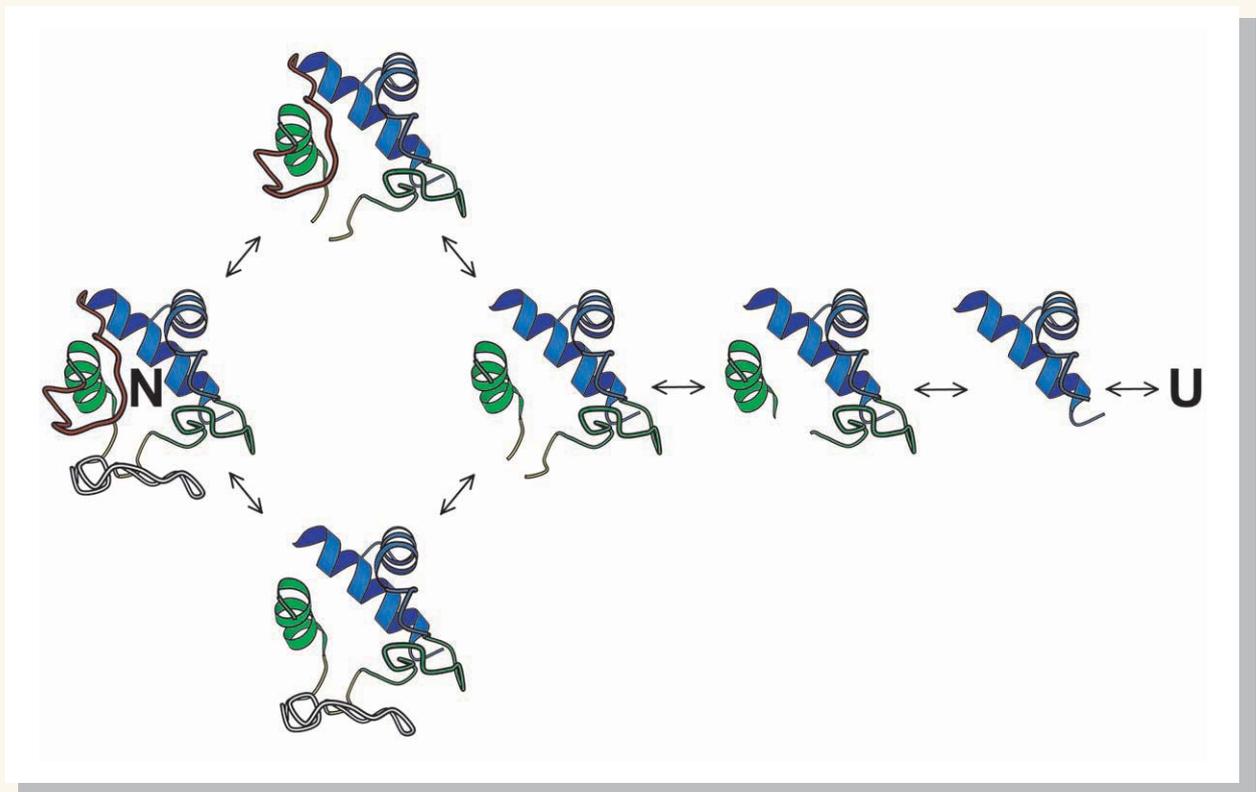# JMB

## JOURNAL OF MOLECULAR BIOLOGY

## JMB

# Order of Steps in the Cytochrome *c* Folding Pathway: Evidence for a Sequential Stabilization Mechanism

## Mallela M. G. Krishna*, Haripada Maity, Jon N. Rumbley Yan Lin and S. Walter Englander

*Johnson Research Foundation Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia PA 19104-6059, USA*

Previous work used hydrogen exchange (HX) experiments in kinetic and equilibrium modes to study the reversible unfolding and refolding of cytochrome *c* (Cyt *c*) under native conditions. Accumulated results now show that Cyt *c* is composed of five individually cooperative folding units, called foldons, which unfold and refold as concerted units in a stepwise pathway sequence. The first three steps of the folding pathway are linear and sequential. The ordering of the last two steps has been unclear because the fast HX of the amino acid residues in these foldons has made measurement difficult. New HX experiments done under slower exchange conditions show that the final two foldons do not unfold and refold in an obligatory sequence. They unfold separately and neither unfolding obligately contains the other, as indicated by their similar unfolding surface exposure and the specific effects of destabilizing and stabilizing mutations, pH change, and oxidation state. These results taken together support a sequential stabilization mechanism in which folding occurs in the native context with prior native-like structure serving to template the stepwise formation of subsequent native-like foldon units. Where the native structure of Cyt *c* requires sequential folding, in the first three steps, this is found. Where structural determination is ambiguous, in the final two steps, alternative parallel folding is found.

© 2006 Elsevier Ltd. All rights reserved.

*Keywords:* hydrogen exchange; folding pathway; mutation; *m* value; stability labeling

*Corresponding author

Present addresses: J.N. Rumbley, Department of Chemistry, University of Minnesota, 313 Chemistry, 1039 University Dr., Duluth, MN 55812, USA; Y. Lin, Ennova MedChem Group Inc., 675 US Highway 1, North Brunswick, NJ 08902, USA.

Abbreviations used: Cyt *c*, cytochrome *c*; WT, wild-type Cyt *c*; pWT, pseudo wild-type recombinant equine Cyt *c* (H26N, H33N); HX, hydrogen exchange; NHX, native-state HX; EX1, monomolecular exchange where the HX rate is equal to the determining structural opening rate; EX2, bimolecular exchange where the HX rate is proportional to the concentration of catalyst and can yield the equilibrium constant and free energy of the determining structural unfolding; foldon, cooperative folding/unfolding unit; PUF, partially unfolded form; GdmCl, guanidinium chloride; GdmSCN, guanidinium thiocyanate; $\Delta G_{HX}$, unfolding free energy measured by NHX; *m* value, d($\Delta G_{HX}$)/d[GdmCl] given by the slope of the NHX data plot; pDr, pH of $^2H_2O$ solution read by glass electrode.

E-mail address of the corresponding author: kmallela@mail.med.upenn.edu

## Introduction

To understand any biochemical pathway, one must identify and order the important pathway intermediates. The same is true for protein folding pathways. Work in recent years has traced the folding pathways of several proteins in some structural detail (cytochrome *c*,[1–12] ribonuclease H,[13–15] apocytochrome *b*562,[16–20] OspA of Borrelia,[21,22] and triosephosphate isomerase[23]). These proteins have been found to fold through distinct intermediates in distinct stepwise pathways. A major finding is that the pathway steps are determined by the formation of cooperative native-like structural units, called foldons. The order of pathway steps appears to be determined by a sequential stabilization process in which earlier native-like structure serves to guide and stabilize the formation of subsequent native-like foldons in a sequence that progressively constructs the native protein.

The identification of protein folding intermediates has been difficult because they generally live for less than one second and cannot be isolated for study by the usual structural methods. An alternative approach is possible. Although protein molecules under native conditions predominantly occupy their lowest free energy native state,[24] it is a thermodynamic truism that proteins must populate each higher energy state according to its Boltzmann factor and over time must kinetically cycle through all of these states. Thus, protein molecules must unfold and refold repeatedly, even under native conditions. Accordingly, folding intermediates and pathways might be studied directly over long periods of time under native conditions. Unfortunately, the infinitesimally populated high-energy states that define folding pathways under native conditions are normally invisible, because most experimental measurements are dominated by signals from the overwhelmingly populated native state. The opposite is true for hydrogen exchange (HX) experiments. HX rates that one measures receive no contribution from sites that are protected by H-bonding in the native state but are wholly determined by the cycling of molecules through their H-bond-broken higher energy states. Among the high-energy states are the partially unfolded forms (PUFs) that construct protein folding pathways. This is so whether one views the high-energy manifold in terms of a classical pathway with distinct intermediates,[25–28] or as a funnel-shaped landscape.[29,30]

In favorable cases native state HX (NHX) experiments can determine the structure and properties of the dominant high-energy forms and so can help to define potential folding intermediates and pathways. The experiment exploits conditions that selectively promote large unfolding reactions so that measured HX becomes dominated by them.[2,31,32] Cytochrome *c* (Cyt *c*) has proven to be a particularly favorable model due to its high global stability and the unusually separable nature of the cooperative folding/unfolding units (foldons) that form it, shown in Figure 1(a). Figure 1(b) summarizes the NHX results that originally defined the Cyt *c* foldon units. Some hydrogen atoms, called marker protons, are very sensitive to denaturant. Their exchange is governed by sizeable unfolding reactions that expose considerable new surface to solvent, encoded in the slope, $m = d\Delta G_{HX}/d[Den]$.[33] The exchange of many other hydrogen atoms is independent of denaturant. They exchange by way of small local fluctuation reactions that expose little new surface to the solvent ($m \approx 0$).[2,34,35] When any given unfolding reaction is sufficiently promoted so that it provides a faster HX pathway than the local fluctuations, it comes to dominate the exchange of the hydrogen atoms that it exposes. This is seen as a merging of multiple flat HX curves for different amides into a common HX isotherm. Figure 1(b) shows only one marker proton and one merging proton for each HX isotherm.

More complete data[2,4,10] show that each HX isotherm merges one or two sets of residues that are adjacent in the Cyt *c* sequence. These results identify the five cooperative unfolding/refolding units (foldons) in Figure 1(a), evaluate the free energy level of each of the partially unfolded forms (PUFs) that they produce, and provide a measure of the size of each unfolding (*m* value). The properties listed in Table 1 reveal a ladder of PUFs defined by a concomitant increase in $\Delta G$ and *m*. This ladder was initially taken to suggest a sequential pathway in which the foldons unfold and refold one discrete step at a time. Related results have been obtained in HX pulse labeling experiments done during kinetic folding,[1,7] in NHX experiments done in equilibrium (EX2) and kinetic (EX1) modes as Cyt *c* unfolds and
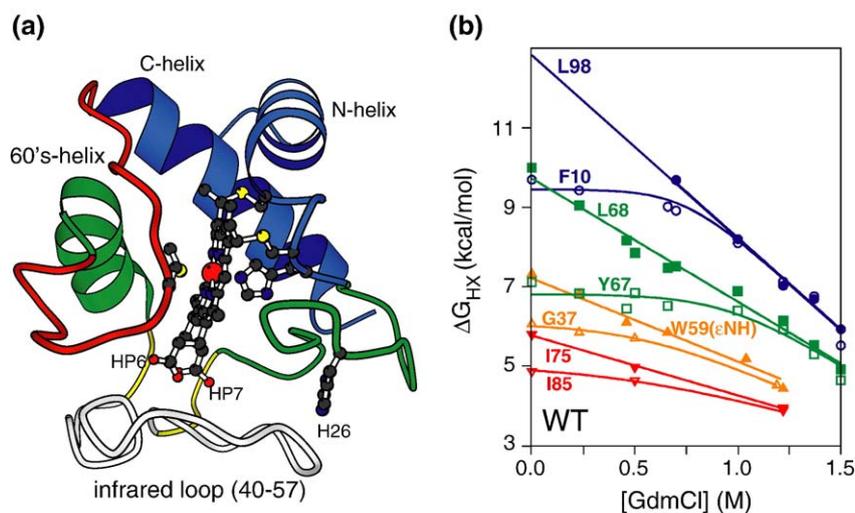


**Figure 1.** (a) Cyt *c* structure (1HRC.pdb[39] and MOLSCRIPT[58]), color-coded to indicate the foldon units previously identified by HX experiments and ranked in spectral order of decreasing $\Delta G_{HX}$ and *m* value[2–7,9,31], as specified in Table 1. (b) Illustration of NHX experiments on oxidized WT equine Cyt *c*.[2–4] Native Cyt *c* is placed into $^2H_2O$ and the H/D exchange of many amides is measured by recording 2D NMR spectra in time. The experiment is repeated with increasing concentrations of guanidinium chloride (GdmCl) but still far below the melting transition ($C_m = 2.75$ M GdmCl). The free energy of the high-energy state that determines the HX of each amide is calculated and plotted using the standard EX2 formalism[59] and the known unprotected chemical exchange rate.[56,57] For each PUF, only one marker proton and one merging proton are shown. Complete data sets for oxidized[2,4] and reduced[4] Cyt *c* and several mutants[10] are shown elsewhere.

**Table 1.** Properties of five foldons in oxidized Cyt *c*

| Foldon | Identity | $\Delta G_{HX}$ (kcal/mol) | $m$ (kcal/mol/M [GdmCl]) |
|---|---|---|---|
| Blue bihelix | N and C-terminal α-helices | 12.8 [a] | 4.7 [a] |
| Green helix and Green loop | 60's α-helix and 19–36 Ω-loop | 10.0 [a] | 3.2 [a] |
| Yellow β-sheet | 37–39, 58–61 short anti-parallel β-strands | 7.4 [b] | 2.1 [b] |
| Red loop | 71–85 Ω-loop | 6.3 [b] | 1.5 [b] |
| Infrared loop | 40–57 Ω-loop | 3.8 [b] | 1.5 [b] |

[a] From Bai *et al.*[2]
[b] At pDr 5.5, 20 °C from Figure 4.

refolds spontaneously under native conditions,[2,4,5] and in stability labeling experiments in which the stability of a given foldon is altered and the effect on the other foldons is measured.[3,9,10] All of these results have consistently found the same cooperative foldon units and ordered them in the same sequential unfolding/refolding pathway. In the folding direction, the N-terminal and C-terminal helices (blue foldon) form in an initial step,[1,7,36,37] the 60s helix and Ω-loop (green foldon) fold next, followed by the small yellow unit, and finally the two large Ω-loops.[2,4–6]

A compelling explanation for this pathway order is suggested by the arrangement of the foldons in the native Cyt *c* structure.[38] In the native context, the first folding blue bihelical unit contacts only the two green segments, therefore can only guide and stabilize their formation. The two short yellow strands grow from the ends of the two green units and, therefore, can only follow the two green units. This entire scaffold must be in place in order to template and stabilize the formation of the final two loops. However, the sequential stabilization mechanism seems to allow the assembly of these two loops in either order.

Here, we describe further HX experiments that clarify the position of these two loops in the Cyt *c* folding sequence. The results indicate that they can unfold and refold in arbitrary order, just as suggested by the sequential stabilization mechanism.

## Results

### The nested-yellow (infrared) loop is a separate foldon

In earlier work, the grouped HX behavior of three slow marker protons protected by the yellow segments (Lys60, Leu64, Trp59 εNH) was taken to mark the concerted unfolding of the yellow neck region together with its entire nested Ω-loop.[2] Hydrogen atoms in the nested loop segment exchanged too fast to measure. Subsequent results obtained under slower HX conditions indicated that the nested loop can behave as a separate foldon.[6]

Figure 2 shows more complete HX results as a function of destabilizing pH or denaturant, at lower pH where HX is slowed, in both reduced and oxidized Cyt *c*. Nine measurable hydrogen atoms in the nested loop exchange initially at quite different rates, evidently by way of localized structural fluctuations, except for the slowest protons, which appear to act as foldon markers. As pDr is lowered or denaturant is increased, a large concerted unfolding is promoted and comes to dominate the exchange of all of the nested loop hydrogen atoms, which merge into a common HX isotherm. Figure 3 shows that protons in other parts of the protein, including the yellow and the red foldon markers, are not entrained by the unfolding of the nested loop. These results identify the nested loop as a separate foldon, different from the yellow foldon, and show that it can unfold independently and at lower $\Delta G_{HX}$ than all of the other foldons. We designate this loop as infrared (previously called nested yellow) to imply that its unfolded state is at lower free energy than even the red unfolded state.

The destabilization of the infrared loop by decreasing pH is due to two structurally protected groups that have decreased $pK_a$ values in the native protein; namely, a heme propionate and His26.[6] Both are buried in their unprotonated form and interact directly with groups in the infrared loop.[39] No other candidate group in Cyt *c* titrates in the measured pH range. Sizeable general charge effects seem unlikely, since these experiments were done in 0.5 M KCl, no salt dependence was seen,[6] and hydrogen atoms that exchange by way of local fluctuations show no pH effect, even though HX is catalyzed by a hydroxyl ion, the local effective concentration of which would respond to increasing electrostatic potential.

### Relationship of the red and infrared foldons: *m* values

Prior results show that the red, yellow, green, and blue foldon units unfold in a sequential manner, up the ladder of increasing $\Delta G_{HX}$ and increasing *m* values (Figure 1; Table 1).[10] Each higher unfolded state obligately includes the lower ones to produce PUFs with identity as follows: red unfolded, red+yellow unfolded, red+yellow+ green unfolded, and red+yellow+green+blue unfolded signifying the U state. The unfolding of the infrared loop has even lower $\Delta G_{HX}$, suggesting that
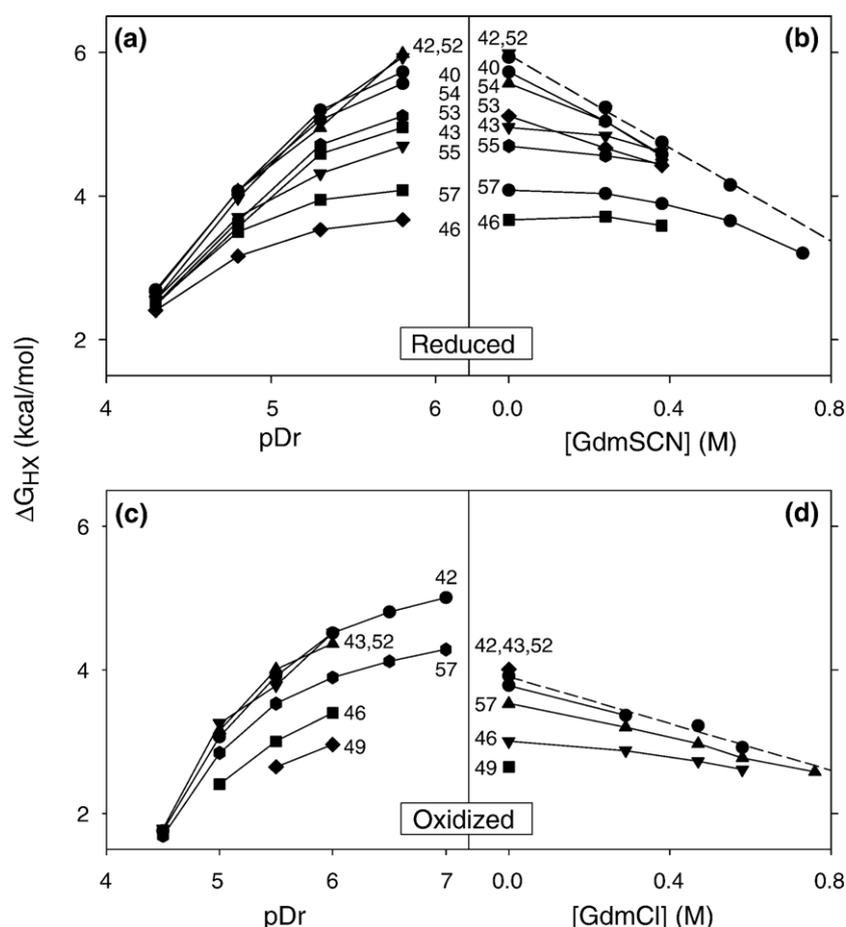
**Figure 2.** NHX results for the infrared (nested yellow) loop residues in reduced (pDr 5.8) and oxidized (pDr 5.5) WT Cyt *c* (20 °C; 0. 5 M KCl), showing the dependence of $\Delta G_{HX}$ of the measurable amide exposure reactions on solution pH and denaturant. Broken lines show the straight line fit to the marker protons. The markers include residues 42 and 52 in reduced Cyt *c*, and 43 in oxidized Cyt *c*, where the loop is further destabilized. Residues 40, 53, and 54 can also be expected to appear as markers in oxidized Cyt *c* but their crosspeaks are not seen in the oxidized NMR spectra. Errors in $\Delta G_{HX}$ are generally less than 0.1 kcal/mol.

it might unfold first and that the subsequent red loop unfolding includes it. It now appears that this is not the case.

Figure 4 compares the denaturant dependence (*m* values) of marker protons in the lower lying Cyt *c* foldons. It is clear that the red and infrared loops do not unfold together in a single concerted step, since they have very different HX rates and $\Delta G_{HX}$ for unfolding. Further, the *m* values measured for their unfolding are both 1.5 kcal/mol/M [GdmCl] (Table 1). This result is against the view that the unfolding of either loop obligately includes the other. If they unfolded in a prescribed order, one would have a larger *m* value than the other, as for the other units (Table 1).

(Distinct PUFs having similar *m* values were observed before in RNase H,[13] which may similarly indicate that their unfolding/refolding is not obligately sequential in either order.)

**Relationship of the red and infrared foldons: stability labeling**

In further tests we constructed several stability labeling experiments in which the stability of one of the foldon units was selectively perturbed and the effect on it and the other foldons was measured by NHX experiments. The results are illustrated by Figure 5.

Figure 5(a) shows the results for the non-perturbing surface mutation Glu62Gly, placed in the yellow neck (see Figure 1(a)).[9] Glycine exerts a destabilizing effect by entropically favoring the unfolded state. The mutation directly destabilizes the yellow foldon and higher foldons about equally (~0.9 kcal/mol) but the lower red and infrared foldons are unaffected. This is the expected result for a linear unfolding ladder in which each higher, partially unfolded state includes all of the lower unfoldings.

Figure 5(b) compares results for oxidized and reduced Cyt *c*. In reduced Cyt *c*, the heme iron to Met80-S ligation is strengthened by 3.2 kcal/mol, as measured by direct binding experiments with a model heme peptide.[3] The red loop is stabilized by the same 3.2 kcal/mol, evidently because its unfolding must break the Fe to S ligation. The higher yellow and green units are equally stabilized, as expected if their unfolding includes the red unfolding. The global blue unfolding has an additional stabilization because the heme iron reduction has an additional stabilizing effect, not limited to the Fe to S bond, as described below. The important point for the present discussion is that the infrared loop is stabilized by much less than 3.2 kcal/mol.[6] This result confirms that the unfolding of the infrared loop does not obligately include the red loop unfolding. The same conclusion is
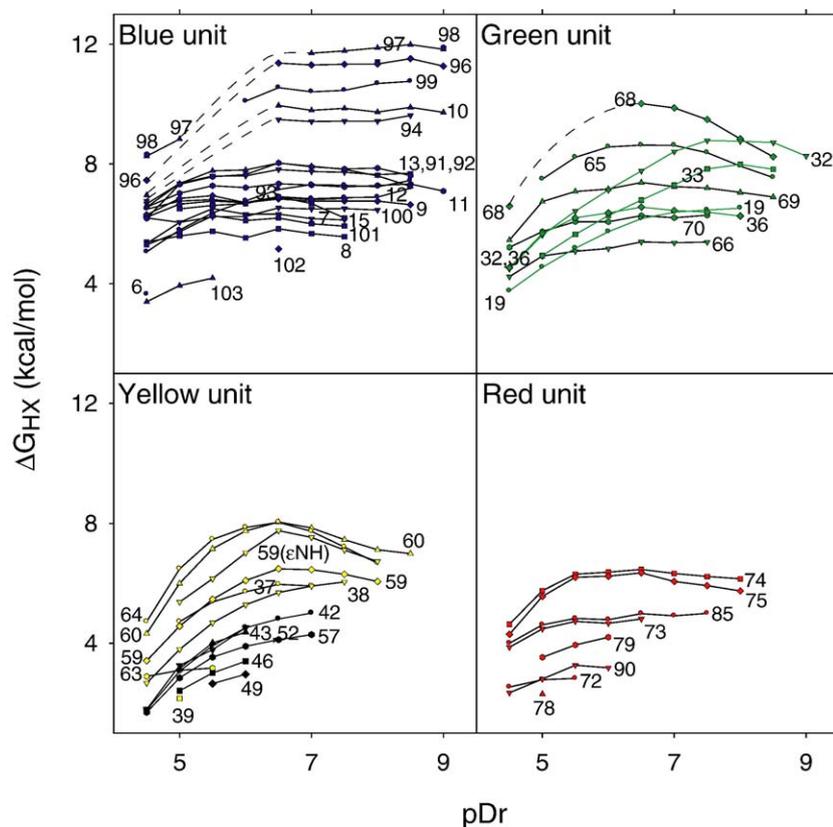
**Figure 3.** NHX results for all of the measurable amides in oxidized WT Cyt *c* as a function of solution pH (20 °C; 0.5 M KCl). The measured amide protons are grouped and color-coded according to the foldons in Figure 1(a) (infrared loop in black). NB: Previous denaturant-dependent NHX experiments suggested that the green helix and the green loop unfold and refold together. However, the upper right panel now indicates that these two units can unfold separately (see the text).

obvious from Figure 4, which shows that the infrared loop has much larger $K_{op}$ (lower $\Delta G_{HX}$).

An analogous situation is seen when the infrared loop is destabilized (Figure 5(c)). The Thr47Gly mutation placed in the infrared loop directly
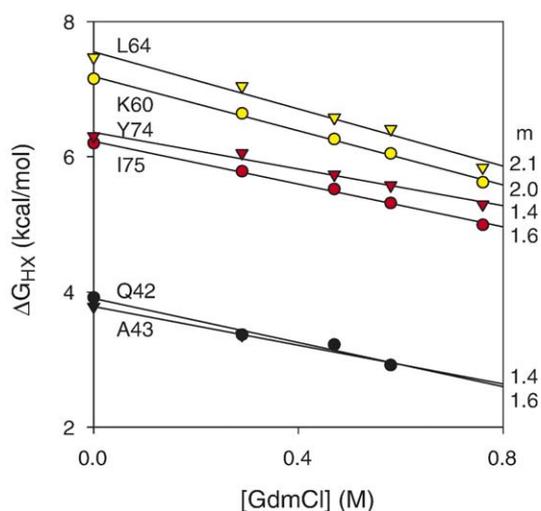


**Figure 4.** NHX results for some marker protons in oxidized WT Cyt c (pDr 5.5, 20 °C, 0.5 M KCl). Infrared (black) and red foldons (red) have the same denaturant dependence (slope = *m* value), suggesting that neither of these two loops is obligatorily included in the unfolding of the other. Errors in $\Delta G_{HX}$ and *m* values are generally less than 0.1 kcal/mol.

destabilizes the loop itself by 0.8 kcal/mol. The higher foldons are destabilized equally, as expected if their unfolding necessarily includes the infrared unfolding. However, the red unfolded state is promoted by significantly less, indicating that it does not obligately include the infrared unfolding.

The same effect is observed in a comparison of foldon destabilization at decreasing pH (Figure 5(d)), taken from the $\Delta G_{HX}$ results for foldon marker protons in Figure 3. At pDr 4.5 the infrared unit is destabilized by 3.3 kcal/mol (relative to pDr 7.0) because it is supported by the buried protonatable groups noted before (His26 and heme propionate). Marker protons for the higher yellow and green unfoldings are destabilized equally. This is the expected result if the infrared loop is on the unfolding pathway so that the higher unfoldings contain it. The red loop also appears to be destabilized by low pH but by much less. In fact, the red foldon destabilization may be even less. The apparent destabilization observed may simply reflect the fact that its free energy level is ultimately limited by the higher yellow foldon, insofar as the yellow unfolded state obligately includes the red unfolding. In either case, this result (Figure 5(d)), as for the Thr47Gly result just described (Figure 5(c)), is against the view that the red loop unfolded state obligately includes the unfolded infrared loop. (Here also, the global unfolding marker feels an additional effect, which is considered below.)

In summary, the results described here show that the two lowest free energy foldons in the
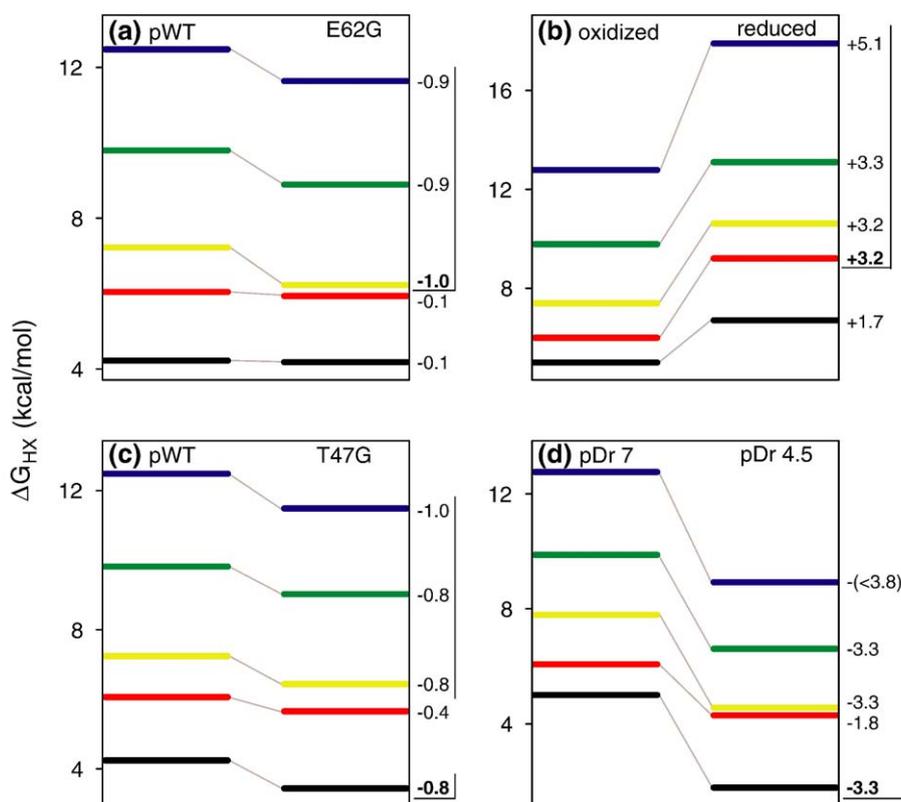
**Figure 5.** Stability labeling results testing the sequential nature of the individual foldons in Cyt *c* unfolding. NHX experiments were done at pDr 7, 20 °C unless noted otherwise. The effect of each stabilizing or destabilizing perturbation is shown in terms of the change in $\Delta G_{HX}$ of each PUF, measured from the marker protons for each PUF. $\Delta \Delta G_{HX}$ of the targeted foldon in each case is shown in boldface. Values are in kcal/mol. pWT refers to recombinant pseudo wild-type Cyt *c*. (a) E62G–pWT.[9] The direct effect of the E62G mutation is focused on the yellow foldon that contains it. Infrared loop stability was measured at pDr 5.4. (b) Reduced–oxidized.[3,6] The direct oxidation/reduction effect focuses on the heme iron to Met80-S ligand bond in the red loop. (c) T47G–pWT. The direct destabilizing effect of the T47G mutation is on the infrared loop that contains it. Infrared loop stability was measured at pDr 5.4. The change in blue unit unfolding was measured from change in global stability by equilibrium GdmCl-induced unfolding (at both pH 5.0 and 7.0). (d) pDr 4.5–pDr 7. The direct effect of low pH is to protonate two buried groups that support the infrared loop (His26 and a heme propionate when they are transiently exposed). The $\Delta \Delta G_{HX}$ value listed for the global blue unfolding has been approximately corrected for an additional effect due to His26 protonation as described in the text.

Cyt *c* folding/unfolding ladder are able to fold and unfold separately and in no obligatory order.

**Additional effects**

In two of the trials just described, the global unfolding (blue foldon) registers a larger effect than the foldon that was targeted. The increased stability change occurs because the imposed perturbation is not limited to the targeted low-lying foldon but has an additional effect at some higher point on the unfolding ladder. These observations do not challenge the present demonstration that the unfolding of either the red or infrared loop does not obligatorily include the other but they do have interest in respect to the use of stability labeling experiments.

In the oxidized to reduced comparison (Figure 5(b)), heme iron reduction strengthens the Fe ligand bond

by 3.2 kcal/mol and stabilizes the red unit and all higher foldons by the same amount, but the global unfolding is stabilized by an additional 1.9 kcal/mol. This occurs because the effect of heme iron reduction is not limited to the Fe to S bond. The buried heme iron is charged in oxidized Cyt *c* and is therefore destabilizing but it is neutral in the reduced form. Evidently the buried charge is unmasked and expresses its destabilizing effect (1.9 kcal/mol)[40] only upon global unfolding, so that only the final blue unfolding registers this additional contribution.

In the pH perturbation experiment (Figure 5(d)), low pH (pDr 4.5) destabilizes the infrared loop by 3.3 kcal/mol because it is supported by the buried protonatable His26 and propionate groups.[6] The higher foldons are destabilized similarly. The global unfolding is promoted by an additional 1.2 kcal/mol (4.5 kcal/mol total, calculated from the known 12.8 kcal/mol of global

stability at pDr 7[2,3,34] minus 8.3 kcal/mol at pDr 4.5 from Figure 3). This extra destabilization is due to the fact that protonation of the transiently exposed His26 has an additional effect. The effect can be detected directly and evaluated approximately in the behavior of the green loop, which contains His26. As pH is decreased below the normal histidine $pK_a$ (~6.5), the green loop is sharply destabilized and can be induced to unfold separately from the green helix (highlighted in green in Figure 3, top right). The effect is evidently due to the suppressed $pK_a$ of His26 ($pK_a$ < 3.6),[41] buried in the green loop in the folded state, since it is not seen in the mutant His26Asn (H.M., unpublished results). The size of the effect on the green loop (>0.7 kcal/mol) approximately accounts for the additional global destabilization that is observed (1.2 kcal/mol), as follows. A minimum estimate for the total green loop destabilization, measured using Leu32 in the green loop, is 4 kcal/mol (pDr 4.5–pDr 7; Figure 3), higher than the expected 3.3 kcal/mol by 0.7 kcal/mol. This value underestimates the additional effect because Leu32 has not yet joined the isotherm at pDr 7, so its measured $\Delta G_{HX}$ is lower than the correct value for green loop unfolding. Since the global unfolding must contain both effects, the comparison in Figure 5(d) lists < 3.8 kcal/mol for the global blue unit ($-(<3.8) = -(4.5 - >0.7)$).

It is also interesting that, for each perturbation of the red or the infrared loop, the other loop may be fractionally affected. One possibility is that the unfolding of either loop can promote but not require the other. The two loops are in extensive mutually stabilizing contact.[6,42] Alternatively, some additional effect, not limited to the targeted loop, may occur for the different perturbants used (mutations, pH, redox state), just as for the additional effects just noted.
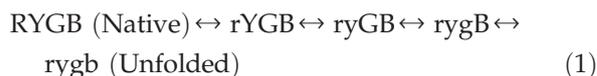
A similar behavior appears in previous work designed to test the sequential folding model in mutational experiments with yeast iso-1-Cyt *c*. Kristinsson and Bowler[42] imposed a number of mutations at the buried Asn52 position, then measured the pH-dependent alkaline transition as a proxy for red loop stability,[43] and denaturant-dependent global melting as a proxy for the blue unit. These measures were closely equivalent, consistent with the ladder model. However, the result seems at odds with our present results, insofar as the red loop unfolding seemed to include the infrared loop. The resolution seems to be that stability changes due to Asn52 mutations are likely to be widespread,[44–46] and likely to exert an additional effect on the red loop. Asn52 is an absolutely conserved residue;[47] it occupies a critical internal position between the red and infrared loops; it interacts with a heme propionate and with Ile75; and it helps to stabilize an internal water molecule that participates in an extensive hydrogen bonding network,[44,48] all of which can influence the red loop directly rather than just indirectly through a common unfolding with the infrared loop.

As a general observation, these results underline the difficulty of imposing a truly localized perturbation.

## Discussion
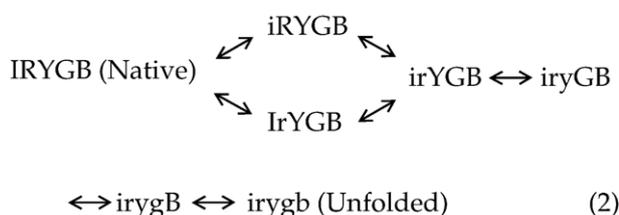
### Foldons and protein folding

Previous results indicated that the different Cyt *c* foldons unfold and refold in a sequential pathway manner as in equation (1) (letters refer to the color coding in Figure 1; upper case folded, lower case unfolded). Starting from the folded native protein, the red unit unfolds as an initial step (equation (1) left to right). The yellow unit unfolds only when the red unit is already unfolded, so that the measured yellow unfolding reflects a partially unfolded form with the red and yellow units open together. The higher green unfolding produces a PUF with green + yellow + red unfolded. The final blue unfolding produces the globally unfolded state:[49]

$$\text{RYGB (Native)} \leftrightarrow \text{rYGB} \leftrightarrow \text{ryGB} \leftrightarrow \text{rygB} \leftrightarrow$$
$$\text{rygb (Unfolded)} \tag{1}$$

The HX experiments that led to this picture were done under equilibrium native conditions (pDr 7, 20 °C or 30 °C), far below the melting condition ($C_m = 2.75$ M GdmCl). To maintain each component in equation (1) at equilibrium, each unfolding arrow must be matched by an equal arrow in the refolding direction. (If side-reactions exist, the same requirement holds for each individual reaction.) Therefore, the demonstration of a stepwise sequential unfolding pathway requires that refolding proceeds in the opposite direction through the same set of intermediates. Related results for other proteins consistently support analogous foldon-based pathways.[10,13–23]

It seems most interesting that the experimentally derived folding sequence summarized in equation (1) is just as expected if the pathway is directed by the foldon arrangement in the native Cyt *c* protein.[38] The present work pursued one of the implications of this sequential stabilization principle. Unlike the higher energy foldons, the structure of Cyt *c* does not appear to require the red and infrared Ω-loop foldons to unfold and refold in a pre-defined sequential manner. The results obtained here show that these two loops can unfold separately, with different free energy. They expose very similar amounts of new surface, indicating that they do not predominantly unfold in an obligatory sequential order, in which case one unfolding would necessarily include the other. The same conclusion comes from specific stabilization and destabilization experiments. The loops can unfold and refold alternatively, in parallel, as in equation 2, just as implied by the

sequential stabilization principle (i represents infrared):

$$\begin{array}{c} \text{iRYGB} \\ \text{IRYGB (Native)} \quad \nwarrow\!\!\nearrow \quad \swarrow\!\!\searrow \quad \text{irYGB} \leftrightarrow \text{iryGB} \\ \text{IrYGB} \quad \searrow\!\!\swarrow \quad \nearrow \end{array}$$

$$\leftrightarrow \text{irygB} \leftrightarrow \text{irygb (Unfolded)} \qquad (2)$$

### General implications

Results now available bear on the nature and ordering of the steps that construct protein folding pathways. Cyt *c* unfolds and refolds in stepwise manner. The steps are determined by the cooperative nature of its constituent foldon units. The generality of this behavior is suggested by the fact that similar results have been obtained for several other proteins.

The component foldons of Cyt *c* are essentially identical with its native secondary elements. This need not be completely true in general,[16] because secondary structural elements are not usually independently stable. In native proteins, they are stabilized and shaped by their supporting native tertiary interactions. Partially folded forms are known to incorporate non-native energy-minimizing interactions.[20,50–53] Accordingly, given the loss of some native interactions and the presence of non-native interactions, foldons in incompletely folded states need not accurately duplicate the elements seen in the native protein.

Previous results suggest that the native-like foldon arrangement dictates the order of folding steps according to a sequential stabilization mechanism. The present results support this conclusion in an interesting way. Where Cyt *c* structure clearly dictates a linear folding sequence, this is found. Where the structure allows alternatives, in the final two folding steps, this is found. The indication is that pathways may be linear, branched or parallel, depending on how the foldon units interact with each other in the native-like context.

These conclusions seem unlikely to be limited to Cyt *c* or to depend on some artifact of the conditions or methods used. They rest on detailed structural information, obtained under native conditions, for Cyt *c* and other proteins, with and without heme groups, with modest destabilization by different destabilants (urea, GdmCl, pH, temperature, pressure) or none at all, in HX and non-HX experiments.[1–10,13–23]

Different inferences have often been drawn from other experimental observations and theoretical simulations.[29,30] We note that most experimental methods are intrinsically unable to characterize, or even to detect, folding intermediates. This is true for kinetic experiments where intermediates occur after the initial rate-limiting step, for melting experiments where intermediates are unstable, and for experiments at equilibrium where intermediates are minimally populated. Even the NHX experiment succeeds in distinguishing minimally populated, partially folded states only in favorable cases.[32] Similarly, many theoretical simulations depend on the still incomplete treatment of cooperative protein interactions.[54] This is the property that fundamentally determines the identity of foldon units and their pathway formation by the sequential stabilization mechanism.

## Materials and Methods

The purity of WT equine Cyt *c* (type VI, Sigma Chemical Co.) was checked and it was further purified when necessary by reverse phase HPLC.[55] Recombinant pWT Cyt *c* (H26N, H33N) and its mutants were expressed in a high-yield *Escherichia coli* system and purified as described.[55] All other chemicals were as described.[6,9] Appropriate pH buffers (0. 1 M) were used with p$K_a$ close to the pH in each NHX experiment. All experiments were done at 20 °C. Solutions contained 0. 5 M KCl.

NHX experiments and data analysis were done as described.[9,32] Typical protein concentration was ~6 mM. The dead time from the start of the HX reaction was ~15 min. To collect faster time-points, short gradient correlated spectroscopy (COSY) spectra (two scans, ~23 min) were collected initially back-to-back.

$\Delta G_{HX}$ values were calculated from the equation:

$$\Delta G_{HX} = -RT\ln K_{op} = -RT \ln(k_{ex}/k_{ch})$$

which holds in the EX2 region below pH 10. Here, $k_{ex}$ is the measured exchange rate and $k_{ch}$ is the chemical exchange rate calculated for unprotected amides.[56,57] Errors in reported $\Delta G_{HX}$ and $m$ values are of the order of 0.1 kcal/mol for fast-exchanging amides and less for slow-exchanging amides.

## References

1. Roder, H., Elöve, G. A. & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR. *Nature*, **335**, 700–704.
2. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.
3. Xu, Y., Mayne, L. & Englander, S. W. (1998). Evidence for an unfolding and refolding pathway in cytochrome c. *Nature Struct. Biol.* **5**, 774–778.
4. Milne, J. S., Xu, Y., Mayne, L. C. & Englander, S. W. (1999). Experimental study of the protein folding landscape: unfolding reactions in cytochrome *c*. *J. Mol. Biol.* **290**, 811–822.

5. Hoang, L., Bédard, S., Krishna, M. M. G., Lin, Y. & Englander, S. W. (2002). Cytochrome *c* folding pathway: kinetic native-state hydrogen exchange. *Proc. Natl Acad. Sci. USA,* **99**, 12173–12178.

6. Krishna, M. M. G., Lin, Y., Rumbley, J. N. & Englander, S. W. (2003). Cooperative omega loops in cytochrome *c*: role in folding and function. *J. Mol. Biol.* **331**, 29–36.

7. Krishna, M. M. G., Lin, Y., Mayne, L. & Englander, S. W. (2003). Intimate view of a kinetic protein folding intermediate: residue-resolved structure, interactions, stability, folding and unfolding rates, homogeneity. *J. Mol. Biol.* **334**, 501–513.

8. Krishna, M. M. G., Lin, Y. & Englander, S. W. (2004). Protein misfolding: optional barriers, misfolded intermediates, and pathway heterogeneity. *J. Mol. Biol.* **343**, 1095–1109.

9. Maity, H., Maity, M. & Englander, S. W. (2004). How cytochrome *c* folds, and why: submolecular foldon units and their stepwise sequential stabilization. *J. Mol. Biol.* **343**, 223–233.

10. Maity, H., Maity, M., Krishna, M. M. G., Mayne, L. & Englander, S. W. (2005). Protein folding: the stepwise assembly of foldon units. *Proc. Natl Acad. Sci. USA,* **102**, 4741–4746.

11. Weinkam, P., Zong, C. & Wolynes, P. G. (2005). A funneled energy landscape for cytochrome *c* directly predicts the sequential folding route inferred from hydrogen exchange experiments. *Proc. Natl Acad. Sci. USA,* **102**, 12401–12406.

12. Pletneva, E. V., Gray, H. B. & Winkler, J. R. (2005). Snapshots of cytochrome *c* folding. *Proc. Natl Acad. Sci. USA,* **102**, 18397–18402.

13. Chamberlain, A. K., Handel, T. M. & Marqusee, S. (1996). Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nature Struct. Biol.* **3**, 782–787.

14. Chamberlain, A. K. & Marqusee, S. (2000). Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms. *Advan. Protein Chem.* **53**, 283–328.

15. Cecconi, C., Shank, E. A., Bustamante, C. & Marqusee, S. (2005). Direct observation of the three-state folding of a single protein molecule. *Science,* **309**, 2057–2060.

16. Chu, R., Pei, W., Takei, J. & Bai, Y. (2002). Relationship between the native-state hydrogen exchange and folding pathways of a four-helix bundle protein. *Biochemistry,* **41**, 7998–8003.

17. Fuentes, E. J. & Wand, A. J. (1998). Local stability and dynamics of apocytochrome b562 examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry,* **37**, 9877–9883.

18. Fuentes, E. J. & Wand, A. J. (1998). Local dynamics and stability of apocytochrome b562 examined by hydrogen exchange. *Biochemistry,* **37**, 3687–3698.

19. Feng, H., Takei, J., Lipsitz, R., Tjandra, N. & Bai, Y. (2003). Specific non-native hydrophobic interactions in a hidden folding intermediate: Implications for protein folding. *Biochemistry,* **42**, 12461–12465.

20. Feng, H., Zhou, Z. & Bai, Y. (2005). A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc. Natl Acad. Sci. USA,* **102**, 5026–5031.

21. Yan, S., Kennedy, S. D. & Koide, S. (2002). Thermodynamic and kinetic exploration of the energy landscape of *Borrelia budgdorferi* OspA by native-state hydrogen exchange. *J. Mol. Biol.* **323**, 363–375.

22. Yan, S., Gawlak, G., Smith, J., Silver, L., Koide, A. & Koide, S. (2004). Conformational heterogeneity of an equilibrium folding intermediate quantified and mapped by scanning mutagenesis. *J. Mol. Biol.* **338**, 811–825.

23. Silverman, J. A. & Harbury, P. B. (2002). The equilibrium unfolding pathway of a (β/α)₈ barrel. *J. Mol. Biol.* **324**, 1031–1040.

24. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science,* **181**, 223–230.

25. Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44–45.

26. Levinthal, C. (1969). How to fold graciously. *Mossbauer Spectroscopy in Biological Systems, Proceedings, University of Illinois Bulletin,* vol. 67, pp. 22–24, Urbana, IL 61801: University of Illinois Press.

27. Kim, P. S. & Baldwin, R. L. (1982). Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**, 459–489.

28. Kim, P. S. & Baldwin, R. L. (1990). Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631–660.

29. Plotkin, S. S. & Onuchic, J. N. (2002). Understanding protein folding with energy landscape theory. Part I: basic concepts. *Quart. Rev. Biophys.* **35**, 111–167.

30. Plotkin, S. S. & Onuchic, J. N. (2002). Understanding protein folding with energy landscape theory. Part II: quantitative concepts. *Quart. Rev. Biophys.* **35**, 205–286.

31. Bai, Y. & Englander, S. W. (1996). Future directions in folding: the multi-state nature of protein structure. *Proteins: Struct. Funct. Genet.* **24**, 145–151.

32. Krishna, M. M. G., Hoang, L., Lin, Y. & Englander, S. W. (2004). Hydrogen exchange methods to study protein folding. *Methods,* **34**, 51–64.

33. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.

34. Milne, J. S., Mayne, L., Roder, H., Wand, A. J. & Englander, S. W. (1998). Determinants of protein hydrogen exchange studied in equine cytochrome *c*. *Protein Sci.* **7**, 739–745.

35. Maity, H., Lim, W. K., Rumbley, J. N. & Englander, S. W. (2003). Protein hydrogen exchange mechanism: local fluctuations. *Protein Sci.* **12**, 153–160.

36. Krantz, B. A., Moran, L. B., Kentsis, A. & Sosnick, T. R. (2000). D/H amide kinetic isotope effects reveal when hydrogen bonds form during protein folding. *Nature Struct. Biol.* **7**, 62–71.

37. Krantz, B. A., Srivastava, A. K., Nauli, S., Baker, D., Sauer, R. T. & Sosnick, T. R. (2002). Understanding protein hydrogen bond formation with kinetic H/D amide isotope effects. *Nature Struct. Biol.* **9**, 458–463.

38. Rumbley, J., Hoang, L., Mayne, L. & Englander, S. W. (2001). An amino acid code for protein folding. *Proc. Natl Acad. Sci. USA,* **98**, 105–112.

39. Bushnell, G. W., Louie, G. V. & Brayer, G. D. (1990). High-resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.* **214**, 585–595.

40. Cohen, D. S. & Pielak, G. J. (1995). Entropic stabilization of cytochrome *c* upon reduction. *J. Am. Chem. Soc.* **117**, 1675–1677.

41. Moore, G. R. & Pettigrew, G. W. (1990). *Cytochromes c: Evolutionary, Structural and Physicochemical Aspects,* Berlin and Heidelberg: Springer-Verlag.

42. Kristinsson, R. & Bowler, B. E. (2005). Communication of stabilizing energy between substructures of a protein. *Biochemistry,* **44**, 2349–2359.

43. Maity, H., Rumbley, J. N. & Englander, S. W. (2006). Functional role of a protein foldon - An Ω-loop foldon

controls the alkaline transition in ferricytochrome *c*. *Proteins: Struct. Funct. Bioinform.* **63**, 349–355.

44. Redzic, J. S. & Bowler, B. E. (2005). Role of hydrogen bond networks and dynamics in positive and negative cooperative stabilization of a protein. *Biochemistry*, **44**, 2900–2908.

45. Linske-O'Connell, L. I., Sherman, F. & McLendon, G. (1995). Stabilizing amino acid replacements at position 52 in yeast iso-1-cytochrome *c*: in vivo and in vitro effects. *Biochemistry*, **34**, 7094–7102.

46. Lett, C. M., Berghuis, A. M., Frey, H. E., Lepock, J. R. & Guillemette, J. G. (1996). The role of a conserved water molecule in the redox-dependent thermal stability of iso-1-cytochrome *c*. *J. Biol. Chem.* **271**, 29088–29093.

47. Banci, L., Bertini, L., Rosato, A. & Varani, G. (1999). Mitochondrial cytochromes *c*: a comparative analysis. *J. Biol. Inorg. Chem.* **4**, 824–837.

48. Berghuis, A. M., Guillemette, J. G., McLendon, G., Sherman, F., Smith, M. & Brayer, G. D. (1994). The role of a conserved internal water molecule and its associated hydrogen bond network in cytochrome *c*. *J. Mol. Biol.* **236**, 786–799.

49. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1994). Protein stability parameters measured by hydrogen exchange. *Proteins: Struct. Funct. Genet.* **20**, 4–14.

50. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nature Struct. Biol.* **9**, 209–216.

51. Sosnick, T. R., Dothager, R. S. & Krantz, B. A. (2004). Differences in the folding transition state of ubiquitin indicated by $\phi$ and $\psi$-analysis. *Proc. Natl Acad. Sci. USA*, **101**, 17377–17382.

52. Religa, T. L., Markson, J. S., Mayor, U., Freund, S. M. V. & Fersht, A. R. (2005). Solution structure of a protein denatured state and folding intermediate. *Nature*, **437**, 1053–1056.

53. Nishimura, C., Dyson, H. J. & Wright, P. E. (2006). Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin. *J. Mol. Biol.* **355**, 139–156.

54. Kaya, H. & Chan, H. S. (2005). Explicit-chain model of native-state hydrogen exchange: implications for event ordering and cooperativity in protein folding. *Proteins: Struct. Funct. Bioinform.* **58**, 31–44.

55. Rumbley, J. N., Hoang, L. & Englander, S. W. (2002). Recombinant equine cytochrome *c* in *Escherichia coli*: high-level expression, characterization, and folding and assembly mutants. *Biochemistry*, **41**, 13894–13901.

56. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 75–86.

57. Connelly, G. P., Bai, Y., Jeng, M.-F. & Englander, S. W. (1993). Isotope effects in peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 87–92.

58. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 945–949.

59. Hvidt, A. & Nielsen, S. O. (1966). Hydrogen exchange in proteins. *Advan. Protein Chem.* **21**, 287–386.

*Edited by C. R. Matthews*