Protein Science (1997), 6:706–716. Cambridge University Press. Printed in the USA.

Folding propensities of peptide fragments of myoglobin

MARTINE T. REYMOND,[1] GENE MERUTKA,[2] H. JANE DYSON,[1] AND PETER E. WRIGHT[1]

[1]Department of Molecular Biology and Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

(RECEIVED November 27, 1996; ACCEPTED January 2, 1997)

Abstract

Myoglobin has been studied extensively as a paradigm for protein folding. As part of an ongoing study of potential folding initiation sites in myoglobin, we have synthesized a series of peptides covering the entire sequence of sperm whale myoglobin. We report here on the conformational preferences of a series of peptides that cover the region from the A helix to the FG turn. Structural propensities were determined using circular dichroism and nuclear magnetic resonance spectroscopy in aqueous solution, trifluoroethanol, and methanol. Peptides corresponding to helical regions in the native protein, namely the B, C, D, and E helices, populate the α region of (ϕ,ψ) space in water solution but show no measurable helix formation except in the presence of trifluoroethanol. The F-helix sequence has a much lower propensity to populate helical conformations even in TFE. Despite several attempts, we were not successful in synthesizing a peptide corresponding to the A-helix region that was soluble in water. A peptide termed the AB domain was constructed spanning the A- and B-helix sequences. The AB domain is not soluble in water, but shows extensive helix formation throughout the peptide when dissolved in methanol, with a break in the helix at a site close to the A-B helix junction in the intact folded myoglobin protein. With the exception of one local preference for a turn conformation stabilized by hydrophobic interactions, the peptides corresponding to turns in the folded protein do not measurably populate β-turn conformations in water, and the addition of trifluoroethanol does not enhance the formation of either helical or turn structure. In contrast to the series of peptides described here, earlier studies of peptides from the GH region of myoglobin show a marked tendency to populate helical structures (H), nascent helical structures (G), or turn conformations (GH peptide) in water solution. This region, together with the A-helix and part of the B-helix, has been shown to participate in an early folding intermediate. The complete analysis of conformational properties of isolated myoglobin peptides supports the hypothesis that spontaneous secondary structure formation in local regions of the polypeptide may play an important role in the initiation of protein folding.

Keywords: myoglobin; nuclear magnetic resonance; peptide conformation; protein folding; reverse turn; secondary structure

The mechanism by which proteins fold into their native conformations remains one of the central unsolved problems in molecular biology. One view is that folding occurs along defined pathways, often with structured intermediates (for reviews, see Creighton, 1988; Wright et al., 1988; Baldwin, 1989; Kuwajima, 1989; Montelione & Scheraga, 1989; Kim & Baldwin, 1990; Jaenicke, 1991; Ptitsyn, 1991). An alternative view is that of the folding funnel, where a multitude of unstructured conformations coalesces into a smaller population of partly folded conformations that finally proceed through a few defined intermediates to the folded state (Bryngelson et al., 1995). For many proteins, compact intermediates have been identified during the later stages of the folding process. Much progress has been made in characterization of these intermediates, which are formed on a millisecond to second time scale (Matthews, 1993; Dyson & Wright, 1996). However, little is currently known about events that precede the formation of compact intermediates, i.e., about the processes involved in initiation of protein folding. To gain insights into the earliest events of protein folding, thought to involve transient formation of rudiments of secondary structure and local hydrophobic clusters (Wright et al., 1988), the spontaneous propensities of peptide fragments to adopt local structures in aqueous solution have been examined (reviewed in Dyson & Wright, 1991). The secondary structural elements found in peptides, studied under conditions where they are devoid of tertiary interactions, are in rapid dynamic equilibrium with unfolded states (Wright et al., 1988; Dyson & Wright, 1991). In

Reprint requests to: Peter E. Wright or H. Jane Dyson, Department of Molecular Biology MB2, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037; email: wright@scripps.edu; dyson@scripps.edu.

[2]Present address: Trimeris Inc., 4727 University Drive, Durham, North Carolina 27707.

Abbreviations: CD, circular dichroism; HPLC, high-performance liquid chromatography; Mb, myoglobin; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; 2Q, double quantum; 2QF COSY, double quantum filtered scalar correlated spectroscopy; ROESY, rotating-frame NOESY; TFE, 2,2,2-trifluoroethanol; TOCSY, two-dimensional total correlated spectroscopy.

previous studies, we have determined the conformational propensities of peptide fragments spanning the entire sequence of the all-β protein plastocyanin (Dyson et al., 1992b) and the helical protein myohemerythrin (Dyson et al., 1992a). Here we describe the extension of this approach to apomyoglobin, the folding pathway of which has been extensively characterized by both experimental and theoretical methods (Bashford et al., 1987; Hughson et al., 1990; Jennings & Wright, 1993)

Kinetic studies of apomyoglobin folding show an initial association of the A, G, and H helices, followed by stabilization of secondary structure in the B helix and then the rest of the protein (Jennings & Wright, 1993). According to recent temperature-jump experiments, secondary structure formation and compaction of the chain in these regions occurs on a sub-millisecond time scale (Ballew et al., 1996). The A, G, and H helices are also folded and packed within a compact molten globule state formed at equilibrium at pH 4 (Hughson & Baldwin, 1989; Hughson et al., 1990). Studies on the conformation of several peptide fragments of myoglobin have already been carried out (Epand & Scheraga, 1968; Hermans & Puett, 1971; Waltho et al., 1989) and an extensive study has been made of the conformational preferences of synthetic peptides of the GH region (Shin et al., 1993a, 1993b; Waltho et al., 1993). We present here the continuation of these studies with the aim of examining peptide fragments that together constitute the complete sequence of apomyoglobin to obtain insights into the location of potential folding initiation sites. A series of peptides was synthesized to span the entire myoglobin sequence from the A helix to the FG turn, residues 4 to 99 (Table 1), complementing our earlier work on the GH helical hairpin region (Shin et al., 1993a, 1993b; Waltho et al., 1993). As in previous studies (Dyson et al., 1992a, 1992b), the peptides were designed to represent elements of secondary structure and their conformational preferences were examined by nuclear magnetic resonance (NMR) spectroscopy and circular dichroism (CD) spectroscopy.

Results

Peptide selection and solubility

Peptides corresponding to the A, B, D, E, and F helices and the EF and FG turns were synthesized blocked at the N- and C-termini with acetyl and amide groups, respectively. Neither the A-helix peptide nor a longer peptide corresponding to the A and B helices connected by the AB turn sequence (the AB-domain peptide) was soluble in water, but the conformational propensities of both could be studied by circular dichroism spectroscopy in methanol solution and in water/TFE mixtures. Because the C helix is so short, it was represented by a peptide spanning the C helix and the CD turn (the CCD-domain peptide). The BC-turn peptide includes the C-terminus of the B helix and the N-terminus of the C helix. Apart from the peptides containing the A helix, all peptides were water soluble. For each peptide, a series of one-dimensional NMR spectra was acquired for samples at a number of different concentrations (from 8 to 1 mM). The shortest peptides showed no concentration dependence and could be used at quite high concentrations (BC-turn peptide 7.5 mM; EF-turn peptide 7.5 mM; FG-turn peptide 7.5 mM), but the longer ones had to be studied at lower concentrations (B-helix peptide 1 mM; CCD-domain peptide 4 mM; D-helix peptide 2 mM; E-helix peptide 2 mM; F-helix peptide 2 mM) to avoid complications due to aggregation. The NMR spectra of the AB-domain peptide were acquired in deuterated methanol at 0.5 mM. The peptide sequences are shown in Table 1.

CD of monomeric species

A series of ultraviolet CD spectra were acquired in water and water/TFE mixtures at low peptide concentration (10 to 15 μM). The dependence of the CD spectra of the peptides on TFE concentration in water/TFE mixtures is shown in Figure 1. None of the peptides corresponding to the helical segments of myoglobin,

Table 1. *Amino acid sequence of sperm whale myoglobin showing secondary structure and the sequences of the peptides*

	1 5 10 15 20 25 30 35 40 45 50
Mb Sequence	V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P E T L E K F D R F K H L K
2° Structure	A Helix | B Helix | C Helix | CD Turn
Peptides	
A-Helix	Ac-V L S E G E W Q L V L H V W A K V E A-NH_2
AB-Domain	Ac-E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K-NH_2
B-Helix	Ac-D V A G H G Q D I L I R L F K S-NH_2
BC-Turn	Ac-K S H P E T-NH_2
CCD-Domain	Ac-H P E T L E K F D R F K H L K-NH_2
	51 55 60 65 70 75 80 85 90 95 100
	T E A E M K A S E D L K K H G V T V L T A L G A I L K K K G H H E A E L K P L A Q S H A T K H K I P
	D Helix | E Helix | EF Turn | F Helix | FG Turn
D-Helix	Ac-T E A E M K A
E-Helix	Ac-S E D L K K H G V T V L T A L G A I L K-NH_2
EF-Turn	Ac-K K G H H E A E-NH_2
F-Helix	Ac-E L K P L A Q S H A-NH_2
FG-Turn	Ac-A T K H K I P-NH_2
	101 105 110 115 120 125 130 135 140 145 150
	I K Y L E F I S 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 A L E L F R K D I A A K Y K E L G Y Q G
G/H Helices[a]	G Helix | GH Turn | H Helix

[a]Studies performed previously (Shin et al., 1993b; Waltho et al., 1993).

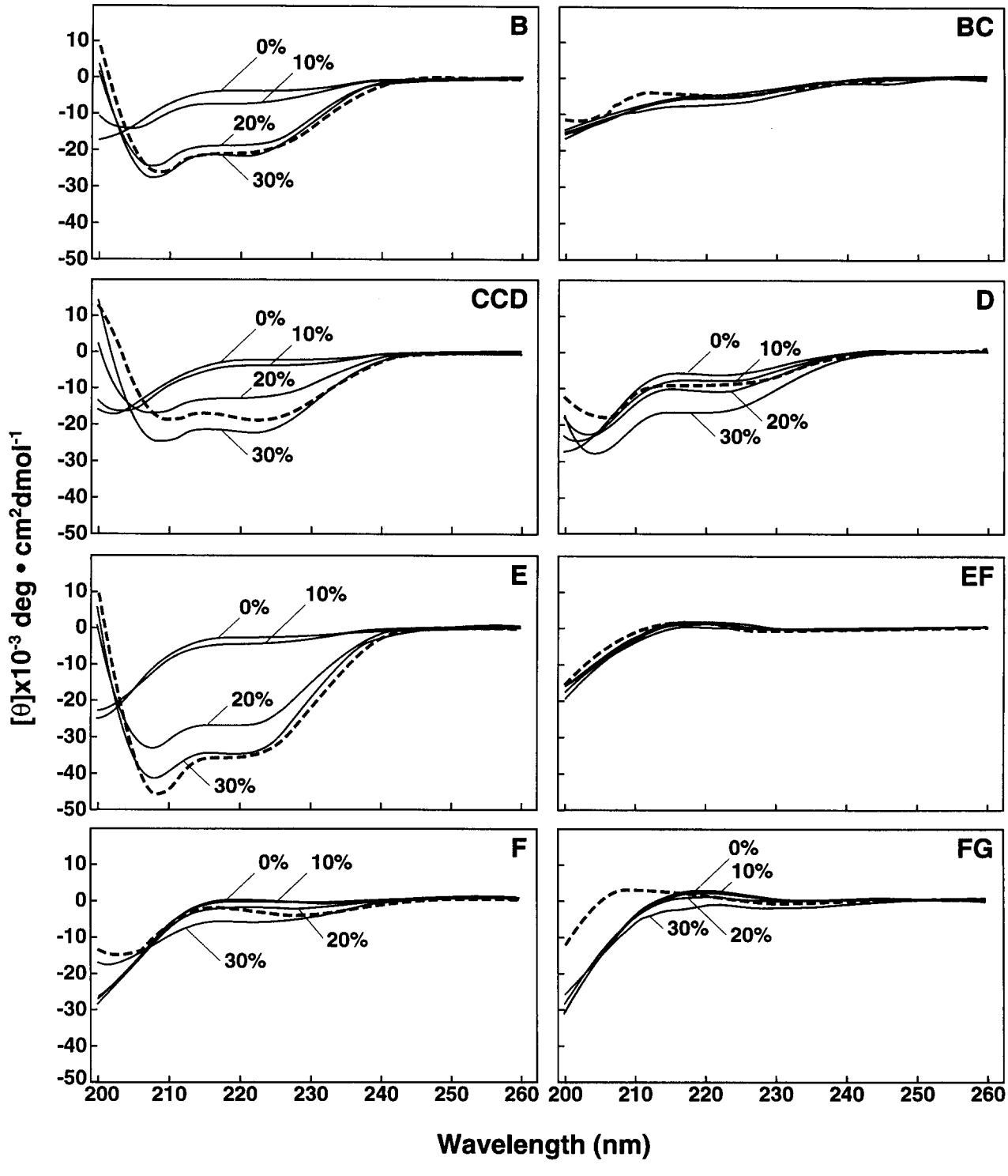

Fig. 1. Ultraviolet CD spectra of each of the peptides in various concentrations of trifluoroethanol (% v/v as marked) in 10 mM phosphate buffer pH 5 (solid lines) and in 100% methanol (dashed lines). Peptide concentrations were: B-helix peptide 15 μM, BC-turn peptide 12 μM, CCD-domain peptide 16 μM, D-helix peptide 12 μM, E-helix peptide 14 μM, EF-turn peptide 16 μM, F-helix peptide 15 μM, FG-turn peptide 10 μM.

i.e., the B-helix, the CCD-domain, the D-helix, the E-helix, and the F-helix peptides, shows the double minima at 208 and 222 nm characteristic of ordered helix in the absence of TFE. However, addition of a small amount of TFE results in an increase in the negative ellipticity at 222 nm, with maximal helix formation at about 30% TFE for the B-helix, D-helix, E-helix, and CCD-domain peptides. The E-helix peptide reaches the highest helical content of all peptides (Θ_{222} = −35,000 deg cm^2 dmol^{-1} in 30%

TFE), comparable with the values found for the G- and H-helix peptides (Shin et al., 1993a). The BC- and FG-turn peptides have CD spectra characteristic of unfolded conformations at all TFE concentrations, while the EF-turn peptide shows a slight positive ellipticity, which may be due to the formation of a turn. Interestingly, the F-helix peptide fails to significantly populate ordered helical conformations, even in TFE; we, therefore, conclude that binding to heme and packing interactions in the folded holoprotein play a major role in formation and stabilization of helix in this region of the polypeptide. Even in native apomyoglobin, the F-helix region does not adopt stable secondary structure but appears to fluctuate between folded and unfolded conformations (Eliezer & Wright, 1996).

The A- and B-helices are stabilized early in the folding pathway of apomyoglobin (Jennings & Wright, 1993) and there is, consequently, much interest in their secondary structural propensities. It is unfortunate, therefore, that all of the peptides that contain the A helix sequence are water insoluble and cannot be directly compared to the results for the rest of the molecule. In view of the potential interest of the results, considerable effort was expended to determine solution conditions where these peptides could be studied. The ultraviolet CD spectra of the A-helix and AB-domain peptides were recorded in methanol (Fig. 2); the CD spectra of the other peptides were also recorded in methanol for comparison (Fig. 1). The spectrum of the A-helix peptide shows evidence for a mixture of species (Fig. 2A), with an apparent predominance of β-structure as indicated by the ellipticity minimum at 215 nm. In TFE/water mixtures there is a decrease in the ellipticity at 215 nm, with the appearance of signals at <210 nm and >220 nm. The minima are not well defined, and are in any case of quite low intensity. These results indicate that a small amount of helix may be present in the A-helix peptide in TFE/water solutions, but the ellipticity is low and the results are complicated by the solubility problems of the peptide, especially at low concentrations of TFE.

The AB-domain peptide is highly helical in methanol with a minimum ellipticity of $-33{,}000$ deg cm^2 dmol^{-1} at 222 nm (Fig. 2B). The measurements made on the other peptides (Fig. 1) show that stabilization of helix by methanol is comparable to that of 30% TFE in 10 mM phosphate buffer, the conditions under which most of the peptides have reached their maximum helicity. The per-residue helical content of the AB-domain peptide appears to be greater in methanol than that of the B-helix peptide, reflecting a contribution of the A-helix sequence to the overall helix in the molecule.

Although the AB-domain peptide is completely insoluble in water, CD spectra could be measured at low peptide concentration in TFE/water mixtures (Fig. 2B). At the lowest concentrations of TFE, the CD data are unreliable due to difficulties associated with the low solubility of the peptide. However, at higher concentrations of TFE, the spectrum becomes more characteristic of helix and at 100% TFE approaches the signal of the methanol spectrum. Concentration-dependence measurements made for the AB-domain peptide in methanol indicate that the peptide is monomeric, even at high concentrations. The similarity in the shape of the CD spectra in methanol and at TFE concentrations >20% v/v indicates that the aggregated state of the peptide is dispersed at higher concentrations of TFE and that the peptide is monomeric.

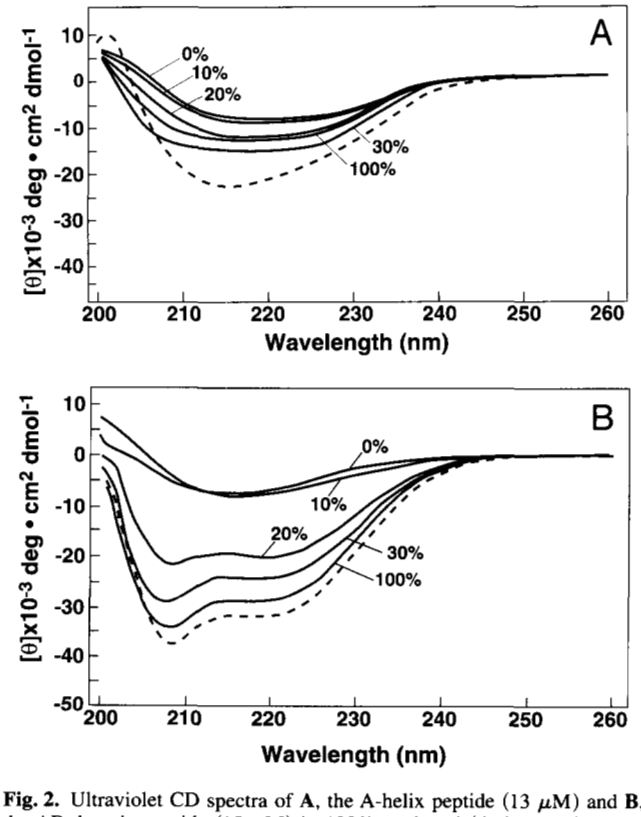

Fig. 2. Ultraviolet CD spectra of **A**, the A-helix peptide (13 μM) and **B**, the AB-domain peptide (15 μM) in 100% methanol (dashed line) and in mixtures of various concentrations (% v/v as marked) of trifluoroethanol in 10 mM phosphate buffer, pH 5 (solid lines).

NMR spectra

^{1}H NMR spectra were recorded for all peptides, except the A-helix peptide, in order to obtain site-specific information on conformational propensities in water solution. The A-helix peptide was not sufficiently soluble to permit NMR measurements. A representative NOESY spectrum, of the CCD-domain peptide, is shown in Figure 3, and a summary of the NOE connectivities for all peptides is given in Figure 4. The sequential $d_{\alpha N}(i,i+1)$ connectivities together with scalar connectivities observed in TOCSY experiments permit complete sequence-specific assignment to be made for all peptides; these assignments are summarized in Table 2. The interpretation of NOE connectivities to provide information about conformational preferences of small linear peptides that fluctuate rapidly over an ensemble of conformations in solution has been discussed in detail elsewhere (Wright et al., 1988; Dyson & Wright, 1991).

The peptides derived from helical segments of myoglobin (the B-, CCD-, D-, E-, and F-helix peptides) are largely unfolded in water solution, and in general no medium range $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+2)$, or $d_{\alpha\beta}(i,i+3)$ cross-peaks characteristic of folded helix or nascent helix (Dyson et al., 1988) are observed. However, for all of these peptides both $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ sequential NOE cross-peaks are observed, indicating that the peptides populate both the α- and β-regions of (ϕ,ψ) space. This is in contrast to peptides corresponding to β-strand regions of plastocyanin, which populate predominantly the β-strand regions of (ϕ,ψ) space (Dyson et al., 1992b). However, the propensities of the myoglobin peptides for helix formation are significantly lower than those of the A-, B-, C-, and D-helix peptides from myohemerythrin (Dyson et al., 1992a), all of which populate either ordered helical states or nascent helix in aqueous solution.

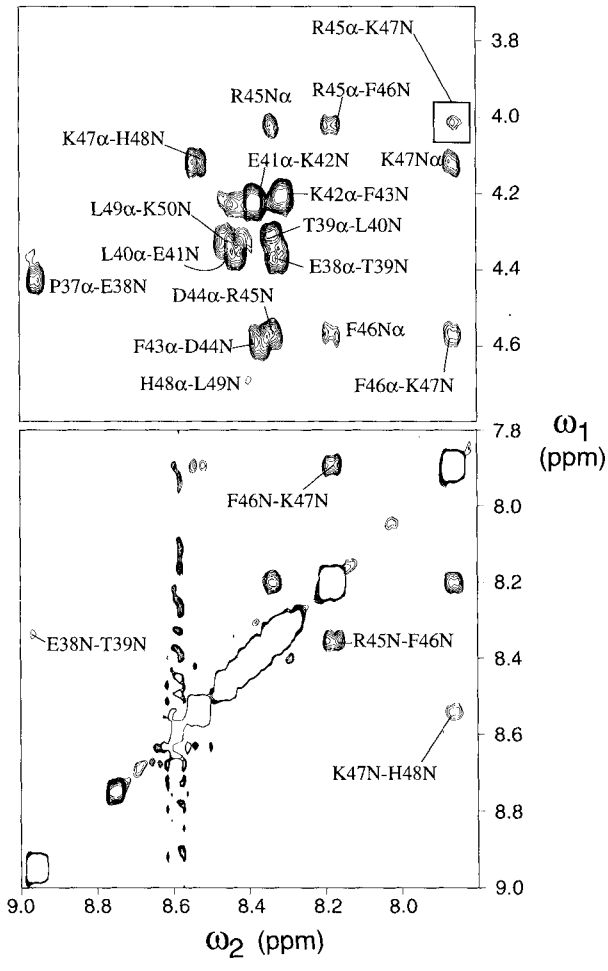

Fig. 3. Portion of a 500 MHz NOESY spectrum of the CCD-domain peptide in 10 mM phosphate pH 5, in 90% ^1H_2O/10% ^2H_2O at 278 K, showing NOEs between NH and C$^\alpha$H (upper panel) and between NH and NH (lower panel). The $d_{\alpha N}(i,i+2)$ NOE between Arg 45 and Lys 47 is shown plotted at a lower contour level in the box.

The peptides corresponding to interhelical segments (BC, EF, and FG) also show no measurable tendency to adopt folded structures in water solution. No $d_{NN}(i,i+1)$ cross-peaks were observed for the BC-turn or EF-turn peptides. The BC-turn peptide contains a proline at position 37; the NMR spectra show that this proline is predominantly *trans*, with a small proportion (10%) of the *cis* isomer. The spectrum of the EF-turn peptide, like that of the F-helix peptide, is heavily overlapped in the αN and NN regions. The FG-turn peptide has very weak $d_{NN}(i,i+1)$ cross peaks, indicating that the population of backbone conformations in the α region of (ϕ,ψ) space is small.

The NOESY spectrum of the AB-domain peptide in methanol shows short and medium range NOE connectivities characteristic of a highly populated helical secondary structure in methanol (Fig. 5), in accord with its helical CD spectrum. Medium range $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha N}(i,i+3)$ cross-peaks are present throughout the peptide, including connectivities (ambiguous due to overlap) between Val 17 and Ala 22 and a $d_{\alpha\beta}(i,i+3)$ connectivity between Val 21 and His 24 (Fig. 4). Although the presence of this unambiguous NOE suggests that the helix continues through the A-B connection, there are no other $d_{\alpha\beta}(i,i+3)$ or $d_{\alpha N}(i,i+3)$ NOEs in this region, and, together with evidence from the C$^\alpha$H chemical shifts, these results indicate that there is a break in the helix at this point. If this is the case, then the isolated Val 21–His 24 $d_{\alpha\beta}(i,i+3)$ connectivity may be indicative of a kink in the peptide. The C$^\alpha$H resonances are shifted significantly upfield from random coil values for most amino acids in the peptide (Fig. 6), consistent with formation of highly populated helical structure. From the mean C$^\alpha$H shifts, we estimate an average population of 75% helix from Trp 7 to Ala 22 and 57% helix between His 24 and Phe 33 (Merutka et al., 1993; Rizo et al., 1993). From both the C$^\alpha$H chemical shifts and the $d_{\alpha N}(i,i+3)$ NOEs, there appears to be a break in the helix at Ala 22 and Gly 23. This does not correspond exactly to the crystallographic estimate of the A/B helix boundary, which occurs from Ala 19–Asp 20 (Table 1), although it can be seen from Figure 6 that the upfield shifts, with the exception of Val 21, are generally quite small in this region, a further indication that the helix is not continuous throughout the peptide.

Temperature coefficients and hydrogen bonding

The temperature coefficients for the peptide amide protons in water solution are overall very high and reflect the general absence of secondary structure observed in the CD and NOESY spectra. There are three exceptions: Lys 47 in the CD-turn peptide (1.8 ppb/K), Glu 54 in the D-helix peptide (3.9 ppb/K), and Lys 63 in the E-helix peptide (4.3 ppb/K). The pattern of NOEs in the vicinity of Lys 47 indicates the presence of local structure. Strong $d_{NN}(i,i+1)$ NOEs are present between Arg 45 and Phe 46, and between Phe 46 and Lys 47 (Figs. 3 and 4), and a weak $d_{\alpha N}(i,i+2)$ NOE is observed between Arg 45 and Lys 47 (inset to Fig. 3). In addition, the C$^\alpha$H resonances of Arg 45 and Lys 47 are significantly upfield-shifted from random-coil values (Merutka et al., 1995), and the amide proton resonance of Lys 47 is also highly upfield-shifted (Fig. 3). These observations suggest that the amide proton of Lys 47 is protected from solvent (as shown by the exceptionally low temperature coefficient) by two structural elements, a hydrogen-bonded turn conformation between residues 44 and 47 and a local hydrophobic interaction mediated by the side chains of Lys 47 and Phe 46. Such an interaction has been previously observed for glycine amide protons (Dyson et al., 1992a). Possible close contacts between the side chains of Arg 45 and Lys 47, and that of Phe 46 are suggested by the observation of NOEs between the aromatic protons of the Phe and side chain protons corresponding to Lys and Arg resonances (data not shown). However, these contacts are not unambiguous, because there are two other lysine residues with similar chemical shifts in the CCD-domain peptide. The low temperature coefficient for Glu 54 does not appear from the pattern of NOEs to arise from the presence of structure; it possibly arises from a local interaction involving the glutamic acid side chain itself (Ebina & Wüthrich, 1984). No unusual NOEs are observed that could explain the lowered temperature coefficient of Lys 63 and the upfield shift relative to random coil of the amide proton resonance. These presumably reflect local interactions, perhaps of a hydrophobic nature, which cannot be identified by the present analysis.

Discussion

Implications for initiation of apomyoglobin folding

There is a striking correlation between the folding propensities of the isolated peptide fragments of myoglobin and the kinetic fold-

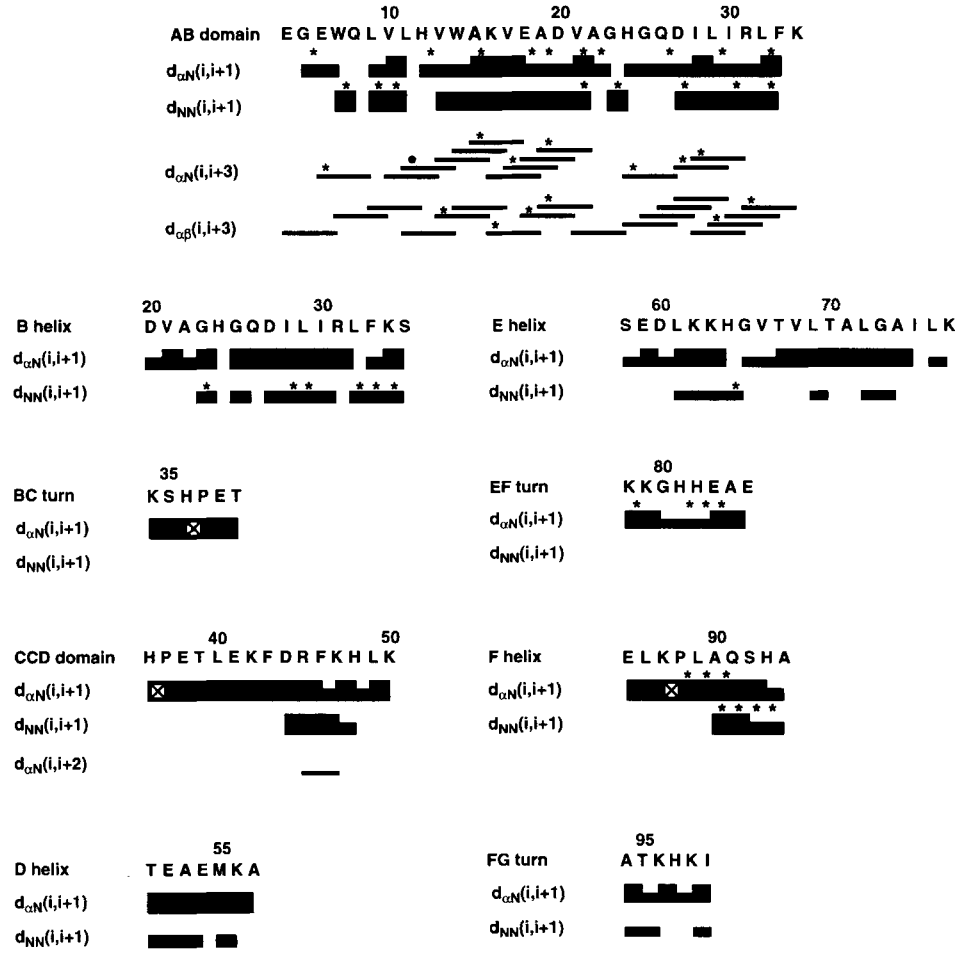

Fig. 4. Summary of the NOEs observed for each of the peptides. Asterisks indicate NOE connectivities that are ambiguous due to resonance overlap. Hatched boxes indicate NOE connectivities to proline $C^{\delta}H$ resonances in place of the NH.

ing pathway (Jennings & Wright, 1993). As discussed above, apomyoglobin rapidly forms a molten globule folding intermediate in which the A, G, and H helices, and possibly part of the B helix, are folded and stabilized. This intermediate is formed within the six millisecond dead-time of the quench-flow apparatus. The remainder of the B helix becomes stabilized on a slower time scale, followed by the C and E helices and the CD loop. Recent temperature-jump fluorescence studies of refolding of cold-denatured apomyoglobin (Ballew et al., 1996) indicate that a compact state that appears to correspond to the AGH intermediate is formed within 20 μs. This implies that the initial formation of secondary structure in the A-, G- and H-helices must be extremely rapid; indeed, it is quite likely that folding is initiated through rapid formation of helical conformations in one or more of these segments during the earliest stages of polypeptide chain collapse.

Extensive studies of the peptide sequences at the C-terminus of apomyoglobin, comprising the G-helix, the H-helix, and the GH turn, were carried out previously using CD and NMR spectroscopy (Shin et al., 1993a, 1993b; Waltho et al., 1993). The conformational ensemble of the 27-residue H-helix peptide was found to contain a high population of ordered helical forms, particularly in the center of the peptide, with fraying at the termini (Waltho et al., 1993). The 19-residue G-helix peptide showed only marginal preferences for helix formation in water, but was folded to a highly helical state at low concentration (20% v/v) of the helix-promoting solvent TFE (Shin et al., 1993a). In addition, the 5-residue GH-turn peptide was found to populate a turn conformation in aqueous solution (Shin et al., 1993b). This turn is retained in peptides of varying length (5, 25, and 51 residues). These results imply that initiation of folding in the intact protein can occur through formation of secondary structure in two locations, the H helix, and the GH turn. Further, the G-helix peptide can be readily induced to form helix either by the addition of TFE or through intermolecular association with the H-helix peptide in a four-helix bundle (Shin et al., 1993a). Thus there is a marked propensity for spontaneous secondary structure formation in the H-helix and GH-turn regions of the polypeptide. These structures are expected to be in rapid dynamic equilibrium with unfolded states. The rates of helix-coil transitions are on the order of 10^6–10^7 s^{-1} in homopolypeptides (Lumry et al., 1964; Cummings & Eyring, 1975), and molecular dynamics simulations suggest that turn formation is likely to be significantly faster (Soman et al., 1993; Hirst & Brooks, 1995). Thus secondary structure formation in the H-helix and GH-turn is expected to occur on a time scale that is significantly shorter than that required for collapse to form the AGH molten globule intermediate (~20 μs; Ballew et al., 1996).

The helical structure formed spontaneously by the H-helix peptide is only marginally stable (Waltho et al., 1993). Our previous

Table 2. *Resonance assignments and temperature coefficients of peptide fragments of myoglobin*

Peptides[a]	Amino acid residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Chemical shift (ppm at 278 K) $C^{\gamma}H$	$C^{\delta}H$	Other	$-\Delta\delta/\Delta T$ (ppb/K)
AB-domain	Glu 4	8.02[b]	4.31	2.18	2.61			n.d.
	Gly 5	8.35[b]	4.17					n.d.
	Glu 6	8.82[b]	4.25	2.15	2.25, 2.48			n.d.
	Trp 7	8.26[b]	4.52	3.43			7.30($C^{\delta 1}H$) 7.59($C^{\epsilon 3}H$)	n.d.
						7.10($C^{\zeta 3}H$)	7.20($C^{\eta}H$) 7.43($C^{\zeta 2}H$)	
	Gln 8	8.35	3.91	2.14	2.31, 2.49			n.d.
	Leu 9	7.98	4.21	1.90, 2.01	1.79	1.02, 1.06		n.d.
	Val 10	8.25	3.55	2.30	0.91, 1.18			n.d.
	Leu 11	8.37	4.16	1.88, 1.90	1.85	0.88, 0.93		n.d.
	His 12	8.69[b]	4.46	3.51, 3.58		8.99	7.55($C^{\epsilon}H$)	n.d.
	Val 13	8.62	3.68	2.42	1.00, 1.22			n.d.
	Trp 14	8.84	4.38	3.45, 3.48			7.26($C^{\delta 1}H$) 7.64($C^{\epsilon 3}H$)	n.d.
						7.08($C^{\zeta 3}H$)	7.20($C^{\eta}H$) 7.43($C^{\zeta 2}H$)	
	Ala 15	8.68	4.06	1.66				n.d.
	Lys 16	8.02	4.09	2.05, 2.23	1.55	1.88	3.06($C^{\epsilon}H$)	n.d.
	Val 17	8.55	3.70	2.36	1.00, 1.10			n.d.
	Glu 18	8.77	4.02	2.03, 2.12	2.37			n.d.
	Ala 19	8.48	4.17	1.64				n.d.
	Asp 20	8.60	4.56	2.96, 3.34				n.d.
	Val 21	8.81	3.75	2.36	1.10, 1.21			n.d.
	Ala 22	8.82	4.25	1.62				n.d.
	Gly 23	7.98	4.22					n.d.
	His 24	8.28	4.62	3.53		8.96	7.62($C^{\epsilon}H$)	n.d.
	Gly 25	8.07[b]	3.82					n.d.
	Gln 26	8.28[b]	4.12	2.26, 2.42	2.46, 2.60			n.d.
	Asp 27	8.62[b]	4.49	2.96, 3.23				n.d.
	Ile 28	8.05	3.71	2.20	1.21, 1.80	0.99	0.91($C^{\gamma}H_3$)	n.d.
	Leu 29	8.12	4.12	2.11	1.98	1.02, 1.03		n.d.
	Ile 30	8.61	3.81	2.11	1.25, 2.03	1.05	1.02($C^{\gamma}H_3$)	n.d.
	Arg 31	8.07	4.07	2.00, 2.07	1.58, 1.72	3.24		n.d.
	Leu 32	8.36	4.15	1.66, 1.79	1.43	0.87		n.d.
	Phe 33	8.28	4.60	3.23, 3.32		7.30	7.35,7.47($C^{\epsilon}H$,$C^{\zeta}H$)	n.d.
	Lys 34	8.22	4.39	2.06	1.60	1.72, 1.76	3.02($C^{\epsilon}H$)	n.d.
B-helix	Asp 20	8.45	4.59	2.61, 3.67				9.1
	Val 21	8.32	4.09	2.07	0.91			10.0
	Ala 22	8.53	4.27	1.37				9.9
	Gly 23	8.46	3.91					8.7
	His 24	8.51	4.75	3.16, 3.35		8.60	7.28($C^{\epsilon}H$)	7.3
	Gly 25	8.72	3.95					8.7
	Gln 26	8.47	4.32	1.97, 2.09	2.34			6.9
	Asp 27	8.52	4.60	2.64, 2.71				5.5
	Ile 28	8.12	4.07	1.86	1.18, 1.46	0.87	0.82($C^{\gamma}H_3$)	11.4
	Leu 29	8.24	4.30	1.63	1.69	0.91, 0.84		10.5
	Ile 30	8.13	4.06	1.82	1.18, 1.46	0.86	0.83($C^{\gamma}H_3$)	9.2
	Arg 31	8.35	4.25	1.73	1.55	3.15	7.24($N^{\eta}H$), 6.44($N^{\epsilon}H$)	9.0
	Leu 32	8.25	4.30	1.56	1.44	0.89	0.81($C^{\gamma}H_3$)	9.0
	Phe 33	8.33	4.58	2.64, 3.06		7.23	7.30,7.32($C^{\epsilon}H$,$C^{\zeta}H$)	11.2
	Lys 34	8.34	4.27	1.72	1.38	1.65	2.96($C^{\epsilon}H$), 7.59($N^{\zeta}H$)	11.2
	Ser 35	8.33	4.33	3.87				9.2
BC-turn	Lys 34	8.43	4.29	1.80, 1.86	1.45	1.70	2.99($C^{\epsilon}H$)	9.0
	Ser 35	8.47	4.37	3.81				9.2
	His 36	8.59	5.02	3.15, 3.25		8.59	7.31($C^{\epsilon}H$)	8.7
	Pro 37		4.44	2.30	1.99	3.72, 3.63		
	Glu 38	9.05	4.38	1.99, 2.11	2.26, 2.36			8.0
	Thr 39	8.29	4.32	4.29	1.21			9.2
CCD-domain	His 36	8.53	4.69	3.15, 3.26		8.60	7.30($C^{\epsilon}H$)	8.7
	Pro 37		4.43	2.26, 2.30	1.99	3.62, 3.74		
	Glu 38	8.96	4.37	2.03, 2.11	2.41			6.4
	Thr 39	8.34	4.31	4.22	1.21			9.7
	Leu 40	8.35	4.36	1.62	1.62	0.92	0.86($C^{\gamma}H_3$)	9.2
	Glu 41	8.44	4.22	1.97	2.33			8.0
	Lys 42	8.39	4.22	1.66	1.31	1.68	2.94($C^{\epsilon}H$), 7.60($N^{\zeta}H$)	10.5
	Phe 43	8.32	4.60	2.31, 3.09		7.19	7.27, 7.27($C^{\epsilon}H$,$C^{\zeta}H$)	11.0
	Asp 44	8.38	4.56	2.67				6.7
	Arg 45	8.34	4.02	1.29, 1.58	1.11	3.03	7.16($N^{\eta}H$) 6.56,6.81($N^{\epsilon}H$)	8.7

(*continued*)

Table 2. *Continued*

Peptides[a]	Amino acid residue	NH	$C^{\alpha}H$	$C^{\beta}H$	Chemical shift (ppm at 278 K) $C^{\gamma}H$	$C^{\delta}H$	Other	$-\Delta\delta/\Delta T$ (ppb/K)
CCD-domain (*continued*)	Phe 46	8.18	4.56	2.93, 3.23		7.20	7.27,7.27($C^{\epsilon}H$,$C^{\zeta}H$)	5.5
	Lys 47	7.89	4.11	1.62	1.32	1.72	2.94($C^{\epsilon}H$), 7.60($N^{\zeta}H$)	1.8
	His 48	8.53	4.69	3.15, 3.26		8.60	7.30($C^{\epsilon}H$)	6.7
	Leu 49	8.41	4.32	1.60	1.60	0.91	0.85($C^{\gamma}H_3$)	9.9
	Lys 50	8.46	4.22	1.79, 1.91	1.44	1.68	2.94($C^{\epsilon}H$), 7.60($N^{\zeta}H$)	9.4
D-helix	Thr 51	8.28	4.28	4.28	1.21			7.7
	Glu 52	8.69	4.24	2.02, 2.06	2.40			8.2
	Ala 53	8.42	4.19	1.37				8.2
	Glu 54	8.22	4.21	2.01, 2.04	2.38			3.9
	Met 55	8.39	4.39	2.02, 2.06	2.55, 2.58			9.2
	Lys 56	8.32	4.27	1.80	1.45	1.66	2.98($C^{\epsilon}H$), 7.59($N^{\zeta}H$)	6.4
	Ala 57	8.27	4.23	1.39				6.4
E-helix	Ser 58	8.45	4.36	3.86				10.0
	Glu 59	8.74	4.25	1.96, 2.08	2.33			7.0
	Asp 60	8.42	4.53	2.62, 2.75				5.7
	Leu 61	8.22	4.22	1.84	1.62	0.92, 0.82		8.2
	Lys 62	8.07	4.20	1.74	1.35, 1.41	1.64	2.96($C^{\epsilon}H$), 7.58($N^{\zeta}H$)	6.9
	Lys 63	8.16	4.20	1.80	1.38, 1.43	1.66	2.98($C^{\epsilon}H$), 7.58($N^{\zeta}H$)	4.3
	His 64	8.51	4.70	3.16, 3.28		8.61	7.31($C^{\epsilon}H$)	8.7
	Gly 65	8.51	3.97					7.0
	Val 66	8.22	4.19	2.07	0.92			8.2
	Thr 67	8.50	4.34	4.09	1.16			9.4
	Val 68	8.46	4.08	2.03	0.92			11.4
	Leu 69	8.53	4.43	1.60	1.60	0.84, 0.91		9.1
	Thr 70	8.24	4.28	4.20	1.20			9.4
	Ala 71	8.50	4.29	1.38				10.0
	Leu 72	8.23	4.08	1.64	1.64	0.82, 0.88		10.0
	Gly 73	8.45	3.88					8.7
	Ala 74	8.10	4.29	1.35				6.4
	Ile 75	8.33	4.28	1.76	1.62	0.87	0.92($C^{\gamma}H_3$)	10.0
	Leu 76	8.45	4.35	1.59	1.59	0.85, 0.91		10.0
	Lys 77	8.39	4.23	1.80	1.40, 1.46	1.68	2.98($C^{\epsilon}H$), 7.58($N^{\zeta}H$)	9.6
EF-turn	Lys 78	8.44	4.22	1.71, 1.73	1.40	1.64	2.96($C^{\epsilon}H$)	8.0
	Lys 79	8.58	4.32	1.63, 1.76	1.44	1.64	2.96($C^{\epsilon}H$)	8.7
	Gly 80	8.58	3.91					7.5
	His 81	8.50	4.64	3.11, 3.23		8.55	7.23($C^{\epsilon}H$)	6.4
	His 82	8.76	4.64	3.11, 3.23		8.53	7.21($C^{\epsilon}H$)	7.0
	Glu 83	8.78	4.21	1.95, 2.07	1.91, 2.30			6.4
	Ala 84	8.59	4.25	1.39				8.7
	Glu 85	8.44	4.21	1.97, 2.10	1.95, 2.31			8.0
F-helix	Glu 85	8.48	4.18	1.88, 1.93	2.28			10.5
	Leu 86	8.53	4.34	1.63	1.63	0.85, 0.91		10.5
	Lys 87	8.48	4.60	1.77	1.45	1.68	2.98($C^{\epsilon}H$), 7.60($N^{\zeta}H$)	10.5
	Pro 88		4.37	1.87, 2.26	1.99	3.61, 3.79		
	Leu 89	8.52	4.26	1.60	1.60	0.88, 0.92		9.9
	Ala 90	8.45	4.27	1.36				8.7
	Gln 91	8.47	4.31	1.95, 2.08	2.35		6.98, 7.65($N^{\epsilon}H$)	11.1
	Ser 92	8.46	4.37	3.83				9.3
	His 93	8.68	4.72	3.27, 3.19		8.60	7.29($C^{\epsilon}H$)	9.2
	Ala 94	8.45	4.27	1.37				9.4
FG-turn	Ala 94	8.48	4.29	1.38				9.4
	Thr 9	8.30	4.28	4.18	1.17			9.9
	Lys 96	8.39	4.27	1.76	1.42	1.67	2.97($C^{\epsilon}H$), 7.59($N^{\zeta}H$)	9.9
	His 97	8.63	4.66	3.23, 3.13		8.59	7.27($C^{\epsilon}H$)	10.3
	Lys 98	8.58	4.30	1.75	1.41	1.66	2.98($C^{\epsilon}H$), 7.58($N^{\zeta}H$)	9.2
	Ile 99	8.47	4.10	1.83	1.21, 1.52	0.88	0.94($C^{\gamma}H_3$)	11.7

[a] AB-domain peptide N-terminal CH_3 and C-terminal NH_2 are not assigned; B-helix peptide N-terminal CH_3 2.00, C-terminal NH_2 6.94, 7.65; BC-turn peptide N-terminal CH_3 2.04, C-terminal NH_2 7.68, 7.30; CCD-domain peptide N-terminal CH_3 1.95, C-terminal NH_2 7.21, 7.66; D-helix peptide N-terminal CH_3 2.09, C-terminal NH_2 7.59, 7.13; E-helix peptide N-terminal CH_3 2.02, C-terminal NH_2 7.62, 7.21; EF-turn peptide N-terminal CH_3 1.99, C-terminal NH_2 7.70, 7.19; F-helix peptide N-terminal CH_3 2.00, C-terminal NH_2 7.75, 7.17; FG-turn peptide N-terminal CH_3 2.02, C-terminal NH_2 7.86, 7.24.

[b] These amide protons were assigned using the Hahn-Echo NOESY spectrum of the AB domain. Concentration and pH of solutions: AB-domain peptide 0.5 mM (methanol); B-helix peptide 1 mM pH 4.3; BC-turn peptide 7.5 mM pH 5.0; CCD-domain peptide 4 mM pH 5.0; D-helix peptide 4 mM pH 4.3; E-helix peptide 2 mM pH 5.0; EF-turn peptide 7.5 mM pH 4.5; F-helix peptide 2 mM pH 4.3; FG-turn peptide 7.5 mM pH 4.3.

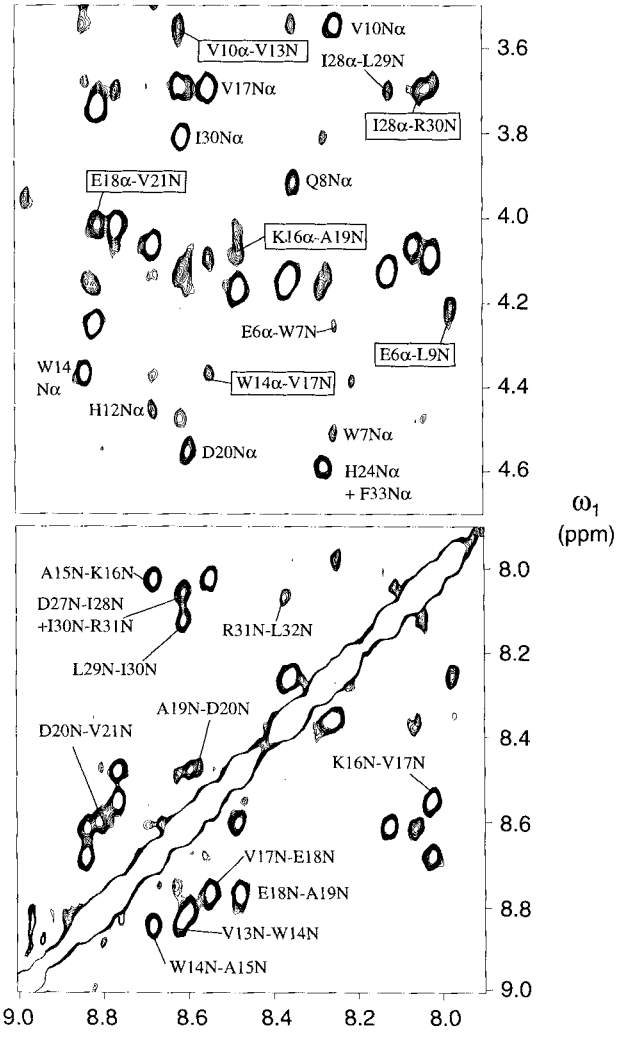

Fig. 5. Portion of the 500 MHz NOESY spectrum of the AB-domain in 100% methanol, showing NOEs between NH and $C^{\alpha}H$ (upper panel) and between NH and NH (lower panel). Representative NOE connectivities are labeled. Boxed labels refer to medium-range NOEs.

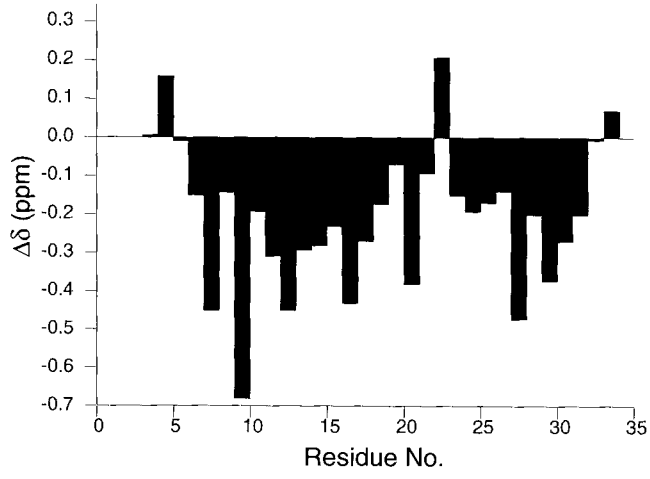

Fig. 6. Plot of the difference in the observed chemical shift for the $C^{\alpha}H$ for the AB-domain peptide in 100% methanol and the "random coil" chemical shift obtained from a series of peptides GGXGG (Merutka et al., 1995). Values for the $C^{\alpha}H$ chemical shifts of these "random" peptides were invariant in water and in TFE (Merutka et al., 1995) and are therefore likely to remain the same in methanol.

studies of peptide fragments corresponding to the G- and H-helices of myoglobin clearly implicate long-range packing interactions, presumably involving the A-helix, in stabilization of helical structure in the AGH molten globule (Shin et al., 1993b). The intrinsic conformational propensities of the A-helix region are therefore of great interest. Although a great deal of effort was expended in design and synthesis of this peptide, it proved extremely difficult to study due to its insolubility in water and most other solvents. Attempts to form complexes by mixing the 51-residue GH-hairpin peptide and the A-helix peptide were not successful, probably due to the insolubility and strong tendency towards aggregation of the A-helix peptide. Attempts to form the complex through specific formation of disulfide bonds such as had been used for the G- and H-peptides (Shin et al., 1993a) were also frustrated by low solubility and the difficulty of characterizing and purifying the peptide and the complex. We were finally able to form a construct of the A helix together with the B-helix sequence and the linker AB sequence. Although insoluble in water, this peptide is soluble in methanol, and we were able to determine its conformational preferences under those conditions. A comparison of the CD spectra of the other component peptides of apomyoglobin (Fig. 1) in water and methanol shows that the effect of methanol is similar to that of TFE; hence, the observation of helical conformational preferences in the AB-domain peptide in methanol cannot be taken as a direct implication of helical structure in water. While the AB peptide is strongly helical in methanol and 30% TFE, the question of whether the A-helix region spontaneously forms helix in aqueous solution remains open.

In contrast to the H-, G-, and GH-peptides, the other peptides derived from apomyoglobin exhibit little or no measurable tendency for spontaneous secondary structure formation in water solution. However, with the exception of part of the F-helix, all peptides corresponding to helical regions in the protein preferentially populate the α region of (ϕ,ψ) space, as indicated by the presence of $d_{NN}(i,i+1)$ NOEs in the NMR spectra. It is striking that the peptides with the strongest tendencies for spontaneous secondary structure formation in aqueous solution come from the GH region, consistent with a likely role in initiation of folding of the AGH molten globule intermediate.

Populations of helix in small concentrations of TFE

The addition of TFE causes an increase in the CD signal characteristic of helix in peptides that are capable of forming helix, generally those that contain a significant number of dihedral angles that populate the α region of (ϕ,ψ) space and/or can be described as "nascent helix" (Dyson et al., 1988). The percentage of helix can be roughly quantitated by assuming that a peptide with 0% helix has a Θ_{222} of 0, while 100% helix has Θ_{222} of $-40{,}000$ deg cm^2 dmol^{-1} (Chakrabartty et al., 1994). The results of these calculations are shown in Table 3 for all of the peptides corresponding to helical segments of apomyoglobin, in 0% and 30% TFE. All of the peptides except the H-helix peptide have negligible amounts of helix in the absence of TFE; the H-helix peptide has a significant proportion of helix in water solution, which is visible by NMR (Waltho et al., 1993). However, substantial populations of helix are induced

Table 3. *Helical content at 0% and 20% TFE*

Peptide fragment	% helix at 0% TFE[a]	% helix at 30% TFE[a]
B	10	53
CCD	5	55
D[a]	18	43
E	8	88
F[a]	0	18
G[b]	5	70
H[b]	23	68

[a]Obtained by dividing the molar ellipticity at 222 nm by −40,000, representing 100% helix.
[b]From Shin et al., 1993b; Waltho et al., 1993.

by quite small amounts of TFE in all but one of the peptides corresponding to helices in the native protein but not in the inter-helical peptide fragments (Fig. 1, Table 3). It is of great interest that the F-helix peptide shows no propensity to form helix in water solution and little, if any, propensity in TFE. We conclude that whereas local interactions are sufficient to stabilize the H-helix, long-range interactions that occur during later stages of apomyoglobin folding play an important role in stabilization of the other myoglobin helices.

Conclusion

Consistent with the equilibrium and kinetic results obtained for the folding of intact apomyoglobin, very little evidence for pre-formed secondary structure is observed in peptide fragments comprising secondary structure elements in the N-terminal segment (residues 4–99) of the folded protein. However, helical structure can readily be induced in a number of the peptides corresponding to helices in the folded protein by the addition of small quantities of helix-stabilizing solvents such as TFE, suggesting that these helices are stabilized through packing interactions within compact states formed late in the folding process. Although helical structure is observed in a peptide including the A-helix sequence in methanol solution, this evidence for structure must be regarded as inconclusive for the presence of pre-formed structure in the A helix sequence under folding conditions in aqueous solution. Even in helix-promoting solvents, the AB peptide is not uniformly helical. The helix is broken at a site close to the A-B helix junction in native myoglobin, suggesting that the bend between the A and B helices in the folded protein arises at least in part from intrinsic properties of the local amino acid sequence.

Materials and methods

Peptide synthesis and purification

Peptides were manually synthesized by standard solid phase Boc chemistry using previously described protocols (Reymond et al., 1992; Schnölzer et al., 1992). Peptides were blocked at the N-terminus with an acetyl group and at the C-terminus with a primary amide group. The natural sequence of sperm whale myoglobin is retained throughout, and the residues are numbered accordingly.

Crude peptides were purified by preparative reverse-phase HPLC in one step to yield pure compounds. A Waters PrepLC4000 System, with a Waters PrePak module 1000 holding a C_{18} Waters cartridge was used. The peptides were eluted using a linear gradient (1% B per minute) where A buffer was trifluoroacetic acid 0.1% and B buffer was 60% acetonitrile in 0.1% TFA. Purity was assessed by analytical HPLC using a Hitachi system and a Vydac C_{18} column. The identity of the peptides was confirmed by mass spectrometry and amino acid analysis.

NMR spectroscopy

^{1}H-NMR spectra were acquired at 278.8 K using a Bruker AMX300 spectrometer. Chemical shifts were referenced to an internal dioxane standard (3.75 ppm), and the probe temperature was calibrated using methanol (Van Geet, 1968). Except for the AB domain, all peptides were dissolved in 90% ^1H_2O/10% ^2H_2O. The pH was adjusted to 4–5 by addition of small amounts of NaOH or HCl. Concentrations ranged between 1 and 7.5 mM and were chosen so that peptides were monomeric in solution. Double quantum filtered scalar correlated spectroscopy (2QF COSY) experiments (Rance et al., 1983) were recorded with a t_{1max} of 100 ms and a t_{2max} of 400 ms, with 48 to 96 scans per t_1 increment. Total correlated spectroscopy (TOCSY) experiments (Braunschweiler & Ernst, 1983; Rance, 1987) were recorded with mixing time of 50 to 70 ms and 16 to 48 scans. For the B-helix, CCD-domain and E-helix peptides, two-dimensional nuclear Overhauser effect spectra (NOESY) (Jeener et al., 1979) were recorded with a mixing time of 250 ms. NOESY data were acquired with a Hahn-echo pulse to give a flat baseline (Rance & Byrd, 1983). For the other peptides, rotating-frame NOESY (ROESY) spectra were recorded with mixing time of 300–400 ms and 48 or 96 scans. Spectra were routinely acquired with 4096 complex points and spectral widths were 6125 Hz in both dimensions.

NMR spectra for the AB-domain peptide in deuterated methanol were recorded at 273 K on a Bruker AMX500. 2QF-COSY (4K complex points and 6250 Hz spectral width in both dimensions), double quantum (2Q) (4K complex points, 700 experiments, 6250 Hz spectral width in $\omega 2$ and 12500 Hz in $\omega 1$), TOCSY (8K complex points and 6250 Hz spectral width in both dimensions) and NOESY (8K complex points and 12500 Hz spectral width in both dimensions) spectra were acquired.

Data were processed on a CONVEX C2 computer and on Sun workstations using a modified version of FTNMR (Hare Research). Spectra were Fourier transformed in both dimensions using Lorentz-to-Gauss and/or phase-shifted sine bell window functions. The final 2D spectra were stored in 2K × 2K real matrices. Where necessary, baseline corrections were applied in ω_2. TOCSY and 2QF COSY spectra were used for side-chain assignment and ROESY or Hahn-echo NOESY were used for sequential assignment.

Amide proton temperature coefficients were measured from short mixing time TOCSY spectra recorded at various temperatures for all peptides except the AB domain.

CD spectroscopy

CD spectra were recorded at 5 °C on an AVIV 61DS spectropolarimeter calibrated with a standard solution of D-camphorsulfonic acid at 290.5 nm (Johnson, 1985). The wavelength was calibrated using the peaks of benzene vapor. Samples were prepared at a concentration of about 15 μM in a 10 mM phosphate buffer at pH 5, except for the A-helix and AB-domain peptides, which were dissolved in methanol. Peptide concentrations were determined by

quantitative amino acid analysis. A quartz cell with a 10 mm path length was used. The mean residue ellipticity $[\Theta]$, in deg cm^2 $dmol^{-1}$, was calculated from the relationship $[\Theta] = \Theta \times 1/(10lcN)$, where Θ is the observed ellipticity, l is the pathlength in cm, c is the molar concentration of the peptide, and N is the number of residues in the peptide. Data were obtained using a 1.5 nm bandwidth, a 0.5 nm step size, and a 4.0 s time constant and were smoothed with a third-order polynomial function with a window size typically between 5 and 20 points.

Acknowledgments

We thank Dr. Jian Yao and Dr. Shohei Koide for assistance with the NMR experiments. This work was supported by Grant GM38794 from the National Institutes of Health.

References

Baldwin RL. 1989. How does protein folding get started? *Trends Biochem Sci 14*:291–294.

Ballew RM, Sabelko J, Gruebele M. 1996. Direct observation of fast protein folding: The initial collapse of apomyoglobin. *Proc Natl Acad Sci USA 93*:5759–5764.

Bashford D, Chothia C, Lesk AM. 1987. Determinants of a protein fold. Unique features of the globin amino acid sequence. *J Mol Biol 196*:199–216.

Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson 53*:521–528.

Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. 1995. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins 21*:167–195.

Chakrabartty A, Kortemme T, Baldwin RL. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci 3*:843–852.

Creighton TE. 1988. Toward a better understanding of protein folding pathways. *Proc Natl Acad Sci USA 85*:5082–5086.

Cummings AL, Eyring EM. 1975. Helix-coil transition kinetics in aqueous poly(α,L-glutamic acid). *Biopolymers 14*:2107–2114.

Dyson HJ, Rance M, Houghten RA, Wright PE, Lerner RA. 1988. Folding of immunogenic peptide fragments of proteins in water solution. II The nascent helix. *J Mol Biol 201*:201–217.

Dyson HJ, Wright PE. 1991. Defining solution conformations of small linear peptides. *Annu Rev Biophys Biophys Chem 20*:519–538.

Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE. 1992a. Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding I. Myohemerythrin. *J Mol Biol 226*:795–817.

Dyson HJ, Sayre JR, Merutka G, Shin H-C, Lerner RA, Wright PE. 1992b. Folding of peptide fragments comprising the complete sequence of proteins: Models for the initiation of protein folding II Plastocyanin. *J Mol Biol 226*:819–835.

Dyson HJ, Wright PE. 1996. Insights into protein folding from NMR. *Annu Rev Phys Chem 47*:369–395.

Ebina S, Wüthrich K. 1984. Amide proton titration shifts in bull seminal inhibitor IIA by two-dimensional correlated ^{1}H nuclear magnetic resonance (COSY). Manifestation of conformational equilibria involving carboxylate groups. *J Mol Biol 179*:283–288.

Eliezer D, Wright P. 1996. Is apomyoglobin a molten globule? Structural characterization by NMR. *J Mol Biol 263*:531–538.

Epand RM, Scheraga HA. 1968. The influence of long-range interactions on the structure of myoglobin. *Biochemistry 7*:2864–2872.

Hermans J Jr, Puett D. 1971. Relative effects of primary and tertiary structure on helix formation in myoglobin and α-lactalbumin. *Biopolymers 10*:895–914.

Hirst JD, Brooks CL III. 1995. Molecular dynamics simulations of isolated helices of myoglobin. *Biochemistry 34*:7614–7621.

Hughson FM, Baldwin RL. 1989. Use of site-directed mutagenesis to destabilize native apomyoglobin relative to folding intermediates. *Biochemistry 28*:4415–4422.

Hughson FM, Wright PE, Baldwin RL. 1990. Structural characterization of a partly folded apomyoglobin intermediate. *Science 249*:1544–1548.

Jaenicke R. 1991. Protein folding: Local structures, domains, subunits, and assemblies. *Biochemistry 30*:3147–3161.

Jeener J, Meier BH, Bachmann P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys 71*:4546–4553.

Jennings PA, Wright PE. 1993. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science 262*:892–896.

Johnson WC. 1985. Circular dichroism and its empirical application to biopolymers. *Methods Biochem Anal 31*:61–163.

Kim PS, Baldwin RL. 1990. Intermediates in the folding reactions of small proteins. *Annu Rev Biochem 59*:631–660.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins 6*:87–103.

Lumry R, Legare R, Miller WG. 1964. The dynamics of the helix-coil transition in the poly-α,L-glutamic acid. *Biopolymers 2*:489–500.

Matthews CR. 1993. Pathways of protein folding. *Annu Rev Biochem 62*:653–683.

Merutka G, Morikis D, Brüschweiler R, Wright PE. 1993. NMR evidence for multiple conformations in a highly helical model peptide. *Biochemistry 32*:13089–13097.

Merutka G, Dyson HJ, Wright PE. 1995. "Random coil" ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J Biomol NMR 5*:14–24.

Montelione GT, Scheraga HA. 1989. Formation of local structures in protein folding. *Acc Chem Res 22*:70–76.

Ptitsyn OB. 1991. How does protein synthesis give rise to the 3-D structure? *FEBS Lett 285*:176–181.

Rance M, Byrd RA. 1983. Obtaining high-fidelity spin-1/2 powder spectra in anisotropic media: Phase-cycled Hahn echo spectroscopy. *J Magn Reson 52*:221–240.

Rance M, Sorensen OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K. 1983. Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering. *Biochem Biophys Res Commun 117*:479–485.

Rance M. 1987. Improved techniques for homonuclear rotating-frame and isotropic mixing experiments. *J Magn Reson 74*:557–564.

Reymond MT, Delmas L, Koerber SC, Brown MR, Rivier JE. 1992. Truncated, branched and/or cyclic analogues of neuropeptide Y: Importance of the pancreatic peptide fold in the design of specific Y_2 receptor ligands. *J Med Chem 35*:3653–3659.

Rizo J, Blanco FJ, Kobe B, Bruch MD, Gierasch LM. 1993. Conformational behavior of *Escherichia coli* OmpA signal peptides in membrane mimetic environments. *Biochemistry 32*:4881–4894.

Schnölzer M, Alewood P, Jones A, Alewood D, Kent SBH. 1992. In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *Int J Pept Protein Res 40*:180–193.

Shin H-C, Merutka G, Waltho JP, Tennant LL, Dyson HJ, Wright PE. 1993a. Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin. *Biochemistry 32*:6356–6364.

Shin H-C, Merutka G, Waltho JP, Wright PE, Dyson HJ. 1993b. Peptide models of protein folding initiation sites. 2. The G-H turn region of myoglobin acts as a helix stop signal. *Biochemistry 32*:6348–6355.

Soman KV, Karimi A, Case DA. 1993. Molecular dynamics analysis of a ribonuclease C-peptide analogue. *Biopolymers 33*:1567–1580.

Van Geet AL. 1968. Calibration of the methanol and glycol nuclear magnetic resonance thermometers with a static thermistor probe. *Anal Chem 40*:2227–2229.

Waltho JP, Feher VA, Lerner RA, Wright PE. 1989. Conformation of a T cell stimulating peptide in aqueous solution. *FEBS Lett 250*:400–404.

Waltho JP, Feher VA, Merutka G, Dyson HJ, Wright PE. 1993. Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin. *Biochemistry 32*:6337–6347.

Wright PE, Dyson HJ, Lerner RA. 1988. Conformation of peptide fragments of proteins in aqueous solution: Implications for initiation of protein folding. *Biochemistry 27*:7167–7175.