Biopolymer Conformational Distributions from Solid-State NMR: α-Helix and 3_{10}-Helix Contents of a Helical Peptide

Henry W. Long‡ and Robert Tycko*

Contribution from the Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received December 17, 1997

Abstract: A new approach to the quantitative experimental analysis of conformational distributions in partially ordered biopolymers is described and used to determine the fractional contents of α-helical, 3_{10}-helical, and random coil components in the conformational distributions of a 17-residue helix-forming peptide [Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci., U.S.A.* **1987**, *84*, 8898−8902]. The approach is based on solid-state NMR measurements that probe ϕ and ψ dihedral angles at specific isotopically labeled sites, in particular two-dimensional magic-angle spinning NMR exchange spectra, carried out on glass-forming frozen solutions in glycerol/water. The data analysis employs a combination of constrained molecular dynamics simulations, to generate model dihedral angle distributions, and Bayesian statistics, to permit a quantitative determination of the relative probabilities of the possible fractional contents of α-helical, 3_{10}-helical, and random coil components. The peptide is found to be highly α-helical in pure glycerol/water at low temperatures. Addition of urea has the primary effect of converting α-helical conformations to 3_{10}-helical conformations, rather than to random coil conformations. These results provide experimental evidence for the role of 3_{10}-helices as thermodynamic intermediates in the folding and unfolding of α-helical segments in peptides and proteins.

Introduction

The formation of helical structures by polypeptides is a subject of enduring interest[1,2] because of the importance of helices as structural elements of proteins. Recent investigations[3−8] have focused on the factors that influence the stability of helical structures in relatively short synthetic peptides, following the discovery that peptides consisting of roughly 15−20 amino acids can be designed to have high helix contents in aqueous solutions.[3,4] A considerable number of designed helical peptides have been developed with the explicit or implicit assumption that they form α-helices, that is, helices in which hydrogen bonds connect the carbonyl oxygen of residue *i* to the amide hydrogen of residue *i*+4.[3−8] However, on the basis of evidence from electron spin resonance (ESR) measurements on doubly spin-labeled helical peptides,[9−12] Millhauser and co-workers have proposed that short, polyalanine-based, helical peptides may in fact exist in solution partly as 3_{10}-helices, that is, helices in which the hydrogen bonds connect residue *i* to residue *i*+3. This proposal has proven to be controversial. For example, ESR measurements by Smythe et al. on a 17-residue, polyalanine-based peptide indicate an α-helical conformation.[13] Recent liquid-state nuclear magnetic resonance (NMR) and hydrogen exchange measurements on a 16-residue peptide by Millhauser et al. provide evidence for a distribution between α-helical and 3_{10}-helical conformations, with the 3_{10}-helical conformations more highly populated at the C- and N-termini of the peptide.[14] In general, definitive experiments have been difficult to carry out because of the lack of well-established physical measurements that distinguish reliably between α-helical and 3_{10}-helical conformations and because of the inherent difficulty of making quantitative structural measurements on systems where a distribution of conformations is likely to be present. On the theoretical side, Sheinerman and Brooks have developed a modified Zimm−Bragg model for the helix−coil transition which allows for the two distinct types of helices.[15] For reasonable choices of the initiation and propagation parameters in their model, Sheinerman and Brooks find that 3_{10}-helical hydrogen bonds may predominate in short peptides. This behavior is expected if the entropic cost of initiating a 3_{10}-helix is lower than that of initiating an α-helix, even though the α-helix may be the energetically preferred conformation. The

* To whom correspondence should be addressed. Phone: 301-402-8272. Fax: 301-496-0825. E-mail: tycko@helix.nih.gov.

‡ Current address: Curagen Corporation, New Haven, CT 06511.

(1) Chakrabartty, A.; Baldwin, R. L. *Adv. Protein Chem.* **1995**, *46*, 141−176.

(2) Kallenbach, N. R.; Lyu, P.; Zhou, H. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996; pp 201−259.

(3) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci., U.S.A.* **1987**, *84*, 8898−8902.

(4) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci., U.S.A.*, **1989**, *86*, 5286−5290.

(5) Padmanabhan, S.; Marqusee, S.; Ridgeway, T.; Laue, T. M.; Baldwin, R. L. *Nature* **1990**, *344*, 268−270.

(6) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646−651.

(7) Lyu, P. C.; Liff, M. I.; Marky, L. A.; Kallenbach, N. R. *Science* **1990**, *250*, 669−673.

(8) Bradley, E. K.; Thomason, J. F.; Cohen, F. E.; Kosen, P. A.; Kuntz, I. D. *J. Mol. Biol.* **1990**, *215*, 607−622.

(9) Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, *359*, 653−655.

(10) Fiori, W. R.; Miick, S. M.; Millhauser, G. L. *Biochemistry* **1993**, *32*, 11957−11962.

(11) Fiori, W. R.; Lundberg, K. M.; Millhauser, G. L. *Nat. Struct. Biol.* **1994**, *1*, 374−377.

(12) Hanson, P.; Martinez, G.; Millhauser, G.; Formaggio, F.; Crisma, M.; Toniolo, C.; Vita, C. *J. Am. Chem. Soc.* **1996**, *118*, 271−272.

(13) Smythe, M. L.; Nakaie, C. R.; Marshall, G. R. *J. Am. Chem. Soc.* **1995**, *117*, 10555−10562.

(14) Millhauser, G. L.; Stenland, C. J.; Hanson, P.; Bolin, K. A.; van de Ven, F. J. M. *J. Mol. Biol.* **1997**, *267*, 963−974.

(15) Sheinerman, F. B.; Brooks, C. L. *J. Am. Chem. Soc.* **1995**, *117*, 10098−10103.

S0002-7863(97)04277-7
Published on Web 07/03/1998

ratio of 3_{10}-helical to α-helical hydrogen bonds increases as the overall helix content is reduced, lending support to the idea that 3_{10}-helical structures may be thermodynamic intermediates in the folding and unfolding of α-helical structures.[16−18] In addition, a number of molecular dynamics (MD) simulations have identified 3_{10}-like conformations as kinetic and thermodynamic intermediates in the unfolding of α-helical peptides.[19−23] MD simulations on polyalanine-based, helical peptides under conditions that favor a fully folded structure have found α-helical conformations to be preferred over 3_{10}-helical conformations,[24] although 3_{10}-helical conformations may be preferred in polyaminoisobutyric acid-based peptides under certain conditions.[25,26]

The purpose of this paper is 2-fold. First, we present experimental measurements that address the question of whether 3_{10}-helical conformations are significantly populated in a designed helical peptide developed previously by Marqusee and Baldwin.[3] Second, in the process of addressing this question, we demonstrate a new approach to the quantitative experimental investigation of conformational distributions in biopolymers that are structurally disordered. This approach employs two-dimensional (2D) solid-state NMR measurements, in particular 2D magic-angle spinning (MAS) NMR exchange spectroscopy, which we have recently introduced for studies of peptide backbone conformations;[27−30] constrained MD simulations to generate model conformational distributions; and a quantitative, probabilistic analysis of the solid-state NMR data based on Bayes' theorem.[31,32] Since central problems in biophysical chemistry and structural biology frequently involve partially disordered biopolymers (e.g., protein folding intermediates) and since the quantitative characterization of conformational distributions in disordered biopolymers by other structural techniques is difficult, the approach demonstrated in this paper may be of general interest and utility.

Experimental and Computational Methods

MB(i+4)EK peptides (see Results section below) were synthesized using standard FMOC solid-phase methods on a Perkin-Elmer/Applied Biosystems model 433A peptide synthesizer. ^{13}C- and ^{15}N-labeled FMOC-alanine was obtained from Cambridge Isotopes. MBHA Rink Amide Resin (Nova Biochem) was used to create C-terminal amides. Acetic anhydride in dimethylformamide was used to acetylate the N-termini. Peptides were deprotected and cleaved from the resin with 95% TFA/5% H_2O. Purification was accomplished with preparative scale HPLC, using a reverse-phase C8 column and a 5−35% CH_3CN/H_2O gradient, with 0.1% TFA. Peptide purity (>95%) and amino acid sequence were verified by FAB mass spectrometry.

Stock solutions of unlabeled MB(i+4)EK in pure H_2O, at a concentration of 1 mM, were diluted into 1:1 (vol/vol) H_2O/glycerol, with or without urea, to a final concentration of 30 μM, pH ≈ 4.5. Circular dichroism (CD) spectra were measured on a Jasco J-720 spectropolarimeter, calibrated with (+)-10-camphorsulfonic acid, using 1.0 and 0.1 cm cuvettes. Low-temperature measurements were made by placing the cuvettes in a home-built, foam-insulated enclosure with optical windows which was cooled by a temperature-regulated N_2(g) flow and placed in the optical path.

MB(i+4)EK samples for solid-state NMR measurements were 6−8 mM, pH ≈ 3.5 solutions in 1:1 (vol/vol) H_2O/glycerol, with or without ^{13}C-depleted urea (Cambridge Isotopes). Peptide solutions were not buffered because of reduced peptide solubility in pH 7 buffer and because of uncertainties regarding the effects of buffering agents in H_2O/glycerol at low temperatures. Also, 15 mM $CuSO_4$ was added to the peptide solutions to reduce the low-temperature proton spin−lattice relaxation times from approximately 15 s to approximately 0.7 s, thereby speeding up data collection by a factor of about 20. NMR measurements were carried out at 9.39 T (100.4 MHz ^{13}C NMR frequency) using a Chemagnetics CMX-400 spectrometer, a Chemagnetics magic-angle spinning (MAS) probe with 6 mm rotors (240 μL sample volume), a home-built spinning speed controller, and a home-built temperature control system. MAS rotors were filled with the peptide solution, immersed in N_2(l) to freeze the sample, and then transferred rapidly (several seconds) into the precooled MAS probe at −140 °C. Measurements on lyophilized powders were carried out at room temperature.

2D MAS NMR exchange measurements were carried out as previously described.[27,28] The pulse sequence for these measurements is CP−t_1−90_ϑ−τ−90_x−t_2, where CP represents standard Hartmann−Hahn cross-polarization with 50 kHz radio frequency (rf) fields, and proton decoupling with 80 kHz rf fields is applied during t_1 and t_2. Four separate 2D data sets are acquired, with the phase ϑ of the first 90° pulse set to either x or y and the pulses synchronized with the sample rotation such that either the exchange period τ or $\tau + t_1$ is a multiple of the rotation period. Additional phase cycling is used to reduce spectral artifacts. These data sets are combined[28,30] to give a single, phase-sensitive 2D spectrum in which cross-peaks between spinning sideband lines are present only if an exchange process operates. Exchange of nuclear spin polarization between the two labeled carbonyl ^{13}C or amide ^{15}N sites, driven by nuclear magnetic dipole−dipole couplings, results in *intersite* cross-peaks. In the limit of complete exchange, which is reached at the values of $\tau \approx 500$ ms and $\tau \approx 8$ s employed in our ^{13}C and ^{15}N measurements, respectively, the intersite cross-peak intensities are determined by the relative orientations and principal values of the chemical shift anisotropy (CSA) tensors of the labeled sites and are independent of the internuclear distance. Exchange of carbonyl ^{13}C frequencies due to spin−lattice relaxation of directly bonded ^{14}N nuclei results in *intrasite* cross-peaks.[27,28] The intensity of the intrasite cross-peaks depends on the ^{14}N relaxation time. Measurements of cross-peak intensities on a singly labeled (carbonyl of Ala-9) MB(i+4)EK sample, where only intrasite exchange can occur, indicate a relaxation time for ^{14}N of Ala-10 of approximately 0.6 s in frozen solution at −140 °C. In the fully exchanged limit, the calculated intensities of the largest intrasite cross-peaks are less than 20% of the calculated intensities of the largest intersite cross-peaks.

Typically, sweep widths of 50 kHz in t_2 and 25 kHz in t_1 were used for ^{13}C measurements. Complete 2D MAS NMR exchange data sets with between 32 and 96 t_1 points were acquired in 1−7 days, depending on the line widths, using a recycle delay of 1 s. Block averaging and ramped cross-polarization field amplitudes were used in the longer runs to minimize effects of instabilities in rf pulse amplitudes and phases. Data were processed with the NMRPipe software routines.[48] Cross-peak intensities for spinning sidebands from −2 to +2 in each dimension (where 0 is the isotropic line) were measured by integrating the 2D spectra over rectangular areas corresponding to the full widths at half-maximum of the cross-peaks. Root-mean-squared noise levels

(16) Sundaralingam, M.; Sekharudu, Y. C. *Science* **1989**, *244*, 1333−1337.
(17) Sung, S.-S. *Biophys. J.* **1995**, *68*, 826−834.
(18) Millhauser, G. L. *Biochemistry* **1995**, *34*, 3873−3877.
(19) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864−3871.
(20) Soman, K. V.; Karimi, A.; Case, D. A. *Biopolymers* **1991**, *31*, 1351−1361.
(21) Tobias, D. J.; Brooks, C. L. *Biochemistry* **1991**, *30*, 6059−6070.
(22) Hirst, J. D.; Brooks, C. L. *Biochemistry* **1995**, *34*, 7614−7621.
(23) Young, W. S.; Brooks, C. L. *J. Mol. Biol.* **1996**, *259*, 560−572.
(24) Tirado-Rives, J.; Maxwell, D. S.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1993**, *115*, 11590−11593.
(25) Zhang, L.; Hermans, J. *J. Am. Chem. Soc.* **1994**, *116*, 11915−11921.
(26) Smythe, M. L.; Huston, S. E.; Marshall, G. R. *J. Am. Chem. Soc.* **1995**, *117*, 5445−5452.
(27) Weliky, D. P.; Tycko, R. *J. Am. Chem. Soc.* **1996**, *118*, 8487−8488.
(28) Tycko, R.; Weliky, D. P.; Berger, A. E. *J. Chem. Phys.* **1996**, *105*, 7915−7930.
(29) de Jong, A. F.; Kentgens, A. P. M.; Veeman, W. S. *Chem. Phys. Lett.* **1984**, *109*, 337−342.
(30) Luz, Z.; Spiess, H. W.; Titman, J. J. *Isr. J. Chem.* **1992**, *32*, 145−160.
(31) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C. The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: New York, 1992; pp 818−826.
(32) Press, S. J. *Bayesian Statistics: Principles, Models, and Applications*; Wiley: New York, 1989.

σ were determined by integrating the spectra over 50−100 equal areas devoid of NMR signals.

2D MAS NMR exchange cross-peak intensities were calculated numerically as functions of the dihedral angles ϕ and ψ using specialized software (available on request) as previously described.[28] The peptide bond geometry and carbonyl ^{13}C CSA tensor orientation described in ref 29 were used in these calculations. For ^{15}N calculations, the δ_{11} and δ_{33} principal axes of the amide nitrogen CSA tensor are taken to lie in the peptide plane, with δ_{33} at an angle of 100° to the N−CO bond. Because the NMR signals of the two labeled sites, for example, Ala-8 and Ala-9, in our MB(i+4)EK samples were unresolved, calculated intersite cross-peak intensities from Ala-8 → Ala-9 exchange were added to cross-peak intensities from Ala-9 → Ala-8 exchange before comparison with experimental cross-peak intensities. In the ^{13}C case, calculated intrasite cross-peak intensities, which are independent of ϕ and ψ, were added to the intersite cross-peak intensities after multiplication by a fitting parameter between 0 and 1 to account for the uncertain extent of ^{14}N spin−lattice relaxation. Fits to experimental data represented in Figures 5 and 7 were carried out by specialized software (available on request). All cross-peaks among sidebands numbered from −2 to +2 were included in these fits, but diagonal peaks were excluded. Intensities of pairs of cross-peaks that were symmetrically disposed about the diagonal of the experimental 2D MAS NMR exchange spectra, which are theoretically equal, were added together before comparison with calculated cross-peak intensities.

MD simulations were carried out using the CHARMm routines in the QUANTA 96 modeling program (Molecular Simulations). Distributions of ϕ and ψ angles for α- and 3_{10}-helices were extracted from MD simulations on MB(i+4)EK in which van der Waals, electrostatic, bond, angle, torsion, and improper potential energy terms were active, in addition to artificial potentials that constrained the distances between the carbonyl oxygen of residue i and the amide hydrogen of residue i+4 to be less than 2.0 Å (for α-helices) or the distances between the carbonyl oxygen of residue i and the amide hydrogen of residue i+3 to be less than 2.1 Å (for 3_{10}-helices). Excursions outside the distance constraints were limited by harmonic restoring forces with force constants of 5×10^3 kcal/mol·Å^2, which value was determined empirically. The distance constraints are in accordance with known structures of helical peptides and proteins. Planarity of peptide bonds was maintained by restoring torques about all N−CO bonds with a force constant of 100 kcal/mol·rad^2, which were active when the ω dihedral angles deviated by more than 2° from 180°. Simulations were carried out in a vacuum, with a distance-dependent dieletric constant. To mimic low pH conditions, the side chains of glutamic acid residues were neutralized by adjustments of the atomic charges. Distributions of ϕ and ψ angles for random coil conformations are from MD simulations on the tripeptide Ac-AAA-NH$_2$, in which only van der Waals, bond, angle, and dihedral potential energy terms were active, in addition to the ω dihedral constraints. The Verlet algorithm with time steps of 1 fs was used in all simulations. Coordinates were stored after 50 fs intervals. Simulations of α- and 3_{10}-helices were run at 200 K for 200 ps, after energy minimization, heating, and equilibration for 20 ps. Simulations for random coil conformations were run at 400 K for 400 ps, after energy minimization, heating, and equilibration for 40 ps.

Results

Description of the System. Experiments were carried out on the 17-residue, polyalanine-based peptide Ac-AEAAA-KEAAAKEAAAKA-NH$_2$ (acetylated at the N-terminus and amidated at the C-terminus) designed by Marqusee and Baldwin[3] and called (i+4)EK by them. We refer to this peptide as MB(i+4)EK to acknowledge its designers. Marqusee and Baldwin developed this peptide with the idea that the oppositely charged side chains of the lysine and glutamate residues would interact electrostatically in such a way as to stabilize an α-helical structure. The peptide does in fact exhibit a high helix content in aqueous solution, as shown by CD.[3] However, the side chain electrostatic interactions are not essential for helix formation, as demonstrated by the fact that the helix content remains high at low pH, where the glutamatic acid side chains are not ionized, or in solutions of high ionic strength, where electrostatic interactions are screened, and by the observation of similarly high helix contents in polyalanine-based peptides with only glutamate or only lysine residues.[3,4] A quantitative study of glutamate−lysine interactions in model peptides by Scholtz et al.[33] indicates a helix-stabilizing effect per pair of residues of approximately 0.3 kcal/mol at pH = 2 and 0.5 kcal/mol at pH = 7. The charged side chains of MB(i+4)EK certainly do increase the solubility of the peptide in aqueous solutions, a factor that is critical in our experiments.

Solid-state NMR measurements on MB(i+4)EK were carried out in frozen solutions, using 1:1 (vol/vol) glycerol/H$_2$O solvent mixtures. These solutions form homogeneous glasses (i.e., amorphous solid solutions) rather than crystalline or phase-separated solids at low temperatures, with glass transition temperatures $T_g \approx -60$ °C. A rigid solid environment is a prerequisite for the solid-state NMR techniques described below, because large-amplitude molecular motions on time scales <1 s would otherwise prevent the extraction of structural information from the solid-state NMR data. The use of a glass-forming solvent permits such a rigid environment while preventing perturbations of the peptide structure that may be associated with crystallization of the solvent. The packing and interactions of solvent molecules in a glass-forming solvent at temperatures below T_g do not differ significantly from those at temperatures immediately above T_g. Only the time scale for molecular motions differs, becoming much longer than any experimentally relevant time scale at temperatures below T_g. We therefore expect the peptide conformations and conformational distributions probed by our solid-state NMR measurements, performed at temperatures well below T_g, to reflect the conformations and conformational distributions present in the liquid phase at an effective temperature T_{eff} that is slightly above T_g. The precise value of T_{eff}, which is roughly the temperature at which the solution falls out of thermodynamic equilibrium as the temperature is decreased, depends on the experimental sample cooling rate. For the relatively low cooling rates employed in our experiments (estimated to be 20 deg/s), we expect $T_{eff} - T_g$ to be less than 10 °C.

The 2D MAS NMR exchange measurements described below probe the dihedral angles ϕ and ψ that specify the peptide backbone conformation at a particular residue between two isotopically labeled sites. ^{13}C NMR measurements on MB(i+4)EK samples with ^{13}C labels at the carbonyl carbons of Ala-8 and Ala-9 probe the ϕ and ψ angles of Ala-9. ^{13}C NMR measurements on samples with ^{13}C labels at the carbonyl carbons of Ala-13 and Ala-14 probe the ϕ and ψ angles of Ala-14. These two samples were chosen to allow comparisons between the conformations at the center and the conformations at the C-terminus of the peptide, motivated by the theoretical prediction[15] and observations in simulations[17] and in protein crystal structures[34] that 3_{10}-helical conformations concentrate toward the C-termini of helical segments. Ala-14 is the most C-terminal residue of which the carbonyl oxygen can participate in both α-helical and 3_{10}-helical hydrogen bonds. 2D ^{15}N NMR measurements on samples with ^{15}N labels at the amide nitrogens of Ala-9 and Ala-10 also probe the ϕ and ψ angles of Ala-9.

CD Measurements. The temperature dependence of CD

(33) Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 9668−9676.

(34) Richardson, J. S. In *Advances in Protein Chemistry*; Anfinsen, C. B., Edsall, J. T., Richards, F. M., Eds.; Academic Press: New York, 1981; pp 167−339.

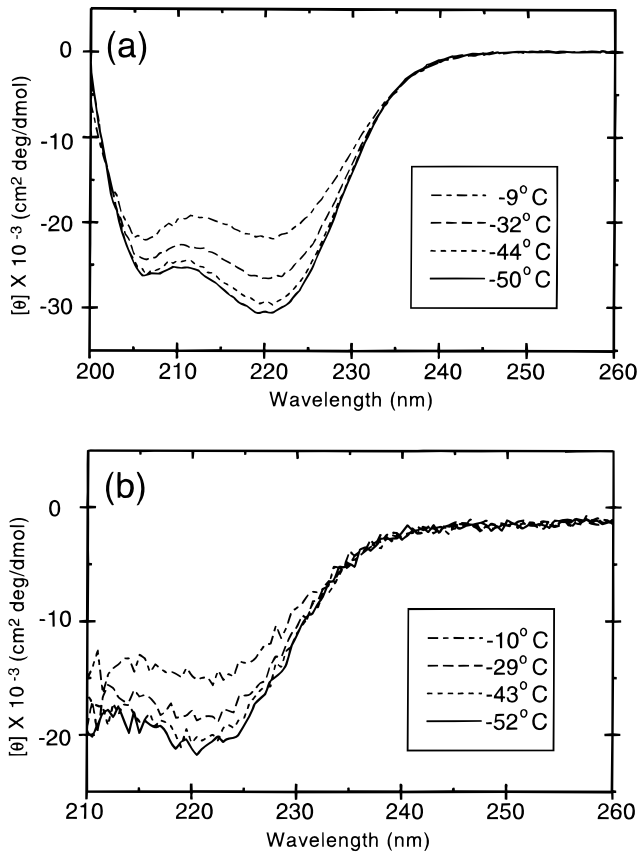

Figure 1. Circular dichroism spectra of the peptide MB(i+4)EK in glycerol/H_2O solution at the indicated temperatures, without (a) and with (b) 5.1 M urea.

spectra of MB(i+4)EK in glycerol/H_2O at pH = 4.5 is shown in Figure 1a. The helix content increases with decreasing temperature, as previously observed in purely aqueous solutions. Taking a molar ellipticity value at 222 nm ($[\theta]_{222}$) of -34×10^3 deg·cm^2/dmol to represent 100% helix content and 600 deg·cm^2/dmol to represent 100% random coil,[33] the helix content at the lowest temperature appears to be 90%. We observe very similar CD spectra for MB(i+4)EK in H_2O and in glycerol/H_2O at 1 °C (less than 10% difference in $[\theta]_{222}$), indicating that glycerol does not strongly stabilize or destabilize helical structures in this system.

Figure 1b shows CD spectra of MB(i+4)EK in glycerol/H_2O with 5.1 M urea. Addition of urea clearly affects the CD spectra, leading to an apparent helix content of 63% at the lowest temperature.

Solid-State NMR Spectroscopy. Solid-state NMR measurements were carried out in a 9.39 T field on MB(i+4)EK samples with isotopic labels at specific sites as described above. Figures 2a−2c show one-dimensional (1D) ^{13}C NMR spectra of samples with carbonyl ^{13}C labels at Ala-8 and Ala-9 in lyophilized form (2a), in glycerol/H_2O at −140 °C (2b), and in glycerol/H_2O with 5.1 M ^{13}C-depleted urea at −140 °C (2c). These spectra were acquired with MAS at a 2.5 kHz spinning rate. As a result, the carbonyl signals appear as series of spinning sideband lines,[35] separated by 2.5 kHz (25 ppm) intervals. The carbonyl line widths in Figures 2a and 2b are 2.6 and 2.3 ppm, respectively, while the carbonyl line widths in Figure 2c are 6.1 ppm. These line widths are due to inhomogeneous broadening from the dependence of the isotropic chemical shifts δ_{iso} on the peptide conformation,[36] the packing of solvent molecules around the peptide, and intra- and intermolecular hydrogen bonding.[37,38] The line widths in Figures 2a and 2b are comparable to the smallest values we have observed for peptides in frozen solutions (D. P. Weliky and R. Tycko, unpublished results), including well-structured systems such as melittin (1.9 ppm carbonyl line width for Ala-4) and a peptide derived from the V3 loop of the HIV-1 envelope glycoprotein gp120 bound to the Fab fragment of an anti-gp120 antibody (1.5 ppm carbonyl line widths for residues in the epitope region). Thus, the 1D MAS spectra suggest that MB(i+4)EK is well-structured both in frozen glycerol/H_2O, in agreement with the CD results, and in lyophilized form. Carbonyl line widths of 2.0 ppm are also

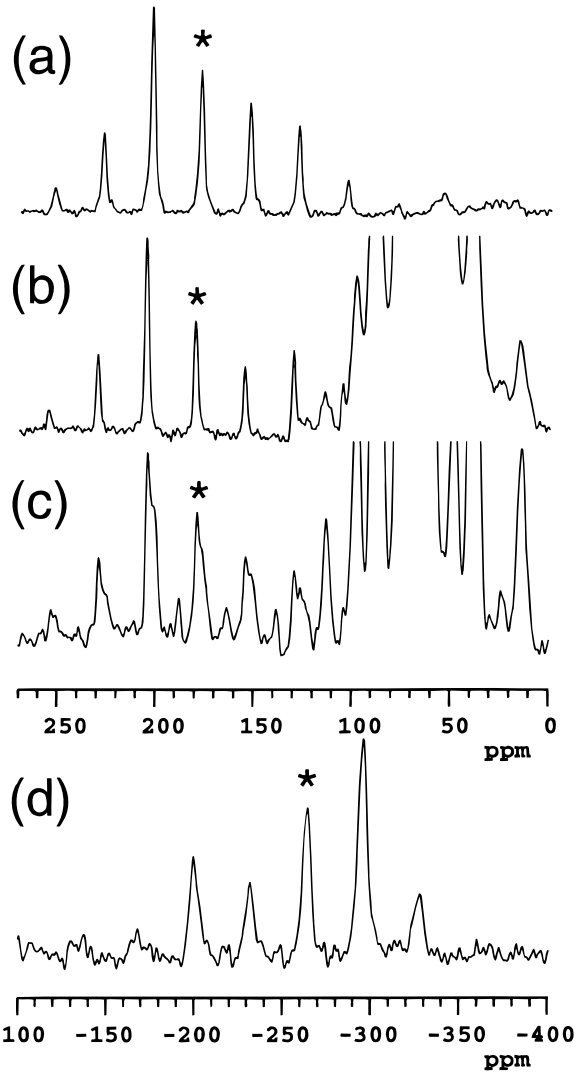

Figure 2. One-dimensional solid-state NMR spectra of specifically labeled MB(i+4)EK samples at 9.39 T, obtained with magic-angle spinning. (a) ^{13}C NMR of lyophilized powder at 21 °C, with carbonyl ^{13}C labels at Ala-8 and Ala-9. (b) ^{13}C NMR of frozen glycerol/H_2O solution at −140 °C, with carbonyl ^{13}C labels at Ala-8 and Ala-9. (c) Same conditions as (b), but with 5.1 M ^{13}C-depleted urea. (d) ^{15}N NMR of lyophilized powder at 21 °C, with amide ^{15}N labels at Ala-9 and Ala-10. Asterisks indicate the isotropic, or centerband, NMR lines. Series of equally spaced lines on either side are spinning sidebands. Signals between 0 and 120 ppm in ^{13}C spectra are primarily from glycerol.

(35) Herzfeld, J.; Berger, A. E. *J. Chem. Phys.* **1980**, *73*, 6021−6030.
(36) Jiao, D.; Barfield, M.; Hruby, V. J. *J. Am. Chem. Soc.* **1993**, *115*, 10883−10887.
(37) Asakawa, N. et al. *J. Am. Chem. Soc.* **1992**, *114*, 3261−3265.
(38) de Dios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1994**, *116*, 11485−11488.

observed in a frozen glycerol/H_2O solution of an MB(i+4)EK sample with a carbonyl ^{13}C label at Ala-9 alone, indicating that isotropic shift differences between the Ala-8 and Ala-9 carbonyls do not contribute greatly to the apparent line widths. The increased line width from the addition of urea in Figure 2c indicates qualitatively an increase in structural disorder.

We determine principal values δ_{11}, δ_{22}, and δ_{33} of the carbonyl CSA tensors from the relative intensities of the spinning sideband lines.[35] Sideband intensities in all frozen solution spectra, including spectra of samples with and without urea and samples with labels at Ala-8 and Ala-9, Ala-13 and Ala-14, and Ala-9 alone, are well fit by values of the anisotropy $\delta \equiv \delta_{11} - \delta_{33}$ and asymmetry parameter $\eta \equiv (\delta_{11} - \delta_{22})/(\delta_{iso} - \delta_{33})$ of 153 ppm and 0.374, respectively. Sideband intensities in all spectra of lyophilized ^{13}C-labeled MB(i+4)EK samples, including spectra of samples with labels at Ala-8 and Ala-9 and samples with labels at Ala-13 and Ala-14, are well fit by $\delta =$ 154 ppm and $\eta = 0.581$. These carbonyl ^{13}C CSA principal values are obtained with fitting programs that take into account contributions to the observed sideband intensities from the magnetic dipole−dipole coupling to directly bonded ^{14}N nuclei.[28]

Figure 2d shows a ^{15}N MAS NMR spectrum of lyophilized MB(i+4)EK, with ^{15}N labels at the amide nitrogens of Ala-9 and Ala-10, at a spinning rate of 1.30 kHz. The spinning sideband intensities are well fit by $\delta = 156$ ppm and $\eta = 0.039$.

Figure 3 shows examples of 2D MAS NMR exchange spectra of doubly labeled MB(i+4)EK samples. We have developed 2D MAS NMR exchange spectroscopy[27−30] as a means of making local structural measurements on biopolymers in unoriented, noncrystalline solids such as frozen solutions, lyophilized powders, and unoriented fibers. Structural information in 2D MAS NMR exchange spectra is contained in the relative intensities of intersite cross-peaks that connect the spinning sideband lines of the two different labeled sites. These cross-peaks arise from exchange of nuclear spin polarization between the labeled sites during the delay period τ that separates the t_1 and t_2 periods of the 2D NMR measurement. In the fully exchanged limit, spinning sideband cross-peak intensities are determined primarily by the relative orientations of CSA tensors associated with two different labeled functional groups in a molecule, that is, by the relative orientations of the two labeled functional groups (e.g., carbonyl or amide groups). The 2D MAS NMR exchange technique therefore provides information that is directly angular in nature, in contrast to dipolar recoupling techniques such as rotational resonance, REDOR, and DRAMA which provide information primarily in the form of internuclear distances.[39,40] The cross-peaks of interest are clearly seen in Figure 3. Figure 4 shows 1D slices from 2D MAS NMR exchange spectra, to illustrate the experimental signal-to-noise ratios and the effects of urea on the line widths.

In the case of a peptide with labels at two successive carbonyl or amide sites, the spinning sideband cross-peak intensities can be simulated for any assumed values of the intervening ϕ and ψ angles by evaluating exact mathematical expressions[28] that depend only on the known sample spinning rate, the known external magnetic field strength, the two CSA tensors (whose orientations are known from studies of model compounds[41−43] and whose principal values can be determined experimentally as described above), and the values of other bond angles in the peptide backbone (which are known from crystallographic studies of peptides and proteins). When a distribution of ϕ and ψ angles at a single residue is present, the 2D MAS NMR exchange spectrum is simply the sum of spectra for the individual ϕ,ψ pairs in the distribution weighted by the populations of each ϕ,ψ pair. Thus, we can readily simulate the 2D MAS NMR exchange spectrum for any conformational distribution.

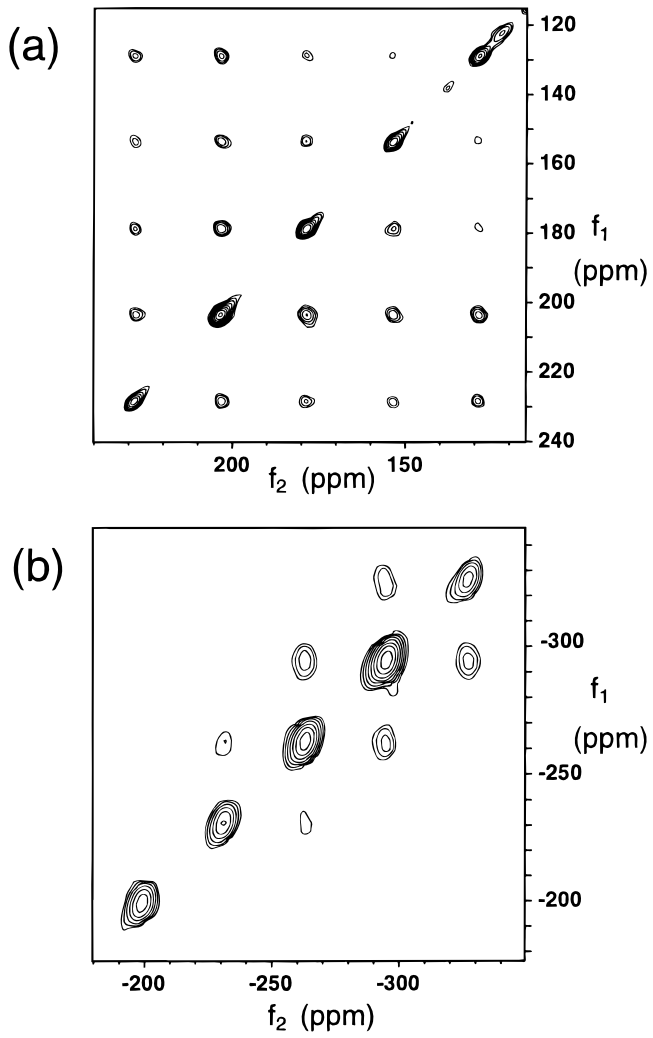

Figure 3. Contour plots of two-dimensional magic-angle spinning (2D MAS) NMR exchange spectra, showing the spinning sideband cross-peaks that contain conformational information. (a) ^{13}C spectrum of MB(i+4)EK with carbonyl ^{13}C labels at Ala-8 and Ala-9 in frozen glycerol/H_2O solution at −140 °C. (b) ^{15}N spectrum of MB(i+4)EK with amide ^{15}N labels at Ala-9 and Ala-10 as lyophilized powder at 21 °C.

Preliminary Analyses of 2D Solid-State NMR Data. Figure 5 shows results of preliminary analyses of 2D MAS NMR exchange data that assume the existence of a single peptide backbone conformation, that is, a single ϕ,ψ pair, at the labeled site. We calculate the deviation $\chi^2 = S^2/\sigma^2$ between experimental and simulated spinning sideband cross-peak intensities as a function of the ϕ and ψ angles assumed in the simulations. Here S^2 is the sum of the squared differences between cross-peak intensities, and σ^2 is the mean squared noise per cross-peak in the experimental 2D spectra. In the case of 2D ^{13}C MAS NMR exchange data, χ^2 is minimized with respect to two fitting parameters for each choice of ϕ and ψ, namely (1) the fractional contribution of intrasite cross-peaks (see Methods

(39) Smith, S. O. *Curr. Opin. Struct. Biol.* **1993**, *3*, 755−759.
(40) Bennett, A. E.; Griffin, R. G.; Vega, S. *NMR* **1994**, *33*, 1−77.
(41) Oas, T. G.; Hartzell, C. J.; McMahon, T. J.; Droby, G. P.; Dahlquist, F. W. *J. Am. Chem. Soc.* **1987**, *109*, 5956−5962.
(42) Oas, T. G.; Hartzell, C. J.; Dahlquist, F. W.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5962−5966.
(43) Teng, Q.; Iqbal, M.; Cross, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 5312−5321.

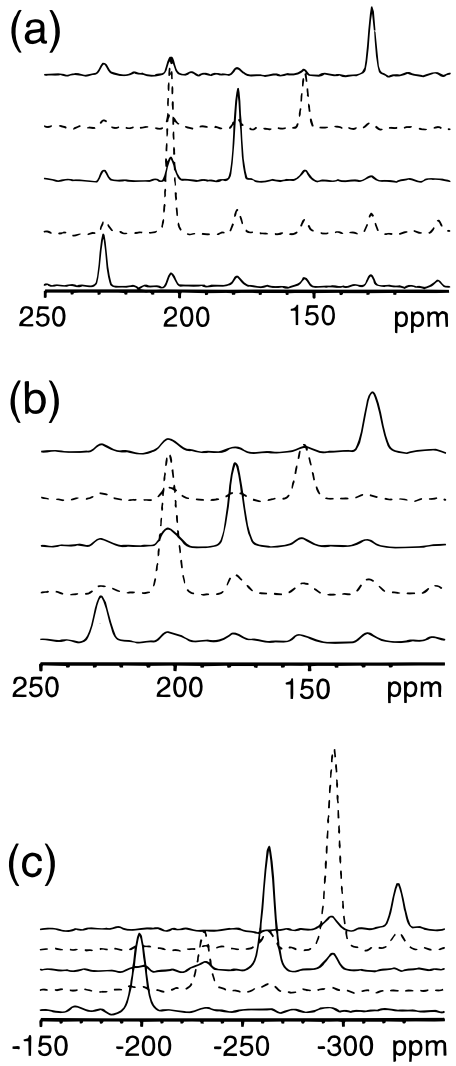

Figure 4. One-dimensional slices of 2D MAS NMR exchange spectra of MB(i+4)EK, taken through the centerband frequency of the isotopic label signal in the f_1 dimension and through the first two sidebands on either side. Solid and dashed lines alternate for clarity. (a) ^{13}C spectrum of sample with carbonyl ^{13}C labels at Ala-8 and Ala-9 in frozen glycerol/H_2O solution at −140 °C. (b) Same conditions as (a), but with 5.1 M urea. (c) ^{15}N spectrum of sample with amide ^{15}N labels at Ala-9 and Ala-10, as lyophilized powder at 21 °C.

section) to the experimental intensities and (2) the overall scaling of the simulated cross-peak intensities. In the case of 2D ^{15}N MAS NMR exchange data, χ^2 is minimized with respect to the overall scaling of the simulated cross-peak intensities for each choice of ϕ and ψ. The minimized χ^2 is plotted as a function of ϕ and ψ for MB(i+4)EK labeled at the carbonyl carbons of Ala-8 and Ala-9 in frozen glycerol/H_2O without (Figure 5a) and with (Figure 5b) urea and in lyophilized form (Figure 5c), as well as for MB(i+4)EK labeled at the amide nitrogens of Ala-9 and Ala-10 in lyophilized form (Figure 5d). Only half of the possible ϕ,ψ space is shown, with $-180° \leq \phi \leq 0°$, because the 2D MAS NMR exchange spectra are invariant to the substitution $\phi,\psi \rightarrow -\phi,-\psi$.[28]

The global minima in χ^2 in Figures 5a and 5c lie in the helical region of ϕ,ψ space, as expected. Local minimum in χ^2 also occur in a region centered around $\phi \approx -110°$, $\psi \approx 100°$. These local minima are a consequence of the symmetry properties of CSA tensors, as discussed in detail in ref 28, which can lead to similar but not identical spinning sideband cross-peak intensities in distinct regions of ϕ,ψ space. Because of differences between the ^{15}N and ^{13}C CSA tensors, Figure 5d does not show a local minimum in χ^2 near $\phi \approx -110°$, $\psi \approx 100°$, while the global minimum region in Figure 5d does include the helical region of ϕ,ψ space. Thus, the 2D MAS NMR exchange data are consistent with a predominance of helical conformations for MB(i+4)EK, both in frozen glycerol/H_2O and in lyophilized form.

Standard models of α-helices correspond to backbone dihedral angles $\phi \approx -60°$ and $\psi \approx -40°$. Standard models of 3_{10}-helices correspond to backbone dihedral angles $\phi \approx -50°$, $\psi \approx -30°$, although there is considerable variation among different models. In Figures 5a, 5c, and 5d, χ^2 is significantly lower at this choice of α-helical ϕ,ψ values than at this choice of 3_{10}-helical values. Results very similar to those of Figures 5a and 5c are obtained when the same analysis is applied to 2D MAS NMR exchange data on MB(i+4)EK samples labeled at the carbonyl carbons of Ala-13 and Ala-14. Thus, the preliminary analyses in Figure 5 suggest that α-helical conformations are more highly populated than 3_{10}-helical conformations, in both frozen glycerol/H_2O and lyophilized samples and at both Ala-9 and Ala-14. However, these analyses are not definitive for at least three reasons. First, the α-helix and 3_{10}-helix are defined in practice by hydrogen-bonding patterns, not by ϕ and ψ values. In a solution (whether frozen or not) or in a lyophilized sample, we expect each type of helix to encompass a distribution of ϕ and ψ values because of the inherent flexibility of the peptide and the absence of crystalline order. No single choice of ϕ and ψ can be an adequate description. Second, in addition to the two helical conformations under consideration, it is possible that a significant fraction of the peptide molecules exhibit nonhelical, random-coil-like conformations at the labeled positions. Third, we would like to make quantitative statements about the fractional populations of the two types of helices and the ability of our measurements to distinguish between them, as well as about the fractional population of random coil conformations. In particular, we would like to determine the probability that any possible set of populations of α-helix, 3_{10}-helix, and random coil is the correct set, given our experimental data. A method of analysis that addresses these problems is described in the following section.

The χ^2 plot in Figure 5b for the frozen solution containing 5.1 M urea differs substantially from that in Figure 5a. The minima in χ^2 are shifted, χ^2 in the global minimum region (centered on $\phi \approx -45°$, $\psi \approx -45°$) is not significantly different from χ^2 in the local minimum regions, and the smallest value of χ^2 in Figure 5b is greater than the smallest value of χ^2 in Figure 5a by a factor of 3 (i.e., the quality of the fit is lower). These differences are attributable to the disordering effects of urea. Since our CD measurements (Figure 1b) and 1D ^{13}C NMR spectra (Figure 2c) indicate that MB(i+4)EK adopts more than a single conformation in the presence of 5.1 M urea, an analysis that assumes a single pair of ϕ and ψ values is most certainly inadequate.

Quantitative Analyses of Conformational Distributions. We adopt a model for the conformational distributions in which the peptide conformation at each labeled site can be in one of three classes. A fraction $f(\alpha)$ of the peptides are in an α-helical conformation at the labeled site, a fraction $f(3_{10})$ are in a 3_{10}-helical conformation, and a fraction $f(c) = 1 - f(\alpha) - f(3_{10})$ are in a random coil conformation. We derive separate distributions of ϕ and ψ values for each of the three classes of conformations by carrying out constrained molecular dynamics simulations. The resulting distributions are shown in Figure 6. The ϕ,ψ distributions for the helical conformations (Figures 6a−

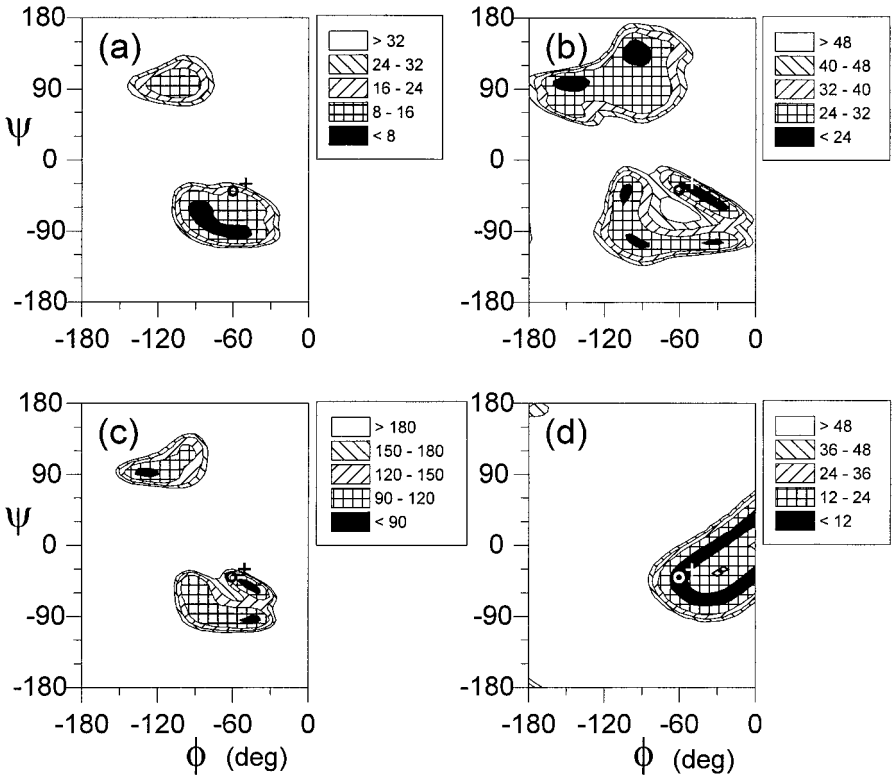

Figure 5. Preliminary analyses of 2D MAS NMR exchange data, assuming a single preferred conformation for MB(i+4)EK. Deviations χ^2 (defined in the text) between experimental and calculated spinning sideband cross-peak intensities as functions of the dihedral angles ϕ and ψ assumed in the calculations are represented as contour plots. Circles and crosses on each plot indicate ϕ,ψ pairs that correspond roughly to standard α-helical and 3_{10}-helical conformations, respectively. (a) Analysis of ^{13}C spectrum of MB(i+4)EK with carbonyl ^{13}C labels at Ala-8 and Ala-9 in frozen glycerol/H_2O solution at −140 °C. (b) Same conditions as (a), but with 5.1 M urea. (c) Analysis of ^{13}C spectrum of sample with carbonyl ^{13}C labels at Ala-8 and Ala-9, as lyophilized powder at 21 °C. (d) Analysis of ^{15}N spectrum of sample with amide ^{15}N labels at Ala-9 and Ala-10, as lyophilized powder at 21 °C.

6d) are in good agreement with reported distributions for helical segments in protein crystal structures.[44,45]

Simulated intersite spinning sideband cross-peak intensities for 2D MAS NMR exchange spectra of α-helical, 3_{10}-helical, and random coil conformational distributions are obtained by binning the corresponding ϕ,ψ distributions in Figure 6 into 5° increments in both ϕ and ψ, generating a table of simulated cross-peak intensities with 5° increments in both ϕ and ψ, and finally coadding the simulated cross-peak intensities with weights given by the binned conformational distributions. The resulting sets of cross-peak intensities then form a three-component basis which we use to fit the experimental data.

Our quantitative fitting procedure is based on Bayes' theorem, a fundamental theorem of probability theory.[31,32] To our knowledge, analyses based on Bayes' theorem have not been employed previously in structural studies by solid-state NMR. In general, for a model with N fitting parameters $p_1, p_2, ..., p_N$, each of which has a discrete set of allowed values, Bayes' theorem states that the probability $P(p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}|D;I)$ that the correct values of these parameters are in fact $p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}$, given a set of experimental data D and any other prior information I, can be evaluated from the expression

$$P(p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}|D;I) = \frac{1}{Q}P(D|p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N};I)P(p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}|I) \quad (1)$$

where $P(D|p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N};I)$ is the probability that the experiments would yield data D given I and, assuming that $p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}$ were in fact the correct values of the fitting parameters, $P(p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}|I)$ is the probability that $p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}$ are the correct values given prior information I but no experimental data, and Q is a normalization constant determined from

$$Q = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_N}P(p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}|D;I) \quad (2)$$

where the summations are over all possible values of each fitting parameter.

In our model for conformational distributions of MB(i+4)-EK, the fitting parameters are $f(\alpha)$, $f(3_{10})$, a parameter ξ that represents the fractional contribution of the intrasite cross-peaks to the observed cross-peaks relative to their maximum possible contribution (see Materials and Methods section), and a parameter ζ that represents the overall scaling of the simulated cross-peak intensities. We take $P(f(\alpha), f(3_{10}), \xi, \zeta|I)$ to be a constant, meaning that all combinations of values of the fitting parameters are assumed to be equally likely in the absence of the experimental data, and we take $P(D|f(\alpha), f(3_{10}), \xi, \zeta; I)$ to be proportional to $\exp[-{}^1/_2\chi^2(f(\alpha), f(3_{10}), \xi, \zeta)]$, where $\chi^2(f(\alpha), f(3_{10}), \xi, \zeta)$ is the total squared deviation between experimental and calculated 2D MAS NMR exchange cross-peak intensities for the given values of the parameters divided by the mean squared noise σ^2 in the experimental 2D spectrum. The assumption that $P(D|f(\alpha), f(3_{10}), \xi, \zeta;I)$ is proportional to $\exp[-{}^1/_2\chi^2(f(\alpha), f(3_{10}), \xi, \zeta)]$ is justified if the noise is Gaussian and if errors in the experimental cross-peak intensities due to random noise exceed systematic errors due to pulse sequence

(44) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601−619.
(45) Karplus, P. A. *Protein Sci.* **1996**, *5*, 1406−1420.

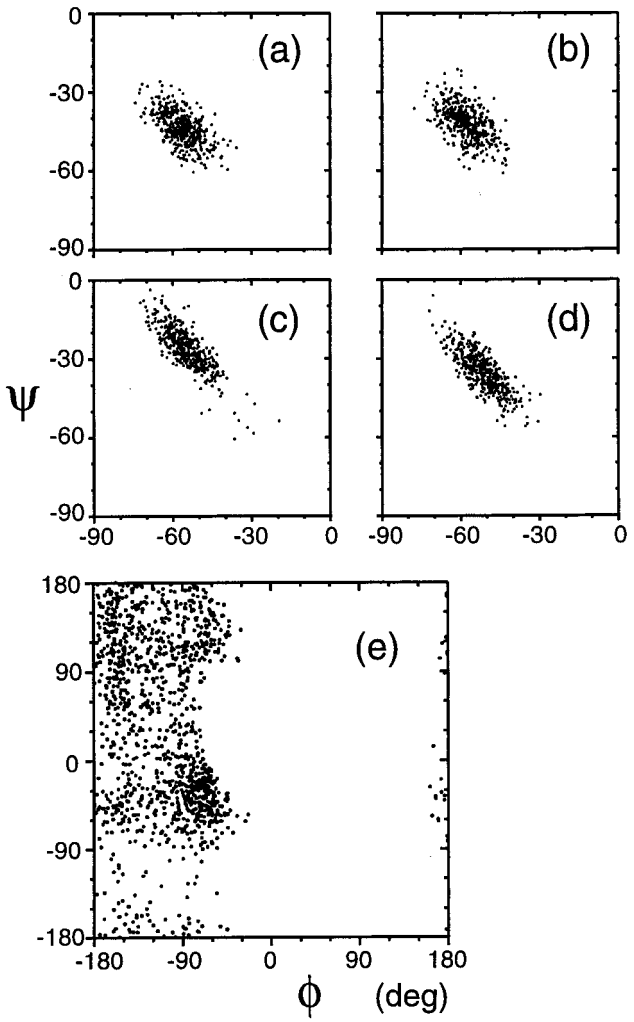

Figure 6. Distributions of ϕ and ψ dihedral angles from constrained molecular dynamics simulations, represented as scatter plots. The dots indicate the ϕ,ψ values for a specific residue after successive 500 fs simulation intervals. These distributions are used to construct calculated 2D MAS NMR exchange spectra representative of α-helical, 3_{10}-helical, and random coil conformations, which are used in the quantitative data analyses in Figure 7. (a) Ala-9 in MB(i+4)EK in α-helix. (b) Ala-14 in MB(i+4)EK in α-helix. (c) Ala-9 in MB(i+r)EK in 3_{10}-helix. (d) Ala-14 in MB(i+4)EK in 3_{10}-helix. (e) Ala-2 in Ac-AAA-NH_2 in random coil.

imperfections, imperfections in processing of the 2D NMR spectra, errors in the CSA tensors or peptide geometry assumed in the simulations, or other sources. Our experience with model compounds[27,28] and the high degree of symmetry observed in the experimental 2D MAS exchange spectra of frozen solutions of MB(i+4)EK indicate that systematic errors do not contribute significantly to χ^2.

Following eq 1, we evaluate $P(f(\alpha), f(3_{10}), \xi, \zeta|D;I)$ numerically over a grid of values of the fitting parameters, with $f(\alpha)$ and $f(3_{10})$ varying between 0.0 and 1.0 in increments of 0.05 (subject to the condition $f(c) \geq 0$), ξ varying between 0.0 and 1.0 in increments of 0.01, and ζ varying between $0.5\zeta_{min}$ and $1.5\zeta_{min}$ in increments of $0.01\zeta_{min}$, where ζ_{min} is the value of ζ that minimizes χ^2. The quantity of ultimate interest in the context of our three-component model for the conformational distributions is the probability $P(f(\alpha), f(3_{10}))$ that particular values of $f(\alpha)$ and $f(3_{10})$ are correct, regardless of the values of ξ and ζ. This is given by

$$P(f(\alpha), f(3_{10})) = \sum_i \sum_j P(f(\alpha), f(3_{10}), \xi_i, \zeta_j|D;I)$$

$$= \frac{1}{Q'} \sum_i \sum_j \exp\left[-\frac{\chi^2(f(\alpha), f(3_{10}), \xi_i, \zeta_j)}{2\sigma^2}\right] \quad (3)$$

where the summations are over all values of ξ and ζ on the grid and Q' is a normalization constant.

The results of numerical evaluations of eq 3 are shown in Figure 7 for MB(i+4)EK samples labeled at the carbonyl carbons of Ala-8 and Ala-9 (Figures 7a and 7b) and the carbonyl carbons of Ala-13 and Ala-14 (Figures 7c and 7d) in frozen glycerol/H_2O solutions. $P(f(\alpha), f(3_{10}))$ is represented as a contour plot of the probability density over the triangular region of the $f(\alpha)$,$f(3_{10})$ plane that includes the allowed values of $f(\alpha)$ and $f(3_{10})$. The value of $f(c)$ at a point in this plane is given by $1 - f(\alpha) - f(3_{10})$. In particular, points on the diagonal edge of the triangular region represent $f(c) = 0$, that is, conformational distributions comprised entirely of α-helical and 3_{10}-helical conformations.

In the absence of urea (Figures 7a and 7c), the most probable conformational distributions occur very close to $f(\alpha) = 1.0$, indicating that the peptide conformation at Ala-9 and Ala-14 is predominantly α-helical. The total probability that the correct conformational distribution contains less than 20% 3_{10}-helical or random coil conformations, obtained by summing all values of $P(f(\alpha), f(3_{10}))$ with $f(\alpha)$ greater than or equal to 0.80, is 0.74 for Ala-9 and 0.79 for Ala-14. This result demonstrates that our measurements and analysis are capable of distinguishing α-helical from 3_{10}-helical conformations, despite the relatively small differences in ϕ and ψ angles.

With the addition of 5.1 M urea (Figures 7b and 7d), the conformational distributions change dramatically. The most probable distributions now include substantial fractions of both α-helical and 3_{10}-helical conformations. Interestingly, the most probable distributions lie close to the diagonal edge of the triangular $P(f(\alpha), f(3_{10}))$ plots, that is, they contain relatively small fractions of random coil conformations. The total probability that the correct conformational distribution contains less than 30% random coil conformations, obtained by summing all values of $P(f(\alpha), f(3_{10}))$ with $f(c)$ less than 0.30, is 0.74 for Ala-9 and 0.81 for Ala-14. The total probability that the correct conformational distribution contains a greater fraction of 3_{10}-helical conformations than random coil conformations, obtained by summing all values of $P(f(\alpha), f(3_{10}))$ with $f(3_{10})$ greater than $f(c)$, is 0.71 for Ala-9 and 0.69 for Ala-14.

Table 1 summarizes the results in Figure 7 in terms of probability-weighted average values, defined for example by $\langle f(\alpha)\rangle = \sum_i \sum_j f(\alpha)_i P(f(\alpha)_i, f(3_{10})_j)$, and probability-weighted uncertainties, defined for example by $\delta f(\alpha) = [\sum_i \sum_j (f(\alpha)_i - \langle f(\alpha)\rangle)^2 P(f(\alpha)_i, f(3_{10})_j)]^{1/2}$ for the fractions of α-helical, 3_{10}-helical, and random coil conformations in each sample. The summations in these expressions are over all values of $f(\alpha)$ and $f(3_{10})$ on the grid. Table 1 also includes the results of analyses of 2D MAS NMR exchange measurements on lyophilized MB-(i+4)EK samples.

Discussion

3_{10}-Helices as Folding Intermediates. That frozen solutions of MB(i+4)EK containing 5.1 M urea are likely to contain substantial fractions of 3_{10}-helical conformations but smaller fractions of random coil conformations is the most significant result of this study. At both Ala-9 and Ala-14, the most

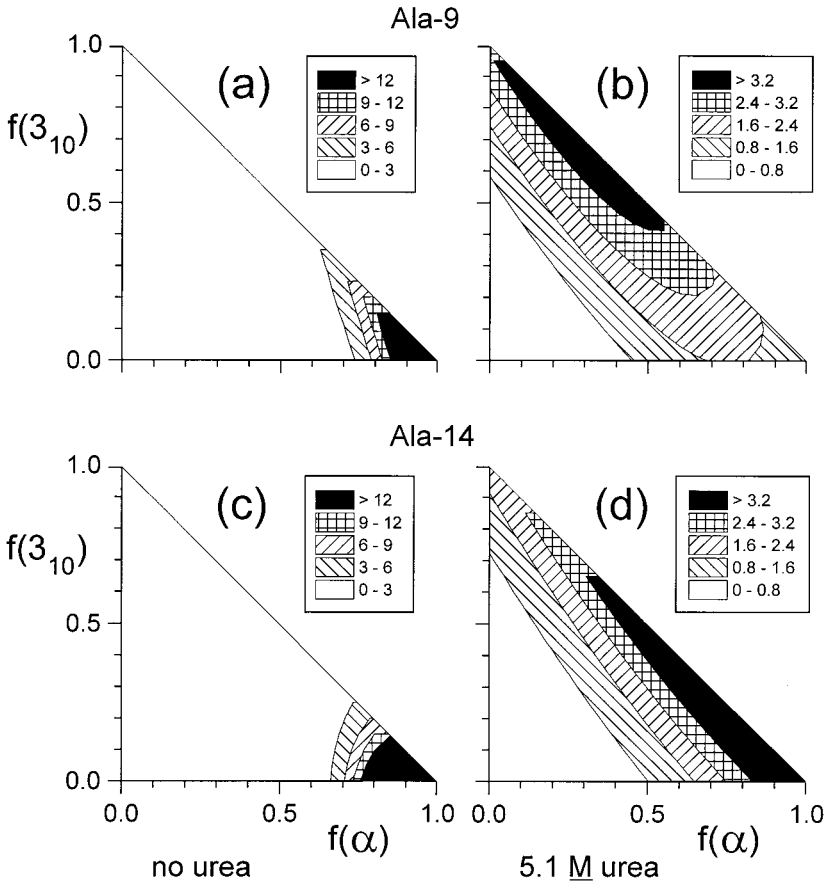

Figure 7. Quantitative analyses of 2D MAS NMR exchange data. Probability densities that the conformational distribution of MB(i+4)EK contains a fraction $f(\alpha)$ of α-helical conformations, a fraction $f(3_{10})$ of 3_{10}-helical conformations, and a fraction $f(c) = 1 - f(\alpha) - f(3_{10})$ of random coil conformations are represented as contour plots. These probability densities result from Bayesian analyses in which the data are fit to a three-component model for the conformational distributions, based on the ϕ,ψ distributions for each component shown in Figure 6. (a) Probability density for Ala-9 in MB(i+4)EK in frozen glycerol/H_2O at −140 °C. (b) Same conditions as (a), but with 5.1 M urea. (c) Probability density for Ala-14 in MB(i+4)EK in frozen glycerol/H_2O at −140 °C. (d) Same conditions as (c), but with 5.1 M urea.

Table 1. Probability-Weighted Average Values $\langle f \rangle$ and Uncertainties δf of the Fractional Populations of α-Helical, 3_{10}-Helical, and Random Coil Conformations at Specific Residues in the Peptide MB(i+4)EK in Frozen Solutions and in Lyophilized Form, As Determined from the Quantitative Analysis of 2D MAS NMR Exchange Spectra

MB(i+4)EK sample condition	residue	$\langle f(\alpha)\rangle$	$\delta f(\alpha)$	$\langle f(3_{10})\rangle$	$\delta f(3_{10})$	$\langle f(c)\rangle$	$\delta f(c)$
glycerol/H_2O, −140 °C	Ala-9	0.84	0.13	0.10	0.11	0.06	0.07
glycerol/H_2O, −140 °C	Ala-14	0.85	0.10	0.06	0.07	0.09	0.09
glycerol/H_2O, 5.1 M urea, −140 °C	Ala-9	0.38	0.24	0.44	0.26	0.19	0.16
glycerol/H_2O, 5.1 M urea, −140 °C	Ala-14	0.49	0.26	0.36	0.26	0.15	0.15
lyophilized powder, 24 °C	Ala-9	0.83	0.06	0.00	0.01	0.17	0.06
lyophilized powder, 24 °C	Ala-14	0.90	0.07	0.02	0.03	0.08	0.07

probable conformational distributions contain no random coil. The probability-weighted average values $\langle f(c)\rangle$ are 0.19 and 0.15, respectively. The probability-weighted average values $\langle f(3_{10})\rangle$ are 0.44 and 0.36. Thus, the primary effect of adding urea, that is, of changing from solvent conditions that favor an ordered helical conformation to conditions that favor a less-ordered conformation, is the conversion of α-helical conformations to 3_{10}-helical conformations. This finding is consistent with suggestions that 3_{10}-helical conformations are intermediates on the thermodynamic folding and unfolding pathway of α-helical peptides. Experimental evidence for substantial 3_{10}-helix contents in helical peptides at equilibrium in solution has previously come primariily from the EPR and liquid-state NMR studies of Millhauser and co-workers.[9−12,14] Theoretical evidence for 3_{10}-helical conformations as thermodynamic folding intermediates comes from the modified Zimm−Bragg model of Sheinerman and Brooks,[15] the Monte Carlo (MC) simulations of Sung,[17] the MD simulations of Brooks and co-workers,[21,23] and (possibly) the crystallographic analyses of Sundaralingam and Sekharudu.[16] Theoretical evidence for 3_{10}-helical conformations as *kinetic* intermediates comes from the MD simulations of Tirado-Rives and Jorgensen,[19] Soman et al.,[20] and Hirst and Brooks,[22] but these kinetic simulations do not necessarily imply that a measurable 3_{10}-helix content would be present at equilibrium under conditions of partial helix formation.

Sheinerman and Brooks[15] predict that 3_{10}-helical segments may exist principally at the C-terminal end of helical peptides under conditions of high helix content, as a result of the statistical weight of zero assigned to configurations in which 3_{10}-helical units occur at the N-terminal end of α-helical segments in their modified Zimm−Bragg model. Concentration of 3_{10}-helical segments toward the C-termini of helical peptides is also reported in the MC simulations of Sung[17] and the MD simulations of Young and Brooks[23] and is observed in protein crystal structures.[34] Our experimental results do not reveal a pronounced effect of this sort. $\langle f(3_{10})\rangle$ at Ala-14 is not

significantly greater than $\langle f(3_{10})\rangle$ at Ala-9 under any of the conditions studied. In fact, the conformational distributions at Ala-9 and Ala-14 are remarkably similar under all conditions studied. This result is in contrast to the evidence of Fiori et al. from EPR measurements[10] and Millhauser et al.[14] from liquid-state NMR measurements, on different helical peptides under different conditions, that the 3_{10}-helix content near the termini exceeds the 3_{10}-helix content in the middle.

As explained above, the conformational distributions observed in our solid state NMR measurements are the equilibrium distributions at approximately −60 °C. The low temperatures in these measurements may be expected to reduce the importance of entropic effects relative to experiments and simulations above 0 °C. Thus, one explanation for our observation of apparently high α-helix contents at both Ala-9 and Ala-14 in the absence of urea may be the relative unimportance of entropically driven 3_{10}-helix or random coil formation at the C-terminus.

Comparison of CD and Solid-State NMR Results. Our CD measurements on urea-containing glycerol/H_2O solutions slightly above T_g (Figure 1b) might conventionally be interpreted to mean that the helix content, averaged over all residues in the peptide, is roughly 0.63. Our 2D MAS NMR exchange measurements indicate that the total helix content, including both α-helices and 3_{10}-helices, is likely to be significantly higher. The sample conditions for the two types of measurements differ in three ways. First, the pH of the CD solutions is slightly higher. The higher pH would be expected to increase, rather than decrease, the helix content.[3] Second, the solid-state NMR measurements are carried out at lower temperatures. However, no significant changes in $[\theta]_{222}$ are observed below the lowest temperature in Figure 1b, and conformational distribution of MB(i+4)EK in frozen solutions can be no different from the distribution at $T_g \approx -60$ °C. Third, the peptide concentrations in the solid-state NMR measurements are much higher than the concentrations in the CD measurements. We cannot rule out the possibility of concentration-dependent effects at temperatures near T_g, although such effects have been shown to be absent from CD measurements in aqueous solution at 20 °C for peptide concentrations up to at least 100 μM.[3] If we exclude concentration-dependent effects for the moment, our experimental results suggest a real underestimation of the total helix content by CD measurements interpreted in a naive manner. One possible explanation for such an underestimation may be that 3_{10}-helical segments make a significantly less negative contribution to $[\theta]_{222}$ than α-helical segments, as predicted by theory[46] and observed in experimental measurements on a model peptide.[47] Another possible explanation may lie in the difficulty of quantifying helix contents from CD measurements when the helical segments are conformationally disordered and of uncertain length, since both conformational variations and length variations have strong effects on CD spectra.[46]

Extensions of the Structural Approach. The experimental measurements and methods of analysis employed in this study of the conformational distributions of MB(i+4)EK represent a new approach to structural investigations of partially ordered biopolymers. This approach yields quantitative information, including estimated uncertainties, within the context of a well-defined model for a conformational distribution. The use of solid-state NMR measurements on frozen glassy solutions of specifically labeled molecules permits the examination of a static distribution, thus avoiding the dynamic averaging and conformational exchange effects that complicate liquid-state NMR and other measurements, and the examination of structural features at specific sites. The 2D MAS NMR exchange measurements are particularly well-suited for structural studies of partially ordered biopolymers because of our ability to simulate the 2D spectra of any desired conformational distributions with minimal assumptions and very few (i.e., one or two) extraneous adjustable parameters. A single 2D MAS NMR spectrum of a single doubly labeled peptide places strong constraints on both the ϕ and the ψ dihedral angles at the labeled site.

The Bayesian analysis described above presents the information content of the measurements in a clear, quantitative, and flexible manner. Experimental data from other types of measurements, including solid-state NMR measurements of internuclear distances that can also be compared quantitatively with simulations,[49] can be combined with or substituted for the 2D MAS NMR exchange data in a straightforward way that properly accounts for differences in the uncertainties (i.e., the σ values) of the measurements. Structural information from other sources, such as MD simulations, calculated potential energy surfaces, or statistical analyses of structural databanks, can be included as well, for example, by introducing nonuniform values of $P(p_{1,i_1}, p_{2,i_2}, ..., p_{N,i_N}|I)$ in eq 1.

Biopolymers that are only partially ordered are important in a number of areas of biophysics and structural biology. Examples include protein folding intermediates, molten globule states, and fibrous peptide aggregates that occur in amyloid diseases and diseases involving prion proteins. We expect the methods described above to be of value in structural investigations of these and other systems.

JA974277J

(46) Manning, M. C.; Woody, R. W. *Biopolymers* **1991**, *31*, 569−586.
(47) Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J. *J. Am. Chem. Soc.* **1996**, *118*, 2744−2745.
(48) Delaglio, F. et al. *J. Biomol. NMR* **1995**, *6*, 277−293.
(49) Bennett, A. E.; Weliky, D. P.; Tycko, R. *J. Am. Chem. Soc.* **1998**, *120*, 4897−4898.