with PCM model approach is able to generate potential energy contour maps of AD in gas phase and in water that is at least as accurate as the previous MP2 approaches. The cost-efficient hybrid DFT method has also allowed us to generate FESs by calculating free energy corrections on a  $50 \times 50$  grid base. The error range, from the free energy corrections, is estimated at around 0.2–0.4 kcal/mol. The average free energy difference of 0.9 kcal/mol between the first two conformational basins, LeRI ( $C_5\text{-}\beta\text{-}C7_{eq}$ ) and LeRII ( $\beta_2\text{-}\alpha_R$ )

on the ab initio-DFT FES in aqueous phase agrees excellently with a 80–20 probability distribution from recent experimental/computational investigations.<sup>27,38</sup>

Metadynamics simulations, in combination with five commonly used MM force fields, have been carried out to obtain FESs of AD in both gas phase and water. The error range of these FESs is calculated to be 0.3–0.4 kcal/mol and in some conformational areas goes up to 0.7 kcal/mol due to sampling. Quantitative assessment of these force field FESs in three low-energy conformational basins, LeRI ( $C_5\text{-}\beta\text{-}C7_{eq}$ ), LeRII ( $\beta_2\text{-}\alpha_R$ ), and LeRIII( $\alpha_L\text{-}C7_{ax}\text{-}\alpha_D$ ) as well as three transition-state regions was made according to the ab initio DFT FESs. The average free energy differences between the LeRI and LeRII basins, among several force fields, agree well with previous MD simulations,<sup>24</sup> which gives validation to the partition of conformational basin and the assessment method. The quantitative assessment reveals variations in performance from one conformational region to another and from force field to force field or from gas to aqueous phase. Although not one MM force field is able to outperform all others in all conformational areas, the overall best performer is the OPLSAA/L force field, followed by OPLSAA and AMBER03. The results also indicate a certain degree of transferability of performance from gas to aqueous phase. However, there are also areas where a better fit in one phase leads to a decrease performance of another phase, such as AMBER03 or OPLSAA/L.

In summary, we have presented a new method of assessing force field performance not only in the energies of conformers but also transition-state barriers. The method presented and results obtained here should be useful for improving the parametrization of force field. More specifically, the ab initio-DFT FESs, the partitions of conformational basins, and the transition-state regions as well as the method of quantitative assessment can be used in force field parametrization to improve the force field description of all areas of the ( $\phi\text{-}\psi$ ) conformational map.

**Acknowledgment.** Z.L. and P.B.M. acknowledge the financial support from the National Science Foundation (EMT CCF-0622162 and MRI CHE-0420556), the National Health Institutes (GM075990) and the H. O. West Foundation. B.E. acknowledges financial support through a VIDI grant from the Chemical Sciences Division of The Netherlands Organization for Scientific Research (NWO-CW). We also thank the TeraGrid Project for providing computational resources.

**Supporting Information Available:** Details on the series of partial optimization to obtain the  $\alpha_D$  conformer and the TS3 transition state for gas phase and the  $\beta$  conformer for aqueous phase. Details on the error analysis of the force field FESs. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, U.K., 1987.
- (2) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press: San Diego, CA, 1996.

- (3) Cheatham, T. E., III; Kollman, P. A. Molecular dynamics simulation of nucleic acids. *Annu. Rev. Phys. Chem.* **2000**, *51*, 435–471.
- (4) Leach, A. R. *Molecular Modelling: Principles and Applications*. Prentice Hall: Upper Saddle River, NJ, 2000.
- (5) Becker, O. M.; MacKerell, A. D., Jr.; Roux, B.; Watanabe, M. *Computational Biochemistry and Biophysics*, 1st ed.; Marcel Dekker: New York, 2001.
- (6) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211–243.
- (7) Becker, O. M.; Karplus, M. *Guide to Biomolecular Simulations*. Springer: Dordrecht, The Netherlands, 2005.
- (8) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197.
- (9) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.
- (10) Schuler, L. D.; Daura, X.; van, G. W. F. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *J. Comput. Chem.* **2001**, *22* (11), 1205–1218.
- (11) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236.
- (12) Anderson, A. G.; Hermans, J. Microfolding - conformational probability map for the alanine dipeptide in water from molecular-dynamics simulations. *Proteins* **1988**, *3* (4), 262–265.
- (13) Wu, X. W.; Sung, S. S. Simulation of peptide folding with explicit water--a mean solvation method. *Proteins* **1999**, *34* (3), 295–302.
- (14) Brooks, C. L., III. Protein and peptide folding explored with molecular simulations. *Acc. Chem. Res.* **2002**, *35* (6), 447–454.
- (15) Zhou, R. Free energy landscape of protein folding in water: Explicit vs. implicit solvent. *Proteins* **2003**, *53* (2), 148–161.
- (16) Swope, W. C.; Pitera, J. W.; Suits, F.; Pitman, M.; Eleftheriou, M.; Fitch, B. G.; Germain, R. S.; Rayshubski, A.; Ward, T. J. C.; Zhestkov, Y.; Zhou, R. Describing protein folding kinetics by molecular dynamics simulations. 2. Example applications to alanine dipeptide and a  $\beta$ -hairpin peptide. *J. Phys. Chem. B* **2004**, *108* (21), 6582–6594.
- (17) Scheraga, H. A.; Khalili, M.; Liwo, A. Protein-folding dynamics: Overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* **2007**, *58*, 57–83.
- (18) Lei, H.; Duan, Y. Protein folding and unfolding by all-atom molecular dynamics simulations. *Methods Mol. Biol.* **2008**, *443*, 277–295.
- (19) van, d. K. M. W.; Schaeffer, R. D.; Jonsson, A. L.; Scouras, A. D.; Simms, A. M.; Toofanny, R. D.; Benson, N. C.; Anderson, P. C.; Merkley, E. D.; Rysavy, S.; Bromley, D.; Beck, D. A. C.; Daggett, V. Dynameomics: A comprehensive database of protein dynamics. *Structure* **2010**, *18* (4), 423–435.
- (20) Velez-Vega, C.; Borrero, E. E.; Escobedo, F. A. Kinetics and reaction coordinate for the isomerization of alanine dipeptide by a forward flux sampling protocol. *J. Chem. Phys.* **2009**, *130* (22), 225101–225112.
- (21) Gaigeot, M. P. Unraveling the conformational dynamics of the aqueous alanine dipeptide with first-principle molecular dynamics. *J. Phys. Chem. B* **2009**, *113* (30), 10059–10062.
- (22) Han, W.-G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. Theoretical study of aqueous N-Acetyl-L-alanine N'-Methylamide: Structures and raman, VCD, and ROA spectra. *J. Phys. Chem. B* **1998**, *102* (14), 2587–2602.
- (23) Smith, P. E. The alanine dipeptide free energy surface in solution. *J. Chem. Phys.* **1999**, *111* (12), 5568–5579.
- (24) Hu, H.; Elstner, M.; Hermans, J. Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine “dipeptides” (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. *Proteins* **2003**, *50* (3), 451–463.
- (25) Wang, Z.-X.; Duan, Y. Solvation effects on alanine dipeptide: A MP2/cc-pVTZ//MP2/6-31G\*\* study of  $(\phi, \psi)$  energy maps and conformers in the gas phase, ether, and water. *J. Comput. Chem.* **2004**, *25* (14), 1699–1716.
- (26) Feig, M. Kinetics from implicit solvent simulations of biomolecules as a function of viscosity. *J. Chem. Theory Comput.* **2007**, *3* (5), 1734–1748.
- (27) Kwac, K.; Lee, K.-K.; Han, J. B.; Oh, K.-I.; Cho, M. Classical and quantum mechanical/molecular mechanical molecular dynamics simulations of alanine dipeptide in water: Comparisons with IR and vibrational circular dichroism spectra. *J. Chem. Phys.* **2008**, *128* (10), 105106–105119.
- (28) Liu, C.; Zhao, D.-X.; Yang, Z.-Z. ABEM $\sigma\pi$  fluctuating charge force field applied to alanine dipeptide and alanine dipeptide-water systems. *J. Theor. Comput. Chem.* **2010**, *9* (Supp. 1), 77–97.
- (29) Jono, R.; Watanabe, Y.; Shimizu, K.; Terada, T. Multicanonical *ab initio* QM/MM molecular dynamics simulation of a peptide in an aqueous environment. *J. Comput. Chem.* **2010**, *31* (6), 1168–1175.
- (30) Weise, C. F.; Weisshaar, J. C. Conformational analysis of alanine dipeptide from dipolar couplings in a water-based liquid crystal. *J. Phys. Chem. B* **2003**, *107* (14), 3265–3277.
- (31) Lavrich, R. J.; Plusquellic, D. F.; Suenram, R. D.; Fraser, G. T.; Hight, W. A. R.; Tubergen, M. J. Experimental studies of peptide bonds: Identification of the C7eq conformation of the alanine dipeptide analog N-acetylalanine N'-methylamide from torsion-rotation interactions. *J. Chem. Phys.* **2003**, *118* (3), 1253–1265.
- (32) Madison, V.; Kopple, K. D. Solvent-dependent conformational distributions of some dipeptides. *J. Am. Chem. Soc.* **1980**, *102* (15), 4855–63.
- (33) Deng, Z.; Polavarapu, P. L.; Ford, S. J.; Hecht, L.; Barron, L. D.; Ewig, C. S.; Jalkanen, K. Solution-phase conformations of N-Acetyl-N'-methyl-L-alaninamide from vibrational raman optical activity. *J. Phys. Chem.* **1996**, *100* (6), 2025–34.

- (34) Poon, C.-D.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. Do bridging water molecules dictate the structure of a model dipeptide in aqueous solution. *J. Am. Chem. Soc.* **2000**, *122* (23), 5642–5643.
- (35) Takekiyo, T.; Imai, T.; Kato, M.; Taniguchi, Y. Temperature and pressure effects on conformational equilibria of alanine dipeptide in aqueous solution. *Biopolymers* **2004**, *73* (2), 283–290.
- (36) Bohr, H. G.; Jalkanen, K. J. Spectroscopic studies of small biomolecules in condensed phases. *Condens. Matter Theor.* **2006**, *20*, 375–391.
- (37) Kim, Y. S.; Wang, J.; Hochstrasser, R. M. Two-dimensional infrared spectroscopy of the alanine dipeptide in aqueous solution. *J. Phys. Chem. B* **2005**, *109* (15), 7511–7521.
- (38) Grdadolnik, J.; Golic, G. S.; Avbelj, F. Determination of conformational preferences of dipeptides using vibrational spectroscopy. *J. Phys. Chem. B* **2008**, *112* (9), 2712–2718.
- (39) Mukhopadhyay, P.; Zuber, G.; Beratan, D. N. Characterizing aqueous solution conformations of a peptide backbone using Raman optical activity computations. *Biophys. J.* **2008**, *95* (12), 5574–5586.
- (40) Schweitzer-Stennet, R.; Measey, T.; Kakalis, L.; Jordan, F.; Pizzanelli, S.; Forte, C.; Griebenow, K. Conformations of alanine-based peptides in water probed by FTIR, raman, vibrational circular dichroism, electronic circular dichroism, and NMR spectroscopy. *Biochemistry* **2007**, *46* (6), 1587–1596.
- (41) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L. III Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25* (11), 1400–1415.
- (42) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24* (16), 1999–2012.
- (43) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105* (28), 6474–6487.
- (44) Chipot, C.; Pohorille, A. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*. Springer-Verlag: Berlin, Germany, 2007.
- (45) Smiatek, J.; Heuer, A. Calculation of free energy landscapes: a histogram reweighted metadynamics approach. *Physics* **2010**, 1–11; arXiv:1006.4308v1 [physics.comp-ph].
- (46) Dickson, B. M.; Legoll, F.; Lelievre, T.; Stoltz, G.; Fleurat-Lessard, P. Free energy calculations: An efficient adaptive biasing potential method. *J. Phys. Chem. B* **2010**, *114* (17), 5823–5830.
- (47) Bartels, C.; Karplus, M. Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations. *J. Comput. Chem.* **1997**, *18* (12), 1450–1462.
- (48) Maragakis, P.; van der Vaart, A.; Karplus, M. Gaussian-mixture umbrella sampling. *J. Phys. Chem. B* **2009**, *113* (14), 4664–4673.
- (49) Rosso, L.; Abrams, J. B.; Tuckerman, M. E. Mapping the backbone dihedral free-energy surfaces in small peptides in solution using adiabatic free-energy dynamics. *J. Phys. Chem. B* **2005**, *109* (9), 4162–4167.
- (50) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (20), 12562–12566.
- (51) Ensing, B.; De, V. M.; Liu, Z.; Moore, P.; Klein, M. L. Metadynamics as a tool for exploring free energy landscapes of chemical reactions. *Acc. Chem. Res.* **2006**, *39* (2), 73–81.
- (52) Vymetal, J.; Vondrasek, J. Metadynamics as a tool for mapping the conformational and free-energy space of peptides: The alanine dipeptide case study. *J. Phys. Chem. B* **2010**, *114* (16), 5632–5642.
- (53) Seabra, G. d. M.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the amber molecular dynamics package. *J. Phys. Chem. A* **2007**, *111*, 5655–5664.
- (54) Mu, Y. G.; Kosov, D. S.; Stock, G. Conformational dynamics of trialanine in water. 2. Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments. *J. Phys. Chem. B* **2003**, *107* (21), 5064–5073.
- (55) Laio, A.; Rodriguez-Forte, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the accuracy of metadynamics. *J. Phys. Chem. B* **2005**, *109* (14), 6714–21.
- (56) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. *J. Am. Chem. Soc.* **1997**, *119* (25), 5908–5920.
- (57) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37* (2), 785–9.
- (58) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157* (3), 200–6.
- (59) Becke, A. D. A new mixing of Hartree-Fock and local-density-functional theories. *J. Chem. Phys.* **1993**, *98* (2), 1372–7.
- (60) Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the revision of solvent effects. *Chem. Phys.* **1981**, *55* (1), 117–29.
- (61) Miertus, S.; Tomasi, J. Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. *Chem. Phys.* **1982**, *65* (2), 239–45.
- (62) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Ab initio* study of solvated molecules: a new implementation of the polarizable continuum model. *Chem. Phys. Lett.* **1996**, *255* (4,5,6), 327–335.
- (63) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.;

- Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03*, revision B.05; Gaussian, Inc.: Wallingford, CT, 2004.
- (64) Wong, M. W. Vibrational frequency prediction using density functional theory. *Chem. Phys. Lett.* **1996**, *256* (4,5), 391–399.
- (65) <http://www.westcenter.usp.edu/code>.
- (66) Toukan, K.; Rahman, A. Molecular-dynamics study of atomic motions in water. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *31* (5), 2643–8.
- (67) Ahlborn, H.; Ji, X.; Space, B.; Moore, P. B. A combined instantaneous normal mode and time correlation function description of the infrared vibrational spectrum of ambient water. *J. Chem. Phys.* **1999**, *111* (23), 10622–10632.
- (68) Siu, S. W. I.; Vacha, R.; Jungwirth, P.; Boeckmann, R. A. Biomolecular simulations of membranes: Physical properties from different force fields. *J. Chem. Phys.* **2008**, *128* (12), 125103–125115.
- (69) Feig, M.; Pettitt, B. M. Experiment vs force fields: DNA conformation from molecular dynamics simulations. *J. Phys. Chem. B* **1997**, *101* (38), 7361–7363.
- (70) Feller, S. E.; Pastor, R. W.; Rojnuckarin, A.; Bogusz, S.; Brooks, B. R. Effect of electrostatic force truncation on interfacial and transport properties of water. *J. Phys. Chem.* **1996**, *100* (42), 17011–17020.
- (71) Norberg, J.; Nilsson, L. On the truncation of long-range electrostatic interactions in DNA. *Biophys. J.* **2000**, *79* (3), 1537–1553.
- (72) Feig, M. Is Alanine dipeptide a good model for representing the torsional preferences of protein backbones. *J. Chem. Theory Comput.* **2008**, *4* (9), 1555–1564.
- (73) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. DFT studies on helix formation in N-acetyl-(L-alanyl) $n$ -N'-methylamide for  $n = 1\text{--}20$ . *Chem. Phys.* **2000**, *256* (1), 15–27.
- (74) Jalkanen, K. J.; Elstner, M.; Suhai, S. Amino acids and small peptides as building blocks for proteins: comparative theoretical and spectroscopic studies. *THEOCHEM* **2004**, *675* (1–3), 61–77.
- (75) Grenie, Y.; Avignon, M.; Garrigou-Lagrange, C. Molecular structure study of dipeptides isolated in an argon matrix by infrared spectroscopy. *J. Mol. Struct.* **1975**, *24* (2), 293–307.
- (76) Ramachandran, G. N.; Sasisekharan, V. Conformation of polypeptides and proteins. *Adv. Protein Chem.* **1968**, *23*, 283–438.

CT100395N