Side Chain Packing of the N- and C-Terminal Helices Plays a Critical Role in the Kinetics of Cytochrome *c* Folding†

Wilfredo Colón,‡ Gülnur A. Elöve,‡,§ L. Paul Wakem,||,⊥ Fred Sherman,|| and Heinrich Roder*,‡,#

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, Department of Biochemistry, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, and Department of Biochemistry & Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059

Received January 9, 1996; Revised Manuscript Received March 4, 1996⊗

ABSTRACT: The pairing of two α-helices at opposite ends of the chain is a highly conserved structural motif found throughout the cytochrome *c* family of proteins. It has previously been shown that association of the N- and C-terminal helices is a critical early event in the folding process of horse cytochrome *c* and is responsible for the formation of a partially folded intermediate (I_{NC}). In order to gain further insight into the structural and energetic basis of helix packing interactions and their role in folding, we prepared a series of horse cytochrome *c* variants in which Leu94, a critical residue at the helix contact site, was replaced by Ile, Val, or Ala. The Ile and Val substitutions resulted in minor changes in the stability of the native state, indicating that conservative mutations can be accommodated at the helix interface with only minor structural perturbations. In contrast, the L94A mutation resulted in a 3.5 kcal/mol decrease in unfolding free energy, suggesting that the smaller Ala side chain causes severe packing defects at the helix interface. The effect of these mutations on the kinetics of folding and unfolding as a function of denaturant concentration was studied by a systematic series of stopped-flow fluorescence measurements. The proteins with Leu, Ile, or Val at position 94 exhibit a major unresolved fluorescence change during the 1-ms dead time of the stopped-flow refolding measurements, while this effect is less pronounced in L94A, indicating that the rapid formation of a compact state (I_C) with largely quenched Trp59 fluorescence is favored by a large hydrophobic side chain at the helix−helix interface. Despite their small effects on overall stability, the L94I and L94V mutations result in a substantial reduction in the relative amplitude of the fastest observable folding phase (formation of I_{NC}) consistent with a strong decrease in the population of I_{NC} compared to the wild-type protein. This effect is amplified in the case of the destabilizing L94A variant, which exhibits slower folding kinetics and negligible accumulation of I_{NC}. Whereas the presence of a large hydrophobic side chain at position 94 is sufficient for the stabilization of I_C, the subsequent partially folded intermediate, I_{NC}, is stabilized by specific interactions that are responsible for the proper packing of the two α-helices.

This study addresses the role of helix−helix packing interactions in protein folding, using horse cytochrome *c* (cyt *c*)[1] as a model. This 104-residue protein is a member of an extended family of heme proteins involved in electron transfer functions in mitochondrial respiratory chains and bacterial photosynthesis (Meyer & Kamen, 1982; Moore & Pettigrew, 1990). Despite their highly divergent amino acid sequences, these proteins all share some common structural features, including the covalent linkage of the heme group to two cysteine residues (14 and 17 in horse cyt *c*), the axial ligands of the heme iron (His18 and Met80), and a cluster of three α-helices grouped around one edge of the heme. The structure of horse cyt *c* (Bushnell et al., 1990) illustrated in Figure 1 is representative of the less diverse group of eukaryotic cytochromes *c* (Mathews, 1985). The perpendicular arrangement of the N- and C-terminal helices and close-packed hydrophobic residues at the interface, including the Leu94 residue targeted in this study, are particularly well conserved (Pielak et al., 1995).

Chothia et al. (1981) classified helix packing interactions in globular proteins according to the relative orientation of the helix axes and the pattern of side chain intercalation (ridges into grooves) at the interface. In the most frequently observed arrangement (class 4−4 packing), the helices cross at an angle of about −50°, while other cases (class 1−4 or 3−3 packing) exhibit a nearly perpendicular orientation with a mean angle of −105° (Chothia et al., 1981). The mitochondrial cytochromes *c* feature examples from each of these categories. In tuna cytochrome *c* (Takano & Dickerson, 1981a,b), which is structurally very similar to the horse protein (Bushnell et al., 1990) illustrated in Figure 1, the N-terminal (residues 5−11) and C-terminal (90−101) helices

† This work was supported by NIH Grants GM35926 (to H.R.) and GM12702 (to F.S.), by NSF Grants MCB9306367 (to H.R.) and CHE9123792 (to F.S.), and by NIH Grant CA06927 and an appropriation from the Commonwealth of Pennsylvania to the Institute for Cancer Research. W.C. is supported by a National Science Foundation Minority Postdoctoral Fellowship (BIR-9406859) awarded in 1994.

* Correspondence should be addressed to this author at the Fox Chase Cancer Center. Phone: (215) 728-3123. Fax: (215) 728-3574. E-mail: H_Roder@fccc.edu.

‡ Fox Chase Cancer Center.

§ Current address: Regeneron Pharmaceuticals, Inc., 777 Old Saw Mill River Road, Tarrytown, NY 10591-6707.

|| University of Rochester School of Medicine and Dentistry.

⊥ Current address: Department of Dermatology, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642.

University of Pennsylvania.

⊗ Abstract published in *Advance ACS Abstracts,* April 15, 1996.

[1] Abbreviations: cyt *c*, horse heart cytochrome *c*; GuHCl, guanidine hydrochloride; CD, circular dichroism; PCR, polymerase chain reaction.

S0006-2960(96)00052-9 CCC: $12.00

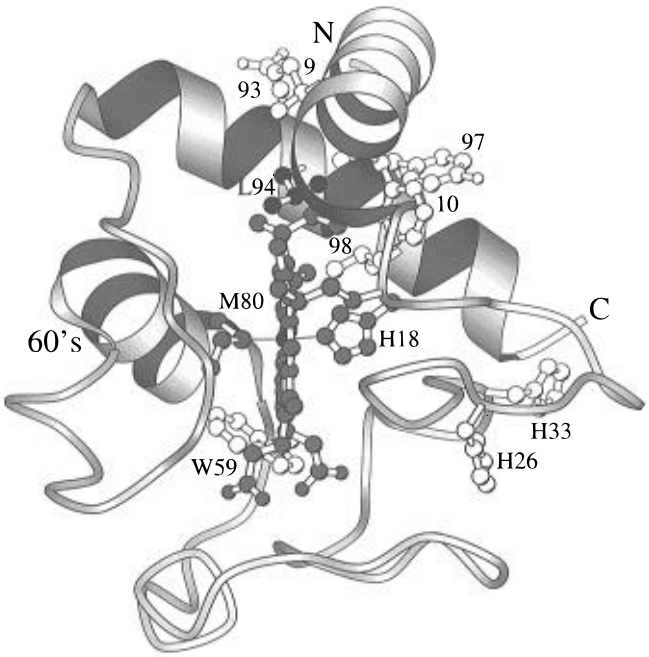

FIGURE 1: Ribbon diagram of horse cytochrome *c* based on the crystal structure (Bushnell et al., 1990). Side chains involved in the interaction between the N- and C-terminal helices and other key residues are shown explicitly. The figure was prepared by Michael Sauder, using the program Molscript (Kraulis, 1991).

cross each other at an angle of −102° and exhibit class 3−3 packing (Chothia & Lesk, 1985). Due to the presence of a glycine at position 6, which is found in 105 of 106 known eukaryotic cyt *c* sequences [compiled by Pielak et al. (1995)], the contact is unusually tight with a minimal interaxis distance of 6.8 Å. A third helix (the 60's helix, residues 60−70) packs against the C-terminal helix from the opposite side, forming a class 4−4 contact with an angle of −49° and interaxis distance of 9.2 Å (Chothia & Lesk, 1985).

Previous structural and kinetic studies on the cyt *c* folding mechanism [reviewed in Roder and Elöve (1994)] have shown that the pairing of the N- and C-terminal helices is among the earliest detectable structural events in cyt *c* folding, while packing interactions with the 60's helix are established in a subsequent step. As a result, a partially folded intermediate (termed I_{NC}) accumulates within 10−20 ms under certain refolding conditions (pH above 6). This conclusion is supported by the following observations: (1) under the conditions of our initial hydrogen exchange labeling studies (pH 6.3, 10 °C, 0.7 M GuHCl), hydrogen-bonded amide protons in or adjacent to the N-terminal helix (residues 8−15 and His18) and in the C-terminal helix (residues 90−100) are protected against deuterium-hydrogen exchange in a common folding phase with a rate of ∼50 s^{-1} (Roder et al., 1988); (2) all other amide probes, including those in the 60's helix, acquire significant protection only in a subsequent phase of folding with a rate of ∼2 s^{-1}; (3) amide protons on either side of the helix−helix interface have very similar protection factors (Elöve & Roder, 1991); (4) a heme-containing N-terminal fragment (1−38) and a synthetic peptide comprising the C-terminal helix (87−104) readily associate in solution and form a noncovalent 1:1 complex with enhanced helix content (Wu et al., 1993).

By using Trp59−heme fluorescence energy transfer and far-UV CD to monitor the formation of compact states and secondary structure during cyt *c* folding, we found evidence for a compact intermediate (I_C) with native-like helix content, which accumulates during the first few milliseconds of refolding prior to the formation of I_{NC} (Elöve et al., 1992). The absence of an aromatic CD band and the lack of significant amide protection indicate that I_C is a loosely folded state without stable and persistent hydrogen-bonded structure. Some of these properties are similar to those of the A-state of cyt *c*, a compact equilibrium state observed at low pH and high ionic strength (Stellwagen & Babul, 1975; Goto et al., 1990, 1993; Jeng et al., 1990; Kataoka et al., 1993). However, in contrast to both I_C and I_{NC}, the A-state contains highly protected amide protons in all three helices and some nonhelical regions (Jeng et al., 1990). Thus, the A-state appears to have a more stable and more native-like conformation than either of the kinetic intermediates.

Under the refolding conditions used in most of our earlier studies (pH 6.3, 10 °C, 0.7 M GuHCl), the folding mechanism can be adequately described by a simple mechanism with two sequential intermediates: I_C, which accumulates on the submillisecond time scale and is in rapid equilibrium with unfolded molecules, and I_{NC}, which is well populated in the 10−100 ms time regime. However, the folding mechanism of cyt *c* is complicated by the presence of the heme and its axial ligands, which affect both the kinetics of folding and unfolding (cf. Figure 7). The involvement of heme ligation in folding was investigated by measuring the effects of pH and addition of extrinsic ligands on the cyt *c* folding kinetics (Brems & Stellwagen, 1983; Elöve & Roder, 1991; Sosnick et al., 1994; Elöve et al., 1994). In the GuHCl-denatured state at neutral pH, one of the axial ligands to the heme iron, His18, remains bound, but the second one, Met80, is replaced by another deprotonated histidine, such as His26 or His33 (cf. Muthukrishnan & Nall, 1991). Due to its slow rate of dissociation (20−100 s^{-1}), this nonnative ligand remains bound during the initial stages of folding, including formation of I_{NC}, and its dissociation represents the rate-limiting step for completing the folding process. If binding of nonnative His ligands is prevented (e.g., by lowering the pH to 5, or addition of imidazole), the protein can bypass this slow step and the folding process becomes more cooperative. The linkage of His33 (or His26) to the heme iron does not interfere with the docking of the N- and C-terminal helices, but it may prevent the proper folding of the 60's helix, which is in close contact with the Met80 ligand in the native structure (Figure 1). In contrast to Sosnick et al. (1994), who suggested that I_{NC} represents a misfolded state, we feel that the presence of an alternative histidine ligand does not perturb the early conformational events but traps an intrinsic structural intermediate and renders it experimentally observable. Our kinetic model predicts that a structural intermediate analogous to I_{NC} occurs even in the absence of nonnative ligands. However, since the subsequent structural events are fast, this intermediate does not accumulate to detectable levels (Elöve et al., 1994). Our view is supported by the fact that the hydrogen exchange labeling experiments revealed only native-like hydrogen-bonded interactions (Roder et al., 1988).

In order to gain further insight into the structural and energetic basis of helix pairing interactions in cyt *c* and their role in folding, we prepared a series of horse cyt *c* variants in which Leu94, a central residue at the helix−helix interface was replaced by either Ile, Val, or Ala. The results of our equilibrium and kinetic experiments show that the stability of the native state and the population of an early folding intermediate (I_C) are relatively insensitive to conservative

changes at position 94 (L94I and L94V). In contrast, the subsequent I_{NC} intermediate is highly destabilized by these mutations, indicating that the pairing of the N- and C-terminal helices relies on specific interactions that are sensitive to even minor stereochemical changes of interfacial residues. These observations indicate that the initial compaction of the polypeptide chain is driven by relatively nonspecific hydrophobic interactions, while subsequent partially folded intermediates are stabilized by more specific tertiary packing interactions.

MATERIALS AND METHODS

Materials. Horse heart ferricytochrome *c* (type VI from Sigma Chemical Co.) was used without further purification. Amberlite CG-50 resin was purchased from Sigma. All other chemicals were reagent grade.

Preparation of Cyt c Mutants. The gene *CYC1-HORSE*, encoding horse cytochrome *c*, was previously constructed with synthetic oligodeoxyribonucleotides having preferred codons and with 5′ and 3′ untranslated regions of the iso-1-cytochrome *c* gene from the yeast *Saccharomyces cerevisiae* (Hickey et al., 1991). *CYC1-HORSE* was expressed in *S. cerevisiae* at approximately 66% of the normal wild-type level of yeast cytochrome *c*. The recombinant cyt *c* was shown to corresponded to the horse sequence, except that only approximately 70% of the N-terminal glycine residue was acetylated, and that approximately 80% of lysine 72 was trimethylated (Hickey et al., 1991).

The availability of *CYC1-HORSE* provides a convenient means to construct altered forms of horse cytochrome *c* by oligonucleotide-directed mutagenesis. In this study, we have prepared horse cytochrome mutants where Leu94 was replaced by Ile (*CYC1−1100*), Val (*CYC1−1099*), and Ala (*CYC1−1098*) using the method of Kunkel et al. (1987) and the plasmid *pAB831* containing *CYC1-HORSE*. The oligonucleotides contained the following codon changes at position 94 (Leu-TTG): Ile-ATT, Val-GTT, and Ala-GCT. Each of the altered genes were integrated in the *CYC1* region of the yeast strain B-6748 as previously described (Hickey et al., 1991). The sequences of these *CYC1* genes were verified by amplifying the appropriate region by PCR and subsequently sequencing the double-stranded DNA.

Purification of Cyt c Expressed in Yeast. Horse cyt *c* expressed in yeast was purified by a modification of the method of Sherman et al. (1968). A 12-L culture of yeast was grown at 30 °C in YPD media (1% yeast extract, 2% peptone, and 3% dextrose) for 3 days. Cyt *c* was then extracted from yeast by adding 0.5 mL of a 1 M NaCl solution per gram of cell paste and 0.25 mL/g ethyl acetate and stirring at 5 °C for at least 18 h. Approximately 2.5 mL of water per g of yeast and 10% (w/w) of Amberlite CG-50 coarse resin (based on harvested yeast weight) were added. After 2 h the resin containing the bound cytochrome *c* was decanted and washed with distilled water until the supernatant was clear. The resin was loaded onto a column and eluted with 0.8 M KCl in 0.1 M sodium phosphate buffer at pH 7. The colored eluate was brought to 90% saturation with ammonium sulfate. The cytochrome *c* remained in solution while most of the impurities precipitated. Final purification was obtained by ultrafiltration through a YM10 and YM30 Amicon membranes, with 10 and 30 kDa MW cutoffs, respectively. Final protein stocks were obtained by concentration/dilution cycles using an Amicon concentrator with YM5 membrane. Protein concentration was determined spectrophotometrically using an extinction coefficient for oxidized cyt *c* of 1.06×10^5 at 409 nm (Babul & Stellwagen, 1972). The yield of pure oxidized cyt *c* from a 12-L culture was 4−8 mg.

Equilibrium Fluorescence Measurements. Fluorescence measurements were performed on an Aminco-Bowman Series 2 fluorescence spectrometer (SLM-Aminco, Urbana, IL), using a 1-cm quartz cuvette thermostated at 10 °C. Excitation and emission wavelengths were 280 and 350 nm, respectively, and both slits were set to 4 nm. A 2−4 μM solution of folded (100 mM sodium phosphate, pH 7.0) and unfolded (5−6 M GuHCl, 100 mM sodium phosphate, pH 7.0) cyt *c* was prepared. The fluorescence was measured for the unfolded protein and solutions at progressively lower concentrations of GuHCl, achieved by repetitive withdrawal of unfolded cyt *c* from the cuvette followed by addition of the corresponding volume of folded cyt *c* solution. The fluorescence signal of cyt *c* was normalized relative to an equimolar solution of *n*-acetyl tryptophanamide in 6 M GuHCl and 0.1 M sodium phosphate, pH 7.0. The free energy of unfolding in the absence of denaturant, ΔG_{H_2O}, was obtained by nonlinear least-squares analysis of the data, based on a two-state model with a linear dependence of unfolding free energy, ΔG, on denaturant concentration, $\Delta G = \Delta G_{H_2O} - mC = m(C_m - C)$ (Pace, 1986; Tanford, 1970; Santoro & Bolen, 1992).

Fluorescence-Detected Stopped-Flow Kinetics. Most kinetic experiments were performed on a Bio-Logic SFM-4/QS stopped-flow instrument (Molecular Kinetics, Pullman, WA) equipped with a 0.8×0.8 mm^2 cuvette. The dead time was determined to be in the range of 1−1.5 ms (depending on the flow rate used). A 150 W Xe/Hg arc lamp and a monochromator (0.5 mm slit widths) were used for excitation at 280 nm and a high-pass glass filter with 324 nm cutoff (No. 51255, Oriel Corp., Stratford, CT) for measuring the tryptophan fluorescence emission. For some experiments, we used a PQ/SF-53 stopped-flow instrument (Hi-Tech, Salisbury, England) as described in Elöve et al. (1994). On both instruments, the fluorescence changes from 64 μs to ∼100 s were recorded in a single kinetic trace, using a previously described logarithmic averaging method (Khorasanizadeh et al., 1993).

The kinetics of refolding was measured at 10 °C following 6-fold dilution (Hi-Tech) or 6−20-fold dilution (Biologic) of an unfolded solution of cyt *c* in 4.2 M GuHCl (100 mM phosphate, pH 7.0) to various final GuHCl concentrations within and below the unfolding transition. Final protein concentrations were 10−30 μM. The kinetics of unfolding was measured by 6-fold dilution of native cyt *c* (100 mM phosphate, pH 7.0) into various GuHCl concentrations within or above the transition region.

The amplitudes of the unresolved process occurring in the dead time (burst phase) and the observable kinetic phases were normalized with respect to the total expected fluorescence change at the final GuHCl concentrations used. The initial step was to normalize all the data with respect to the signal of the unfolded protein (in 4.2 M GuHCl). An equilibrium denaturation curve of cyt *c* was obtained from the kinetic data by plotting the final signal at long refolding times (>60 s) of every kinetic experiment. The expected signal of native and unfolded cyt *c* at the various GuHCl concentrations was obtained by extrapolation of the native and unfolded baselines. The fractional amplitude of unre-

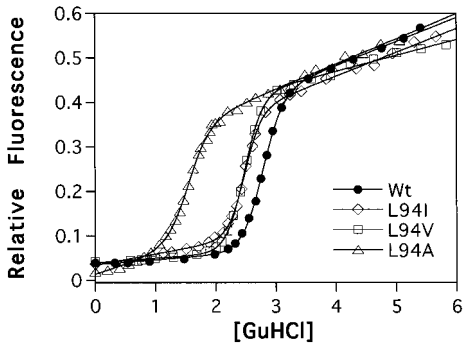

FIGURE 2: Comparison of the fluorescence-detected GuHCl denaturation transitions for wild-type and L94 variants of oxidized horse cyt *c* at 10 °C in 0.1 M sodium phosphate, pH 7.0. The tryptophan fluorescence emission at 350 nm was normalized relative to an equimolar solution of *N*-acetyl-L-tryptophanamide.

Table 1: Stability of Horse Cytochrome *c* Mutants[a]

	C_m (M)	m (kcal M^{-1} mol^{-1})	ΔG_{H_2O} (kcal mol^{-1})
WT	2.77 ± 0.02	3.5 ± 0.4	9.7 ± 1.1
L94I	2.47 ± 0.02	4.0 ± 0.3	9.9 ± 0.7
L94V	2.47 ± 0.02	3.6 ± 0.3	8.9 ± 0.7
L94A	1.67 ± 0.02	3.7 ± 0.3	6.2 ± 0.5

[a] All measurements were performed at 10 °C. Cyt *c* solutions contained 0.1 M sodium phosphate, pH 7.0. Equilibrium transition curves were calculated by nonlinear least-square analysis of the fluorescence signal as a function of GuHCl concentration as described in Materials and Methods. The slopes and intercepts of the pre- and posttransition baselines were used as fitting parameters, in addition to C_m and m. The errors are the standard deviation from 3−5 independent experiments.

solved kinetic events (burst phase), A_0, is given by the ratio $[S_U - S(0)]/[S_U - S_N]$, where S_U and S_N correspond to the denaturant-dependent signal of the unfolded and native protein, respectively, and $S(0)$ represents the initial signal at $t = 0$. The total observable amplitude (A_{obs}) is given by $1 - A_0$. The fractional amplitudes for the individual observable phases, A_i, were obtained by dividing the fluorescence amplitude of each phases, S_i by the total expected fluorescence change, $S_U - S_N$.

RESULTS

Effect of L94 Mutations on the Stability of the Native State. The reversible GuHCl-induced unfolding transition of cyt *c* was monitored by using tryptophan fluorescence as a probe of conformational changes. In the native structure of cyt *c* (Bushnell et al., 1990), the indole ring of Trp59 (the only tryptophan in the protein) lies within 9 Å of the heme iron (cf. Figure 1), resulting in almost complete quenching of its fluorescence. The unfolding transition is accompanied by a large increase in fluorescence indicative of a major increase in the average heme−tryptophan distance (Tsong, 1976). Figure 2 compares the GuHCl-induced denaturation transitions for wild-type cyt *c* and the variants with L94I, L94V, and L94A mutations, and Table 1 lists the corresponding two-state equilibrium parameters (midpoint concentrations, C_m, slopes, m, and the derived free energy in the absence of denaturant, $\Delta G_{H_2O} = mC_m$). The L94I and L94V mutations result in a ∼0.3 M decrease in C_m, suggesting a slight decrease in stability. However, due to a compensating increase in the m values, ΔG_{H_2O} remains unchanged within experimental error (Table 1). On the other hand, the L94A variant shows a C_m of 1.7 M, compared to 2.8 M for the wild-type, which corresponds to a 3.5 kcal/mol decrease in ΔG_{H_2O} (Table 1).

Burst Phase. Early stages of folding were investigated by monitoring the fluorescence changes occurring during the dead time of stopped-flow refolding experiments. In Figure 3, the initial signal (extrapolated to $t = 0$) and the final baseline signal for each kinetic trace is plotted as a function of the final GuHCl concentration. The fluorescence was scaled relative to that of unfolded cyt *c* in 4.2 M GuHCl, where the fluorescence of Trp59 reaches about 50% of that of free tryptophan (Tsong, 1976). The GuHCl dependence of the final signal (Figure 3) is in close agreement with the equilibrium unfolding curves in Figure 2, indicating that equilibrium is reached at the end of each kinetic measurement. The initial signal expected for fully unfolded molecules under refolding conditions is approximated by the dashed lines in Figure 3 obtained by linear extension of the posttransition (unfolded) baseline to lower GuHCl concentrations. At low denaturant concentrations, the observed initial signal falls far below the expected level, indicating that an early intermediate, I_C, with low fluorescence yield accumulates during the dead time of the stopped-flow instrument (∼1 ms). The fact that the initial fluorescence continues to decrease down to 0 M GuHCl without leveling off suggests that the fluorescence of Trp59 is largely quenched in I_C, consistent with a compact conformation having a relatively short average donor−acceptor distance. The results for wild-type cyt *c* (Figure 3A) confirm and extend previous measurements at a single GuHCl concentration of 0.7 M (Elöve et al., 1992), where we measured 35% "missing amplitude" ($A_0 = 0.35$). Toward lower GuHCl concentration, A_0 increases further, reaching a value of about 0.7 in the absence of denaturant (cf. solid triangles in Figure 5A). Since I_C is formed rapidly (>1000 s^{-1}) compared to subsequent steps (<300 s^{-1}), it can reach a pre-equilibrium with unfolded states. Thus, the normalized amplitude, A_0, approximates the population of I_C, and the decrease in A_0 observed in refolding experiments at higher final denaturant concentrations reflects the shift in the pre-equilibrium toward the unfolded state as the intermediate is destabilized.

All of the variants with Leu94 substitution show evidence for accumulation of a burst-phase intermediate under sufficiently stabilizing conditions. The initial signal for the L94I variant is similar to the wild type (Figure 3B), whereas L94V shows small deviations below 1 M GuHCl (Figure 3C). On the other hand, the L94A mutation (Figure 3D) results in a significantly lower burst phase amplitude with a limiting value for A_0 of 0.4 at 0 M GuHCl, compared to 0.65−0.7 for the other variants. Thus, substitution of Leu94 with hydrophobic side chains of similar size has little effect on the population of the compact early intermediate, while the smaller, less hydrophobic Ala side chain is highly destabilizing.

Mutational Effects on the Kinetics of Folding at 0.7 M GuHCl. Figure 4 compares the time course of refolding for the four variants at a final GuHCl concentration of 0.7 M. Each trace was normalized relative to the extrapolated unfolded baseline value at 0.7 M GuHCl (dashed lines in Figure 3), so that the initial fluorescence at short times reflects the missing amplitude effect described above. As in previous studies on wild-type cyt *c* under the same conditions (Roder et al., 1988; Elöve et al., 1992), the observable folding kinetics for all variants can be deconvoluted into three exponential phases: a fast phase with a relaxation time in the 10−100 ms range, an intermediate phase in the 0.5−1 s range, and a minor slow phase in the

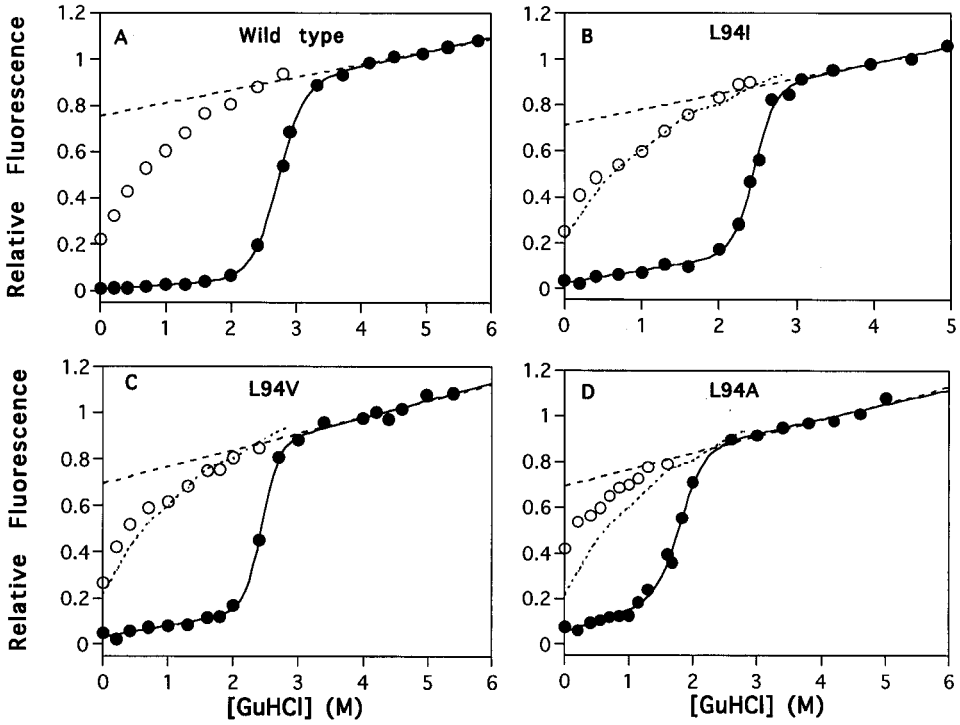

FIGURE 3: Denaturant dependence of the initial and final fluorescence amplitudes from stopped-flow refolding and unfolding experiments on horse cyt *c* and three Leu94 variants at 10 °C, pH 7.0 (0.1 sodium phosphate). The baseline signal at long times (solid symbols) vs [GuHCl] closely corresponds to the equilibrium denaturation curves shown in Figure 2; the corresponding C_m and *m* values obtained by two-state analysis (solid curves) are in close agreement with those listed in Table 1. The dashed line (- - -) indicates the expected signal for fully unfolded cyt *c* at any given [GuHCl]. The deviation of the initial signal (open symbols) from this line toward lower [GuHCl] indicates accumulation of an intermediate during the ~1 ms dead time of the stopped-flow measurement. The data were normalized relative to the unfolded signal at 4.2 M GuHCl. The short dashed lines in panels B−D indicate the amplitude behavior of wild-type cyt *c* for comparison with the mutants.

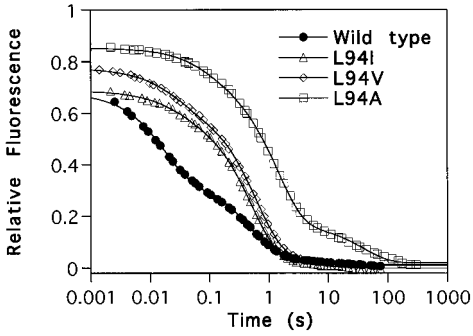

FIGURE 4: Representative fluorescence-detected stopped-flow traces for the refolding of wild-type and Leu94 mutants of horse cyt *c* plotted on a logarithmic scale. Refolding was initiated by rapid dilution of the denatured protein in 4.2 M GuHCl to a final GuHCl concentration of 0.7 M (pH 7.0 and 10 °C). The fluorescence was normalized with respect to the total expected signal (cf. Figure 3). The curves represent nonlinear least-squares fits of three exponential phases with the parameters listed in Table 2.

5−25 s range. The corresponding rates and amplitudes are listed in Table 2. On the basis of our earlier pulsed H/D exchange results, the fast phase reflects the formation of a partially folded intermediate with interacting N- and C-terminal α-helices (Roder et al., 1988), while the intermediate phase can be attributed to the formation of the native state following dissociation of a nonnative histidine−iron ligand (Elöve & Roder, 1991; Elöve et al., 1994; Sosnick et al., 1994). The slow phase is most likely due to *cis*/*trans* isomerization of prolyl peptide bonds (Ridge et al., 1981; Wood et al., 1988).

It is apparent from Figure 4 that the wild-type protein exhibits a much more pronounced fast phase than any of the mutants. Comparison of the kinetic parameters for the L94I and L94V mutants with those of the wild type (Table 2) shows that the amplitude of the fast phase, A_1, is 2−3-fold lower for the mutants, while the amplitude of the intermediate phase, A_2, increases by a similar amount. The rates k_1 and k_2 show a decreasing trend, by about a factor of 2 in the case of L94V. The large changes in A_1 and A_2 caused by the Ile and the Val substitutions is especially surprising, since these mutations have little or no effect on the stability of the native state (Table 1) and the burst-phase intermediate (A_o in Table 2). The destabilizing L94A mutation results in at least 3-fold slower rates for all three kinetic phases and a major amplitude increase for the intermediate phase at the expense of the fast phase.

For a more complete kinetic characterization of the L94 variants, we studied the kinetics of folding and unfolding over a wide range of denaturant concentrations. In Figure 5A−D, the rates of folding and unfolding for each variant are plotted on a logarithmic scale vs GuHCl concentration (upper panel), and the corresponding normalized amplitudes are shown in the lower panel. The fractional amplitude, A_i, of each observable phase was calculated by dividing the observed amplitude by the total expected signal, $S_U - S_N$. The complex denaturant dependence of the rates seen for the three observable refolding phases (open symbols), as well as the single-exponential unfolding process (solid diamonds), is a clear indication of the presence of multiple intermediates, both under refolding and unfolding conditions.

Unfolding Kinetics. The denaturant dependence of the rate of unfolding (solid diamonds in Figure 5A−D) shows similar characteristics for cyt *c* and all variants studied. Above the midpoint of the GuHCl-induced unfolding transition, the rate increases sharply (by nearly two orders of magnitude) but

Table 2: Effect of Leu94 Mutations in Horse Cyt *c* on the Kinetics of Refolding[a]

	k_0 (s^{-1})	A_0	k_1 (s^{-1})	A_1	k_2 (s^{-1})	A_2	k_3 (s^{-1})	A_3
WT	>1000	0.34 ± 0.04	64 ± 6	0.39 ± 0.04	2.2 ± 0.1	0.22 ± 0.04	0.20 ± 0.06	0.05 ± 0.01
L94I	>1000	0.31 ± 0.04	43 ± 10	0.12 ± 0.04	2.4 ± 0.8	0.53 ± 0.04	0.25 ± 0.10	0.04 ± 0.01
L94V	>1000	0.23 ± 0.03	32 ± 5	0.15 ± 0.03	1.4 ± 0.1	0.57 ± 0.02	0.18 ± 0.04	0.05 ± 0.01
L94A	>1000	0.15 ± 0.04	19 ± 6	0.09 ± 0.04	0.66 ± 0.1	0.61 ± 0.03	0.04 ± 0.03	0.15 ± 0.02

[a] Rate constants and normalized amplitudes of the three kinetic phases observed in fluorescence-detected refolding experiments at 10 °C (Figure 4). Refolding conditions: 0.7 M GuHCl and 0.1 M sodium phosphate buffer, pH 7.0. The errors represent the standard deviation from 3−4 independent experiments on fresh samples.

begins to level off at high GuHCl concentrations. The steep linear increase in the log(rate) vs [GuHCl] plot is indicative of a major conformational transition involving a large increase in denaturant-accessible surface area [see, e.g., Matthews (1993), Matouschek et al. (1994), and Khorasanizadeh et al. (1996)]. In contrast to most other proteins, the rate profile for cyt *c* curves downward at higher GuHCl concentrations, approaching a limiting rate of about 50 s^{-1} at 5 M. This behavior suggests a change in the rate-limiting step in unfolding from the typical denaturant-dependent transition to a nearly denaturant-independent process that limits the rate of unfolding under strongly denaturing conditions. This observation provides clear evidence for the presence of a native-like intermediate that precedes the main conformational barrier in unfolding of cyt *c*.

The kinetics of unfolding for L94I and L94V (Figure 5B,C) are very similar to those of the wild type (Figure 5A), indicating that the barriers separating the native and intermediate states from the unfolded state are not significantly perturbed by these mutations. In the case of the L94A variant, the rate of the GuHCl-dependent structural transition above C_m increases nearly 10-fold relative to the wild type, which is consistent with the observed decrease in equilibrium stability (Table 1). However, the rate-limiting process at high GuHCl concentrations shows only a slight increase (from 50 to 65 s^{-1} at 5 M GuHCl).

Denaturant Dependence of the Folding Kinetics. Under most conditions, the kinetics of folding is well described by three exponential phases (upper panels in Figure 5). However, at low GuHCl concentrations (<0.5 M), the fast phase splits into two distinct phases separated by a factor of 5−10 in rate. Since this complication occurs only for the wild type, L94I, and L94V below 0.5 M GuHCl, we chose to simplify the analysis by combining their amplitudes (open circles in lower panels). For all variants, the slowest phase (open diamonds) accounts for only a few percent of the total amplitude at low denaturant concentrations and reaches about 20% in the transition region. The denaturant dependence of this slow phase is roughly parallel to that of the intermediate phase, suggesting that it is due to a minor population of molecules following a parallel pathway. This behavior is characteristic of a slowly folding species containing one or more *cis* prolyl peptide bonds, for which the rates of *cis*/*trans* isomerization and structure formation are weakly coupled (Kiefhaber et al., 1992; Schmid, 1992). This minor process has little influence on the main structural folding events and will not be discussed further.

Wild-type cyt *c* and L94I (Figure 5A,B) exhibit very similar kinetics. In the absence of denaturant, the overall fluorescence change is dominated by the burst phase (solid triangles). Up to about 1 M GuHCl, the fast phase is the dominant observable process. Toward the transition region, the intermediate phase gains amplitude and accounts for ~80% of the total fluorescence change at 2 M GuHCl. In this log rate vs [GuHCl] plot, the rate of the fast phase shows an approximately linear decrease up to ~2 M GuHCl, where its amplitude becomes negligible. The rate of the intermediate phase is nearly constant up to ~1 M GuHCl, where it begins to decrease, joining the rate of unfolding at the midpoint of the equilibrium transition. At this point, the kinetics are essentially two-state (except for the ~20% slow-folding species), indicating that structural intermediates are no longer populated at these denaturant concentrations. According to Elöve at al. (1994), the intermediate phase in the flat region below 1 M GuHCl is due to the dissociation of a nonnative histidine ligand that limits the formation of the native state from a partially folded intermediate, I_{NC}. The small denaturant dependence of this rate indicates that the ligand dissociation step involves little or no change in solvent exposure. The linear decrease in the intermediate-phase rate above 1 M GuHCl can be explained by the destabilization of the I_{NC} intermediate (cf. Khorasanizadeh et al., 1996).

The L94V and L94A mutations cause a 2−3-fold reduction in the rate of the fast phase (Figure 5C), which reflects the formation of I_{NC} (Elöve et al., 1994). This effect can be attributed to a decrease in the population of the burst-phase intermediate. In the absence of denaturant, the rate of the intermediate phase is quite similar for all variants (1−3 s^{-1}). This observation supports our previous assignment of this limiting rate to the dissociation of an alternative histidine ligand from I_{NC}, since this process is not expected to be affected by conservative amino acid changes at the interface between the N- and C-terminal helices. Compared to wild-type cyt *c* and the L94I and L94V variants, L94A exhibits a more linear rate profile for the intermediate phase with a significantly steeper slopes in the pretransition region. This effect appears to be related to the diminished population of I_{NC} in L94A (see below).

More direct insight into the population of intermediates and their response to mutations can be gained by considering the amplitudes of the various phases. In Figure 6, we compare the amplitudes of the three major phases for the different variants. The normalized amplitude of the burst phase in panel A illustrates the effects of the denaturant and mutations on the population of I_C, whereas the amplitude of the fast phase (panel B) is related to the transient accumulation of I_{NC}, which is most highly populated during folding of the wild-type protein at 0.7 M GuHCl. The population of I_{NC} decreases under destabilizing conditions, including both higher denaturant concentrations and progressively more perturbing amino acid changes at the helix interface. The amplitude of the intermediate phase (panel C) follows the opposite trend, i.e., it is more pronounced for the less stable variants at lower GuHCl concentrations.

DISCUSSION

Equilibrium Effects of Leu94 Mutations. Leu94 occupies a central position at the interface between the N- and

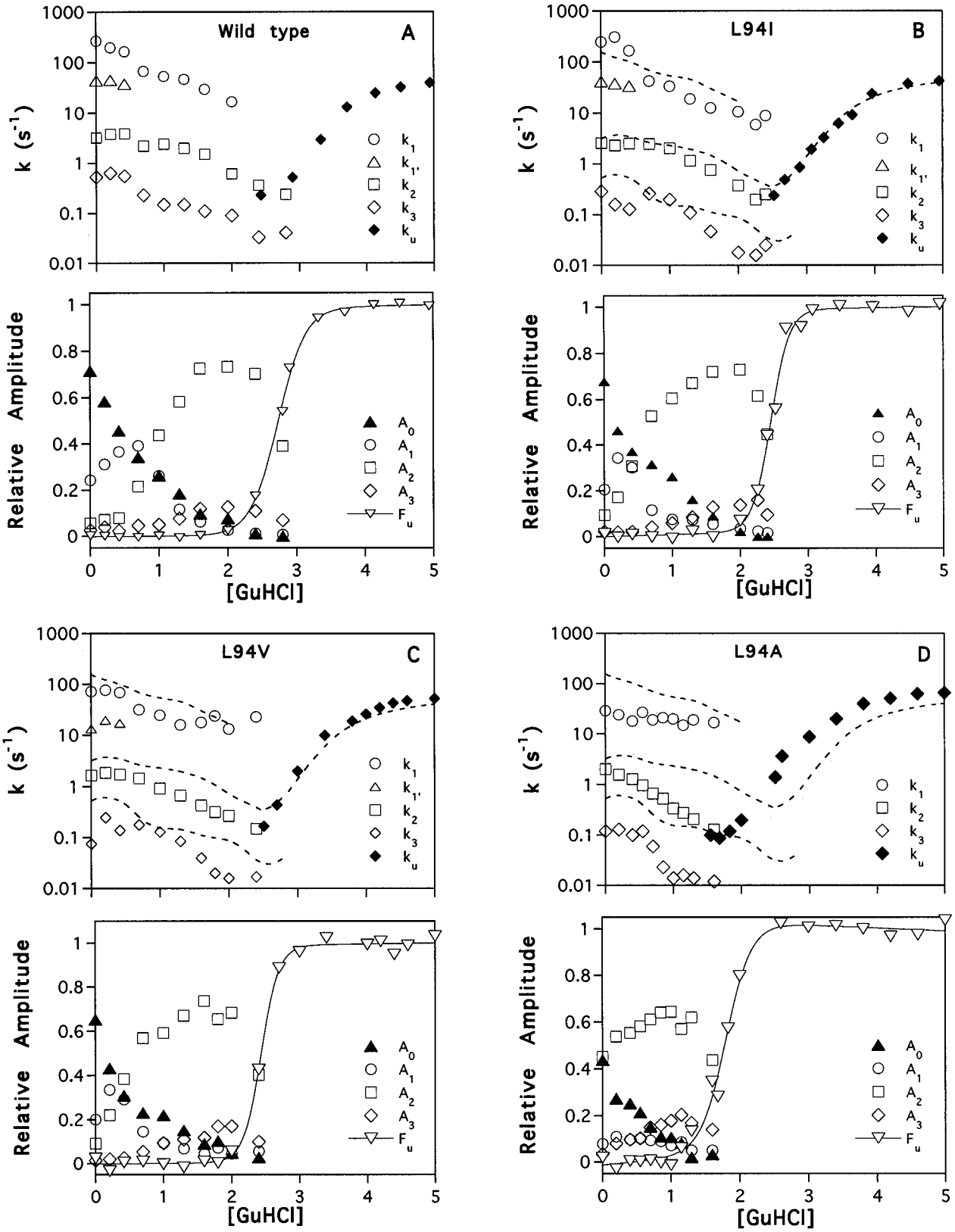

FIGURE 5: Denaturant dependence of the unfolding and refolding rates and amplitudes for wild-type cyt *c* (A) and the L94 variants (B−D). The upper panels in A−D show the GuHCl dependence of the rates for the unfolding phase (k_u) and the fast (k_1), intermediate (k_2), and slow refolding phase (k_3) of WT, L94I, L94V, and L94A, respectively. The lower panels in A−D show the GuHCl dependence of the relative amplitude for each refolding phase (A_0, A_1, A_2, and A_3) and the fraction of unfolded cyt *c* (f_u) obtained from the baseline points (final signal) of each kinetic experiment. A_0 represents the amplitude of the kinetically unresolved fast process (burst phase). The dashed lines represent the rate vs GuHCl behavior of the wild type (A) shown for comparison with each mutant (B−D).

C-terminal helices of cyt *c*, where it is in close contact with the backbone in the vicinity of Gly6 and with the side chains of Ile9 and Phe10 as well as the heme thiol bridge to Cys14 (Bushnell et al., 1990). This type of tight orthogonal packing of the two helices is a common helix−helix pairing motif found in other proteins, including the globins and lysozyme (Lesk & Chothia, 1980; Chothia et al., 1981). Leu94 and Gly6 are among the most highly conserved residues in over 100 known eukaryotic cytochrome *c* sequences (Mathews, 1985; Pielak et al., 1995). We found that cyt *c* is able to

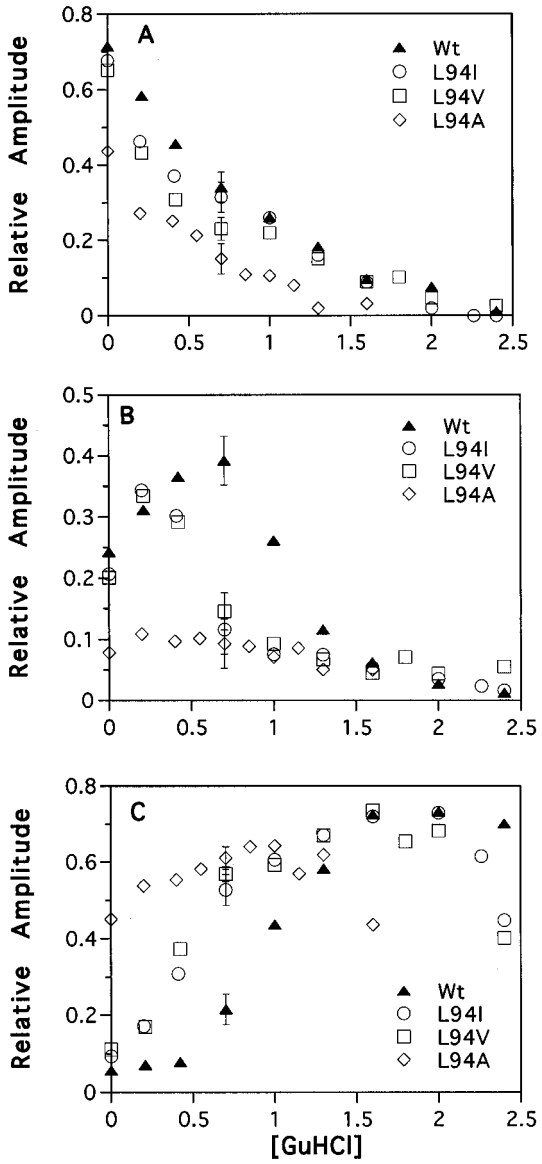

FIGURE 6: Comparison of the relative amplitudes of the burst phase (A), the fast phase (B), and the intermediate phase (C) for the various cyt *c* variants with aliphatic substitutions at position 94 (replotted from Figure 5). The amplitudes in panels A and B reflect the populations of the intermediates I_C and I_{NC}, respectively, and their dependence on denaturant concentration and mutations. Representative error bars are included with the data at 0.7 M GuHCl (cf. Table 2).

accommodate an Ile or Val residue in place of Leu at position 94 without major destabilization of the native structure (Figure 2 and Table 1). However, when Leu 94 is replaced by Ala, the stability of the native state decreases by 3.5 kcal/mol. The reduction in side chain volume apparently removes important interactions and leads to packing defects at the helix interface. Replacement of large hydrophobic side chains by Ala or Gly in the core of globular proteins is generally found to be highly destabilizing (Shortle et al., 1990; Hurley et al., 1991; Lim et al., 1992; Jackson et al., 1993) and has been attributed to the formation of cavities (Eriksson et al., 1992). Our results are consistent with the observations of Pielak and co-workers on a series of mutations at position 94 in yeast iso-1 cyt *c* (Fredericks & Pielak, 1993; Pielak et al., 1995). They found that while the helix−helix interface is flexible and able to accommodate many substitutions without loss of function, a bulky hydrophobic residue is necessary for stability. Our equilibrium

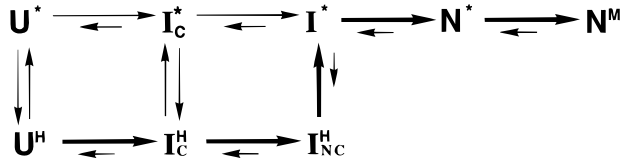

FIGURE 7: Kinetic model of cytochrome *c* folding consistent with the data obtained in this and previous studies (cf. Elöve et al., 1994). N^M represents the native state with intact Met80-iron bond. N* depicts a native-like five-coordinate state without the Met80 ligand. I^H_{NC} represents an intermediate with interacting N- and C-terminal helices and a trapped nonnative ligand (His26 or His33). I* represents a hypothetical partially folded intermediate with a five-coordinate heme. I^H_C and I^*_C denote compact intermediates with and without trapped nonnative ligand. U^H describes the unfolded state with a nonnative His ligand, and U* represents the unfolded five-coordinate state. All states contain His18 as the fifth ligand. The predominant folding pathway at pH 7 is shown in bold arrows.

data (Table 1) agree well with their calorimetric data on the corresponding Leu94 substitutions in yeast iso-1 cyt *c* (Pielak et al., 1995), which also show a slight increase in ΔG for the L94I mutation (by 0.3 kcal/mol) and marked decrease for the L94A mutation (by 3.6 kcal/mol).

Kinetic Mechanism. Our analysis of the folding kinetics of cyt *c* and its variants with aliphatic substitutions at position 94 suggests a rather complex folding mechanism involving at least two structurally distinct intermediates and multiple heme ligation states. Elöve et al. (1994) proposed a kinetic mechanism that takes into account the various ligation states involving alternative histidine and methionine residues bound at the sixth coordination site of the heme iron. This mechanism provides a straightforward explanation for the coupling between heme ligation and conformational steps in folding and unfolding. While this model can account for the kinetic behavior under a variety of conditions (Elöve et al., 1994), it does not predict the observation of a burst phase. Thus, we propose an extended model (Figure 7) that includes an early intermediate, I_C, in each of the parallel folding pathways originating from unfolded molecules with either one (His18) or two (His18 and His26 or His33) axial histidine ligands, designated U* and U^H, respectively. We have also added a native-like state, N*, between the native methionine-ligated state (N^M) and the partially folded five-coordinate state (I*) to account for our evidence that the initial deligation step in unfolding leads to a folded state with native-like structure lacking the Met80-iron linkage (in contrast to I*, which appears to be partially unfolded).

Under the folding conditions of this study, at neutral pH, U^H is the predominant unfolded state, and the protein follows the pathway indicated by bold arrows in Figure 7. The first event in folding at low denaturant concentration (less than ~1.5 M GuHCl) is the rapid formation of a compact intermediate, I^H_C, within the instrument dead time (~1 ms). Folding continues through a partially folded intermediate I^H_{NC} stabilized mainly by the mutually interacting N- and C-terminal helices (Roder et al., 1988). This intermediate still contains a nonnative histidine ligand, which prevents the rapid formation of stable structure in other regions of the protein (Elöve et al., 1994; Sosnick et al., 1994). Dissociation of this His ligand to form a five-coordinate intermediate with similar structural properties (I*) is the rate-limiting step in folding. The folding process is rapidly completed in a subsequent structural event (I* → N*), followed by coordination of the Met80 sulfur to the heme iron.

The mechanism shown in Figure 7 is supported by previous kinetic studies under conditions that prevent the formation of nonnative ligands. For example, at pH 5, where His26 and His33 are protonated and cannot serve as ligands, folding proceeds along a more direct pathway through a series of five-coordinate intermediates to the native state (Elöve et al., 1994). A variety of different conformational probes, including Trp59 fluorescence, heme absorbance in the visible and near-infrared regions, and protection of backbone amide protons, all exhibit a dominant refolding phase with a time constant of ~15 ms (in 0.7 M GuHCl at pH 5, 10 °C), indicative of a highly cooperative folding process (Elöve et al., 1994; Sosnick et al., 1994). This rate-limiting step corresponds to the $I^*_C \rightarrow I^*$ transition in Figure 7. In spite of the complexity of this mechanism, we believe that it contains the minimum number of states required to explain the bulk of the available experimental observations (excluding proline-dependent slow folding phases). Moreover, the model provides a rational explanation for the effects of mutations at the helix−helix interface on the kinetics of folding and unfolding.

The Kinetic Effect of Mutations on an Early Compact Intermediate. Previous kinetic CD and fluorescence experiments showed that a compact intermediate with substantial α-helix content accumulates during the first few milliseconds of the cyt *c* folding reaction at 0.7 M GuHCl (Elöve et al., 1992). On the other hand, earlier pulsed NH exchange studies showed less than 20% protection of amide protons during the 3 ms dead time of the quenched−flow experiment (Roder et al., 1988), suggesting that the helices detected by CD are unstable and lack persistent hydrogen bonds. Our present observation that the initial fluorescence signal decreases sharply toward lower denaturant concentrations (Figure 3A), approaching the level of the native protein, confirms the compact nature of this early intermediate (I_C). Since the equilibrium between U and I_C is fast compared to subsequent steps, the burst-phase amplitude is directly related to the population of I_C, and its denaturant dependence provides information on the stability of I_C. For wild-type cyt *c*, this pre-equilibrium transition has an apparent C_m of 0.4 M and an *m* value of 1 kcal mol^{-1} M^{-1}, suggesting a marginally stable ($\Delta G_u \sim 0.4$ kcal/mol) loosely packed state with a solvent-exposed surface area closer to that of the unfolded than the native conformation.

Our observation (Figure 6A) that substitution of the Leu94 side chain by smaller aliphatic groups affects the burst-phase amplitude, A_0, shows that hydrophobic interactions involving residue 94 contribute to the stability of I_C. The L94A variant exhibits a ~2-fold decrease in A_0 at all GuHCl concentrations studied (Figure 6A), indicating a substantial decrease in the population of I_C. The behavior of L94I and L94V is more similar to that of the wild type, but they also show a small, but significant, decrease in A_0 at intermediate denaturant concentrations (see Table 2 for typical errors). The observation that conservative aliphatic substitutions have a noticeable destabilizing effect on I_C suggests that not only the hydrophobicity of residue 94 but also more specific steric factors contribute to the stabilization of this early intermediate. This is in contrast to the formation of a compact state via random hydrophobic collapse without specific tertiary contacts, which would not be expected to be very sensitive to small changes in a single amino acid. It is interesting that the L94A mutation destabilizes I_C in spite of the fact that Ala favors helical secondary structure, while the helix-destabilizing Ile and Val substitutions are less perturbing. Thus, intrinsic helix propensity does not appear to be an important factor in stabilizing the burst-phase intermediate. In conjunction with previous stopped-flow CD, fluorescence, and pulse labeling H/D exchange experiments (Roder et al., 1988; Elöve et al., 1992), the present results suggest that, in addition to its compact dimensions and fluctuating helical secondary structure, the burst-phase intermediate also relies on some preferred tertiary interactions involving a central core residue that are inconsistent with a nonspecifically collapsed state.

With its native-like secondary structure and compact dimensions (Elöve et al., 1992), I_C exhibits some of the characteristics of a molten globule state [for a recent review, see Ptitsyn (1995)]. However, comparison with the equilibrium A-state of cyt *c* observed at low pH and high salt concentration (Ohgushi & Wada, 1983; Goto et al., 1990) reveals some important differences (Roder & Elöve, 1994). In contrast to I_C, the A-state exhibits well-protected amide protons in all three major α-helices and some nonhelical regions (Jeng et al., 1990). Moreover, the A-state exhibits a cooperative unfolding transition in thermal as well as chemical denaturation experiments (Kuroda et al., 1992; Hagihara et al., 1994), whereas the shallow denaturant dependence of A_0 (Figure 6) indicates a less cooperative unfolding transition in the case of I_C. In recent equilibrium studies on yeast iso-1 cytochrome *c*, Marmorino and Pielak (1995) found that mutation of contact residues on the N- and C-terminal helices lowers the stability of the A-state and the native state to a similar degree, indicating the presence of specific helix−helix contacts in both forms. For example, the A-state was no longer detectable in the L94A variant of yeast cyt *c*, while the same mutation in horse cyt *c* still allows partial formation of I_C at low denaturant concentrations (Figure 6). These observations suggest that the A-state of cyt *c* may be an example of a highly ordered molten-globule state (Dobson, 1994) stabilized by specific tertiary interactions. On the other hand, the early kinetic intermediate with its marginal stability, limited cooperativity and low sensitivity to conservative core mutations appears to be a more representative case of a kinetic molten globule intermediate (Ptitsyn, 1995; Roder, 1995). Uversky and Ptitsyn (1994) recently found that, in addition to the previously described molten globule state, β-lactamase exhibits a second less compact equilibrium intermediate, which appears to be a better analog of early kinetic intermediates than some of the more stable molten globules.

Effect of Mutations on the Stability of I_{NC}. The striking changes in the kinetics of folding caused by conservative amino acid substitutions at position 94 (Figure 4) clearly demonstrate the critical role of helix−helix interactions in the folding of cyt *c*. According to our kinetic mechanism (Figure 7), the observed amplitude gain of the intermediate phase at the expense of the fast phase is primarily due to the destabilizing effect of the mutations on the partially folded intermediate I_{NC}, resulting in a large decrease in its population. As in our kinetic analysis of core mutations in ubiquitin (Khorasanizadeh et al., 1996), the observed rate constants for folding, k_1 and k_2 (Table 2), are relatively insensitive to mutations, suggesting that the depletion of I_{NC} is primarily due to an increase in the unfolding rate constant ($I_{NC} \rightarrow I_C$). The observation that the stability of I_{NC} is highly sensitive to amino acid changes at position 94 represents the most direct demonstration so far that it contains tightly interacting N- and C-terminal helices, confirming our previ-

ous evidence, based on pulsed hydrogen exchange (Roder et al., 1988; Elöve & Roder, 1991) and model peptide (Wu et al., 1993) studies.

The fact that even the conservative Ile and Val substitutions at the helix−helix interface have a pronounced effect on the population of I_{NC} (Figures 4−6) indicates that the pairing of the N- and C-terminal helices in this partially folded intermediate is mediated by very specific packing interactions. The preference for a leucine at position 94 suggests that its tight contact with the N-terminal helix involving Gly6, Ile9, and Phe10 found in the native structure (Bushnell et al., 1990) is already established in I_{NC}. On the other hand, this specific contact does not appear to be formed in I_C, since the L94I and L94V substitutions affect its stability to a much smaller extent, as expected for a compact state with a loosely packed hydrophobic core. The lack of a significant effect of these conservative mutations on the stability of the native state (Table 1) suggests that other interactions within the native structure can compensate for the packing defects at the helix−helix interface.

Since the packing of the N- and C-terminal helices involves a close contact between two aromatic side chains, Phe10 and Tyr97 (Bushnell et al., 1990), one might expect the formation of I_{NC} to be accompanied by a change in the near-UV region of the CD spectrum. However, our previous time-resolved CD studies showed no detectable change in ellipticity at 289 nm prior to the 100 ms time range, in contrast to the helical signal at 222 nm (Elöve et al., 1992). This apparent contradiction may be explained by the fact that the near-UV CD signal of cyt *c* is dominated by Trp59 (Davies et al., 1993). The indole NH of Trp59 forms a hydrogen bond to one of the heme propionate side chains in the native state but is still accessible to the solvent in I_{NC}, as indicated by its lack of protection against H/D exchange (Roder et al., 1988).

The thermodynamic and kinetic effects of mutations at position 94 show no apparent correlation with the intrinsic propensity of this residue for helix formation. For example, although alanine is the most favorable residue for helix formation in model peptides (Scholtz & Baldwin, 1992), it destabilizes I_C, I_{NC}, and N more than any of the other substitutions investigated. This is not unexpected, since hydrophobic interactions and specific van der Waals contacts of this buried residue are expected to outweigh any contributions due to intrinsic helix propensities.

Unfolding Mechanism. The unusual denaturant dependence of the rate of unfolding observed for wild-type cyt *c* and all variants investigated here (Figure 5) is a clear indication of a complex unfolding reaction with at least one intermediate state preceding the main structural unfolding transition. According to the kinetic mechanism shown in Figure 7, the observation of a nearly denaturant-independent process can be explained by the formation of a native-like intermediate, N*, which limits the unfolding reactions at high GuHCl concentrations. Our model suggests that the main distinction of N* from the native state (N^M) is its lack of the bond between the Met80 sulfur and the iron. This interpretation is consistent with the fact that the Met80 ligand can be displaced by addition of extrinsic ligands, such as imidazole, cyanide, or azide, resulting in a folded state with native-like structural properties (Babul & Stellwagen, 1971; Roder et al., 1986). At intermediate denaturant concentrations, above C_m, the main structural unfolding step is slow compared to the rate of Met80 dissociation ($N^M \rightarrow N^*$) and determines the observed rate of unfolding. With increasing GuHCl concentration, this process accelerates and eventually exceeds the ligand dissociation rate, which becomes rate-limiting above 4 M GuHCl. The shallow denaturant dependence of this rate indicates that dissociation of the methionine ligand is accompanied by only small, localized structural perturbations.

The observation that the unfolding behavior is essentially unaffected by the L94I and L94V mutations (Figure 5) is consistent with our recent solution structure refinement of the L94V variant, which revealed only subtle structural rearrangements near the site of mutation (Gochin & Roder, 1995). In the case of L94A, the rate of the GuHCl-dependent structural transition increases nearly 10-fold, which is in part responsible for the large decrease in equilibrium stability (Table 1), whereas the rate-limiting process at high GuHCl concentrations ($N^M \rightarrow N^*$) increases by less than 2-fold. Thus, the L94A mutation affects the rate and equilibrium constant for global unfolding but has little effect on the local structural changes associated with the dissociation of the Met80 ligand, indicating that the two processes are only weakly coupled.

Our interpretation of the cyt *c* unfolding kinetics according to the scheme in Figure 7 implies that the most efficient unfolding pathway involves Met80 deligation followed by the main conformational transition rather than direct unfolding of the methionine-ligated state. The long-range cross link between Met80 and the heme apparently stabilizes the native structure to such an extent that it unfolds at a negligible rate, unless that bond is broken first. This interpretation is supported by our recent finding that addition of imidazole accelerates the rate of unfolding and causes the GuHCl-independent unfolding process to disappear (unpublished results).

Cyt c Folding Mechanism. Previous kinetic studies involving optical probes and amide protection revealed the existence of at least two intermediates in the folding of cyt *c* at neutral pH, an early compact intermediate, I_C (Elöve et al., 1992), and a subsequent partially folded state with interacting N- and C-terminal helices, I_{NC} (Elöve & Roder, 1991; Roder et al., 1988; Wu et al., 1993). The present results provide further insight into the nature of helix packing interactions and their kinetic importance. Our finding that the L94A mutation destabilizes both intermediates (Figure 5D), in addition to the native state (Figure 2), shows that the conserved residue Leu94 is critical for the formation of stable structure already at early stages of folding and remains so throughout the whole process. While the presence of a large hydrophobic side chain, such as Leu, Ile, or Val, is sufficient for the stabilization of I_C, the stability of I_{NC} is remarkably sensitive to small changes in side chain geometry, indicating that its formation relies on more specific packing interactions. Nevertheless, once the native state is reached, conservative amino acid changes at the helix−helix interface (L94I and L94V) have little effect on stability, suggesting that other native interactions that are not present yet in I_{NC} can compensate for minor packing defects. Pielak et al. (1995) recently reported a similar tolerance toward conservative amino acid changes for Leu94 and two other residues at the helix−helix interface (Phe10 and Tyr97) in yeast iso-1 cyt *c*. They further noted that, in spite of this apparent plasticity, leucine is by far the most common residue at position 94, found in 100 out of 106 known eukaryotic cyt *c* sequences; the remaining six sequences have either Ile or

Val at position 94. This observation suggests that the preference for Leu over closely related side chains depends on factors other than the thermodynamic stability of the native state. Functional requirements can also be ruled out in this case, since random mutagenesis studies showed that yeast iso-1 cyt *c* is able to accommodate a wide range of residues at position 94, including Ile and Val, without measurable effects on the function *in vivo* (Fredericks & Pielak, 1993). This leaves us with the intriguing possibility that the kinetics of folding and the stability of folding intermediates may have played some role in the evolution of the cytochromes. The selection of the wild-type sequence with Leu94 over other variants with less stable folding intermediates suggests a possible physiological role of the I_{NC} intermediate, perhaps in the localization of cyt *c* in the mitochondrial intermembrane space and/or in the covalent attachment of the heme.

On the basis of the observation that I_{NC} accumulates only under refolding conditions favoring the formation of nonnative heme ligands (pH >6; cf. Brems & Stellwagen, 1983; Elöve & Roder, 1991), Sosnick et al. (1994) suggested that it represents a structurally misfolded state. On the other hand, we argued that association of the two chain termini reflects a critical structural event in cyt *c* folding regardless of the heme ligation state (Elöve et al., 1994). Our view is supported by several lines of evidence indicating that I_{NC} is a partially folded state with primarily native-like structural elements. In particular, pulsed H/D exchange measurements revealed strongly protected amide protons only in the N- and C-terminal helices and no evidence for stable hydrogen-bonded structure (native-like or nonnative) elsewhere (Roder et al., 1988; Elöve & Roder, 1991). The presence of an independent structural subdomain was confirmed by our finding that the heme-containing N-terminal fragment of cyt *c* (residues 1−38) and a synthetic peptide corresponding to the C-terminal helix (residues 87−104) form a stable helical complex with native-like interactions (Wu et al., 1993). The mutational effects presented here show that the mutual recognition of the two helices is not just mediated by loose hydrophobic contacts but involves specific packing interactions similar to those found in the native structure. If pairing of the chain termini is an essential step in cyt *c* folding, it should also occur under conditions that disfavor the formation of nonnative histidine ligands, such as lowering pH or addition of extrinsic ligands. In terms of our kinetic scheme (Figure 7), this means that on the direct pathway from U*, an intermediate I* should be encountered that is structurally analogous to I^{H}_{NC}, but lacks the nonnative histidine−iron bond. If the subsequent folding steps (I* → N* and N* → N^{M}) are fast compared to formation of I*, this state would not accumulate to detectable levels. However, its existence may be deduced from the effect of mutations on the kinetics of folding at pH 5 or in the presence of imidazole (Colón et al., manuscript in preparation). Furthermore, this hypothesis is consistent with recent hydrogen exchange results on cyt *c* under native conditions as a function of GuHCl concentration, which show that partially unfolded forms with protected amide protons in the N- and C-terminal helices are accessible by equilibrium fluctuations of the native state (Bai et al., 1995). These short-lived species detected by NH exchange may correspond to I* in Figure 7, since the presence of a nonnative histidine ligand can be ruled out on the basis of its slow rate of dissociation (10−100 s^{-1}; cf. Elöve et al., 1994).

The L94I and L94V variants of cyt *c* characterized in this study are striking examples of "kinetic mutants" (Matthews, 1987), since they have a pronounced effect on the kinetics of folding (by selectively destabilizing a folding intermediate) without perturbing the thermodynamic stability. Few such cases have been described in the literature. One example is the V26I mutation in ubiquitin, which increases both the rates of folding and unfolding but has little effect on the unfolding equilibrium (Khorasanizadeh et al., 1996). Other examples have been reported for the folding and assembly of oligomeric proteins. Temperature-sensitive mutants of the tailspike protein P22 can show native-like stability once folded but form aggregates when cells are grown at nonpermissive temperature (Haase-Pettingel & King, 1988). These inclusion bodies have been suggested to be caused by the aggregation of folding intermediates, which has also been observed *in vitro* (Teschke & King, 1995). Similar effects were found for interleukin-1β (Chrunyk et al., 1993; Wetzel & Chrunyk, 1994).

Cyt *c* is not the only protein that exhibits a partially folded intermediate stabilized by tertiary helix−helix contacts. For example, hydrogen exchange labeling experiments on apomyoglobin (Jennings & Wright, 1993) revealed an intermediate containing stable hydrogen-bonded structure only in some of the helices (A, G, H, and parts of B). In the native structure of myoglobin, these form a cluster stabilized by three close-packed helix−helix contacts (Chothia et al., 1981). Another example is lysozyme, where stable hydrogen-bonded structure first appears in the predominantly α-helical domain of this two-domain structure (Radford et al., 1992). It is interesting that in each case a helix near the N-terminus forms a tight perpendicular contact (helix axes distances < 8.5 Å) with another helix near the C-terminus: the N/C helix pair in cyt *c*, the A/H helix pair in myoglobin, and the contact between the B-helix and the C-terminal 3^{10}-helix in the case of lysozyme (Chothia et al., 1981; Chothia & Lesk, 1985). While this may, of course, be a coincidence, these orthogonal helix−helix packing arrangements (class 1−4 or 3−3) may be intrinsically more stable than arrangements with more parallel helix axes (class 4−4 or 3−4). The docking of structural elements near opposite ends of the chain would be an especially efficient mechanism for lowering the conformational entropy at an early stage of folding (cf. Roder et al., 1988).

Summary and Conclusions. Replacement of Leu94 with Ile or Val resulted in only small changes in the GuHCl-induced unfolding equilibrium, indicating that conservative mutations can be accommodated at the helix−helix interface with only minor structural perturbations. However, the L94A mutation resulted in a nearly 3.5 kcal/mol decrease in the unfolding free energy of the native state, indicating that the reduction in side chain volume causes severe packing defects. These equilibrium results show that hydrophobic interactions between the N- and C-terminal helices of cyt *c* are a major factor in stabilizing the native cyt *c* structure. A similar pattern is observed for the burst-phase intermediate I_C, which is favored in the presence of the larger aliphatic side chains but destabilized by the Ala substitution. In contrast, the effects of mutations at position 94 on the kinetics of folding and unfolding monitored by Trp59 fluorescence show a distinctly different trend. In spite of their small effect on overall stability, the L94I and L94V variants exhibited a substantial reduction in the relative amplitude of the fast phase, consistent with a significantly reduced population of

the I_{NC} intermediate compared to the wild-type protein. This trend continues in the case of the less stable L94A variant, which exhibits slower folding kinetics and even lower population of I_{NC}. The observation that a large hydrophobic side chain at this helix−helix contact site is sufficient for stabilizing I_C and N, while even the most conservative changes in side chain geometry lead to major destabilization of I_{NC} indicates that specific packing interactions are responsible for the mutual structural recognition of the two α-helices at intermediate stages of cyt *c* folding. The fact that a leucine is found at this position in 94% of all known eukaryotic cyt *c* sequences suggests a possible functional importance of stabilizing this folding intermediate with tightly packed N- and C-terminal helices.

ACKNOWLEDGMENT

We thank L. Comfort for synthesizing the oligonucleotides and for preparing cyt *c* samples. We are grateful to J. Glusker, E. K. Jaffe, S. D. Luck, and J. M. Sauder for valuable advice and discussions.

REFERENCES

Babul, J., & Stellwagen, E. (1971) *Biopolymers 10*, 2359−2361.
Babul, J., & Stellwagen, E. (1972) *Biochemistry 11*, 1195−1200.
Bai, Y., Sosnick, T. R., Mayne, L., & Englander, S. W. (1995) *Science 269*, 192−197.
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.
Chothia, C., & Lesk, A. M. (1985) *J. Mol. Biol. 182*, 151−158.
Chothia, C., Levitt, M., & Richardson, D. (1981) *J. Mol. Biol. 145*, 215−250.
Chrunyk, B. A., Evans, J., Lillquist, J., Young, P., & Wetzel, R. (1993) *J. Biol. Chem. 268*, 18053−18061.
Davies, A. M., Guillemette, J. G., Smith, M., Greenwood, C., Thurgood, A. G. P., Mauk, A. G., & Moore, G. R. (1993) *Biochemistry 32*, 5431−5435.
Dobson, C. M. (1994) *Curr. Biol. 4*, 636−640.
Elöve, G. A., & Roder, H. (1991) *ACS Symp. Ser. 470*, 50−63.
Elöve, G. A., Bhuyan, A. K., & Roder, H. (1994) *Biochemistry 33*, 6925−6935.
Elöve, G. A., Chaffotte, A. F., Roder, H., & Goldberg, M. E. (1992) *Biochemistry 31*, 6876−6883.
Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P., & Matthews, B. W. (1992) *Science 255*, 178−183.
Fredericks, Z. L., & Pielak, G. J. (1993) *Biochemistry 32*, 929−936.
Gochin, M., & Roder, H. (1995) *Protein Sci. 4*, 296−305.
Goto, Y., Calciano, L. J., & Fink, A. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 573−577.
Goto, Y., Hagihara, Y., Hamada, D., Hoshino, M., & Nishii, I. (1993) *Biochemistry 32*, 11878−11885.
Haase-Pettingel, C., & King, J. (1988) *J. Biol. Chem. 263*, 4977−4983.
Hagihara, Y., Tan, Y., & Goto, Y. (1994) *J. Mol. Biol. 237*, 336−348.
Hickey, D. R., Jayaraman, K., Goodhue, C. T., Shah, J., Fingar, S. A., Clements, J. M., Hosokawa, Y., Tsunasawa, S., & Sherman, F. (1991) *Gene 105*, 73−81.
Hurley, J. H., Baase, W. A., & Mathews, B. W. (1991) *J. Mol. Biol. 224*, 1143−1159.
Jackson, S. E., Moracci, M., elMasry, N., Johnson, C. M., & Fersht, A. R. (1993) *Biochemistry 32*, 11259−11269.
Jeng, M.-F., Englander, S. W., Elöve, G. A., Wand, A. J., & Roder, H. (1990) *Biochemistry 29*, 10433−10437.
Jennings, P. A., & Wright, P. E. (1993) *Science 262*, 892−895.
Kataoka, M., Hagihara, Y., Mihara, K., & Goto. (1993) *J. Mol. Biol. 229*, 591−596.
Khorasanizadeh, S., Peters, I. D., Butt, T. R., & Roder, H. (1993) *Biochemistry 32*, 7054−7063.
Khorasanizadeh, S., Peters, I. D., & Roder, H. (1996) *Nature Struct. Biol. 3*, 193−205.
Kiefhaber, T., Kohler, H., & Schmid, F. X. (1992) *J. Mol. Biol. 224*, 217−229.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−383.
Kuroda, Y., Kidokoro, S., & Wada, A. (1992) *J. Mol. Biol. 223*, 1139−1153.
Lesk, A. M., & Chothia, C. (1980) *J. Mol. Biol. 136*, 225−270.
Lim, W. A., Farruggio, D. C., & Sauer, R. T. (1992) *Biochemistry 31*, 4324−4333.
Marmorino, J. L., & Pielak, G. J. (1995) *Biochemistry 34*, 3140−3143.
Mathews, F. S. (1985) *Prog. Biophys. Mol. Biol. 45*, 1−56.
Matouschek, A., Serrano, L., & Fersht, A. R. (1994) in *Mechanisms of Protein Folding: Frontiers in Molecular Biology* (Pain, R. H., Ed.) pp 137−159, Oxford University Press, Oxford.
Matthews, C. R. (1987) *Methods Enzymol. 154*, 498−511.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
Meyer, T. E., & Kamen, M. D. (1982) *Adv. Protein Chem. 35*, 105−212.
Moore, G. R., & Pettigrew, G. W. (1990) *Cytochrome c: Evolutionary, Structural and Physiochemical Aspects*, Springer-Verlag, Berlin.
Muthukrishnan, K., & Nall, B. T. (1991) *Biochemistry 30*, 4706−4710.
Ohgushi, M., & Wada, A. (1983) *FEBS Lett. 164*, 21−24.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
Pielak, G. J., Auld, D. S., Beasley, J. R., Betz, S. F., Cohen, D. S., Doyle, D. F., Finger, S. A., Fredericks, Z. L., Hilgen-Willis, S., Saunders, A. J., & Trojak, S. K. (1995) *Biochemistry 34*, 3268−3276.
Ptitsyn, O. B. (1992) *Curr. Opin. Struct. Biol. 2*, 13−20.
Ptitsyn, O. B. (1995) *Adv. Protein Chem. 47*, 83−229.
Radford, S. E., Dobson, C. M., & Evans, P. A. (1992) *Nature 358*, 302−307.
Ridge, J. A., Baldwin, R. L., & Labhardt, A. M. (1981) *Biochemistry 20*, 1622−1630.
Roder, H. (1995) *Nature Struct. Biol. 2*, 817−820.
Roder, H., & Elöve, G. A. (1994) in *Mechanisms of Protein Folding: Frontiers in Molecular Biology* (Pain, R. H., Ed.) pp 26−55, Oxford University Press, New York.
Roder, H., Wand, A. J., Milne, J. S., & Englander, S. W. (1986) *Biophys. J. 49*, 57a.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700−704.
Schmid, F. X. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 197−241, W. H. Freeman & Co., New York.
Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95−118.
Sherman, F., Stewart, J. W., Parker, J. H., Inhaber, E., Shipman, N. A., Putterman, G. J., Gardinsky, R. L., & Margoliash, E. (1968) *J. Biol. Chem. 243*, 5446−5456.
Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033−8041.
Santoro, M. M., & Bolen, D. W. (1992) *Biochemistry 31*, 4901−4907.
Sosnick, T. R., Mayne, L., Hiller, R., & Englander, S. W. (1994) *Nature Struct. Biol. 1*, 149−156.
Stellwagen, E., & Babul, J. (1975) *Biochemistry 14*, 5135−5140.
Takano, T., & Dickerson, R. E. (1981a) *J. Mol. Biol. 153*, 95−115.
Takano, T., & Dickerson, R. E. (1981b) *J. Mol. Biol. 153*, 79−94.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.
Teschke, C. M., & King, J. (1995) *Biochemistry 34*, 6815−6826.
Tsong, T. Y. (1976) *Biochemistry 15*, 5467−5473.
Uversky, V. N., & Ptitsyn, O. B. (1994) *Biochemistry 33*, 2782−2791.
Wetzel, R., & Chrunyk, B. A. (1994) *FEBS Lett. 350*, 245−248.
Wood, L. C., White, T. B., & Nall, B. T. (1988) *Biochemistry 27*, 8562−8568.
Wu, L., Laub, P. B., Elöve, G. A., Carey, J., & Roder, H. (1993) *Biochemistry 32*, 10271−10276.

BI960052U