ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/8913295

Tripeptides with Ionizable Side Chains Adopt a Perturbed Polyproline II Structure in Water †

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 2004

Impact Factor: 3.02 · DOI: 10.1021/bi035740+ · Source: PubMed

CITATIONS
64

READS
23

5 AUTHORS, INCLUDING:

Kai Griebenow
University of Puerto Rico at Rio Piedras
140 PUBLICATIONS **5,300** CITATIONS
SEE PROFILE

Laurence A Nafie
Syracuse University
306 PUBLICATIONS **7,703** CITATIONS
SEE PROFILE

Reinhard Schweitzer-Stenner
Drexel University
219 PUBLICATIONS **4,375** CITATIONS
SEE PROFILE

Available from: Reinhard Schweitzer-Stenner
Retrieved on: 02 February 2016

Tripeptides with Ionizable Side Chains Adopt a Perturbed Polyproline II Structure in Water†

Fatma Eker,‡,§ Kai Griebenow,‡ Xiaolin Cao,|| Laurence A. Nafie,|| and Reinhard Schweitzer-Stenner*,‡,⊥

Departments of Chemistry and Biology, University of Puerto Rico, Río Piedras Campus, San Juan, Puerto Rico 00931, Department of Chemistry, Syracuse University, Syracuse, New York 13244, and Department of Chemistry, Drexel University, Philadelphia, Pennsylvania 19104

Received September 25, 2003

ABSTRACT: The present paper reports the conformations of the acidic and basic homotripeptides triglutamate, triaspartate, and trilysine in aqueous solution to better understand their relevance for the structure of disordered proteins and protein segments and for a variety of protein binding processes. The determination of the dihedral angles of the central amino acid residue was achieved by analyzing the amide I band profile of the respective polarized visible Raman, Fourier transform infrared (FT-IR), and vibrational circular dichroism (VCD) spectra by means of recently developed algorithms [Schweitzer-Stenner, R. (2002) *Biophys. J. 83*, 523−532; Eker et al. (2002) *J. Am. Chem. Soc. 124*, 523−532]. The results were validated by measuring the UV electronic circular dichroism (ECD) spectra of the peptides. The analyses revealed that a polyproline II-like conformation is predominant at room temperature. For triaspartate and triglutamate the dihedral angles of $\phi = -70°$, $\psi = 165°$ and $\phi = -60°$, $\psi = 160°$ were obtained, respectively. A similar conformation, i.e., $\phi = -50°$, $\psi = 170°$, was obtained for trilysine, which is at variance with the earlier reported left-handed turn structure. The ECD spectrum of charged tripeptides displayed symmetric negative and positive couplets at 190 and 210 nm, which are interpreted as indicating a somewhat, perturbed polyproline II conformation, in agreement with the obtained dihedral angles. Comparison with literature data shows that the investigated tripeptides are ideal model systems for understanding the local conformation of functionally relevant K_3, K_2X, E_3, and D_3 segments in a variety of different proteins.

One of the basic textbook doctrines of protein chemistry is that the amino acid sequence of a polypeptide polymer is the key determinant of the three-dimensional structure, which is a prerequisite for protein function (*1*). However, this view is at variance with the experimentally established fact that nature has produced numerous so-called disordered or unfolded proteins with very well defined functions (*2*). These proteins are now generally termed intrinsically unstructured proteins (IUPs) (*3*). They have been found to be involved in DNA/RNA−protein interaction, to function as inhibitors and scavengers, and to facilitate the formation and function of multiprotein complexes (*2*, *4*). Others mediate regulatory posttranslational modification processes, such as phosphorylation and proteolysis. The discovery of natively unfolded proteins has led Wright and Dyson to propose a reassessment of the structure−function paradigm (*4*).

The question arising from the above findings is how a protein can have a well-defined function even if it is structurally random. The term random generally implies that each amino acid residue can sample the entire allowed region of the Ramachandran space, which is thus regarded as nearly ergodic. However, this view, which is based on the insights and findings of Tanford (*5*) and Flory (*6*), has been seriously questioned over the last 15 years on the basis of multiple theoretical and experimental findings. In particular, electronic and vibrational circular dichroism (ECD[1] and VCD) spectra of so-called unfolded polypeptides and proteins have provided strong evidence for the existence of local segments of polyproline II (PPII) or 3_1-helices (*7*−*11*). This is an ordered structural motif because it exhibits a perfect 3-fold rotational symmetry for its canonical conformation with $\phi, \psi = -78°$, 146° (*12*). Its most prominent spectroscopic fingerprint is an asymmetric couplet in the UV−ECD spectrum with a strong minimum between 190 and 210 nm and a weaker maximum between 210 and 225 nm (*7*, *9*, *10*). It was observed in the spectrum of the unfolded state of many proteins subjected to denaturing detergents. Thermal denaturation, however, often yields a really disordered state with multiple coexisting extended structures when studied by

† Financial support was provided from the NIH-COBRE II grant for the Center for Research in Protein Structure, Function, and Dynamics (P20 RR16439-01), from the NIH-SCORE grant (S06 GM008102-3052), and from the Fondos Institucionales para la Investigación of the University of Puerto Rico (20-02-2-78-514).

* Address correspondence to this author at the Department of Chemistry, Drexel University (phone, 215-895-2268; fax, 215-895-1265; e-mail, RSchweitzer-Stenner@drexel.edu).

‡ Department of Chemistry, University of Puerto Rico.

§ Department of Biology, University of Puerto Rico.

|| Department of Chemistry, Syracuse University.

⊥ Department of Chemistry, Drexel University.

[1] Abbreviations: FT-IR, Fourier transform infrared spectroscopy; VCD, vibrational circular dichroism; ECD, electronic circular dichroism; D_3, L-aspartyl-L-aspartyl-L-aspartate; E_3, L-glutamyl-L-glutamyl-L-glutamate; K_3, L-lysyl-L-lysyl-L-lysine; PPII, polyproline II helix.

10.1021/bi035740+ CCC: $27.50 © 2004 American Chemical Society
Published on Web 12/25/2003

ECD. The far-UV−ECD signal is nearly symmetrical, showing a weak couplet with a positive maximum between 180 and 210 nm and a negative minimum between 210 and 230 nm (*10*−*13*). ECD and Raman optical activity (ROA) spectra of some IUPs (tau protein, casein, stathmin, Bob1) suggest that their structure contains a substantial amount of PPII (*14*−*16*). This implies that the term "intrinsically unstructured proteins" is inappropriate because it does not account for the high degree of order assignable to a PPII helix (*17*). On the basis of the results presented in this paper we argue below that the term "intrinsically flexible proteins" better reflects the intrinsic structural propensities and dynamic properties of the amino acid residues in a so-called unfolded protein or peptide.

The subject of this paper is directly related to an important finding about naturally unfolded proteins reported in the paper of Tompa (*3*). This author investigated the amino acid composition of "intrinsically disordered protein" regions in globular proteins. His data clearly indicate a very high propensity of E, P, and K and moderate propensities for A, V, S, and D.[2] On the contrary, W, Y, F, C, I, L, and N frequently do not appear in such structures. Theoretical calculations suggest that Flory's isolated pair hypothesis applies at least to some extent to the formation of PPII (*18*). From this it follows that PPII forming segments in proteins should exhibit nearly the same structure if taken out of their protein environment into an aqueous solution. Thus, one expects that even very small peptide fragments such as tri- and tetrapeptides can adopt PPII if they contain the appropriate amino acid residues. As shown in earlier papers from our research group, alanine-based tripeptides such as A_3, KA_2, and SA_2 exhibit indeed as a substantial population of the PPII conformation (*19*, *20*). K_3, however, was reported to adopt a left-handed turn or helical structure with its central dihedral coordinates in the upper right square of the Ramachandran plot (*19*). This result is at variance with the findings of several ECD and VCD studies which clearly show that polylysine or a seven lysine residue-containing peptide has substantial PPII contents (*11*, *21*).

This discrepancy and the apparent relevance of lysine and other ionizable amino acid residues such as glutamic acid and aspartic acid for the formation of PPII in so-called unordered peptides, proteins, and protein segments prompted us to reinvestigate K_3 and to determine the structure of deprotonated E_3 and D_3 in D_2O by the combined use of FT-IR, visible polarized Raman, VCD, and ECD spectroscopy. We show that indeed all of these peptides adopt a structure, which can be described as a somewhat distorted PPII conformation. Our results shed some light on the likely structure of so-called K-, E-, and D-rich protein segments in proteins, which have been described as disordered (*2*, *3*).

MATERIALS AND METHODS

Materials

L-Aspartyl-L-aspartyl-L-aspartate (D_3), L-glutamyl-L-glutamyl-L-glutamate (E_3), and L-lysyl-L-lysyl-L-lysine (K_3) were purchased from Bachem Bioscience Inc. (>98% purity) and used without further purification. $NaClO_4$ was obtained from Sigma-Aldrich Chemical Co. (St. Louis, MO). All chemicals were of analytical grade. The peptides were dissolved in D_2O at a concentration between 0.125 and 0.2 M for IR, Raman, and VCD and 1 mM for ECD. The pD of the solutions was adjusted by adding small aliquots of DCl or NaOD to obtain the cationic, zwitterionic, and anionic state of the peptides. The pD values were determined by utilizing the method of Glasoe and Long (*22*) to correct the values obtained from pH electrode measurements. For the Raman experiments the solvent contained 0.25−0.1 M $NaClO_4$ whose 934 cm^{-1} Raman band was used as an internal standard (*23*).

Methods

Raman Spectroscopy. Excitations at 442 nm (70 mW) and 458 nm (150 mW) were obtained from a HeCd laser (Model IK 4601R-E; Kimmon Electric US) and an argon ion laser (Lexel), respectively. A laser filter was used for 458 nm excitation to eliminate the plasma lines. The polarized exciting laser beam was focused onto the sample with a lens of 100 mm focus length. The Raman scattered light was collected in a 135° backscattering geometry. The scattered radiation was imaged onto the entrance slit (width adjusted to 100 μm) of a triple-grating spectrometer (T64000; Jobin Yvon Inc.). A polarization analyzer followed by an appropriately oriented $\lambda/2$ plate between the collimator and the entrance slit of the spectrometer was employed to measure the Raman intensity polarized parallel (I_x) and perpendicular (I_y) to the scattering plane. The scattering light was dispersed by the spectrometer and then detected by a liquid nitrogen cooled charge-coupled device (CCD) with 256 × 1024 pixels in the chip. The spectral bandwidths were 3.2 cm^{-1} for 458 nm excitation and 4.0 cm^{-1} for 442 nm excitation. The frequency calibration of the recorded Raman spectra was checked by means of the 934 cm^{-1} band of the internal standard, the frequency of which had been determined earlier with high accuracy (*23*).

FT-IR Spectroscopy. FT-IR spectra were measured with a Nicolet Magna-IR System 560 optical bench as described elsewhere (*24*). A total of 256 scans at 2 cm^{-1} resolution using Happ-Ganzel apodization were averaged to obtain each spectrum. For all experiments, a Spectra Tech liquid cell equipped with CaF_2 windows and 6 mm thick Mylar spacers was used. Each peptide sample was measured at least four times. Spectra were corrected for the solvent background in an interactive manner using Nicolet OMNIC 3.1 software.

VCD Spectroscopy. VCD spectra were measured with a Chiralir FT-VCD spectrometer from Bomem/BioTools modified to the setup of dual polarization modulation (DPM). The DPM setup includes two ZeSe photoelastic modulators (PEM), one of which is to create left and right circularly polarized radiation and the other is used to suppress the linear birefringence and the associated VCD artifacts. This spectrometer is equipped with a HgCdTe detector having a cutoff at 800 cm^{-1}. VCD spectra were measured in D_2O with a resolution of 8 cm^{-1} using a CaF_2 cell with a path length of 30 μm. The VCD spectra were collected in blocks for a total collection time of approximately 12 h, depending on the peptide sample investigated. The PEM was optimized for maximum quarter-wave response at 1400 cm^{-1}. Other experimental conditions are provided in the figure captions referring to the VCD spectra.

[2] Unfortunately, Tompa still characterizes the corresponding segments as exhibiting a lack of structure, which obfuscates the fact that their conformation is in fact PPII.

CD Spectroscopy. Far-UV CD spectra (250−190 nm) were measured with an OLIS DSM-10 UV/vis CD spectrophotometer in a 1.0 mm quartz cell with 2 nm resolution. The samples were placed in a nitrogen gas-purged OLIS CD module. The temperature of the cuvette was controlled by means of a Peltier-type heating system (accuracy ±1 °C). For each measurement, the sample in the cuvette was allowed to equilibrate for 5 min at the adjusted temperature prior to data acquisition. For all experiments reported in this paper ΔA (λ,T) was measured by increasing the temperature in increments of 5 °C. The room temperature spectra were obtained by averaging five scans. The solvent reference spectra were used as baseline, which was automatically subtracted from the peptide CD spectra. For the final presentation in this paper the original ΔA (λ,T) spectra were converted to the $\Delta\epsilon$ (λ,T) representation by using the above-mentioned sample concentrations and the path length of the cuvette according to the Beer−Lambert law.

Spectral Analysis. All IR and Raman spectra were analyzed using the program MULTIFIT (*25*). They were normalized to the internal standard, i.e., the ClO_4^- band at 934 cm^{-1}. To eliminate solvent contributions, we measured the solvent reference spectra for both polarizations, which were then subtracted from the corresponding peptide spectra. The intensities of the normalized polarized Raman bands were derived from their band areas. These and the corresponding IR spectrum were self-consistently analyzed in that they were fitted with a set of identical frequencies, half-widths, and band profiles. The isotropic and anisotropic Raman intensities and the depolarization ratios ρ were calculated as

$$I_{iso} = I_x - {}^4/_3 I_y$$

$$I_{aniso} = I_y$$

$$\rho = I_x/I_y$$

It should be mentioned that in principle I_{aniso} should be written as $2{,}33 \cdot I_y$. As indicated in our earlier papers (*19*, *20*, *27*), we prefer to identify it with I_y in the depicted figures so that the polarization properties of different lines can be better inferred.

RESULTS

Concept of Analysis. The methods used to analyze the amide I′ bands in the FT-IR, Raman, and VCD spectra of tripeptides have been described in detail in earlier papers (*19*, *26*, *27*). Briefly, our algorithm utilizes the fact that the two amide I modes are vibrationally coupled by through-bond and through-space transition dipole coupling (TDC) (*28*). The band assignable to the respective out-of-phase combination of the two individual vibrations is generally found at lower wavenumbers and is termed AI^-. The corresponding band of the in-phase combination is named AI^+. The investigated peptides are dissolved in D_2O in order to eliminate the vibrational mixing between amide I and water bending modes, which complicates the analysis (*29*). In a first step, the degree of quantum mechanical mixing is obtained from the intensity ratio $R_{iso} = I_{iso}(AI^-)/I_{iso}(AI^+)$ of the amide I′ bands in the isotropic Raman spectrum. This information is then used to determine the angle $\tilde{\theta}$ between the transition dipole moments of amide I′ from the corresponding intensity ratio $R_{IR} = I_{IR}(AI^-)/I_{IR}(AI^+)$. This angle $\tilde{\theta}$, the intensity ratio $R_{aniso} = I_{aniso}(AI^-)/I_{aniso}(AI^+)$, and the depolarization ratios ρ^- and ρ^+ depend on the dihedral angles ϕ and ψ of the central amino acid residue. We are using a simple formalism based on the combination of matrices accounting for rotation around backbone bonds to simulate $\tilde{\theta}$, R_{aniso}, ρ^-, and ρ^+ as a function of ϕ and ψ. A solution is obtained if all of the experimentally obtained parameters reproduced by the same ϕ, ψ pair in the limit of accuracy. Generally, eight mathematical solutions are obtained, but fortunately most of the solutions can be easily rejected since they correspond to forbidden regions of the Ramachandran space (*34*). A further reduction is then carried out by using the remaining pairs to calculate the VCD signal of amide I′ (*19*, *20*). In most cases this procedure yields a single, final solution. However, in one case to be discussed below, two solutions survived this scrutiny, and a decision had to be made on the basis of the respective electronic ECD spectrum. It should be mentioned that throughout this paper we disregard solutions assignable to the lower part of the Ramachandran plot ($\psi < 0$), because they are all inconsistent with the observed VCD signals.

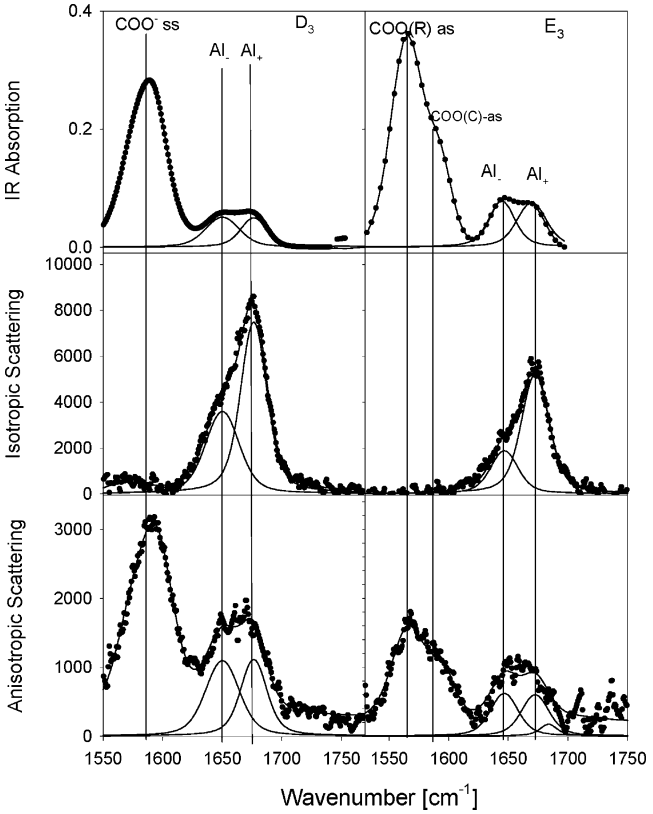

FIGURE 1: IR, isotropic, and anisotropic Raman amide I′ spectra of D_3 and E_3 in D_2O measured at pD = 7.0. The Raman spectra were measured with 442 nm excitation (laser power = 65 mV; slit width = 100 μm). The solid lines are the results of the spectral decomposition of IR and Raman amide I′ bands.

Analysis of FT-IR, Raman, and VCD Spectra of D_3 and E_3. Figure 1 shows the IR, isotropic, and anisotropic Raman spectra of the amide I′ region of D_3 and E_3 measured at pD 7. The absence of a CO stretching band above 1700 cm^{-1} and the presence of strong IR bands of the COO^- antisymmetric stretching modes between 1560 and 1590 cm^{-1} indicate that all side chains are deprotonated. The corresponding Raman bands are completely depolarized and do

Table 1: Spectral Parameters and Raman Tensors of the Amide I′ Mode of L-Aspartyl-L-aspartyl-L-aspartate (D_3) and L-Glutamyl-L-glutamyl-L-glutamate (E_3) Measured at pD 7 and L-Lysyl-L-lysyl-L-lysine (K_3) Measured at pD 1 in D_2O

parameter	D_3	E_3	K_3
Ω_{AI^-} (cm^{-1})[a]	1647	1647	1647
Ω_{AI^+} (cm^{-1})[a]	1675	1672	1674
Γ_{L^-} (cm^{-1})[b]	11	11	11
Γ_{L^+} (cm^{-1})[b]	11	11	11
Γ_{G^-} (cm^{-1})[c]	29	21	20.5
Γ_{G^+} (cm^{-1})[c]	22.6	21	24.5
ρ^-	0.17	0.25	0.14
ρ^+	0.11	0.11	0.08
R_{iso}	0.47	0.36	0.19
R_{aniso}	1.1	1.03	0.69
R_{IR}	1.19	1.05	0.73
$\tilde{\theta}$ (deg)	104	92.9	95.0
ϕ (deg)	−70	−60	−50
ψ (deg)	165	160	170

[a] Wavenumber position. [b] Lorentzian half-width if the Voigtian profile. [c] Gaussian half-width if the Voigtian profile.

therefore appear only in the anisotropic Raman spectrum (*30*). This indicates that vibrational coupling with amide I′ modes is negligible since this would mix isotropic scattering into the Raman tensor. Thus, the two amide I′ bands can be described in terms of the coupled oscillator model outlined in our earlier paper (*26*). The intensity ratios required for our structure analysis were obtained from a spectral decomposition (Figure 1), and the obtained spectral parameters are listed in Table 1.

In view of the earlier results obtained for the fully protonated state of K_3 (*19*), we expected to obtain a solution in the upper right square of the Ramachandran plot, which would be describable as a left-handed turn or left-handed helix. This expectation was supported by the corresponding electronic CD spectra shown in Figure 2a, which very much resemble that observed for K_3 (*31*). However, it turned out that all solutions with positive ϕ and ψ values were far away from any sterically allowed region. We therefore focused on the upper left square ($\phi < 0$, $\psi > 0$), which contains the broad allowed region comprising β-strand, PPII, and β-turn conformations and obtained two solutions for D_3, namely, $\phi, \psi = -30°, 100°$ and $-70°, 160°$. The first one is sterically unlikely and was found to be inconsistent with the VCD signal shown in Figure 3 in that it produces a couplet having a sign opposite to the observed one. The second solution, however, is capable of reproducing the positive signal at the AI^+ position. The AI^- signal is overestimated, but this is not surprising because our model does not assume any internal rotational strength of this mode and, thus, predicts a nearly symmetric couplet. This was found to be a solid approximation for peptides with a deprotonated C-terminus. One might speculate that the less negative signal observed for D_3 reflects some coupling between the C-terminal amide I′ mode (which contributes mostly to the AI^- vibrational state) and the COO^- mode of the nearest aspartate side chain, which has a limited effect on the Raman and IR intensity. This is theoretically possible because, e.g., a perpendicular orientation of the respective transition dipole moment is not very favorable for TDC but can provide a contribution to VCD via the $\vec{\mu}_{COO} \times \vec{\mu}_{AI}$ term in the coupled oscillator formalism (*19*) even in the case of very weak quantum mechanical mixing between the respective vibrational states.

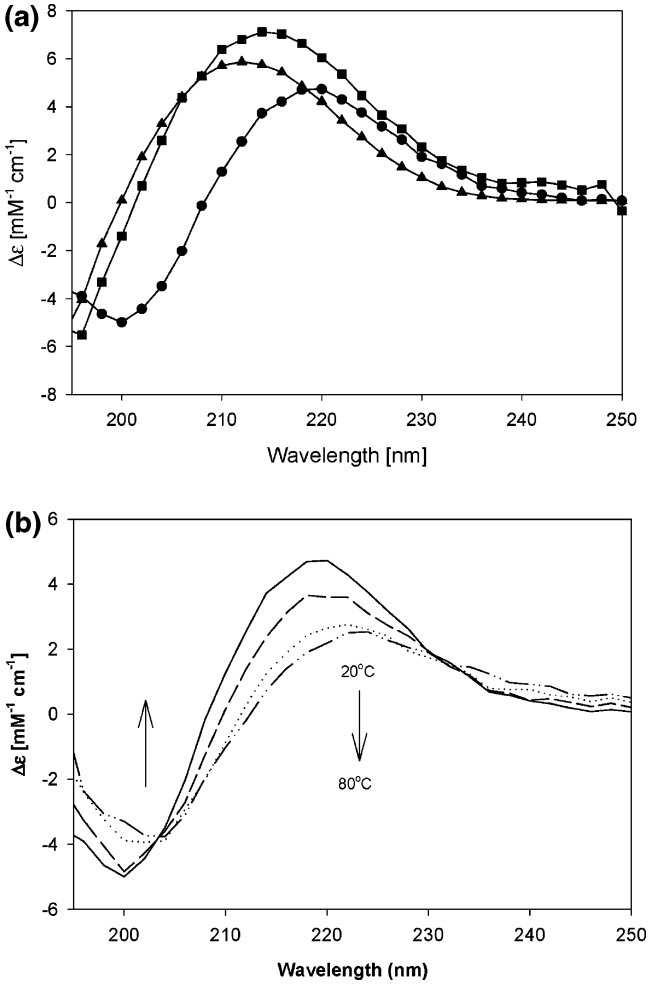

FIGURE 2: (a) ECD spectra of D_3 (●) at pD = 6.0, E_3 (■) at pD = 1.0, and K_3 (▲) at pD = 1.0 between 195 and 250 nm. (b) Temperature-dependent ECD spectra of D_3 at pD = 7.0 between 195 and 250 nm. The arrows show the direction of changes at positive and negative peaks with increasing temperature.

Therefore, we consider our reproduction of the VCD spectrum as satisfactory and, thus, $\phi, \psi = -70°, 160°$ as the correct solution. The uncertainty of the angles is about ±10°. Figure 4 illustrates how this solution reproduces the experimentally obtained values for $\tilde{\theta}$, R_{aniso}, ρ^-, and ρ^+. In the Ramachandran-like plot obtained from Shi et al. (*32*) this conformation is located slightly above the classical PPII position, which is $\phi, \psi = -75°, 145°$, but still located in the general PPII region which is describable by the intervals $-120° \leq \phi \leq 30°$ and $80° < \psi \leq 180°$ (Figure 5).

We also tried to analyze the fully protonated, cationic state of D_3, but this was rendered difficult because our spectra were very noisy owing to a strong Raman fluorescence background, which could not be eliminated using available quenchers. Furthermore, the Raman as well as the CD data indicate some excitonic coupling between the CO modes of the side chains and the amide I modes so that our coupled oscillator model is no longer applicable. Fluorescence also hampered the analysis of the spectra of the fully deprotonated peptide (pD 12), which is even more difficult to analyze even if high-quality spectra are obtained due to severe spectral overlap (*33*). However, we measured the corresponding VCD spectrum and found it reproducible by nearly the same coordinates obtained from the above analysis of the spectra taken at neutral pD (data not shown).

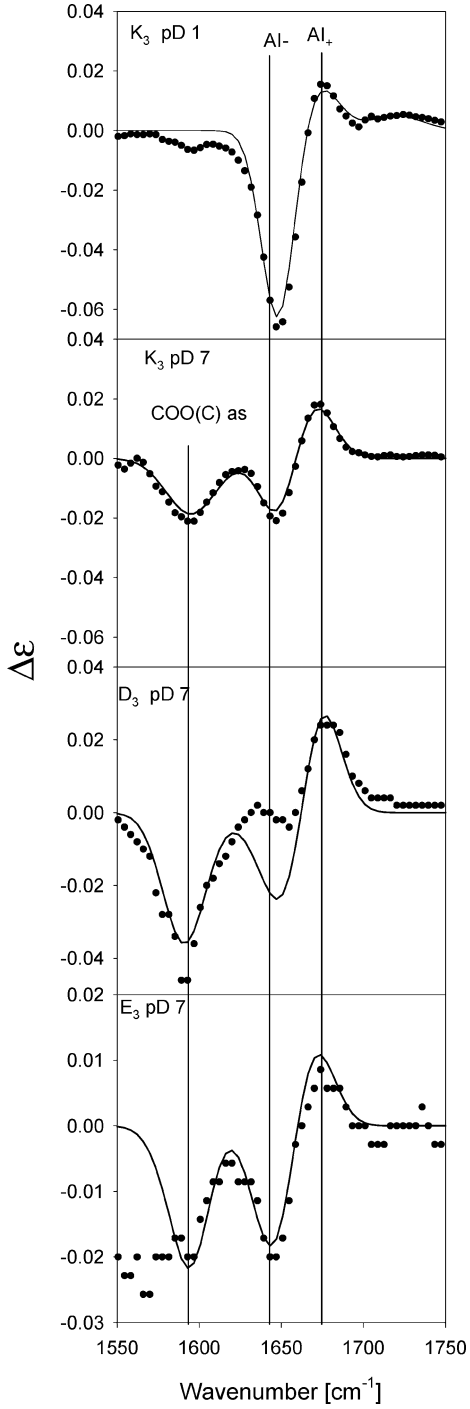

FIGURE 3: VCD spectra of the amide I′ region of K_3 at pD = 1.0 and 7.0, D_3 at pD = 7.0, and E_3 at pD = 7.0. The solid lines result from a theoretical calculation of the VCD signal based on the dihedral angles obtained from the analysis of the amide I′ band profile of the corresponding FT-IR and polarized Raman spectra.

The spectral parameters of E_3 are slightly different from those obtained for D_3 (cf. Figure 1 and Table 1). The R_{iso} value is smaller, indicating stronger excitonic coupling between the amide I′ modes. Our analysis yielded four solutions for $\psi > 0$, namely, ϕ, $\psi = -20°$, 40°; $-180°$, 40°; 120°, 160°; and $-60°$, 160°. Steric considerations suggest that only the last pair reflects a meaningful solution. This is confirmed by the VCD spectrum (Figure 3). Herein, amide I′ displays a symmetric couplet which is nicely reproduced by this pair of coordinates. All of the other solutions do not even qualitatively account for the observed

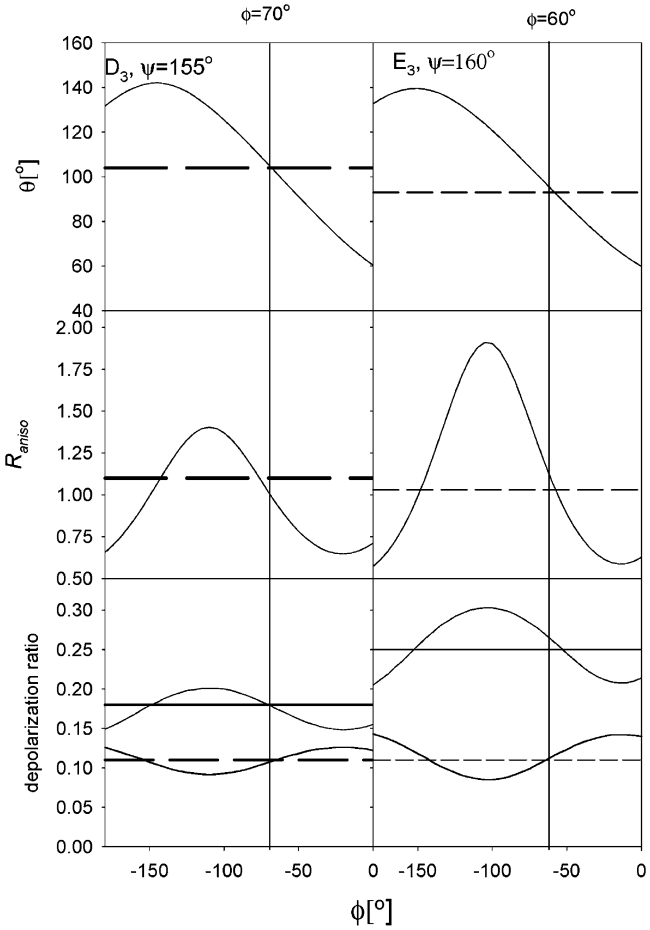

FIGURE 4: $\tilde{\theta}$, R_{aniso}, ρ_-, and ρ_+ calculated as a function of the dihedral angle ϕ for $\psi = -155°$ for D_3 and $\psi = -160°$ for E_3 by using the algorithm described previously (*26*, *27*). The horizontal lines depict the respective experimental values. The solid vertical line indicates the ϕ value by which all experimental values could be simultaneously reproduced in the limit of their accuracy.

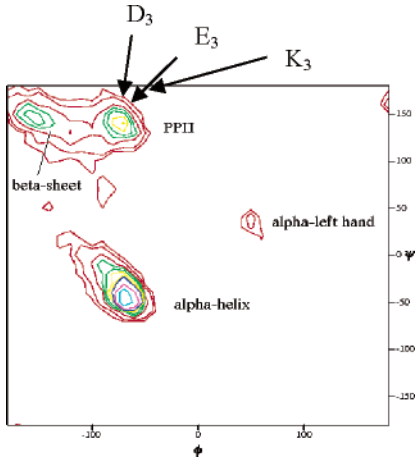

FIGURE 5: Comparison of the dihedral coordinates obtained for K_3, E_3, and D_3 with a frequency contour plot, analogous to the Ramachandran plot, showing the ϕ, ψ distribution of core backbone conformations of all non-glycine residues in high-resolution protein structures (*53*). The plot was obtained and modified from ref *32*. In the upper left square the plot visualizes two maxima corresponding to an extended β-strand and the (nearly) canonical PPII conformation.

VCD signal. The absence of any asymmetry due to coupling with the antisymmetric stretch of the glutamic acid residues reflects the greater distance between the residues' carboxylate

groups and the peptide CO groups. Thus, in the limit of accuracy we obtained a similar structure as for D_3, namely, a slightly perturbed PPII conformation. The correspondence between the obtained dihedral angles and the experimental parameters is depicted in Figure 4.

We were also able to investigate the cationic (pD 1) and the fully protonated state of E_3 (pD 12), even though the low signal to noise ratio particularly of the anisotropic Raman spectrum caused some larger uncertainties of the spectral and consequently also of the geometric parameters (data not shown). For the cationic state we obtained ϕ, $\psi = -45°$, 160° with an uncertainty of about ±20°. In the limit of accuracy the fully protonated conformation was found to be identical with that obtained at neutral pD.

Electronic CD Spectra and Revisiting K_3. We already mentioned the CD spectra of D_3 and E_3 in Figure 2a as exhibiting nearly symmetric couplets with a positive maximum around 215 nm and a negative one at 200 nm. These are not the typical PPII spectra abundantly discussed in the literature (*9*, *10*). Both spectra are very similar to that of the fully protonated K_3 (also shown in Figure 2a), indicating very similar conformations. However, K_3 was recently found to adopt a left-handed turn or helix structure with ϕ, $\psi =$ 15°, 40° because this conformation reproduced the Raman, IR, and VCD spectra of this peptide (*19*).

To resolve this discrepancy, we scanned the $\psi > 0$, ϕ of the Ramachandran space for other solutions of the Raman and IR data. Thus, two further solutions were discovered, namely, ϕ, $\psi = -20°$, 130° and −50°, 170°. The first one fails to reproduce the VCD signal, but the latter does it as perfectly as that reported earlier. Hence, it turns out that the K_3 data can also be explained by a structure, which is very close to that obtained for D_3 and E_3. It is somewhat more displaced from the ideal PPII toward the upper border of the Ramachandran plot (Figure 5). However, the similarity of the CD spectra of D_3, E_3, and K_3 (Figure 2a) clearly shows that their respective conformations must be very similar.

The temperature dependence of the CD spectrum of D_3 at pD 6 is not as pronounced as those earlier obtained for A_3 and K_3 (Figure 2b) (*31*), but it is clearly indicative of a population of a more disordered β-strand conformation at high temperatures, and a dichroic point is clearly discernible. E_3 behaves similarly (data not shown). The temperature dependence of K_3 reported earlier (*31*) is more pronounced and is more likely to reflect the population of an extended conformation at higher temperatures.

Taken together, our results clearly show that the investigated peptides with ionizable side chains adopt a very similar somewhat distorted PPII structure with coordinates located above the position of the classical PPII conformation in the Ramachandran plot. Their CD spectra deviate from that observed for PPII in that they display a much more symmetric couplet between 190 and 220 nm.

DISCUSSION

The torsion angles ϕ and ψ are the essential coordinates determining the backbone conformation of a polypeptide chain (*34*). The ideal PPII helix exhibits ϕ and ψ angles of −75° and 145° (Figure 6a), respectively; this imparts a perfect 3-fold rotational symmetry to the structure (*35*). Originally, the PPII conformation was thought to be restricted

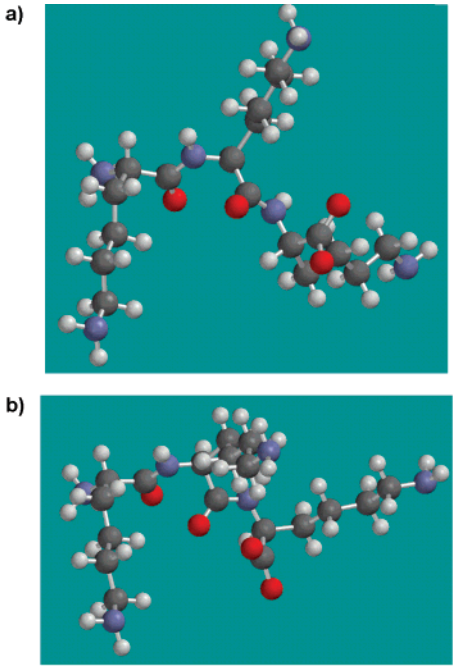

FIGURE 6: Representation of K_3 in the ideal PPII conformation (a) and in the structure inferred from the experimental data (b).

to only *trans*-proline homopolymers (*36*, *37*), but multiple experimental and theoretical evidence was provided over the last 35 years that even polypeptides which do not contain any proline residue can also adopt this conformation (*13*). By exploiting the excitonic character of amide I′ vibrational states of polypeptides, our previous polarized visible Raman, FT-IR, and VCD analysis of trialanine structure in D_2O yielded a nearly 50:50 mixture of PPII and a very extended β-strand conformation (*19*, *20*, *31*). On the contrary, trivaline was found to adopt a single extended β-strand-like conformation, while a left-handed turn or helix like structure was attributed to the fully protonated state of K_3. The latter result was surprising in view of the fact that several lines of evidence suggested a high PPII propensity of lysine. We reinvestigated K_3 in the present study and found an alternative structure being consistent with all experimental data, namely, a somewhat perturbed PPII conformation with ϕ, $\psi = -50°$, 170° (Figure 6b). As can be inferred from a direct comparison of the structures depicted in Figure 6, the obtained K_3 conformation is somewhat less extended than the canonical PPII. The distance between the N- and C-terminus is 7.4 Å for the former and 8.3 Å for the latter. Our result now confirms earlier results on lysine-containing peptides and also fits into the picture emerging from the amide I′ analysis of E_3 and D_3, which were found to exhibit very similar conformations. Interestingly, a comparison of structures obtained from the spectra measured at acid, neutral, and alkaline pD strongly suggests that the structural preference of all these peptides does not depend on the protonation state of the residues. For K_3, this is in full accordance with results reported by Rucker and Creamer (*21*), who employed ECD spectroscopy to investigate the structure of a K_7 peptide in water. Evidence for a significant PPII propensity of D and E was reported by Kelly et al. (*38*). Their host−guest study of Ac-$P_3XA_2P_3$GY-NH_2 by ECD spectroscopy showed that the propensity of asparagine and glutamine residues to form PPII helices was as high as that of proline residues. In line with these results Stapley and Creamer (*17*) presented

a survey of PPII helices from a set of 274 nonhomologous proteins, and they found that glutamine and positively charged residues are mostly favored after proline. A very good example for demonstrating the PPII propensity of E and K is the heart muscle protein titin which is composed of tandem repeats of ca. 100-residue immunoglobulin domains and 28-residue PEVK modules. Ma et al. investigated the latter by ECD and NMR spectroscopy and obtained three short stretches of PPII helices of 4−6 residues that are interrupted by more flexible and disordered segments (*39*). The authors interpreted their results as suggesting that the presence of negatively charged glutamic acid in these helices contributes to their deformation and stability.

In the following we are addressing some specific issues of our study by first discussing the CD spectra in Figure 2 and, subsequently, its implications for the understanding of the structure of K-, D-, and E-containing segments in the ligand binding sites of peptides and proteins.

Interpretation of CD Spectra. CD has proven to be a powerful and discriminating tool for the identification of PPII in peptides and proteins (*9*, *10*, *13*). The positive band around 220 nm assignable to n$-\pi^*$ transition of the peptide linkage was utilized to quantitatively determine the PPII fraction of peptides and proteins (*13*). The position of this positive peak shifts to higher wavelengths in the presence of proline residues in the sequence (*40*, *41*). The negative peak around 190 nm reflects the rotational strength of the peptides' $\pi-\pi^*$ transition. Generally, a PPII spectrum is asymmetric in that the positive n$-\pi^*$ signal is significantly weaker than the negative signal at the position of the $\pi-\pi^*$ band. The CD spectra in Figure 2 are different from the canonical PPII spectrum. Compared with that of A_3 (*31*) at ambient temperature, their positive signal is larger by a factor of 2−3, while the negative $\pi-\pi^*$ signal is nearly reduced by a factor of 2. Thus, the couplet becomes more symmetric and more similar to the basis spectrum of turn structures as reported by Sreerama and Woody (*10*).

While VCD spectra of polypeptides can now be theoretically reproduced with sufficient accuracy by various methods (*19*), the ECD spectra are still poorly understood from a theoretical point of view. However, a qualitative understanding of our results can be provided by invoking the classical study of Bayley et al. (*42*), who calculated the n$-\pi^*$ rotational strength of the alanine dipeptide AcANHMe as a function of ϕ and ψ (Figure 13 of ref *9*). The canonical PPII conformation is in the plus region, but close to the nodal line separating the plus from the minus region, thus indicating a weak positive signal as obtained. The coordinates obtained for K_3, E_3, and D_3 are more distant from this border and should therefore show a stronger signal, in line with our results. The concomitant decrease of the $\pi-\pi^*$ signal can be understood in terms of a coupling mechanism describable by a term proportional to the scalar product $\vec{\mu}_i(\pi\pi^*)\cdot\vec{m}_{i\pm1}(n\pi^*)$ of the electronic dipole moment of the ith peptide's $\pi-\pi^*$ transition and the magnetic moments of the adjacent $i \pm 1$ peptides' n$-\pi^*$ transition, which could transfer rotational strength from the $\pi-\pi^*$ to the n$-\pi^*$ signal.

One might argue that the temperature dependence of the ECD spectrum shown in Figure 2b indicates the presence of at least a second conformer, which has thus far not been taken into consideration. However, a substantial population of an extended (disordered) β-strand structure at room temperature is unlikely for the following reasons. If the obtained coordinates of the investigated peptides reflected a mixture of PPII and β-strand, the obtained ϕ value would be significantly below the canonical PPII value (as obtained for A_3), while our results indicate the opposite. From this it follows that the obtained coordinates (and also the ECD spectra) would be indicative of an even more distorted PPII structure for the dominant conformer (i.e., with ϕ values above −50°) even if a small fraction (20%) of an extended conformation were present.

K-, D-, and E-Containing Segments in Ligand Binding Sites. Besides providing some understanding for the high frequency of E, K, and to a minor extent D in "disordered" proteins or protein segments (*2*, *3*), our results may also be useful for elucidating the function of D_3, K_3, E_3, XK_2, K_2X, and D_3 motifs (X represents an arbitrary residue) in many proteins. According to Dunker et al. IUPs are found to be substantially depleted in W, C, F, I, Y, V, and L, which are called order-promoting amino acids, whereas enriched in A, R, G, Q, S, P, E, and K amino acids, which are called disorder-promoting amino acids. H and M are considered as neither order-promoting nor disorder-promoting amino acids, and D is found to be ambivalent (*2*). In many cases Y_3 motifs are directly involved in the function of their respective protein. Chang and Lee (*43*), for instance, identified the sequences PL_2K_2IKQ, P_2QK_2IKS, $PQPK_3P$, and SKRVSKRKL within different oncoproteins that directed pyruvate kinase hybrids into the cell nucleus. The proline and lysine contents of these fragments indicate that their prominent structure should be PPII. The relevance of the latter for ligand binding is explicitly corroborated by a study of Sleigh et al. (*44*), who investigated the binding of K_2, K_3, and K_3A to the periplasmic oligopeptide binding protein (OppA). The authors obtained the crystal structure of an OppA−K_2 complex. We retrieved the N-terminal ψ angle and the C-terminal ϕ angle from the respective file of the Protein Data Bank (PDB code 2RKM) and obtained the values 144° and −83.8°, which are close to the canonical PPII. It is very likely that this conformation is also adopted by K_3. However, the X-ray structures obtained for various α_1-antichymotrypsin derivatives, which all exhibit a functionally essential K_3 segment (*45*) at positions 212−214, show that the PPII model for such motifs is not generally applicable. The protein structures deposited in the Protein Data Bank rather exhibit a variety of conformational coordinates for K_3, which partly fall somewhat outside of the allowed region of the Ramachandran space, e.g., $\phi, \psi =$ −51.7°, −19.7°; −139°, −32°; and −78°, 96° for a cleaved protein variant (*46*) and $\phi, \psi = -43°, -49°; -112°, 22.5°$; and −131°, 115° for a cleaved antichymotrypsin T345R mutant (*47*). None of these coordinates fall into the PPII region visualized in Figure 5. These data rather suggest that the K_3 segment is subject to constraints imposed by the protein environment, which can tune its structure over a wide range of ϕ and ψ values. This conformational plasticity of K_3, however, might be an intrinsic property of the PPII conformation which has been shown to exhibit a shallow potential in the ϕ, ψ space; the minimum can be easily shifted by a perturbing potential (*48*). Thus, K_3 (as well as E_3 and D_3) can be expected to adopt a PPII-like structure only in an environment which is void of nonlocal protein−protein interactions. An example for such a "stress-free" environment

is the E_3DSD C-terminal activation domain of the B-cell-specific transcription c-activator, Bob1, which remains structurally unaffected upon the formation of the ternary complex with DNA (*14*). This segment certainly contributes to the PPII signal in the CD spectrum of the unbound protein. While the inspection of the remaining sequences of the protein (PDB accession code 1CQT) reveals high fractions of E, P, A, G, and D, which all have a substantial PPII propensity, segments with more than three PPII-preferring residues are not very frequent. This heterogeneity and the high helical propensity of many of its residues are likely to facilitate the transformation of the free protein mixture of PPII and disordered segments into a mostly helical structure in protein−DNA complexes (*49*). This is an unlikely option for the C-terminal segment owing to the Coulombic interactions between its charged side chains.

A similar mixture of a highly basic segment (i.e., E_3-KSE_3T, residues 152−160) and a longer sequence featuring a substantial contribution of residues with a high PPII propensity (residues 101−150) is exhibited by the kinase-inducible transcription domain (KID) of the cyclic AMP response element binding protein (CREB). Radhakrishnan et al. reported NMR experiments as indicating that the former remains disordered upon binding to the coactivator CBP (CREB binding protein), whereas part of the 120−150 segment becomes helical depending on whether the KID is phosphorylated (*50*).

In line with the propensity analysis of Tompa (*3*) numerous proteins that contain regions of intrinsic disorder required for function have been identified which exhibit very long segments predominantly composed of E, D, and K. Calsequestrin, for instance, is a calcium storage protein that has 7-residue disordered loops with five consecutive negative charges (MD_2E_2DL) and a 20-residue carboxy terminus with 16 glutamate and aspartate residues ($EGEINTED_4ED_8$). All of these disordered, negatively charged residues are involved in calcium binding (*51*). The amino-terminal VD_4K sequence of trypsinogen has an inhibitory function in that its proteolytic cleavage is required to convert the protein into trypsin. Unfortunately, the PDB files of all these proteins do not contain information about these disordered segments. Our data suggest that PPII should be the dominant structure but in a thermodynamic equilibrium with a more extended, disordered β-strand-like structure. This implies that disordered proteins and peptides are better described by the term intrinsically flexible proteins. The conformational fluctuation between both conformers is likely to be very fast (*48*) so the respective residues appear disordered in the X-ray structure analysis. NMR spectroscopy is not very helpful in this regard, because it cannot be utilized to discriminate between PPII and a truly random conformation (*52*). Thus far, only ECD (*9*, *10*) and ROA spectroscopy (*15*) provide unambiguous spectral marker signals diagnostic of PPII. Studies on peptide fragments such as the present one on ionizable homopeptides should be very useful in elucidating the structure of IUPs.

SUMMARY

By utilizing a recently developed protocol for determining the central dihedral angles of tripeptides by means of FT-IR, Raman, and VCD spectroscopy, we found that the homotripeptides E_3, D_3, and K_3 predominantly adopt an extended polyproline II helix conformation. The obtained ECD spectra are consistent with this notion. Our result suggests that the frequent appearance of these residues in so-called disordered proteins and peptides as well as in disordered segments of proteins reflects their intrinsic structural propensity for PPII. A search of the literature revealed that the binding sites of a large variety of onco-proteins and DNA binding proteins contain segments with D_3, K_2, and K_3 sequences which are essential for their function. Our results led us to propose that these binding sites prefer a PPII conformation in the absence of any nonlocal protein interactions. The relevance of the PPII conformation for ligand−receptor binding is corroborated by the fact that K_2 (and certainly also K_3) exhibits a PPII conformation in the binding pocket of OppA.

ACKNOWLEDGMENT

R.S.-S. thanks Dr. Robert Woody for a very useful discussion of the ECD spectra reported in this study.

REFERENCES

1. Taniuchi, H., and Anfinsen, C. B. (1969) *J. Biol. Chem. 244*, 3864−3875.
2. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) *J. Mol. Graphics Modelling 19*, 26−59.
3. Tompa, P. (2002) *Trends Biochem. Sci. 27*, 527−533.
4. Wright, P. E., and Dyson, H. J. (1999) *J. Mol. Biol. 293*, 321−331.
5. Tanford, C. (1968) *Adv. Protein Chem. 23*, 121−282.
6. Brant, D. A., and Flory, P. J. J. (1965) *J. Am. Chem. Soc. 87*, 2791−2800.
7. Tiffany, M. L., and Krimm, S. (1968) *Biopolymers 6*, 1767−1770.
8. Dukor, R., and Keiderling, T. (1991) *Biopolymers 31*, 1747−1761.
9. Woody, R. W. (1992) *Adv. Biophys. Chem. 2*, 37−79.
10. Sreerama, N., and Woody, R. (1994) *Biochemistry 33*, 10022−10025.
11. Keiderling, T. A., and Xu, Q. (2002) *Adv. Protein Chem. 62*, 111−161.
12. Cowan, P. M., and McGavin (1955) *Nature 176*, 501−503.
13. Shi, Z., Woody, R. W., and Kallenbach, N. R. (2002) *Adv. Protein Chem. 62*, 163−240.
14. Chang, J.-F., Phillips, K., Lundbäck, T., Gstaiger, M., Ladbury, J. E., and Luisi, B. (1999) *J. Mol. Biol. 288*, 941−952.
15. Blanch, E. W., Morozowa-Roche, L. A., Cochran, D. A. E., Doig, A. J., Hecht, L., and Barron, L. D. (2000) *J. Mol. Biol. 301*, 553−563.
16. Syme, C. D., Blanch, E. W., Holt, C., Ross, J., Goedert, M., Hecht, L., and Barron, L. D. (2002) *Eur. J. Biochem. 269*, 148−156.
17. Stapley, B. J., and Creamer, T. P. (1999) *Protein Sci. 8*, 587−595.
18. Pappu, R. V., and Rose, G. D. (2002) *Protein Sci. 11*, 2437−2455.
19. Eker, F., Cao, X., Nafie, L., and Schweitzer-Stenner, R. (2002) *J. Am. Chem. Soc. 124*, 14330−14341.
20. Eker, F., Cao, X., Nafie, L., Huang, Q., and Schweitzer-Stenner, R. (2003) *J. Phys. Chem. B 107*, 358−365.
21. Rucker, A. L., and Creamer, T. P. (2002) *Protein Sci. 11*, 989−985.
22. Glasoe, P. K., and Long, F. A. (1960) *J. Phys. Chem. 64*, 188−193.
23. Sieler, G., and Schweitzer-Stenner, R. (1997) *J. Am. Chem. Soc. 119*, 1720−1726.
24. Griebenow, K., Diaz Laureano, Y., Santos, A. M., Montañez Clemente, I., Rodriguez, L., Vidal, M., and Barletta G. (1999) *J. Am. Chem. Soc. 121*, 8157.

25. Jentzen, W., Unger, E., Karvounis, G., Shelnutt, J. A., Dreybrodt, W., and Schweitzer-Stenner, R. (1996) *J. Phys. Chem. 100*, 14184−14191.
26. Schweitzer-Stenner, R. (2002) *Biophys. J. 83*, 523.
27. Schweitzer-Stenner, R., Eker, F., Perez, A., Griebenow, K., Cao, X., and Nafie, L. (2003) *Biopolymers (Pept. Sci.)* (in press).
28. Torii, H., and Tasumi, M. (1998) *J. Raman Spectrosc. 29*, 81.
29. Sieler, G., and Schweitzer-Stenner, R. (1997) *J. Am. Chem. Soc. 119*, 1720−1726.
30. Sieler, G., Schweitzer-Stenner, R., Holtz, J. S. W., Pajcini, V., and Asher, S. A. (1999) *J. Phys. Chem. B 103*, 372−384.
31. Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2003) *J. Am. Chem. Soc. 125*, 1878−1885.
32. Shi, Z., Olson, C. A., Rose, G. D., Baldwin, R. L., and Kallenbach, N. R. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 9190−9195.
33. Schweitzer-Stenner, R., Eker, F., Huang, Q., and Griebenow, K. (2001) *J. Am. Chem. Soc. 123*, 9628−9633.
34. Ramachandran, G. N., Ramadrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol. 7*, 95−99.
35. Cowan, P. M., and McGavin, S. (1955) *Nature 176*, 470−478.
36. Tiffany, M. L., and Krimm, S. (1968) *Biopolymers 6*, 1767−1770.
37. Isemura, T., Okabayaski, H., and Sakakibara, S. (1968) *Biopolymers 6*, 307−321.
38. Kelly, M. A., Chellgreen, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A.-F., and Creamer, T. P. (2001) *Biochemistry 40*, 14376−14383.
39. Ma, K., Kan, L., and Wang, K. (2001) *Biochemistry 40*, 3427−3438.
40. Petrella, E. C., Machesky, L. M., and Kaiser, D. A. (1996) *Biochemistry 35*, 16535−16543.
41. Drake, A. F., Siligardi, G., and Gibbons, W. A. (1988) *Biophys. Chem. 31*, 143−146.
42. Bayley, P., Nielsen, E. B., and Schellman, J. A. (1969) *J. Phys. Chem. 73*, 228−243.
43. Dang, C. V., and Lee, W. M. F. (1989) *J. Biol. Chem. 264*, 18019−18023.
44. Sleigh, S. H., Tame, J. R. H., Dodson, E. J., and Wilkinson, A. J. (1997) *Biochemistry 36*, 9747−9758.
45. Naidoo, N., Cooperman, B. S., Wang, Z., Liu, X., and Rubin, H. (1995) *J. Biol. Chem 270*, 14548−14555.
46. Baumann, U., Huber, R., Bode, W., Grosse, D., and Laurell, C. B. (1991) *J. Mol. Biol. 218*, 595−600.
47. Lukacs, C. M., Zhong, J. Q., Plotnick, M. I., Rubin, H., Cooperman, B. S., and Christianson, D. W. (1996) *Nat. Struct. Biol. 3*, 888.
48. Mu, Y., Kosov, D. S., and Stock, G. (2003) *J. Phys. Chem. 107*, 5064−5073.
49. Klemm, J. D., Rould, M. A., Aurora, R., Herr, W., and Pabo, C. O. (1994) *Cell 77*, 21−32.
50. Radhakrishnan, I., Pérez-Alvardo, H., Dyson, J. H., and Wright, P. E. (1998) *FEBS Lett. 430*, 317−322.
51. Wang, S., Trumble, W. R., Liao, H., Wesson, C. R., Dunker, A. K., and Kang, C. H. (1998) *Nat. Struct. Biol. 5*, 476−483.
52. Bochicchio, B., and Tamburro, A. M. (2002) *Chirality 14*, 782−792.
53. Kleywegt, G. J., and Jones, T. A. (1996) *Structure (London) 4*, 1395−1400.

BI035740+