J|A|C|S
ARTICLES
Published on Web 01/17/2007

Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study

Jürgen Graf,† Phuong H. Nguyen,‡ Gerhard Stock,*,‡ and Harald Schwalbe*,†

Contribution from the Institute of Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance, and Institute of Physical and Theoretical Chemistry, Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue Strasse 7, D-60438 Frankfurt/Main, Germany

Received August 19, 2006; E-mail: stock@theochem.uni-frankfurt.de; schwalbe@nmr.uni-frankfurt.de

Abstract: The ϕ,ψ backbone angle distribution of small homopolymeric model peptides is investigated by a joint molecular dynamics (MD) simulation and heteronuclear NMR study. Combining the accuracy of the measured scalar coupling constants and the atomistic detail of the all-atom MD simulations with explicit solvent, the thermal populations of the peptide conformational states are determined with an uncertainty of <5 %. Trialanine samples mainly (~90%) a poly-L-proline II helix-like structure, some (~10%) β extended structure, but no α_R helical conformations. No significant change in the distribution of conformers is observed with increasing chain length (Ala_3 to Ala_7). Trivaline samples all three major conformations significantly. Tryglycine samples the four corner regions of the Ramachandran space and exists in a slow conformational equilibrium between the cis and trans conformation of peptide bonds. The backbone angle distribution was also studied for the segment Ala_3 surrounded by either three or eight amino acids on both N- and C-termini from a sequence derived from the protein hen egg white lysozyme. While the conformational distribution of the central three alanine residues in the 9mer is similar to that for the small peptides Ala_3-Ala_7, major differences are found for the 19mer, which significantly (30−40%) samples α_R helical stuctures.

Introduction

The conformational properties of small, homopolymeric polypeptides is a matter of ongoing interest.[1,2] For example, the sampling of ϕ,ψ space of the polypeptide chain is of considerable interest for the understanding of protein folding. In addition, it has been shown recently that homopolymeric peptides can form fibrils if conditions are chosen properly.[3] Homopolymeric polypeptides have also been used as catalysts for organic reactions; the asymmetric epoxidation of α,β-unsaturated ketones is catalyzed by homopolymeric polypeptides such as polyalanine, but also polyisoleucine.[4] However, it is long known that investigation of the conformation of those polypeptides is challenging. They are ensembles of rapidly interconverting conformers, and a number of different approaches have been used to describe such conformational sampling. Those methods include the use of ϕ,ψ distributions derived from the database of protein structures and also molecular dynamics (MD) descriptions.[5−13]

Due to these difficulties, there is still some controversy about the conformational sampling of alanine-based short peptides. Employing two-dimensional (2D) infrared (IR) spectroscopy[14] in combination with density functional theory (DFT) calculations and MD simulations, it was suggested that cationic trialanine consists of two conformations at room-temperature, mainly a poly-L-proline II helix-like structure (PP_{II}) (≥80%) and some α_R helix.[15] Contrary to these findings, a 50:50 mixture of PP_{II} and an extended β-strand-like conformation was postulated on the basis of Raman, FTIR, and circular dichroism (CD) measurements.[16,17] Tetraalanine as cation was reported to adopt predominantly PP_{II} conformations in water[18] and also as zwitterion in cesium pentadecafluorooctanoate/water.[19] From CD and NMR spectra of the alanine-based peptide

† Institute of Organic Chemistry and Chemical Biology, Center for Biomolecular Magnetic Resonance.
‡ Institute of Physical and Theoretical Chemistry.

(1) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163−240.
(2) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877−1897.
(3) Fändrich, M; Dobson, C. M. *EMBO J.* **2002**, *21*, 5682−5690.
(4) Colonna, S.; Molinari, H.; Banfi, S.; Julia, S.; Masana, J.; Alvarez, A. *Tetrahedron* **1983**, *39*, 1635−1641.
(5) Smith, L. J.; Bolin, K. A.; Schwalbe, H.; MacArthur, M. W.; Thornton, J. M.; Dobson, C. M. *J. Mol. Biol.* **1995**, *255*, 494−506.
(6) Fiebig, K. M.; Schwalbe, H.; Buck, M.; Smith, L. J.; Dobson, C. M. *J. Phys. Chem.* **1996**, *100*, 2661−2666.
(7) Schwalbe, H.; Fiebig, K. M.; Buck, M.; Jones, J. A.; Grimshaw, S. B.; Glaser, S. J.; Smith, L. J.; Dobson, C. M. *Biochemistry* **1997**, *36*, 8977−8991.
(8) Mu, Y.; Kosov, D. S.; Stock, G. *J. Phys. Chem. B* **2003**, *107*, 5064−5073.
(9) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782−2787.
(10) Gnanakaran, S.; Garcia, A. E. *J. Phys. Chem. B* **2003**, *107*, 12555−12557.
(11) Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. *Proteins* **2004**, *55*, 502−507.
(12) Hu, H.; Elstner, M.; Hermans, J. *Proteins* **2003**, *50*, 451−463.
(13) Yoda, T.; Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **2004**, *386*, 460−467.
(14) Woutersen, S.; Hamm, P. *J. Phys. Chem. B* **2000**, *104*, 11316−11320.
(15) Woutersen, S.; Pfister, R.; Hamm, P.; Mu, Y.; Kosov, D. S.; Stock, G. *J. Chem. Phys.* **2002**, *117*, 6833−6840.
(16) Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2002**, *124*, 14330−14341.
(17) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2003**, *125*, 8178−8185.
(18) Schweitzer-Stenner, R.; Eker, F.; Griebenow, K.; Cao, X.; Nafie, L. A. *J. Am. Chem. Soc.* **2004**, *126*, 2768−2776.

Ac-XXAAAAAAAOO-NH_2 (short XAO, where X, A, and O denote diaminobutyric acid, alanine, and ornithine, respectively) it was proposed that PP_{II} is the dominant conformation.[20] This view has been questioned by other experimental results for this peptide which led to the conclusion that PP_{II} is simply one of several accessible conformational states.[21,22] From a comparison of Raman optical activity spectra of cationic Ala_2 to Ala_6 with spectra of the peptide Ac-OOAAAAAAAOO-NH_2 it was concluded that the PP_{II} propensity of alanine increases with the number of residues.[23] Trivaline, on the other hand, supposedly adopts mostly an extended β-sheet conformation.[16,17]

Concerning the theoretical description of small peptides in aqueous solution, numerous groups have performed quantum mechanical *ab initio* calculations[24,25] as well as classical MD simulations using molecular mechanics force fields.[8−13] While most studies reported a similar distribution within the main conformations α_R, β, and PP_{II}, the thermal populations of these states differ considerably, depending on the employed theoretical model.[8,12,13] The latter finding is a consequence of the fact that an accuracy $\Delta\Delta G$ of the relative free energy of ∼1−2 kcal/mol already introduces an uncertainty $\exp(\Delta\Delta G/k_BT)$ of a factor of 10 to the population probability.

The ϕ,ψ backbone angle distribution can be probed experimentally by NMR via spin−spin coupling constants which can be related to specific torsion angles by Karplus relationships.[26] For the backbone angle ϕ, four coupling constants $^3J(H_N,H_\alpha)$, $^3J(H_N,C')$, $^3J(H_\alpha,C')$, $^3J(H_N,C_\beta)$ are readily accessible, while for the backbone angle ψ, the three coupling constants $^1J(N,C_\alpha)$, $^2J(N,C_\alpha)$, and $^3J(H_N,C_\alpha)$ can be measured. The values of the measured coupling constants reflect the ensemble character of the conformational distribution, which has to be taken into account in their analysis.[27]

In this report, we attempt to combine the accuracy of the experimentally measured scalar coupling constants and the atomistic detail of the calculated structures. We therefore adopt the following strategy. Assuming that the force field gives a reasonable description of the structure and its distribution of the main conformations, we perform a global fit of their thermal populations by minimizing the deviation between measured and calculated NMR parameters. In this way, we have examined the side-chain dependence of conformational sampling of the tripeptides Ala_3, Val_3, and Gly_3. In addition, we examined the chain-length dependence of conformational sampling of Ala_3 to Ala_7. We were also interested in the comparison of the homopolymeric segment Ala_3 in the heteropolymeric sequence context of the protein hen egg white lysozyme (HEWL). For this, two peptides were synthesized comprising residues 6−14 and residues 1−19 of the full-length protein, respectively: C(S^{Me})EL**AAA**MKR (short HEWL-9mer) and KVFGRC(S^{Me})-EL**AAA**MKRHGLDN (short HEWL-19mer).

In our analysis, we find that trialanine samples mainly the PP_{II} conformation (∼90%) and the β-strand conformation (∼10%), while α helical structures are not sampled at all. By increasing the chain length from Ala_3 to Ala_7, no substantial change in this distribution of conformers is observed. As expected, the comparison of different trimeric peptides (Ala_3, Val_3, and Gly_3) reveals a significant variation of ϕ,ψ sampling determined by the size of the side chain. The sequence context of a given trimeric segment embedded within longer heteropolymeric peptides, however, modulates significantly the ϕ,ψ sampling of the central trimeric segment. In addition, the length of the polypeptide chain seems to be important, since significant differences are observed between the short (HEWL-9mer) and the long (HEWL-19mer) peptide derived from hen egg white lysozyme.

Materials and Methods

Peptide Synthesis. Peptides were synthesized either manually, by using standard Fmoc solid-phase peptide synthesis protocols,[28] or by using an Applied Biosystems 433A peptide synthesizer with standard Fmoc chemistry. *N*-Fmoc-*S*-methyl-L-cysteine was synthesized as published.[29] 2-Chlorotrityl chloride, H-Xaa-2-chlorotrityl resins, activating reagents and Fmoc amino acids were purchased from Novabiochem and isotopically labeled products from Cambridge Isotope Laboratories or Sigma-Aldrich. All other chemicals and solvents were of analytical grade and if necessary dried with molecular sieves[30] prior to use. The products were purified on reversed-phase HPLC columns and freeze-dried before preparation of samples. Peptide purifications were verified by analytical HPLC and electrospray mass spectrometry analysis on a Fisons Instruments VG Platform II to confirm molecular weight. The synthesized peptides with their isotopic labeling pattern are listed in Table S1 in the Supporting Information. All peptides investigated were free of any N- or C-terminal modifications.

HPLC. Analytical reversed-phase HPLC was performed on a Merck Hitachi system with a Eurospher-100 C18 column (5 μm, 4.6 mm × 250 mm) using conditions of 1 mL/min flow rate and 2%/min linear gradient of solvent B (0.1% trifluoroacetic acid in acetonitrile) in solvent A (0.1% trifluoroacetic acid in water). Semi-preparative reversed-phase HPLC was run on a Bruker LC21 system with a Kromasil-100 C18 column (5 μm, 20 mm × 250 mm) at a flow rate of 8 mL/min and a 2%/min linear gradient of solvent B in solvent A.

NMR Spectroscopy. The freeze-dried samples were dissolved in water, pH 2, containing 10% D_2O. The final concentrations of the NMR samples were determined by the ERETIC method[31] and varied between 0.9 and 88 mmol/L. The saturation concentrations of Ala_6 and Ala_7 were 8.8 and 0.9 mmol/L, respectively. For the XAO peptide, a concentration dependence of the CD spectra was reported.[22] We tested the concentration dependence of the *J*-coupling constants on Ala_3 for the $^3J(H_N,H_\alpha)$ coupling constant in the range from 0.2 to 88 mmol/L, and these were found to be independent of the concentration. Thus, we assume that this also holds true for the other peptides in this study. The NMR data were acquired on a Bruker 400 MHz spectrometer equipped with either a 5-mm $^1H\{^{13}C/^{15}N\}$ *z*-axis gradient probe or a 5 mm $^1H\{BB\}$ *z*-axis gradient probe, an 600 MHz instrument with a

(19) Pizzanelli, S.; Monti, S.; Forte, C. *J. Phys. Chem. B* **2005**, *109*, 21102−21109.
(20) Shi, Z.; Olson, C. A.; Rose, G. D., Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190−9195.
(21) Zagrovic, B.; Lipfert, J.; Sorin, E. J.; Millett, I. S.; van Gunsteren, W. F.; Doniach, S.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11698−11703.
(22) Makowska, J.; Rodziewicz-Motowidlo, S.; Baginska, K.; Vila, J. A.; Liwo, A.; Chmurzynski, L.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1744−1749.
(23) McColl, I. H.; Blanch, E. W.; Hecht, L.; Kallenbach, N. R.; Barron, L. D. *J. Am. Chem. Soc.* **2004**, *126*, 5076−5077.
(24) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908−5920.
(25) Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Mol. Struct. (THEOCHEM)* **2004**, *675*, 61−77.
(26) Karplus, M *J. Chem. Phys.* **1959**, *30*, 11−15.
(27) Wirmer, J.; Schlörb, C.; Schwalbe, H. Conformation and Dynamcis of Nonnative States of Proteins studied by NMR Spectroscopy. In *Protein Folding Handbook. Part I*; Buchner, J., Kiefhaber, T., Eds.; Wiley-VCH: Weinheim, 2005; pp 737−808.
(28) Chan, W. C., White, P. D., Eds. *Fmoc Solid Phase Peptide Synthesis*; Oxford University Press: Oxford, 2000.
(29) Hart, D. J.; Magomedov, N. A. *J. Am. Chem. Soc.* **2001**, *123*, 5892−5899.
(30) Merck KGaA, *Drying Agents*; http://uk.chemdat.info/mda/uk/broch/index.html.
(31) Akoka, S.; Barantin, L.; Trierweiler, M. *Anal. Chem.* **1999**, *71*, 2554−2557.

5-mm ^{1}H{^{13}C/^{15}N} *z*-axis gradient probe, a 700 MHz spectrometer with a 5-mm ^{1}H{^{13}C/^{15}N} *z*-axis gradient cryogenic probe, a 800 MHz spectrometer with a 5-mm ^{1}H{^{13}C/^{15}N} *z*-axis gradient cryogenic probe and a 900 MHz spectrometer with a 5-mm ^{1}H{^{13}C/^{15}N} *z*-axis gradient cryogenic probe. The NMR data were processed with Bruker XWIN-NMR 3.5 and TopSpin 1.3 programs and analyzed with either Bruker programs or Felix2000 (Accelrys). An automated routine[32] was used for extracting the *J*-coupling constants from the E.COSY pattern in Felix2000. The temperature was calibrated by methanol or glycol thermometer[33] for each spectrometer. ^{1}H chemical shifts were referenced to the methyl resonance of internal DSS (3-(trimethylsilyl)-1-propanesulfonic acid sodium salt). ^{13}C and ^{15}N chemical shifts were referenced indirectly to the ^{1}H standard using published[34] conversion factors. Spectral resonance assignment was done with a combination of standard HSQC and HMBC experiments up to Ala$_5$. For Ala$_6$ and Ala$_7$, a semiconstant time version of the HNN[35] experiment was applied. The HEWL-peptides were assigned following the standard method[36] utilizing the program CARA.[37] Chemical shift values are listed in the Supporting Information. Chemical exchange was measured with phase-sensitive[38] ^{1}H,^{1}H NOESY experiments with excitation sculpting[39] for water suppression. The forward and reverse rate constants were obtained from the experimental peak volumes by solving eq 26 in the review article of Perrin and Dwyer[40] using the program EXSYCalc (Mestrelab Research). The given error is the standard deviation of the obtained rate constants for different mixing times.

It was found that the $^3J(H_N,H_\alpha)$ coupling constant measured in ^{1}H one-dimensional (1D) spectra depends on the exact NMR pulse sequences which implement different water suppression schemes. For gradient-based suppression methods such as excitation sculpting, homonuclear scalar coupling evolves during the suppresion schemes which leads to a significant larger coupling constant. Thus, the $^3J(H_N,H_\alpha)$ coupling constants were measured in 1D ^{1}H spectra with presaturation for water suppression. There, the time during which homonuclear coupling can evolve is kept to a minimum and can be safely corrected for by linear phase correction. Direct determination of the $^3J(H_N,H_\alpha)$ coupling constants from the splitting of amide protons underestimates the true coupling constant value whenever the doublet components are not resolved to baseline. This seems to be the case for the reported[8] values of Ala$_3$. However, deconvolution of the spectrum by fitting a Lorentzian function to the peaks or apodization of the FID with a Lorentz-to-Gauss transformation prior to Fourier transformation recovers true *J*-coupling constants. In most cases, the average value derived by both methods is given, except for residue A^3, A^5 of Ala$_5$, A^{11} of HEWL-9mer (same numbering as for HEWL-19mer for easy comparison), and A^9 to A^{11} of HEWL-19mer where we only applied the apodization due to partial overlap of the signals. When applying the deconvolution routine it did not matter if we used a prior Lorentzian broadening factor for the window function between 0 and 1 Hz for spectra processing. In the case of Gly$_3$, the $^3J(H_N,H_\alpha)$ coupling constant was measured on the H$_\alpha$ protons because of line broadening of the amides signals. Note also that for the peptides with low molecular weight, systematic errors as reported by Harbison[41] due to relaxation of the second kind on the coupling constant determination are very small and can safely be neglected.

The temperature dependence of $^3J(H_N,H_\alpha)$ coupling constant of Ala$_3$ were measured on AAA, AA$^\#$A, and AAA$^\#$ (A$^\# = {}^{15}$N isotopic labeled). Coupling constants have been measured both from low to high temperature and from high to low temperature to ensure proper thermal equilibration of the NMR setup.

$^3J(H_N,C')$, $^3J(H_\alpha,C')$, $^3J(C',C')$, $^3J(H_N,C_\beta)$, $^3J(H_N,C_\alpha)$, $^1J(N,C_\alpha)$, and $^2J(N,C_\alpha)$ were measured with soft HNCA-COSY,[42] CO-coupled (H)-NCAHA,[43] (HN)CO(CO)NH,[44,45] HNHB[CB] E.COSY,[46] HNCO[CA] E.COSY,[47] and *J*-modulated ^{1}H, ^{15}N HSQC's,[48] respectively. The signal overlap in the (HN)CO(CO)NH experiment was so severe that it was only possible to measure the $^3J(C',C')$ coupling constant in a few cases. For the other experiments, the signal overlap of intra- and interresidual correlations was avoided either by measuring at higher magnetic fields or by the applied isotope-labeling schemes. The given statistical error is obtained from at least two measurements. Examples for the excellent quality of the experimental data and the summary of the used acquisition and processing parameters are found in Figure 1 and in the Supporting Information.

CD Spectroscopy. CD spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a Jasco PTC-423S temperature control unit using quartz cuvettes with 0.2-mm pathlengths. Data were collected at 0.5-nm increments from 260 to 185 nm with a scanning speed of 100 nm/min. For the measurements the same solvents as for the NMR measurements were used, and the peptide concentrations were the following: c(Ala$_3$) = 1.6 mmol/L, c(Ala$_7$) = 0.9 mmol/L, c(Val$_3$) = 2.6 mmol/L, c(HEWL-9mer) = 0.6 mmol/L, and c(HEWL-19mer) = 0.1 mmol/L. Ten scans were averaged (HEWL-19mer 20 scans), and the solvent baseline was subtracted, but no line smoothing was applied.

Molecular Dynamics Simulation. We used the GROMOS96 force field 43a1[49] to model the peptides and the SPC water model[50] to describe the solvent. The peptides Ala$_n$ were placed in cubic boxes containing 650, 807, 874, 1096, and 1243 water molecules for $n = 3$, 4, 5, 6, and 7, respectively. The simulation boxes of the Gly$_3$ and Val$_3$ were similar to that of Ala$_3$. In all simulations, the GROMACS program suite[51,52] was employed. The equations of motion were integrated by using a leapfrog algorithm with a time step of 2 fs. Covalent bond lengths were constrained via the SHAKE[53] procedure with a relative geometric tolerance of 10^{-4}. We used the particle-mesh Ewald method to treat the long-range electrostatics interactions.[54] The nonbonded interaction pair-lists were updated every 5 fs, using a cutoff of 1.2 nm. The systems were minimized using the conjugate gradient method. Subsequently, the solvated systems were equilibrated for 100 ps at constant pressure (1 atm) and temperature ($T = 300$ K), respectively, using the Berendsen coupling procedure.[55] Each system was then run for 100 ns, and the data were collected every 0.2 ps.

The calculation of the *J*-coupling constants from the MD simulations is based on Karplus relations of the type $J(\varphi) = A\cos^2(\varphi + \theta) + B$

(32) Schwalbe, H.; Rexroth, A.; Eggenberger, U.; Geppert, T.; Griesinger, C. *J. Am. Chem. Soc.* **1993**, *115*, 7878−7879.
(33) Braun, S.; Kalinowski, H.-O.; Berger, S. *150 and More Basic NMR Experiments: A Practical Course*, 2nd ed.; Wiley-VCH: Weinheim, 1998.
(34) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. *Pure Appl. Chem.* **1998**, *70*, 117−142.
(35) Panchal, S. C.; Bhavesh, N. S.; Hosur, R. V. *J. Biomol. NMR* **2001**, *20*, 135−147.
(36) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York 1986.
(37) Keller, R. *The Computer Aided Resonance Assignment Tutorial*; CANTINA: Goldau, Switzerland, 2004. http://www.nmr.ch.
(38) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286−292.
(39) Hwang, T.-L.; Shaka, A. J. *J. Magn. Reson., Series A* **1995**, *112*, 275−279.
(40) Perrin, C. L.; Dwyer, T. J. *Chem. Rev.* **1990**, *90*, 935−967.
(41) Harbison, G. S. *J. Am. Chem. Soc.* **1993**, *115*, 3026−3027.
(42) Weisemann, R.; Rüterjans, H.; Schwalbe, H.; Schleucher, J.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4*, 231−240.
(43) Löhr, F.; Rüterjans, H. *J. Biomol. NMR* **1995**, *5*, 25−36.
(44) Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 8170−8171.
(45) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1997**, 9, 207−211.
(46) Löhr, F.; Rüterjans, H. *J. Biomol. NMR* **1999**, *13*, 263−274.
(47) Hennig, M.; Bermel, W.; Schwalbe, H.; Griesinger, C. *J. Am. Chem. Soc.* **2000**, *122*, 6268−6277.
(48) Wirmer, J.; Schwalbe, H. *J. Biomol. NMR* **2002**, *23*, 47−55.
(49) Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag at the ETH Zürich: Zürich, 1996.
(50) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Reidel: Dordrecht, 1981; pp 331−342.
(51) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(52) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306−317.
(53) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(54) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(55) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F. Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

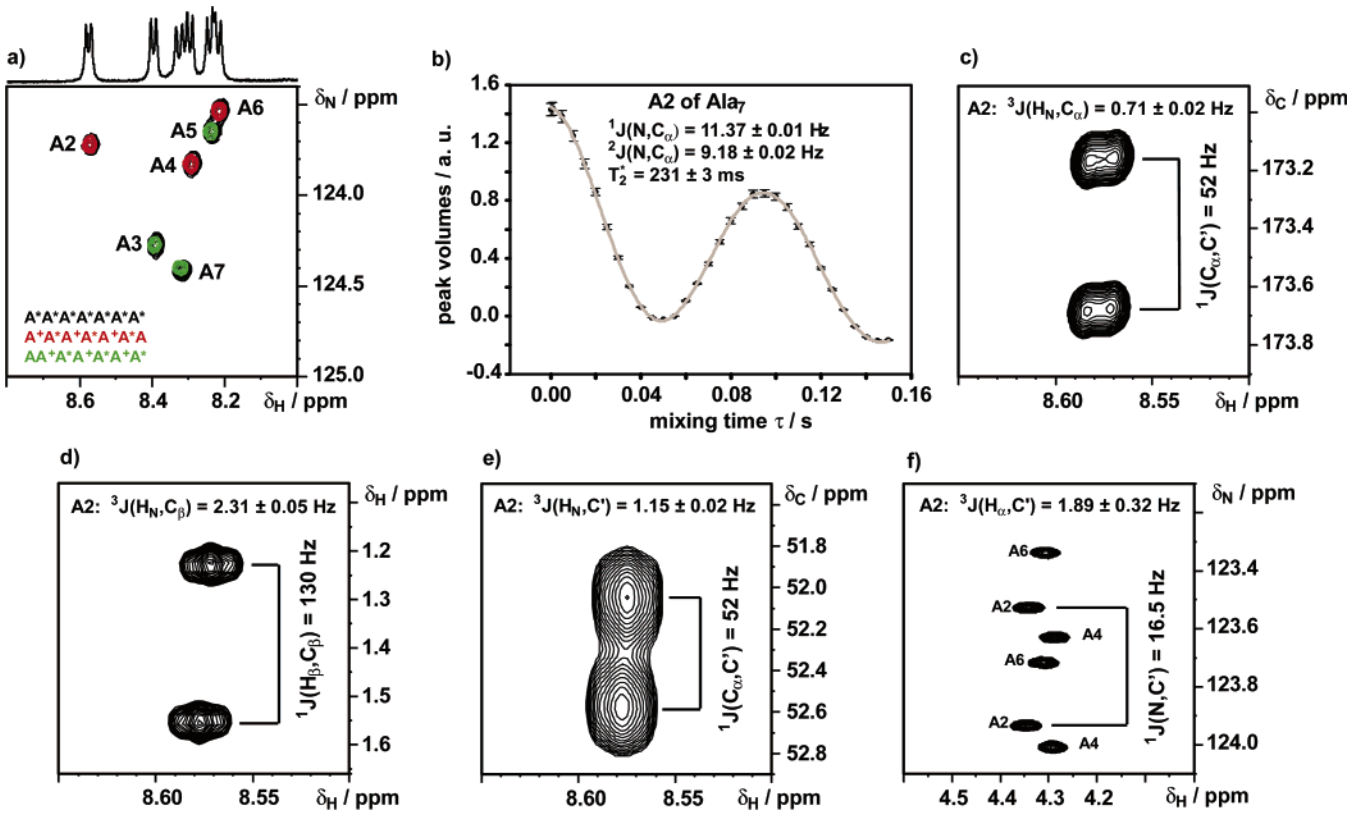

Figure 1. Example of the experimental NMR data for Ala_7. Each spectrum was recorded at 400 MHz and 300 K. Acquisition and processing parameters are listed in the Supporting Information. If not otherwise stated, the measured *J*-coupling constant is indicated together with the statistical error obtained from two measurements. Isotopical labeling pattern: A = unlabeled; A* = fully ^{13}C,^{15}N-labeled; A^+ = $^{13}C-C'$ labeled. (a) ^1H, ^{15}N HSQC spectra of the different isotopically labeled Ala_7-peptides. On top, the proton 1D spectra with presaturation for water suppression of the unlabeled peptide. (b) Determination of $^1J(N,C_\alpha)$ and $^2J(N,C_\alpha)$ coupling constants. Experimental peak volumes (black circles) of the *J*-modulated ^1H, ^{15}N HSQC spectra and fitting the equation $I = A\cos(\pi^1 J\tau)\cos(\pi^2 J\tau)\exp(-\tau/T_2^*)$ to them (gray line) for the peaks of A2. Four spectra with different mixing times were measured twice, and the largest percentage deviation of the peak volumes was taken for the error bars of all peak volumes. The quality of the fit is $R^2 = 0.9999$. Fitting was performed using the program SigmaPlot 9.0. The obtained *J*-coupling constants are indicated together with the error of the fit. (c) Determination of $^3J(H_N,C_\alpha)$ coupling constant from a HNCO-E.COSY spectrum of A*A*A*A*A*A*A*. The contour plot shows a cross section through $^1H_N-^{13}C'$ signals taken at the ^{15}N resonance position of A2. (d) Determination of $^3J(H_N,C_\beta)$ coupling constant from a HNHB-E.COSY spectrum of $A^+A^*A^+A^*A^+A^*A$. The contour plot shows a cross section through $^1H_N-^1H_\beta$ signals taken at the ^{15}N resonance position of residue A2. (e) Determination of $^3J(H_N,C')$ coupling constant from HNCA-COSY spectrum of $A^+A^*A^+A^*A^+A^*A$. The contour plot shows a cross section through $^1H_N-^{13}C_\alpha$ signals taken at the ^{15}N resonance position of residue A2. (f) Determination of $^3J(H_\alpha,C')$ coupling constant from (H)N(CA)HA spectrum of $A^+A^*A^+A^*A^+A^*A$.

$\cos(\varphi + \theta) + C$, where φ denotes either the ϕ or the ψ backbone dihedral angle of the peptide, *A*, *B*, *C* represent the parametrization of the relation, and θ signifies the phase shift for the specific *J* of interest. For all *J*-coupling constants considered, Table S2 in the Supporting Information lists the parameters of the Karplus relation adapted from refs 47, 48, 56, 57, and Figure 2 shows the corresponding Karplus curves $J(\varphi)$ together with the typical conformational distributions (see below) along the dihedral angles ϕ or ψ. The figure nicely shows, e.g., that the $^3J(H_N,H_\alpha)$ coupling constant represents an accurate probe of the peptide ϕ angle, since the Karplus curve varies considerably along the β and the PP_{II} conformations. On the other hand, the $^2J(N,C_\alpha)$ is given as a function of the ψ angle and therefore allows clear discrimination between helical α_R from extended β and PP_{II} conformations.

As explained in the Introduction, we will assume that the force field gives a reasonable description of the structure of the main conformations α_R, β, and PP_{II}, whereas the calculated thermal populations P_α, P_β, and $P_{PP_{II}}$ of these states may deviate considerably from their true value. To determine the correct population probabilities, we therefore perform a global fit of the populations by minimizing the deviation between measured and calculated NMR parameters defined by

$$\chi^2(P_\alpha,P_\beta,P_{PP_{II}}) = \sum_k [J_k^{exp} - (P_\alpha J_k^\alpha + P_\beta J_k^\beta + P_{PP_{II}} J_k^{PP_{II}})]^2 \qquad (1)$$

Here, J_k^{exp}represents the measured coupling constant and J_k^s denotes the calculated coupling constant in conformation $s = \alpha_R$, β, and PP_{II}, respectively. The sum runs over all *J*-coupling constants available for the fit. It is noted that each coupling constant J_k^s is obtained from its corresponding Karplus relation (see Table S2) by averaging the coupling J_k^s overall MD structures pertaining to conformation *s*.

Assuming that only the three main conformational states are thermally populated, we have $P_\alpha + P_\beta + P_{PP_{II}} = 1$ and χ^2 in eq 1 can be expressed in terms of only two variables, e.g.,

$$\chi^2(P_\alpha,P_\beta) = \sum_k [(J_k^{exp} - J_k^{PP_{II}}) - P_\alpha(J_k^\alpha - J_k^{PP_{II}}) - P_\beta(J_k^\beta - J_k^{PP_{II}})]^2 \qquad (2)$$

To determine the correct population probabilities P_s, this function is minimized either analytically (requiring that $0 \le P_s \le 1$) or numerically by simply evaluating $\chi^2(P_\alpha, P_\beta)$ on a 2D grid. As a representative example, Figure S1 in the Supporting Information shows this function for the fourth residue of Ala_7. The global minimum of χ^2 at ($P_\alpha = 0$, $P_\beta = 0.15$) is found to be relatively flat. Typically, we therefore estimate the uncertainty of the populations to be ≤ 5 %. To study the robustness of the fits, moreover, we have performed a number of tests which study the behavior of χ^2 when one or several coupling constants were excluded from the fit or the Karplus curves were shifted

(56) Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 6360−6368.
(57) Ding, K.; Gronenborn, A. M. *J. Am. Chem. Soc.* **2004**, *126*, 6232−6233.

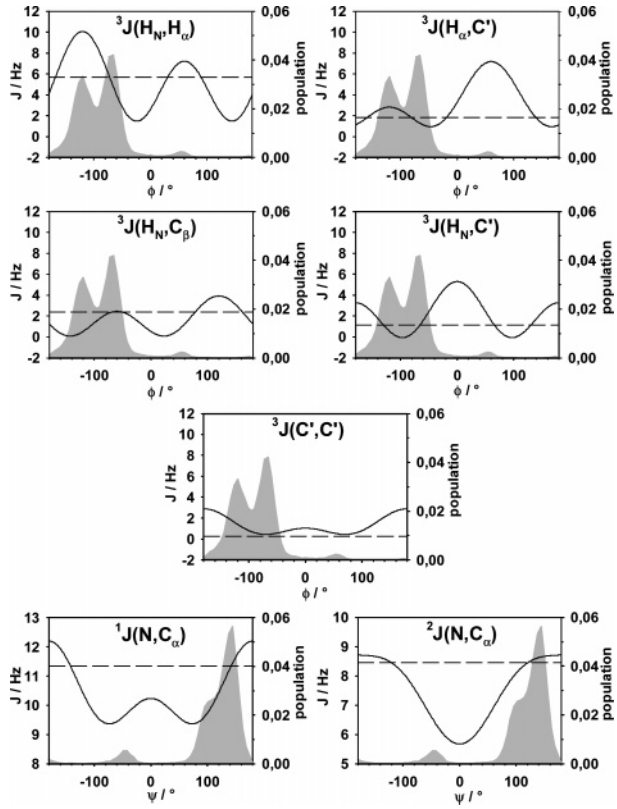

Figure 2. Karplus curves (solid lines) of the measured *J*-coupling constants (dashed lines) for residue A2 of Ala_3, together with the conformational distributions (gray shaded) obtained from the MD simulation.

by an uncertainty of 0.5 Hz. As documented in the Supporting Information, all tests yielded quite similar population probabilities.

Besides the dependence on the Karplus relations, we furthermore studied the reliability of our results with respect to the accuracy of the force field and the convergence of the conformational sampling. Our working assumption, that (apart from thermal weighting factors) the force field gives a reasonable description of the structural distribution within a conformational state, is based on a careful analysis of comparisons of various force fields for polyalanines.[8−10,12] For example, by comparing the six popular MD force fields including AMBER94,[58] CHARMM22,[59] GROMOS96,[49] and OPLS-AA[60] for trialanine, Mu et al.[8] found that the mean values of the (ϕ, ψ)-distributions pertaining to a specific conformational state differed at most by ±20° between the various force fields, while the width of the distributions were quite similar. As shown in Table S14 and Figure S1 in the Supporting Information, the fitted thermal populations are not very sensitive to a 25° shift of the conformational states, demonstrating that our results are quite robust with respect to uncertainties of the force field. To test if our results are converged with respect to the conformational sampling, we have recalculated the (ϕ, ψ)-distributions using either the first or the second half of the 100 ns trajectory. The resulting conformational distributions are virtually the same, thus reflecting the fact that the single-residue distributions $P(\phi_i,\psi_i)$ are highly averaged quantities. We note in passing that the total conformational distribution $P(\phi_1,\psi_2, ..., \phi_n,\psi_n)$ of Ala_n is (by far) not converged in a 100 ns MD simulation, but this information is not required in the present work.

(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(59) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(60) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

Results and Discussion

By NMR spectroscopy, a number of different structure-dependent parameters can be measured, for example NOEs, chemical shifts, residual dipolar and scalar coupling constants. In highly flexible molecules like the homopolymeric peptides studied here, the experimental NMR observables are an average of the conformational distribution with specific averaging time regime. NOEs are averaged by $\langle r^{-6}\rangle$ which biases toward shorter distances in cases of conformational averaging. In addition, strong signal overlap for NOEs is a serious problem in homooligomers. Also, the time regime of internal and overall rotational reorientation influences the NOE cross-peak intensities. Therefore, both chemical shifts and NOEs are rather insensitive experimental restraints to derive information about highly averaged structures. Hence, we have chosen to test the conformational distribution by measuring scalar coupling constants which in most cases report directly on a conformational ensemble around a single torsion, e.g., the backbone angles ϕ or ψ via Karplus parametrizations, which can be directly related to predictions, derived from MD simulations.

The most precise and accurate way to measure vicinal coupling constants is from analyzing the signal splitting in a 1D spectrum. This was possible for the $^3J(H_N,H_\alpha)$ coupling constants for the peptides studied here. The $^3J(H_N,H_\alpha)$ coupling constant is also the most sensitive reporter for the backbone angle ϕ out of the six on ϕ-dependent coupling constants due to the fact that refined Karplus parametrizations are available[56,61] and the value of coupling constant varies by 8 Hz over the interesting range of ϕ. For the angle ψ, the $^2J(N,C_\alpha)$ coupling constant is the most sensitive one, but its smaller variation of about 3 Hz makes the possible error higher.

Comparison of experimental *J*-coupling constants to calculated *J*-couplings from MD simulations reveals a remarkable overall agreement (cf. Table 1, Table 3, and Tables S3 to S13) considering the inherent error of each method. The average deviation for each residue varies between 7 and 25%, with an overall deviation of 14%. The largest deviations are observed for the $^3J(H_N,C_\beta)$, $^1J(N,C_\alpha)$, $^2J(N,C_\alpha)$, and $^3J(H_N,C_\alpha)$ coupling constants, which is partly due to the smaller range of values these coupling constants can adopt. Hence, the overall agreement suggests that the force field gives a reasonable description of the peptide structure and that the above-mentioned strategy combining NMR and MD is justified.

Side-Chain Dependence of Tripeptide Conformational Sampling. To obtain a first impression on the typical conformational states sampled by MD simulations, Figure 3 displays the Ramachandran probability distribution of the central dihedral angles (ϕ,ψ) of the three tripeptides Ala_3, Val_3, and Gly_3. For Ala_3 and Val_3 peptides, there are essentially three populated conformational states: the right-handed helix conformation α_R ($-150° < \phi < -25°$ and $-150° < \psi < 0°$), the β conformation ($-150° < \phi < -90°$ and $80° < \psi < 160°$), and the PP_{II} ($-90° < \phi < -25°$ and $80° < \psi < 160°$) conformation, which are located at (ϕ,ψ) ≈ (−80°, −50°), (−120°, 130°), (−60°, 140°), respectively. In addition, there is a small population (≤3%) of the left-handed helix conformation α_L located at (50°, 100°). For each peptide, the population probability of each conformer is listed in Table 1. The simulations using the GROMOS96 force

(61) Vuister, G. W.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 7772−7777.

Table 1. *J*-Coupling Constants for the Central Residue of the Tripeptides Ala_3, Val_3, and Gly_3 as Obtained from Simulation (MD), Fitting (Fit), and Experiment (Exp) Together with the Corresponding Populations Probability of the Conformations α_R, β, and PP_{II}

		J-coupling constants/Hz								populations/%	
peptide		$^3J(H_N,H_\alpha)$	$^3J(H_N,C')$	$^3J(H_\alpha,C')$	$^3J(C',C')$	$^3J(H_N,C_\beta)$	$^1J(N,C_\alpha)$	$^2J(N,C_\alpha)$	$^3J(H_N,C_\alpha)$	P_{MD}	P_{Fit}
Ala_3	α_R	5.6 ± 2.5	1.3 ± 1.0	1.6 ± 0.6	0.7 ± 0.4	2.0 ± 0.7	9.7 ± 0.2	6.5 ± 0.5	0.6 ± 0.1	15	0
	β	9.4 ± 0.9	0.8 ± 0.7	2.6 ± 0.3	1.5 ± 0.5	0.6 ± 0.4	10.9 ± 0.8	8.4 ± 0.3	0.8 ± 0.1	40	8
	PP_{II}	5.3 ± 1.7	1.1 ± 0.8	1.5 ± 0.4	0.5 ± 0.1	2.3 ± 0.2	10.9 ± 0.8	8.5 ± 0.3	0.6 ± 0.1	41	92
	MD	7.0 ± 2.6	1.1 ± 1.0	2. ± 1.0	0.9 ± 0.6	1.5 ± 0.9	10.8 ± 0.8	8.3 ± 0.6	0.7 ± 0.1		
	Fit	5.6	1.1	1.5	0.6	2.1	10.9	8.5	0.6		
	Exp	5.68	1.13	1.84	0.25	2.39	11.34	8.45	0.70		
Val_3	α_R	7.1 ± 2.5	0.8 ± 0.8	2.0 ± 0.7	0.8 ± 0.4	1.7 ± 0.8	9.6 ± 0.2	6.6 ± 0.5	0.7 ± 0.1	13	19
	β	9.6 ± 0.7	0.6 ± 0.5	2.7 ± 0.2	1.3 ± 0.5	0.7 ± 0.4	10.5 ± 0.8	8.3 ± 0.4	0.8 ± 0.1	35	52
	PP_{II}	5.4 ± 1.8	1.0 ± 0.9	1.5 ± 0.4	0.5 ± 0.1	2.2 ± 0.2	10.5 ± 0.8	8.4 ± 0.3	0.6 ± 0.1	47	29
	MD	7.1 ± 2.5	0.9 ± 0.9	2.2 ± 1.1	0.8 ± 0.5	1.6 ± 0.8	10.4 ± 0.8	8.1 ± 0.7	0.7 ± 0.1		
	Fit	7.9	0.8	2.2	1.0	1.3	10.3	8.0	0.7		
	Exp	7.94	0.58	2.42	0.34	1.38	10.80	7.80	0.77		
Gly_3	MD	5.8 ± 2.7	1.2 ± 1.1	3.3 ± 2.1	1.3 ± 0.8	–	10.4 ± 1.0	8.1 ± 0.6	0.8 ± 0.2		
	Exp	5.89	1.10	4.01	0.26	–	12.17	9.05	0.78		

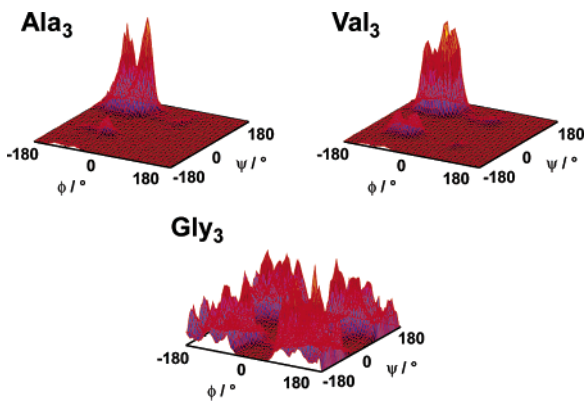

Figure 3. Ramachandran probability distributions for the central residues of Ala_3, Val_3, and Gly_3 as obtained from the MD simulations.

field predict that both peptides exist mainly in the extended conformations β and PP_{II} (~82%) and only little in the right-handed helix conformation α_R (~15%). Also, the populations of the β and PP_{II} conformations are comparable with each other. The Ramachandran plot of Gly_3 is quite different since it exhibits four populated regions with $\phi \leq -50°$ or $\phi \geq 50°$, and $\psi \leq -50°$ or $\psi \geq 50°$. Note that these regions actually are connected with each other due to the periodicity of the dihedral angles. Overall, the agreement between the simulation and experiment seems to be better for Val_3 and Gly_3 than for Ala_3. Note that for glycine residues, some of the Karplus parametrizations may not be too accurate because of very restricted set of data used for the parametrization.

From fitting the calculated to the experimental *J*-coupling constants (see Materials and Methods) we find that trialanine samples mainly the PP_{II} conformation (~90%). The β conformation has a population probability of about 10%, whereas the α_R conformation is not sampled. This is in qualitative agreement with the populations determined previously from a combination of 2D IR spectroscopy and MD simulations,[15] but does not support the postulated 50:50 mixture of PP_{II} and an extended β-strand-like conformation.[16,17] For trivaline with its branched side chain the population shifts more toward β and α_R conformations (cf. Table 1). This is best reflected in the *J*-coupling constants due to an increase in the $^3J(H_N,H_\alpha)$ coupling constant and a decrease in the $^2J(N,C_\alpha)$ coupling constant. The literature for trivaline so far postulates that it samples mostly the extended β-sheet conformation.[16,17]

The amide proton signals in the 1D ^1H spectrum of Gly_3 show significant line broadening. Exchange phenomena can be measured with EXSY[40] experiments related to the classical $^1H,^1H$ NOESY experiments but with longer mixing times. There was no water-to-amide H_N cross-peak detectable which rules out significant contributions of water exchange to the entire line width of 16 Hz observed for the H_N of Gly_3. At longer mixing times chemical exchange cross-peaks of the amide proton signals, with weak signals nearby, become visible (cf. Figure S7). The ratio of the diagonal signals gives a population ratio of 0.995 to 0.005 and 0.997 to 0.003 for H_{N2} and H_{N3}, respectively. From the analysis of the experimental peak volumes, using eq 26 in the review article of Perrin and Dwyer,[40] one obtains the forward and reverse rate constants which are 0.005 ± 0.001 Hz and 0.842 ± 0.132 Hz for H_{N2} and 0.002 ± 0.001 Hz and 0.818 ± 0.137 Hz for H_{N3}. The resultant equilibrium constants and free energy differences are $K_{cis/trans} = 0.006 \pm 0.002$, $\Delta G = 3.1 \pm 0.2$ kcal/mol for H_{N2} and $K_{cis/trans} = 0.002 \pm 0.002$, $\Delta G = 3.6 \pm 0.4$ kcal/mol for H_{N3}. We conclude that this chemical exchange process is a cis−trans isomerization of the peptide amide bond. In the shorter time scale sampled by the MD simulation this process is not visible. Such chemical exchange between cis and trans peptide conformation was not detected for any other peptide under study here.

Chain-Length Dependence of Ala_n Conformational Sampling. The conformational distributions from MD of every residue of the peptides Ala_4, Ala_5, Ala_6, and Ala_7 look quite similar to their counterparts of the Ala_3 peptide (data not shown), although the population probability of the conformations are different. Employing the Karplus relations described above, we have calculated coupling constants directly from the MD simulations for all polyalanines Ala_n ($n = 3-7$). Each coupling constant presented in Table 2 (denoted as MD) was averaged over all values calculated from every snapshot of (i) the entire trajectory and of (ii) each single conformational state α_R, β, and PP_{II}.

Table 2 shows, as an example, the coupling constants $^3J(H_N,H_\alpha)$ and $^2J(N,C_\alpha)$ which reflect the ϕ- and ψ-dependence of a peptide conformation, respectively. Note that the widths of the calculated coupling constants reported in the table account for conformational fluctuations of the peptide. These values are typically large for the $^3J(H_N,H_\alpha)$ coupling constant, because the

Table 2. Selected *J*-Coupling Constants from Simulation (MD), Fitting (Fit), and Experiment (Exp) of the Alanine Peptides as well as the Populations Obtained from Simulation and Fitting

		Ala_3			Ala_4			Ala_5			Ala_6			Ala_7		
res.		MD	Fit	Exp	MD	Fit	Exp	MD	Fit	Exp	MD	Fit	Exp	MD	Fit	Exp
2	$^3J(H_N,H_\alpha)$	7.0 ± 2.6	5.6	5.68	5.9 ± 2.6	5.5	5.62	5.9 ± 2.6	5.5	5.59	5.9 ± 2.6	5.5	5.60	5.5 ± 2.6	5.6	5.61
	$^2J(N,C_\alpha)$	8.3 ± 0.6	8.5	8.45	7.9 ± 0.9	7.8	8.56	7.8 ± 1.0	7.8	8.55	8.0 ± 0.9	7.7	8.52	7.7 ± 1.0	7.6	8.52
	$P_\alpha,P_\beta,P_{PP_{II}}$	15,40,41	0,8,92		25,23,43	0,14,86		31,20,40	0,14,86		22,22,46	0,17,83		40,17,35	0,17,83	
3	$^3J(H_N,H_\alpha)$				6.2 ± 2.6	5.9	5.89	6.0 ± 2.6	5.7	5.74	5.6 ± 2.6	5.6	5.67	5.4 ± 2.5	5.6	5.66
	$^2J(N,C_\alpha)$				7.7 ± 1.0	8.3	8.37	7.6 ± 1.0	7.5	8.40	7.4 ± 1.1	7.5	8.34	7.4 ± 1.0	7.6	8.29
	$P_\alpha,P_\beta,P_{PP_{II}}$				25,23,40	0,19,81		34,20,40	0,16,84		45,15,30	0,14,86		44,15,29	0,16,84	
4	$^3J(H_N,H_\alpha)$							6.4 ± 2.6	5.9	5.98	5.9 ± 2.5	5.7	5.80	5.5 ± 2.4	5.7	5.77
	$^2J(N,C_\alpha)$							7.5 ± 1.1	7.5	8.27	7.3 ± 1.1	7.4	8.26	7.1 ± 1.0	7.3	8.22
	$P_\alpha,P_\beta,P_{PP_{II}}$							33,23,33	0,17,83		45,18,28	0,13,87		57,12,21	0,15,85	
5	$^3J(H_N,H_\alpha)$										6.4 ± 2.7	6.0	6.02	5.9 ± 2.4	5.7	5.92
	$^2J(N,C_\alpha)$										7.5 ± 1.1	7.4	8.18	7.2 ± 1.1	7.4	8.24
	$P_\alpha,P_\beta,P_{PP_{II}}$										34,25,31	0,18,82		52,16,22	0,14,86	
6	$^3J(H_N,H_\alpha)$													6.2 ± 2.4	5.9	6.04
	$^2J(N,C_\alpha)$													7.3 ± 1.1	8.1	8.18
	$P_\alpha,P_\beta,P_{PP_{II}}$													43,21,25	0,17,83	

corresponding Karplus relation is steep in the populated regions, see Figure 2 (top left panel).

Let us first discuss the $^3J(H_N,H_\alpha)$ coupling constant as the most sensitive reporter on the dihedral angle ϕ. Interestingly, the experimental results show that the $^3J(H_N,H_\alpha)$ coupling constants are very much the same (the variation of the results is less than 0.5 Hz) for most peptides. A closer inspection, however, reveals the trend that the value of $^3J(H_N,H_\alpha)$ increases along a peptide chain going from the N-terminal to the C-terminal end, which can be explained by a small change from PP_{II} toward more β conformations. The simulations also show a similar behavior, except for the Ala_3. The calculated values of these constants for Ala_3 are larger than their counterparts of the other polyalanine peptides. This is due to the fact that the β conformation of Ala_3 is much more populated, and according to the Karplus relation (Figure 1, top left panel), the $^3J(H_N,H_\alpha)$ value is large for the β conformation.

The $^2J(N,C_\alpha)$ is the most sensitive coupling constant for the dihedral angle ψ and therefore may provide information on the transition between the extended and helical states. However, its range of values for different conformations is not as large as that for the $^3J(H_N,H_\alpha)$ coupling constant (see the Karplus curve in Figure 1, bottom right panel). Similar to the $^3J(H_N,H_\alpha)$, the experimental values of $^2J(N,C_\alpha)$ for different peptide groups do not exhibit large variations. The values of the simulated $^2J(N,C_\alpha)$, however, are smaller than those obtained from experiment, which suggests that the MD simulations overestimate the helical conformations. The values of the experimentally determined $^2J(N,C_\alpha)$ decrease along peptide chains from the N-terminal to the C-terminal end. This indicates a trend toward a larger population of helical conformation along the peptide chain. Interestingly, the experimental values of $^2J(N,C_\alpha)$ ($\sim$8.4 Hz) suggest that all polyalanine peptides under study hardly exist in the helical conformations. The values are around 6.5 Hz, otherwise, according to the Karplus relation.

Tables 3 and S3−S13 show in detail the simulated results assuming resolved states of all the coupling constants of all peptide groups for all polyalanine peptides, together with the experimental data. Again, we note a good overall agreement between theory and experiment. The values of the coupling constants do not show significant deviation regardless of different peptides or peptide groups. Therefore, the conformational sampling has to be very similar, and an increase of PP_{II} conformation with increasing chain length as published[23,62] seems to be implausible.

Temperature Dependence of Ala_3 Conformational Sampling. To investigate the influence of the temperature on peptide structure, we have measured seven *J*-coupling constants for Ala_3 at four temperatures. The results for $^3J(H_N,H_\alpha)$ increase with temperature (from 5.33 Hz at 275 K to 6.29 Hz at 350 K for residue A2), which is well represented by a linear fit, see Figure 4. While a linear increase of $^3J(H_N,H_\alpha)$ was also reported for the XAO peptide,[20] in Ac-GG(A)$_n$GG-NH$_2$ ($n = 1, 2, 3$) peptides the $^3J(H_N,H_\alpha)$ of the Ala residues were reported to exhibit a transition curve behavior.[63,64] As $^3J(H_N,H_\alpha)$ is significantly higher in the β than in the PP_{II} conformation, its increase with temperature indicates that the β conformational state is more populated, which agrees with the findings of refs 20, 63, and 64. The other three ϕ-dependent *J*-coupling constants, $^3J(H_\alpha,C')$, $^3J(H_N,C_\beta)$, and $^3J(H_N,C')$ show a temperature behavior consistent with that interpretation (cf. Table 4). On the other hand, both ψ-dependent coupling constants $^1J(N,C_\alpha)$ and $^2J(N,C_\alpha)$ show only a small decrease with increasing temperature, suggesting that the population remains in the extended region and only a minor α helical content is populated as the temperature increases.

To support this interpretation by the MD simulations, we have used the state-dependent *J*-coupling constants obtained for Ala_3 at $T_0 = 300$ K (see Table 1) to calculate the thermal population of the various conformational states via a global fit of the calculated to the measured coupling constants $J(T)$. (We note that this assumes that the structural distribution within a conformational state is similar for all considered temperatures.) Table 4 shows that the resulting calculated *J*-coupling constants are in excellent agreement with experiment and also confirms the interpretation above. While the population of the β state increases from 3 to 26% when the temperature is increased from $T = 275-350$ K, the α state is hardly ($\leq$1%) populated even at $T = 350$ K. From the thus obtained thermal populations for the β and the PP_{II} conformers, we may calculate the temperature-

(62) Hagarman, A.; Measey, T.; Doddasomayajula, R. S.; Dragomir, I.; Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Phys. Chem. B.* **2006**, *110*, 6979−6986.
(63) Ding, L.; Chen, K.; Santini, P. A.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2003**, *125*, 8092−8093.
(64) Chen, K.; Liu, Z.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15352−15357.

Table 3. *J*-Coupling Constants for Residues A2 to A6 of Ala_7; Shown Are Results from Simulation (MD), Fitting (Fit), and Experiment (Exp) Together with the Populations of the Conformational States α_R, β, and PP_{II}, as Obtained from Simulation and Fitting

		J-coupling constants/Hz						populations/%
res.	type (angle)	α_R	β	PP_{II}	MD	Fit	Exp	P_α, P_β, $P_{PP_{II}}$
	$^3J(H_N,H_\alpha)$ (ϕ_2)	4.7 ± 2.3	9.3 ± 1.0	4.8 ± 1.7	5. ± 2.6	5.6	5.61	
	$^3J(H_N,C')$ (ϕ_2)	1.7 ± 1.2	0.8 ± 0.8	1.3 ± 0.9	1.5 ± 1.2	1.2	1.15	40, 17, 35 (MD)
	$^3J(H_\alpha,C')$ (ϕ_2)	1.5 ± 0.5	2.6 ± 0.3	1.4 ± 0.4	1.8 ± 1.2	1.6	1.89	
2	$^3J(H_N,C_\beta)$ (ϕ_2)	2.1 ± 0.6	0.6 ± 0.4	2.4 ± 0.2	1.8 ± 0.8	2.0	2.31	
	$^1J(N,C_\alpha)$ (ψ_2)	9.7 ± 0.2	10.7 ± 0.9	11.0 ± 0.8	10.4 ± 0.9	10.9	11.37	0, 17, 83 (Fit)
	$^2J(N,C_\alpha)$ (ψ_2)	7.6 ± 1.0	7.6 ± 1.0	7.7 ± 1.0	7.7 ± 1.0	7.6	8.52	
	$^3J(H_N,C_\alpha)$ (ϕ_2, ψ_1)	0.6 ± 0.1	0.8 ± 0.1	0.6 ± 0.1	0.6 ± 0.1	0.6	0.71	
	$^3J(H_N,H_\alpha)$ (ϕ_3)	4.5 ± 2.1	9.4 ± 1.0	4.9 ± 1.7	5.4 ± 2.5	5.6	5.66	
	$^3J(H_N,C')$ (ϕ_3)	1.7 ± 1.1	0.8 ± 0.8	1.3 ± 0.8	1.5 ± 1.2	1.2	1.20	45, 15, 29 (MD)
	$^3J(H_\alpha,C')$ (ϕ_3)	1.4 ± 0.5	2.6 ± 0.3	1.4 ± 0.4	1.9 ± 1.4	1.6	1.85	
3	$^3J(H_N,C_\beta)$ (ϕ_3)	2.2 ± 0.5	0.7 ± 0.4	2.3 ± 0.2	1.9 ± 0.7	2.0	2.20	
	$^1J(N,C_\alpha)$ (ψ_3)	9.7 ± 0.2	10.6 ± 0.9	10.8 ± 0.8	10.2 ± 0.8	10.8	11.27	0, 16, 84 (Fit)
	$^2J(N,C_\alpha)$ (ψ_3)	7.3 ± 1.0	7.4 ± 1.1	7.6 ± 1.0	7.4 ± 1.0	7.6	8.29	
	$^3J(H_N,C_\alpha)$ (ϕ_3, ψ_2)	0.4 ± 0.2	0.6 ± 0.2	0.4 ± 0.2	0.5 ± 0.2	0.5	0.66	
	$^3J(H_N,H_\alpha)$ (ϕ_4)	4.7 ± 2.0	9.4 ± 1.0	5.0 ± 1.7	5.5 ± 2.4	5.7	5.77	
	$^3J(H_N,C')$ (ϕ_4)	1.5 ± 1.0	0.8 ± 0.8	1.2 ± 0.9	1.5 ± 1.1	1.1	1.20	57, 12, 21 (MD)
	$^3J(H_\alpha,C')$ (ϕ_4)	1.4 ± 0.5	2.6 ± 0.3	1.4 ± 0.4	2.0 ± 1.5	1.6	1.80	
4	$^3J(H_N,C_\beta)$ (ϕ_4)	2.2 ± 0.5	0.7 ± 0.4	2. ± 0.2	1.9 ± 0.7	2.0	2.23	
	$^1J(N,C_\alpha)$ (ψ_4)	9.7 ± 0.2	10.6 ± 0.9	10.7 ± 0.8	10.1 ± 0.7	10.7	11.22	0, 15, 85 (Fit)
	$^2J(N,C_\alpha)$ (ψ_4)	6.9 ± 0.9	7.0 ± 1.1	7.4 ± 1.1	7.1 ± 1.0	7.3	8.22	
	$^3J(H_N,C_\alpha)$ (ϕ_4, ψ_3)	0.3 ± 0.2	0.6 ± 0.2	0.4 ± 0.2	0.4 ± 0.3	0.5	0.56	
	$^3J(H_N,H_\alpha)$ (ϕ_5)	5.2 ± 2.1	9.4 ± 1.0	5. ± 1.7	5.9 ± 2.4	5.7	5.92	
	$^3J(H_N,C')$ (ϕ_5)	1.3 ± 0.9	0.8 ± 0.8	1.2 ± 0.8	1.3 ± 1.0	1.1	1.19	52, 16, 22 (MD)
	$^3J(H_\alpha,C')$ (ϕ_5)	1.5 ± 0.5	2.6 ± 0.3	1.4 ± 0.4	2.0 ± 1.4	1.6	1.56	
5	$^3J(H_N,C_\beta)$ (ϕ_5)	2.2 ± 0.5	0.7 ± 0.4	2.3 ± 0.2	1.9 ± 0.8	2.0	2.23	
	$^1J(N,C_\alpha)$ (ψ_5)	9.8 ± 0.2	10.5 ± 0.9	10.7 ± 0.8	10.1 ± 0.7	10.6	11.29	0, 14, 86 (Fit)
	$^2J(N,C_\alpha)$ (ψ_5)	7.0 ± 1.0	7.0 ± 1.0	7.4 ± 1.1	7.2 ± 1.1	7.4	8.24	
	$^3J(H_N,C_\alpha)$ (ϕ_5, ψ_3)	–	–	–	0.4 ± 0.2	–	–	
	$^3J(H_N,H_\alpha)$ (ϕ_6)	5.5 ± 2.2	9.3 ± 1.2	5.3 ± 1.8	6.2 ± 2.4	5.9	6.04	
	$^3J(H_N,C')$ (ϕ_6)	1.2 ± 0.9	0.9 ± 0.9	1.1 ± 0.9	1.3 ± 1.1	1.1	1.10	43, 21, 25 (MD)
	$^3J(H_\alpha,C')$ (ϕ_6)	1.5 ± 0.5	2.6 ± 0.4	1.5 ± 0.4	2. ± 1.6	1.6	1.67	
6	$^3J(H_N,C_\beta)$ (ϕ_6)	2. ± 0.6	0.7 ± 0.4	2.3 ± 0.2	1.7 ± 0.8	2.0	2.21	
	$^1J(N,C_\alpha)$ (ψ_6)	9.8 ± 0.2	10.4 ± 0.9	10.6 ± 0.8	10.1 ± 0.7	10.6	11.29	0, 17, 83 (Fit)
	$^2J(N,C_\alpha)$ (ψ_6)	6.4 ± 0.4	8.0 ± 0.9	8.2 ± 0.7	7.3 ± 1.1	8.1	8.18	
	$^3J(H_N,C_\alpha)$ (ϕ_6, ψ_5)	–	–	–	0.4 ± 0.2	–	–	

dependent free energy difference ΔG between the two conformers as well as its enthalpic and entropic contributions. Using the values of $P_\beta/P_{PP_{II}} = e^{-\Delta G/k_B T}$ extracted from Table 4, we obtain free energy differences $\Delta G = 1.90$, 1.45, 0.94, and 0.72 kcal/mol for $T = 275$, 300, 325, and 350 K, respectively. Assuming that the differences in enthalpy ΔH and entropy ΔS are temperature-independent (an assumption that is difficult to prove but commonly made for temperatures far from the transition point), the free energy difference $\Delta G = \Delta H - T\Delta S = 1.5$ kcal/mol between the β and the PP_{II} conformations at 300 K consists of $\Delta H = 6$ kcal/mol and $-T\Delta S = -4.5$ kcal/mol. As expected, this finding indicates that the PP_{II} state is favored over the β state due to a lower enthalpy, although the latter is stabilized by entropy.

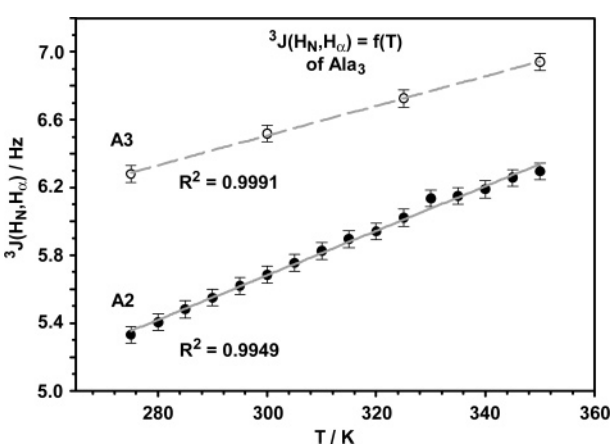

Figure 4. Temperature dependence of the $^3J(H_N,H_\alpha)$ coupling constants of Ala_3, showing the measured values for residue A2 (filled circles) and A3 (open circles), fitted with linear regression (gray line) are plotted versus temperature. Error bars are 0.05 Hz.

Since only the β and the PP_{II} conformation of Ala_3 are populated over the studied temperature range (i.e., $P_\beta + P_{PP_{II}} = 1$), we may analyze the observed temperature dependence of the *J*-coupling constants finding in terms of a two-state model. Using $P_\beta/P_{PP_{II}} = e^{-\Delta G/k_B T}$, we find for the thermal population of the β conformation $P_\beta = 1/(e^{\Delta G/k_B T} + 1)$. This yields the temperature-dependent *J*-coupling constant

$$J(T) = P_\beta J_\beta + P_{PP_{II}} J_{PP_{II}} = \frac{1}{e^{\Delta G/k_B T} + 1}(J_\beta - J_{PP_{II}}) + J_{PP_{II}} \tag{3}$$

where J_β and $J_{PP_{II}}$ are the (presumably temperature independent) *J*-coupling constants pertaining to the β and the PP_{II} conformations, respectively. Linearizing $J(T)$ in the limit $(T - T_0)/T_0 \ll 1$, we obtain

$$J(T) \approx J(T_0) + (J_\beta - J_{PP_{II}})K\frac{(T - T_0)}{T_0} \tag{4}$$

where

$$K = \left(\frac{\Delta G_0}{k_B T_0} - \frac{\Delta G_0'}{k_B} \cdot \frac{e^{\Delta G_0/k_B T_0}}{(e^{\Delta G_0/k_B T_0} + 1)^2}\right)$$

Table 4. Temperature Dependence of the *J*-Coupling Constants for Residue A2 of Ala_3

temperature/K	275		300			325		350	
J-coupling constants type (angle)	Exp	Fit	Exp	MD	Fit	Exp	Fit	Exp	Fit
$^3J(H_N,H_\alpha)$ (ϕ_2)	5.33	5.4	5.68	7.0 ± 2.6	5.6	6.02	6.1	6.29	6.4
$^3J(H_N,C')$ (ϕ_2)	1.08	1.1	1.13	1.1 ± 1.0	1.1	1.11	1.0	1.02	1.0
$^3J(H_\alpha,C')$ (ϕ_2)	1.84	1.5	1.84	2.1 ± 1.0	1.5	2.04	1.7	2.23	1.8
$^3J(H_N,C_\beta)$ (ϕ_2)	2.28	2.2	2.39	1.5 ± 0.9	2.1	2.22	1.9	2.13	1.8
$^1J(N,C_\alpha)$ (ψ_2)	11.43	10.9	11.34	10.8 ± 0.8	10.9	11.29	10.9	11.21	10.9
$^2J(N,C_\alpha)$ (ψ_2)	8.45	8.5	8.45	8.3 ± 0.6	8.5	8.42	8.5	8.34	8.5
$^3J(H_N,C_\alpha)$ (ϕ_2, ψ_1)	0.73	0.6	0.70	0.7 ± 0.1	0.6	0.75	0.7	0.80	0.7

	275 K	300 K		325 K	350 K
populations/%	Fit	MD	Fit	Fit	Fit
P_α	0	15	0	0	1
P_β	3	40	8	19	26
$P_{PP_{II}}$	97	41	92	81	73

Table 5. *J*-Coupling Constants from Experiment (Exp) and Fitting (Fit) Together with the Populations Obtained from Fitting for the HEWL Peptides

		HEWL-9mer			HEWL-19mer		
		J-coupling constants/Hz		populations/%	*J*-coupling constants/Hz		populations/%
res.	*J*-coupling type (angle)	Exp	Fit	P_a, P_β, $P_{PP_{II}}$	Exp	Fit	P_a, P_β, $P_{PP_{II}}$
	$^3J(H_N,H_\alpha)$ (ϕ_9)	5.44	5.5		5.18	5.3	
	$^3J(H_N,C')$ (ϕ_9)	1.32	1.2		1.39	1.3	
	$^3J(H_\alpha,C')$ (ϕ_9)	1.78	1.5		2.06	1.5	
A9	$^3J(H_N,C_\beta)$ (ϕ_9)	2.19	2.1	0, 10, 90 (Fit)	2.26	2.1	34, 7, 59 (Fit)
	$^1J(N,C_\alpha)$ (ψ_9)	10.80	10.7		10.54	10.4	
	$^2J(N,C_\alpha)$ (ψ_9)	7.72	7.6		7.24	7.4	
	$^3J(H_N,C_\alpha)$ (ϕ_9, ψ_8)	–	–		–	–	
	$^3J(H_N,H_\alpha)$ (ϕ_{10})	5.48	5.5		5.10	5.2	
	$^3J(H_N,C')$ (ϕ_{10})	1.29	1.2		1.33	1.3	
	$^3J(H_\alpha,C')$ (ϕ_{10})	1.88	1.6		1.72	1.5	
A10	$^3J(H_N,C_\beta)$ (ϕ_{10})	2.15	2.1	5, 12, 83 (Fit)	2.19	2.2	40, 4, 56 (Fit)
	$^1J(N,C_\alpha)$ (ψ_{10})	10.79	10.7		10.58	10.3	
	$^2J(N,C_\alpha)$ (ψ_{10})	7.46	7.6		7.02	7.4	
	$^3J(H_N,C_\alpha)$ (ϕ_{10}, ψ_9)	0.48	0.3		0.46	0.3	
	$^3J(H_N,H_\alpha)$ (ϕ_{11})	5.70	5.8		5.67	5.8	
	$^3J(H_N,C')$ (ϕ_{11})	1.10	1.1		1.09	1.2	
	$^3J(H_\alpha,C')$ (ϕ_{11})	1.98	1.6		2.20	1.6	
A11	$^3J(H_N,C_\beta)$ (ϕ_{11})	2.15	2.0	0, 17, 83 (Fit)	2.21	2.0	27, 18, 55 (Fit)
	$^1J(N,C_\alpha)$ (ψ_{11})	10.80	10.7		10.57	10.4	
	$^2J(N,C_\alpha)$ (ψ_{11})	7.61	7.6		7.17	7.4	
	$^3J(H_N,C_\alpha)$ (ϕ_{11}, ψ_{10})	0.49	0.3		0.43	0.3	

and

$$\Delta G_0 = \Delta G(T_0)$$

$$\Delta G_0' = \left.\frac{\partial \Delta G}{\partial T}\right|_{T_0}$$

The approximation predicts a slope of $J(T)$ of ~0.011 Hz/K, which is in good agreement with the experimental result of 0.013 Hz/K.

Comparison with the Ala_3 Sequence in HEWL-Peptides. Short homopolymeric peptides are often used as models for the intrinsic properties of a segment of amino acids within longer peptide sequences. We therefore were interested to compare the studied alanine peptides with the natural Ala_3 sequence in lysozyme. The S-methylated full-length protein at pH 2 is well studied and shows the features of an unfolded protein.[65−67] Thus, we decided to study the Ala_3 sequence in two peptides derived from the natural protein sequence. The shorter peptide, HEWL-9mer, comprises the Ala_3 sequence with three added amino acids, and the longer peptide, HEWL-19mer, has added eight amino acids at each side. The measured *J*-coupling constants for the alanine residues show a clear shift toward α_R helical conformations already for HEWL-9mer, which becomes even more pronounced for HEWL-19mer (cf. Table 5).

Since the MD simulation of longer peptides is cumbersome (if converged conformational sampling is required), we employed the following simple procedure in order to estimate the population of the conformational states from the measured *J*-coupling constants. The method is based on the observation that for all residues of peptides Ala_4 to Ala_7—except for the terminal residues—the calculated state-specific *J*-coupling constants are quite similar, see Tables 3 and S3−S9. For example, the $^3J(H_N,H_\alpha)$ coupling constant of Ala_7 in the PP_{II} state is 5.0 ± 0.2 Hz, if one averages over the values of the five inner residues. This finding suggests performing the fitting of the

(65) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. *Science* **2002**, *295*, 1719−1722.

(66) Wirmer, J.; Schlörb, C.; Klein-Seetharaman, J.; Hirano, R.; Ueda, T.; Imoto, T.; Schwalbe, H. *Angew. Chem., Int. Ed.* **2004**, *43*, 5780−5785.

(67) Schlörb, C.; Ackermann, K.; Richter, C.; Wirmer, J.; Schwalbe, H. *J. Biomol. NMR* **2005**, *33*, 95−104.

Table 6. Averaged *J*-Coupling Constants (in Hz) for Alanine Peptides

state	$\langle^3J(H_N,H_\alpha)\rangle$	$\langle^3J(H_N,C')\rangle$	$\langle^3J(H_\alpha,C')\rangle$	$\langle^3J(H_N,C_\beta)\rangle$	$\langle^1J(N,C_\alpha)\rangle$	$\langle^2J(N,C_\alpha)\rangle$	$\langle^3J(H_N,C_\alpha)\rangle$
α	5.2	1.4	1.5	2.0	9.7	7.4	0.5
β	9.4	0.8	2.6	0.7	10.6	7.6	0.7
PP_{II}	5.0	1.2	1.4	2.3	10.8	7.7	0.5

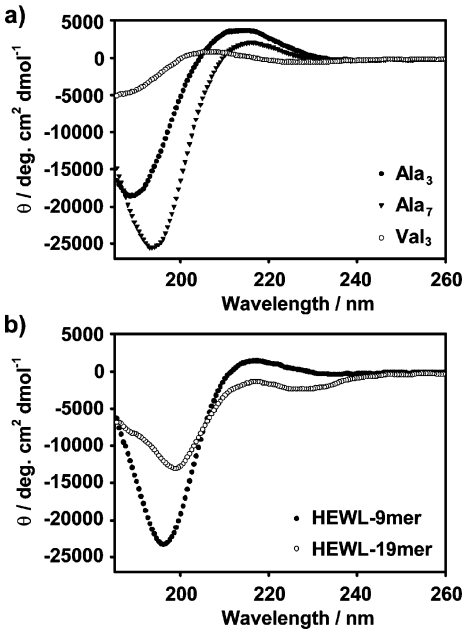

Figure 5. UV−CD spectra measured at 300 K of (a) Ala_3, Ala_7, and Val_3; (b) HEWL-9mer and HEWL-19mer.

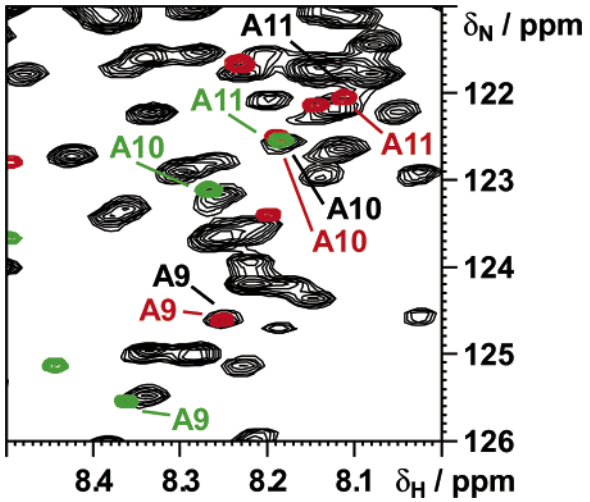

Figure 6. Overlay of ^1H, ^{15}N HSQC spectra of the HEWL-9mer (green), HEWL-19mer (red), and the full length S-methylated HEWL protein (black). Shown is the spectral region around the Ala_3 sequence. Each spectrum was recorded at 700 MHz and at 293 K. Acquisition and processing parameters and the full spectral region overlay is shown in the Supporting Information.

conformational populations (see Materials and Methods) on the basis of state-averaged calculated *J*-coupling constants, which are collected in Table 6. Proceeding this way, we find that the conformational distribution of the central three alanine residues in the 9mer is similar as for the small peptides Ala_3-Ala_7 (i.e., 80−90% PP_{II}). However, major differences are found for the 19mer, which significantly (30−40%) samples α_R helical structures (cf. Table 5).

As an independent test, we consider the UV−CD spectra of these peptides, which are given in Figure 5. The spectra of the HEWL-9mer peptide show no significant difference when compared to the spectra of Ala_3 and Ala_7 despite the value of the ψ-dependent $^2J(N,C_\alpha)$ coupling constant which is about 1 Hz lower than that for Ala_3. In the spectra of HEWL-19mer the shift toward more α_R helical conformations is seen, and one obtains a fractional α helicity of about 8% from the mean residue molar ellipticity at 222 nm for the 19mer peptide.[68] It is known from the full-length S-methylated HEWL protein that the same residues show a high induced helicity,[66] and published[69] $^3J(H_N,H_\alpha)$ coupling constants for these residues are also almost identical to the ones from the HEWL-19mer. In the overlay of the ^1H, ^{15}N HSQC spectra of the two HEWL peptides and the full-length S-methylated HEWL protein one sees that the chemical shifts of the alanine residues of the HEWL-19mer are identical with those from the full-length protein while the chemical shifts for the HEWL-9mer are different (cf. Figure 6). This questions strongly the model character of short homopolymeric peptides for longer peptide sequences. Instead, it seems to be sufficient to have added about eight amino acids at each side to resemble a full-length protein. Interestingly, the same size was found for the persistence length[70,71] in a random coil model for the analysis of ^{15}N relaxation rates of different HEWL mutants.[65]

Conclusions

We have studied the conformational distribution of short alanine peptides in aqueous solution, employing a joint NMR/MD strategy. For each peptide under consideration, we have measured five ϕ-dependent and three ψ-dependent *J*-coupling constants, respectively, and performed all-atom MD simulations including explicit solvent. We have assumed that the MD force field gives a reasonable description of the structure of the main conformations α, β, and PP_{II}, while the corresponding thermal population probabilities P_α, P_β, and $P_{PP_{II}}$ are only predicted with high degrees of uncertainty. In order to obtain accurate results for the thermal populations, we therefore have performed a global fit of P_α, P_β, and $P_{PP_{II}}$ by minimizing the deviation between measured and calculated NMR parameters. Taking into account the uncertainties of the parametrizations of the Karplus relations and of the MD force field, we estimate an overall error of about 5% for the population probabilities. This represents an unprecedented accuracy, which may serve as a benchmark for testing other experimental and computational approaches.

It has been found that the alanine peptides Ala_n ($n = 3-7$) exhibit virtually the same conformational distribution. They mainly populate the PP_{II} (∼90%) and the β (∼10%) conformations, while the α_R helical conformation is not sampled at all. This finding is in qualitative agreement with the results of

(68) Rohl, C. A.; Baldwin, R. L. *Biochemistry* **1997**, *36*, 8435−8442.
(69) Grimshaw, S. B. Ph.D. Thesis, University of Oxford, 1999.
(70) Pappu, R. V.; Srinivasan, R.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12565−12570.
(71) Möglich, A.; Joder, K.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12394−12399.

Kallenbach and Hamm and their co-workers.[1,2,14,15] The comparison of the tripeptides Ala_3, Val_3, and Gly_3 reveals the expected side-chain dependent variation of the (ϕ,ψ) sampling. Considering the temperature dependence of the *J*-coupling constants for the Ala_3, experiment and theory reveals a linear behavior of $J(T)$ for $T = 275, ..., 350$ K, which reflects the increase of the β population at the expense of the PP_{II} population. The free energy difference $\Delta G = 1.5$ kcal/mol between the two conformers at 300 K is caused by the enthalpic contribution $\Delta H \approx 6$ kcal/mol and the entropic contribution $-T\Delta S \approx -4.5$ kcal/mol, indicating that the PP_{II} state is favored over the β state due to a lower enthalpy, although the latter is stabilized by entropy.

Our MD results for short alanine peptides also facilitate the estimation of the conformational distribution of larger alanine-containing peptides. This is because for all residues of peptides Ala_4 to Ala_7 the calculated state-specific *J*-coupling constants are found to be quite similar. Hence, the thermal populations of the conformational states of alanine residues can be obtained on the basis of the averaged calculated *J*-coupling constants collected in Table 6. Applied to a 9mer and a 19mer peptide of HEWL protein, the approach reveals significant differences in the conformational distribution of the Ala_3 segment of these peptides: Only the HEWL-19mer samples a significant amount (30−40%) of α_R helical conformation, while the HEWL-9mer is predominately (80−90%) found in PP_{II} conformation. This finding seriously questions the model character of tripeptides to predict the conformational distribution of longer peptide sequences. Instead, it seems to be sufficient to have about eight amino acids at each side to resemble a full-length protein in agreement with earlier studies of persistence length in unfolded proteins.

As explained in the Introduction, it is extremely difficult to obtain accurate thermal population probabilities from a MD force field calculation (even more so from *ab initio* calculations), since a typical accuracy $\Delta\Delta G$ of the relative free energy of ∼1−2 kcal/mol already introduces an uncertainty $\exp(\Delta\Delta G/k_BT)$ of a factor of 10. Indeed, we have found that the GROMOS96 force field clearly overestimates the population of the α_R helical state of the alanine peptides. In the worst case of Ala_7, the MD simulation predicts about 50% α_R instead of ≤5 %, corresponding to a deviation of the relative free energy of about 1 kcal/mol. As shown in recent comparison studies of various force field,[8,12,13] this overestimation of α_R helical states is quite common and also occurs (often even more) for other popular force field models including AMBER94,[58] CHARMM22,[59] OPLS-AA,[60] as well as in QM/MM calculations.[12] A notable exception is the model AMBER94/MOD put forward by Garcia and co-workers.[10] By modifying (actually removing) the backbone dihedral angle terms of the AMBER94 force field, they obtained conformational distributions for various alanine peptides which appear to be quite similar to our NMR/MD results.

While it is obvious that modifications of the backbone dihedral angle force field terms will change the resulting (ϕ,ψ) probability distribution, it is still unclear whether these terms are the only physical origin of the deviation between theory and experiment. Using the GROMOS96 force field, for example, we have obtained essentially the correct conformational distribution for Val_3, although—by construction—the backbone dihedral angle terms are identical for Val_3 and Ala_3. The joint NMR/MD strategy proposed in this work is capable of providing benchmark data for improving force fields that can well be extended to other side chains.

Acknowledgment. We dedicate this work to Prof. Jürgen Bereiter-Hahn on the occasion of his 65th birthday. We are grateful to Sarah Mensch for help with peptide synthesis and purification. We thank Christian Richter for assistance with NMR measurements and Christian Schlörb for providing the ^1H, ^{15}N HSQC spectra of the full length S-methylated hen egg white lysozyme. Financial support by the state of Hesse (Center for Biomolecular Magnetic Resonance), the Deutsche Forschungsgemeinschaft (SFB 579), the Center for Scientific Computing of Frankfurt, and the EU-project UPMAN is gratefully acknowledged.

Supporting Information Available: Complete ref 59; Table S1 listing the synthesized peptides with their isotopic labeling pattern; Table S2 with the Karplus parametrizations used in this work; Tables S3−S13 with the *J*-coupling constants obtained from NMR and MD together with populations from MD and fitting; Table S14 and Figure S1 showing the results for the test of the robustness of the fitting of the populations; Tables S15−S49 with acquisition and processing parameters of the NMR experiments; Tables S50−S63 with the assigned chemical shifts of the peptides; Figures S2−S5 showing ^1H,^{15}N HSQC and ^1H 1D spectra of the studied peptides; Figure S6: the full spectral region overlay of the ^1H,^{15}N HSQC spectra of HEWL-9mer, HEWL-19mer, and the full length S-methylated HEWL protein; Figure S7: the ^1H,^1H NOESY spectrum of Gly_3. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0660406