Motional timescale predictions by molecular dynamics simulations: Case study using proline and hydroxyproline sidechain dynamics

Abil E. Aliev,* Martin Kulke, Harmeet S. Khaneja, Vijay Chudasama, Tom D. Sheppard, and Rachel M. Lanigan

Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

ABSTRACT

We propose a new approach for force field optimizations which aims at reproducing dynamics characteristics using biomolecular MD simulations, in addition to improved prediction of motionally averaged structural properties available from experiment. As the source of experimental data for dynamics fittings, we use $^{13}$C NMR spin-lattice relaxation times $T_1$ of backbone and sidechain carbons, which allow to determine correlation times of both overall molecular and intramolecular motions. For structural fittings, we use motionally averaged experimental values of NMR $J$ couplings. The proline residue and its derivative 4-hydroxyproline with relatively simple cyclic structure and sidechain dynamics were chosen for the assessment of the new approach in this work. Initially, grid search and simplexed MD simulations identified large number of parameter sets which fit equally well experimental $J$ couplings. Using the Arrhenius-type relationship between the force constant and the correlation time, the available MD data for a series of parameter sets were analyzed to predict the value of the force constant that best reproduces experimental timescale of the sidechain dynamics. Verification of the new force-field (termed as AMBER99SB-ILDNP) against NMR $J$ couplings and correlation times showed consistent and significant improvements compared to the original force field in reproducing both structural and dynamics properties. The results suggest that matching experimental timescales of motions together with motionally averaged characteristics is the valid approach for force field parameter optimization. Such a comprehensive approach is not restricted to cyclic residues and can be extended to other amino acid residues, as well as to the backbone.

INTRODUCTION

Molecular dynamics (MD) simulations are widely employed for structural and dynamics characterizations of peptides and proteins.1–4 These simulations rely mainly on classical force field parameters, such as AMBER,5–7 CHARMM,8 GROMOS9,10 and OPLS-AA.11,12 Amongst different types of force field parameters, backbone, and sidechain torsional potentials have been the subject of extensive reoptimizations, leading to improved modifications of AMBER13–16 and CHARMM17,18 force fields. Based on a number of detailed benchmark studies,19–33 AMBER99SB14 has emerged as one of the force fields which reproduces experimentally measured parameters with better accuracy compared to other force fields. This force field has undergone further useful refinements in recent years.15,16 To predict the correct balance of secondary structure propensities in proteins, a simple backbone energy correction was introduced to reproduce the fraction

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Abil E. Aliev, Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom. E-mail: A.E.Aliev@ucl.ac.uk

Received 29 April 2013; Revised 1 June 2013; Accepted 6 June 2013

Published online 2 July 2013 in Wiley Online Library (wileyonlinelibrary.com).

DOI: 10.1002/prot.24350
of helix measured in short peptides at 300 K, with the modified force field known as AMBER99SB*. Recently, the AMBER99SB force field has been improved further (known as AMBER99SB-ILDN) by refitting the amino acid sidechain torsion potentials of the AMBER99SB force field for four residues: isoleucine, leucine, aspartic acid, and asparagine.

One of the important properties not exploited in the force field optimizations for biomolecular MD simulations is the timescale of motion for a given backbone or sidechain fragment. As a result, while the motionally averaged experimental NMR parameters can be reproduced well by new force fields, the timescale over which this averaging is achieved may deviate significantly from experiment. The reason for the lack of timescale verifications is that either experimental data is not available or it is not clear how the force field parameters can be modified to reproduce better the experimental data. To explore the possibilities that involve experimentally known motional timescales in force field optimizations, we have selected a relatively simple example of the proline (Pro) sidechain dynamics in this work. The simplicity of the Pro dynamics arises from the fact that unlike other amino acid residues the Pro residue has a unique cyclic structure, which interconverts continuously between two conformers, known as C⁷-endo and C⁷-exo. Another factor in favor of the Pro residue is that numerous theoretical and experimental studies have been undertaken in the past focusing mainly on the pyrrolidine ring dynamics. Furthermore, the torsional parameters of the Pro residue have not been optimized in the past and standard force field parameters obtained for open chain fragments are used for proline. The result is that the predicted geometry of the pyrrolidine ring by AMBER force fields is relatively flat compared to single-crystal X-ray diffraction data or quantum-mechanical (QM) calculations, as judged by the value of the endocyclic torsion J₂ (Fig. 1) or the pseudorotation amplitude Jm also known as the maximum puckering angle. In the first approximation, the nonplanarity of the pyrrolidine ring can be assessed by how far atoms Cβ and Cy are placed from the plane formed by the remaining three atoms. The further they are from the plane, the higher the absolute values of J₂ and Jm are. We note that changes in geometry of the ring have also further energetic implications, and, as shown previously, the larger the maximum puckering angle the larger the pyrrolidine ring interconversion barrier in Pro and hydroxyproline (Hyp) residues. The increase in the energy barrier implies less frequent transitions or longer motional timescales. Based on these considerations, force field optimizations may potentially improve the accuracy of MD simulations for predicting both the structure and dynamics of the Pro residue in proteins. Note that one of the important attributes of the Pro residue is its hinge-like function, which enhances the probability of β-turns in proteins. Therefore, accurate predictions of the proline structure and dynamics may have critical implication on the outcome of MD descriptions of proteins.

Returning to the original problem of force field optimizations, we expect that the introduction of an additional dynamics constraint into force field optimizations should be advantageous from a methodological point of view, as multiple solutions are often found in force field optimizations which fit equally well experimental data. This is not surprising, as the experimental data consists of motionally averaged values of NMR J couplings and chemical shifts, which are dependent on the relative populations of conformers, but not on how fast they exchange. Timescale fittings combined with fittings of NMR J couplings and/or chemical shifts are expected to select a correct solution in such cases. Unlike previous optimizations based on the quantum-mechanical calculations, we will use experimentally measured NMR J-couplings in our initial re-optimization of the Pro sidechain torsion potentials. The approach used by us is based on either simple grid search or iterative fittings of experimental NMR data, in which a figure-of-merit function is evaluated using MD trajectories calculated for each trial set of parameters. Once torsional force fields reproducing experimental NMR J-couplings have been identified, we will probe MD-predicted timescales of motions which best match experimental data. NMR spin-lattice relaxation times will be used to estimate both overall and intramolecular timescales of motions. In addition to the Pro residue, we will also reoptimize torsional force field parameters for the trans-4-hydroxy-L-proline residue (Hyp) to match experimental dynamics data.
**MATERIALS AND METHODS**

**NMR Data**

Apart from Ace-Hyp-NHMe (AHM) and Ace-Hyp-Gly (AHG), all other peptides were used as received from Sigma Aldrich and Cambridge Bioscience. The synthesis of AHM and AHG is described in Supporting Information. Experimental values of proton \(^1\)J values determined from the full line-shape analysis. Unless otherwise specified, the trans-orientation about the amide bond preceding the Pro (or Hyp) residue is assumed for a given peptide. For the values of \(^1\)J values determined from the full line-shape analysis, the standard deviation was estimated to be <0.1 Hz. Experimental values of \(^1\)J values for ubiquitin were estimated to be of the order of ~0.1 Hz.

Solution \(^1\)H NMR spectra were recorded on a Bruker Avance III 600 MHz NMR spectrometer equipped with a 5 mm cryoprobe (\(^1\)H 600.13 MHz and \(^13\)C 150.90 MHz). Data acquisition and processing were performed using standard TopSpin (version 2.1) software. \(^1\)H and \(^13\)C chemical shifts were calibrated using deuterostatic and deuteranionic shifts in D\(_2\)O (\(^1\)H 3.75 ppm, \(^13\)C 67.19 ppm). Uncertainties in measured values of \(^1\)H and \(^13\)C chemical shifts were typically better than ±0.01 ppm. Unless otherwise specified, NMR measurements were carried out at 298 K. High (>300 K) and low (<300 K) temperature calibrations were carried out using standard samples of 80% 1,2-ethanediol in DMSO-d_6 and 4% CH\(_3\)OH in CD\(_3\)OD, respectively.

The \(^13\)C spin-lattice relaxation times were measured for solutions of peptides in either D\(_2\)O or H\(_2\)O:D\(_2\)O (9:1) using a standard inversion-recovery technique with the \(^13\)C observation in the presence of proton decoupling. To minimize errors associated with low signal-to-noise ratios, these experiments were carried out on a 600 spectrometer with a dual channel \(^1\)H/\(^13\)C cryoprobe with the sensitivity optimized for \(^13\)C measurements. From five independent measurements carried out at probe ambient temperature (293 K) for the 214 mM solution of GPPG in D\(_2\)O at different dates over 60 days, the standard deviations for \(^13\)C \(T_1\) measurements were within 0.4–1.4% of the corresponding mean values. From three independent measurements carried out at 298 K for the 77 mM VAPG in H\(_2\)O:D\(_2\)O (9:1), the standard deviations for \(^13\)C \(T_1\) measurements were within 0.2–1.1% of the corresponding mean values.

Chemical shift anisotropies (\(\Delta\sigma\), in ppm) of aliphatic carbons were measured using slow MAS measurements (2.5 kHz) on a Bruker AVANCE III 850 spectrometer equipped with a 4 mm CPMAS probe and a solid sample of L-proline. The estimated \(\Delta\sigma\) values (~43 ppm for \(\alpha\) and −30 ppm for \(\gamma\) of L-proline) were used in calculations of correlation times using \(^13\)C \(T_1\) values. From the \(^13\)C \(T_1\) calculations, at \(\Delta\sigma = −43\) ppm, the dipolar relaxation mechanism remains the dominant factor determining \(^13\)C \(T_1\) relaxation at 14.1 T, while chemical shift anisotropy accounts for <1% of \(T_1\) values.

**MD calculations and simplex fittings**

All MD simulations were carried out using GROMACS (version 4.5.5). One molecule of NAcPro molecule (terminated with CO\(_2^-\) and with a Na\(^+\) cation added for neutralization) was solvated with 147 water molecules in a dodecahedral box with a volume of 4.7 nm\(^3\) in MD simulations. Periodic boundary conditions and the TIP3P water model were employed in all MD simulations. An integration step of 2 fs was used and neighbor lists were updated every 5th step. The particle mesh Ewald (PME) method was employed for the electrostatics with fourth-order interpolation. The neighbor list and the real-space cutoff distances were set to 0.9 nm, which is similar to that used in optimizations of the original force field and its recent modifications. The van der Waals interactions in all MD simulations were treated with a twin-range cutoff method using the neighbor list and van der Waals cutoff distances. The value of the van der Waals cutoff distance was 0.9 nm. The temperature at 298 K was controlled using velocity rescaling with a stochastic term (V-rescale) and a time constant of 0.1 ps. A Parrinello–Rahman scheme was employed for pressure control at 1 bar using a coupling constant of 2 ps and an isothermal compressibility of 4.5 × 10\(^{-5}\) bar\(^{-1}\). Prior to production MD runs, including those implemented within downhill simplex optimizations, the system was minimized using steepest-descent and conjugate gradient algorithms. Minimization steps were followed by four steps of equilibration. The system was equilibrated for 40 ps with the positionally restrained solute molecule to allow water molecules to equilibrate around it, followed by a NVT molecular dynamics for 100 ps, NPT dynamics for 200 ps and another NVT dynamics for 200 ps. Reproducible production MD simulations at each step of simplex fittings were performed for 7.5–40.5 ns using NVT ensemble, the first 0.5 ns of which was discarded from the calculations of averaged NMR parameters. For the selected set of parameters from simplex fittings additional 200 ns log MD simulations were carried out.

The vicinal \(3\)J couplings of the five-membered pyrroline ring in NAcPro (as well as in other peptides, see below) in each frame of MD simulations were calculated using empirically optimized Karplus-type equations 8C and 8D of Haasnoot et al. These equations contain...
terms accounting for the differences in electronegativities of α- and β-substituents, and hence are better suited for the analysis of the $^{3}J$ couplings of the pyrrolidine ring than the original Karplus equation.\textsuperscript{47} The precision of equation 8C of Haasnoot et al. (expressed as the rms deviation) for a structural fragment containing 2 substituents (-CHXY-CH$_{2}$Z-) is estimated as 0.367 Hz using a set of 45 experimental $^{3}J_{HH}$ couplings.\textsuperscript{46} The precision of equation 8D of Haasnoot et al. for a structural fragment containing 3 substituents (-CHXY-CH$_{2}$Z-) is estimated as 0.485 Hz using a set of 100 experimental $^{3}J_{HH}$ couplings.\textsuperscript{46}

To analyze MD trajectories, including those obtained at each step of simplex fittings, dihedral angles were extracted for each frame recorded every 0.01 ps during the MD simulation. The calculated values of $^{3}J$ couplings using the corresponding dihedral angles in each frame were used to calculate the averaged values of $^{3}J$ couplings over the duration of the MD simulation. The rms deviation defined as $\sqrt{\frac{1}{N} \sum_{i=1}^{N} (J_{i}^{exp} - J_{i}^{calc})^2}$ (denoted as rms$^{p}$J for the $^{3}J_{HH}$-couplings of the pyrrolidine ring) was used as a figure-of-merit function in simplex fittings, where $J_{i}^{exp}$ and $J_{i}^{calc}$ are conformationally averaged experimental and calculated couplings, respectively, and $N$ is the number of different $J$ couplings available ($N = 10$ for the Pro sidechain). As simplex may in principle lead to a local minimum of the merit function,\textsuperscript{44,45} it is important to consider several sets of starting values of the optimized parameters $x_i$. This was achieved by varying the factor $c$, by which one of the optimized parameters $x_i$ is varied within the first $n + 1$ steps of the simplex run using the following expression: $x_i + c x_i$ (i.e., at step $n = 1$ the initial values of $x_i$ from the original AMBER99SB force field are used followed by $x_i + c x_i$, $x_i + 2c x_i$ . . . $x_i$ at step $n = 2$, etc.). Several simplex fittings were considered with $c$ varied between 0.2 and 5 (see the main text for further details). In addition, for $|c| < 1$, both positive and negative values were considered. An additional constraint requiring $x_i > 0$ was imposed in simplex fittings.

For further optimization and validation of newly derived force field parameters, 800 ns MD simulations of GPGG, VAPG, Gly-Pro-Phe (GPF), 1.5 μs MD simulations of angiotensin II, 1 μs MD simulations of human ubiquitin (PDB entry 1UBQ),\textsuperscript{48} 60 ns and 1.5 μs MD simulations of AHM and 1.5 μs MD simulations of AHG were carried out. One molecule of zwitterionic GPGG was solvated with 253 water molecules in a dodecahedral box with a volume of 8.3 nm$^3$. For VAPG, one molecule of zwitterionic peptide was solvated with 260 water molecules in a dodecahedral box with a volume of 8.4 nm$^3$. In the case of GPF, one molecule of zwitterionic peptide was solvated with 292 water molecules in a dodecahedral box with a volume of 9.3 nm$^3$. Similarly, one molecule of angiotensin II (with a Cl$^-$ anion added for neutralization) was solvated with 1201 water molecules in a dodecahedral box with a volume of 40.8 nm$^3$. One molecule of ubiquitin (with six Na$^+$ cations added for neutralization) was solvated with 2605 water molecules in a cubic box with a volume of 91.1 nm$^3$. For the Hyp parameter optimizations, one molecule of AHM was solvated with 225 water molecules in a dodecahedral box with a volume of 7.4 nm$^3$ and one molecule of AHG (with a Na$^+$ cation added for neutralization) was solvated with 300 water molecules in a dodecahedral box with a volume of 9.4 nm$^3$. Other conditions and parameters of MD simulations were the same as described above for NAcPro. Frames recorded every 1 ps were used in estimating averaged $^{3}J$-couplings from MD simulations of GPGG and ubiquitin.

The calculated $^{3}J_{HH}$ couplings are expected to depend on the length of the MD simulation. To estimate the significance of this dependence, we have considered MD simulations of varying lengths. Calculations of $^{3}J_{HH}$ couplings using 600, 700, and 800 ns long MD simulations of GPGG using the modified force field (referred to as (25), Table 1) showed the largest variation of less than ±0.023 Hz in the calculated $^{3}J_{HH}$ values over 200 ns change in the length of the MD simulation (<0.5% of the value of the $^{3}J_{HH}$ coupling). Two MD simulation of GPGG with 800 ns and 3 μs lengths were available for the parameter set (19), with the third largest value of $V_{i}$ considered (6.92437 kJ mol$^{-1}$, Table 1). These were used for error estimates in MD-predicted quantities. The changes were (see Tables (I–IV) for definitions of parameters): $P_{exp.0}^0$, $P_{endo.0}^0$, $\chi_{exp.0.0.1}^0$, $\chi_{endo.0.0.1}^0$, $\chi_{endo.0.0.9}^0$, rms$^{p}$J$^{0.025}$ Hz, $\rho_{i}^0$ 0 %, $d_{i.0}^0$ 0 Å, $N^{\theta.1}$ +0.41, $N^{\phi.2} +0.04$, $N^{\psi.2} -0.33$, $N^{\psi.1} -0.29$, $N^{\psi.1} -0.02$, $N^{\psi.2} +0.01$, $S^{2} 0$, $\tau_{i} -0.02$ ps. The negative sign here corresponds to the decrease of the value on increasing the length of the MD simulation. The absolute values of these changes can be considered as an estimate of the upper limit of errors involved, as the value of the force constant in parameter set (19) is higher than that in the final selected set (25), hence requiring longer MD simulations for better convergence in calculated parameters in the case of (19).

The motionally averaged $^{3}J$-couplings of the peptide backbone of GPGG and ubiquitin were calculated using quantum-mechanically derived Karplus relationships\textsuperscript{31,49} and empirically parameterized Karplus equations.\textsuperscript{50,51}

Interatomic distances from the MD simulations of GPGG were calculated in a manner similar to that used in NMR measurements\textsuperscript{36}: (i) internuclear distances ($r_{i}$) for pairs of hydrogen atoms were calculated in each MD frame; (ii) a quantity equal to $r^{-6}$ was calculated as a measure of the expected NOE in each frame, $\eta$; (iii) the sum of $r^{-6}$ were used as a measure of the expected total NOE over the full length of the MD run; (iv) using $r = 2.4$ Å as the reference $H^{a}$-$H^{b}$ distance in the Pro residue,\textsuperscript{36} internuclear distances for other proton pairs were calculated using the $\eta \sim r^{-6}$ relationship.

As shown by Tropp,\textsuperscript{52} when overall molecular motions are relatively slow and intramolecular motions are
where \( h \) the sequence of four bonded peptide backbone atoms and \( \mathbf{C}(Pro) - C(Pro) - C(Pro) - C(Pro) \) are the second order Legendre polynomial. Prior to the \( C(t) \) calculations, the overall rotational and translational motions of the solute molecule were removed from the MD trajectory. This was accomplished by superimposing the sequence of four bonded peptide backbone atoms \( C(Pro) - C(Pro) - C(Pro) - C(Pro) \) on the corresponding atoms of the snapshot at the midpoint of the production run, chosen as the reference structure. A similar approach was used by Showalter and Brüschweiler in their detailed analysis of NMR relaxation data (for a detailed discussion see Section 2.3 of Ref. 25). The Lipari–Szabo model was used to fit the initial 20 ns of the autocorrelation \( C(t) \) functions.

\[
C(t) = S^2 + (1 - S^2) e^{-t/\tau}
\]

In Eq. (2) above, \( S^2 \) denotes the order parameter and \( \tau \) is the autocorrelation time for the intramolecular \( C-H \) bond reorientations.

Quantum-mechanical calculations

All quantum-mechanical calculations were carried out using Gaussian 09. Geometry optimizations were carried using various combinations of QM methods and basis sets, as described in the main text. The “norsym” keyword of Gaussian 09 was employed to carry out QM calculations with the symmetry of molecules disabled. For DFT M06-2X geometry optimizations, the ultrafine numerical integration grid (with 99 radial shells and 590 angular points per shell) was used, combined with the “verytight” convergence condition (requesting the root-mean-square forces to be smaller than 1 \( \times 10^{-6} \)).
Conformational Populations and Geometries of the Pro ring in GPGG in Water as Predicted by NMR and by 800-ns Long MD Simulations Using Various Sets of Torsional Parameters for the Pro residue

|        | $P_{\text{exo}}$ (%) | $P_{\text{endo}}$ (%) | $\chi_m$ (%) | $\chi_{\text{endo}}$ (%) | $\text{rms}_{\text{rel}}$ (Hz) |
|--------|-----------------------|-----------------------|--------------|--------------------------|-------------------------------|
| AMBER99SB | 14 | 180 | 35.3 | 58.9 | 0.662 |
| 1      | 14 | 181 | 36.5 | 59.0 | 0.618 |
| 2      | 13 | 180 | 36.4 | 59.2 | 0.601 |
| 3      | 14 | 180 | 36.7 | 59.1 | 0.588 |
| 4      | 13 | 180 | 36.4 | 59.2 | 0.611 |
| 5      | 13 | 181 | 37.7 | 59.3 | 0.561 |
| 6      | 11 | 182 | 41.3 | 56.3 | 0.562 |
| 7      | 12 | 181 | 40.0 | 59.3 | 0.562 |
| 8      | 13 | 181 | 37.0 | 58.9 | 0.544 |
| 9      | 12 | 182 | 39.3 | 57.3 | 0.486 |
| 10     | 12 | 181 | 38.7 | 58.9 | 0.520 |
| 11     | 13 | 181 | 38.3 | 58.7 | 0.517 |
| 12     | 12 | 181 | 38.3 | 58.6 | 0.511 |
| 13     | 12 | 181 | 38.4 | 58.2 | 0.499 |
| 14     | 13 | 181 | 37.6 | 59.1 | 0.547 |
| 15     | 12 | 180 | 37.8 | 59.1 | 0.540 |
| 16     | 13 | 181 | 37.9 | 58.4 | 0.513 |
| 17     | 11 | 182 | 39.8 | 58.3 | 0.523 |
| 18     | 12 | 181 | 40.2 | 59.3 | 0.572 |
| 19     | 11 | 181 | 40.0 | 57.9 | 0.524 |
| 20     | 12 | 181 | 38.8 | 59.1 | 0.526 |
| 21     | 12 | 181 | 38.1 | 59.2 | 0.530 |
| 22     | 12 | 181 | 38.0 | 59.0 | 0.525 |
| 23     | 12 | 181 | 38.5 | 58.4 | 0.502 |
| 24     | 13 | 180 | 37.7 | 57.7 | 0.483 |
| 25     | 12 | 181 | 38.7 | 59.2 | 0.529 |
| NMR35  | 11 | 189 | 41   | 54.3 | 0.49* |

*From least-squares fittings of the vicinal $^3J$-couplings using Eqs. (8C) and (8D) of Haasnoot et al.66

Hartree Bohr$^{-1}$). Additional frequency calculations were also undertaken to verify that the optimized geometries correspond to true minima. The reaction field field IEPCM63,64 was used to account for water solvent effects. The jump angles $\Delta \theta$ of the C–H bonds as a result of the pyrrolidine ring interconversion were determined using Python Molecular Viewer (version 1.5.4).65 Calculations employing MP2 and M06-2X methods were also carried out in which a selected dihedral angle was incremented or decremented in 5° steps. Basis sets considered are specified in the main text. At each step the selected dihedral angle was fixed with all the remaining degrees of freedom optimized using MP2 or M06-2X QM calculations. A relaxed 1D potential energy surface scan was performed in this manner and minimized QM energies at each step were obtained. The QM-optimized structures were then used in molecular mechanics (MM) calculations using AMBER99SB force field to obtain the corresponding MM energies (see the main text for further details).

**Conformational notation**

The original conformational notation proposed by Haasnoot et al. for L-prolines are used in this work.66 The exo- and endo-orientations of the Pro ring carbon $C'$ are defined relative to the substituent (COO or CONH groups) at the $C''$ carbon of the Pro ring. The definition of endo- and exocyclic torsional angles is shown in Figure 1.

The pseudorotation phase angle, $P$, which identifies a given conformation on the pseudorotation circle, is defined by the maximum value attained by $\chi_m$, which is the maximum value attained by $\chi_m$, $\chi_m = (A^2 + B^2)^{1/2}$, (3)

where

$$A = \frac{2}{5} \sum_{i=1}^{5} \chi_i \cos \left( \frac{4\pi}{5} (i-2) \right)$$

$$B = -\frac{2}{5} \sum_{i=1}^{5} \chi_i \sin \left( \frac{4\pi}{5} (i-2) \right)$$  

Note that 180° is added to the calculated value of $P$ if $\chi_2 < 0$. From the distributions of endocyclic torsional angles, a two-site exchange between $C', \text{endo}$ and $C', \text{exo}$ conformations of the pyrrolidine ring of Pro and Hyp residues was observed in MD simulations of the peptides considered. The populations of these ring conformations are denoted as $x_{\text{endo}}$ and $x_{\text{exo}}$ (in % with $x_{\text{endo}} + x_{\text{exo}} = 100\%$).

**RESULTS**

**Initial simplex MD fittings of experimental NMR data**

In our initial revision of the AMBER99SB force field we undertook simplex fittings of $^3J_{HH}$-couplings, which comprised the optimization of the C-C-C-C dihedral parameters for the endocyclic carbons in the Pro residue of N-acetyl-L-proline (NACPro) and Gly-Pro-Gly-Gly (GPGG). The choice here is dictated by the fact that accurate experimental data is available for NACPro and GPGG.34–36 In particular, full lineshape analysis was employed to derive accurate experimental values of $^3J_{HH}$-couplings in $D_2O$ solutions, with the estimated standard deviation $\leq 0.03$ Hz for vicinal couplings.34 As for the choice of the force field, the analysis of >10 different force fields applied to GPGG, identified AMBER99SB as the force field which reproduces best experimentally measured NMR parameters in aqueous solutions.31 Thus, further improvement of this force field presents a challenging task for the simplex fittings of $^3J_{HH}$ couplings.

While AMBER99SB predicts satisfactorily the relative energies of $C', \text{exo}$ and $C', \text{endo}$ conformations (as
The geometry of the pyrrolidine ring as predicted by AMBER99SB MD simulations is flatter (\( \chi_m \approx 35^\circ \)), where \( \chi_m \) is approximately the same as the largest of the ring endocyclic torsions \( \chi_1-\chi_5 \), which is usually \( \chi_2 \) compared to NMR, X-ray and QM calculations (\( \chi_m = 37^\circ-42^\circ \)).\textsuperscript{34,35} The reason for such a difference is that the same set of dihedral C—C—C—C parameters is used in AMBER force fields for both the cyclic (e.g., C\( ^\alpha -\)C\( ^\beta -\)C\( ^\gamma -\)C\( ^\delta \) in Pro corresponding to the endocyclic torsion \( \chi_2 \)) and open chain systems (see Ref. 69 for details of how the C—C—C—C parameter was derived).

For our initial simplex optimizations, a standard AMBER dihedral energy term of the following form was used:

\[
E_{\text{dih}}(\theta) = \frac{1}{2} \sum_{n=1}^{3} V_n \left( 1 + \cos (n\theta - \gamma_n) \right)
\]

where \( V_n \) represents dihedral force constant (amplitude), \( n \) is dihedral periodicity and \( \gamma_n \) with the value of either 0\(^\circ \) or 180\(^\circ \) is a phase of the dihedral angle \( \theta \). The dihedral force constants, \( V_n \), were optimized to obtain the best agreement between experimental and MD-predicted values of \( J_{HH} \)-couplings of NAcPro. These are optimized for the angle \( \chi_2 \) (Fig. 1), which is usually the largest amongst the endocyclic dihedral angles \( \chi_1-\chi_5 \) for the Pro sidechain in peptides and proteins. There are three non-zero \( V_n \) values (\( V_1 \), \( V_2 \), and \( V_3 \)) for the \( \chi_2 = \text{CT-CT-CT} \) torsion (CT denotes tetrahedral carbon) in the original AMBER99SB force field. Thus, three parameters \( V_1 \), \( V_2 \), and \( V_3 \) were optimized in our simplex fittings, each step of which consisted of MD simulation followed by the calculation of the MD-averaged \( J_{HH} \).
Table IV
Intramolecular Autocorrelation Times \( \tau_c \) (in ps) and Order Parameters \( S^2 \) for the C\( ^=\)H Bond Reorientations of Pro in GPGG as Predicted by 800-ns MD simulations

| Parameter set | \( V_p \) (kJ mol\(^{-1}\)) | \( S^2 \) | \( \tau_c \) ps | rms\(^a\) |
|---------------|----------------|---------|-----------------|---------|
| AMBER99SB     | 0.75312        | 0.33    | 4.1             | 0.0017  |
| 1             | 1.37935        | 0.32    | 5.6             | 0.0020  |
| 4             | 1.06169        | 0.32    | 5.9             | 0.0020  |
| 2             | 1.29055        | 0.31    | 6.4             | 0.0022  |
| 3             | 1.75728        | 0.31    | 7.1             | 0.0023  |
| 5             | 1.95811        | 0.69    | 11.0            | 0.0011  |
| 8             | 2.25938        | 0.30    | 11.1            | 0.0029  |
| 14            | 2.39584        | 0.30    | 11.6            | 0.0037  |
| 15            | 2.6885         | 0.30    | 12.9            | 0.0029  |
| 16            | 3.028          | 0.29    | 15.2            | 0.0033  |
| 11            | 3.30976        | 0.29    | 16.9            | 0.0034  |
| 12            | 3.58557        | 0.29    | 20.0            | 0.0040  |
| 13            | 3.79243        | 0.28    | 21.6            | 0.0040  |
| 23            | 4.06           | 0.28    | 25.4            | 0.0043  |
| 10            | 4.17167        | 0.29    | 26.2            | 0.0040  |
| 20            | 4.42712        | 0.28    | 29.3            | 0.0046  |
| 22            | 4.6633         | 0.28    | 34.0            | 0.0049  |
| 21            | 4.816241       | 0.28    | 36.6            | 0.0048  |
| 9             | 5.51626        | 0.27    | 54.3            | 0.0059  |
| 17            | 6.35714        | 0.27    | 52.3            | 0.0076  |
| 7             | 6.61951        | 0.28    | 51.7            | 0.0079  |
| 19            | 6.92437        | 0.27    | 112.0           | 0.0085  |
| 18            | 7.17114        | 0.27    | 124.2           | 0.0091  |
| 6             | 9.31503        | 0.25    | 440.2           | 0.0147  |
| 24            | 7.4146         | 0.30    | 531.5           | 0.0177  |
| 25            | 4.3474         | 0.29    | 28.7            | 0.0045  |
| NMR           | —              | 0.27(1) | 29.7(4)         | —       |

\( ^a \)The fitting errors (rms, arbitrary units with \( C(t) = 1 \) at \( t = 0 \) ps) are shown.

\( ^b \)The sum of \( V_1, V_5 \) and \( V_6 \) is shown.

couplings using the modified Karplus equation of Haasnoot et al.\(^{46} \)

Prior to deciding the length of MD simulation within simplex fittings, we examined the convergence of the population of endo conformation (\( \chi^{endo} \), in %) using a 500 ns long MD simulation (Fig. S1, Supporting Information). The results indicate that after the initial \( \sim 10 \) ns the populations of the two conformers have converged sufficiently. In particular, after 10 ns MD run the population of the endo conformer is 56.6% compared to 56.3% after 20 ns, 56.5 after 100 ns and 56.7% after 500 ns. Even in the region between 1.5 and 10 ns, the population deviations are within less than \( \pm 2.0\% \) (Fig. S1, Supporting Information). We have therefore used 7.5 ns long MD simulations at each step in our simplex fittings. The first 0.5 ns were considered as equilibration period and the corresponding data were discarded from calculations of averaged \( ^3J_{HH} \)-couplings. Up to 10 different simplex MD simulations were carried out using different scaling factors \( c \) between \( -0.5 \) and 5, with 50–200 steps of 7.5 ns long MD simulations in each case.

The original AMBER99SB values of force field parameters, together with those derived from our simplex fittings of experimental \( ^3J_{HH} \)-couplings are shown in Table I. Five sets of optimized parameters (1)–(5) were selected from simplex fittings, showing the rms deviations from the experimental \( ^3J_{HH} \)-couplings (rms\(_p\), in Hz) less than 0.8 Hz based on 7 ns long MD simulations. For comparison, rms\(_p\) = 0.96 Hz for the original AMBER99SB force field. Considering that the increase in force constants during simplex optimizations may lead to longer convergence times, we used additional 200-ns long MD simulations for final estimates of merit functions (rms\(_p\)) for parameter sets (1)–(5) and AMBER99SB. The results of these simulations are summarized in Table I.

As can be seen from Table I, parameter sets (1)–(5) obtained from simplex MD simulations show 5–14\% improvements in rms values compared to the original AMBER99SB force field. The \( \chi_m \) values in (1)–(5) have slightly increased compared to that in the original force field, which are in better agreement with the NMR, XRD and QM results (37–42).\(^{34,35} \) From the \( E_{\text{in}}(\gamma_2) \) graphs for the CT-CT-CT-CT fragment (Fig. S2, Supporting Information), it can be seen that the \( E_{\text{in}}(\gamma_2) \) graphs for the parameter sets (1)–(4) show higher maxima at \( \gamma_2 = 0^\circ \), the values of which correspond to the value of \( V_3 \), since \( n = 1 \) and \( n = 2 \) terms of Eq. (5) are zero at \( \gamma_2 = 0^\circ \) as \( \gamma_1 = \gamma_2 = 180^\circ \). In the transition state between the \( C^\alpha\)-endo and \( C^\gamma\)-exo conformations of the pyrrolidine ring, the value of \( \gamma_2 \) is \( 0^\circ \). Thus, the increase of the \( V_3 \) value here corresponds to the decrease of the activation energy of the interconversion. Based on the Arrhenius relationship, the increase of the activation energy is expected to lead to the decrease of the frequency of transitions between the \( C^\alpha\)-endo and \( C^\gamma\)-exo states.

The above results suggest that relatively short MD simulations combined with subsequent long MD simulations using selected sets can be applied for the refinement of force field parameters provided that the force constants do not increase significantly. Note that the simplex fittings described in this work generate a new MD trajectory for each trial set of parameters to evaluate the rms deviation between experimental and MD-predicted NMR data, that is, new conformations are created at each step of fittings (see Single Trajectory Reweighting Approach section below). However, the disadvantage of the current method is that it is computationally expensive and relatively large increase in optimized parameters may not be described adequately by short MD simulations used in simplex fittings.

### QM optimizations of force field parameters

After initial simplex MD simulations, we considered QM optimizations of force field parameters followed by iterative MD simulations for further refinement of the force field parameters obtained from QM fittings. Four sets of QM calculations were considered to estimate the dependence of the results on the choice of the basis set.
and the QM method, as well as to assess the level of uncertainty involved: M06-2X/def2-TZVP, M06-2X/6-31G(d,p), M06-2X/cc-pVTZ and MP2/6-31+G(d). Based on previous studies,\cite{35,70} these QM methods and basis sets reproduce relative conformational energies and geometries in good agreement with experimental data. Calculations of 31 conformers of NAcPro were carried out in which the {C}α-{C}β-{C}γ-{C}δ dihedral angle ($\chi_2$) was varied in 5° steps between $-75°$ and $+75°$. The QM predicted energy profiles in the gas phase and in water (using IEFPCM)\cite{63,64} are compared in Figure 2. Considering relative energies of Cγ-endo and Cγ-exo conformers (with the corresponding $\chi_2$ values at $-40°$ and $+40°$, respectively), the experimentally measured ratio of two conformers in water ($x_{endo}=61%$ and $x_{exo}=39%$)\cite{34} are best reproduced by IEFPCM(H2O) MP2/6-31+G(d) and M06-2X/def2-TZVP calculations [Fig. 2(c)]. The predicted populations of the Cγ-endo were 66 and 71%, respectively, by IEFPCM(H2O) MP2/6-31+G(d) and M06-2X/def2-TZVP calculations. Thus, the results from these two sets of calculations were used in our further analysis.

The following merit function of Lindorff-Larsen et al.\cite{16} was used in our fittings\cite{16}:

$$\Phi = \sum_{i=1}^{M} \left( E_{i}^{QM} - E_{i}^{MM} \right)^2 e^{-\beta E_{i}^{QM}}$$  \hspace{1cm} (6)

where $E_{i}^{QM}$ and $E_{i}^{MM}$ are the QM and molecular mechanics (MM) energies, respectively, and $M$ is the number of conformations optimized at the QM level (31 in this case). The inverse temperature, $\beta$, is set to 1.0 mol kcal$^{-1}$ (see discussion in Ref. 16 regarding the choice of $\beta$ value). Adopting the approach developed by Lindorff–Larsen et al.,\cite{16} the $E_{i}^{MM}$ energy is given by the

---

**Figure 2**

QM-predicted energy profiles as a function of the endocyclic pyrrolidine torsion angle $\chi_2$ in NAcPro (a) in water and (b) in the gas phase. Expansions of (a) and (b) are shown in (c) and (d), respectively.
AMBER99SB energy, $E^{AMBER99SB}$, plus a new torsion term, that replaces the existing AMBER99SB torsion, $V^{AMBER99SB}(\theta)$:

$$E^{MM} = E^{AMBER99SB} - V^{AMBER99SB}(\theta) + k_0 \sum_{n=1}^{N} V_n [1 + \cos \left( n\theta - \gamma_n \right)]$$  

(7)

where $k_0$ is a constant, the $V_n$s are force constants in the cosine expansion including $N$ terms and the $\gamma_n$s are corresponding phases of the dihedral angle $\theta$.

Simulated annealing fittings were employed to minimize $\Phi$ as a function of $\theta = \chi_2$ with $N = 3$ by varying $V_n$ and $\gamma_n$ values in the torsional force-field term. In line with the approach used to modify the AMBER99SB backbone potential, we have assumed that $V_n \geq 0$ kJ mol$^{-1}$ and $\gamma_n$ is either $0^\circ$ or $180^\circ$. However, on fitting the gas phase data the predicted values of $V_1$, $V_2$, and $V_3$ were $0$ kJ mol$^{-1}$ for both the MP2/6-31+G(d) and M06-2X/def2-TZVP data. We therefore consider only the IEPPCM(H$_2$O) data below and any further reference to MP2/6-31+G(d) and M06-2X/def2-TZVP calculations assumes the use of the IEPPCM(H$_2$O) method.

The values of the merit function $\Phi$ for the original AMBER99SB force field compared to the MP2/6-31+G(d) and M06-2X/def2-TZVP profiles were 2.84 and 2.24 kcal mol$^{-1}$ (after $k_0$-corrections according to Eq. (7)). On using simulated annealing fittings, these reduce to 2.39 kcal mol$^{-1}$ for the parameter set (6) obtained from fittings to the MP2/6-31+G(d) profile and 1.93 kcal mol$^{-1}$ for the parameter set (7) obtained from fittings to the M06-2X/def2-TZVP profile (Table I and Fig. 2). In both cases, $V_1 = V_2 = 0$ and $V_3 \neq 0$ kJ mol$^{-1}$. Such a result with only $V_3 \neq 0$ kJ mol$^{-1}$ is not surprising, considering that $\chi_2$ in the pyrrolidine ring varies between $\sim -40^\circ$ and $+40^\circ$. Only the $V_3$ term (with $\gamma_3 = 0^\circ$) will have a maximum equal to $V_3$ kJ mol$^{-1}$ at $\chi_2 = 0^\circ$, while the $V_1$ and $V_2$ terms (with $\gamma_1 = \gamma_2 = 180^\circ$) will show minima equal to 0 kJ mol$^{-1}$ at $\chi_2 = 0^\circ$. The value of $V_3$ increases significantly compared to the original force field, which is in qualitative agreement with earlier results from simplexed MD simulations (parameter sets (1)-(5)) indicating to better agreement with experiment on increasing $V_3$. From 200 ns MD simulations of NAcPro (Table I), the parameter set (7) from M06-2X calculations shows significantly better agreement with experiment than (6) derived from MP2 calculations.

Using the QM-derived parameter set (7) as a starting point, simplexed MD simulations were carried out to optimize the value of $V_3$. Initially, 40 ns long MD simulations were used at each step of simplexed MD simulations for merit function calculations. Parameter sets (8)-(23) were selected from these fittings with lowest merit function values for further 200 ns long MD simulations (Table I). On increasing the length of MD runs from 40 ns to 200 ns, the rmsd$^2$ values increase from 0.65–0.85 Hz to 0.73–1.08 Hz for parameter sets (8)-(23). The MD-averaged $\chi_m$ values predicted by these parameter sets (Table I) show better agreement with the experimental NMR value compared to the original AMBER99SB force field. However, it is likely that at relatively high values of $V_3$ the short MD simulations used in simplexed MD fittings were not converged sufficiently. We therefore retain all new parameter sets (6)-(23) in our further analysis, as these provide sufficiently fine distribution of $V_3$ values between 1.9 and 9.3 kJ mol$^{-1}$. In addition, parameter sets (1)-(5) were also included in our further analysis.

**Single trajectory reweighting approach**

The first application of the method relying on the energy-based reweighting approach$^{71–74}$ to fittings of $^3J$-couplings of NAcPro with optimizations of three parameters $V_1$, $V_2$, and $V_3$ led to unusually large values of $V_2$ and $V_3$ on using a 500-ns long MD trajectory with frames recorded every 0.04 ps: $V_1 = 0.0419$, $V_2 = 22.3835$, and $V_3 = 22.5864$ kJ mol$^{-1}$ with the rms of the fitting 0.79 Hz. As the predicted value of $V_1$ is very high, significantly smaller number of the pyrrolidine ring transitions are expected in MD simulations compared to, for example, the number of peptide backbone transitions, which does not agree with experiment.$^{68}$ As discussed by Li and Brückschweiler,$^{73}$ the effectiveness of the reweighting scheme critically depends on the degree of overlap between the parent and the reweighted trajectories, since the reweighted procedure does not create any new conformations. On introducing a collectivity parameter $\kappa$ with the requirement $\kappa > 50\%$ (see Eq. (2) and the discussion following it in Ref. 73), a physically plausible solution was obtained from the 500 ns long parent trajectory of NAcPro in water: $V_1 = 0$, $V_2 = 0.0009$, and $V_3 = 2.3891$ kJ mol$^{-1}$. This set of parameters is essentially the same as (14) (Table I) and therefore is not included into our further analysis. On increasing the number of terms from three to six in Eq. (5), an alternative set of parameters was derived using the reweighting approach, which is included in Table I as (24). Based on 200 ns MD simulations of NAcPro, this set of parameters performs slightly better than AMBER99SB and is therefore included into our further analysis.

**MD simulations of Gly-Pro-Gly-Gly**

For further examination, we carried out MD simulations of GPGG (Fig. 3) using force field parameters (1)-(24) and the original AMBER99SB force field. Note that AMBER99SB* or AMBER99SB-ILDN simulations would be the same in this case as the AMBER99SB simulation, as there are only Gly and Pro residues in GPGG. The recent study verifying different force fields using GPGG used 2 ms long MD simulations, which were sufficient for the majority of the force fields considered.$^{31}$
However, on examination of the convergence of the population of the folded form against the length of the MD run for the AMBER99SB force field (Fig. 6 in Ref. 31), it is clear that no significant change occurs in the population of the folded form after 600 ns. Thus, we carried out 800 ns long MD simulations for our analysis.

From the results obtained for the Pro ring in GPGG (Table II), all new parameter sets show better agreement with the experimental data with rms$_{dp}$ in the range between 0.48 and 0.62 Hz compared to the original force field (0.66 Hz). More importantly, all the tested sets provide higher values of $\chi_v$ (36.4°–41.3°) compared to the original set of parameters (35.3°). These results confirm that new parameter sets predict pyrrolidine ring geometries in better agreement with NMR, XRD and QM data compared to the original force field.

We have also analyzed NMR parameters dependent on the backbone conformation of GPGG. In particular, based on the analysis of NOE data for GPGG internuclear distances for seven proton pairs were measured previously. Averaged values of internuclear distances from MD runs were estimated over 800 ns time length for each of three MD simulations. The rms deviation between experiment and the MD predictions of distances (rms$_{dN}$) were calculated (Table SI, Supporting Information). In addition, four $^3J_{CH}$ and two $^3J_{HH}$ were available from NMR measurements for the GPGG backbone, which were used for NMR versus MD comparisons. As described previously, two empirical (corresponding to rms$_{dCH}$ and rms$_{dHH}$ in Table SI) and two QM-derived equations (corresponding to rms$_{dCH}$ and rms$_{dHH}$ in Table SI) were used to exclude possible model dependent deficiencies. For rms$_{dCH}$ and rms$_{dHH}$, we used B972 and B3LYP-predicted Karplus relationships, which have been shown to be sufficiently accurate. The results summarized in Table SI confirm that modifications of the Pro $\chi_v$ dihedral parameters do not cause any significant changes in the backbone conformations as there is a very good agreement for all the MD simulations when considering parameters averaged over backbone conformations. Similarly, the population of the U-shaped folded conformation of GPGG ($P_v, \%$) and the mean terminal N...C' distance ($d_{NC}$ Å) predicted by new parameter sets are in agreement with those predicted by the original AMBER99SB force field (Table III).

**Matching relative motional timescales from MD simulations and experiment**

To identify which of the new parameter sets is likely to reproduce both structural and dynamics properties of Pro residues more accurately, we have considered timescales of motions in GPGG. First, we consider the number of the $\psi_2$, $\phi_3$, $\psi_3$ (see definitions of angles in Fig. 3), $\chi_1$ and $\chi_2$ (see definitions of angles in Fig. 1) torsional transitions per nanosecond ($N^{\psi_2}$, $N^{\phi_3}$, $N^{\psi_3}$, $N^{\chi_1}$, and $N^{\chi_2}$ in Table III). As expected, the backbone transition numbers ($N^{\psi_2}$, $N^{\phi_3}$ and $N^{\psi_3}$) are not affected by the change of the Pro torsional parameters, whereas moderate ($N^{\psi_2} \approx 41$–66) and significant ($N^{\psi_2} \approx 1$–37) decrease in the $N^{\psi_2}$ values are observed for parameter sets (1)-(5) and (6)-(24), respectively, compared to the original force field ($N^{\psi_2} \approx 81$). For new force field parameter sets containing only a single $V_3$ term there is a linear relationship $ln (N^{\psi_2})$ versus $V_3$ (Fig. 4), as well as $ln (N^{\psi_2})$ versus $\chi_m$ (Fig. S3, Supporting Information) and $\chi_m$ versus $V_3$ (Fig. S4, Supporting Information). Thus, we can adjust the Pro sidechain torsional force field such that the timescale of the sidechain dynamics matches that from experiment.

Using $^{13}$C spin-lattice relaxation times measured for GPGG in water at 303 K, Mikhailov et al. have estimated that the auto-correlation time of the C=H bonds...
of Pro-2 is 27 ± 1.5 ps. As the accuracy of this type of measurements is critically dependent on the signal-to-noise ratio, we have repeated 13C spin-lattice relaxation time measurements of GPGG using a higher-field NMR spectrometer (14.1 T, 600 MHz 1H frequency) and a cryoprobe (Tables SII and SIII, Supporting Information). For the analysis of the T1 values and deriving correlation times, we have used the approach developed by Ernst et al.,82,83 which is different to that used by Mikhailov et al.84 The following equations were used to derive the correlation times for the overall (τe) and the intramolecular ring interconversion (τc) processes from the measured T1 relaxation times82–84:

\[ T_1^{-1} = \left[ 1 - 3x_\text{endo}x_\text{exo}\sin^2\Delta\theta \right] T_1^{-1}(\tau_c) + 3x_\text{endo}x_\text{exo}\sin^2\Delta\theta T_1^{-1}(\tau_{tot}) \]  

(8)

where \( \Delta\theta \) is the jump angle of the C–H bond on conformational transition, \( \gamma_H \) and \( \gamma_C \) are gyro magnetic ratios of 1H and 13C, h is Planck’s constant divided by 2π, \( r_{CH} = 1.09 \text{ Å} \) is the C–H bond length, \( \Delta\sigma \) is the chemical anisotropy of the 13C nucleus considered (see Experimental), N is the number of H atoms attached to the C atom. Note that in Eq. (8), the sum of populations \( x_\text{endo} + x_\text{exo} = 1 \) (not in %).

The correlation time \( \tau_c \) can be determined using the NT1 value (where N is the number of H atoms bonded to C) of the backbone Cα carbons, which are least affected by the intramolecular motions, hence better describe the overall motion of the molecule.82–86 In GPGG, NT1 values of Cα carbons are 1.146 s (Gly-1), 0.995 s (Pro-2), 1.106 s (Gly-3), and 1.836 s (Gly-4) (Table SIII, Supporting Information). The end residue backbone Cα carbons of Gly-1 and Gly-4 show the largest values, which suggest additional intramolecular dynamics for this carbon compared to mid-chain Cα carbons of Pro-2 and Gly-3. The minimum value of NT1 is observed for the Cα carbon of Pro-2, therefore we have used T1 of this backbone carbon to determine the correlation time \( \tau_c \) for the overall motion. The likely intramolecular motion that can influence the T1 value for this carbon is the pyrrolidine ring interconversion. However, as estimated previously the jump angle \( \Delta\theta \) is <5° for the Cα carbon of the pyrrolidine ring (see Table IX in Ref. 82). Using Eqs. (8)–(11), it can be estimated that \( \Delta\theta = 5° \) leads to only ∼0.4% increase in the T1 value and therefore can be neglected. From the T1 value of 995 ± 6 ms for the Cα carbon of Pro-2 in GPGG measured at 298 K for the 57 mM solution in D2O, the correlation time \( \tau_c \) is 48.2 ± 0.3 ps. This value was used in the analysis of the T1 value for the Cγ carbon of Pro-2 in GPGG to determine the correlation time \( \tau_e \) for the intramolecular ring interconversion (see below).

In Eq. (8), two terms are weighted by factors dependent on the populations of Cγ-endocarbonyl and Cγ-exo carbons conformers (\( x_\text{endo} \) and \( x_\text{exo} \) with \( x_\text{endo} + x_\text{exo} = 1 \)) and on the jump angle \( \Delta\theta \) for a given C–H bond direction on changing the ring conformation. The largest jump angles are expected for Cγ carbon of the pyrrolidine ring. Thus, the T1 relaxation times of Cγ carbons (Tables SII and SIII, Supporting Information) were used for \( \tau_e \) determinations. Madi et al. determined \( \Delta\theta \) values using dihedral angles, which they estimated using the Karplus relationship.82 Because the accuracy of the Karplus relationship for predicting dihedral angles is relatively poor, we have taken a different approach, in which QM predicted geometries are used. Such an approach is supported by the finding that in the absence of relatively strong intramolecular interactions QM geometries reproduce accurately experimental molecular geometries derived from X-ray and neutron diffraction measurements.70 We used the two lowest energy conformations of NAcPro from M06-2X/def2-TZVP IEPFCM(H2O) calculations described above, the geometries of which were optimized without any restrictions. Additional frequency calculations were carried out to verify that the final structures correspond to true minima. The obtained structures correspond to Cγ-endocarbonyl and Cγ-exo conformations of the pyrrolidine ring with \( \psi_{\chi_m} \) values of 171.5°/39.3° and 16.5°/39.0°, respectively. As discussed previously,82,83 the most rigid part of the pro ring in peptides is the C-N-Cα-C fragment, where Cs are carbonyl carbons of COMe and COO in the case of NAcPro (see Fig. 5). We therefore overlaid the Cγ-endocarbonyl and Cγ-exo conformations such that the rms deviations in the positions of four atoms of the C-N-Cα-C fragment are minimal (Fig. S5, Supporting Information).65 The angle \( \Delta\theta \) was then estimated as the angle between the corresponding Cγ-H bond directions in two conformations. The values of \( \Delta\theta \) determined for the Cγ-H12 and Cγ-H13 bonds were 82.65° and 82.47° with the average value of 82.56°, which was used as a fixed value of \( \Delta\theta \) in our fittings using T1 relaxation times of Cγ carbons. The populations of Cγ-endocarbonyl and Cγ-exo ring conformers are known from the analysis of 3JHH coupling constants measured at 298 K (Table II)35 and are assumed to be temperature independent. With these restrictions in place, the correlation time \( \tau_e \) for the intramolecular ring interconversion process were determined using the measured T1 values for Cγ carbons at different temperatures. From the comparison of the above Eq. (8) and Eq. (37) of Lipari and Szabo,59 the generalized order parameter is dependent on the populations of conformers and the jump angle \( \Delta\theta \) in the case of the two-site
jump model and can be calculated using the following relationship:

\[ S^2 = 1 - 3x_{\text{endo}}(1 - x_{\text{endo}})\sin^2\Delta \theta \]  
(12)

For \( x_{\text{endo}} = 0.543 \) and \( \Delta \theta = 82.56^\circ \), the calculated experimental value of \( S^2 \) is 0.27.

Using the measured \( T_1 \) values for \( C^\alpha \) and \( C^\gamma \) carbons of Pro in GPGG for the 214 mM solution of \( D_2O \) (Table SII, Supporting Information) the values of correlation times \( \tau_e \) and \( \tau_c \) were determined at different temperatures (Table SIV, Supporting Information). Assuming Arrhenius dependence of correlation times \( [\tau = \tau_e^0 \exp(E_a/RT)] \), activation parameters are \( E_a = 16.4 \pm 1.2 \text{ kJ mol}^{-1} \) and \( \tau_e^0 = (4.1 \pm 1.6) \times 10^{-14} \text{ s} \) for the pyrroliidine ring interconversions. To estimate errors in activation parameters, we have excluded two highest and two lowest temperatures from consideration which led to \( E_a \) variations between 15.6 and 17.6 kJ mol\(^{-1}\) and \( \tau_e^0 \) variations between 2.5 \times 10^{-14} \text{ s} and 5.5 \times 10^{-14} \text{ s}. The estimated correlation time \( \tau_e \) for the \( C^\gamma \)-H bond movements in Pro-2 of GPGG as a result of the pyrroliidine ring interconversion is 27.2 ps at 303 K, which is in good agreement with the value of 27 ps reported by Mikhailov et al.\(^{88} \)

Our MD simulations were carried out at 298 K. Using the \( T_1 \) value of 898 \pm 4 ms for the \( C^\gamma \) carbon of Pro-2 in GPGG measured at 298 K for the diluted 57 mM solution of GPGG in \( D_2O \), we have estimated the correlation time \( \tau_e \) for the \( C^\gamma \)-H bond reorientations in Pro-2 of GPGG as a result of the pyrroliidine ring interconversion as 29.7 \pm 0.4 \text{ ps}, which is slightly smaller than the value of \( \tau_e \) calculated as 30.3 ps using the activation parameters reported above for the 214 mM solution. As higher concentrations may in principle lead to partial self-associations of peptides,\(^{87} \) we have used the experimental value of \( \tau_e = 29.7 \text{ ps} \) at 298 K as a reference point for our MD simulations. From the analysis of \( \tau_e \) calculated for 14 parameter sets with a single non-zero \( V_3 \) term (\( \gamma_3 = 0^\circ \), Table IV), there is a linear correlation (Fig. S6, Supporting Information): \( V_3 \) (in kJ mol\(^{-1}\)) = 1.9272 \( \ln \tau_e \) (in ps) – 2.1881 (with \( r^2 = 0.9975 \)). Using this relationship, we estimate \( \tau_e = 29.7 \text{ ps} \). For backward verification, the 800-ns long MD simulation at 298 K with \( V_3 = 4.3474 \text{ kJ mol}^{-1} \) (\( \gamma_3 = 0^\circ \)) predict \( \tau_e = 28.7 \text{ ps} \) and \( S^2 = 0.29 \), in close agreement with the experimentally measured values of \( \tau_e = 29.7 \text{ ps} \) and \( S^2=0.27 \). This parameter set (denoted as (25) in Tables (I–V)) is selected as the final solution which reproduces the experimental structural (Tables II and III) and dynamic properties (Tables III and IV, Fig. S7) of the sidechain of the Pro residue significantly better than the original AMBER99SB force field.

**Force field phase variations**

In another set of optimizations we considered variations of both the \( V_5 \) force constant and the phase \( \gamma_3 \). The value of \( V_5 \) was varied between 1 and 5 kJ mol\(^{-1}\) with a step of 1 kJ mol\(^{-1}\), while the value of \( \gamma_3 \) was varied between \(-50 \) and \( 50^\circ \) with a step of \( 10^\circ \). The results of 700 ns long MD simulations for each pair of \( V_5 \) and \( \gamma_3 \) values are summarized in Tables SV–SVIII in Supporting Information. Over four parameters considered (rmsF, \( x_{\text{endo}}, \tau_e \) and \( S^2 \)), the force field with \( V_5 = 4.0 \text{ kJ mol}^{-1} \) and \( \gamma_3 = 0^\circ \) shows the best agreement with experiment. This additional grid search analysis allowed us to confirm that the above optimization leading to \( V_3 = 4.3474 \text{ kJ mol}^{-1} \) and \( \gamma_3 = 0^\circ \) is the unique solution in the two-dimensional \( (V_3, \gamma_3) \)-parameter space.

**Influence on the backbone conformation**

To examine the influence of the new sidechain parameter set on the protein backbone conformations and dynamics, we have carried out 1-\( \mu \)s long MD simulations of ubiquitin. Three Pro residues of ubiquitin—Pro-19, Pro-37, and Pro-38—were considered, conformational
Table V
Conformational Populations and Geometries of the Pro ring in Aqueous Solutions of Peptides from NMR and MD Simulations Using Different Sets of Torsional Parameters for the Pro residue

| Peptide            | Force field | $P_{\text{exo}}$ (%) | $P_{\text{endo}}$ (%) | $\chi_m$ (%) | $\chi_{\text{endo}}$ (%) | $\text{rms}_{\text{HH}}$ (Hz) | $S^g$ | $\tau_p$ (ps) |
|--------------------|-------------|----------------------|-----------------------|--------------|------------------------|-------------------------------|-------|-------------|
| Val-Ala-Pro-Gly    | AMBER99SB   | 14                   | 178                   | 36.3         | 62.6                   | 0.867                         | 0.25  | 4.2         |
| NMR$^{25}$         | 13          | 180                   | 38.8                   | 62.6         | 0.801                   | 0.31                           | 26.6  |             |
| NMR$^{35}$         | 14(4)       | 187(2)                | 41.0(4)                | 52.3(2)      | 0.47$^b$                | 0.26(1)$^b$                    | 30.7(5)|             |
| cis-VAPG           | AMBER99SB   | 23                   | 174                   | 36.8         | 71.1                   | 1.318                         | 0.041 | 3.3         |
| NMR$^{25}$         | 19          | 176                   | 39.4                   | 73.8         | 1.004                   | 0.41                           | 20.9  |             |
| Gly-Pro-Phe        | AMBER99SB   | 15                   | 179                   | 35.9         | 61.2                   | 0.864                         |       |             |
| (GPF)              | 15          | 179                   | 35.9                   | 61.6         | 0.802                   |                               |       |             |
| NMR$^{25}$         | 22(6)       | 183(2)                | 38.9(8)                | 68(1)        | 0.31$^b$                | 0.58(3)$^b$                    | 22(2) |             |
| Angiotensin II     | AMBER99SB   | 15                   | 178                   | 35.5         | 68.2                   | 1.320                         | 0.26  | 8.4         |
| NMR$^{25}$         | 12          | 180                   | 38.8                   | 65.0         | 1.033                   | 0.23                           | 33.1  |             |
| NMR$^{25}$         | 14(8)       | 198(6)                | 42(2)                  | 53(1)        | 0.38$^b$                | 0.26(1)$^c$                    | 32(4)$^c$|             |

$^a$1.5 μs MD simulations for angiotensin II and 800 ns for other peptides were analyzed.
$^b$The rms deviation for NMR is for fittings of experimental $J_{HH}$ values using Eqs. (8C) and (8D) of Haasnoot et al.$^{46}$ on the assumption of a two-site conformational exchange between C\textsuperscript{\text{-}endo} and C\textsuperscript{\text{-}exo} conformers and $\chi_{\text{endo}} = \chi_{\text{exo}}$.
$^c$The values and uncertainties were determined using $T_{1,\text{e}} = 386 \pm 12$ ms for $^1$C\textsuperscript{\text{u}} of Pro-7. From M06-2X/def2-TZVP calculations of GPF, the jump angle $\Delta \theta$ was 83.16°.

characteristics of which are compared in Table SIX (Supporting Information). Compared to the original force field, the parameter set (25) lead to higher $\chi_m$ values (38.8–39.5°), which are in better agreement with experimental XRD data.$^{48,88}$ In particular, the solid-state values of $\chi_m$ are 42.5° (Pro-19), 44.2° (Pro-37), and 45.2° (Pro-38).$^{88}$

Unlike Pro-19 and Pro-37, the pyrrolidine ring of Pro-38 in ubiquitin is in predominantly C\textsuperscript{\text{-}exo} conformation according to MD simulations (Table SIX), which is in agreement with the finding in that in Xaa-Yaa-Gly triplets of collagen the Pro ring prefers the endo pucker (i.e., C\textsuperscript{\text{-}endo} conformation) in the X position, while in the Y position it prefers the exo pucker.$^{89,90}$ In principle, this can be verified experimentally by measuring accurate values of $\gamma_{HH}$-couplings of the pyrrolidine rings in ubiquitin. However, pyrrolidine cyclic protons usually show strongly-coupled $^1$H NMR spectra due to small chemical shift differences for methylene protons in $\beta$ and $\gamma$ positions.$^{34}$ Accurate measurements of $\gamma_{HH}$-couplings would therefore require a full lineshape analysis, which is complicated by strongly overlapping spectra in the case of proteins.

The values of $N_{\text{av}}^2$ in ubiquitin prolines are in good agreement with those predicted for the Pro residue in GPGG, although the number of $\gamma_2$ transitions decreases significantly in Pro-38, which is likely caused by the Pro-37 residue preceding Pro-38. We have compared three experimental $J(C',H\alpha)$ couplings of 1.22 Hz (Pro-19), 1.71 Hz (Pro-37), and 1.06 Hz (Pro-38) in ubiquitin$^{37,38}$ with those calculated from MD simulations of ubiquitin using Karplus parameters, derived empirically$^{30}$ and from DFT B3LYP/EPR-III calculations.$^{31}$ Compared to the AMBER99SB*-ILDN calculations, the parameter set (25) lead to only small variations in $\gamma_{HH}$ values (Table SIX, Supporting Information). This result confirms that the changes in the sidechain dynamics interchanging the C\textsuperscript{\text{u}} atom position below and the above the C\textsuperscript{\text{u}}-N-C\textsuperscript{\text{u}} plane cause only small changes in the torsional angle H\textsuperscript{\text{alpha}}-C\textsuperscript{\text{u}}-N-C\textsuperscript{\text{u}} (Fig. 5 and Fig. S5).

Finally, the performance of parameter sets AMBER99SB*-ILDN and (25) were compared using experimental values of five different types of backbone $\gamma_{HH}$-couplings, each of which has been determined for 60–67 amino acid residues in ubiquitin.$^{37,38}$ On calculating the MD-predicted averaged $\gamma_{HH}$-couplings we have considered up to four different sets of Karplus parameters for each type of $\gamma_{HH}$ coupling.$^{49,50}$ From the results summarized in Table SX (Supporting Information), both force fields reproduce $\gamma_{HH}$ couplings equally well, confirming that the new Pro torsion potential does not cause undesirable side effects on the backbone conformations compared to the original force field, the performance of which has been verified extensively.$^{3,4,14,15,19–33}$

**Force field validation**

As an independent test, we have used NMR data and MD simulations of Val-Ala-Pro-Gly (VAPG). In Table V, we compare conformational populations and geometries of the Pro ring in VAPG in water as predicted by NMR and by 800-ns long MD simulations. The rms$_{\text{p}}$ values relative to experimental values of $\gamma_{HH}$-couplings show that the new force field (25) reproduces better the experimentally measured values than the original force field. The value of $\chi_m$ serves as a measure of non-planarity of the five-membered ring. The results confirm that the new force field (25) leads to significantly improved agreement with experiment compared to the original force field AMBER99SB.

In terms of motional dynamics, the predicted values of the correlation time and generalized order parameter for
the Pro ring interconversion at 298 K are 4.2 ps and 0.35, respectively, according to the 800-ns MD simulations at 298 K using the original AMBER99SB force field. The predicted value of $\tau_e$ is significantly different from the value measured experimentally in this work using $T_1$ values of 30.7 $\pm$ 0.5 ps for the 77 mM solution of VAPG in H2O:D2O (9:1). For $\chi_{endo} = 0.523$ and $\Delta \theta = 82.56^\circ$, the estimated experimental value of $S^2$ is 0.26. Note that in VAPG, the NT$_1$ values of C$^a$ carbons are 0.751 s (Val-1), 0.614 s (Ala-2), 0.641 s (Pro-3) and 1.142 s (Gly-4) (Table SXI, Supporting Information). Judging by NT$_1$ values, the C$^a$ site of Ala is least affected by intramolecular motions, thus the $T_1$ value of this carbon was used to determine the correlation time for the overall molecular motion ($\tau_c = 82.8 \pm 0.7$ ps). The corresponding values predicted by the new force field are $\tau_c = 28.6$ ps and $S^2 = 0.31$, which are in good agreement with experiment.

Although we have primarily focused on force field optimizations for the trans-rotamer about the bond preceding the Pro residue, it would be interesting to verify whether the new force field would offer any improvements for the cis-rotamer compared to the original force field. In the case of cis-VAPG (with the cis-orientation of the CH$_2$ group of Gly and the CO group of Pro), the MD-predicted $J_{HH}$ couplings by the new force field (25) show improved agreement with experimental values of $J_{HH}$ couplings compared to the original force field as judged by the rms$_{dp}$ values: 1.00 Hz and 1.32 Hz for force fields (25) and AMBER99SB. However, the agreement with the experiment is not as good as for the trans-VAPG considered above due to the lower value of the predicted population of the C$^\gamma$-endo conformer by the new force field (74%, as opposed to the experimental value of 83%). The difference in the predicted population by the new force field is further amplified in the predicted value of $S^2 = 0.41$ (experimental value 0.58), as $S^2$ is proportional to the product of $\chi_{endo}$ and $(1 - \chi_{endo})$. At the same time, the predicted value of $\tau_c = 20.9$ ps by the new force field is in good agreement with the experimental value of 22 $\pm$ 2 ps. For comparison, the predicted values of $S^2$ and $\tau_e$ by AMBER99SB are 0.41 and 3.3 ps, respectively.

The change of the amino acid residue proceeding the Pro residue to Phe has been shown to lead to the increased population of the C$^\gamma$-endo conformer. We have re-determined conformational characteristics of the Pro residue in Gly-Pro-Phe (GPF) using experimental values of all ten $J_{HH}$ couplings reported by Anteunis et al. $^{91}$ and the least squares fitting procedure described previously. $^{34}$ The results summarized in Table V confirm that the content of the C$^\gamma$-endo conformer increases in GPF ($\chi_{endo} = 68.0\%$) compared to that in GPGG and VAPG. However, the degree of change is not as significant as previously predicted ($\chi_{endo} = 85\%$) using Karplus relations of Pogliani et al. $^{92}$ In Table V, we compare conformational populations and geometries of the Pro ring in GPF in water from 800 ns long MD simulations and experiment. As in the case of tetrapeptide VAPG above, the rms$_{dp}$ values relative to experimental values of ten $J_{HH}$-couplings show that the new force field (25) reproduces better the experimentally measured values than the original force field. The higher values of $\chi_{endo}$ and $\tau_m$ compared to the original force field are also in better agreement with experiment (Table V).

We have also analyzed NMR data and MD simulations of octapeptide angiotensin II (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe, Fig. S8 in Supporting Information). After initial assignments of peaks in $^1H$ and $^{13}C$ spectra of 16 mM solution of angiotensin II in D$_2$O using 2D NMR spectra (Tables SXII and SXIII, Supporting Information), full $^1H$ NMR lineshape analysis was carried out to determine vicinal $J_{HH}$ couplings of the Pro-7 side-chain (Fig. S9 and Table SXIV, Supporting Information), which were subsequently analyzed to estimate conformational characteristics of the pyrrolidine ring of Pro-7 in angiotensin II. In addition, $^{13}C$ spin-lattice relaxation times were measured at 298 K (Table SXV in Supporting Information), which allowed to measure values of $S^2$ and $\tau_e$. As in the case of GPGG and VAPG discussed above, the $T_1$ values of the backbone C$^a$ carbons show clear decrease towards the mid-chain residues (in ms next page):
Asp Arg Val Tyr Ile His Pro Phe
520 355 347 310 324 327 372 448

The minimum value of $T_1$ observed for the C\(^\alpha\) carbon of Tyr-4 suggests that this site is least affected by intramolecular motions. It is therefore best suited for determining the correlation time $\tau_c$ of the overall molecular motion. From Eqs. (8)–(11), the value of $\tau_c$ corresponding to $T_1 = 310 \pm 3$ ms is 246 ± 6 ps. This value was used in the analysis of the $T_1$ value for the C\(^\alpha\) carbon of Pro-7 in angiotensin II to determine the correlation time $\tau_e$ for the intramolecular ring interconversion (see below).

To estimate the jump angle $\Delta \theta$ in angiotensin, we have used M06-2X/def2-TZVP calculations of GPF with the Phe residue following Pro as in angiotensin II. After overlaying the C\(^\gamma\)-endo- and C\(^\gamma\)-exo-conformations of GPF such that the rms deviations in the positions of four atoms of the C-N-C\(^\alpha\)-C fragment are minimal, the jump angle $\Delta \theta$ was determined as 83.16° (82.97° for C\(^\gamma\)-H\(^2\) and 83.34° for C\(^\gamma\)-H\(^3\)), which was used as a fixed value of $\Delta \theta$ in our fittings $T_1$ relaxation data.

In Table V, we compare conformational populations and geometries of the pyrrolidine ring of angiotensin II in water determined by NMR and by 1500-ns long MD simulations. The rms$_{pp}$ values relative to experimental values of 10 $^3$J$_{HH}$-couplings show that the new force field (25) with rms$_{pp} = 1.03$ Hz reproduces the experimentally measured values better than the original force field with rms$_{pp} = 1.32$ Hz. For the pseudorotation amplitude $\chi_m$, the results confirm that the new force field (25) leads to significantly improved agreement ($\chi_m = 38.8^\circ$) with experiment ($\chi_m = 42^\circ \pm 2^\circ$) compared to AMBER99SB ($\chi_m = 35.5^\circ$). Regarding motional dynamics (Table V), the timescale of motion is reproduced significantly better by the new force field (25). The corresponding values of $\tau_e$ are 8.4, 33.1 and 32 ± 4 ps for AMBER99SB, the new force field (25) and experiment, respectively.

Finally, the relative experimental values of overall and internal correlation times $\tau_c/\tau_e$ were 48.2 ps/29.7 ps in GPGG, 82.8 ps/30.7 ps in VAPG and 246 ps/32 ps in angiotensin II. These clearly show that despite the five-fold increase in the correlation time of the overall motion, the timescales of the internal motion remains essentially unchanged in these peptides of varying size. Thus, it is likely that the overall molecular motions and the intramolecular dynamics of the Pro ring are independent in the peptides considered.

### Force field parameters of hydroxyproline

Together with Pro and Gly, the 4-hydroxy-L-proline residue (Hyp) is one of the main building blocks in collagen,\(^{89,90,93}\) although it is not included in the list of 20 natural amino acid residues. In the GROMACS implementation of AMBER99SB, the force field parameters of Mooney \textit{et al.} \cite{Mooney} is used for the N-C\(^\alpha\)-C\(^\gamma\)-O torsion of Hyp,\(^{89}\) although reparameterization by Park \textit{et al.}\(^{90}\) has been shown to reproduce the experimentally observed preference of the C\(^\gamma\)-exo conformer in Hyp over the C\(^\gamma\)-endo conformer better than that of Mooney \textit{et al.}\(^{89}\) Our MD simulations carried out for Ace-Hyp-NHMe (AHM, Fig. 6) are in agreement with these findings (Table VI). The predicted population of the C\(^\gamma\)-endo conformer is 51.4% on using parameters of Mooney \textit{et al.}, while the smaller value of 6.7% predicted by the Hyp parameters of Park \textit{et al.} is in good agreement with the experimental value of 12%. Similarly, the experimental $^3$J$_{HH}$ couplings

### Table VI

Conformational Populations and Geometries of the Hyp Ring in AHM in Water from NMR and 1.5-\textmu s Long MD Simulations Using Various Sets of Torsional Parameters for the Hyp Residue

| Force field     | $V_j$ (kJ mol$^{-1}$) | $P_{endo}$ (%) | $P_{endo}$ (%) | $\chi_m$ (°) | $\chi_{endo}$ (%) | rms$_{pp}$ (Hz) |
|-----------------|-----------------------|----------------|----------------|--------------|-------------------|----------------|
| AMBER99SB       | 0.65084               | 14             | 177            | 35.0         | 51.4              | 2.721          |
| AMBER99SB*      | $^h$                  | 14             | 167            | 34.6         | 6.7               | 1.046          |
| h1              | 1.7                   | 13             | 179            | 38.4         | 50.8              | 2.852          |
| h2              | 2.7                   | 14             | 181            | 38.9         | 48.6              | 2.528          |
| h3              | 3.7                   | 15             | 182            | 39.3         | 45.0              | 2.321          |
| h4              | 4.7                   | 15             | 183            | 39.8         | 44.6              | 2.296          |
| h5              | 5.7                   | 16             | 184            | 40.1         | 41.0              | 2.085          |
| h6              | 6.7                   | 16             | 185            | 40.6         | 38.6              | 1.946          |
| h7              | 7.7                   | 16             | 186            | 40.9         | 35.9              | 1.941          |
| h8              | 8.7                   | 17             | 186            | 41.2         | 33.4              | 1.641          |
| h9              | 9.7                   | 18             | 187            | 41.4         | 32.3              | 1.579          |
| h11             | 10.7                  | 18             | 188            | 41.6         | 36.6              | 1.828          |
| h12             | 11.7                  | 18             | 188            | 41.8         | 22.7              | 1.026          |
| NMR             | —                     | 12(1)          | 215(9)         | 42(2)        | 11.9(8)           | 0.344*         |

*Apart from the original AMBER99SB force fields using the Hyp force field parameters of Mooney \textit{et al.}\(^{89}\) and Park \textit{et al.}\(^{90}\) all other models use $V_j=4.3474$ kJ mol$^{-1}$ ($\chi_j = 0^\circ$) for the endocyclic C-N-C\(^\alpha\)-C (\(\chi_j\)) torsion of the Hyp residue of AHM.

The modified Hyp force field parameters of Park \textit{et al.} were used as a Ryckaert-Bellemans function with $C_0 = 0.6527$ kJ mol$^{-1}$ and $C_1 = 12.46832$ kJ mol$^{-1}$.

The rms deviation for NMR is for fittings of experimental $^3$J$_{HH}$ values using Eqs. (8C) and (8D) of Haasnoot \textit{et al.}\(^{96}\) assuming a two-site exchange between C\(^\gamma\)-endo and C\(^\gamma\)-exo conformers and $\chi_{endo} = \chi_{endo} + \chi_{endo}$.34
of the Hyp ring are better reproduced by parameters of Park et al. (rms$_{BP}$=1.05 Hz) compared to that of Mooney et al. (rms$_{BP}$=2.72 Hz). However, the $\chi_m$ values by both parameter sets show flattened ring geometries compared to experiment (Table VI). Furthermore, the predicted motional characteristics of the ring dynamics by both parameter sets are in sharp contrast with experiment, showing significantly higher frequencies of ring interconversions. In particular, the correlation times of the ring interconversions ($\tau_c$) are 7.8 ps (Mooney et al.), 1.5 ps (Park et al.) and 82.6 ps (experiment).

We have optimized the force field parameters for the hydroxyproline N-C$^\alpha$-C$^\gamma$-O torsional angle (denoted as $\chi_h$) to better match the dynamics characteristics of the Hyp sidechain. The new force field (25) for the C$\cdots$C$\cdots$C$\cdots$C ($\chi_2$) torsion was used as a fixed constant ($V_3 = 4.3474$ kJ mol$^{-1}$ and $\gamma_3 = 0^\circ$) in these optimizations for the Hyp residue. In the original AMBER99SB force field $V_3 = 0.65084$ kJ mol$^{-1}$ and $\gamma_3 = 0^\circ$ for the hydroxyproline N-C$^\alpha$-C$^\gamma$-O ($\chi_h$) torsion. Initially, 1.5-µs MD simulations were considered in which the value of $V_3$ for $\chi_h$ was gradually increased (Table VI). This showed that the population $x_{endo}$ approaches the experimental value at only very high values of $V_3$ (see Table VI), at which even 1.5 µs MD simulations may not be sufficient for the convergence of the predicted population.

Similar to the Pro residue considered above, we used QM calculations to fit the $\chi_h$ parameters in Hyp. The M06-2X/def2-TZVP IEFPCM(water) calculations of 26 conformers of AHM were carried out in which the N-C$^\alpha$-C$^\gamma$-O dihedral angle was varied in $5^\circ$ steps between 52.8$^\circ$ and 177.8$^\circ$. Simulated annealing fittings were employed to minimize the value of merit function $\Phi$ [Eq. (6)] as a function of $\theta \equiv \chi_h$ by varying $V_3$ values ($\gamma_3 = 0^\circ$) and $K_0$ [Eq. (7)]. This led to $V_3 = 5.5574$ kJ mol$^{-1}$ with only small improvement in the value of $\Phi$ (0.44 kcal mol$^{-1}$) compared to the original force field with the Hyp parameters of Mooney et al. (0.46 kcal mol$^{-1}$). The QM-optimized value is close to the value of $V_3 = 5.7$ kJ mol$^{-1}$ in Table VI, which predicts very high value of $x_{endo}$ compared to experiment. Therefore, no new MD simulations were carried out.

In a new set of optimizations we considered variations of both the $V_3$ force constant and its phase $\gamma_3$. The results of 600-ns long MD simulations for each pair of $V_3$ and $\gamma_3$ values are summarized in Tables SXVI–SXIX (Supporting Information). Over four parameters considered (rms$_{BP}$, $x_{endo}$, $\tau_c$ and $S^\theta$), the force field with $V_3 = 5.3$ kJ mol$^{-1}$ and $\gamma_3 = 30^\circ$ shows the best agreement with experiment. From the spin-lattice relaxation time measurements for a 59 mM solution of AHM in D$_2$O at 298 K, $\tau_c = 32.8 \pm 0.5$ ps, $\tau_c = 82.6 \pm 2.8$ ps and $S^\theta = 0.69 \pm 0.01$ (full NMR data for AHM is included in Tables SX–SXXII in Supporting Information). The $\tau_c$ values for the force constants $V_3 = 4.3, 5.3, 6.3$ and 6.7 kJ mol$^{-1}$ at $\gamma_3 = 30^\circ$ show a satisfactory linear relationship: $V_3$ (in kJ mol$^{-1}$) = $3.6404 \log \tau_c$ (in ps) – $10.555$ (with $r^2 = 0.9968$). Using this relationship, we estimate $V_3 = 5.5138$ kJ mol$^{-1}$ for the experimental value of $\tau_c = 82.6$ ps. This value of $V_3$ together with the phase $\gamma_3 = 30^\circ$ was used for our further verifications (referred to as parameter set (h13)). A 1.5-µs long MD simulation using force field (h13) for $\chi_h$ of Hyp (with force field (25) for the $\chi_h$ potential) confirmed the improvement of the parameterization of the $\chi_h$ potential, as $S^\theta$ is 0.69 and $\tau_c = 77.6$ ps compared to the original AMBER99SB force field with $S^\theta = 0.34$ and $\tau_c = 7.8$ ps and the experimental values of $S^\theta = 0.69$ and $\tau_c \approx 83$ ps (Table VII). Also, the predicted $x_{endo}$ population is 9.6%, which is in close agreement with the experimental value of 11.9%. In addition, the $\chi_m$ value increases from 35.0$^\circ$ for AMBER99SB to 39.5$^\circ$ for (h13), which compares better to the experimental estimate of $42^\circ \pm 2^\circ$. As expected,

| Peptide | Force field | $P_{exo}$ (%) | $P_{endo}$ (%) | $\chi_m$ (%) | $x_{endo}$ (%) | rms$_{BP}$ (Hz) | $S^\theta$ | $\tau_c$ (ps) |
|---------|-------------|---------------|---------------|-------------|---------------|----------------|-----------|-------------|
| AHM     | AMBER99SB$a$ | 14            | 177           | 35.0        | 51.4          | 2.721          | 0.34      | 7.8         |
|         | AMBER99SB$a,b$ | 14            | 167           | 34.6        | 6.7           | 1.046          | 0.78      | 1.5         |
| h13     | 14           | 183           | 39.5          | 9.6         | 0.624         | 0.69           | 77.6      |             |
| NMR     | 12(1)        | 215(9)        | 42(2)         | 11.9(8)     | (0.34)$c$     | 0.69(1)$d$     | 83(3)$d$  |             |
| AHG     | AMBER99SB$a$ | 14            | 176           | 35.0        | 51.2          | 2.597          | 0.34      | 8.8         |
|         | AMBER99SB$a,b$ | 15            | 162           | 34.6        | 6.6           | 1.018          | 0.78      | 1.7         |
| h13     | 14           | 183           | 39.6          | 9.4         | 0.635         | 0.70           | 79.9      |             |
| NMR     | 12(1)        | 213(8)        | 42(2)         | 13.9(6)     | (0.36)$c$     | 0.65(1)$d$     | 80(4)$d$  |             |

$a$Apart from the original AMBER99SB force fields using the Hyp force field parameters of Mooney et al. 89 and Park et al. 90 all other models use $V_3$=4.3474 kJ mol$^{-1}$ ($\gamma_3 = 0^\circ$) for the endocyclic C$\cdots$C$\cdots$C$\cdots$C ($\chi_2$) torsion of the Hyp residue.

$b$The modified Hyp force field parameters of Park et al. were used as a Ryckaert–Bellemans function with $C_0 = 0.6527$ kJ mol$^{-1}$ and $C_2 = 12.46832$ kJ mol$^{-1}$.

$c$The rms deviation for NMR is for fittings of experimental $J_{1H1H}$ values assuming a two-site exchange between C$^\gamma$-endo and C$^\gamma$-exo conformers and $\Delta \nu_{0}^{endo-exo} = 34$.

$d$The values and uncertainties were determined using $T_1$ for $^1$H of Hyp in 59 mM D$_2$O solutions. From M06-2X/aug-cc-PVTZ calculations of AHM, the jump angle $\Delta \theta$ used for determining $S^\theta$ and $\tau_c$ in AHM and AHG was 82.64$^\circ$. The $\tau_c$ values determined using $T_1$ for $^3$C of Hyp were 32.8 ± 0.5 ps for AHM and 43.5 ± 0.6 ps for AHG.
these improvements are reflected in the considerable reduction in the $r_{pB}^\text{rms}$ value, which decreases from 2.72 Hz for AMBER99SB with the Hyp parameters of Mooney et al. to 0.62 Hz for model (h13).

Further independent validation for the hydroxyproline parameters was carried out using 1.5-μs long MD simulations of N-acetyl-4-hydroxy-L-proline-glycine (Ace-Hyp-Gly, AHG, Fig. 6; full NMR data is included in Tables SXX–SXXII, Supporting Information). The new force field (h13) for the $\gamma_2$ endocyclic torsion shows a much improved agreement with experiment compared to the original force field AMBER99SB (Table VII). The value of $\gamma_m$ increases from 35° and 34.6° for the AMBER99SB force field with the Hyp parameters of Mooney et al. and Park et al. respectively, to 39.6°. For comparison, $\gamma_m = 42° \pm 2°$ based on the analysis of the experimental NMR data. The predicted value of $\chi_{endo}$ also shows improved agreement with experiment, that is, the experimental value of 13.8% ± 0.5% is reproduced as 9.4% by the new force field. This is also reflected in the reduced $r_{pB}^\text{rms}$ value which is 0.64 Hz (Table VII). By far the largest improvement is obtained for dynamics characteristics of the hydroxyproline ring interconversion. For example, the original force field using the Hyp parameters of Mooney et al. predicts $\tau_e = 8.7$ ps and $S^2 = 0.34$, while the experimental values are $\tau_e = 80 \pm 4$ ps and $S^2 = 0.65 \pm 0.01$. The new force field predicts $\tau_e = 79.9$ ps and $S^2 = 0.70$, in quantitative agreement with the experimental values and significantly better than the original force field (Table VII).

**DISCUSSION**

We propose a new approach for force field optimizations which aims at reproducing experimental dynamics characteristics using biomolecular MD simulations, in addition to improved prediction of motionally averaged structural properties available from experiment. As the source of experimental data for dynamics fittings, we use $^{13}$C NMR spin-lattice relaxation times $T_1$ of various backbone and sidechain carbon atoms, which allow to selectively determine correlation times of both overall molecular reorientations and intramolecular motions. For relative conformational stability and structural fittings, we use motionally averaged experimental values of NMR $J$ couplings over three bonds. The proline residue and its derivative 4-hydroxyproline with relatively simple structure and sidechain dynamics were chosen for the assessment of the new approach in this work. Initially, the grid search and simplex MD simulations identified large number of parameter sets which fit equally well experimental $J$ couplings. Using the Arrhenius-type exponential relationship between the force constant and the correlation time, the available MD data for a series of different parameter sets were analyzed to determine the value of the force constant that best reproduces experimental timescale of the sidechain dynamics. Verification of the new force-field parameters against NMR $J$ couplings and correlation times showed consistent and significant improvements compared to the original force field in reproducing both structural and dynamics properties. These results suggest that matching experimental timescales of motions together with motionally averaged characteristics is a valid and robust approach for force field parameter optimization. Such a comprehensive approach is not restricted to cyclic proline and 4-hydroxyproline residues and can be extended to sidechain structure and dynamics of other amino acid residues, as well as to the protein backbone. In cases more complex than the Pro or Hyp sidechain dynamics, QM methods may also prove successful in providing information regarding the barrier heights of conformational changes, especially when the interpretation of the NMR relaxation data is not straightforward.

**ACKNOWLEDGMENTS**

The authors thank University College London (UCL) for the provision of computational facilities. The authors acknowledge the use of the UCL Legion High Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work. The work presented here made use of the *Emerald* High Performance Computing facility. The Center is owned and operated by the e-Infrastructure South Consortium formed from the universities of Bristol, Oxford, Southampton and UCL in partnership with STFC Rutherford Appleton Laboratory. Prof Ad Bax is thanked for the provision of the details of their $J$-coupling measurements. Dr. Viktor Hornak is thanked for the details of their force field optimizations. Helpful and stimulating suggestions by the reviewers are gratefully acknowledged.

**REFERENCES**

1. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Biol 2002;9:846.
2. Mackerell AD, Jr. Empirical force fields for biological macromolecules: overview and issues. J Comput Chem 2004;25:1584.
3. Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE. Long-timescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol 2009;19:120.
4. Best RB. Atomic molecular simulations of protein folding. Curr Opin Struct Biol 2012;22:52.
5. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Jr., 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.
6. Sorin EJ, Pande VS. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. Biophys J 2005;88:2472.
7. Garcia AE, Sanbonmatsu KY. α-Helix stabilization by side chain shielding of backbone hydrogen bonds. Proc Natl Acad Sci USA 2002, 99, 2782.
8. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586.

9. Daura X, Mark AE, van Gunsteren WF. Parametrization of aliphatic CHn united atoms of GROMOS96 force field. J Comput Chem 1998;19:535.

10. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 2004;25:1656.

11. Jorgensen WL, Maxwell DS, 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;118:11225.

12. 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.

13. 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.

14. Hornak V, Abel R, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins Struct Funct Bioinf 2006, 65, 712.

15. Best RB, Hummer G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J Phys Chem B 2009, 113, 9004.

16. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins Struct Funct Bioinf 2010; 78:1950.

17. MacKerell AD, Jr., Feig M, Brooks CL, 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 Comp Chem 2004;25:1400.

18. Klauda JB, Venable RM, Freitas JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, Vorobyov I, MacKerell AD, Jr, Pastor AW. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J Phys Chem B 2010;114:7830.

19. Price DJ, Brooks CL, III. Modern protein force fields behave comparably in molecular dynamics simulations. J Comput Chem 2002;23:1045.

20. 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;50:451.

21. Mu Y, Kosov DS, 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;107:5064.

22. Hess B, van der Vegt NFA. Hydration thermodynamic properties of amino acid analogues: A systematic comparison of biomolecular force fields and water models. J Phys Chem B 2006;110:17616.

23. Kaminski J, Jensen F. Force field modeling of amino acid conformational energies. J Chem Theory Comput 2007;3:1774.

24. Rueda M, Ferrer-Costa C, Meyer T, Perez A, Camps J, Hospital A, Gelpí JL, Orozco M. A consensus view of protein dynamics. Proc Natl Acad Sci USA 2007;104:796.

25. Showalter SA, Frischweiler R. Validation of molecular dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: application to the AMBER99SB force field. J Chem Theory Comput 2007;3:961.

26. Best RB, Buchete N-V, Hummer G. Are current molecular dynamics force fields too helical? Biophys J 2008;95:L707.

27. Li Z, Yu H, Zhuang W, Mukamel S. Geometry and excitation energy fluctuations of NMA in aqueous solution with CHARMM, AMBER, OPLS, and GROMOS force fields: implications for protein ultraviolet spectra simulation. Chem Phys Lett 2008;452:78.

28. Wickström L, Okur A, Simmerling, C. Evaluating the performance of the FF99SB force field based on NMR scalar coupling data Biophys J. 2009;97:853.

29. Seabra GM, Walker RC, Roitberg AE. Are current semi-empirical methods better than force fields? A study from the thermodynamics perspective. J Phys Chem A 2009;113:1958.

30. Project E, Nachliel E, Gutman M. Force field-dependant structural divergence revealed during long time simulations of Calbindin d9k. J Comput Chem 2010;31:1864.

31. Aliev AE, Courtier-Murias D. Experimental verification of force fields for molecular dynamics simulations using Gly-Pro-Gly-Gly. J Phys Chem B 2010;114:12358.

32. Beauchamp KA, Lin Y-S, Das R, Pande VS. Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. J Chem Theory Comput 2012;8:1409.

33. Best RB, de Sancho D, Mittal J. Biophys J 2012;102:1462.

34. Aliev AE, Courtier-Murias D. Conformational analysis of L-prolines in water. J Phys Chem B 2007;111:14054.

35. Aliev AE, Bhandal S, Courtier-Murias D. Quantum mechanical and NMR studies of ring puckering and cis/trans-rotameric interconversion in prolines and hydroxyprolines. J Phys Chem A 2009;113:10858.

36. Aliev AE, Courtier-Murias D, Bhandal S, Zhou S. A combined NMR/MD/QM approach for structure and dynamics elucidations in the solution state: pilot studies using tetrapeptides. Chem Commun 2010;46:695.

37. Wang AC, Bax A. Reparametrization of the Karplus relation for $^3\text{H}(\text{Na})$ and $^3\text{H}(-\text{C})$ in peptides from uniformly $^{15}\text{C}$/$^2\text{H}$-enriched human ubiquitin. J Am Chem Soc 1995;117:1810.

38. Wang AC, Bax A. Determination of the backbone dihedral angles $\phi$ in human ubiquitin from reparametrized empirical Karplus Equations. J Am Chem Soc 1995;117:1810.

39. Hed B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4:435.

40. 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.

41. Essman U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103:8577.

42. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101.

43. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys 1981;52:7182.

44. Press WH, Flannery BP, Teukolsky SA. Numerical recipes in FORTRAN: The art of scientific computing, Cambridge: Cambridge University Press; 1992.

45. Aliev AE, Harris KDM. $^2\text{H}$ NMR lineshape analysis using automated fitting procedures based on local and quasi-global optimization techniques. Magn Reson Chem 1998;36:855.

46. Haasnoot CAG, DeLeeuw FAA, Altona C. The relationship between proton-proton NMR coupling constants and substituent electronegativities—I : An empirical generalization of the Karplus equation. Tetrahedron 1980;36:2783.

47. Karplus M. Vicinal proton coupling in nuclear magnetic resonance. J Am Chem Soc 1963;85:2870.

48. Vijay-Kumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 Å resolution. J Mol Biol 1987;194:531.

49. Case DA, Scheurer C, Brüschweiler R. Static and dynamic effects on vicinal scalar J couplings in proteins and peptides: A MD/DFT analysis. J Am Chem Soc 2000;122:10390.
50. Hu J-S, Bax A. Determination of $\varphi$ and $\chi_1$ angles in proteins from $^{13}$C--$^{13}$C three-bond $J$ couplings measured by three-dimensional heteronuclear NMR. How planar is the peptide bond? J Am Chem Soc 1997;119:3630.

51. Schmidt JM, Blumel M, Löh N, Rüterjans H. Self-consistent $J$ coupling constants for the calibration of Karplus coefficients and evaluation of torsion angles. J Biomol NMR 1999;14:1.

52. Tropp J. Dipolar relaxation and nuclear Overhauser effects in non-rigid molecules: The effect of fluctuating intermolecular distances. J Chem Phys 1980;72:6035.

53. Neuhaus D, Williamson MP. The nuclear Overhauser effect in structural and conformational analysis, 2nd ed. New York: Wiley-VCH; 2000.

54. Reginelli M, Hoffman H, Köck M, Mierke DF. Determination of conformation and relative configuration of a small, rapidly tumbling molecule in solution by combined application of NOESY and restrained MD calculations. J Am Chem Soc 1992;114:3272.

55. Lafitte VGH, Aliev AE, Horton PN, Hursthouse MB, Hailes HC. Highly stable cyclic dimers based on non-covalent interactions. Chem Commun 2006;20:2173.

56. Aliev AE, Mia ZA, Busson MJ, Fitzmaurice RJ, Caddick S. Diasteromeric configurations from joint experimental-computational analysis. J Org Chem 2012;77:6290.

57. Ferner I, Villa A, Duchardt E, Widjakaskusma E, Wohntz J, Stock G, Schwabe K. NMR and MD-studies of the temperature dependent dynamics of RNA YRNMG-tetraloops. Nucleic Acids Res 2008:36:1928.

58. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc 1982;104:4546.

59. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RJ, Fox DJ. Gaussian 09, Revision C.01. Wallingford, CT: Gaussian Inc.; 2009.

60. Zhao Y, Truhlar DG. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc 2008;120:215.

61. Zhao Y, Truhlar DG. Exploring the limit of accuracy of the global hybrid meta density functional for main-group thermochemistry, kinetics, and noncovalent interactions. J Chem Theory Comput 2008;4:1849.

62. Cancès E, Mennucci B. New applications of integral equations methods for solvation continuum models: ionic solutions and liquid crystals. J Math Chem 1998;23:309.

63. Cossi M, Rega N, Scalmani G, Barone V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. J Comp Chem 2003;24:669.

64. Sanner MF. Python: a programming language for software integration and development. J Mol Graph Mod 1999;17:57.

65. Haasnoot CAG, DeLeeuw FAAM, DeLeeuw HPM, Altona C. Relationship between proton–proton NMR coupling constants and substituent electronegativities. III. Conformational analysis of proline rings in solution using a generalized Karplus equation. Biopolymers 1981;20:1211.

66. Westhof G, Sundaralingam M. A method for the analysis of puckering disorder in five-membered rings: the relative mobilities of furanose and proline rings and their effects on polynucleotide and polypeptide backbone flexibility. J Am Chem Soc 1983;105:970.

67. Mikhailov D, Daragan VA, Mayo KH. $^{13}$C-multiplet NMR relaxation-deried ring puckering and backbone dynamics in proline-containing glycine-based peptides. Biophys J 1995;68:1540.

68. Wang J, 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.

69. Aliev AE, Mia ZA, Khaneja HS, King FD. Structures in solutions from joint experimental-computational analysis: Applications to cyclic molecules and studies of noncovalent interactions. J Phys Chem A 2012;116:1093.

70. Norgaard AB, Ferkinghoff-Borg J, Lindorff-Larsen K. Experimental parameterization of an energy function for the simulation of unfolded proteins. Biophys J 2008;94:182.

71. Li DW, Brüschweiler R. NMR-Based protein potentials. Angew Chem 2010;122:6930.

72. Li DW, Brüschweiler R. Iterative optimization of molecular mechanics force fields from NMR data of full-length proteins. J Chem Theory Comput 2011;7:1773.

73. Riana S, Lindorff-Larsen K, Shaw DE. How robust are protein folding simulations with respect to force field parameterization? Biophys J 2012;100:147.

74. Krivdin LR, Contreras RH. Recent advances in theoretical calculations of indirect spin–spin coupling constants. Annu Rep NMR Spectros 2007;61:133.

75. Bagno A, Saielli G. Computational NMR spectroscopy: reversing the information flow. Theor Chem Acc 2007;117:603.

76. Helgaker T, Jaszuzski M, Pecul M. The quantum-chemical calculation of NMR indirect spin–spin coupling constants. Prog Nucl Magn Reson Spectrosc 2008;53:249.

77. Yongye AB, Lachele Foley B, Woods RI. On achieving experimental accuracy from molecular dynamics simulations of flexible molecules: aqueous glycerol. J Phys Chem A 2008;112:2634.

78. Di Micco S, Chini MG, Riccio R, Bifulco G. Quantum mechanical calculation of NMR parameters in the stereostructural determination of natural products. Eur J Org Chem 2010;1411.

79. Taha HA, Castillo NC, Sears DN, Wasylifshen RE, Lowary TL, Roy P-N. Conformational analysis of arabinofuranosides: Prediction of $^{3}J_{H,H}$ using MD simulations with DFT-derived spin–spin coupling profiles. J Chem Theory Comput 2010;6:212.

80. Bally T, Rablen PR. Quantum-chemical simulation of 1H NMR spectra. 2. Comparison of DFT-based procedures for computing proton-proton coupling constants in organic molecules. J Org Chem 2011;76:4818.

81. Mádi ZL, Griesinger C, Ernst RR. Conformational dynamics of proline residues in antamanide. J coupling analysis of strongly coupled spin systems based on ECOYS spectra. J Am Chem Soc 1990;112:2908.

82. Schmidt JM, Brüschweiler R, Ernst RR, Dunbrack RL, Joseph D, Karplus M. Molecular dynamics simulation of the proline conformational equilibrium and dynamics in antamanide using the CHARMM force field. J Am Chem Soc 1993;115:8747.

83. Showalter SA, Brüschweiler R. Validation of molecular dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: application to the AMBER99SB force field. J Chem Theory Comput 2007;3:961.

84. London RE. The interpretation of carbon-13 spin-lattice relaxation resulting from ring puckering in proline. J Am Chem Soc 1978;100:2678.

85. Shekar SC, Easwaran KR. Proline ring conformations corresponding to a bistable jump model from $^{13}$C spin-lattice relaxation times. Biopolymers 1982;21:1479.
87. Atkinson CE, Aliiev AE, Motherwell WB. Binding site optimisation for artificial enzymes by diffusion NMR of small molecules. Chem Eur J 2003;9:1714.

88. Huang KY, Amodeo GA, Tong L, McDermott A. The structure of human ubiquitin in 2-methyl-2,4-pentanediol: a new conformational switch. Protein Sci 2011;20:630.

89. Mooney SD, Kollman PA, Klein TE. Conformational preferences of substituted prolines in the collagen triple helix. Biopolymers 2002;64:63.

90. Park S, Radmer RJ, Klein TE, Pande VS. A new set of molecular mechanics parameters for hydroxyproline and its use in molecular dynamics simulations of collagen-like peptides. J Comput Chem 2005;26:1612.

91. Anteunis MJO, Borremans FAM, Stewart JM, London RE. 360-MHz hydrogen-1 NMR conformational analysis of Gly-Pro-X peptides (X = Ala, Cha, Phe). J Am Chem Soc 1981;103:2187.

92. Pogliani L, Ellenberger M, Valat J. A NMR investigation of proline and its derivatives. II—Conformational implications of the 1H NMR spectrum of L-proline at different pH. Org Magn Reson 1975;7:61.

93. Aliiev AE. Solid-state NMR studies of collagen-based parchments and gelatin. Biopolymers 2005;77:230.