A Noncovalent Peptide Complex as a Model for an Early Folding Intermediate of Cytochrome *c*†

Lawren C. Wu,‡,§ Paul B. Laub,∥,⊥ Gülnur A. Elöve,∥ Jannette Carey,*,‡ and Heinrich Roder*,∥,⊥

Chemistry Department, Princeton University, Princeton, New Jersey 80544, Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111, and Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received March 31, 1993; Revised Manuscript Received July 16, 1993•

ABSTRACT: Horse heart cytochrome *c* is one of a small number of proteins for which the folding pathway has been elucidated in structural detail by pulsed hydrogen exchange and NMR. Those studies indicated that a partially folded intermediate with interacting N- and C-terminal helices is formed at an early stage of folding when most of the chain is still disordered. This report describes a peptide model for this early intermediate, consisting of a noncovalent complex between a heme-containing N-terminal fragment (residues 1–38) and a synthetic peptide corresponding to the C-terminal helix (residues 87–104). Far-UV circular dichroism and proton NMR indicate that the isolated peptides are largely disordered, but when combined, they form a flexible, yet tightly bound complex with enhanced helical structure. These results emphasize the importance of interactions between marginally stable elements of secondary structure in forming tertiary subdomains in protein folding.

A major goal in the study of protein folding is to characterize intermediate structures along the folding pathway, in hope of elucidating the mechanism of protein folding. This effort is hindered, however, by the cooperative nature of folding and the short lifetimes of transient intermediates, which make it difficult to isolate and characterize partially folded states. Several approaches have been developed to circumvent this problem, including trapping of disulfide-bonded intermediates (Creighton, 1978; Weissman & Kim, 1991), pulsed hydrogen exchange with NMR analysis of refolding kinetics (Udgaonkar & Baldwin, 1988; Roder et al., 1988), and design of protein fragments that correspond to folding intermediates (Oas & Kim, 1988; Staley & Kim, 1990). A detailed view of the formation of hydrogen-bonded structure during folding can be obtained with the pulsed hydrogen-exchange method, which combines quenched-flow techniques with two-dimensional NMR analysis to monitor protection of amide protons against hydrogen–deuterium exchange during the time course of refolding. Fragment design guided by the structure of the native protein has also been used to study folding intermediates and structural subdomains. Alternatively, proteolytic methods allow identification of potentially important protein fragments without prior knowledge of intermediate structures (Tasayco & Carey, 1992).

The folding pathway of cytochrome *c* (cyt *c*) has been studied extensively (Ikai et al., 1973; Tsong, 1976; Ridge et al., 1981; Brems & Stellwagen, 1983; Ramdas & Nall, 1986; Roder et al., 1988; Jeng et al., 1990; Pryse et al., 1992; Elöve et al., 1992; Roder & Elöve, 1993). Cyt *c* (Takano & Dickerson, 1981; Bushnell et al., 1990) is a structurally well-characterized, globular protein of 104 residues with a heme group covalently attached via thioether linkages to cysteines 14 and 17 (Figure 1). The native protein contains three prominent α-helices grouped around a central hydrophobic core that includes one edge of the heme group. The N- and C-terminal helices cross each other at an angle of 86°, forming a tightly packed hydrophobic interface with a contact area of about 220 Å^2 (computed with a probe radius of 1.4 Å) and with some contact between the C-terminal helix and the heme group (~40 Å^2). This helix-pairing interaction is a highly conserved structural motif among the *c*-type cytochromes (Matthews, 1985), and the C-terminal peptide sequence is predicted consistently to be helical in the family of cytochrome *c* (Rooman & Wodak, 1992). Pulsed hydrogen-exchange studies (Roder et al., 1988) revealed an early (20 ms) folding event in which amide groups become protected only in chain segments corresponding to the N- and C-terminal helices. In the first 100 ms of folding, all other probes remain unprotected, including those in chain segments corresponding to other helical regions or other hydrogen-bonded but nonhelical regions of the native protein. The rate and degree of amide proton protection were similar in the N- and C-terminal helices, even when pulse-exchange conditions were varied (Roder et al., 1988; Elöve & Roder, 1991). These observations suggested early formation of a partially folded intermediate with interacting helical segments near the chain termini and no stable hydrogen-bonded structure elsewhere. Although evidence for this interaction is indirect, it is also consistent with the observation that isolated helices (e.g., in short peptides) are only marginally stable and provide little or no protection from hydrogen exchange (Scholtz & Baldwin, 1992).

Apocytochrome *c* has hydrodynamic and spectroscopic properties characteristic of a disordered coil (Stellwagen & Rysavy, 1972; Fisher et al., 1973; Damaschun et al., 1991), indicating that formation of the native cyt *c* structure relies on the covalently attached heme group. Involvement of the heme group in the contact between N- and C-terminal regions of the folded state (Takano & Dickerson, 1981; Bushnell et al., 1990) suggests it might play a role in helix–helix association during folding. Indeed, previous studies (D. Barrick, S.

† This work was supported by grants from the NIH (GM43558 to J.C., GM35926 to H.R., and CA06927 to Fox Chase Cancer Center) and the Pew Science Program in Undergraduate Education, Mid-Atlantic Cluster, to J.C.

* Address correspondence to these authors.
‡ Princeton University.
§ Present address: Whitehead Institute, Cambridge, MA 02142.
∥ Fox Chase Cancer Center.
⊥ University of Pennsylvania.

• Abstract published in *Advance ACS Abstracts,* September 1, 1993.

0006-2960/93/0432-10271$04.00/0

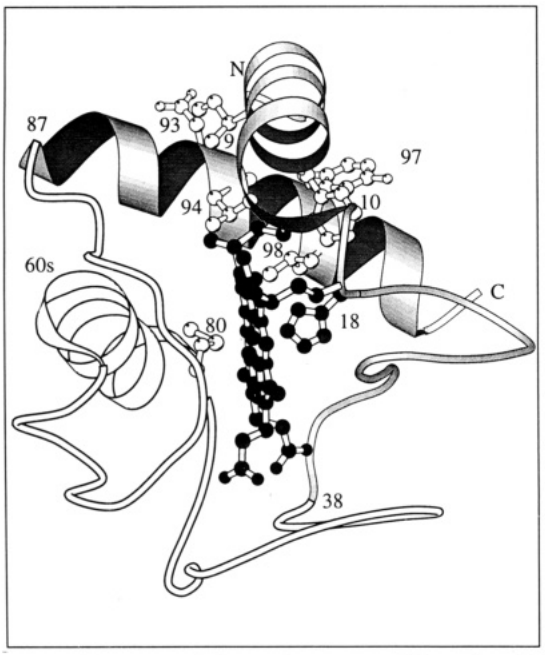

FIGURE 1: Structure of horse heart cyt *c* illustrating N- and C-terminal helices, the heme group, and side chains of residues 9, 10, 93, 94, 97, and 98 in the helix pairing site. Shaded regions identify the locations of the N-fragment (from the N-terminus to residue 38 including the heme) and the C-terminal peptide (from residue 87 to residue 104). Filled circles represent atoms of the heme group and its linkages to the polypeptide (residues 14, 17, and 18). This figure was created with the coordinates of Bushnell et al. (1990) and the program Molscript (Kraulis, 1991).

Marqusee, and R. L. Baldwin, personal communication; Kuroda, 1993) of synthetic cyt *c* peptides lacking the heme group showed no evidence for interaction in the absence of covalent linkage. Thus, we have prepared a proteolytic heme-containing cyt *c* fragment and a synthetic C-terminal peptide to mimic the early folding intermediate. For a detailed and comprehensive review of the extensive literature on cyt *c* fragments, nearly all of it resulting from the elegant work of Taniuchi and co-workers, see Fisher and Taniuchi (1992). Fragment complementation systems have been investigated for several other proteins, including classically ribonuclease A (Richards & Vithayathil, 1959) and staphylococcal nuclease (Taniuchi & Anfinsen, 1969); for general reviews, see Zabin and Villarejo (1975) and Jaenicke (1991). Many of these studies have involved at least one relatively large fragment, and in most cases the entire polypeptide chain is represented among the fragments, sometimes with partial redundancies. In contrast, the present work illustrates that two rather small peptides can interact noncovalently, with concomitant enhancement of helix content, even though each is only marginally helical in isolation and much of the intervening polypeptide chain is absent.

MATERIALS AND METHODS

Preparation of the heme-containing fragment of residues 1–38 from cyt *c* followed the method of Juillerat et al. (1980). This procedure limits trypsin digestion to the two arginine residues at positions 38 and 91 of cyt *c* by reversibly blocking lysines. Briefly, cyt *c* (Sigma, type VI) was derivatized by addition of citraconic anhydride in the presence of 6 M guanidine hydrochloride at pH 8.5 and was purified by dialysis against 0.1 M ammonium bicarbonate, pH 8.3. Trypsin was added in weight ratio of cyt *c* of 1 to 200, and after 5 h phenylmethanesulfonyl fluoride was added to 1 mM. Fragments were separated on a Sephadex G-50 (superfine) column in dialysis buffer. Fractions in only one peak absorbed at both 276 and 530 nm; these were pooled, lyophilized, dissolved in 30% acetic acid, and incubated at ~25 °C for 24 h. Deblocked N-fragment thus obtained was repurified on a G-50 column in 10% formic acid, followed by a sulfopropyl-Sephadex SP-25 column with a gradient from 20 mM to 0.3 M ammonium bicarbonate, pH 7.0. Purity of N-fragment was assessed by analytical gel electrophoresis, amino acid analysis, and HPLC. Concentrations of N-fragment were calculated using the previously determined extinction coefficient, $\epsilon_{406} = 1.14 \times 10^5\ M^{-1}\ cm^{-1}$ (Parr et al., 1978). Heme fragment 11–21 of cyt *c* (microperoxidase-11) from Sigma was used without further purification.

Hemeless N-fragment (corresponding to residues 1–38 of apocytochrome *c*) was prepared by chemically removing the heme group from N-fragment using methods developed for intact cyt *c* (Stellwagen et al., 1972; Fisher et al., 1973; S. Hu, personal communication). Briefly, N-fragment was incubated in 8 M urea, 0.1 M sodium chloride, and 80 mg/mL mercuric chloride in the dark for 4 h at 40 °C and for a further 8 h at ~25 °C and then purified on a G-50 column in 0.1 M acetic acid. Fractions with absorbance at 276 and 394 nm in a ratio greater than 2.5 were pooled, lyophilized, and incubated in 50 mM ammonium acetate, pH 5.0, 8 M urea, and 0.4 M dithiothreitol in the dark at ~25 °C for ~9 h and then purified on a G-50 column in 50 mM ammonium acetate, pH 5.0. Purity was assessed by analytical gel electrophoresis and by UV–visible spectroscopy.

The C-terminal peptide [(Ac)EPTEREDLIAYLKEATNK(NH_2)] was synthesized by R. Pidikiti (University of Pennsylvania) on a Milligen 9050 peptide synthesizer using Fmoc chemistry. The peptide was purified by reverse-phase HPLC. Purity was assessed by HPLC and by peptide sequencing on a Milligen 6600 solid-phase sequencer and by NMR. Concentrations were calculated using the extinction coefficient for free tyrosine, $\epsilon_{275} = 1420\ M^{-1}\ cm^{-1}$.

CD spectra were obtained on an Aviv Model 62DS spectropolarimeter equipped with a Peltier thermal-controlled cell holder in a 1- or 0.1-mm path-length cell. Spectra were acquired with 1.5-nm bandwidth and a step size of 0.2 nm. ^1H NMR T_1 relaxation measurements were performed at 600 MHz on 1.5 mM C-terminal peptide (0.05 M sodium phosphate and 0.05 M sodium acetate buffer in H_2O, pH 7.4, 10 °C) before and after addition of N-fragment to final concentrations of 100 or 625 μM N-fragment. Inversion–recovery data were measured by using a composite 180° pulse (Levitt & Freeman, 1981), recording spectra (80 acquisitions) at 13 delays ranging from 0.01 to 2.5 s. The water resonance was suppressed by presaturation during a 3-s period prior to the 180° pulse. T_1 relaxation times were determined for 25 fully or partially resolved resonances, based on a three-parameter exponential fit of the normalized resonance intensities (peak heights) as a function of the relaxation delay (10 ms–2.5 s).

RESULTS AND DISCUSSION

The N-terminal fragment obtained by trypsin digestion of citraconylated cyt *c* (Juillerat et al., 1980) consists of the first 38 amino acid residues, corresponding to the N-terminal helix (residues 4–14), additional turn structures, and the heme group covalently attached via its native thioether bonds (Parr & Taniuchi, 1979). The synthetic C-terminal peptide includes residues 87–104 of cyt *c*, which encompass the C-terminal helix (residues 87–102). Sequence changes (Lys 87 to Glu,

Lys 88 to Pro, Lys 100 to Glu, and Glu 104 to Lys) were introduced in an attempt to enhance helix propensity (Scholtz & Baldwin, 1992) while preserving potential helix–helix contact points. The Lys 87 to Glu and Glu 104 to Lys substitutions, together with the N- and C-terminal blocking groups, are expected to enhance helix stability by providing favorable electrostatic interactions with the partial charges of the exposed amide and carbonyl groups in the first and last turns of the helix, respectively (Shoemaker et al., 1987; Fairman et al., 1989). Substitution of Pro for Lys 88 was based on statistical data indicating a preference for proline in the first helical turn in globular proteins and the fact that Pro is frequently found in this position in *c*-type cytochromes from plants, insects, and bacteria (Moore & Pettigrew, 1990). The Lys 100 to Glu replacement was intended to permit formation of a salt bridge with the new Lys 104. This charge reversal between positions 100 and 104 occurs in several cytochromes *c* from birds and reptiles (Moore & Pettigrew, 1990). All of these amino acid changes are in the first and last turns of the C-terminal helix (Figure 1), and thus are well removed from the interface with the N-terminal helix in native cyt *c* (Bushnell et al., 1990).

CD spectroscopy of the N-fragment at 10 °C, in 20 mM sodium phosphate, pH 7, showed slight (~15–20%) helix content. The N-fragment has a molar ellipticity, $[\theta]_{222}$, of −6700 deg cm^2 dmol^{-1}, indicating a helix content of ~17% of the theoretical maximum ($[\theta]_{222}$ = −39 500 deg cm^2 dmol^{-1} for all residues in helical conformation in all molecules; Scholtz et al., 1991). For comparison, the synthetic N-terminal peptide of cyt *c* studied by Kuroda (1993) comprising residues 1–15 was only ~3% helical at 5 °C in water. The very different helix contents of these two peptides could be due to an influence of the heme group on stability of the N-terminal helix. The CD signal of N-fragment is linearly dependent on peptide concentration up to 0.3 mM (Figure 2a), indicating the absence of aggregation over this concentration range. The heme absorption spectrum of N-fragment (Parr & Taniuchi, 1979) at neutral pH shows a Soret band at 406 nm, typical for a low-spin ferric heme with two strong-field axial iron ligands, probably including His 18 as the fifth ligand as in native cyt *c* and His 26 or His 33 as the (non-native) sixth ligand. Below neutral pH, the Soret band shifts toward 396 nm, consistent with dissociation of at least one axial ligand due to protonation of the His side chains (Parr & Taniuchi, 1979). The native sixth ligand of intact cyt *c* is Met 80, which is not present on the fragments studied here.

CD spectroscopy of the C-terminal peptide at 10 °C, in 20 mM sodium phosphate, pH 7, indicated a helix content of ~15% ($[\theta]_{222}$ = −5800 deg cm^2 dmol^{-1}). By contrast, the C-terminal peptide (residues 89–103 of horse cyt *c*) studied by Kuroda (1993) is ~27% helical at 5 °C in water. The lower helix content of the peptide studied here may be due largely to the Pro 88 substitution. A recent study of the isolated C-peptide of ribonuclease A showed its helix content was unaffected by Pro substitution in the first helical residue but reduced by substitution in all other positions examined (Strehlow et al., 1991). The CD signal of the present C-terminal peptide was independent of peptide concentration up to 2 mM (Figure 2b). NMR line broadening indicative of aggregation was observed for the C-terminal peptide at concentrations above 2 mM only. The proton NMR spectrum of the C-terminal peptide at 1 mM is consistent with a largely disordered conformation.

CD was used to follow changes in secondary structure upon addition of C-terminal peptide to N-fragment (Figure 3). The

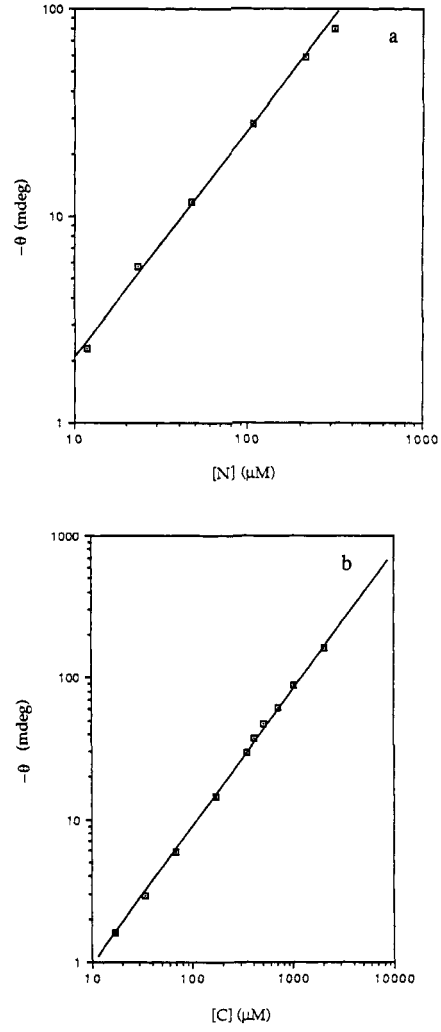

FIGURE 2: Concentration dependence of the CD signal for N-terminal (a) and C-terminal (b) peptide fragments. The signal intensity at 222 nm was measured for samples prepared in 20 mM sodium phosphate, pH 7.0, at 10 °C. Data are presented on a log–log scale in order to better distribute the points.

diagnostic helical signals at 208 and 222 nm increased in intensity (became more negative) when the two peptides were combined. In particular, CD difference spectra (obtained by subtracting the spectra of the individual peptides from the spectrum of the mixture) are markedly helical at high peptide concentrations (Figure 3a). These observations provide direct evidence for association between the N- and C-terminal regions, with concomitant stabilization of helical structure in at least one of the peptides. The ability of the isolated peptides to form a binary complex even without covalent linkage, and with approximately half of the polypeptide chain absent, provides strong support for the idea that the terminal helices can interact in the early folding intermediate of intact cyt *c* (Roder et al., 1988).

Figure 3b shows the increase in negative ellipticity at 222 nm (corresponding to an increase in helicity) as a function of C-terminal peptide concentration, with N-fragment concentration held constant at ~300 μM. The binding isotherm implies that the increase in helicity reports on the concentration of the peptide–fragment complex. The value of $[\theta]_{222}$ begins to increase at ~20 μM C-terminal peptide, has an inflection near 150 μM, and reaches a plateau above ~500 μM. The maximum increase in helicity achieved at the plateau ($[\theta]_{222}$ = −9300 deg cm^2 dmol^{-1}) corresponds to approximately 50% of the theoretical maximum ($[\theta]_{222}$ = −19 000 deg cm^2 dmol^{-1})

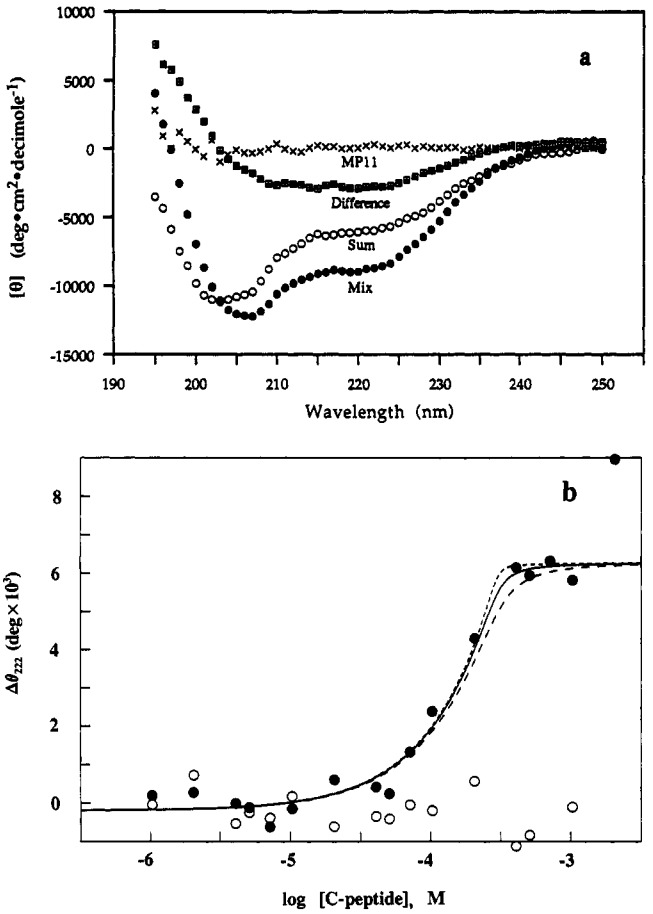

FIGURE 3: Far-UV circular dichroism evidence for complex formation. (a) Mixture sum, and difference spectra of 300 μM N-fragment and 400 μM C-terminal peptide (0.1-mm path length). The sum spectrum was obtained by adding the individual spectra of 300 μM N-fragment and 400 μM C-terminal peptide. The mixture spectrum was obtained by preparing a sample with the indicated concentrations of both N-fragment and C-terminal peptide. The difference spectrum was obtained by subtracting the sum spectrum from the mixture spectrum. The spectrum labeled MP11 is a control difference spectrum for 75 μM C-terminal peptide and 30 μM microperoxidase-11 (1-mm path length). (b) CD titration of C-terminal peptide into N-fragment. Individual samples were prepared for each titration point. Increases in the CD signal of mixture spectra vs sum spectra were calculated at 222 nm from difference spectra recorded in a 0.1-mm cuvette and are plotted against the total concentration of C-terminal peptide added. Symbols: closed circles, 300 μM N-fragment; open circles 200 μM hemeless N-fragment; solid line, theoretical binding curve calculated for a dissociation constant, K_d, of 5×10^{-6} M by using the relationship $NC = (1/2)(C_t + N_t + K_d - [(C_t + N_t + K_d)^2 - 4N_tC_t]^{1/2})$, where NC is the concentration of the 1:1 complex of N- and C-terminal fragments and N_t and C_t are the total concentrations of the individual fragments; dashed lines, theoretical curves calculated for $K_d = 1 \times 10^{-6}$ M (to the left of the solid line) and $K_d = 10 \times 10^{-6}$ M (right). The outlying point at very high C-terminal peptide concentration is presumed to result from aggregation (cf. Figure 2) and was not included in the fitting.

expected for a native complex with 27 of 56 residues in helical conformation. The shape of the binding isotherm, with an inflection where the concentration of C-terminal peptide is approximately half that of N-fragment, and a plateau above, suggests formation of a 1:1 complex with a dissociation constant (K_d) far below 300 μM (Weber, 1992). Indeed, the titration data can be fit with a theoretical binding curve calculated by assuming 1:1 complex formation and $K_d \cong 5$ μM. This magnitude of K_d is comparable to that of many specific protein–protein interactions.

The CD spectrum of hemeless N-fragment shows very little helix content and is insensitive to addition of C-terminal peptide over a similar concentration range (Figure 3b). In another control experiment, we found no evidence for complex formation between the C-terminal peptide and heme-containing fragment 11–21 of horse cyt *c* (microperoxidase-11), which lacks most of the N-terminal helix. The far-UV CD spectrum of a mixture of 30 μM microperoxidase and 75 μM C-terminal peptide was, within error, identical to the sum of the spectra of the individual peptides (Figure 3a). Because of the limited solubility of microperoxidase, the spectra were recorded at lower concentrations and longer path length (1 mm) than spectra involving N-fragment shown in Figure 3a. A parallel experiment on 43 μM N-fragment and 75 μM C-terminal peptide resulted in a difference spectrum very similar to that obtained at higher peptide concentration in a 0.1-mm cuvette. Together, these control experiments indicate that peptide–fragment complex formation requires both an intact N-terminal segment and a covalently attached heme group, suggesting a relatively specific interaction.

Addition of excess C-terminal peptide (up to ~1 mM) had no measurable effect on the heme absorption spectrum of N-fragment in the Soret and visible regions, indicating that the non-native sixth heme ligand is retained upon complex formation and that the heme group remains largely solvent-exposed in the peptide–fragment complex. Solvent exposure of the heme suggests only limited contact between the C-terminal peptide and the heme group in the complex. Kinetic and spectroscopic studies on cyt *c* folding provide evidence for a solvent-exposed heme group and a non-native histidine ligand to the iron atom during the early stages of folding, including the intermediate with interacting N- and C-terminal helices (Brems & Stellwagen, 1983; Elöve & Roder, 1991; Roder & Elöve, 1993).

Complex formation has been observed previously with much larger fragments of cyt *c* [reviewed in Fisher and Taniuchi (1992)] and other proteins (Zabin & Villarejo, 1975; Jaenicke, 1991). Models for protein folding intermediates with small peptides, however, have generally been assembled from covalently connected fragments (Oas & Kim, 1988; Staley & Kim, 1990; Kuroda, 1993). Previous work with synthetic peptides corresponding to the N- and C-terminal regions of cyt *c* but lacking the heme group showed no evidence of interaction unless the peptides were covalently linked (Barrick, S. Marqusee, and R. L. Baldwin, personal communication; Kuroda, 1993). In fact, Kuroda found that a covalent homodimer of C-terminal peptides is more helical than the covalent heterodimer between one N-terminal and one C-terminal peptide. In the present study, we observe association between the heme-containing N-fragment and the C-terminal peptide at concentrations far below those where the separate peptides begin to aggregate, indicating that heterodimer formation is strongly favored in the presence of the covalently attached heme group.

NMR studies that could provide structural details on the peptide complex are complicated by line broadening caused by the paramagnetic heme iron in the N-fragment (cf. Figure 4C). A substantial increase in line widths was observed in the proton NMR spectrum of the C-terminal peptide upon addition of N-fragment. This line broadening can be attributed in part to bulk paramagnetic effects, but several resonances showed excess broadening compared to the chemical shift standard TSP. To take advantage of the paramagnetic center, we measured the effect of added N-fragment on longitudinal relaxation rates ($1/T_1$) for protons of the C-terminal peptide

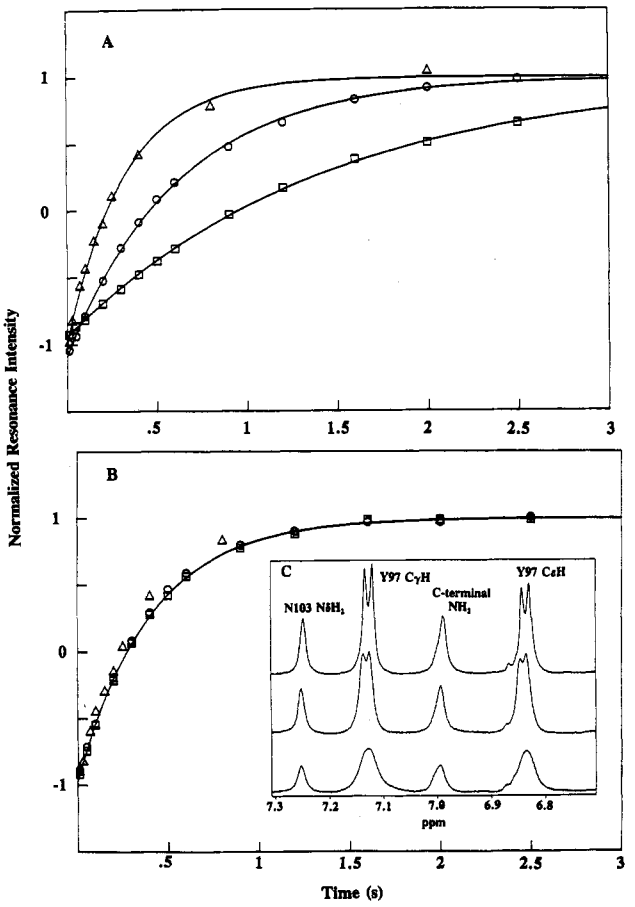

FIGURE 4: Evidence for complex formation, based on paramagnetic effects of the N-terminal heme fragment on longitudinal relaxation rates for proton resonances of the C-terminal peptide. Two extreme examples are illustrated in panels A and B. The relaxation rate for the Tyr 97 $C_\epsilon H$ resonance is accelerated by factors of 2.3 and 4.6 in the presence of 100 and 625 μM N-peptide, respectively (panel A). Among the 25 resonances analyzed, only two others showed significant paramagnetically enhanced relaxation rates, namely, the Tyr 97 $C_\delta H$ protons (enhanced 1.6-fold at 100 μM and 2.5-fold at 625 μM N-peptide) and a composite line at 0.88 ppm containing methyl resonances from Leu 94, Ile 95, and Ile 98 (enhanced 1.3-fold at 100 μM and 1.9-fold at 625 μM N-peptide). The data for Asn 103 $N_\delta H_2$ resonances are representative for the remaining 22 resonances, which experienced an average paramagnetic T_1 enhancement of 1.11 ± 0.09 at 100 μM and 1.23 ± 0.15 at 625 μM N-fragment. These values are similar to the corresponding enhancements of 1.12 and 1.37 measured for the resonance of the chemcial shift standard [TSP, 3-(trimethylsilyl)tetradeuteriopropionate]. Panel C is an expanded region from the equilibrium spectra of the C-terminal peptide alone (upper) and in the presence of 100 μM (middle) or 625 μM N-fragment (lower), showing Tyr 97 ring proton resonances, Asn 103 side-chain amide resonances, and a resonance from the NH_2 blocking group on the C-terminus.

by using conventional proton NMR relaxation techniques (Farrar & Becker, 1971). Due to the inverse sixth power distance dependence of electron–nuclear dipolar interactions, enhancement of proton relaxation rates by paramagnetic centers provides a sensitive probe of molecular structure and interactions (Mildvan & Cohn, 1970; Wüthrich, 1976). Longitudinal relaxation rates were measured for 25 at least partially resolved C-terminal peptide resonances in the presence of 0, 100, or 625 μM N-fragment (Figure 4). Some resonances, most notably those from the ring protons of Tyr 97 and a composite line containing methyl proton resonances of Leu 94, Ile 95, and Leu 98, showed large paramagnetic enhancement of relaxation rates (up to 4.6-fold at 625 μM N-fragment; Figure 4A), whereas the average enhancement observed for all other resonances (1.23 ± 0.15 at 625 μM N-fragment; Figure 4B) was similar to that for TSP (1.37), consistent with the bulk paramagnetism of the solution.

In native cyt c, the methyl groups of Leu 94 and Leu 98 are in direct contact with one edge of the heme group and within ~10 Å of the iron atom, with the side chains of Leu 95 and Tyr 97 nearby and within ~12 Å of the iron atom (Bushnell et al., 1990; cf. Figure 1). Proton resonances of these residues show the largest paramagnetic effects, with Tyr 97 being the most strongly affected. All other residues of the C-terminal helix lie more than 13 Å from the heme iron. Although the observed paramagnetic relaxation effects in the peptide–fragment complex are generally consistent with these distances, their magnitudes do not provide a precise meter stick. Thus, the peptide–fragment complex appears to be a good model for the corresponding region of the intact protein, although at present we cannot conclude that it is an exact replica.

These results provide strong evidence for formation of a complex in which certain C-terminal peptide side chains are brought into close proximity with the heme group, while most residues of the C-terminal peptide reside at a large average distance from the heme. Consistent with this picture is the observation that peptide–fragment association is accompanied by only minor changes in chemical shifts, suggesting that although the two peptides are tightly associated, they do not appear to form a rigid structure. This flexibility, plus rapid association–dissociation rates, results in extensive conformational averaging, reducing the difference in chemical shifts between bound and free forms. Nevertheless, complex formation appears specific in that the C-terminal peptide interacts preferentially with the heme-containing N-fragment, indicating that binding is not due merely to nonspecific interactions between two amphiphatic helices nor between the hydrophobic region of the C-terminal peptide and the heme, as indicated by the control experiment with the shorter heme fragment. Such tightly associated, yet flexible peptide complexes have been observed, e.g., in the designed four-helix bundle protein α_2, where two pairs of antiparallel helices form a high-affinity complex but remain relatively flexible (Regan & DeGrado, 1988; Handel et al., 1993).

Demonstration of a noncovalent peptide complex mimicking an early folding intermediate of cytochrome c confirms previous evidence that this intermediate represents a structural subdomain with mutually interacting N- and C-terminal helices. Association of the peptide fragments leads to enhanced helicity and requires the covalently attached heme group, suggesting that subdomains are stabilized by tertiary contacts between marginally stable elements of secondary structure. Together with hydrogen-exchange studies on the refolding kinetics of cyt c (Roder et al., 1988), these data suggest that formation and association of the N- and C-terminal helices early in the folding process may direct subsequent steps in the folding pathway. Association of chain termini may be a common early step in protein folding that could explain the observation (Thornton & Sibanda, 1983) that many proteins have their termini nearby in the native state. Further study of protein subdomains, and of the spontaneous, noncovalent interactions among their constituent structural elements, will be useful in elucidating the protein folding problem.

REFERENCES

Brems, D. N., & Stellwagen, E. (1983) *J. Biol. Chem. 258*, 3655–3660.

Bushnell, G. W., Louie, G. V., & Brayer, G. D. (1990) *J. Mol. Biol. 214*, 585–595.

Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231–297.
Damaschun, G., Damaschun, H., Gast, K., Gernat, C., & Zirwer, D. (1991) *Biochim. Biophys. Acta 1078*, 289–295.
Elövé, G. A., & Roder, H. (1991) *ACS Symp. Ser. 470*, 50–63.
Elöve, G. A., Chaffotte, A. F., Roder, H., & Goldberg, M. F. (1992) *Biochemistry 31*, 6876–6883.
Fairman, R., Showmaker, K. R., York, E. J., Stewart, J. M., & Baldwin, R. L. (1989) *Proteins: Struct., Funct., Genet. 5*, 1–7.
Farrer, T. C., & Becker, E. D. (1971) *Pulse and Fourier Transform NMR*, Academic Press, New York.
Fisher, A., & Taniuchi, H. (1992) *Arch. Biochem. Biophys. 296*, 1–16.
Fisher, W. R., Taniuchi, H., & Anfinsen, C. B. (1973) *J. Biol. Chem. 248*, 3188–3195.
Handel, T. M., Williams, S. A., & DeGrado, W. F. (1993) *Science 261*, 879–885.
Ikai, A., Fish, W., & Tanford, C. (1973) *J. Mol. Biol. 73*, 165–184.
Jaenicke, R. (1991) *Biochemistry 30*, 3147–3161.
Jeng, M. F., Englander, S. W., Elöve, G. A., Wand, A. J., & Roder, H. (1990) *Biochemistry 29*, 10433–10437.
Juillerat, M., Parr, G. R., & Taniuchi, H. (1980) *J. Biol. Chem. 255*, 845–853.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Kuroda, Y. (1993) *Biochemistry 32*, 1219–1224.
Levitt, M. H., & Freeman, R. (1981) *J. Magn. Reson. 43*, 64–80.
Mathews, F. S. (1985) *Prog. Biophys. Mol. Biol. 45*, 1–56.
Mildvan, A. S., & Cohn, M. (1970) *Adv. Enzymol. Relat. Areas Mol. Biol. 33*, 1–70.
Moore, G. R., & Pettigrew, G. W. (1990) in *Cytochromes: Evolutionary, Structural, and Physicochemical Aspects*, pp 116–127, Springer-Verlag, New York.
Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42–48.
Parr, G. R., & Taniuchi, H. (1979) *J. Biol. Chem. 254*, 4836–4842.
Parr, G. R., Hantgen, R. R., & Taniuchi, H. (1978) *J. Biol. Chem. 253*, 5381–5388.
Pryse, K. M., Bruckman, T. G., Maxfield, B. W., & Elson, E. L. (1992) *Biochemistry 31*, 5127–5136.
Ramdas, L., & Nall, B. T. (1986) *Biochemistry 25*, 6959–6964.
Regan, L., & DeGrado, W. F. (1988) *Science 241*, 976–978.
Richards, F. M., & Vithayathil, P. J. (1959) *J. Biol. Chem. 234*, 1459–1465.
Ridge, J. A., Baldwin, R. L., & Labhardt, A. M. (1981) *Biochemistry 20*, 1622–1630.
Roder, H., & Elöve, G. A. (1993) in *Frontiers in Molecular Biology*, IRL/Oxford Press, Oxford, England (in press).
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700–704.
Rooman, M. J., & Wodak, S. J. (1992) *J. Mol. Biol. 221*, 961–979.
Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95–118.
Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M., & Baldwin, R. L. (1991) *Biopolymers 31*, 1463–1470.
Staley, J. P., & Kim, P. S. (1990) *Nature 334*, 685–688.
Stellwagen, E., Rysavy, R. J., & Babul, G. (1972) *J. Biol. Chem. 274*, 8074–8077.
Strehlow, K. G., Robertson, A. D., & Baldwin, R. L. (1991) *Biochemistry 30*, 5810–5814.
Takano, T., & Dickerson, R. E. (1981) *J. Mol. Biol. 153*, 79–94.
Tasayco, M. L., & Carey, J. (1992) *Science 255*, 594–597.
Thornton, J. N., & Sibanda, B. L. (1983) *J. Mol. Biol. 167*, 443–460.
Tsong, T. Y. (1976) *Biochemistry 15*, 5467–5473.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694–699.
Weber, G. (1992) *Protein Interactions*, pp 31–70, Chapman-Hall, London.
Weissman, J., & Kim, P. S. (1991) *Science 253*, 1386–1393.
Wüthrich, K. (1976) *NMR in Biological Research: Peptides and Proteins*, North-Holland/American Elsevier, New York.
Zabin, I., & Villarejo, M. R. (1975) *Annu. Rev. Biochem. 44*, 295–313.