Peptide Models of Protein Folding Initiation Sites. 2. The G–H Turn Region of Myoglobin Acts as a Helix Stop Signal†

Hang-Cheol Shin,‡ Gene Merutka, Jonathan P. Waltho,§ Peter E. Wright,* and H. Jane Dyson*

Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received February 8, 1993; Revised Manuscript Received April 2, 1993

ABSTRACT: A series of peptide fragments of sperm whale myoglobin, corresponding to segments of the region between the G- and H-helices of the protein, have been synthesized and their conformational preferences investigated using circular dichroism and nuclear magnetic resonance spectroscopy in aqueous solution and in solvent mixtures containing water and trifluoroethanol. The smallest fragment, Mb-GH5, a five-residue peptide with the sequence HPGDF corresponding to the connecting loop between the two helices in the folded protein, adopts highly populated turn conformations in aqueous solution. A 25-residue peptide, Mb-GH25, containing the same sequence flanked by contiguous segments of the G- and H-helix sequences, was also found to contain a high proportion of conformers with a turn in this region. No helix formation was observed in the flanking sequences in water solution, either in Mb-GH25 or in control 10-residue peptides (Mb-G10 and Mb-H10) with sequences corresponding to the G- and H-helix segments. No additional helicity above that of the sum of the components was observed for Mb-GH25, indicating that a helical hairpin structure is not formed in the monomeric peptide in aqueous solution. In the presence of TFE, ordered helix is formed in Mb-GH25 according to the CD spectrum, and NMR spectra indicate that this is localized in the N-terminal portion of the peptide. NOESY spectra clearly show that the turn conformation is retained under these conditions. A peptide, Mb-AA25, in which the central Pro-Gly sequence of the turn are replaced with Ala-Ala was found to contain significantly more helix by CD in the presence of TFE, indicating that the turn sequence acts as a helix stop signal at the C-terminus of the G-helix under these conditions. The effect of TFE on the conformational preferences of these peptides confirms that helix is induced only in sequences which have a high propensity for helix formation—there is no effect on the turn conformation. These results suggest that the turn sequence may be important for helix termination during the initial stages of folding as well as in the native protein.

INTRODUCTION

Reverse turns are likely sites for initiation of protein folding, since they are determined by short-range interactions, they limit the conformational space available to the polypeptide chain, and by bringing more distant parts of the polypeptide chain together, they may be instrumental in directing subsequent folding events (Lewis et al., 1971; Zimmerman & Scheraga, 1977; Dyson et al., 1988a; Wright et al., 1988). It has long been recognized that certain amino acid sequences have a high probability of being part of a turn conformation in proteins (Chou & Fasman, 1977), and this has more recently been shown to be true also for peptides in aqueous solution (Dyson et al., 1988a). Turns have also been implicated in folding processes involving helices: the structure known as "nascent helix" formed in short peptides in aqueous solution (Dyson et al., 1988b) consists of a series of interlinked turnlike conformations, which readily form helix in the presence of solvents such as trifluoroethanol or in the intact protein.

The last 50 amino acids of myoglobin form a helical hairpin consisting of the G- and H-helices connected by a loop that adopts a turn-like structure (Takano, 1977a,b; Hanson & Schoenborn, 1981; Kuriyan et al., 1986). Since, as explained in detail in the previous paper in this series (Waltho et al., 1993), there is considerable evidence that this region of the protein is involved in early folding events, we reasoned that the very earliest events are likely to involve the G–H interhelical sequence. We therefore synthesized the five-residue sequence spanning the two helices and investigated its conformational preferences in aqueous solution. If this sequence forms turnlike structures in solution as an initiating event in the folding process, then it is likely that the structure will persist in longer fragments containing the sequence. Peptides were therefore synthesized to include portions of the G- and H-helix sequences. It is also possible that helix stabilization may occur due to intramolecular interactions at the ends of the peptide. We describe in this paper the results of a conformational analysis of several peptides which include the G–H turn sequence. Results for the isolated G- and H-helix peptides and for peptides which model the next stage of protein folding, coalescence of secondary structure, are described in the accompanying papers (Waltho et al., 1993; Shin et al., 1993).

MATERIALS AND METHODS

Peptide Synthesis. Peptides were synthesized on an Applied Biosystems Model 430A peptide synthesizer by stepwise solid-phase procedures using *t*-Boc chemistry [see review by Kent (1988)]. The peptides were acetylated with acetic anhydride at the N-terminus in dimethylformamide containing an equivalent of diisopropylethylamine and cleaved from the resin using standard procedures (HF:*p*-cresol 90:10, 0 °C for 1 h). The C-terminal amide group was introduced by use of *p*-methylbenzhydrylamine resin. The peptides were purified on a Hitachi HPLC system with a Vydac C18 semipreparative column, using a linear water/acetonitrile gradient containing

† Supported by Grant GM38794 (P.E.W. and H.J.D.) and postdoctoral award GM14526 (G.M.) from the National Institutes of Health.

‡ Present address: Institute of Biological Sciences, Hanhyo Institutes of Technology, CPO Box 1751, Seoul, Korea.

§ Present address: Department of Biology and Biotechnology, University of Sheffield, P.O. Box 594, Sheffield, S10 2UH, U.K.

0006-2960/93/0432-6348$04.00/0

Table I: Peptides Synthesized

	110 115 120 125 130	overall helix probability[c]
Mb (SW)	...E A I I H V L H S R H P G D F G A D A Q G A M N K...	
predicted[b] 2° struct	H H h h I h H I i i I B B I h B H I H h B H H b h	
Mb-G10	Ac-Q[a] A I I H V L H S R-NH_2	1.78
Mb-GH5	Ac-H P G D F-NH_2	0.37
Mb-H10	Ac-G A D A Q G A M N K-NH_2	1.18
Mb-GH25	Ac-Q[a] A I I H V L H S R H P G D F G A D A Q G A M N K-NH_2	0.78
Mb-AA25	Ac-Q[a] A I I H V L H S R H A A D F G A D A Q G A M N K-NH_2	4.83

[a] A Gln residue was substituted for the correct Glu residue in these peptides following early literature sequences which contained an error at this position. [b] Using the method of Chou and Fasman (1978). [c] Calculated by multiplying together the numerical values (Chou & Fasman, 1978) for the helix propensities for all of the residues.

0.1% trifluoroacetic acid. Amino acid analyses of the HPLC-purified peptides were in good agreement with the theoretical values.

Circular Dichroism Spectroscopy. CD spectra were recorded as described in the preceding paper (Waltho et al., 1993).

Proton NMR Spectroscopy. NMR samples were prepared in 90% ^1H_2O/10% ^2H_2O or in 99% ^2H_2O, and the pH adjusted to a value between 4 and 5 with small additions of concentrated HCl or NaOH, or DCl and NaOD. Reported pH values are meter readings uncorrected for deuterium isotope effects.

NMR experiments were performed on Bruker MSL300 or AM500 spectrometers equipped with digital phase-shifting hardware. The chemical shifts for spectra in aqueous solution or in TFE/water mixtures were referenced to internal dioxane at 3.75 ppm[1] relative to TSS. Most spectra were recorded at 278 K; probe temperature was calibrated using methanol by the method of VanGeet (1969). Two-dimensional NMR spectra were recorded in the phase-sensitive mode using the method of time-proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983).

Double-quantum filtered COSY (2QF COSY; Rance et al., 1983) and phase-sensitive NOESY (Jeener et al., 1979; Bodenhausen et al., 1984) or ROESY (Bothner-By et al., 1984; Bax & Davis, 1985) experiments were used to make complete sequence-specific assignments of all proton resonances for peptides Mb-GH5 and Mb-GH25. NOESY spectra were recorded using mixing times of 100, 150, 300, 400, and 500 ms. Solvent saturation was achieved by continuous low-power irradiation of the water resonance during the relaxation delay, as well as during the mixing time in the case of NOESY experiments. For all experiments, the transmitter offset was placed on the water resonance.

Data were processed on CONVEX C2 or Sun computers using the program FTNMR (Hare Research). Lorentzian–Gaussian or shifted sine-bell window functions were applied prior to Fourier transformation in both dimensions, and the final matrix generally contained 2K real points in both dimensions. Baseline corrections were applied in the ω_2 dimension for NOESY, ROESY, and 2Q spectra.

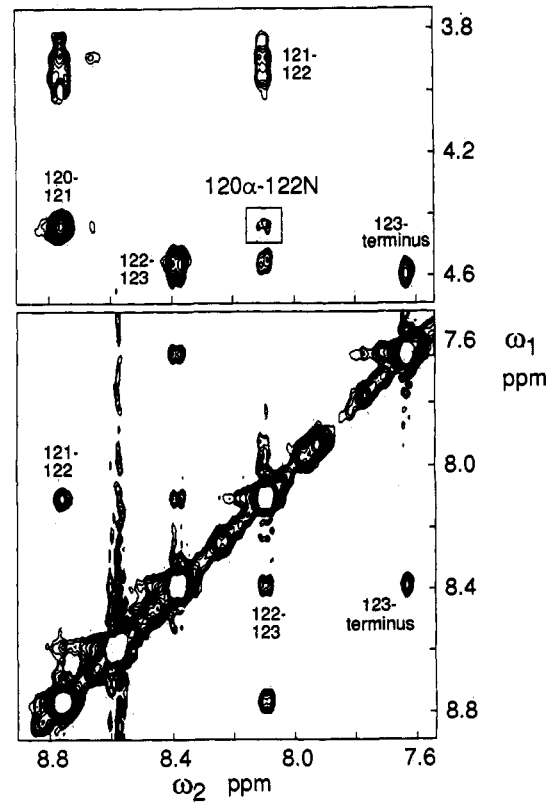

FIGURE 1: Portion of a 300-MHz ROESY spectrum (τ_m = 300 ms) of Mb-GH5 at pH 5.0, 280 K. The $d_{\alpha N}(i,i+2)$ NOE between Pro^{120} and Asp^{122} is shown in the box, plotted at a slightly lower contour level than the rest of the figure.

RESULTS AND DISCUSSION

Selection of Peptides. The sequences of all of the peptides are shown in Table I, together with the portion of the sperm whale myoglobin sequence from which they are derived and a predicted probability of helix formation calculated by multiplying together the helix probabilities for the residues, obtained from Chou and Fasman (1978). The central two residues of the turn, Pro^{120} and Gly^{121}, were substituted with Ala in the peptide Mb-AA25, to investigate the effect of the turn on helix propagation. In peptides Mb-G10, Mb-GH25, and Mb-AA25, a glutamate residue in the sperm whale

[1] Abbreviations: NMR, nuclear magnetic resonance; CD, circular dichroism; Mb, myoglobin; Mb-G10, Mb-G19, Mb-G21, Mb-GH5, Mb-GH25, Mb-GH51, Mb-H10, Mb-H27, Mb-H30, synthetic peptides of varying lengths corresponding to fragments of the G- and H-helices and the G–H hairpin region of myoglobin; Mb-GssG, Mb-HssH, Mb-GssH, disulfide-bridged homo- and heterodimers, respectively, of the G- and H-helix peptides of myoglobin; NOE, nuclear Overhauser effect; HPLC, high-performance liquid chromatography; 2QF COSY, double quantum-filtered two-dimensional correlated spectroscopy; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; ROESY, rotating-frame NOESY spectroscopy; 2Q, double quantum spectroscopy; R-COSY, relayed COSY spectroscopy; $d_{\alpha N}(i,j)$, $d_{NN}(i,j)$, etc., intramolecular distance between the protons $C^{\alpha}H$ and NH, NH, and NH, etc., on residues i and j; $^3J_{HN\alpha}$, NH–$C^{\alpha}H$ coupling constant; ppm, parts per million; TSS, (trimethylsilyl)propanesulfonic acid; TFE, 2,2,2-trifluoroethanol.

Table II: Proton Resonance Assignments for Mb-GH5 (278 K, pH 5.0, 90% ^1H_2O/10% ^2H_2O)

residue	chemical shift (ppm) NH	$C^\alpha H$	$C^\beta H$	$C^\gamma H$	$C^\delta H$	other	$-\Delta\delta/\Delta T$ (ppb/K)
H^{119}	8.59	5.00	3.08, 3.08		7.33	8.57 ($C^\epsilon H$)	6.9
P^{120}		4.45	2.32, 1.98	2.03	3.85, 3.69		
G^{121}	8.77	3.85, 3.95					8.6
D^{122}	8.10	4.56	2.57, 2.57				3.3
F^{123}	8.39	4.60	3.22, 3.00		7.28	7.37 ($C^\epsilon H$) 7.33 ($C^\zeta H$)	8.3

N-terminal CH_3 1.96; C-terminal NH_2 7.63, 7.21

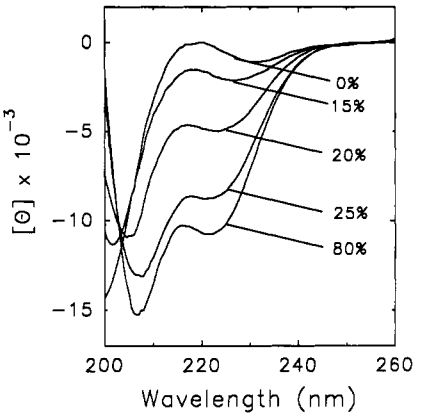

FIGURE 2: CD spectra at 278 K and pH 5.0 of Mb-GH25 (16 μM) in water and in water–TFE mixtures at concentrations (%v/v TFE) of 0%, 15%, 20%, 25%, and 80%.

myoglobin sequence was inadvertently replaced by glutamine. This substitution was made due to reference to an erroneous primary sequence and is the same sequence as was used for the initial study of Mb-G (Waltho et al., 1993). The substitution is of no consequence to the studies reported here; a comparison of the CD spectra of Mb-G(Q^{109}) (Waltho et al., 1993) and Mb-G19(E^{109}) (Shin et al., 1993) reveals no detectable difference.

Conformation of Mb-GH5 in Water. The CD spectrum of Mb-GH5 in aqueous solution shows a positive band at 220 nm and shoulders at ~210 and 197 nm (data not shown) and is apparently dominated by the dichroism of the phenylalanine residue (Sears & Beychok, 1973). A difference spectrum obtained by subtracting the CD spectrum measured in 6 M guanidine hydrochloride (GuHCl) contains a broad positive band centered around 214 nm, similar to that of a type-II β-turn (Woody, 1974; Brahms & Brahms, 1980).

Complete ^1H resonance assignments of Mb-GH5 were made using standard two-dimensional methods (Billeter et al., 1982) and are summarized in Table II. A low concentration of a minor form (<5%) was observed by NMR spectra, probably due to *cis–trans* isomerism of the proline. A similar proportion of *cis* isomer has been observed for the peptide HPGDV (unpublished observations). Amide proton chemical shifts were measured over the temperature range 278–298 K and the temperature coefficients $\Delta\delta/\Delta T$ (ppb/K) are listed in Table II. The low temperature coefficient of the Asp^{122} amide proton (−3.3 ppb/K) indicates that it is protected from solvent to a considerable extent. The other amide proton temperature coefficients are high, indicating relative exposure to the solvent.

The ROESY spectrum of Mb-GH5 (Figure 1) shows strong $d_{\alpha N}(i,i+1)$ NOEs throughout the peptide and $d_{NN}(i,i+1)$ NOE connectivities between Gly^{121}-Asp^{122} and Asp^{122}-Phe^{123}. A weak $d_{\alpha N}(i,i+2)$ NOE connectivity is observed between Pro^{120} and Asp^{122}, which, together with the low temperature coefficient of the Asp^{122} amide proton, is consistent with the formation of a hydrogen-bonded reverse turn (Dyson et al., 1988a).

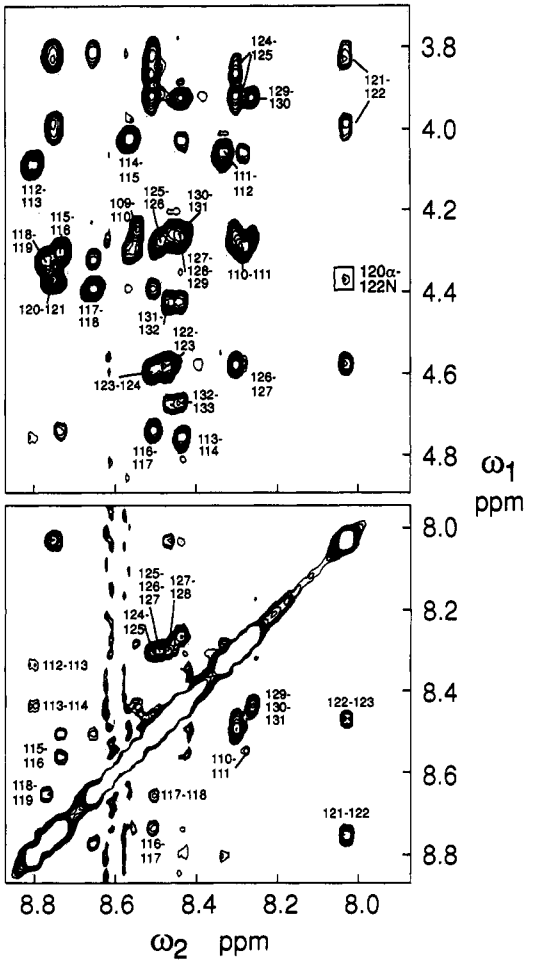

FIGURE 3: Part of the 500-MHz NOESY spectrum (τ_m = 300 ms) of Mb-GH25 in 90% ^1H_2O/10% ^2H_2O at 278 K, pH 4.6. Sequential $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ NOE cross peaks are indicated. Other cross peaks are intraresidue $d_{N\alpha}(i,i)$ and $d_{N\beta}(i,i)$ connectivities. The $d_{\alpha N}(i,i+2)$ NOE between Pro^{120} and Asp^{122}, boxed, is plotted at the same contour level as the rest of the figure.

Conformation of Mb-GH25 in Water. The lack of significant negative ellipticity at 222 or 208 nm in the CD spectrum of Mb-GH25 in water solution (Figure 2) indicates that if any ordered helical conformations are present they are at very low population. Mb-GH25 contains 10 residues (109–118 and 124–133) on either side of the turn that correspond to 2–3 turns of helix in the folded protein. Given the short length of these potential helical sequences and the presence of a possible helix stop signal at the center of the peptide, the lack of ordered helix detectable by CD is perhaps not surprising. Addition of TFE to Mb-GH25 causes a substantial increase in the negative ellipticity at 208 and 222 nm (Figure 2), indicating the formation of helix. This is treated in detail in a later section.

Table III: Proton Resonance Assignments for Mb-GH25 (278 K, pH 4.6, 90% ^1H_2O/10% ^2H_2O)

	chemical shift (ppm)					
residue	NH	$C^\alpha H$	$C^\beta H$	$C^\gamma H$	$C^\delta H$	other
Q^{109}	8.45	4.25	2.15, 1.95	2.38, 2.38		7.73, 7.00 ($N^\delta H$)
A^{110}	8.55	4.29	1.35			
I^{111}	8.29	4.06	1.76	1.46, 1.17	0.85	0.75 ($C^\gamma H_3$)
I^{112}	8.33	4.09	1.77	1.43, 1.15	0.81	0.83 ($C^\gamma H_3$)
H^{113}	8.81	4.76	3.19, 3.12		7.29	8.62 ($C^\epsilon H$)
V^{114}	8.44	4.03	1.98	0.90, 0.84		
L^{115}	8.57	4.30	1.59, 1.59	1.52	0.90, 0.83	
H^{116}	8.75	4.75	3.17, 3.25		7.30	8.62 ($C^\epsilon H$)
S^{117}	8.51	4.40	3.81, 3.81			
R^{118}	8.66	4.32	1.76, 1.69	1.57	3.16	6.94 ($N^\epsilon H$)
H^{119}	8.75	5.00	3.23, 3.13		7.33	8.58 ($C^\epsilon H$)
P^{120}		4.39	2.32, 1.95	2.03	3.81, 3.59	
G^{121}	8.74	3.80, 4.00				
D^{122}	8.08	4.61	2.65, 2.65			
F^{123}	8.47	4.60	3.21, 3.02		7.24	7.32 ($C^\epsilon H$), 7.29 ($C^\zeta H$)
G^{124}	8.51	3.92, 3.98				
A^{125}	8.31	4.27	1.39			
D^{126}	8.50	4.61	2.79, 2.71			
A^{127}	8.31	4.27	1.39			
Q^{128}	8.46	4.27	2.15, 2.02	2.38, 2.38		
G^{129}	8.45	3.92				
A^{130}	8.27	4.28	1.40			
M^{131}	8.45	4.43	2.07, 2.03	2.61, 2.52		2.09 ($C^\epsilon H_3$)
N^{132}	8.47	4.67	2.84, 2.73			7.75, 7.03 ($N^\gamma H$)
K^{133}	8.45	4.25	1.87, 1.76		1.68	2.99 ($C^\epsilon H_2$), 7.63 ($N^\epsilon H_3$)
N-terminal CH_3 2.02; C-terminal NH_2 7.70, 7.26						

Complete ^1H resonance assignments for Mb-GH25 were made using 2QF COSY and NOESY spectroscopy at 500 MHz. Proton resonance assignments are summarized in Table III. The NOESY spectrum of Mb-GH25 (Figure 3) shows all of the expected $d_{\alpha N}(i,i+1)$ NOE connectivities. The intensity of these NOEs indicates that the peptide significantly populates the β-region of (ϕ,ψ) conformational space. In addition, $d_{NN}(i,i+1)$ and $d_{\beta N}(i,i+1)$ NOEs are observed throughout the peptide, implying that the peptide also contains a significant population of conformers with dihedral angles in the α-region of (ϕ,ψ) space (Wüthrich, 1986; Dyson & Wright, 1991). Additional evidence is required before the presence of a folded conformation can be deduced. Medium- and long-range NOE connectivities in the NMR spectrum can provide such evidence, as can the CD spectrum. A $d_{\alpha N}(i,i+2)$ NOE between Pro^{120} and Asp^{122} indicates that this region contains a well-populated β-turn conformation, as it does in the shorter peptide Mb-GH5. Chemical shifts in this region of the peptide are similar to those found for Mb-GH5, consistent with the presence of similar conformations in the two peptides. Neither $d_{N\delta}(i,i+1)$ nor $d_{\delta N}(i,i+1)$ connectivities were observed between Pro^{120} and His^{119} and Gly^{121}, respectively, again consistent with the results for Mb-GH5, and indicating that the turn type is probably close to type II.

Unambiguous medium-range $d_{\alpha N}(i,i+3)$ or $d_{\alpha\beta}(i,i+3)$ NOEs are generally absent from the spectra of Mb-GH25 in aqueous solution, indicating that helical conformations are not highly populated. No evidence was found in water solution for long-range NOEs between the two ends of the peptide, which would indicate that compact folded structures were being formed. To test whether such a conformation could be induced by forcing the peptide into helix, the solution structure of the peptide was examined in the presence of trifluoroethanol (TFE). This experiment also tests whether TFE is capable of inducing helix propagation through the turn region.

Conformation of Mb-GH25 in TFE Solutions. The conformation of Mb-GH25 was studied as a function of trifluoroethanol concentration. CD spectra show that the peptide becomes more helical upon addition of TFE (Figure 2), to a maximum helix population of approximately 30%. The conformation of the peptide was studied by NMR at two concentrations of TFE, 25% and 40% (v/v) and portions of NOESY spectra recorded under these conditions are shown in Figures 4 and 5. Extensive medium-range NOE connectivities are observed even at 25% TFE. These are intensified at 40% TFE, and additional connectivities are observed.

The NOE connectivities for Mb-GH25 under these conditions are summarized in Figure 6. It is clear that TFE stabilizes an ordered helical conformation from Gln^{109} to Arg^{118}, as indicated by the extensive network of medium-range NOE connectivities in the N-terminal half of the peptide, i.e., in the region corresponding to the C-terminal portion of the G-helix of myoglobin. The presence of both $d_{\alpha N}(i,i+2)$ and $d_{\alpha N}(i,i+4)$ NOE connectivities, between Ile^{111} and His^{113} and Leu^{115} respectively (see Figure 4), suggests that the conformational ensemble may include both α-helical and 3_{10} helical forms in this region. The C-terminal half of Mb-GH25, corresponding to the N-terminal portion of the H-helix of myoglobin, does not appear to adopt stabilized helical conformations in TFE. A few isolated medium-range NOE connectivities are observed in water/TFE mixtures (Figure 6); however, most of the other medium-range NOEs that would normally be observed for a helical peptide are unambiguously absent from the spectra. The lack of an extensive network of medium-range NOEs argues against formation of a significant population of ordered helix. Furthermore, the $d_{NN}(i,i+1)/d_{\alpha N}(i,i+1)$ NOE intensity ratios for residues in the C-terminal region of the peptide remain approximately constant upon addition of TFE, whereas those for the N-terminal portion increase (Figures 4–6). We therefore conclude that there is little increase in the proportion either of ordered helical forms or of backbone dihedral angles in the α region of conformational space in the C-terminal region of Mb-GH25.

A further illustration of the localized nature of the ordered helical conformations induced by TFE is provided by the change in the chemical shift of the NH, $C^\alpha H$, and $C^\beta H$ resonances of the peptide, as shown in Figure 7. A consistent

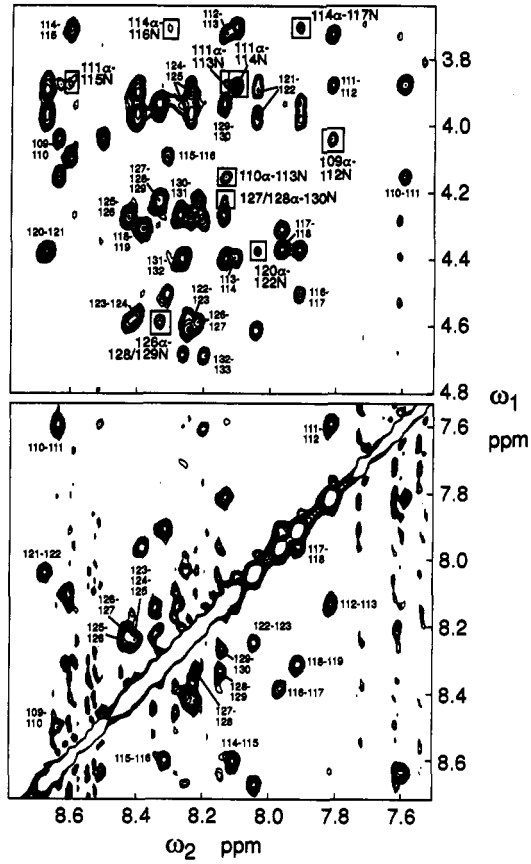

FIGURE 4: Part of the 500-MHz NOESY spectrum (τ_m = 150 ms) of Mb-GH25 in 25% (v/v) TFE-d_3 in ^1H_2O, at 278 K, pH 4.6. Medium-range $d_{\alpha N}(i,i+2)$, $d_{\alpha N}(i,i+3)$, and $d_{\alpha N}(i,i+4)$ NOE connectivities, boxed, are plotted at the same contour level as the rest of the figure.

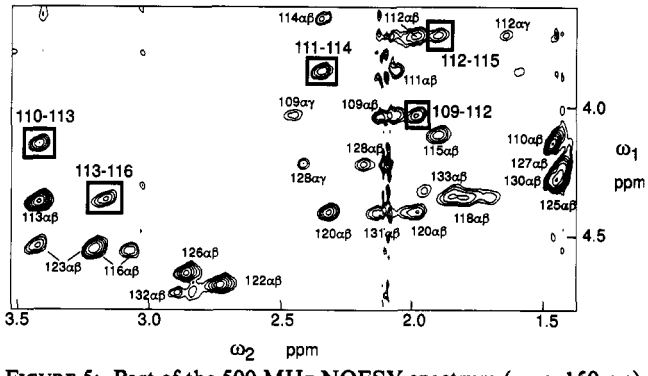

FIGURE 5: Part of the 500 MHz NOESY spectrum (τ_m = 150 ms) of Mb-GH25 in 40% (v/v) TFE, showing medium-range $d_{\alpha\beta}(i,i+3)$ NOE connectivities, boxed, plotted at the same contour level as the rest of the figure.

upfield shift (to lower ppm) of the C$^\alpha$H resonance is observed for residues 109–116 upon addition of TFE; C$^\alpha$H chemical shifts in the remainder of the peptide are little affected by TFE. Upfield shifts in C$^\alpha$H resonances have been correlated with the regions of helix in proteins (Dalgarno et al., 1983; Pastore & Saudek, 1990; Wishart et al., 1991) and helix formation in peptides (Peña et al., 1989; Chandrasekhar et al., 1991; Lee et al., 1992; Zhou et al., 1992). Most of the C$^\beta$H resonances of Mb-GH25 are shifted downfield in the N-terminal portion (residues 109–119) of the molecule (Figure 7), consistent with the small but systematic downfield shift observed for the C$^\beta$H resonances of helical residues in proteins (Wishart et al., 1991).

The largest change in chemical shift upon addition of TFE to solutions of Mb-GH25 occurs for the NH resonances of the

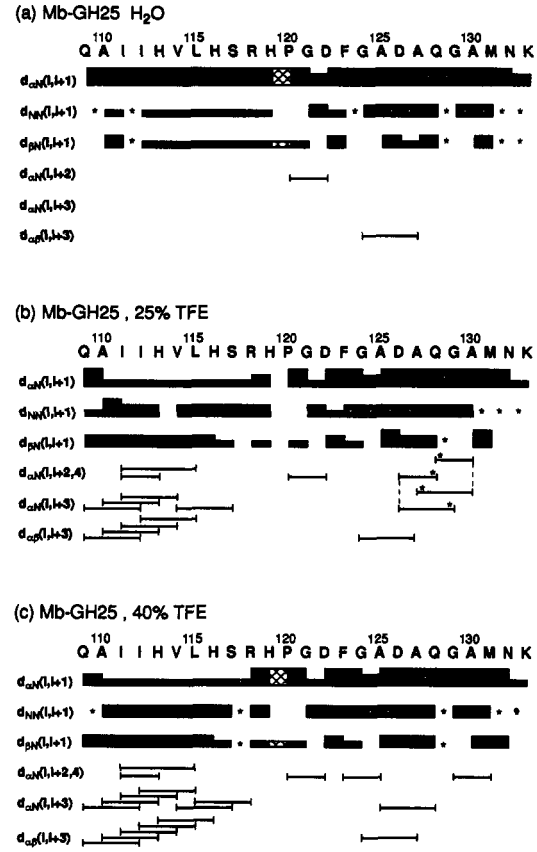

FIGURE 6: Summary of sequential and medium range NOE connectivities for Mb-GH25 (a) in 90% ^1H_2O/10% ^2H_2O, (b) in 25% TFE (v/v), and (c) in 40% TFE (v/v). The thickness of the bars corresponds to NOE intensities, hatched bars represent connectivities to the C$^\beta$H of proline, as a substitute for the NH. Connectivities that were obscured by resonance overlap are denoted by a black asterisk; connectivities for which a cross peak is present, but which are ambiguous due to the resonance overlap of residues on either side are denoted by a white asterisk inside a bar of estimated average size. Medium-range NOEs which are unambiguously present, but which cannot be unambiguously assigned due to specific resonance overlaps, are shown joined by dotted lines. Possible medium-range NOE connectivities that are unobservable due to overlap with strong sequential or intraresidue cross peaks are omitted, but the majority of the connectivities absent from this figure, especially in (a), are in fact unambiguously absent from the spectrum.

residues in the N-terminal half of the peptide. Most, but not all, of these resonances are shifted significantly upfield, as has been reported for helical regions of other peptides in TFE (Zhou et al., 1992; Blanco et al., 1992) these shifts were attributed to changes in hydrogen bonding associated with helix formation. Our results are fully consistent with this interpretation. The amide proton resonances of the first two residues, Gln^{109} and Ala^{110}, do not shift significantly in TFE, despite the fact that this region is helical as indicated by both the C$^\alpha$H and C$^\beta$H resonance shifts and the presence of medium-range NOEs (Figure 6). Since these are the N-terminal residues, no helical hydrogen bonds are possible for their amide protons. The NH resonance of Ile^{111} experiences a large upfield shift, which we attribute to intramolecular hydrogen-bond formation, presumably to the acetyl blocking group in a 3_{10} hydrogen-bonding pattern. The observed downfield shift of the amide proton resonance of Leu^{115} probably reflects a change in hydrogen-bond strength (Kuntz et al., 1991; Zhou et al., 1992) rather than disruption of the helix, since both the network of medium-range NOEs and the C$^\alpha$H resonance shifts indicate continuous helix through this region. The chemical shifts (Figure 7) suggest that the hydrogen-bonded helix terminates at His^{119} in TFE. We do

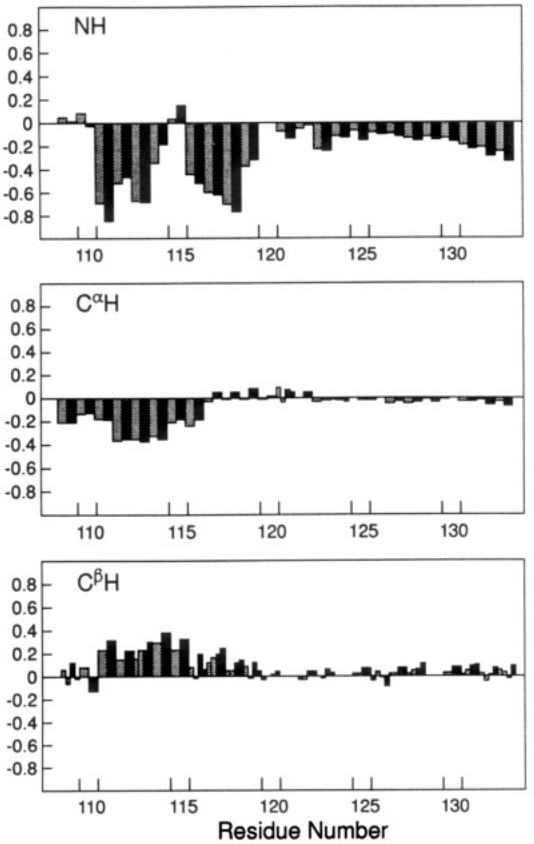

FIGURE 7: Histogram showing the change in chemical shift of the NH, $C^{\alpha}H$ and $C^{\beta}H$ resonances of Mb-GH25 upon addition of TFE. The chemical shifts in 25% and 40% TFE were subtracted from the shifts in 0% TFE to give the hatched and filled bars respectively. Values for the $C^{\beta}H$ and for the $C^{\alpha}H$ of glycine residues are divided where two chemical shifts were available; in these cases, the $C^{\beta}H$ chemical shifts were subtracted as though the high- and low-field values had simply moved, rather than considering the possibility that they had been interchanged.

not observe the periodicity in amide proton chemical shifts reported by Zhou et al. (1992), probably because extensive hydrophobic packing leading to helix bending does not occur in Mb-GH25. Thus, both medium-range NOEs and changes in chemical shift upon addition of TFE are consistent in indicating helix formation in TFE between residues 109 and 119, all of which are helical in intact myoglobin. Residues 124–133 show little tendency to adopt helical structures even in TFE but are helical in myoglobin.

The chemical shifts of residues 119–122, which comprise the turn conformation in Mb-GH5 and Mb-GH25 in water solution, are little changed upon addition of TFE. Indeed, the proton chemical shifts of these residues are remarkably consistent for peptides of greatly varying lengths and under significantly different conditions; for example, the $C^{\alpha}H$ chemical shift of Pro^{120} varies only between 4.38 (Mb-GH25, 25% TFE) and 4.45 ppm (Mb-GH5). Significantly, the $d_{\alpha N}(i,i+2)$ NOE connectivity between Pro^{120} and Asp^{122} is present under all of these conditions, and no additional medium-range NOEs characteristic of helix are observed between His^{119} and Gly^{124}, indicating that the β-turn conformation is not influenced by TFE nor by the helical conformations induced in the neighboring regions.

Conformation of Peptides Lacking the Turn Sequence. To assess the influence of the turn sequence on helix formation in Mb-GH25, three control peptides were synthesized, Mb-G10, Mb-H10, and Mb-AA25 (Table I), representing respectively the portions of the G- and H-helices present in

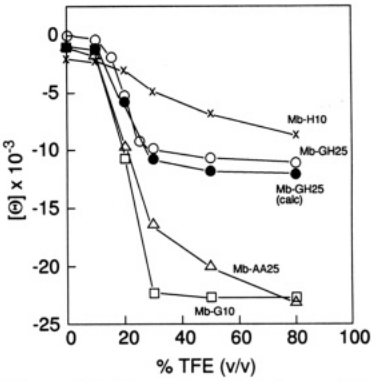

FIGURE 8: Effect of TFE on the ellipticity at 222 nm ([Θ]) for Mb-GH25 (16 μM) (○), Mb-G10 (23 μM) (●), Mb-H10 (54 μM) (×) and Mb-AA25 (11 μM) (△). All of the CD spectra were acquired at pH 5.0 ± 0.1 and 278 K. The filled circles show the expected ellipticity values for Mb-GH25 calculated from the per-residue ellipticities of Mb-G10 and Mb-H10 under the same conditions, assuming no effect on the CD spectrum of the extra five residues in the center of the peptide.

Mb-GH25, and a sequence identical to that of Mb-GH25 except that the central Pro^{120}-Gly^{121} sequence was replaced by Ala^{120}-Ala^{121}. The behavior of all three peptides was comparable to that of Mb-GH25: little helical structure was visible by CD spectroscopy in water solution, but helix could be induced by the addition of TFE. These results are shown in Figure 8.

Mb-G10 becomes about 70% helical upon addition of TFE. Mb-H10 becomes progressively more helical upon addition of TFE, to a maximum helix content of about 20%, even at 80% TFE. This is consistent with the results seen on addition of TFE to Mb-GH25. The maximum helix content (~30%) of Mb-GH25 in TFE solutions (circles of Figure 8) is a little smaller than the average of the maximum per-residue helix contents of Mb-G10 and Mb-H10 under the same conditions (filled circles). There is apparently a consistent systematic difference of about 800 deg cm^2 dmol^{-1} between the two values in water solution and at each TFE concentration. This may well be due to the presence of some positive ellipticity at 222 nm associated with the turn conformation, which is apparently populated to the same extent at all TFE concentrations according to the NMR spectra. We do not observe by CD a measurable increase in helix content in Mb-GH25 relative to the component peptides which would be expected if additional helix were stabilized in the longer peptide by intramolecular interactions between the N- and C-terminal regions.

The maximum helix content (~65%) observed for Mb-AA25 in TFE solutions is considerably greater than that for Mb-GH25 under the same conditions (Figure 8). Given the limited tendency of residues 124–133 to adopt helical conformations in Mb-GH25 or in Mb-H10, even in TFE, the greater helicity of Mb-AA25 in TFE indicates extension of the G-helix in the latter peptide. These results are consistent with the suggestion that the turn conformation seen in residues 119–122 acts as a helix stop signal in Mb-GH25.

Central Turn Conformation. A major finding of the present work is that a turn conformation is present at significant populations in the same sequence in peptides of different lengths. The four-residue sequence His^{119}-Pro^{120}-Gly^{121}-Asp^{122} consistently shows evidence of turn formation in peptides Mb-GH5 and Mb-GH25 under a variety of conditions, as well as in Mb-GH51 (Shin et al., 1993), and the proton chemical shifts of these residues are virtually identical.

We infer from this that the turn conformation and population are similar under all of these conditions.

A distinction between type I (or III) and type II turn conformations in folded proteins can be made by NMR on the basis of the NOE connectivities between residues 2 and 3 of the turn: a $d_{NN}(i,i+1)$ connectivity between these two residues is expected for a type I turn, and a $d_{\alpha N}(i,i+1)$ NOE connectivity indicates type II (Wagner et al., 1986). The identification of turn type in peptides in solution by this means is ambiguous, however, since the conformational ensemble of the peptide invariably contains a substantial population of unfolded forms with dihedral angles in the β-region of (ϕ,ψ) space, giving rise to a strong $d_{\alpha N}(i,i+1)$ NOE connectivity between Pro^{120} and Gly^{121}, for example. Since proline does not have an amide proton, a $d_{NN}(i,i+1)$ NOE connectivity cannot be observed. However, the $C^{\delta}H$ protons of the proline are located in a similar position to the amide proton of other residues; connectivities involving the $C^{\delta}H$ protons can be used to determine backbone conformational preferences in the vicinity of proline (Wüthrich et al., 1984). Thus, a $d_{\delta N}(i,i+1)$ connectivity between Pro^{120} and Gly^{121} is equivalent to a $d_{NN}(i,i+1)$ NOE connectivity between other residues. No NOE connectivities are observed between the $C^{\delta}H$ resonances of Pro^{120} and the NH resonances of His^{119} or Gly^{121} in any of the peptides. As suggested previously (Dyson et al., 1985), we infer from this that the population of type I turn conformations is rather low, and that the turn is closer to type II. The CD data for Mb-GH5 also suggest the presence of type II turn conformations.

The crystal structures of myoglobin (Takano, 1977a,b; Hanson & Schoenborn, 1981; Kuriyan et al., 1986) show that a turn is formed by His^{119}-Pro^{120}-Gly^{121}-Asp^{122}; the chain reversal is not of a standard β-turn type but involves a bifurcated hydrogen bond between the CO of His^{119} and the amide protons of Asp^{122} and Phe^{123}. Similar hydrogen bonded structures have been observed in theoretical simulations of similar peptides in water solution (Wright et al., 1990).

On the basis of previous work (Dyson et al., 1988a) we had expected the sequence HPGD to form a turn. Conformational preferences of a series of peptides for which the residues in the initial YPYDV sequence were replaced by all other amino acids in turn [series XPGDV, AXGDV (H. J. Dyson, J. J. Osterhout, L. Bolinger, and P. E. Wright, unpublished), YPXDV, and YPYXV (Dyson et al., 1988a)] show that the sequence APGD has the highest turn propensity under the conditions of the NMR experiments. Further, the identity of residue 1 is less crucial than the others; the sequence HPGD therefore contains a sequence which we would regard as one of high potential for turn formation in water solution. The presence of the turn at high population in Mb-GH5 is therefore entirely consistent with former results. The presence of the turn in the longer peptides is an important observation, since it implies that the turn will form in the intact, unfolded myoglobin polypeptide under folding conditions.

Many examples now exist to show that the conformational propensities of short peptides can reflect quite closely the secondary structure of the same sequence in the folded protein. Apart from the myoglobin system described in this and the accompanying papers, we have observed such a correspondence between structure in protein and conformational proprensities of peptide fragments for myohemerythrin (Dyson et al., 1992a) and plastocyanin (Dyson et al., 1992b). Turn structures observed in the same sequence present in peptides of different length have also been found for fragments corresponding to the V3 loop sequence of the gp120 envelope glycoprotein of HIV1 (Chandrasekhar et al., 1991) and of HIV2 (N. Assa-Munt, E. Norrby and H. J. Dyson, unpublished observations).

Recent theoretical and experimental studies have strongly implicated turns in the initiation of protein folding. Small but significant local intrinsic turn preferences have been shown to be essential for the formation of correct protein topologies in lattice Monte Carlo computations of the folding of many different structural motifs (Skolnick & Kolinski, 1990, 1991) and in the folding of helical hairpin structures using off-lattice Brownian dynamics (Rey & Skolnick, 1991). Turns have been shown to be essential for in vivo folding in the P22 tailspike protein, where a temperature sensitive folding defect (Yu & King, 1988) has been correlated with reduced capacity for turn formation in a short peptide model of the mutant protein (Stroup & Gierasch, 1990). The time scale for the local conformational search for turn formation is probably of the order of 1 ns or less, based on molecular dynamics simulations of peptides in water (Tobias et al., 1991; Wright et al., 1990). Even the formation of helices may be initiated by turn formation, as modeled, for example, by the nascent helix observed for peptides derived from the four-helix bundle protein myohemerythrin (Dyson et al., 1988b; Dyson et al., 1992a). Such structures have been observed in some protein crystal structures, where water molecules inserted between turns of helices give conformations identical with reverse turns (Sundaralingam & Sekharudu, 1989). Turns have also been observed as intermediates in helix-coil transitions in molecular dynamics simulations of helix unfolding (DiCapua et al., 1990; Tirado-Rives & Jorgensen, 1991; Soman et al., 1991) and initiation (Tobias & Brooks, 1991), with time scales of turn formation quite consistent with those that would be required for the initiation of protein folding.

Turn Conformation in TFE: A Helix Stop Signal. Addition of TFE to Mb-GH25 results in helix formation to a population of ~30%, as shown by an increase in the negative ellipticity at 208 and 222 nm. The plateau reached at high concentrations of TFE suggests that this is the maximum helicity possible under the conditions used. The difference in the maximal helicity measured by CD for Mb-GH25 and Mb-AA25 shows that the β-turn acts as a "helix stop signal", even in TFE. Similar behavior has been observed in the C-peptide of ribonuclease A, where residue 12 has been proposed to act as a helix stop signal both in aqueous solution (Kim & Baldwin, 1984), and in TFE (Nelson & Kallenbach, 1989). The behavior of each of the myoglobin peptides in TFE is qualitatively consistent with the predictions of helix probability by the method of Chou and Fasman (1978) shown in Table I. The substitution of the central Pro^{120}-Gly^{121} with Ala^{120}-Ala^{121} in Mb-AA25 gives a significantly enhanced probability of helix formation compared with that of Mb-GH25, which is reflected in a greater ellipticity at 222 nm in the CD spectrum (Figure 8) in TFE but not in water solution. The turn sequence is predicted to be a strong helix breaker; this is reflected in the behavior of Mb-GH5 and Mb-GH25 [and Mb-GH51; see Shin et al. (1993)] in water and TFE solutions. It is clear that TFE neither induces helix in this sequence nor influences the population of the turn conformation. This is consistent with previous observations (Dyson et al., 1992b; Sönnichsen et al., 1992) on the effect of TFE on peptide conformation. Helix is induced in a peptide sequence only in regions for which there is a significant preference for helical conformations; in the absence of such a preference, no evidence for secondary structure formation, helix or otherwise, has been found. Although stabilization of turn conformations in peptides by TFE has been previously

reported (Cann et al., 1987; Siligardi et al., 1987), we find no evidence for any effect of TFE on turn stability for the myoglobin peptides.

CONCLUSION

The amino acid sequence connecting the G- and H-helices of myoglobin, His^{119}-Pro^{120}-Gly^{121}-Asp^{122}, has a high propensity to adopt a reverse turn conformation both in water solution and at high concentrations of the helix-stabilizing solvent, TFE. The turn propensity is unaltered in peptides varying widely in length. Studies of a peptide in which the central residues of the turn are replaced with alanine indicate that the turn may function as a helix stop signal, limiting helix propagation at the C-terminal end of the G-helix. The G–H turn region of the myoglobin sequence may thus play a role in the earliest protein folding initiation events as well as function as a helix stop signal during the folding process. The G-helix is stabilized to a much greater extent than the N-terminal segment of the H-helix by TFE. Addition of TFE does not affect turn formation for residues 119–122 and thus appears to stabilize helix only in regions of the peptide which have high intrinsic helical propensity.

ACKNOWLEDGMENT

We thank Linda Tennant and Steve Schiebel for excellent technical assistance.

REFERENCES

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 63*, 207.
Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.
Blanco, F. J., Herranz, J., González, C., Jiménez, M. A., Rico, M., Santoro, J., & Nieto, J. L. (1992) *J. Am. Chem. Soc. 114*, 9676–9677.
Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.
Bothner-By, A. A., Stephens, R. L., Lee, J., Warren, C. D., & Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811–813.
Brahms, S., & Brahms, J. (1980) *J. Mol. Biol. 138*, 149–178.
Cann, J. R., London, R. E., Unkefer, C. J., Vavrek, R. J., & Stewart, J. M. (1987) *Int. J. Pept. Protein Res. 29*, 486–496.
Chandrasekhar, K., Profy, A. T., & Dyson, H. J. (1991) *Biochemistry 30*, 9187–9194.
Chou, P. Y., & Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135–175.
Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. 47*, 45–147.
Dalgarno, D. C., Levine, B. A., & Williams, R. J. P. (1983) *Biosci. Rep. 3*, 443–452.
DiCapua, F. M., Swaminathan, S., & Beveridge, D. L. (1990) *J. Am. Chem. Soc. 112*, 6769–6771.
Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985) *Nature (London) 318*, 480–483.
Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988a) *J. Mol. Biol. 201*, 161–200.
Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201–217.
Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., & Wright, P. E. (1992a) *J. Mol. Biol. 226*, 795–817.
Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A., & Wright, P. E. (1992b) *J. Mol. Biol. 226*, 819–835.
Hanson, J. C., & Schoenborn, B. P. (1981) *J. Mol. Biol. 153*, 117–146.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kent, S. B. H. (1988) *Annu. Rev. Biochem. 57*, 957–989.
Kim, P. S., & Baldwin, R. L. (1984) *Nature (London) 307*, 329–334.
Kuntz, I. D., Kosen, P. A., & Craig, E. C. (1991) *J. Am. Chem. Soc. 113*, 1406–1408.
Kuriyan, J., Wilz, S., Karplus, M., & Petsko. G. A. (1986) *J. Mol. Biol. 192*, 133–154.
Lewis, P. N., Momany, F. A., & Scheraga, H. A. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2293–2297.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Nelson, J. W., & Kallenbach, N. R. (1989) *Biochemistry 28*, 5256–5261.
Peña, M. C., Rico, M., Jiménez, M. A., Herranz, J., Santoro, J., & Nieto, J. L. (1989) *Biochim. Biophys. Acta 957*, 380–389.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Rey, A., & Skolnick, J. (1991) *Chem. Phys. 158*, 199–219.
Sears, D. W., & Beychok, S. (1973) in *Physical Principles and Techniques of Protein Chemistry, Part C* (Leach S. J., Ed.) pp 445–593, Academic Press, New York.
Shin, H. C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6356–6364.
Siligardi, G., Drake, A. F., Mascagni, P., Neri, P., Lozzi, L., Niccolai, N., & Gibbons, W. A. (1987) *Biochem. Biophys. Res. Commun. 143*, 1005–1011.
Skolnick, J., & Kolinski, A. (1990) *Science 250*, 1121–1125.
Skolnick, J., & Kolinski. A. (1991) *J. Mol. Biol. 221*, 499–531.
Soman, K. V., Karimi, A., & Case, D. A. (1991) *Biopolymers 31*, 1351–1361.
Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S., & Sykes, B. D. (1992) *Biochemistry 31*, 8790–8798.
Stroup, A. N., & Gierasch, L. M. (1990) *Biochemistry 29*, 9765–9771.
Sundaralingam, M., & Sekharudu, Y. C. (1989) *Science 244*, 1333–1337.
Takano, T. (1977a) *J. Mol. Biol. 110*, 537–568.
Takano, T. (1977a) *J. Mol. Biol. 110*, 569–584.
Tirado-Rives, J., & Jorgensen, W. L. (1991) *Biochemistry 30*, 3864–3871.
Tobias, D. J., & Brooks, C. L. (1991) *Biochemistry 30*, 6059–6070.
Tobias, D. J., Mertz, J. E., & Brooks, C. L. (1991) *Biochemistry 30*, 6054–6058.
VanGeet, A. L. (1969) *Anal. Chem. 40*, 2227–2229.
Wagner, G., Neuhaus, D., Wörgötter, E., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 131–135.
Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6337–6347.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311–333.
Woody, R. W. (1974) in *Peptides, Polypeptides and Proteins* (Blout, E. R., Bovey, F. A., Goodman, M., & Lotan, N. Eds) pp 338–348, Wiley, New York.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167–7175.
Wright, P. E., Dyson, H. J., Feher, V. A., Tennant, L. L., Waltho, J. P., Lerner, R. A., & Case, D. A. (1990) in *Frontiers of NMR in Molecular Biology, UCLA Symposium on Molecular Biology*, new series (Live, D., Armitage, I. M., & Patel, D., Eds.) Alan R. Liss Inc., New York, Vol. 109, pp 1–13.
Wüthrich, K. *NMR of Proteins and Nucleic Acids* (1986) John Wiley & Sons, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.
Yu, M. H., & King, J. (1988) *J. Biol. Chem. 263*, 1424–1431.
Zhou, N. E., Zhu, B.-Y., Sykes, B. D., & Hodges, R. S. (1992) *J. Am. Chem. Soc. 114*, 4320–4326.
Zimmerman, S. S., & Scheraga, H. A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4126–4129.