

Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters

Viktor Hornak,¹ Robert Abel,² Asim Okur,³ Bentley Strockbine,⁴ Adrian Roitberg,² and Carlos Simmerling^{1,3,5*}

¹Center for Structural Biology, Stony Brook University, Stony Brook, New York 11794

²Department of Chemistry and Quantum Theory Project, University of Florida, Gainesville, Florida 32611

³Department of Chemistry, Stony Brook University, Stony Brook, New York 11794

⁴Department of Pharmacological Sciences, Stony Brook University, Stony Brook, New York 11794

⁵Computational Science Center, Brookhaven National Laboratory, Upton, New York 11973,

ABSTRACT The ff94 force field that is commonly associated with the Amber simulation package is one of the most widely used parameter sets for biomolecular simulation. After a decade of extensive use and testing, limitations in this force field, such as over-stabilization of α -helices, were reported by us and other researchers. This led to a number of attempts to improve these parameters, resulting in a variety of “Amber” force fields and significant difficulty in determining which should be used for a particular application. We show that several of these continue to suffer from inadequate balance between different secondary structure elements. In addition, the approach used in most of these studies neglected to account for the existence in Amber of two sets of backbone ϕ/ψ dihedral terms. This led to parameter sets that provide unreasonable conformational preferences for glycine. We report here an effort to improve the ϕ/ψ dihedral terms in the ff99 energy function. Dihedral term parameters are based on fitting the energies of multiple conformations of glycine and alanine tetrapeptides from high level *ab initio* quantum mechanical calculations. The new parameters for backbone dihedrals replace those in the existing ff99 force field. This parameter set, which we denote ff99SB, achieves a better balance of secondary structure elements as judged by improved distribution of backbone dihedrals for glycine and alanine with respect to PDB survey data. It also accomplishes improved agreement with published experimental data for conformational preferences of short alanine peptides and better accord with experimental NMR relaxation data of test protein systems. *Proteins* 2006;65:712–725. © 2006 Wiley-Liss, Inc.

Key words: trialanine; dihedral parameters; molecular dynamics; molecular mechanics; decoy analysis; NMR order parameters; α -helix

INTRODUCTION

Molecular modeling studies of biologically important molecules usually require the evaluation of potential energies for alternate conformations. The most accurate approach would encompass the use of the quantum me-

chanical (QM) wave function. Unfortunately, because of the large sizes of biological macromolecules, such calculations are extremely time consuming and can only be applied to a limited number of conformations. Therefore, most simulation projects employ classical molecular mechanics (MM) energy functions. Because of the need to evaluate these functions a large number of times during the simulation, these functions are relatively simple and utilize many adjustable empirical parameters. These are most often obtained by fitting to data from experiments or from high level QM calculations. The energy function together with the set of empirical parameters is known as a force field.

In spite of persistent yet slow emergence of force fields that explicitly account for charge polarization,¹ the fixed charge additive force fields remain popular because of computational efficiency. For a more in-depth overview of current force fields and trends in force field development, the reader is referred to recent review articles.^{2–4} In this report, we focus on the existing Amber fixed charge additive force fields.

The “Cornell et al.” force field⁵ (denoted as ff94 in Amber) has been the most widely used with the Amber suite of programs^{6,7} since its publication over a decade ago. It introduced the set of parameters for all-atom simulations suitable for protein simulations in the condensed phase, largely inspired by the optimized potential for liquid simulations (OPLS). Some characteristic features of ff94 include fixed partial charges on atom centers, explicit use of all hydrogen atoms, no specific functional form for hydrogen bonding, and dihedral parameters fit to relative QM energies of alternate rotamers of small molecules. In particular, the protein ϕ/ψ dihedrals have specific rota-

The Supplementary Material referred to in this article can be found at <http://www.interscience.wiley.com/jpages/0887-3585/suppmat/>

Grant sponsor: National Institutes of Health; Grant number: GM6167803; Grant sponsor: Department of Energy; Grant number: DE-AC02-98CH10886.

*Correspondence to: Carlos Simmerling, Department of Chemistry, Stony Brook University, Stony Brook, NY 11794.
E-mail: carlos.simmerling@stonybrook.edu

Received 23 February 2006; Revised 4 May 2006; Accepted 15 June 2006

Published online 15 September 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21123

tional parameters that affect relative energies of alternate backbone conformations. These were fit to optimize agreement with QM relative energies for several conformations of glycine and alanine. The partial atomic charges were derived by fitting the (gas phase) electrostatic potential calculated at the Hartree-Fock 6-31G* level. This approach intentionally “overpolarizes” bond dipoles as present in the gas phase, such that the resulting charge distribution approximates those occurring in aqueous condensed phase.⁹

Because of limited computational resources at the time, dihedral parameters were fit to a small number of low-energy conformations of glycine and alanine dipeptides. A possible limitation of using dipeptides is that their gas phase energy surfaces do not have a local minimum in the α -helical region, which occurs with high frequency in protein structures. This was addressed in subsequent modifications of ff94, such as ff96¹⁰ and, more recently, ff99.¹¹ In ff96, identical parameters were used for ϕ and ψ terms, which were empirically adjusted to reproduce the energy difference between extended and constrained α -helical energies for alanine tetrapeptide. ff99 reflected another attempt at refitting backbone dihedral parameters by including 11 representative structures of alanine tetrapeptide along with the alanine dipeptides. During this process, however, a new problem was introduced in the way ff96/ff99 ϕ/ψ dihedral parameters were optimized (discussed in detail later), resulting in incorrect conformational preferences for glycine. In addition to changes in backbone dihedral parameters, ff99 also generalized and extended atom types to compounds beyond amino and nucleic acids.

Even though both ff99 and, particularly, ff94 parameter sets have been successfully used for many years, improvement in conformational sampling due to increased computer power and algorithmic advances revealed that both force fields over-stabilize α -helical peptide conformations.^{12,13} On the other hand, the changes introduced in the less frequently used ff96 were observed to overestimate β -strand propensity.^{14–17} Because the backbone dihedral parameters are shared by all amino acids, regardless of the type of side-chain, their cumulative effect is most likely responsible for the bias towards the specific secondary structure. This recently motivated us^{12,18} and others^{13,19} to revisit backbone dihedral parameterization for ff94/ff99. Garcia and Sanbonmatsu¹³ simply zeroed the torsion potential for ϕ and ψ (although as we explain later, only some of the ϕ/ψ terms were removed) and noted improved agreement with experimental helix-coil parameters. More recently, Sorin and Pande^{19,20} modified ff99 by replacing its ϕ dihedral parameters with the ones from ff94. This improved the agreement with experimental kinetic and thermodynamic measurements for folding of two 21-residue α -helical peptides. Our own modification of ff99, published previously,¹⁸ specifically addressed strong helical bias and, in this respect, succeeded to achieve better balance of major secondary structure types.

All these previous modifications were aimed at correcting an apparent problem encountered with a specific system. The heuristic approaches that were employed might

have improved that specific weakness but a more systematic revision was clearly necessary to quantify the problem and address the issue with generality and transferability in mind.

Much of the confusion and suboptimal performance of recent Amber protein force field variants may have been caused by a somewhat nonintuitive parameterization of protein backbone dihedral angles. In Amber, each dihedral profile is defined by a set of four atoms. The set of atoms used to define ϕ and ψ for glycine is as expected, following ϕ and ψ along the main chain ($\phi = \text{C}-\text{N}-\text{C}^\alpha-\text{C}$, $\psi = \text{N}-\text{C}^\alpha-\text{C}-\text{N}$). Importantly, for other amino acids that have a side chain, an *additional* set of dihedrals also influences rotation about the ϕ/ψ bonds connecting the C^α atom to the amide C and N atoms. This extra set of terms corresponds to dihedral angles branched out to the C^β carbon, which in this work we identify as $\phi' = \text{C}-\text{N}-\text{C}^\alpha-\text{C}^\beta$ and $\psi' = \text{C}^\beta-\text{C}^\alpha-\text{C}-\text{N}$.

This definition of dihedrals was originally implemented in ff94, with ϕ and ψ that fit independently to glycine data and ϕ' and ψ' used to adjust the behavior for alanine. Later modifications of ff94, such as ff96, ff99, Garcia's¹³, Pande's¹⁹ and ours^{12,18} only changed the first set of ϕ/ψ terms, using them to adjust backbone preferences for alanine. Thus the new ϕ and ψ parameters were fit in the presence of the ff94 alanine-based ϕ' and ψ' and will only give their intended behavior, when they accompany these ϕ' and ψ' dihedrals. Importantly, ϕ' and ψ' are not present in glycine. For example, ff99 modified the backbone ϕ and ψ terms in order to reproduce relative energies for alanine. However, the alanine residue also had the ϕ' and ψ' dihedral parameters present during the ϕ/ψ modification. These optimizations may perform well for nonglycine residues that possess the ϕ'/ψ' terms. However, when these modified backbone dihedral parameters are applied to glycine, the result has little physical justification, since they were fit in the presence of ϕ' and ψ' . A similar inconsistency is present in Garcia's modification; only the ϕ and ψ terms were zeroed; thus, while glycine has no backbone dihedral potential all other amino acids retain the ff94 ϕ' and ψ' terms, which were fit in the presence of the glycine-based ϕ and ψ terms, and thus have no meaning when applied without them.

We show below that these modifications result in unreasonable sampling of dihedral space for several post-ff94 Amber force fields. These problems are in addition to the over-stabilization of helical conformations present in ff94 and ff99. Amber and other popular force fields (e.g. Charmm22 and Gromos96 variants and OPLS-AA) were previously compared^{21,22} in terms of conformational sampling of blocked glycine and alanine dipeptide and were noted to disagree to various degrees (particularly for glycine²¹) with simulations employing combined QM/MM force field as well as with statistical analysis of high resolution protein crystal structures. In another study, all force fields were also shown to perform very differently with respect to a number of properties of trialanine compared with NMR and infrared observables.²³

We also show that our reparameterization of backbone dihedral parameters improves conformational preferen-

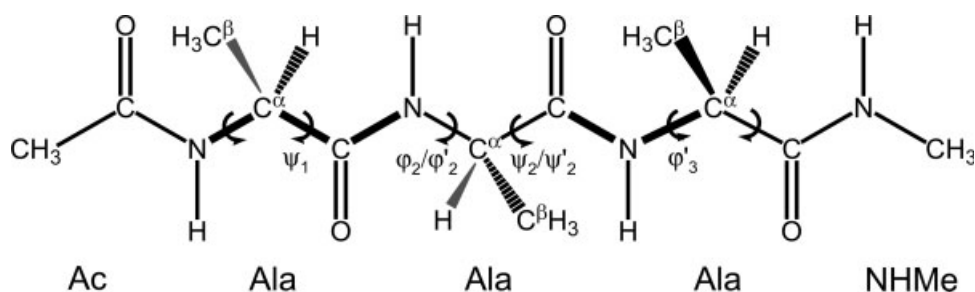


Fig. 1. Blocked alanine tetrapeptide (three alanine residues but four peptide bonds) used in the optimization procedure. Note the definition of dihedral angles: $\phi = \text{C}-\text{N}-\text{C}^\alpha-\text{C}$, $\psi = \text{N}-\text{C}^\alpha-\text{C}-\text{N}$, $\phi' = \text{C}-\text{N}-\text{C}^\alpha-\text{C}^\beta$, $\psi' = \text{C}^\beta-\text{C}^\alpha-\text{C}-\text{N}$. There are three ϕ/ψ pairs and three additional ϕ'/ψ' pairs in this tetrapeptide. Glycine tetrapeptide would only have the ϕ/ψ dihedrals due to the absence of C^β carbon. As an example, ψ_1 and ϕ'_3 are shown in bold.

ces for typical secondary structures. It should be noted that this work does not attempt to create yet another variant of Amber force field but strives to improve the existing ff94 and ff99 force fields, as well as specifically address the problem with glycine sampling. Thus rather than contributing to the diverging set of Amber force fields, we extend the evolution started with ff94, followed by ff99 and subsequently refined in this report. This is also reflected by naming this modified force field ff99SB. We attempted to maximally improve the methodology that we used to derive our new parameters while still following the same general philosophy as was used with ff94/ff99.

A different approach was taken by Duan et al.,²⁴ who recently introduced a more extensive modification of ff94/ff99 (called ff03), in which a fundamentally different concept to derivation of partial atomic charges was used. Instead of relying on the HF/6-31G* approach to provide aqueous-phase charges, a low-dielectric continuum model corresponding to an organic solvent environment was included directly in the QM calculation of the dihedral parameters and electrostatic potential (from which the charges are obtained). Because of these differences, ff03 should be considered a distinct force field model rather than extension of previous Amber force fields.

Amber force fields are not the only ones undergoing such evolutionary changes brought about by discovery of limitations through improved conformational sampling capabilities. Similar efforts have recently been reported by MacKerell et al.,²⁵ improving the backbone parameters in the Charmm force field, specifically removing the bias toward π -helical peptide conformations.^{26,27} In this case, the Charmm22 alanine dipeptide dihedral energy surface was corrected using a grid-based difference map (CMAP) to achieve an almost perfect match with LMP2/cc-pVQZ(-g) energy surface. Likewise, dihedral terms in the original OPLS-AA force field²⁸ were also reparameterized to improve the agreement with high level QM data.²⁹

METHODS

Optimization of Backbone Dihedral Parameters

To better understand the procedure we used for obtaining a new set of backbone dihedral parameters, we first outline the general procedure originally used in ff94.⁵

Even though it may at first seem more straightforward to have a single set of backbone ϕ/ψ parameters for all amino acids, the different nature of glycine (no C^β carbon) motivated the following approach: there is indeed a single set of backbone ϕ/ψ parameters for glycine (defined as $\phi = \text{C}-\text{N}-\text{C}^\alpha-\text{C}$, $\psi = \text{N}-\text{C}^\alpha-\text{C}-\text{N}$), but any other amino acid has an additional set of ϕ'/ψ' parameters which are *added* to the glycine ones (here defined as $\phi' = \text{C}-\text{N}-\text{C}^\alpha-\text{C}^\beta$, $\psi' = \text{C}^\beta-\text{C}^\alpha-\text{C}-\text{N}$). In other words, for any nonglycine amino acid, the dihedral energy is the sum of dihedral energies calculated for ϕ/ψ and ϕ'/ψ' . Note however, that ϕ'/ψ' is calculated for a dihedral angle shifted by $\sim 120^\circ$ because that is the offset of the two torsions (as follows from their definitions using different sets of four atoms). Figure 1 demonstrates the definition of these two sets of dihedral angles on Ala₃ tetrapeptide.

The glycine ϕ/ψ parameters were optimized first, usually based on the best reasonable fit of QM and MM energies for a set of glycine conformers. These are total potential energies for the molecule, where backbone dihedral parameters are the adjustable variables in the optimization. For example, if we optimized the absolute difference of QM and MM energies for glycine tetrapeptide (Gly₃), we would minimize the absolute error (AE) between the MM and QM energies [Eq. (1)]:

$$\text{AE} = 1/N \sum_{i=1}^N |E_{\text{QM}}(i) - E_{\text{MM}}(i)| \quad (1)$$

where $E_{\text{QM}}(i)$ and $E_{\text{MM}}(i)$ correspond to the QM and MM energies respectively for i -th glycine tetrapeptide conformer, and N is the number of all glycine conformers. In our case, MM energy for a given conformer is given by the Amber energy function (see ff94⁵):

$$E_{\text{MM}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{non-bond}} + E_{\text{dihedral}} \quad (2)$$

where the dihedral energy term is

$$E_{\text{dihedral}}(\theta) = \sum_{n=1}^3 V_n (1 + \cos(n\theta - \gamma_n)) \quad (3)$$

V_n is dihedral force constant (amplitude), n is dihedral periodicity, and γ_n is a phase of the dihedral angle θ

(which would be either ϕ or ψ for backbone dihedral terms). The Fourier series in E_{dihedral} is approximated using a small number of terms. Our choice of three terms in the expansion is consistent with the AMBER philosophy of using dihedral terms that can be physically rationalized³⁰ and yields a reasonable number of terms per single dihedral (six) that need to be optimized.

Dihedral energy parameters in Eq. (3) corresponding to backbone ϕ/ψ dihedrals (i.e., V_n and γ_n) are the adjustable parameters. We seek values of these parameters that minimize the function in Eq. (1). Once the glycine parameters are optimized, they are held fixed and used in a second round of fitting where ϕ'/ψ' parameters are optimized. This is carried out by fitting the parameters to best reproduce QM and MM energies for a set of alanine tetrapeptide (Ala₃) conformations. This second round of ϕ'/ψ' optimization was initially misunderstood by us and by others who modified the original ff94. The use of a single round fitting in which the glycine ϕ/ψ dihedral parameters were modified to improve agreement for alanine data, where the ‘real’ ϕ'/ψ' alanine parameters were untouched, led to incorrect parameterization (primarily for glycine) and overall confusion about the effect of dihedral parameters on the backbone conformation.

In the following, we describe the specifics of this general procedure that were employed for the present study. All optimizations were carried out with blocked (acetyl and *N*-methylamine groups at the N- and C-termini, respectively) glycine and alanine tetrapeptides (Gly₃, Ala₃) as shown in Figure 1. A set of Gly₃ and Ala₃ tetrapeptide conformers was chosen to represent local minima on the ff94 energy surface (ff99 was not used here because of apparent flaws in the ff99 energy function for glycine). Because the six-dimensional dihedral space of tetrapeptides is too large to be represented exhaustively (as would be possible for simple two-dimensional surface of dipeptides), we chose to optimize the relative depths of local minima on the energy surface, because these are most relevant to the thermodynamic stability of alternate secondary structure types. Each point in this dihedral space corresponds to a specific tetrapeptide conformation. Local minima were identified through a stochastic search in ϕ/ψ dihedral space followed by local minimization.

All conformations corresponding to the proposed minima were optimized at the Hartree-Fock level with 6-31G* basis set. After discarding the conformers that converged to the same geometry, we ended up with 28 glycine and 51 alanine tetrapeptide conformers (with a maximum range ff94 energy cutoff of 10 kcal/mol). The energies of these conformations were calculated using higher level (gas phase) QM LMP2/cc-pVTZ(-f) in Jaguar,³¹ employing the methodology described by Beachy et al.,³² who showed that inclusion of electron correlation (in this case using LMP2) is important and has a significant effect on relative energies of peptide conformations. Tables of all Gly₃/Ala₃ conformers and their QM energies are provided in the supplementary material (Tables S1 and S2). Note that we only used these energies to refit backbone dihe-

dral parameters, and no change was made to the ff94 partial atomic charges.

The optimization of dihedral parameters was done in two steps. First, the ϕ/ψ torsion parameters for Gly₃ were optimized as follows. We used an exhaustive grid search in 12-dimensional dihedral parameter space: three amplitudes (V_1, V_2, V_3) and three phases ($\gamma_1, \gamma_2, \gamma_3$) for each ϕ and ψ . Amplitudes were systematically varied between 0 and 2 kcal/mol with a step of 0.2, and phases were set to either 0 or π radians. Thus, the total number of grid points in the initial search was $11^6 \times 2^6$ (over 100 million).

The initial grid spacing was 0.2 kcal/mol for amplitudes (V_n), which would be the maximum precision of optimized parameters. To improve this, we employed ‘grid focusing’ by repeating the grid search with finer grid spacing centered on the parameters resulting from the initial course grid. Our refined grid had a spacing of 0.01 kcal/mol for amplitudes, and thus, this is the ultimate precision of our optimized grid parameters.

The function we used for optimization is somewhat more complicated than that shown in Eq. (1). Since the zero of the MM energy function is arbitrary, the optimization should be performed using energy differences between alternate conformations. This was achieved in ff99 by setting one of the energies (the lowest one) to zero, and all other energies were assigned values relative to this “zero energy” reference. However, we observed that the optimized parameter set changes depending on which conformer’s energy is used as a reference. Therefore, the function we optimized was calculated as an average of QM and MM energy differences with each conformer’s energy set as a reference in turn. This gives an average absolute error (AAE) defined as follows:

$$\text{AAE} = 1/N(N-1) \sum_{i=1}^N \sum_{j>i}^N |E_{\text{QM}}^{o,i}(j) - E_{\text{MM}}^{o,i}(j)| \quad (4)$$

where $E_{\text{QM}}^{o,i}(j)$ is the QM energy of conformer j with conformer i as a reference, and $E_{\text{MM}}^{o,i}(j)$ is the MM energy of conformer j with conformer i as a reference, N is number of conformers, which is 28 for Gly₃. Another possibility to define a function for optimization is to concentrate on what the maximum absolute error (MAE) might be when we consider all individual QM and MM differences and, again, take into account that any conformer may serve as a reference “zero energy.” This would give a function of the following form:

$$\text{MAE} = \max_{i,j>i} |E_{\text{QM}}^{o,i}(j) - E_{\text{MM}}^{o,i}(j)| \quad (5)$$

Here, we would search for such set of dihedral parameters that give the smallest MAE. As it turns out, there are many sets of dihedral parameters with very similar AAE. However, inspecting MAE reveals that this parameter is much more sensitive to different parameter sets and therefore the selection of the final set was based on

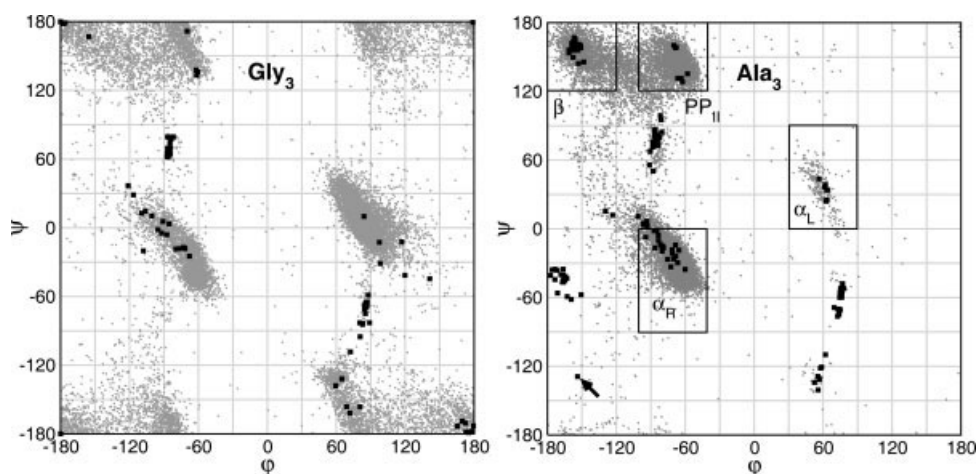


Fig. 2. Ramachandran plot of 28 Gly₃ (left) and 51 Ala₃ (right) conformers as obtained from QM geometry optimization of structures that were acquired using stochastic search with MM energies. All three dihedral pairs are shown for each conformer (black squares). The gray points in the background are ϕ/ψ values of glycine and alanine residues collected from a subset of the PDB. Typical secondary structure regions are outlined by boxes in the Ala₃ plot (see Methods for additional details). Ala₃ conformer number 16 (outlier) in the lower left of Ala₃ plot is designated by an arrow.

this error estimate. In general, small values of MAE always correlate with small values of AAE, such that the best parameter set picked based on MAE will also have one of the lowest AAE values.

For Gly₃ optimization, an additional constraint was introduced, requiring the first phase for ϕ (i.e. γ_1 for ϕ) to be 0. This was necessary to obtain physically reasonable dihedral functions with energy barrier located at $\phi \approx 0^\circ$ on the glycine Ramachandran map. Even though glycine parameter sets without this constraint achieved slightly better fit of QM and MM data, the resulting dihedral energy function would have a physically unreasonable minimum in the $\phi \approx 0^\circ$ region of Ramachandran map (where we also did not have any QM data). This arises from our use of only local minima to train the dihedral correction terms.

Once the Gly₃ parameter set was obtained, the same procedure was repeated with 51 Ala₃ conformers, and parameter sets corresponding to ϕ'/ψ' dihedrals were obtained. Although the resulting set of parameters appeared satisfactory, a more detailed analysis of errors revealed that Ala₃ conformer number 16 was repeatedly responsible for the largest fitting errors (as measured by AAE or MAE). This conformer also falls into the unusual region of the Ramachandran map (Fig. 2 right), and therefore, it was excluded from the optimization of Ala₃ dihedral parameters.

The resulting parameters in the form suitable for use with Amber are provided in supplementary material (Table S3).

Simulations in Explicit Water

All molecular dynamics simulations for Gly₃ and Ala₃ tetrapeptides were carried out with the *sander* module in Amber⁸⁷ using several different force fields as discussed

in the main text: ff94,⁵ ff99,¹¹ Garcia's modification of ff94¹³ with C—N—C $^\alpha$ —C and N—C $^\alpha$ —C—N terms zeroed (denoted ff94gs in the text), Pande's modified ff99¹⁹ with C—N—C $^\alpha$ —C term replaced by the one from ff94 (denoted ff99 ϕ), ff03²⁴ (as present in Amber8 distribution) and ff99SB developed as described earlier. The time step was 2 fs, and all bonds involving hydrogen were constrained by SHAKE with a tolerance of 10^{-4} Å. Glycine and alanine tetrapeptide systems were solvated by ~ 520 TIP3P³³ water molecules in a periodic box. Simulations were carried out in the NPT ensemble at 300 K. A cutoff of 8 Å was used for nonbonded interactions, and long-range electrostatic interactions were treated with the particle mesh Ewald (PME) method.³⁴

Replica exchange molecular dynamics (REMD) simulations,³⁵ as implemented in Amber8, were run for Ala₃ using explicit water solvation and the ff99SB and ff03 force fields. REMD simulations were set up under similar conditions as the corresponding standard MD runs. Ala₃ was solvated in a truncated octahedral box using 595 TIP3P water molecules. Twenty six replicas were used to span the temperature range of 267–571 K. Temperatures were optimized to give a uniform exchange acceptance ratio of $\sim 30\%$. Exchange between neighboring temperatures was attempted every picosecond and each REMD simulation was run for $\sim 30,000$ exchange attempts (30 ns). Because of the high temperatures of some replicas, the NVT ensemble was employed.

Calculation of NMR Order Parameters

Hen egg white lysozyme (PDB code 6LYT³⁶) and ubiquitin (PDB code 1UBQ^{37,38}) simulations were run under the same conditions as the short peptides described earlier, but with truncated octahedral periodic boundary conditions including 4350 and 3300 water molecules, respec-

tively. Equilibration was done in multiple steps, with positional restraints on all heavy atoms in the crystal structure gradually released from 2 to 0 kcal/mol Å.² The production run trajectories were 30 ns long. These trajectories were used to calculate the N–H internuclear vector autocorrelation functions with *ptraj* module. Order parameters (S^2) were obtained from a plateau region of autocorrelation functions.³⁹ Specifically, the autocorrelation function was calculated up to time of half of trajectory length and the mean of last 5 ns was taken as S^2 . Standard deviation of the mean was used to plot the error bars. Since autocorrelation functions for a few backbone amides typically do not converge well (meaning that they do not show clear plateau), the mean will be a less reliable estimate of S^2 for these amides, which will be reflected by larger standard deviations (i.e., larger error bars).

Dihedral Plots and Free Energy Surfaces

A Ramachandran plot showing the distribution of glycine or alanine ϕ/ψ angles in protein crystal structures was constructed based on a subset of proteins from a PDB database, with high resolution (<1.6 Å), small R-factor (<0.25) and less than 20% sequence homology.^{40,41} Dihedral plots based on this statistical survey of the PDB provide experimental indication of the values of backbone torsions that are commonly found in proteins.

Gly₃/Ala₃ simulations were used to produce free energy surfaces (or potential of mean force, PMF, maps) for backbone dihedral angles, assuming that ϕ/ψ are represented in the ensembles according to a Boltzmann distribution. First, the values of all ϕ/ψ dihedral pairs over the production portions of simulations were collected. Then, two-dimensional ϕ/ψ normalized histograms were constructed from these values and converted to free energies using Eq. (6):

$$\Delta G_i = -RT \ln(N_i/N_0) \quad (6)$$

where N_i is the population of a particular histogram bin for specified values of ϕ/ψ , and N_0 is the most populated bin. Thus, the global free energy minimum always has a value of 0 kcal/mol.

The ϕ/ψ histograms were also used to evaluate populations in different regions of secondary structures. The definitions of the four principal regions were as follows: right-handed α -helix (α_R), (ϕ , ψ) $\approx (-70^\circ \pm 30^\circ, -45^\circ \pm 45^\circ)$; left handed α -helix (α_L), $+60^\circ \pm 30^\circ/+45^\circ \pm 45^\circ$; polyproline II (PP_{II}), $-70^\circ \pm 30^\circ, +150^\circ \pm 30^\circ$; extended β -strand conformation (β), $-150^\circ \pm 30^\circ, +150^\circ \pm 30^\circ$. Most ranges are $\pm 30^\circ$ except from the ψ range of α_R and α_L , which were slightly wider ($\pm 45^\circ$) to capture the shift in these regions that resulted from using the ff94gs parameter set. The number of structures in individual regions were summed and divided by the total number of structures to obtain population fractions. Relative fractions are also reported; these differ in that the populations in

the secondary structure basins were normalized by the total number of structures in just the four regions.

The convergence of Ala₃ simulations was estimated by comparing the first and second half of the trajectories in terms of differences of population fractions. The fractions were calculated independently for the two parts of the trajectory and the resulting values are given as averages, with deviations reflecting the differences of the two independent calculations.

Decoy Analysis

Similar to the work that we previously published¹² for ff94 and ff99, we used “decoy analysis” to test the capability of the force field to identify the experimentally determined native structure as that with the lowest potential energy. Three systems were used for decoy analysis: trpzip2⁴² (SWTWENGKWTWK-NH₂), Baldwin type⁴³ alanine based α -helical peptide⁴⁴ (Ace-GGG(KAAAA)₃K-NH₂), and trpcage miniprotein⁴⁵ (NLYIQWLKDGGPSSGRPPPS). The decoys for the three systems were generated by a number of unrestrained as well as forced MD simulations at 300 K as described previously.¹² The sets cover structures generated in simulations started from various initial structures, such as native, extended, and multiple unfolded or partially folded structures from intermediary stages of folding trajectories. The decoy ensembles thus contain local and global structural variation with a greater diversity of structures than could be obtained from a single simulation. The sizes of the three decoy sets are as follows: 177,000 for trpzip, 270,000 for the helical peptide, and 118,000 for trpcage. The energies of these decoy sets were analyzed with different force fields using a Generalized Born solvation model⁴⁶ implemented in Amber. For each force field, the potential energy and root mean square difference (RMSD) from the experimentally determined structure were plotted for each of the decoy structures. The expectation for a “correct” force field is that the lowest energies are obtained for structures with low RMSDs, that is, for native folds.

Because many different structures can have similar RMSD values, and because different force fields may favor slightly different conformations in each RMSD range, we performed a further simplification of the data. Since we are primarily interested in structures with low energies, we only plot one energy value, which is the average of the 20 lowest energies in that particular RMSD bin. This results in a curve that outlines the lowest energies sampled for the RMSD range (“lowest energy profile”), which can be more easily interpreted as being representative of the underlying energy surface for the force field. This approach has several limitations, including neglect of entropic contributions to the actual free energy and also uses an implicit solvation model which may have its own effects on the stability of different conformations. Nevertheless, this methodology is useful for fast qualitative screening of many variations of a force field on larger peptides.

RESULTS AND DISCUSSION

The unique conformations of 28 Gly₃ and 51 Ala₃ that resulted from QM optimization are shown in Figure 2. In general, these conformers overlap the regions outlined by PDB survey data. There are several distinct clusters that fall within the regions of typical protein secondary structures, such as right handed (α_R) and left handed (α_L) α -helix, polyproline II (PP_{II}), and extended β -strand conformations.

Following the procedure described in Methods, we optimized a set of dihedral parameters (Table I). The RMSD between the QM and MM energies for glycine tetrapeptide using these parameters is 1.17 kcal/mol. The average absolute error and maximum error are 0.93 and 2.71 kcal/mol, respectively. Alanine tetrapeptide calculations used the glycine ϕ/ψ backbone parameters as well as the ϕ'/ψ' dihedral terms. In this case, the RMSD, absolute average error and maximum error are 1.31, 1.05, and 3.43 kcal/mol. This is comparable to similar fitting errors obtained in ff99¹¹ and ff03.²⁴ For comparison, the values of backbone dihedral parameters for many other variants of Amber force field are given in the supplementary information (Table S4). It should be noted again that RMSD, absolute average error (AAE) and maximum absolute error (MAE) are calculated such that each conformer in turn is set as a “zero energy” reference, and thus all pairs of energy differences are used in the average. The sensitivity of errors (RMSD, AAE, MAE) to the choice of reference can be demonstrated by calculating these errors for all 51 Ala₃ conformers with a reference conformation fixed to each conformer in turn. Such evaluation of errors produces best RMSD of 0.94 if conformer #31 is a reference and worst RMSD of 1.97 for conformer #45 as a reference. Similarly, depending on the choice of conformer as a reference, AAE ranges between 0.77 and 1.82 and MAE between 1.74 and 3.43 kcal/mol. Thus, averaging over all pair-wise differences should provide a more stringent error estimation (e.g., MAE is always the largest of all pair wise differences when using all pairs) and yields an independence of results with respect to which conformer is chosen as a reference.

Force Field Validation

Explicit water simulations of Gly₃/Ala₃

We performed molecular dynamics simulations on the same systems we used for parameter optimization, that is, Gly₃ and Ala₃ peptides, with different Amber force fields. All simulations were fully solvated with explicit water and extended to ~ 80 ns. Histogram analysis was used to calculate relative free energies of different conformers with respect to ϕ/ψ backbone angles. These free energy (PMF) surfaces are shown in Figure 3 for both Gly₃ and Ala₃.

The correspondence of the distributions obtained from PDB and simulations has been used extensively in force field validation and development.^{21,24–26} This is the basis for our first test and qualitatively evaluates our dihedral parameter set. PDB survey data are presented here as simple Ramachandran maps, not as free energy surfaces.

TABLE I. Optimized Dihedral Parameters: Amplitudes (V) and Phases (γ) for Backbone Dihedral Angles

		V_1	γ_1	V_2	γ_2	V_3	γ_3
ϕ	C—N—C $^\alpha$ —C	0.00	0	0.27	0	0.42	0
ψ	N—C $^\alpha$ —C—N	0.45	π	1.58	π	0.55	π
ϕ'	C—N—C $^\alpha$ —C $^\beta$	2.00	0	2.00	0	0.40	0
ψ'	N—C—C $^\alpha$ —C $^\beta$	0.20	0	0.20	0	0.40	0

The ϕ/ψ terms are always used by all residues, ϕ'/ψ' are used by all residues except glycine. Phases are in radians and amplitudes are in kcal/mol.

One might expect that the ϕ/ψ regions sampled in crystal structures should be also accessible to solvated Gly₃/Ala₃ peptides. However, it would be wrong to assume that the relative populations of ϕ/ψ angles in static crystal structures should match populations in small tetrapeptides solvated in water. Not only are the two environments very different but also populations in crystal structures will certainly be biased by occurrence of common secondary structures, such as α -helices and β -sheets. Furthermore, there is no obvious choice for temperature associated with the PDB distribution, which would make any free energy calculation problematic.

We first discuss the plots for Gly₃. The plots obtained from the Gly₃ simulations are symmetric with respect to the origin, while alanine plots are not due to the presence of the chiral center on the alanine α -carbon. A certain amount of asymmetry in the glycine PDB data arises from the influence of chiral centers in other residues, which are not present in the Gly₃ simulations.

The analysis of glycine conformations in the PDB data reveals that both right- and left-handed α -helical regions are densely populated. PP_{II} and extended β -strand regions are also found with high frequency. One should expect similar conformational preferences resulting from MD simulations (Fig. 3). This is largely true for ff99SB and ff03 and to a lesser extent for ff94 (which is missing PP_{II} and β , and has strong bias favoring α). However, ff99 and ff94gs both show unexpected patterns for glycine conformations, sampling regions which are not represented in PDB survey data. As we noted before, these likely arise from lack of treatment of both sets of backbone parameters (ϕ/ψ and ϕ'/ψ'). The ff99 glycine parameters were fit using alanine data and favor conformations that show no population in the PDB data ($\phi = -180^\circ$ and $\psi = 0^\circ$). The ff94gs surface is also in significant disagreement with PDB data, with a surface that has minima and low barrier heights that appear highly dissimilar to the PDB data and to any of the force fields that do include ϕ and ψ terms. This may artificially increase the glycine conformational transition rates. The plots for ff96, ff99 ϕ as well as the plots obtained from simulations that used our previous modification of dihedral parameters^{12,18} are shown in Figure S6 (supplementary material) and, similarly to ff99 and ff94gs, exhibit incorrect sampling for glycine dihedrals due to the lack of explicit treatment of both sets of backbone parameters.

Alanine free energy maps for the different force fields (Fig. 3) show more similarity to each other, with major regions represented by right handed α -helix, PP_{II} and extended β -conformations. Once again, differences in the relative free energies of the basins are readily apparent in the different force fields. Although ff99SB and ff03 show a reasonable balance between α_R and PP_{II} , ff94 and ff99 show clear bias in favor of α_R . ff99 also shows the same tendency to adopt too-low (close to -180°) values of ϕ that was seen in its corresponding Gly₃ data. ff94gs samples both α_R and PP_{II} , but little population of extended β is present; α_R is also shifted low in ψ . The α_L conformation (with a positive backbone ϕ value) is sampled most in ff99SB and least in ff03. This conformation occurs in several types of β -turns,⁴⁷ particularly those found in β -hairpins, and thus reproducing the energy of α_L region may be important for modeling β -hairpins.

To better interpret such significant differences in relative populations of various secondary structure elements, we turned to available experimental data that provide more direct evidence concerning the dominant structure of short alanine peptides in aqueous solution. All of these observations have important implications for the varied conformational states of proteins and therefore should be captured in the force field if it should ever be used for the description of unfolded proteins or for protein folding studies.

A recent two-dimensional infrared spectroscopy study of alanine dipeptide in aqueous solution⁴⁸ estimated backbone dihedral angles in the range ($-70^\circ \pm 25^\circ$, $+120^\circ \pm 25^\circ$) corresponding to PP_{II} -like conformation. Another structural study of alanine dipeptide using ¹³C NMR⁴⁹ confirmed the presence of PP_{II} conformation, but also suggests that a mixture of PP_{II} and α_R is more likely in water. The dominance of PP_{II} in trialanine was reiterated by two-dimensional vibrational spectroscopy studies of Woutersen and Hamm^{50,51} and later modified by the same authors to include around 20% of α_R apart from PP_{II} .⁵² The experimental studies of Schweitzer-Stenner & Eker on trialanines^{53,54} and tetraalanines⁵⁵ in water using polarized Raman, FTIR, and VCD spectroscopy confirm the dominance of PP_{II} conformation in tetraalanines, while a 50:50 mixture of PP_{II} and extended β -strand-like conformations was observed for trialanines. Other studies,^{56–58} mostly from Kallenbach and colleagues, have confirmed that short alanine peptides form predominantly PP_{II} conformation, which is in temperature dependent equilibrium with extended β -strand conformations. Taken together, these experimental findings suggest that short alanine peptides are found predominantly in PP_{II} -like structures with varying degree of extended β -strand, and possibly smaller fractions of α -helical conformations.

We analyzed the Ala₃ simulation data obtained from each force field in terms of fractional population of local conformational basins corresponding to the four prevalent secondary structure elements. The frequent transitions between PP_{II} , α_R , and β regions result in relatively

rapid thermalization of these basins during the simulations. However, the α_L region is separated from the other basins by a higher barrier near $\phi = 0$, resulting in much less frequent transitions from other regions (see Figure S5 in the supplementary information for further details). To ensure that the relative populations in that region were determined reliably, we ran REMD for ff99SB and ff03, which were the two force fields that provided conformational ensembles most consistent with the PDB data. All results are summarized in Table II. Comparison of standard MD and REMD demonstrates that relative populations in the four regions are essentially unchanged, which indicates that the results are well-converged and that the differences do not arise from poor sampling.

Simulations with ff99SB produce ensembles that are largely PP_{II} (38%) with significant fraction of β (20%) and lower population of α_R (13%). Compared with other force fields, this appears to be in best agreement with experimental observations, in that PP_{II} is the most favorable conformation with lower presence of extended β and even lower fraction of α -helical structures.

ff94 and ff99 show clear preference towards α_R (72% and 36% respectively), with the relative fractions of α_R when compared with the other basins being over 90% for both force fields. The lower absolute fraction in ff99 (36%) arises from the unusually high population of structures around (-150° , 0°) which do not correspond to any of the four canonical secondary structure basins and are also much less populated in other force fields. Overall, the data from ff94 and ff99 are both at variance with experimental observations.

ff03 behaves most similarly to ff99SB, but it may still be slightly over stabilizing α -helices as the most populated basin is the α_R conformation (34%). It is also somewhat unsettling that ff03 samples very poorly the α_L region that is important for β -turn conformations. We obtained no sampling of α_L in standard 80-ns-long MD simulations (0.0% fraction), indicating that this region is at least 5 kcal/mol (this is based on our cutoff for free energy plots, Fig. 3) less stable than the most populated α_R region. However, REMD simulations do sample α_L region with 0.06% population fraction. The sampling is insufficient to make a reliable estimate (see ϕ versus time plots in Figure S5 of supplementary material) of free energy difference between α_R and α_L , but based on the populations in those regions, α_L is at least 3.8 kcal/mol higher than α_R . Even though ff99SB only adopts 2–3% population in that region, it is still the largest compared with other force fields.

Simulations with ff99 ϕ result in essentially the same behavior as was obtained for ff94, with the relative basin populations overwhelmingly dominated by α_R conformations (97 vs. 95% for ff94, see relative fractions of individual secondary structures in Table II and free energy surfaces in Figure S6 of supplementary material). Thus, at least at the level of glycine/alanine tetrapeptide, the small change of a single ϕ term in ff99 ϕ does not seem to result in significantly different or improved behavior.

TABLE II. Fractions (Percentages) of the Four Secondary Structures in Ala₃ MD Simulations in Explicit Water

Force field	Region			
	PP _{II}	β	α_R	α_L
ff99SB	38.1 \pm 1.2 (51.9 \pm 1.7)	20.1 \pm 0.7 (27.4 \pm 0.5)	13.0 \pm 1.6 (17.8 \pm 2.6)	2.1 \pm 0.9 (2.8 \pm 1.2)
ff99SB (REMD)	36.9 \pm 0.1 (50.1 \pm 0.3)	20.5 \pm 0.8 (27.9 \pm 1.0)	12.7 \pm 0.5 (17.2 \pm 0.7)	3.4 \pm 0.1 (4.6 \pm 0.1)
ff03	28.0 \pm 0.4 (37.8 \pm 0.4)	12.4 \pm 0.5 (16.7 \pm 0.5)	33.6 \pm 0.5 (45.4 \pm 0.9)	0.0 \pm 0.0 (0.0 \pm 0.0)
ff03 (REMD)	28.6 \pm 0.6 (38.2 \pm 0.6)	12.6 \pm 0.4 (16.9 \pm 0.5)	33.4 \pm 0.7 (44.7 \pm 1.1)	0.1 \pm 0.0 (0.1 \pm 0.0)
ff94	2.7 \pm 0.5 (3.5 \pm 0.7)	0.8 \pm 0.1 (1.1 \pm 0.2)	72.0 \pm 0.5 (94.6 \pm 0.6)	0.5 \pm 0.2 (0.7 \pm 0.3)
ff99	1.0 \pm 0.0 (2.5 \pm 0.1)	2.4 \pm 0.3 (6.3 \pm 0.8)	35.8 \pm 0.4 (90.8 \pm 0.5)	0.1 \pm 0.1 (0.3 \pm 0.3)
ff94gs	20.7 \pm 3.1 (30 \pm 6.0)	2.3 \pm 0.3 (3.4 \pm 0.7)	46.3 \pm 6.8 (66.0 \pm 6.5)	0.4 \pm 0.2 (0.6 \pm 0.2)
ff99 ϕ	1.6 \pm 0.7 (2.2 \pm 0.9)	0.4 \pm 0.2 (0.5 \pm 0.2)	68.8 \pm 0.5 (96.6 \pm 0.6)	0.4 \pm 0.4 (1.2 \pm 1.2)

See Methods for definitions of secondary structure regions. Numbers in parentheses are relative fractions, only considering populations within the four regions.

Decoy analysis of longer peptides

A critical property of the energy function is whether it can distinguish the native state such that it corresponds to the lowest (free) energy. We previously employed “decoy analysis”¹² as a rapid test of how well the energy function identifies the native structure. By plotting the potential energy against RMSD from the native structure one can readily see if the native state generally has lowest energies. This has the advantage that energy profiles for multiple force fields can be compared without obtaining converged ensembles for each peptide/force field combination. As a result of neglecting entropy and the use of an implicit water model, however, the results do not provide quantitative stability data for any of the peptides simulated.

We show the “lowest profile” (see Methods) of energies plotted versus RMSD from the native structure (Fig. 4) with the expectation of seeing lowest potential energies (including solvation free energy) for small RMSD values, reflecting that the native basin indeed can be distinguished based merely on the energy function. We performed decoy screening using several peptides that have been experimentally demonstrated to adopt stable secondary and/or tertiary structures. Trpzip2 represents an exceptionally well-defined β -hairpin,⁴² stabilized by cross-strand pairs of indole rings. The Baldwin type alanine-based peptide⁴⁴ was designed to adopt α -helix. Lastly, the trpcage mini-protein⁴⁵ has a mixture of secondary structure elements (α , 3^{10} and polyproline).

We previously used decoy analysis of the trpzip2 β -hairpin to demonstrate the extreme helical bias of the ff94/ff99 force fields,¹² showing that nonnative α -helical decoys had the lowest energies and were also the most populated in simulations. Therefore it comes as no surprise that the present decoy plots again clearly reveal the inadequacy of the ff94 and ff99 force fields to identify the native state of the hairpin. Indeed the pronounced minimum around 4–5 Å [Fig. 4(A)] represents a large set of helical structures. Consistent with our observations from the tetrapeptide simulations, ff99 ϕ behaves much like ff94. The native trpzip2 structure scores as lowest in potential energy only with ff99SB. ff03 also performs

much better than ff94/ff99 but, as we noted before, may still be slightly too favorable for helices. ff94gs behaves similarly to ff03.

On the other hand, all tested force fields successfully score the α -conformation of the Baldwin helix with lowest energy. As expected, the energy gap between the helical conformation and those with larger RMSD values is greater for the force fields with strong helical bias (ff94, ff99, ff99 ϕ). Even though ff94gs performed almost as well as ff99SB for trpzip2, it scores the Baldwin helix even more favorably than the ff94/ff99 force fields that are known to over stabilize α -helix conformations. Likewise, all of the force fields correctly identify the trpcage native fold.

Overall, both ff99SB and ff03 perform well in all three decoy cases. ff99 ϕ follows very closely the behavior of ff94. Since ff94gs and ff99 ϕ have not been used as extensively as the original Amber force fields upon which they are based (ff94/ff99), and because they have not performed well on the tetrapeptide and decoy screens, we will not consider them in further validation.

Comparison with NMR parameters for lysozyme and ubiquitin

Hen egg white lysozyme has become a standard for evaluating the quality of force fields by comparing internal dynamics parameters calculated from MD simulations to NMR relaxation experiments.^{59–63} The degree of backbone flexibility is specified by experimentally derived order parameters S^2 , which correspond to amide bond N–H librational motion. Lower S^2 values reflect increased backbone flexibility. We generated 30-ns MD simulations of lysozyme with explicit water, using four different parameter sets: ff94, ff99, ff99SB and ff03. Order parameters calculated from these simulations are compared with experimentally derived values in Figure 5(A). The RMSD and the correlation coefficient between experimental and calculated S^2 are given in Table III. It can be seen that the agreement of ff99SB derived values with the experiment is very good, with the best correlation to experimental data and only half of the RMSD compared with all other force fields.

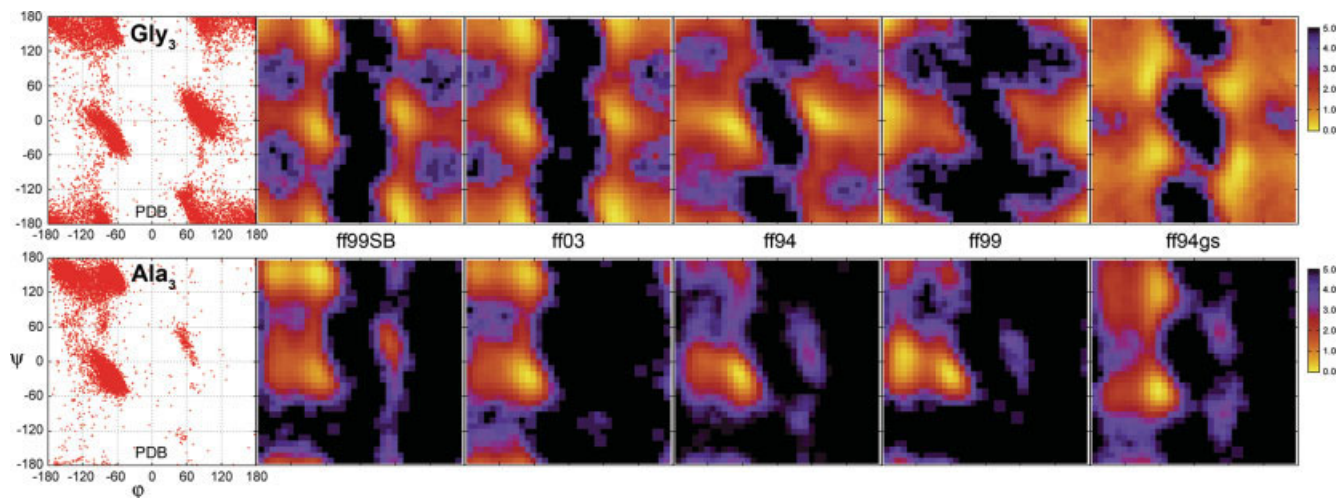


Fig. 3. Free energy ϕ/ψ maps for Gly₃ (top row) and Ala₃ (bottom row) from 80-ns simulations with explicit TIP3P water. The energies are color coded from 0 up to 5 kcal/mol. PDB survey data is represented by a simple Ramachandran plot. Individual force fields are designated as: ff99SB (this work), ff03,²⁴ ff94,⁵ ff99,¹¹ and ff94gs.¹³

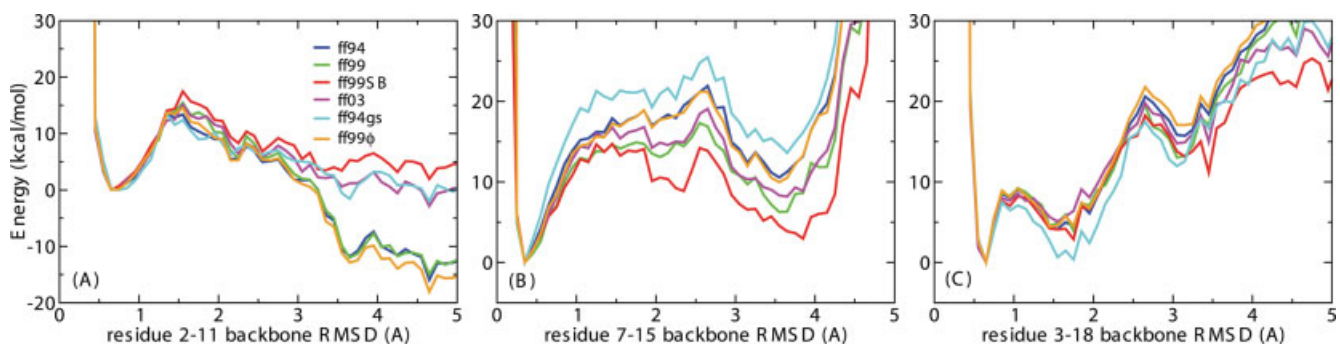


Fig. 4. "Lowest energy" profiles for three prototypical decoy systems, each tested with six Amber force field variants (ff94, ff99, ff99SB, ff03, ff94gs, ff99 ϕ): (A) trpzip2, which is known to adopt a β -hairpin, (B) a Baldwin-type sequence as a representative of an α -helix, and (C) trpcage as a representative of mixed secondary structures. RMSD values are calculated using the experimentally determined structure as a reference. Ideally, a force field should show lowest energies for the lowest RMSD values.

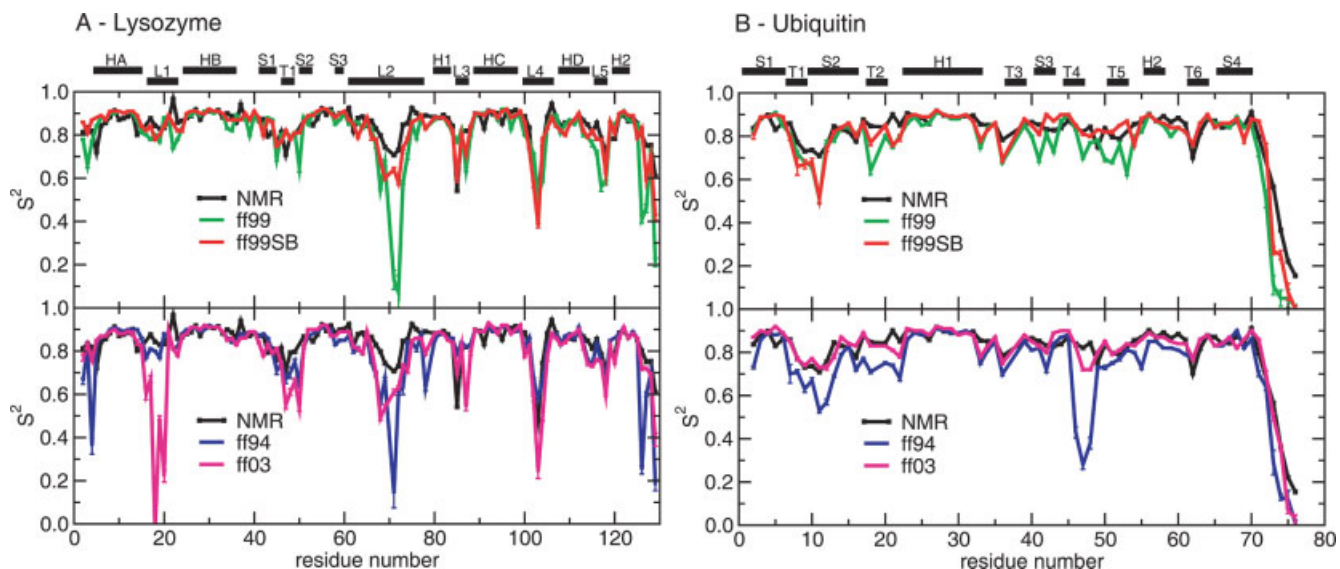


Fig. 5. Order parameters (S^2) derived from experiment (black line) and calculated from MD simulations in explicit water with different force field parameter sets: ff99 (green), ff99SB (red), ff94 (blue) and ff03 (magenta). Error bars reflect the convergence of calculated S^2 values. (A) Lysozyme. Secondary structures of lysozyme are labeled: helix A (HA: residues 4–15), loop 1 (L1: 16–23), helix B (HB: 24–36), strand 1 (S1: 41–45), turn 1 (T1: 46–49), strand 2 (S2: 50–53), strand 3 (S3: 58–60), long loop 2 (L2: 61–78), 3_{10} helix 1 (H1: 80–84), loop 3 (L3: 85–89), helix C (HC: 89–99), loop 4 (L4: 100–107), helix D (HD: 108–115), loop 5 (L5: 116–119) and 3_{10} helix 2 (H2: 120–124). (B) Ubiquitin. Secondary structure elements in ubiquitin: strand 1 (S1: residues 1–7), turn 1 (T1: 7–10), strand 2 (S2: 10–17), turn 2 (T2: 18–21), helix 1 (H1: 23–34), turn 3 (T3: 37–40), strand 3 (S3: 41–44), turn 4 (T4: 45–48), turn 5 (T5: 51–54), helix 2 (H2: 56–59), turn 6 (T6: 62–65) and strand 4 (S4: 66–71). Differences are described in the text.

TABLE III. Correlation of Experimental and Calculated Order Parameters (S^2) for Several AMBER Force Fields

	ff94	ff99	ff99SB	ff03
Lysozyme				
RMSD	0.11	0.12	0.06	0.13
R ^a	0.61	0.62	0.83	0.62
Ubiquitin				
RMSD	0.13	0.10	0.07	0.05
R	0.83	0.93	0.95	0.96

^aLinear correlation coefficient.

ff99SB describes well the increased flexibility of residues in loops L2 and L4 as well as residue 84 in L3 (which is not described well by other force fields). As has been noted previously,⁶⁴ molecular mechanics (MM) force fields perform most poorly in the loop regions, usually exaggerating the flexibility of loop residues. This is confirmed by our simulations where the most obvious disagreements with experimental order parameters occur in loops. Some of these discrepancies might be explained by a particular bias/weakness of the specific force field that was revealed in our model system studies. For example, we demonstrated that ff03 did not sample the α_L region of ϕ/ψ space for alanine tetrapeptide. The pronounced failure of ff03 to match experimental data in the L1 loop (with simulated S^2 values much too low) may be a result of the α_L conformation that is adopted by several L1 residues in the crystal structure. The local instability in this conformation may give rise to increased structural dynamics in the ff03 simulations. Interestingly, significantly lower order parameters specifically in regions with glycines occupying the α_L conformation were also noted for Charmm22 force field by its developers.²⁶

We also showed that ff94/ff99 are both biased towards helical structures, lacking sampling in PP_{II} and β regions. This is likely the cause for increased flexibility in the middle of the long L2 loop, where several residues occupy the extended β -like conformation. This problem may be exacerbated in ff99 because of the presence of two glycines; we showed earlier that the glycine free energy surface in ff99 is highly distorted. Overall, many of the problematic regions are found to contain glycines, further supporting our assertion of the inadequacy of glycine parameters in many Amber force field variants.

We repeated the same type of analysis for another protein, human ubiquitin. This is another small and well-characterized protein of 76 residues with a known crystal structure,^{37,38} for which NMR relaxation parameters were also determined experimentally.⁶⁵ Based on the crystal structure, the C-terminal region extends away from and makes few contacts with the remainder. This is reflected by very low S^2 parameters for the four terminal residues. Again, Figure 5(B) graphically compares the ability of the four force fields to reproduce experimental S^2 parameters. The quantitative indicators for the agreement with experimental data are given in Table III. For this system, all four Amber force fields work reasonably

well, with no large discrepancies with the exception of the region near turn T4 when using ff94. All force fields but ff03 exaggerate flexibility of turn T1 connecting strands S1 and S2. Overall, ff03 performs best, even though ff99SB is nearly the same and both provide slight improvement over ff94/ff99. Examination of trajectories did not reveal any specific structural differences; rather the variations appear to arise from the magnitude of fluctuations about similar average structures.

Overall, these simulations serve as a good example of sensitivity of internal dynamics, as measured by NMR order parameters, to particular details of the MM force field. Also, it is clear from these two sets of simulations that the quality of agreement to experimental relaxation data is somewhat system dependent. All force fields achieved better agreement with S^2 values for ubiquitin than for lysozyme.

CONCLUSIONS

We have shown that several variants of AMBER force field have deficiencies arising from inadequate backbone dihedral term parameterization. The original ff94 force field introduced backbone dihedral treatment to be able to reproduce relative energies of selected glycine and alanine dipeptide conformers. Because of limited computational resources at the time, only a few conformers were included in the fitting procedure. We have shown here and previously that ff94 strongly favors helical conformations. More recent modifications of ff94 attempted to rectify the problem, but only partially succeeded. All of them failed to recognize that not just one, but two sets of dihedral parameters control backbone preferences in ff94. The first set is optimized solely for glycine. Any other amino acid will use both glycine set of parameters and a second set of dihedral parameters which are usually parameterized using alanine. Prior modifications of ff94, such as ff96, ff99, ff94gs, ff99 ϕ , and our previous modification of ff99¹⁸ only changed the first set of dihedrals, effectively overwriting glycine dihedral parameters with the ones intended for alanine, unaware that there existed a set of alanine dihedral parameters that do not use the traditional definition of ϕ and ψ . This resulted in incorrect parameterization of glycine. Our data demonstrate (Fig. 3, Figure S6) that these parameter sets result in conformational ensembles for glycine peptides that are in significant disagreement with both the PDB data and with force fields that explicitly include glycine in the fitting process. We therefore recommend against using these parameter sets for protein simulations.

The modification of ff94/ff99 force field presented in this work not only aims to rectify the problems caused by misunderstanding of parameterization procedure, but also endeavors to improve the parameterization initially done in ff94. This was primarily achieved through two courses of action. First, we used a large number of glycine and alanine tetrapeptide local minima instead of a small number of dipeptide conformations when fitting QM and MM energies. We believe that, unlike the dipeptide, tetrapeptide dihedral space bears closer resemblance to that

found in proteins, and also has a local minimum for the α -helical conformation, permitting direct fitting to avoid over stabilization that was observed in earlier variants. Second, we used a procedure that avoids biasing the outcome of the dihedral parameters optimization with respect to a single a priori chosen reference conformation.

The resulting parameter set was shown to result in better balance of prevalent secondary structures, such as PP_{II} , α -helix, and extended/ β -strand in simulations of short glycine and alanine peptides in water. This was supported by better agreement of dihedral angle regions sampled in the simulations with regions generated from PDB survey data. Here, the improvement for glycine residue is most noticeable, where some of the force fields produced free energy surfaces that bore no resemblance to those observed for real proteins. The relative sampling of the typical secondary structure regions in simulations of alanine tetrapeptide using the new force field is in good agreement with experimental data which used various techniques to infer the prevalent states of short alanine peptides. Beyond short glycine and alanine peptides, the qualitative performance of the new parameters was tested on several larger systems representing typical secondary structures: trpzip2 β -hairpin, Baldwin type α -helix, and the trpcage miniprotein. ff99SB correctly identified native structure as that with the lowest energy in all three cases, which was not observed for any of the other force field variants. Finally, we performed long explicit water simulations of lysozyme and ubiquitin and tested the performance of different force field variants by comparing relaxation order parameters calculated from simulations to experimental NMR values. Once again, the newly derived parameters performed the best. The changes we introduced with ff99SB modification are extremely minimal and therefore excellent compatibility should be preserved with existing parameters developed to be compatible with ff94/ff99. This is in contrast to the ff03 protein parameters, for which no compatible parameters currently exist for small molecules and most importantly nucleic acids. It is therefore unclear whether ff03 could be used for simulations of these complexes (which would include potentially incompatible charge models).

As usual, caution should be taken when any force fields are used beyond their range of validity. For example, one can expect a substantial effect of solvation model used. Implicit solvation models, such as Generalized Born, may shift the balance between different secondary structure elements. If, for example, PP_{II} is indeed stabilized by specific interactions with explicit water as recently suggested,^{57,66–68} such stabilization is not possible with implicit solvents, and thus structural preference may change. In fact, we recently reported that GB solvation models increase α -helical propensity.⁶⁹ Different approaches in force field development might be needed to account for these effects, and force fields should likely be tested with the particular characteristics (and perhaps idiosyncrasies) of these models in mind.

Force field development is a complex process in which many approximations are typically required. In the pres-

ent case, conformational energies from high-level gas-phase QM calculations were used to fit dihedral parameters for a force field in which the partial charges should more closely represent those in aqueous solution. Additionally, the large difference in the size of systems that are used to train biomolecular force fields and those for which they are intended to be used results in a slow maturation period, often with deficiencies becoming apparent only after extensive use. For this reason, we extensively used these parameters on a variety of systems not discussed here before making this report. For example, the improved glycine parameterization was important in our recent simulations of HIV-1 protease,⁷⁰ where glycine rich “flaps” play an essential role in protease dynamics. Another study⁶⁹ reported our hybrid replica exchange method, where ff99SB was used to test the methodology based on conformational sampling of polyalanine peptides of varying lengths. Our recent simulations⁷¹ of fragments of the villin headpiece with ff99SB also achieved good agreement with experimental trends⁷² in calculated J-coupling constants and helical propensities. These studies have supplied further evidence that the ff99SB parameter set provides a reasonable parameter set for protein simulations, in the context of an efficient, additive molecular mechanics model.

ACKNOWLEDGMENTS

The authors thank Alex MacKerell, for useful discussions concerning order parameters, and Angel Garcia and Vijay Pande, for helpful discussions and providing details for their AMBER force field modifications. Supercomputer time at NCSA (NCSA MCA02N028 to CS and MCA05S010 to AER) is gratefully acknowledged. C.S. is a Cottrell Scholar of Research Corporation.

REFERENCES

1. Rick SW, Stuart SJ. Potentials and algorithms for incorporating polarizability in computer simulations. In: Lipkowitz KB, Boyd DB, editors. Reviews in computational chemistry, Vol. 18. New York: VCH; 2002; pp 89–146.
2. Ponder JW, Case DA. Force fields for protein simulations. *Adv Protein Chem* 2003;66:27–85.
3. Mackerell AD. Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem* 2004;25:1584–1604.
4. Jorgensen WL, Tirado-Rives J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc Natl Acad Sci USA* 2005;102:6665–6670.
5. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 1995;117:5179–5197.
6. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P. Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. *Comput Phys Commun* 1995;91:1–41.
7. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. *J Comput Chem* 2005;26:1668–1688.
8. Jorgensen WL, Tiradorives J. The opls potential functions for proteins—energy minimizations for crystals of cyclic-peptides and crambin. *J Am Chem Soc* 1988;110:1657–1666.

9. Kuyper LF, Hunter RN, Ashton D, Merz KM, Kollman PA. Free-energy calculations on the relative solvation free-energies of benzene, anisole, and 1,2,3-trimethoxybenzene—theoretical and experimental-analysis of aromatic methoxy solvation. *J Phys Chem* 1991;95:6661–6666.
10. Kollman P, Dixon R, Cornell W, Fox T, Chipot C, Pohorille A. The development/application of the “minimalist” organic/biochemical molecular mechanic force field using a combination of ab initio calculations and experimental data. In: van Gunsteren WF, Weiner PK, Wilkinson AJ, editors. *Computer simulations of biomolecular systems*, Vol. 3. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1997; pp 83–96.
11. Wang JM, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 2000;21:1049–1074.
12. Okur A, Strockbine B, Hornak V, Simmerling C. Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins. *J Comput Chem* 2003;24:21–31.
13. Garcia AE, Sanbonmatsu KY. α -Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci USA* 2002;99:2782–2787.
14. Kamiya N, Higo J, Nakamura H. Conformational transition states of a β -hairpin peptide between the ordered and disordered conformations in explicit water. *Protein Sci* 2002;11:2297–2307.
15. Higo J, Ito N, Kuroda M, Ono S, Nakajima N, Nakamura H. Energy landscape of a peptide consisting of α -helix, 3(10)-helix, β -turn, β -hairpin, and other disordered conformations. *Protein Sci* 2001;10:1160–1171.
16. Ono S, Nakajima N, Higo J, Nakamura H. Peptide free-energy profile is strongly dependent on the force field: comparison of C96 and AMBER95. *J Comput Chem* 2000;21:748–762.
17. Wang L, Duan Y, Shortle R, Imperiali B, Kollman PA. Study of the stability and unfolding mechanism of BBA1 by molecular dynamics simulations at different temperatures. *Protein Sci* 1999;8:1292–1304.
18. Simmerling C, Strockbine B, Roitberg AE. All-atom structure prediction and folding simulations of a stable protein. *J Am Chem Soc* 2002;124:11258–11259.
19. Sorin EJ, Pande VS. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys J* 2005;88:2472–2493.
20. Sorin EJ, Pande VS. Empirical force-field assessment: the interplay between backbone torsions and noncovalent term scaling. *J Comput Chem* 2005;26:682–690.
21. 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-Struct Funct Genet* 2003;50:451–463.
22. Zaman MH, Shen MY, Berry RS, Freed KF, Sosnick TR. Investigations into sequence and conformational dependence of backbone entropy, inter-basin dynamics and the flory isolated-pair hypothesis for peptides. *J Mol Biol* 2003;331:693–711.
23. Mu YG, Kosov DS, Stock G. Conformational dynamics of trialanine in water, Part 2: Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments. *J Phys Chem B* 2003;107:5064–5073.
24. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong GM, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang JM, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 2003;24:1999–2012.
25. MacKerell AD, Feig M, Brooks CL. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 2004;126:698,699.
26. Mackerell AD, Feig M, Brooks CL. 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;25:1400–1415.
27. Feig M, MacKerell AD, Brooks CL. Force field influence on the observation of π -helical protein structures in molecular dynamics simulations. *J Phys Chem B* 2003;107:2831–2836.
28. Jorgensen WL, Maxwell DS, TiradoRives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 1996;118:11225–11236.
29. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. 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;105:6474–6487.
30. Bowen JP, Allinger NL. Molecular mechanics: the art and science of parametrization. In: Lipkowitz KB, Boyd DB, editors. *Reviews in computational chemistry*, Vol. 2. New York: VCH; 1991; pp 81–98.
31. Jaguar. Version 4.1. Portland, OR: Schrodinger; 2000.
32. Beachy MD, Chasman D, Murphy RB, Halgren TA, Friesner RA. Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. *J Am Chem Soc* 1997;119:5908–5920.
33. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–935.
34. Darden T, York D, Pedersen L. Particle mesh Ewald—an N Log(N) method for Ewald sums in large systems. *J Chem Phys* 1993;98:10089–10092.
35. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 1999;314:141–151.
36. Young ACM, Dewan JC, Nave C, Tilton RF. Comparison of radiation-induced decay and structure refinement from X-ray data collected from Lysozyme crystals at low and ambient-temperatures. *J Appl Crystallogr* 1993;26:309–319.
37. Vijaykumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* 1987;194:531–544.
38. Vijaykumar S, Bugg CE, Wilkinson KD, Cook WJ. 3-Dimensional structure of ubiquitin at 2.8 Å resolution. *Proc Natl Acad Sci USA* 1985;82:3582–3585.
39. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules, Part 1: Theory and range of validity. *J Am Chem Soc* 1982;104:4546–4559.
40. Wang GL, Dunbrack RL. PISCES: a protein sequence culling server. *Bioinformatics* 2003;19:1589–1591.
41. Wang GL, Dunbrack RL. PISCES: recent improvements to a PDB sequence culling server. *Nucleic Acids Res* 2005;33:W94–W98.
42. Cochran AG, Skelton NJ, Starovasnik MA. Tryptophan zippers: stable, monomeric β -hairpins. *Proc Natl Acad Sci USA* 2001;98:5578–5583.
43. Marqusee S, Baldwin RL. Helix stabilization by Glu–... Lys+ salt bridges in short peptides of denovo design. *Proc Natl Acad Sci USA* 1987;84:8898–8902.
44. Fesinmeyer RM, Peterson ES, Dyer RB, Andersen NH. Studies of helix fraying and solvation using C-13’ isotopomers. *Protein Sci* 2005;14:2324–2332.
45. Neidigh JW, Fesinmeyer RM, Andersen NH. Designing a 20-residue protein. *Nat Struct Biol* 2002;9:425–430.
46. Hawkins GD, Cramer CJ, Truhlar DG. Pairwise solute descreening of solute charges from a dielectric medium. *Chem Phys Lett* 1995;246:122–129.
47. Wilmot CM, Thornton JM. Analysis and prediction of the different types of β -turn in proteins. *J Mol Biol* 1988;203:221–232.
48. Kim YS, Wang JP, Hochstrasser RM. Two-dimensional infrared spectroscopy of the alanine dipeptide in aqueous solution. *J Phys Chem B* 2005;109:7511–7521.
49. Mehta MA, Fry EA, Eddy MT, Dedeo MT, Anagnost AE, Long JR. Structure of the alanine dipeptide in condensed phases determined by C-13 NMR. *J Phys Chem B* 2004;108:2777–2780.
50. Woutersen S, Hamm P. Structure determination of trialanine in water using polarization sensitive two-dimensional vibrational spectroscopy. *J Phys Chem B* 2000;104:11316–11320.
51. Woutersen S, Hamm P. Isotope-edited two-dimensional vibrational spectroscopy of trialanine in aqueous solution. *J Chem Phys* 2001;114:2727–2737.
52. Woutersen S, Pfister R, Hamm P, Mu YG, Kosov DS, Stock G. Peptide conformational heterogeneity revealed from nonlinear vibrational spectroscopy and molecular-dynamics simulations. *J Chem Phys* 2002;117:6833–6840.

53. Eker F, Cao XL, Nafie L, Schweitzer-Stenner R. Tripeptides adopt stable structures in water. A combined polarized visible Raman, FTIR, and VCD spectroscopy study. *J Am Chem Soc* 2002;124:14330–14341.
54. Eker F, Griebenow K, Schweitzer-Stenner R. Stable conformations of tripeptides in aqueous solution studied by UV circular dichroism spectroscopy. *J Am Chem Soc* 2003;125:8178–8185.
55. Schweitzer-Stenner R, Eker F, Griebenow K, Cao XL, Nafie LA. The conformation of tetraalanine in water determined by polarized Raman, FT-IR, and VCD spectroscopy. *J Am Chem Soc* 2004;126:2768–2776.
56. Shi ZS, Chen K, Liu ZG, Ng A, Bracken WC, Kallenbach NR. Polyproline II propensities from GGXGG peptides reveal an anticorrelation with β -sheet scales. *Proc Natl Acad Sci USA* 2005;102:17964–17968.
57. McColl IH, Blanch EW, Hecht L, Kallenbach NR, Barron LD. Vibrational Raman optical activity characterization of poly(L-proline) II helix in alanine oligopeptides. *J Am Chem Soc* 2004;126:5076,5077.
58. Shi ZS, Olson CA, Rose GD, Baldwin RL, Kallenbach NR. Polyproline II structure in a sequence of seven alanine residues. *Proc Natl Acad Sci USA* 2002;99:9190–9195.
59. Soares TA, Daura X, Oostenbrink C, Smith LJ, van Gunsteren WF. Validation of the GROMOS force-field parameter set 45A3 against nuclear magnetic resonance data of hen egg lysozyme. *J Biomol NMR* 2004;30:407–422.
60. Stocker U, van Gunsteren WF. Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data. *Proteins-Struct Funct Genet* 2000;40:145–153.
61. Buck M, Boyd J, Redfield C, Mackenzie DA, Jeenes DJ, Archer DB, Dobson CM. Structural determinants of protein dynamics—analysis of 15 N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg-white lysozyme. *Biochemistry* 1995;34:4041–4055.
62. Buck M, Karplus M. Internal and overall peptide group motion in proteins: molecular dynamics simulations for lysozyme compared with results from X-ray and NMR spectroscopy. *J Am Chem Soc* 1999;121:9645–9658.
63. Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell AD. Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys J* 2006;90:L36–L38.
64. Case DA. Molecular dynamics and NMR spin relaxation in proteins. *Acc Chem Res* 2002;35:325–331.
65. Tjandra N, Feller SE, Pastor RW, Bax A. Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation. *J Am Chem Soc* 1995;117:12562–12566.
66. Fleming PJ, Fitzkee NC, Mezei M, Srinivasan R, Rose GD. A novel method reveals that solvent water favors polyproline II over β -strand conformation in peptides and unfolded proteins: conditional hydrophobic accessible surface area (CHASA). *Protein Sci* 2005;14:111–118.
67. Mezei M, Fleming PJ, Srinivasan R, Rose GD. Polyproline II helix is the preferred conformation for unfolded polyalanine in water. *Proteins* 2004;55:502–507.
68. Mezei M, Srinivasan R, Fleming PJ, Rose GD. Solvent effect on the conformational preference of poly alanine. *Biophys J* 2004;86:630A.
69. Okur A, Wickstrom L, Layten M, Geney R, Song K, Hornak V, Simmerling C. Improved efficiency of replica exchange simulations through use of a hybrid explicit/implicit solvation model. *J Chem Theory Comput* 2006;2:420–433.
70. Hornak V, Okur A, Rizzo RC, Simmerling C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc Natl Acad Sci USA* 2006;103:915–920.
71. Wickstrom L, Okur A, Song K, Hornak V, Raleigh DP, Simmerling C. The unfolded state of the villin headpiece helical subdomain: computational studies of the role of locally stabilized structure. *J Mol Biol* 2006;360:1094–1107.
72. Tang YF, Rigotti DJ, Fairman R, Raleigh DP. Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: The villin headpiece subdomain. *Biochemistry* 2004;43:3264–3272.