Contribution of Increased Length and Intact Capping Sequences to the Conformational Preference for Helix in a 31-Residue Peptide from the C Terminus of Myohemerythrin †

ARTICLE *in* BIOCHEMISTRY · MAY 1997

Impact Factor: 3.02 · DOI: 10.1021/bi970038x · Source: PubMed

Contribution of Increased Length and Intact Capping Sequences to the Conformational Preference for Helix in a 31-Residue Peptide from the C Terminus of Myohemerythrin†

Martine T. Reymond, Shouqin Huo, Brendan Duggan, Peter E. Wright,* and H. Jane Dyson*

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

Received January 7, 1997; Revised Manuscript Received March 6, 1997⊗

ABSTRACT: In order to examine the effects of chain length on the propensity of short peptides to form helix-like structures in aqueous solution, we have studied a peptide of 31 residues consisting of the C-terminal sequence (residues 88−118) of the four-helix bundle protein myohemerythrin from *Themiste zostericola*. This peptide, termed MDC, represents the final two elements of secondary structure in the protein, the D-helix and the C-terminal loop sequence, together with a five-residue sequence at the N terminus corresponding to the linker between the C- and D-helices. An N-capping sequence, VDAKNV, immediately precedes the D-helix sequence, and a C-capping sequence, VNHIKGT, corresponding to the α_L termination motif, occurs at the C-terminal end. The effect of replacement of a cysteine residue in the middle of the sequence with an alanine was explored by the comparison of the MDC peptide and a 16-residue peptide representing the sequence of the D-helix alone, both containing the change Cys99Ala. Significant changes in the NMR and CD spectra were seen for both peptides compared to the wild-type sequence. A comparison of the fluorescence spectra of the wild-type and Cys99Ala peptides indicated that a specific interaction between the side chains of Cys 99 and Trp 102 acts to quench the fluorescence of the tryptophan ring and probably contributes a component that distorts the CD spectrum of the wild-type peptide at ∼220−235 nm. The effect of an increase in the length of the peptide, with the incorporation of capping sequences derived from the native sequence, was explored by NMR and CD spectroscopy of the 31-residue and 16-residue peptides in aqueous solution and in TFE/water mixtures. Evidence for the formation of a significant population of helical conformers in the region of the MDC peptide corresponding to the D-helix was observed in aqueous solution using CD and NMR spectroscopy. The C-terminal 10 residues of the MDC peptide behave in solution in a manner identical to that of a 10-residue peptide with the same sequence; a highly specific local interaction between an aromatic ring and a glycine amide proton appears to be retained in the longer peptide. Upon addition of trifluoroethanol (TFE), significant shifts are observed in a number of resonances in the NMR spectrum, and both chemical shifts and NOEs provide evidence for a higher population of helix in the D-helix region of the peptide in TFE. However, TFE is unable to promote the propagation of helix beyond the N-cap or α_L termination motifs, and the specific local interaction observed in the C-terminal sequence is retained in TFE. The CD spectrum in TFE shows an increase in the proportion of helix, to an overall maximum of approximately 55% helix at 50% v/v TFE, corresponding to approximately 100% helix in the D-helix sequence of the peptide, since the N and C termini of the MDC peptide are not helical according to the NMR spectra. The high proportion of helix observed in the D-helix sequence of the longer MDC peptide demonstrates that the presence of intact capping sequences can constrain the peptide conformational ensemble to resemble that seen in the native protein. A compendium of results from this and previous peptide studies has also led to a novel observation, the existence of a correlation between the amide proton chemical shift and temperature coefficient.

The facile folding of many proteins *in vitro* is an indication that both the process of folding and the nature of the final folded structure are determined solely by the amino acid sequence. It follows that, during the earliest stages of the folding of the protein, local structures are formed that depend only on the local amino acid sequence. Since the process of folding, particularly the early steps, is so rapid in many globular proteins, we have undertaken the task of dissecting these stages by examining peptide fragments of proteins. In these systems, the driving forces that initiate protein folding would be operative, but no long-range tertiary interactions, which would lead to the cooperative folding of the protein, are possible. Therefore, under equilibrium conditions, we can determine the nature of likely protein folding initiation structures and the sources of their stabilization. These ideas have been recently reviewed (Dyson & Wright, 1991, 1993).

The conformational preferences of peptide fragments of proteins reflect at least to some extent the structures of the cognate sequences in the folded protein. This has been

† This work was supported by Grant GM38794 from the National Institutes of Health.

* Address correspondence to Peter E. Wright or H. Jane Dyson, The Scripps Research Institute, Department of Molecular Biology MB2, 10550 N. Torrey Pines Rd., La Jolla, CA 92037.

⊗ Abstract published in *Advance ACS Abstracts,* April 15, 1997.

S0006-2960(97)00038-X CCC: $14.00

Table 1: Amino Acid Sequence of the Synthesized Peptide

	86 90 95 100 105 110 118
Myohemerythrin[a]	G L S A P V D A K N V D Y C K E W L V N H I K G T D F K Y K G K L
CD Loop[b]	G L S A P V D
D-Helix[b]	A K N V D Y C K E W L V N H I K
C-Terminal Loop[b]	G T D F K Y K G K L
D(C99A)-Helix[c]	A K N V D Y A K E W L V N H I K
MDC Peptide[c]	S A P V D A K N V D Y A K E W L V N H I K G T D F K Y K G K L

[a] Amino acid sequence from Klippenstein et al. (1976). [b] From Dyson et al. (1992a). [c] This work; the underlined Ala residue is the only change in the amino acid sequence. Brackets denote the N- and C-capping sequences.

shown by a number of studies of peptides comprising the complete sequences of proteins (Dyson et al., 1992a,b; Waltho et al., 1993; Kemmink & Creighton, 1993; Kippen & Fersht, 1995). For example, peptides derived from the sequence of *Themiste zostericola* myohemerythrin, a four-helix bundle protein, consistently show evidence for high populations of helix and turn conformations (Dyson et al., 1992a), while those derived from the sequence of French bean plastocyanin, a β-sandwich protein, show no such preference but appear to prefer extended conformations (Dyson et al., 1992b). The peptides in these two studies were in general designed to encompass a single element of secondary structure, helix, β-strand, β-hairpin, or the intervening turn or loop. It has however become apparent in subsequent studies (Waltho et al., 1993; Shin et al., 1993a,b) that the length of the peptide fragment may be critical to the population of the structured conformer observed, particularly in the case of helices, where end effects, especially where the termini are unblocked, may act to destabilize the helical conformations by interacting unfavorably with the helix dipole. In other cases, populations of folded forms may be increased by the influence of the terminal charges, for example the type VI turn conformation found in five- and six-residue peptides (Yao et al., 1994a,b). We therefore decided to revisit the myohemerythrin sequence by studying a peptide, termed MDC, that comprises two of the secondary structure elements, the D-helix and the C-terminal loop, together with a portion of the loop sequence between the C- and D-helices. The same sequence was studied previously (Dyson et al., 1992a) as three peptides, termed the CD-loop, D-helix, and C-terminal peptides. The sequences of all of these peptides are shown in Table 1.

We were interested in examining the effects on the population of helix and of other structural elements found in the three component peptides (Dyson et al., 1992a) as a function of chain length and the presence and absence of specific capping sequences. Several factors must be taken into account. A longer peptide, because of its greater rotational correlation time, might be expected to give rise to NOEs[1] of greater intensity than a shorter peptide, even when the populations of structured conformers giving rise to the NOEs are similar in the two cases (Ernst et al., 1987). The propensity for helix could be increased in longer peptides due to the Zimm−Bragg propagation effect (Zimm & Bragg, 1959), as long as the helix itself can propagate. The absence of end effects and of the N- and C-terminal charges present in unblocked peptides could serve to increase the length of the helix observed, and in particular, the presence of intact capping sequences in the longer peptide that are absent in the shorter D-helix peptide studied previously could give rise to a measurable increase in local helical propensity in this region. Capping sequences have been shown in a number of cases to increase the observed helicity of model peptides (Zhou et al., 1994a,b; Lyu et al., 1992, 1993) and fragments of proteins (Odaert et al., 1995). Capping was shown to increase the stability of proteins (Serrano & Fersht, 1989), although it appears that capping sequences do not necessarily function as stop signals for the helix (Jiménez et al., 1994). The MDC peptide studied in the present work represents an example of a sequence from a native protein where data are available not only on the structure of the intact protein (Sheriff et al., 1987) but also on the conformational preferences of short peptides lacking both the N- and C-capping sequences that are present in the native protein (Dyson et al., 1992a). The effect of the addition of the native sequence containing these capping residues is evaluated by examination of the conformational preferences of the MDC peptide in aqueous solution and in 30% TFE and comparison with similar results for shorter peptide components of the C-terminal sequence of myohemerythrin.

MATERIALS AND METHODS

Peptide Synthesis and Purification. The MDC peptide (residues 88−118 of *T. zostericola* myohemerythrin), and the D(C99A)-helix peptide (residues 94−108), were manually synthesized by standard solid-phase Boc chemistry and purified using previously described protocols (Reymond et al., 1992; Schnölzer et al., 1992). The N and C termini of MDC were unblocked to provide maximum solubility, while the termini of the D(C99A)-helix peptide were blocked (N-terminal acetyl, C-terminal amide) to give greater similarity to the longer peptide. The cysteine residue at position 99 was replaced in each peptide by alanine to avoid the possibility of forming a disulfide-bridged dimer by air oxidation. The purity was assessed by analytical HPLC, and the identity of the peptides was confirmed by mass spectrometry and by amino acid analysis.

CD Spectroscopy. CD spectra were recorded at 5 °C on an AVIV 61DS spectropolarimeter. Samples were prepared at a concentration of 17−25 μM in 10 mM potassium phosphate at pH 4.5. Peptide concentrations were determined by quantitative amino acid analyses. Data were acquired using a 10 mm path length quartz cell with various concentrations (% v/v) of 2,2,2-trifluoroethanol (TFE).

NMR Measurements. The peptides were dissolved in 90% ^1H_2O/10% ^2H_2O or in 30% (v/v) trifluoroethanol-d_4/70% ^1H_2O at a concentration of 3−5 mM, and the pH was adjusted to 4.5 by addition of small amounts of NaOH or HCl. ^{1}H NMR spectra were acquired at 278 K using a Bruker AMX500 spectrometer. The NMR spectrum of the MDC peptide is well-resolved, with relatively narrow line widths characteristic of a monomeric peptide of the expected molecular weight; one-dimensional spectra acquired at a

[1] Abbreviations: NMR, nuclear magnetic resonance; CD, circular dichroism; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; 2QF COSY, two-dimensional double-quantum filtered correlation spectroscopy; TOCSY, two-dimensional total correlation spectroscopy; 2Q, two-dimensional double-quantum spectroscopy; TFE, 2,2,2-trifluoroethanol; HPLC, high-performance liquid chromatography.

concentration of 0.3 mM showed no difference in the line widths or chemical shifts of the resonances, indicating that the peptide is monomeric under the conditions studied. Data were processed on a Convex C240 computer using a modified version of FTNMR (Hare Research). TOCSY (Braunschweiler & Ernst, 1983; Rance, 1987) and 2QF COSY (Rance et al., 1983) spectra were used for side chain assignment, and a NOESY spectrum was used for sequential assignment. Chemical shifts were referenced to an internal standard of dioxane at 3.75 ppm. The TOCSY experiment was recorded with a mixing time of 64 ms. The NOESY spectrum (Bodenhausen et al., 1984), acquired with a Hahn echo for optimal baseline shape (Rance & Byrd, 1983), was recorded with a mixing time of 300 ms; solvent saturation was applied during the mixing time. Amide proton temperature coefficients were determined at four different temperatures, using double-quantum (2Q) (Braunschweiler et al., 1983; Rance & Wright, 1986) or 2QF COSY spectra.

Fluorescence Measurements. Fluorescence measurements of the two shorter (16-residue) peptides were made on a SLM 8100 spectrofluorimeter at 25 °C. Excitation spectra were recorded by monitoring emission at 336 nm, and emission spectra were recorded with excitation at 282 nm. Peptides were dissolved in water at pH 4.5 to give concentrations of 12.7 μM (wild-type D-helix peptide) and 9.5 μM [D(C99A)-helix peptide]. The concentration of each peptide was determined by amino acid analysis. Spectra were corrected for concentration, water fluorescence, and instrument sensitivity.

RESULTS AND DISCUSSION

Design of Peptides

The sequence of the MDC peptide (Myohemerythrin D-helix and C terminus) corresponds exactly to that of the C-terminal 31 residues of the four-helix bundle protein myohemerythrin from *T. zostericola* (Klippenstein et al., 1976), except that the cysteine residue at position 99 is replaced by alanine (Table 1). This substitution was made to remove the possibility of intermolecular disulfide bond formation. Since previous work on shorter peptides from myohemerythrin was carried out with the native cysteine-containing sequence (Dyson et al., 1992a), a short peptide of 16 residues was synthesized corresponding to the sequence of the D-helix with the Cys-to-Ala substitution at position 99, and its conformational preferences were examined as a control. This peptide is termed the D(C99A)-helix peptide.

CD Spectra

The CD spectrum of the MDC peptide does not show the well-marked double minimum at 208 and 222 nm characteristic of an ordered helix in water solution, but the shape of the curve does suggest a measurable population of helix. The spectrum resembles that of the D-helix peptide (Dyson et al., 1992a), but with greater negative ellipticity in the region between 205 and 230 nm (Figure 1A). The Θ_{222} of -5000 deg cm^2 dmol^{-1} corresponds to about 10−15% helix on average throughout the molecule. This in itself is an indication that parts of the longer peptide may be more helical than the corresponding regions in the shorter peptides. If only the D-helix region is helical, as indicated by the NMR data (see below), the observed Θ_{222} corresponds to about 30% helix in this part of the sequence.

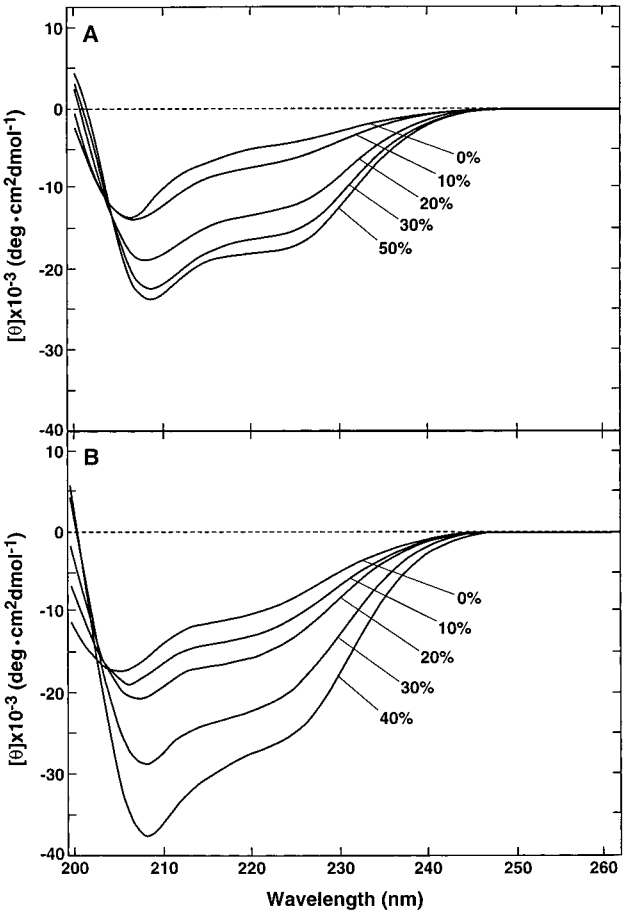

FIGURE 1: Ultraviolet circular dichroism spectra of (A) the MDC peptide (17 μM) and (B) the D(C99A)-helix peptide in various concentrations (% v/v) of trifluoroethanol in 10 mM potassium phosphate at pH 4.5.

In order to determine whether the helix could be made to propagate beyond the D-helix region, the helix-promoting solvent TFE was added. Addition of TFE to the solution results in a greater negative ellipticity at 208 and 222 nm, to a maximum at 50% TFE that represents approximately 50% helix overall in the peptide, using $\Theta_{222} = -40000$ deg cm^2 dmol^{-1} to represent 100% helix (Chen et al., 1974; Chakrabartty et al., 1991). The NMR data unequivocally indicate that the C-terminal sequence is not helical in aqueous solution (see below), and addition of TFE to the shorter peptide representing the C-terminal sequence gives no detectable difference in the CD spectrum (Dyson et al., 1992a). If it can be assumed that the majority of the increased negative ellipticity in 50% TFE arises from the D-helix part of the MDC peptide, then the CD spectrum indicates nearly complete helix formation in this sequence. On the other hand, it appears that TFE can induce helix formation that continues through capping sequences present in the center of helices (Jiménez et al., 1994). In order to determine whether TFE can induce propagation of helix through capping sequences located at the ends of a natural sequence, we have analyzed the MDC peptide by NMR in 30% TFE solution (see a later section).

The CD spectrum of the D(C99A)-helix peptide is shown in Figure 1B as a function of TFE concentration from 0 to 50% v/v. The shapes of the curves are very similar to those of the MDC peptide. The Θ_{222} for the sample in H_2O is about twice that of MDC, giving a population of helix of about 25%, which is very similar to that observed for the

same residues in the longer MDC peptide. The population of helix in the D(C99A)-helix peptide in 40% TFE is approximately 65%, less than that deduced under the same conditions for the D-helix portion of the MDC peptide by assuming that the terminal sequences are not helical. The published CD spectrum of the D-helix peptide in 70% TFE (Dyson et al., 1992a) indicates that the wild-type sequence is about 40% helical under these conditions. These results suggest a significant increase in the population of ordered helical conformations in the D-helix part of the longer peptide in 40−50% TFE compared to the two shorter peptides in TFE solution.

There are two possible explanations for the lower helicity observed for the wild-type D-helix peptide compared to that for the D(C99A)-helix peptide: (i) the substitution of Ala for Cys and (ii) the wild-type peptide has charged, unblocked termini (Dyson et al., 1992a), which can influence the population of helix in peptides (Fairman et al., 1989), while the termini of the D(C99A)-helix peptide are blocked. An examination of the CD spectrum of the wild-type D-helix peptide in water and TFE [Figure 5 in Dyson et al. (1992a)] reveals a distortion in the spectrum in the region between 215 and 235 nm which appears to be due to the superposition of components with positive ellipticity onto a generally helical CD spectrum. This distortion is absent in the spectrum of the D(C99A)-helix peptide (Figure 1B), which is more typical of a peptide populating both helical and random coil states. The distortion in the CD spectrum of the wild-type peptide may arise from an interaction between the side chains of Cys and Trp (see a later section).

NMR Spectra of the MDC and D(C99A)-Helix Peptides

The ^{1}H NMR spectrum of MDC is relatively well-dispersed and has been completely assigned from 2QF COSY, TOCSY, 2Q, and NOESY spectra in 90% ^1H_2O/10% ^2H_2O. The resonance assignments are listed in Table 2 of the Supporting Information. The addition of TFE to the solution causes a significant change in the NMR spectrum of the peptide. This is illustrated in Figure 2, which shows a comparison of the fingerprint regions of the 2QF COSY spectra of the MDC peptide in aqueous solution and in 30% TFE. The NMR spectrum of the peptide in 30% TFE was assigned from 2QF COSY, TOCSY, and NOESY spectra, and the assignments are listed in Table 3 of the Supporting Information.

Three regions of the NOESY spectrum of MDC in aqueous solution containing sequential and medium-range NOE connectivities are shown in Figure 3, and the corresponding regions are shown in Figure 4 for the peptide in 30% TFE. The NOEs observed for the MDC peptide in aqueous solution and in 30% TFE are summarized in panels A and B of Figure 5, respectively. The NMR spectrum of the shorter D(C99A)-helix peptide has also been studied by NMR; the chemical shifts and NOE patterns are very similar to those of the original D-helix peptide (Dyson et al., 1992a) and are listed in Table 4 of the Supporting Information. Figure 5C shows a summary of the NOEs observed for the D(C99A)-helix peptide.

Propagation of Helix in Long and Short Peptides: Evidence from NOEs and Chemical Shifts

The pattern of NOE connectivities suggests that well-formed helix is present in water solution in the central portion

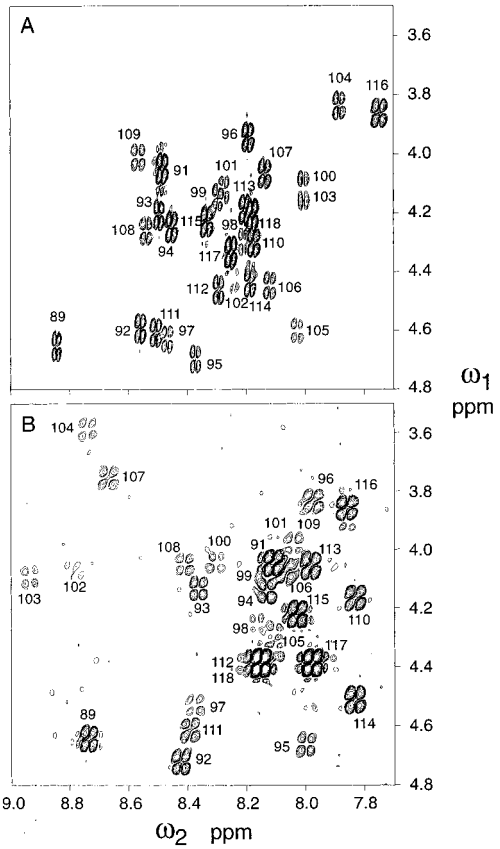

FIGURE 2: Fingerprint regions of the NMR spectra of the MDC peptide in (A) 90% ^1H_2O/10% ^2H_2O at pH 4.5 and 278 K and (B) 30% v/v trifluoroethanol-d_3 at pH 4.5 and 278 K.

of the MDC peptide, corresponding to the D-helix sequence of the protein. This is indicated by the density of both $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs throughout this part of the sequence. Although overlap is a problem, as expected for a peptide of this length without a unique conformation, the population of helical conformations appears to be quite high from the NOE patterns. A comparison of panels A and C of Figure 5 shows considerably greater formation of helix in the MDC peptide compared with that in the D(C99A)-helix peptide, with a greater density of $(i,i+3)$ connectivities. The helix is apparently better formed at the N-terminal end of the D-helix sequence in the longer peptide, and a number of the medium-range NOEs observed in both peptides toward the C terminus of this sequence are of greater intensity relative to the corresponding inter-residue $d_{\alpha N}(i,i+1)$ NOE in the MDC peptide than in the D(C99A)-helix peptide, suggesting greater helix formation in the longer peptide.

Other NMR-derived quantities can be used to estimate the population of helix in peptides in a site-specific manner (Dyson & Wright, 1991, 1993). Among the most useful quantities are the chemical shift and the temperature coefficient. The $C^\alpha H$ chemical shift can provide a measure of the secondary structure propensity at a given site in the peptide sequence (Wishart et al., 1991; Rizo et al., 1993). Figure 6A shows the difference in the $C^\alpha H$ chemical shifts from "random" values (Merutka et al., 1995) for the MDC and D(C99A)-helix peptides plotted as a function of residue number, together with those from the previous study of shorter peptide fragments (Dyson et al., 1992a). Several observations can be made. First, the $C^\alpha H$ resonances at the unblocked N and C termini of all of the peptides are shifted relative to those of the same residues at nonterminal

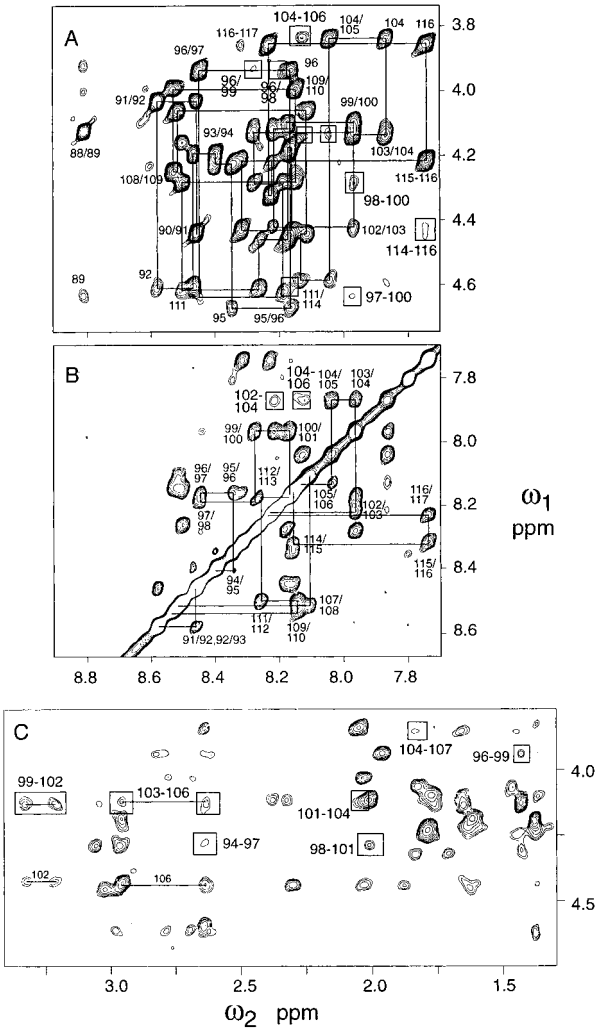

FIGURE 3: Portions of a 500 MHz NOESY spectrum of the MDC peptide in 90% ^1H_2O/10% ^2H_2O at 278 K, showing the NOE connectivities between (A) NH and C$^\alpha$H (upper panel), (B) NH and NH (middle panel), and (C) C$^\alpha$H and C$^\beta$H (lower panel). Selected sequential and medium-range NOEs are labeled; medium-range NOEs are indicated with boxes.

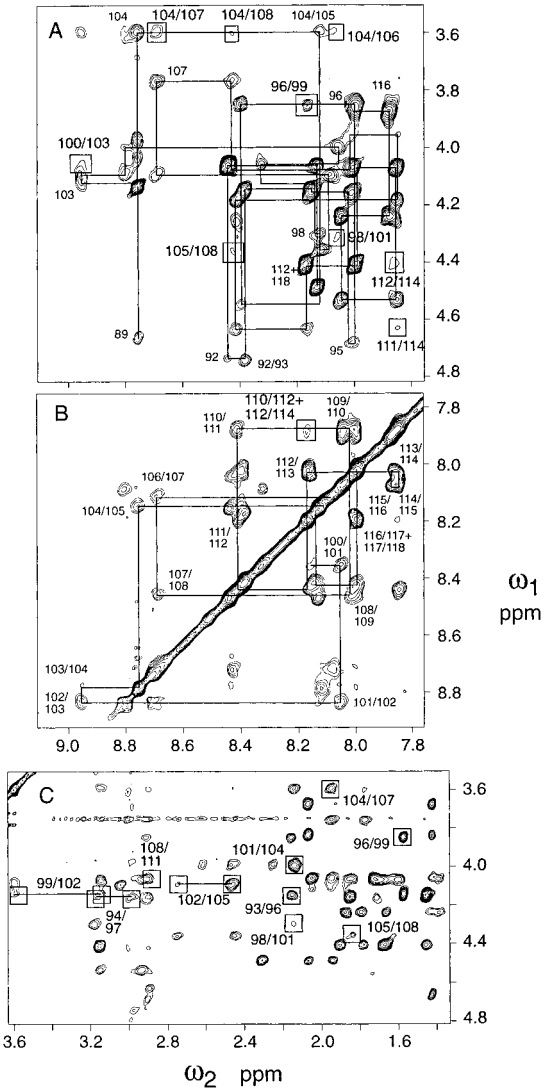

FIGURE 4: Portions of a 500 MHz NOESY spectrum of the MDC peptide in 30% v/v TFE-d_3 at 278 K, showing the NOE connectivities between (A) NH and C$^\alpha$H (upper panel), (B) NH and NH (middle panel), and (C) C$^\alpha$H and C$^\beta$H (lower panel). Selected sequential and medium-range NOEs are labeled; medium-range NOEs are indicated with boxes.

positions. In general, it appears that the C$^\alpha$H resonance moves upfield (decreased chemical shift, i.e. a more negative value in the figure) for both N- and C-terminal residues. For the blocked D(C99A)-helix peptide, the C$^\alpha$H chemical shift is identical at the C terminus but is significantly different at the N terminus from that of the unblocked D-helix peptide (Dyson et al., 1992a), indicating that much of the chemical shift difference at this position is due to end effects, most probably of the charged N-terminal amino group. Second, the portion of the MDC peptide corresponding to the C-terminal loop of the protein shows C$^\alpha$H chemical shifts that are almost identical with those of the shorter peptide corresponding to this sequence. This observation is discussed more fully in a later section. Finally, the C$^\alpha$H resonances of residues 93−107 in the central portion of the MDC peptide, which corresponds to the mutated D(C99A)-helix sequence and for which there is strong evidence for helix formation from NOEs, show a small but consistent upfield shift relative to those in the D(C99A)-helix peptide, particularly in the 98−107 region. A calculation can be made of the percentage of helix according to the method of Rizo et al. (1993). By this calculation, the comparable helical portion (residues 95−107) of the MDC peptide contains a slightly greater proportion of helix than the D(C99A)-helix peptide (48 compared to 36%), while if only the region between residues 98 and 107 is considered, the MDC peptide contains significantly more helix (62 compared to 46%). These numbers can be compared with those estimated from the CD spectra: 30% for MDC (D-helix sequence only) and 25% for D(C99A). The pattern of NOEs and the smaller $\Delta\delta(H^\alpha)$ at the N terminus indicate N-terminal fraying in the shorter peptide.

It is apparent from the NMR results that the helix cannot propagate in the manner described by Zimm and Bragg (1959) beyond the N-terminal and C-terminal capping sequences. These therefore act as effective termination signals for the helix.

Hydrogen Bonding Effects

The NH chemical shift can be used to deduce the extent of intramolecular hydrogen bonding (Wishart et al., 1991). The differences in NH chemical shifts from random values at the same temperature (Merutka et al., 1995) are shown in Figure 6B. These too show a general upfield shift in the

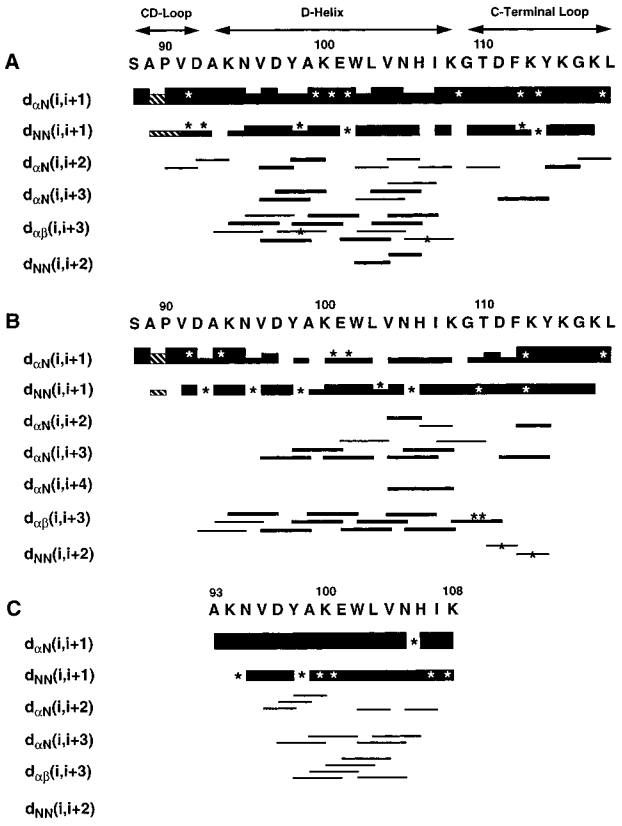

FIGURE 5: Summary of NOE connectivities observed for (A) the MDC peptide in 90% ^1H_2O/10% ^2H_2O, derived from NOESY data under a number of conditions as well as from the NOESY spectrum shown in Figure 3; (B) the MDC peptide in 30% v/v TFE-d_3, derived from the NOESY spectrum of Figure 4; and (C) the D(C99A)-helix peptide (N-terminal acetyl, C-terminal amide). The height of the box indicates the relative intensity of the NOE. Hatched boxes indicate connectivities to the $C^{\delta}H$ resonances of proline in place of the NH. Asterisks indicate connectivities that are ambiguous due to resonance overlap (for sequential NOEs). For medium-range NOEs, asterisks indicate that a resolved cross-peak is observed, but could belong to more than one medium-range NOE; where such NOEs would be overlapped with strong sequential or intraresidue NOEs, they are omitted from the figure. The double asterisk in part B indicates a connectivity that may contain contributions from an NOE between Lys 108 $C^{\alpha}H$ and $C^{\epsilon}H$. In parts A and B, thicker lines for medium-range NOEs indicate those that are clearly visible in Figures 3 and 4, respectively.

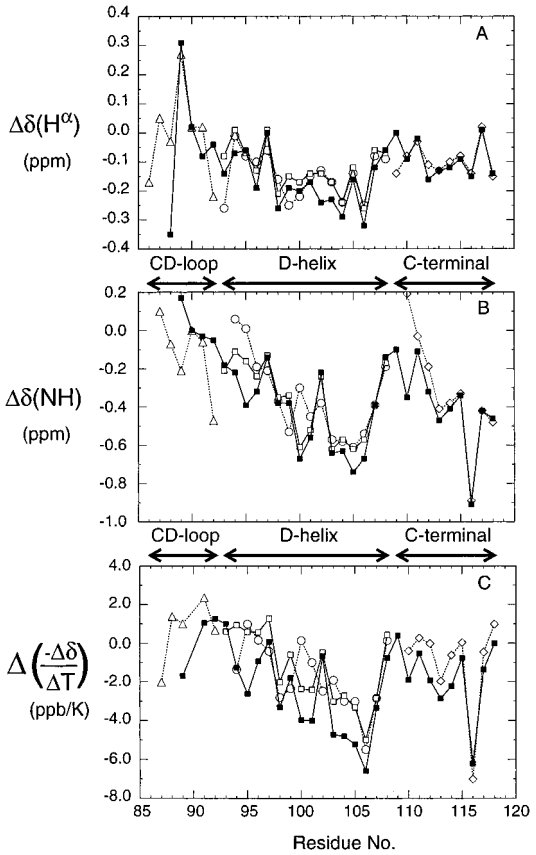

FIGURE 6: Plots as a function of the residue number of (A) the difference between the chemical shift of the $C^{\alpha}H$ resonance and the corresponding "random coil" value (Merutka et al., 1995), (B) the difference between the chemical shift of the NH resonance and the corresponding random coil value (Merutka et al., 1995), and (C) the difference between the temperature coefficient ($-\Delta\delta/\Delta T$ in units of parts per billion per kelvin) and the corresponding random value (Merutka et al., 1995). Data are shown for the MDC peptide (■), the D(C99A)-helix peptide (□), and the three unblocked component peptides (Dyson et al., 1992a): the CD-loop (△), the D-helix (○), and the C-terminal loop (◇). Small corrections have been made to the previously published values for the D-helix peptide, which were slightly in error in the region surrounding Trp 102 due to contamination with a species containing a formyltryptophan side chain. Some of the temperature coefficients are larger than those for random coil peptides, possibly reflecting end effects in the short GGXGG peptides used to obtain the random coil data (Merutka et al., 1995).

middle of the MDC peptide, in the region corresponding to the D-helix. Analogous to the behavior of the $C^{\alpha}H$ chemical shifts, the NH resonances in the unblocked D-helix peptide (Dyson et al., 1992a) are shifted relative to those in the blocked D(C99A)-helix peptide or in the longer MDC peptide. Again, end effects on the chemical shift appear to be greater at the N terminus than at the C terminus.

The temperature coefficients observed for the MDC and D(C99A)-helix peptides (Tables 2 and 4, respectively, of the Supporting Information) are plotted in Figure 6C as a function of residue number, together with the published values for the shorter peptides (Dyson et al., 1992a) and for the D(C99A)-helix peptide. Temperature coefficients can give a measure of the protection of the amide protons from solvent. Lowered temperature coefficients are usually taken to indicate some degree of intramolecular hydrogen bonding, generally as part of an element of secondary structure such as a turn, helix, or β-hairpin (Rose et al., 1985). A number of the temperature coefficients observed for the MDC peptide are significantly lower than those observed for the same residues in the shorter peptides. There are two reasons for this. One is that end effects are more significant in the shorter peptides; for the three short peptides comprising the same sequence as the MDC peptide, there are six termini, compared with two for the longer peptide. The second reason is that there is a greater population of helical conformations in the center of the MDC peptide (residues 98−107) than in the corresponding short D-helix and D(C99A)-helix peptides according to both NOE and chemical shift data, which should cause a general decrease in the temperature coefficient due to hydrogen bonding within the helix and consequent protection from solvent. One of the major differences between the MDC peptide and the shorter peptides is in the chemical shift and temperature coefficient of the amide proton of Asn 95, which indicate that this proton is hydrogen bonded to a significantly greater extent in the longer peptide. This is consistent with the presence of a capping interaction at the N terminus (see a later section).

Evidence for the Interaction of Trp and Cys in the Peptide with Native Sequence

Interestingly, one of the amide protons, Trp 102, appears from Figure 6 to have become more exposed to solvent in the MDC and D(C99A)-helix peptides than in the wild-type D-helix peptide (Dyson et al., 1992a); Figure 6 shows that both the NH chemical shift and the temperature coefficient are higher. This is most probably due to the replacement of Cys 99 by the less bulky Ala 99; since, in a helical structure, the side chain of residue 99 will be packed against Trp 102, the smaller alanine side chain might be expected to allow greater exposure of the Trp 102 amide proton to solvent. Further, a comparison of the CD data for the MDC and D(C99A)-helix peptides shows behavior significantly different from the published spectrum of the D-helix peptide (Dyson et al., 1992a), which shows a distorted curve with some negative ellipticity at 222 nm. The distortion appears to be due to a contribution from a component with positive ellipticity in this region, which was attributed to aromatic residues (Dyson et al., 1992a). The same complement of aromatic residues is also present in the MDC and D(C99A)-helix peptides; the only difference between the D-helix peptide (Dyson et al., 1992a) and the D(C99A)-helix peptide is the substitution of Cys for Ala at position 99. These comparisons indicate that the positive ellipticity at ~230 nm in the D-helix peptide (Dyson et al., 1992a) may well arise from a direct interaction of the tryptophan at position 102 with the cysteine at position 99. Such an interaction is not seen in the crystal structure of the intact protein (Sheriff et al., 1987), although it could easily be accommodated in an isolated helix by a rotation about the $\beta-\gamma$ bond of the tryptophan, an example of the conformational heterogeneity that is quite likely to occur in peptides. Aromatic residues themselves can influence the far-UV CD spectrum (Woody, 1978; Chakrabartty et al., 1993b), and positive ellipticity contributions at 225−230 nm have been associated with disulfide bonds in folded protein domains (Hider et al., 1988). Interactions have also been noted between Cys and Phe side chains in helices (Viguera & Serrano, 1995). The monomeric state of the D-helix peptide has been confirmed by ion spray mass spectrometry; the unusual CD spectrum of the D-helix peptide is not due to the formation of a disulfide bond. We therefore hypothesize that the unusual CD spectrum reflects the presence of an interaction between the side chains of cysteine and tryptophan in the D-helix peptide.

Interaction between Tryptophan and Cysteine Shown by Fluorescence

If there is an interaction between the side chains of Trp 102 and Cys 99 in the wild-type D-helix peptide sufficient to cause changes in the CD spectrum, then the proximity of the cysteine thiol may be sufficient to cause some quenching of the fluorescence of the tryptophan residue. In order to test this hypothesis, a comparison was made of the fluorescence spectra of the D-helix peptide (with blocked termini) and of the D(C99A)-helix peptide. Figure 7 shows the fluorescence emission spectra of the two peptides. The fluorescence of the wild-type peptide is clearly significantly diminished from that of the D(C99A)-helix peptide, an indication that the presence of the thiol affects the local electronic environment of the tryptophan ring. In order to test whether the interaction was due to the formation of structure in the peptide, a comparison was made of the

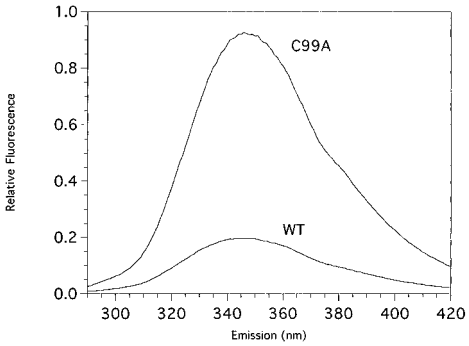

FIGURE 7: Fluorescence emission spectra of the wild-type D-helix peptide (blocked termini) (12.7 μM) and the D(C99A)-helix peptide (9.5 μM) at 25°C. Both peptides were in water solution at pH 4.5. The excitation wavelength was 282 nm. The intensities of the spectra have been corrected to account for concentration, water fluorescence, and instrument sensitivity.

D-helix peptide in the presence and absence of 2 M urea (data not shown). The fluorescence emission increased approximately 20% in urea, confirming that the tryptophan fluorescence quenching is due at least in part to its interactions with the cysteine side chain in a structured conformation.

Effect of TFE on the Helix Population in the MDC Peptide

NOE Information. A comparison of panels A and B of Figure 5 gives a good indication of the local effects of TFE on the conformational preferences of the MDC peptide. It is clear that the N- and C-terminal regions of the peptide do not form measurable populations of ordered helix, even in TFE. From the magnitude of the negative ellipticity at 222 nm in the CD spectrum in 30−50% TFE (Figure 1), it is evident that the population of helix in the central (D-helix) region of the peptide approaches 100% in TFE. This is consistent with the data in Figure 5. Allowing for the inevitable gaps in the series of medium-range NOEs, occasioned by the overlap or near overlap of resonances, panels A and B of Figure 5 show a similar density of (i,i+3) NOEs indicative of helix. There are, however, three major differences that indicate a greater population of helix in TFE, specifically in the center of the peptide. First, there is a clear $d_{\alpha N}(i,i+4)$ NOE between Val 104 and Lys 108 (also marked in Figure 4), indicating a high population of α-helix (Merutka et al., 1993). Second, the density of $d_{\alpha N}(i,i+2)$ NOEs is lower in the TFE solution than in the aqueous solution. These NOEs are characteristic of 3_{10} helix and turn conformations (Dyson et al., 1988a) and have been found in the NMR spectra of unfolded proteins (Smith et al., 1996). Groups of $d_{\alpha N}(i,i+2)$ NOEs are also observed in nascent helix, where interlinked turnlike conformations exist in equilibrium with unfolded states (Dyson et al., 1988b). The conformational ensemble in nascent helix contains a high proportion of backbone dihedral angles in the α region of (ϕ,ψ) space but shows little evidence of ordered helix by CD spectroscopy (Dyson et al., 1988b). Although $d_{\alpha N}(i,i+2)$ NOEs are sometimes observed in α-helices in proteins, it appears from the accumulated evidence that their absence in the TFE solution of the MDC peptide is most likely due to an increase in the proportion of ordered α-helical conformations at the expense of nascent helix or 3_{10} helix. The third piece of evidence for nearly complete formation of helix under these

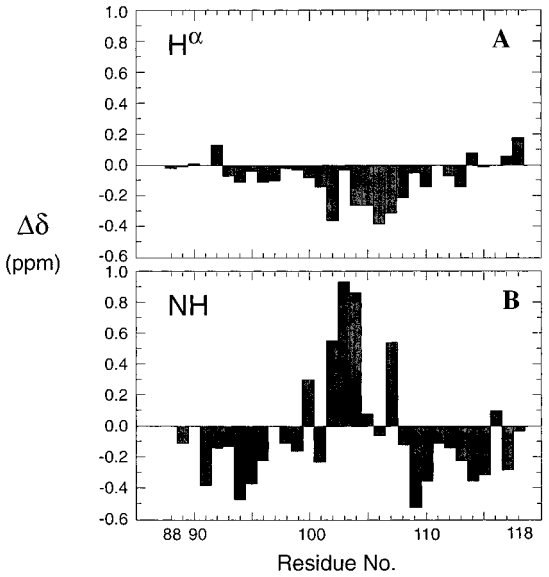

FIGURE 8: Difference between the chemical shift of a given resonance in 30% v/v TFE and the shift in 90% ^1H_2O/10% ^2H_2O plotted against residue number for (A) $C^{\alpha}H$ and (B) NH resonances.

conditions is the low intensity or even the absence of $d_{\alpha N}(i,i+1)$ NOEs in the region between Asp 97 and Thr 110 (Figure 4). This indicates that the population of unfolded states at these positions is very low indeed and forms a distinct contrast with the data from aqueous solution, where $d_{\alpha N}(i,i+1)$ NOEs are quite strong throughout the peptide.

Chemical Shift Information. The chemical shifts of many NH, $C^{\alpha}H$, and $C^{\beta}H$ resonances change dramatically upon addition of 30% v/v TFE. This is illustrated in Figure 8, which shows a plot of the difference in the $C^{\alpha}H$ and NH chemical shifts for each resonance in the two solvent conditions. The $C^{\alpha}H$ resonances have very little intrinsic sensitivity to TFE (Merutka et al., 1995); we therefore attribute these changes to differences in the population of helical states. In the helical region, the $C^{\alpha}H$ resonances move upfield, indicating increased helicity (Wishart et al., 1991) in TFE; interestingly, the largest changes are observed in the C-terminal half of the D-helix region, for residues 102−108. The $C^{\alpha}H$ resonance for the N-cap residue, Asp 92, moves downfield in TFE, perhaps an indication that the peptide is more firmly capped as the helix is more highly populated, giving (ϕ,ψ) values closer to the β region, and hence a downfield shift. Many of the $C^{\beta}H$ resonances move downfield (Table 3 of the Supporting Information), also consistent with an increased population of helix (Wishart et al., 1991).

The behavior of the amide proton resonances upon addition of TFE is complicated. Many resonances, including those of all amide protons of residues outside the helical region except that of Gly 116, shift upfield upon addition of TFE. Similar behavior has been observed in "random coil" peptides and has been attributed to the change of solvent from a good NH hydrogen bond acceptor (water) to a poor hydrogen bond acceptor (TFE) (Merutka et al., 1995). The observed solvent perturbations suggest that the amides of residues 88−96, 110−115, and 117 and 118 of the MDC peptide in water solution are largely exposed to, and hydrogen bonded to, the solvent. In contrast, most of the residues between Asp 97 and Lys 108, all of which are in the helical region, undergo much smaller upfield shifts or even experience downfield shifts upon changing the solvent to 30% TFE (Figure 8). The diminished upfield shift is attributed to the participation of these amide protons in intramolecular hydrogen bonding interactions due to formation of secondary structure in water solution (Urry & Long, 1976). The large downfield shifts experienced by some of the amide protons in the presence of TFE probably reflect formation of short helical hydrogen bonds in the highly populated helix formed in TFE; the magnitude and direction of the amide proton chemical shift are known to be highly sensitive to hydrogen bond length (Wagner et al., 1983). It is of interest that the downfield amide shifts are associated with the most hydrophobic region of the helix; similar correlations between amide proton chemical shift, hydrogen bond length, and hydrophobicity have been recognized in protein helices and other helical peptides (Kuntz et al., 1991; Zhou et al., 1992). It thus appears that local hydrophobic interactions in the MDC peptide play an important role in stabilizing the helical structure and promoting the formation of strong hydrogen bonds.

Indirect Coupling Constant Information. The NH resonances are generally broader in the TFE sample (Figure 2), and the cross-peaks in Figure 2B vary greatly in their intensity. The strongest cross-peaks are those for residues at the N and C termini, and the weakest are for those in the middle of the helical region in the center of the peptide. The low intensity of these COSY cross-peaks is most probably due to a decrease in the $^3J_{HN,H\alpha}$ coupling constant, consistent with the formation of a high proportion of helix in this part of the sequence; cancellation of COSY peaks is a consequence of small coupling constants (Neuhaus et al., 1985). Even in the peptide in aqueous solution, some of the cross-peaks, notably those of residues 102 and 105, are also of low intensity, consistent with their position in the center of the helical region of the peptide.

Extension of the Helix in TFE

According to the pattern of $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs (Figure 5), the helical region in aqueous solution spans residues Ala 93−Ile 107 or −Lys 108. The addition of TFE does not extend the helix significantly at either the N or the C terminus. The helix unambiguously ends at Lys 108 in TFE, while the evidence that Lys 108 is part of the helix in water consists only of a single overlapped $d_{\alpha\beta}(i,i+3)$ NOE. A weak $d_{\alpha N}(i,i+3)$ connectivity is observed between Ile 107 and Thr 110 in TFE (Figure 5). Such a connectivity is to be expected if Gly 108 is the C-capping residue in an α_L motif (Aurora et al., 1994). An apparent $d_{\alpha\beta}(i,i+3)$ NOE is observed in TFE between Lys 108 and Asp 111 (Figure 4), but this must be viewed with caution because of the possibility of overlap with the intraresidue NOE between α and ϵ protons [of the two other possible $d_{\alpha\beta}(i,i+3)$ NOEs involving the $C^{\alpha}H$ of lysine residues, 94/97 shows an unambiguous peak at 3.11 ppm, distinct from any $C^{\epsilon}H$ chemical shift, and 100/103 is not observable due to overlap (Figure 4)].

Relationship between the Amide Proton Chemical Shift and the Temperature Coefficient

The chemical shift of the amide proton is greatly influenced by its accessibility to solvent, mainly due to intramolecular hydrogen bonding interactions. The temperature

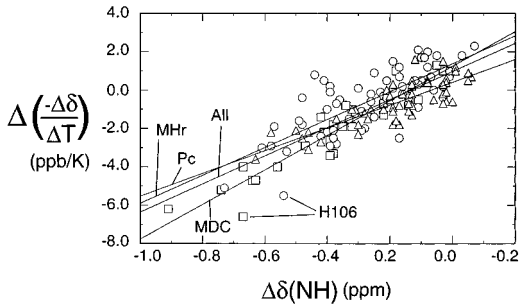

FIGURE 9: Plot of the difference between the temperature coefficient ($-\Delta\delta/\Delta T$ in units of parts per billion per kelvin) and the corresponding random value (Merutka et al., 1995) as a function of the difference between the chemical shift of the NH resonance and the corresponding random coil value (Merutka et al., 1995) for residues in the MDC peptide (□), residues in the four helical peptide segments from myohemerythrin (Dyson et al., 1992a) (○) and residues in nine of the peptides from plastocyanin (Dyson et al., 1992b) (Pc1, Pc3, Pc5, Pc7, Pc9, Pc13, Pc15, Pc17, and Pc18) (△). The outlying values for His 106 in the MDC peptide and in the D-helix are indicated. In all cases, the values for the two residues at the N terminus and for the C-terminal residue were omitted from the list, to avoid end effects. The straight lines represent linear least-squares fits to the data as indicated.

dependence of the chemical shift of this proton is also heavily dependent on hydrogen bonding interactions and the extent of protection of the amide proton from solvent. In Figure 9, we show the temperature coefficient plotted as a function of the amide proton chemical shift, both quantities corrected for sequence effects by subtracting random coil values (Merutka et al., 1995). The data shown are from a total of 14 peptides from myohemerythrin (Dyson et al., 1992a; present work) and plastocyanin (Dyson et al., 1992b). A striking linear dependence is observed between the two quantities, which appears to be independent of the set of peptides. To our knowledge, this is the first observation of such a relationship, which would perhaps be expected given that the two quantitites are both sensitive to the same influences, namely hydrogen bonding and solvation.

The C-Terminal Loop Retains Its Configuration

An unusual conformation was observed in the 10-residue peptide corresponding to the C-terminal loop of myohemerythrin (Dyson et al., 1992a), and the evidence from NOEs and temperature coefficients is that this structure is present unchanged in the longer MDC peptide. This conformation apparently consists of a close association between the hydrophobic portions of several lysine residues and the side chain of Tyr 114, which act to sequester from solvent the amide proton of the nearby Gly 116. The Gly 116 amide proton has an extremely low temperature coefficient, lower, in fact, than expected even for well-formed secondary structure. There is some evidence from the NOESY spectrum that there is a turnlike conformation in this region of the peptide (Figures 3 and 5A), but the major contribution to the lowering of the temperature coefficient is apparently the sequestration of the amide proton from solvent by the nearby residues. The chemical shift and temperature coefficient fit the correlation shown in Figure 9, suggesting that the observed upfield shift of the Gly NH resonance is dominated by intramolecular interactions other than ring current contributions from the tyrosine. Similar results have recently been reported in a series of peptides of the sequence Gly-Gly-Aro-Gly-Gly (Merutka et al., 1995), where Aro represents Phe, Tyr, or Trp, and for a peptide from BPTI (Kemmink et al., 1993). The chemical shifts and NOEs obtained for the MDC peptide in 30% TFE provide evidence that the conformation of the C-terminal sequence is retained in TFE. In particular, the Gly 116 NH resonance is still strongly shifted upfield relative to the random coil value. However, the $d_{\alpha N}(i,i+2)$ NOE between Tyr 114 and Gly 116 is not observable due to overlap of NH resonances.

Capping of the D-Helix

Sequence motifs that commonly appear at the N and C termini of helices in protein crystal structures have been termed N- and C-capping sequences (Baker & Hubbard, 1984; Harper & Rose, 1993). At the N- and C-terminal ends of the helix, the backbone NH and CO groups cannot form helical hydrogen bonds; instead, these groups are hydrogen bonded by the side chains of capping residues. Hydrophobic interactions also appear to play a role in capping (Seale et al., 1994; Muñoz et al., 1995). Capping facilitates initiation and termination of the helix, and also appears to stabilize helices in peptides and proteins (Harper & Rose, 1993; Lyu et al., 1992, 1993; Zhou et al., 1994a; Bruch et al., 1991; Chakrabartty et al., 1993a).

A typical N-capping sequence incorporating both the hydrogen bonding component and the hydrophobic interaction takes the form

$$N'\text{-}N_{cap}\text{-}N1\text{-}N2\text{-}N3\text{-}N4$$

where N′ and N4 are hydrophobic residues, N_{cap} and N3 are hydrophilic residues whose side chains participate in the hydrogen bonding interaction, and N1 and N2 can be any residue. Myohemerythrin, and hence the MDC peptide, contains such a sequence immediately preceding the N terminus of the D-helix (Sheriff et al., 1987)

$$V_{91}\text{-}D\text{-}A\text{-}K\text{-}N\text{-}V_{96}$$

where Ala 93 is the N-terminal residue assigned to the D-helix from the crystal structure of the intact protein (Sheriff et al., 1987) and constitutes the N terminus of the short D(C99A)-helix peptide (Dyson et al., 1992a); this residue also appears to serve as the helix start site in the MDC peptide (Figure 5). The N-capping interaction is potentially enhanced by a hydrophobic interaction between the side chains of Val 91 (N′) and Val 96 (N4) (Muñoz et al., 1995). Interactions typical of N-cap structures are seen in the crystal structure of the intact myohemerythrin protein (Sheriff et al., 1987) and include a hydrogen bond between the side chain of Asp 92 and the backbone amide proton of Asn 95 and an apparent hydrophobic interaction between the side chains of Val 91 and Val 96. The C-capping interactions in the crystal structure (Sheriff et al., 1987) appear to consist of a backbone $(i,i+5)$ hydrogen bond between Ile 107 CO and the amide proton of Phe 112 in the sequence

$$V_{104}\text{-}N\text{-}H\text{-}I\text{-}K\text{-}G\text{-}T\text{-}D\text{-}F_{112}$$

This sequence appears to follow the rules for an α_L helix termination motif (Aurora et al., 1994).

Effect of an Intact N-Capping Sequence on the Population of Helix in the MDC Peptide

The NOE (Figure 5) and other NMR data (Figure 6) provide evidence for an increase in helix in the region

corresponding to the D-helix sequence in the longer MDC peptide in water solution compared to the shorter D(C99A)-helix peptide. Two factors potentially contribute to this increase. The length of the peptide is increased, with consequent displacement of potential end effects to a greater distance from the helical sequence, and intact capping or termination sequences are now present at both the N and C termini of the D-helix region. A greater number of well-defined ($i,i+3$) NOE connectivities are present in the N-terminal portion of the D-helix sequence in MDC than in the D(C99A)-helix peptide (Figure 5), and the chemical shifts and amide proton temperature coefficients (Figure 6) show small but significant changes consistent with greater populations of helix.

A significant decrease is observed in the MDC peptide in both the chemical shift and the temperature coefficient of the backbone amide proton of Asn 95 (Figure 6), compared both to the shorter D(C99A)-helix peptide and to the surrounding residues in the MDC peptide. This may well reflect the influence of hydrogen bonding to the side chain of the N-cap residue, Asp 92, which would be typical of a capped structure and is one of the interactions present at the N terminus of the D-helix in the crystal structure (Sheriff et al., 1987). The NH backbone chemical shift and temperature coefficient of Asp 92 itself provide no evidence for the reciprocal hydrogen bond to the Asn 95 side chain carbonyl group expected in a typical N-terminal capping motif; however, we note that this hydrogen bonding interaction is not present in the crystal structure. Indirect evidence of hydrogen bonding from NMR thus supports the notion that an N-cap structure is present in the MDC peptide in aqueous solution.

On the other hand, direct NOE evidence for an N-capping structure (Lyu et al., 1993) in the MDC peptide is weak. The diagnostic $d_{NN}(i,i+3)$ NOE noted by Lyu et al. (1993) should be observable for Asp 92 and Asn 95 in this peptide but is absent in aqueous solution (Figure 3) and would be overlapped in 30% TFE (Figure 4). $d_{\beta N}(i,i+1)$ and $d_{\gamma N}(i,i+1)$ NOEs are observed between Val 91 and Asp 92, but these are to be expected, and are found, for most residues in peptides. NOEs are also expected between the amide proton of Asp 92 and the $C^{\beta}H$ and $C^{\delta}H$ of Asn 95 and between the NH of Ala 93 and the $C^{\beta}H$ of Asn 95; these are absent in aqueous solution and overlapped in 30% TFE. The only NOE evidence that a capping structure may be present in this sequence is the presence of a weak connectivity between the $C^{\beta}H_3$ of Ala 93 and the NH of Asp 92. None of the other NOEs mentioned by Lyu et al. (1993) as being characteristic of capping motifs are observed. Although some of the NMR data are consistent with formation of an N-cap motif, the absence of many of the expected NOEs suggests that its population is relatively small. Nevertheless, the fact that the population of helix is increased significantly, as judged by temperature coefficients, at the N terminus of the MDC peptide relative to the D(C99A)-helix peptide, suggests that it does help to stabilize helical structure.

Effect of an Intact C-Capping Sequence on the Population of Helix in the MDC Peptide

The helix appears to be more highly populated in the longer peptide in the C-terminal half (residues 101−108) of the D-helix sequence. A more consistent increase in the population of helical conformers appears to occur in this region, according to the NOEs [a greater density of overlapping ($i,i+3$) NOEs and the presence of two $d_{NN}(i,i+2)$ connectivities] and the chemical shifts and temperature coefficients. A $d_{\alpha N}(i,i+3)$ NOE is observed in TFE between Ile 107 and Thr 110 as expected for an α_L capping motif, but the evidence for the Ile 107−Phe 112 hydrogen bond seen in the crystal structure of the intact protein (Sheriff et al., 1987) is rather weak. The temperature coefficient of the Phe 112 amide proton is not particularly low (6.2 ppb/K), indicating that this amide is largely exposed to solvent. However, this value is significantly lower than that observed for the same residue in the short C-terminal peptide (Dyson et al., 1992a) (Figure 6), indicating that a small population of hydrogen-bonded structured forms may be present.

CONCLUSIONS

A number of lines of evidence suggest that the N- and C-capping sequences present in the longer MDC peptide function to prevent helix fraying and as helix stop signals in both water and TFE solutions. The C-capping α_L sequence motif appears to be particularly effective in terminating the helix, and the unusual conformation seen in a short peptide representing the C-terminal loop of myohemerythrin is retained in the MDC peptide, providing evidence that structures present in short peptide fragments of proteins persist in longer peptides and can be regarded as potential initiation sites for protein folding. The limitation of the observed helical structure to the region of the peptide corresponding to helix in the intact protein gives validity to the notion that capping sequences operate simply through local interaction dictated by the amino acid sequence and are not dependent on long-range tertiary interactions in the intact protein. The accumulation of NMR data for these and other peptide systems has allowed us to show that there is a linear relationship between two of the NMR parameters that are most indicative of hydrogen bonding, the chemical shift and temperature coefficient of the amide proton. A comparison of the NMR, CD, and fluorescence spectra of the wild-type sequence (Cys 99) with those of the MDC peptide (Ala 99) indicates an interaction between the side chains of Cys 99 and Trp 102 in the D-helix peptide in solution.

The most significant conclusion is that the inclusion of natural N- and C-capping sequences in the peptide both increases the population of helix in the central portion of the peptide and limits helix formation to this sequence. This result validates the use of peptides as models for the early events in protein folding; not only are local structures observed in folded proteins present in short peptides in aqueous solution, but these structures act in the peptides in the way that they are thought to act in the protein structure, stabilizing secondary structure in certain sequences while limiting the secondary structure to those sequences alone. These results have profound implications for the mechanism of protein folding. It is clear that local conformational propensities can and do bias the conformational search and facilitate the folding process.

ACKNOWLEDGMENT

We thank David Millar and Carlos Garcia for helpful advice concerning the fluorescence measurements.

SUPPORTING INFORMATION AVAILABLE

Three tables containing complete resonance assignments and temperature coefficients for the MDC peptide in aqueous solution (Table 2), for the MDC peptide in 30% TFE (Table 3), and for the D(C99A) peptide in aqueous solution (Table 4) (5 pages). Ordering information is given on any current masthead page.

REFERENCES

Aurora, R., Srinivasan, R., & Rose, G. D. (1994) *Science 264*, 1126−1130.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535−560.

Bruch, M. D., Dhingra, M. M., & Gierasch, L. M. (1991) *Proteins 10*, 130−139.

Chakrabartty, A., Schellman, J. A., & Baldwin, R. L. (1991) *Nature 351*, 586−588.

Chakrabartty, A., Doig, A. J., & Baldwin, R. L. (1993a) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11332−11336.

Chakrabartty, A., Kortemme, T., Padmanabhan, S., & Baldwin, R. L. (1993b) *Biochemistry 32*, 5560−5565.

Chen, Y.-H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.

Dyson, H. J., & Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519−538.

Dyson, H. J., & Wright, P. E. (1993) *Curr. Opin. Struct. Biol. 3*, 60−65.

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.

Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) in *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.

Fairman, R., Shoemaker, K. R., York, E. J., Stewart, R. L., & Baldwin, R. L. (1989) *Proteins 5*, 1−7.

Harper, E. T., & Rose, G. D. (1993) *Biochemistry 32*, 7605−7609.

Hider, R. C., Kupryszewski, G., Rekowski, P., & Lammek, B. (1988) *Biophys. Chem. 31*, 45−51.

Jiménez, M. A., Muñoz, V., Rico, M., & Serrano, L. (1994) *J. Mol. Biol. 242*, 487−496.

Kemmink, J., & Creighton, T. E. (1993) *J. Mol. Biol. 234*, 861−878.

Kemmink, J., van Mierlo, C. P. M., Scheek, R. M., & Creighton, T. E. (1993) *J. Mol. Biol. 230*, 312−322.

Kippen, A. D., & Fersht, A. R. (1995) *Biochemistry 34*, 1464−1468.

Klippenstein, G. L., Cote, J. L., & Ludlam, S. E. (1976) *Biochemistry 15*, 1128−1136.

Kuntz, I. D., Kosen, P. A., & Craig, E. C. (1991) *J. Am. Chem. Soc. 113*, 1406−1408.

Lyu, P. C., Zhou, H. X., Jelveh, N., Wemmer, D. E., & Kallenbach, N. R. (1992) *J. Am. Chem. Soc. 114*, 6560−6562.

Lyu, P. C., Wemmer, D. E., Zhou, H. X., Pinker, R. J., & Kallenbach, N. R. (1993) *Biochemistry 32*, 421−425.

Merutka, G., Morikis, D., Brüschweiler, R., & Wright, P. E. (1993) *Biochemistry 32*, 13089−13097.

Merutka, G., Dyson, H. J., & Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.

Muñoz, V., Blanco, F. J., & Serrano, L. (1995) *Nat. Struct. Biol. 2*, 380−385.

Neuhaus, D., Wagner, G., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257−273.

Odaert, B., Baleux, F., Huynh-Dinh, T., Neumann, J. M., & Sanson, A. (1995) *Biochemistry 34*, 12820−12829.

Rance, M. (1987) *J. Magn. Reson. 74*, 557−564.

Rance, M., & Byrd, R. A. (1983) *J. Magn. Reson. 52*, 221−240.

Rance, M., & Wright, P. E. (1986) *J. Magn. Reson. 66*, 372−378.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.

Reymond, M. T., Delmas, L., Koerber, S. C., Brown, M. R., & Rivier, J. E. (1992) *J. Med. Chem. 35*, 3653−3659.

Rizo, J., Blanco, F. J., Kobe, B., Bruch, M. D., & Gierasch, L. M. (1993) *Biochemistry 32*, 4881−4894.

Rose, G. D., Gierasch, L. M., & Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1−106.

Schnölzer, M., Alewood, P., Jones, A., Alewood, D., & Kent, S. B. H. (1992) *Int. J. Pept. Protein Res. 40*, 180−193.

Seale, J. W., Srinivasan, R., & Rose, G. D. (1994) *Protein Sci. 3*, 1741−1745.

Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 296−299.

Sheriff, S., Hendrickson, W. A., & Smith, J. L. (1987) *J. Mol. Biol. 197*, 273−296.

Shin, H.-C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J., & Wright, P. E. (1993a) *Biochemistry 32*, 6356−6364.

Shin, H.-C., Merutka, G., Waltho, J. P., Wright, P. E., & Dyson, H. J. (1993b) *Biochemistry 32*, 6348−6355.

Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., & Dobson, C. M. (1996) *J. Mol. Biol. 255*, 494−506.

Urry, D. W., & Long, M. M. (1976) *CRC Crit. Rev. Biochem. 4*, 1−45.

Viguera, A. R., & Serrano, L. (1995) *Biochemistry 34*, 8771−8779.

Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948−5949.

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. (1978) *Biopolymers 17*, 1451−1467.

Yao, J., Dyson, H. J., & Wright, P. E. (1994a) *J. Mol. Biol. 243*, 754−766.

Yao, J., Feher, V. A., Espejo, B. F., Reymond, M. T., Wright, P. E., & Dyson, H. J. (1994b) *J. Mol. Biol. 243*, 736−753.

Zhou, H. X., Lyu, P., Wemmer, D. E., & Kallenbach, N. R. (1994a) *Proteins 18*, 1−7.

Zhou, H. X., Lyu, P. C., Wemmer, D. E., & Kallenbach, N. R. (1994b) *J. Am. Chem. Soc. 116*, 1139−1140.

Zhou, N. E., Zhu, B.-Y., Sykes, B. D., & Hodges, R. S. (1992) *J. Am. Chem. Soc. 114*, 4320−4326.

Zimm, B. H., & Bragg, K. J. (1959) *J. Chem. Phys. 31*, 526−535.

BI970038X