ELSEVIER

Biophysical Chemistry 100 (2003) 409–419

Biophysical Chemistry

www.elsevier.com/locate/bpc

Cytochrome c_{551} as a model system for protein folding☆

Maurizio Brunori*, Maria Giulia Bigotti, Francesca Cutruzzolà, Stefano Gianni[1], Carlo Travaglini-Allocatelli

Istituto Pasteur-Fondazione Cenci Bolognetti, Department of Biochemical Sciences, Università di Roma 'La Sapienza', P. le A. Moro 5, 00185 Roma, Italy

Received 19 February 2002; accepted 28 April 2002

Abstract

Considerable progress was made over the last few years in understanding the mechanism of folding of cytochrome c_{551}, a small acidic hemeprotein from *Pseudomonas aeruginosa*. Comparison of our results with those obtained by others on horse heart cytochrome *c* allows to draw some general conclusions on the structural features that are common determinants in the folding of members of the cytochrome *c* family.

Keywords: Folding kinetics; Intermediates; Cytochromes

1. Introduction

Protein folding is the process by which a polypeptide chain achieves its specific biologically active three-dimensional structure, starting from a disordered random coil state. Since the pioneering work of Anfinsen [1], the mechanism of protein folding has attracted a lot of attention but still is a major unresolved problem of modern biochemistry and structural biology. The classical approach to unravel the mechanism of protein folding envisages that in order to be complete in a finite time, refolding of a denatured protein should occur through a specific reaction pathway, and several models have been proposed and discussed [2]. Experimentally, this hypothesis focused primarily on the identification and characterisation of folding intermediates, i.e. partially folded states crucial to achieve the native conformation by reducing the search of its conformational space [3]. Elucidation of the folding mechanism would require that all the conformational states on the pathway (including the native and denatured states, possible folding intermediates and the relevant transition states) are known in detail at the level of individual residues. It is accepted that this goal can only be achieved by a kinetic investigation of the folding and unfolding reactions of the wild-type and many different site-directed mutants, applying extensive-

☆ This paper is dedicated to the memory of Prof. John T. Edsall and Prof. Jeffries Wyman, two giants of *Biophysical Chemistry* that contributed enormously to science in the Rome Biochemistry Institute.

*Corresponding author. Tel.: +39-6-445-0291; fax: +39-6-444-0062.

E-mail address: maurizio.brunori@uniroma1.it (M. Brunori).

[1] Cambridge Centre for Protein Engineering, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, UK.

0301-4622/03/$ - see front matter
PII: S0301-4622(02)00295-8

ly a protein engineering strategy, as pioneered by Fersht [4].

The characterisation of proteins which fold according to a two-state process [5], the discovery of different types of kinetic traps [6,7] and the development of new theoretical models of protein folding [8] have, however, opened an intense debate about the significance of folding intermediates. Only very recently, with the development of techniques that allow to investigate the folding reaction on a sub-millisecond time range, the kinetic role of partially folded states has been directly addressed [9–11]. The distinction between two- and three-state folders may appear to be a major generalisation brought about by protein folding studies over the last few years. From an experimental viewpoint, this distinction is generally based on the shape of the chevron plot (i.e. the log of the observed rate constant vs. denaturant concentration [D]). While the chevron plot of 2-state proteins is V-shaped (because of the linear dependence of $\log k_{obs}$ on [D]), 3-state proteins typically exhibit deviations from linearity (*roll-over*) at low denaturant concentrations. This behaviour was initially taken as the best evidence for the accumulation of an intermediate under refolding conditions. However, more recently, other equally valid explanations to account for curvatures in chevron plots have been proposed based (for example) on a variable transition state ensemble (TSE) [12]: according to this view, even a 2-state protein may present a *roll-over* effect reflecting a broad and changeable free-energy barrier, rather than accumulation of an intermediate.

Mitochondrial cytochrome *c* (cyt *c*) has been a widely employed model system for equilibrium and kinetic folding studies which provided a wealth of information on its putative folding pathway and led to the identification of possible intermediate states [13,14]. Studies on the kinetics of refolding and unfolding of oxidised horse heart cyt *c*, followed using optical probes and hydrogen exchange labelling methods, have revealed a complex folding landscape. Some kinetic traps were attributed (i) to *cis*/*trans* isomerization of prolyl peptide bonds [15], or (ii) to non-native coordination to the heme iron of His(26) or His(33) in place of Met(80), occurring at approximately neutral pH and slowing down the refolding rate [16–18]. Interestingly, a partially structured intermediate with tertiary contacts between the N- and C-terminal α-helices (a common motif in class-I cytochrome *c* fold) was claimed to be an early, on-pathway, intermediate in the cyt *c* folding mechanism [3,19]; a related question is whether this intermediate can be populated even under acidic conditions, i.e. in the absence of miscoordination events [3,16,17]. In spite of extensive studies, however, the folding mechanism of cyt *c* is still controversial because it is debated whether it follows a two-state [20–22] or a multi-state model with specific, on-pathway, intermediates [3,13,19].

We decided to address the protein folding problem employing a comparative approach, i.e. making use of the simple strategy of comparing intermediates and transition states of proteins having the typical cytochrome *c* fold, but widely different sequences. Thus, we have extensively investigated the cytochrome c_{551} from the bacterium *Pseudomonas aeruginosa* (cyt c_{551}). This is a periplasmic protein belonging (like horse cyt *c*) to class I cytochrome *c* family, and characterised by the typical fold of this family with only ~30% homology in amino acid sequence compared to that of its mammalian counterpart. Peculiar properties of cyt c_{551} (Fig. 1), making it an attractive molecule and a suitable model for folding studies are: (i) its minimal size (82 residues vs. 104); (ii) an acidic isoelectric point (pI=4.7), which allows to investigate refolding at zero net charge; (iii) the presence of two tryptophanyl residues, Trp(56) [homologous to Trp(59) of horse cyt *c*] and Trp(77) located in the C-terminal α-helix; (iv) the presence of only one His(16) (providing proximal co-ordination to the heme Fe), which reduces the chances of kinetic traps due to miscoordination events. In the following we illustrate how each of these features permitted to obtain important insights on the mechanism of cyt c_{551} folding.

2. Results and discussion

2.1. Folding mechanism of wild-type cyt c_{551} in urea

Estimates of protein stability from GdnHCl and urea denaturation equilibrium data may differ

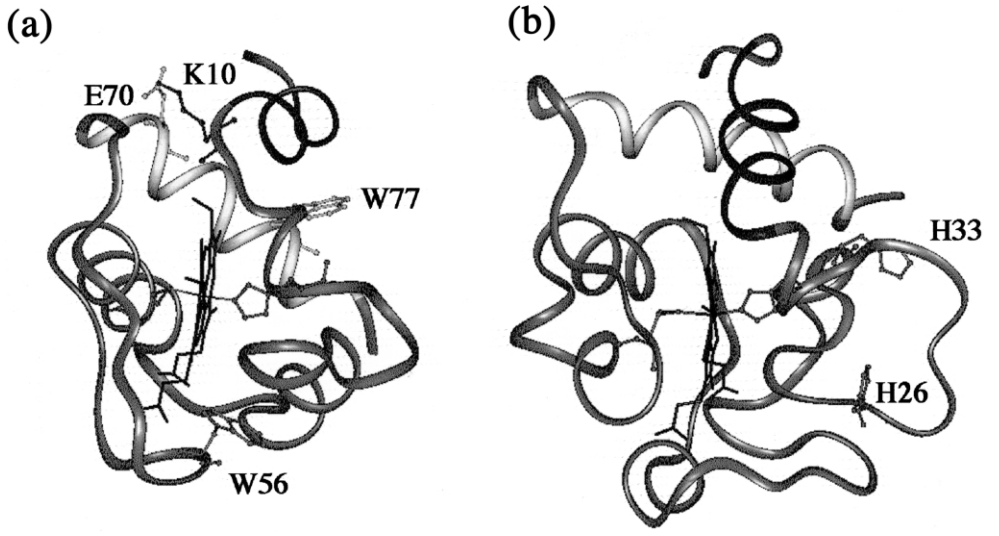

Fig. 1. Backbone structures of (left) *Pseudomonas aeruginosa* cytochrome c_{551} (PDB 351c) and (right) horse heart cytochrome *c* (PDB 1hrc). The N-terminal α-helix and the heme group are shown in black for both proteins. Relevant residues are shown as ball-and-stick representation; the axial ligands in the native proteins are His(16) and Met(61) for cyt c_{551}, and His(18) and Met(80) for cyt *c*. For cyt c_{551}, residues Lys(10) and Glu(70) forming a salt bridge, and residues Trp(56) and Trp(77) are also highlighted. For cyt *c*, the side chains of other His residues that can coordinate to the heme iron during refolding experiments (miscoordination) are highlighted.

depending on the hydrophobic and electrostatic make-up of any given protein [23,24]. In the course of our study on cyt c_{551}, we realised that the ionic nature of GdnHCl affects the stability of this cytochrome, and therefore introduces an extra variable in the experiments, as observed for other proteins with an acidic isoelectric point (see for example, Staniforth et al. [25]). In fact, we determined that assuming a linear relationship between Gibbs free-energy and denaturant concentration [26,27], at pH 4.7 the stability of wt cyt c_{551} is ~1.4 kcal mol^{-1} lower using GdnHCl ($\Delta G^{\circ}_{UN} = -6.6$) as compared to urea ($\Delta G^{\circ}_{UN} = -8.0$).

Upon rapid dilution of a denaturant initially at high concentration, the time course of cyt c_{551} refolding observed by tryptophan fluorescence in the millisecond time range can be fitted with a single exponential process (when corrected for slow events due to proline *cis*/*trans* isomerisation, accounting for <10% of the observed amplitude). Comparison of the chevron plots obtained at pH 4.7 in urea and GdnHCl [28] indicates that the *roll-over* effect seen in the refolding branch is more pronounced using the former denaturant. In view of what is seen for other proteins [29], we suspected that part of this *roll-over* effect may be related to the formation of transient aggregates, rather than accumulation of intermediate state(s). The presence of two Trp residues, each contributing approximately 50% of the fluorescence emission of the unfolded state, permitted us to carry out refolding experiments at very low protein concentrations. Indeed at the isoelectric point (pH 4.7) the refolding rates observed with very dilute protein are consistent with those expected for a simple two-state model; however, at other pH values, the deviation from linearity observed in refolding experiments could not be explained solely on the basis of aggregation effects. The results at pH 7.0 (see Fig. 2) are representative of these conditions and are herein discussed.

While the presence of a *roll-over* effect in the refolding branch of the chevron plot has been often interpreted in terms of *on-pathway* transient intermediates [3,30], Oliveberg et al. [12] recently suggested that it may reflect a variability in the

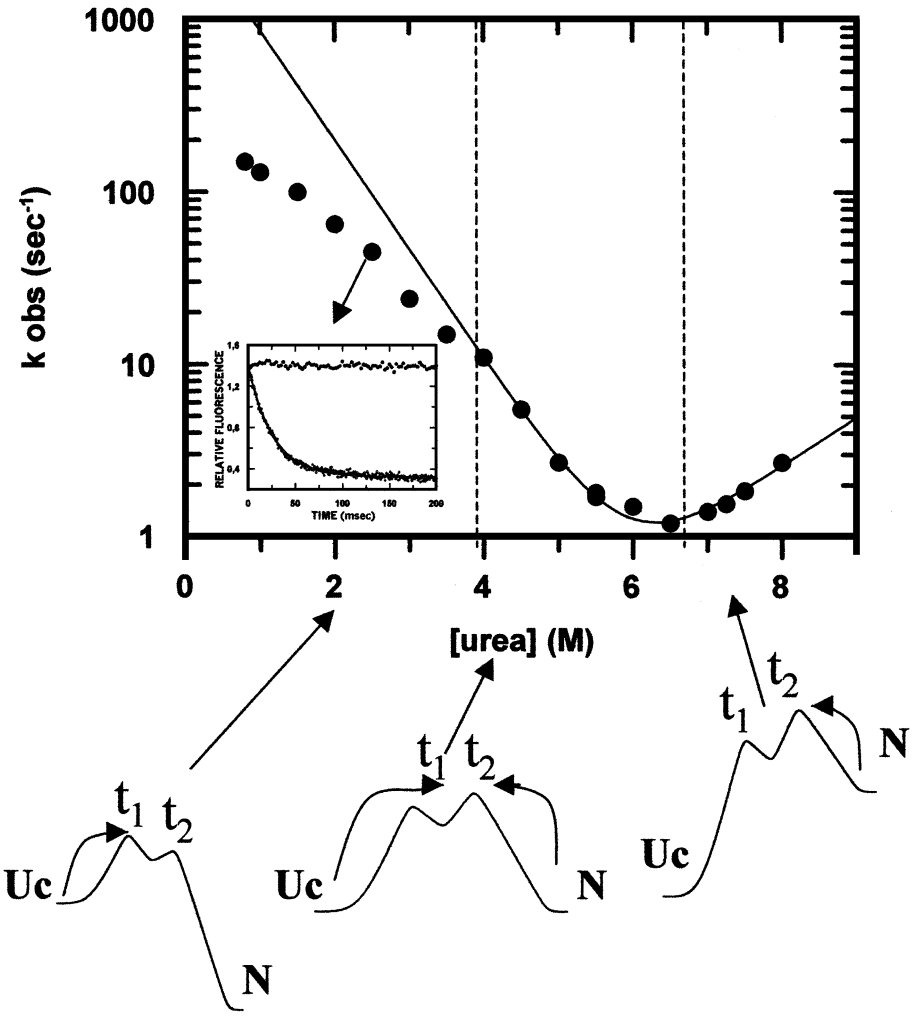

Fig. 2. Kinetics of refolding and unfolding of wt cyt c_{551} as a function of urea at pH 7.0 and 10 °C. An interpretation of the results in terms of transition state(s) is depicted below the graph. We have suggested that the change in slope of the refolding branch is due to an additional transition state (called t_1), which is less compact than transition state t_2, the dominant barrier at [urea] > 4 M. A representative refolding time course measured at 2.5 M urea (shown in the inset) is described by a single exponential without 'burst phase' (reprinted with permission from Gianni et al. [28]).

transition state on top of a broad activation barrier. The amplitude behaviour predicted by this model postulates that the refolding kinetics should be single-exponential (no 'missing' amplitude). In the case of cyt c_{551}, where a burst phase is evident below 2 M urea, the 'broad barrier' model may appear unsuitable; however, we observed that the *roll-over* was noticeable in a denaturant concentration range (2 M < [urea] < 4 M) where the refolding kinetics is single-exponential with no loss in signal recovery (inset of Fig. 2). Therefore, we assumed that the cyt c_{551} folding pathway may proceed through two different transition states, called t_1 and t_2.

In Fig. 2 we have depicted the situation where the rate limiting barrier at the lower denaturant concentrations is associated to a less compact transition state (t_1), consistently with the Ham-

Table 1
Kinetic parameters calculated from urea denaturation experiments at pH 7.0

Protein	$k(t_1)_F$ (s^{-1})	$k(t_2)_F$ (s^{-1})	ϕ[a]	k_F^{wt}/k_F^{MUT}	k_F^{wt}/k_F^{MUTb}
Wild type	~200	3900	–	–	–
Glu(70)Gln	~200	640	1	6.1	1.8
Glu(70)Val	~200	280	0.7	14	2.2

[a] $\phi = [-RT\ln(k_F^{MUT}/k_F^{wt})]/(\Delta G^{wt} - \Delta G^{MUT})$.
[b] In the presence of GdnHCl = 0.5 M.

mond postulate [31]. The observation that at pH 4.7 (zero net charge and maximal stability) the *roll-over* effect disappears at very low protein concentrations, suggests that the less compact transition state t_1 takes origin from electrostatic repulsion, which may limit the observed refolding rate. Thus, the proposal that the net charge above and below the p*I* controls the degree of compactness of the transition state t_1 and of the collapsed state [28,32], was possible because the acidic p*I* of cyt c_{551} allowed us to perform kinetic experiments at zero net charge and to probe the role of electrostatics in the folding process.

2.2. Characterisation of the Glu(70) mutants

We have carried extensive experiments to characterise the transition state t_2, in order to identify tertiary interactions that may be involved in its stabilisation. Initially, we proposed [32] that the salt bridge between Glu(70) and Lys(10) (located, respectively, in the first turn of the C-terminal α-helix and immediately after the end of the N-terminal α-helix, see Fig. 1) may be one of the key interactions formed in this transition state, guiding refolding to the native state. In order to test this hypothesis we have studied the two site-directed mutants Glu(70)Gln and Glu(70)Val [28].

The chevron plot of the wt protein and these two mutants in urea shows that at pH 7.0, $k(t_2)_F$ extrapolated to [urea] = 0 is much smaller for Glu(70)Gln and Glu(70)Val than that calculated for the wt protein (see Table 1). On the other hand, at pH 3.0, where the carboxylate of Glu(70) should be fully protonated in the unfolded state and therefore ineffective in forming a salt bridge with Lys(10), the $k(t_2)_F$ calculated for the wt protein and the two mutants is approx. the same. These observations support the hypothesis of a role of Glu(70)–Lys(10) salt bridge, and suggest that this specific interaction is formed in the transition state t_2.

The kinetic parameters reported in Table 1 allow to calculate a φ-value [4] which provides an estimate of the fractional contribution of a specific interaction to the stability of the transition state. The φ-value ~ 1 calculated for the mutant Glu(70)Gln indicates that the Glu(70)–Lys(10) salt bridge has a primary role in the stabilisation of transition state t_2 (~1.1 kcal/mol); on the other hand, mutation to a hydrophobic residue in Glu(70)Val introduces an additional destabilisation of the native state, as deduced from the lower φ-value (φ = 0.7).

As shown in Table 1, both Glu(70) mutations have no effect on the properties of the less compact transition state t_1 suggesting that this charged residue plays no detectable role in the stability of this transition state ensemble. Notice that the isoelectric point of these two mutants is ~5.7 as compared to p*I* = 4.7 for wt (data not shown). On the whole, although we have evidence that this barrier is controlled by the charge of the unfolded protein, it is likely that electrostatic effects involving some specific residues in the collapsed state play a dominant role. This hypothesis is consistent with the pH dependence of the amplitude of the kinetic phase related to the early collapse, previously reported by Travaglini-Allocatelli et al. [32].

2.3. Refolding kinetics reveals a mechanistic difference between urea and guanidine

Analysis of the chevron plot of wt cyt c_{551} revealed that the k_F^{wt} extrapolated to zero GdnHCl [k_F^{wt}(GdnHCl) ~ 300 s^{-1}] is much smaller than that calculated from urea experiments [k_F^{wt}(urea) ~ 4000 s^{-1}] [28,32]. Moreover, we observed that in urea, the ratio (k_F^{wt}/k_F^{mut}) is ~6 for Glu(70)Gln and ~14 for Glu(70)Val, while in GdnHCl it is ~2 for both mutants.

In order to analyse this effect of GdnHCl and to understand its mechanistic and kinetic significance, we carried out urea refolding experiments in the presence of a constant concentration of

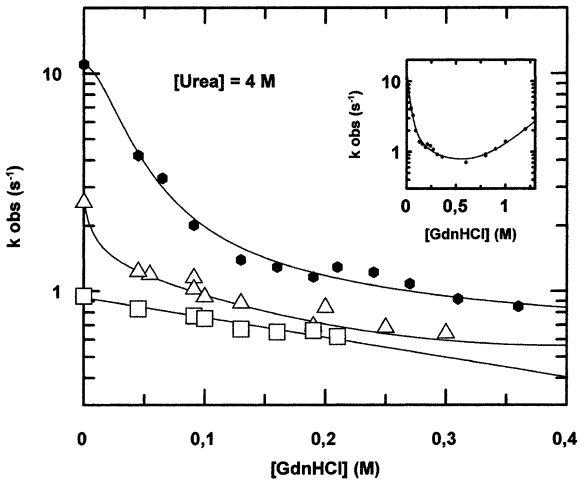

Fig. 3. Refolding kinetics of wt cyt c_{551} (closed circles) and of mutants Glu(70)Gln (open triangles) and Glu(70)Val (open squares) measured by fluorescence stopped-flow in GdnHCl dilution experiments at pH 7.0 and 10 °C, but in the presence of [urea] = 4 M. Inset: complete chevron plot for wt cyt c_{551}. It may be noticed that for the wt protein the refolding branch has an atypical 'upper curvature' that becomes evident at very low [GdnHCl]; this curvature is possibly still present for the Glu(70)Gln mutant, but is not seen in the Glu(70)Val mutant (reprinted with permission from Gianni et al. [33]).

GdnHCl [24,33,34]. At [GdnHCl] = 0.5 M, the ratio k_F^{wt}/k_F^{mut} is ~2 for both mutants (see Table 1), i.e. similar to that calculated from GdnHCl refolding experiments of the wt protein [32] and different from those obtained in urea [28]. It is important to notice that the calculated m_U and m_F values are the same as those calculated from 'classical' urea experiments (i.e. without added GdnHCl; data not shown), confirming that mixing a constant low concentration of a chaotropic denaturant with another does not alter the refolding/unfolding mechanism. The fact that for both mutants the values of k_F^{wt}/k_F^{mut} obtained from urea experiments in the presence of [GdnHCl] = 0.5 M are very similar to those obtained from dilution of GdnHCl alone, strongly suggests that the Glu(70)–Lys(10) salt-bridge is fully broken at 0.5 M GdnHCl.

To further test this hypothesis and to investigate the folding kinetics of these proteins below 0.5 M GdnHCl, we decided to measure the refolding rate constants of the wt and mutants of cyt c_{551} as a function of [GdnHCl], but this time in the presence of 4 M urea: at the latter concentration, the three proteins are still native as judged from far UV-CD and fluorescence spectroscopy, but destabilised. It is evident from the results reported in Fig. 3 that the refolding branch of the wt protein has an atypical and surprising 'upper curvature' which is seen only at very low [GdnHCl] (at or below 0.1 M); this behaviour is barely detectable in the Glu(70)Gln mutant, but absent in the Glu(70)Val mutant.

Since we have shown from urea dilution experiments that the Glu(70)–Lys(10) salt bridge is fully formed in the refolding transition state t_2 (see above), we suggested that this unusual 'upper-curved' chevron behaviour is linked to the breakage of this salt bridge which occurs already at [GdnHCl] ≅ 0.1 M. This hypothesis is consistent with the observation that in these experiments (i.e. at 4 M urea), k_F^{wt}/k_F^{mut} values extrapolated to zero [GdnHCl] are the same as those calculated from 'classical' urea dilution experiments.

2.4. Characterisation of the Trp(56) and Trp(77) mutants

The presence of two Trp residues located in different regions of the three-dimensional structure of cyt c_{551} (Fig. 1) offers the opportunity to probe the different environments experienced by these two aromatics. Trp(77) is part of the conserved hydrophobic cluster proposed by Ptytsin [35] to represent a folding nucleus for all cytochromes *c*, and is likely to be involved in the early collapse. Trp(56), located at the bottom of the heme pocket, is H-bonded to the heme propionate 17 (HP-17), as shown by NMR [36] and mutagenesis [37]. Therefore, two site-directed mutants, namely Trp(56)Phe and Trp(77)Phe, were expressed and characterised.

Since the chevron plot in urea suggested for both mutants a significant destabilisation of the native state but only a minor, if any, effect on $k(t_2)_F$, we carried out kinetic experiments on the wt and the two mutants under mildly destabilising conditions, following the urea dependence of the observed unfolding/refolding rates in the presence of [GdnHCl] = 0.5 M. Analysis of the data indi-

cates that both mutations have no effect on the stability of the refolding transition state t_2, since the calculated Φ values are very low. Thus, $\phi \sim 0.3$ obtained for Trp(77)Phe suggests that this aromatic does not significantly stabilise the transition state t_2 of the refolding reaction observed in the millisecond range, although we cannot exclude a role in the formation of the folding nucleus during early collapse, as predicted by Ptytsin [35]. Likewise, substitution of Trp(56) with Phe further destabilises the native state, the calculated Φ value for Trp(56)Phe being ~ 0.1.

For cyt c_2 [38] and horse cyt *c* [19], unfolding at high denaturant concentrations is rate limited by an additional energy barrier, which has been assigned to the breakage of the Fe^{3+}-Met distal bond; a similar behaviour was observed for cyt c_{551} in GdnHCl [32]. Data at pH 7.0 in the presence of GdnHCl = 0.5 M indicated that the rate limiting step at high [urea] (i.e. = 8 M) can be assigned a rate constant of $\sim 8\ s^{-1}$ for the wt protein. On the other hand, for the mutant Trp(56)Phe, the measured k_U at this denaturant concentration is higher ($\sim 50\ s^{-1}$). The p*K* of the alkaline transition measured following the 695 nm band is diagnostic of the stability of the Fe^{3+}–Met bond in the oxidised protein [39]. Since the p*K* of the alkaline transition is ~ 1.5 pH units lower in the Trp(56)Phe mutant compared to wt cyt c_{551} ($pK_a = 9.5$), it was proposed that the strength of the Fe^{3+}–Met(61) bond in the mutant is reduced because the H-bond between Trp(56) and the HP-17 propionate is lost [37]. In the same work, Molecular Dynamics computations carried out on wt cyt c_{551} and this mutant indicated a displacement of the loop containing Met(61) in the latter protein, a prediction consistent with the observed increase to $50\ s^{-1}$ of the limiting unfolding rate at high denaturant concentrations.

2.5. Initial collapse and the hypothesis of parallel pathways

Starting from the folding mechanism proposed for horse cyt *c* [13], our results on cyt c_{551} are consistent with a sequential process:

$$U \rightarrow I_C \rightarrow M \rightarrow N$$

in which the initial collapsed state (I_C), is formed very rapidly (sub-millisecond) as deduced from fluorescence stopped-flow experiments by amplitude analysis and burst phase quantification, and thereafter it folds to a native-like species [*M*, lacking the Met(61)–Fe bond] through a broad activation barrier characterised by two transition states (t_1 and t_2). Whereas the stabilisation of transition state t_1 is controlled by the electrostatic properties of the molecule, transition state t_2 is stabilised by some native-like tertiary contacts at the interface between the two terminal α-helices of the protein [e.g. notably the salt bridge between Lys(10) and Glu(70), see above]. The formation of the Met(61)–Fe bond, indicated as $M \rightarrow N$, may form without a lag following the population of *M*.

Because of its redox and ligand binding properties, the covalently bound heme group provides unique experimental opportunities for rapid initiation of folding [40,41]. In the presence of CO, reduced cyt *c* unfolds at lower denaturant concentrations than in the absence of this ligand; therefore, rapid photochemical removal of CO from unfolded cyt *c*, was expected to trigger a (reversible) folding process with transient rebinding of the distal Met to the heme iron. In view of the excellent time resolution of the photochemical experiment, this approach appeared well suited to provide direct information on the kinetics of the collapse sub-millisecond phase(s). However, in the case of horse cyt *c*, the time course following photolysis of CO was very complex and analysis of the multiphasic kinetics detected by optical absorption was heavily complicated by distal miscoordination events with additional His residues [40].

When we applied this photochemical approach to CO-bound cyt c_{551} in the presence of GdnHCl [42], we observed a similarly complex time course which was surprising because simpler kinetics was expected (in view of the absence of His residues that may miscoordinate to the heme iron). Careful kinetic analysis of the time course of CO recombination as a function (among others) of denaturant concentration, together with control experiments carried out with microperoxidase, heme and/or different ligands, indicated that some of these kinetic phases are not related to folding.

Arcovito et al. [42] demonstrated that this approach is complicated by the transient breakage of the proximal His(16)–Fe bond that may occur as a consequence of CO photodissociation in unfolded cyt *c*, because of the so-called *base elimination mechanism* [43]. Rebinding of CO to a transient four-coordinate heme yields optically detectable kinetic intermediates, unrelated to folding. Once again, a peculiar property of cyt c_{551}, namely the lack of additional His residues that may be involved in miscoordination events, allowed us to clarify some published observations and, ultimately, to explain why in the case of horse cyt *c*, so little fluorescence or CD evidence for the collapse of the polypeptide chain was recorded after photolysis of CO [44], despite the large changes in heme absorbance.

As a necessary next step towards a more complete description of the folding mechanism of cyt c_{551}, we have started an investigation of the initial sub-millisecond collapse of the protein by means of a continuous-flow capillary mixing method in collaboration with the group of Roder [13,45]. In the case of horse cyt *c*, the sub-millisecond time course was found to be biphasic, even the presence of imidazole [13] suggesting a minimal three-state sequential mechanism ($U^{Im} \rightarrow I_C^{Im} \rightarrow N^{Im}$). In the case of cyt c_{551}, a clear interpretation of the complex kinetics observed in the sub-millisecond time range (to be published) will demand a careful analysis of the urea dependencies of the fast processes observed at pH 4.0; nevertheless, this experiment suggested that the cyt c_{551} folding landscape may involve parallel pathways. To test this hypothesis, we carried out a double-jump interrupted refolding experiment following Kiefhaber [46]; unfolded cyt c_{551} was first mixed against a refolding buffer (*first mix*) and then, after a controlled delay time, refolding was interrupted by mixing rapidly against high concentrations of the same denaturant (*second mix*). Thus, the fraction of native molecules formed during the delay time is monitored by the relative amplitude of the slowest unfolding decay. This approach may allow to estimate the formation, during refolding, of intermediates and native molecules since unfolding induced by the *second mix* is expected to occur slowly only for the native protein, with the Met(61)–Fe bond formed.

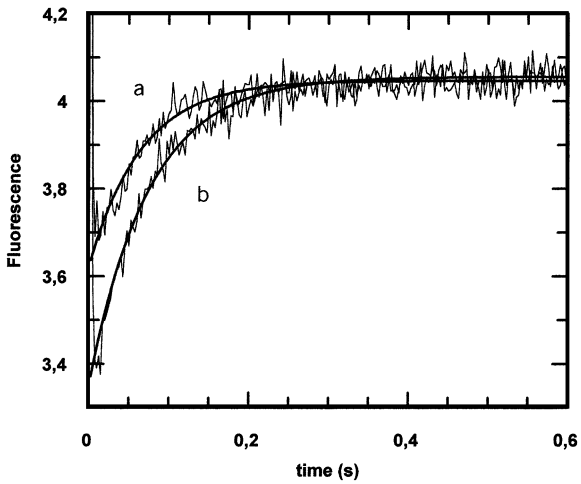

Fig. 4. Double mixing interrupted-refolding experiments on wt cyt c_{551} measured at pH 3.0 and 10 °C. Representative time courses recorded at 17 ms [trace (a)] and 4 s [trace (b)] delay time after the *first mix*, were obtained by a *second mixing* with high concentration GdnHCl (6 M). The observed kinetic traces obtained at different delay times are single-exponentials, as expected if only the unfolding of the native state is detectable in the millisecond time range.

Stopped-flow traces obtained at different delay times (Fig. 4) conform to single-exponential kinetics, as expected if only the unfolding of the native state was detectable in the millisecond time range. Surprisingly, amplitude analysis of the GdnHCl unfolding process indicates the formation of ~50% of the native protein even at very short delay times (<10 ms); moreover, the rate obtained from the time-dependence of the amplitude is consistent with the refolding kinetics measured from classical 'single mix' dilution experiment under identical conditions [28]. This suggests that a fast refolding pathway that allows a fraction of the unfolded protein to reach the fully native state via a rapid parallel process is active. Thus, we envisage that a (significant) fraction of the unfolded protein folds very rapidly (sub-milliseconds) and directly to N through a pathway involving no intermediate states. The demonstration of parallel folding pathways for cyt c_{551} agrees with Kiefhaber's experiments on lysozyme [46], and it is by-

and-large consistent with the new theories of protein folding [47]. The experimental conditions controlling the fraction of U folding directly to N are largely unknown. Although we have some evidence that identical experiments carried out in urea yield a much smaller fraction of molecules folding directly to N, structural features of the unfolded state controlling the partition between 'fast' and 'slow' folding pathways are unknown and a subject for future experiments.

3. Concluding remarks

After more than a decade of folding studies on a limited number of model proteins (including horse cyt *c*) and on Φ analysis of their site-directed mutants, experimental and computational efforts are nowadays directed towards the elucidation of the folding mechanism of structurally related homologous proteins [48]; it is generally believed that such a strategy may be helpful to the identification of the sequence determinants of a given fold.

Comparison of the folding kinetics of cyt c_{551} (see above) with extensive data on horse cyt *c* [13,14] and other cytochromes [49] suggests that more than one folding mechanism may be operative, in this as well in other protein families [50,51]. Incidentally, it may be remarked that the apparent diversity of models describing the folding of different class I cytochromes *c* contradicts the idea that the covalently bound heme group dominates the folding process in this protein family.

Pseudomonas cyt c_{551} folds to the native state through a broad energy barrier characterised by a variable transition state ensemble, and thus according to our interpretation no millisecond intermediate is populated. The transition state from I_C to *M* involves a specific tertiary interaction between Lys(10) and Glu(70) at the interface between N- and C- terminal helices [28]. In the case of horse cyt *c*, stopped flow experiments on wt and mutants have been interpreted [19] in terms of millisecond intermediate I_{NC} preceding the formation of the native protein. Notwithstanding some apparent differences between the two cytochromes, it seems interesting that both I_{NC} of horse cyt *c* and the transition state t_2 of cyt c_{551} are stabilised by the same network of tertiary interactions, i.e. those involving multiple contacts between the N- and C-terminal helices, which are homologous in the two proteins. This observation is at the outset consistent with the analysis published by Ptytsin [35], with the additional interesting finding that in cyt c_{551} a conserved topology (N- to C-terminal helices) is stabilised by a salt bridge rather than hydrophobic interactions. One should ask why in the case of horse cyt *c* this network of tertiary interactions seems to populate a millisecond intermediate I_{NC}, which is not so evident in the case of cyt c_{551}. A possibility is related to the much higher helical propensity of horse cyt *c*, which may control the extent to which a given, helical intermediate is populated, as pointed out also in the case of the globin family [52]. Alternatively the lack of such a species in cyt c_{551} (which is devoid of kinetic traps due to His miscoordination) may be taken as an evidence that the 'misfold–reorganization barrier' proposed by Englander [14] is in fact responsible for the accumulation of I_{NC} in the case of horse cyt *c*.

Finally, it may be concluded that the strategy based on a comparative approach to the folding of cytochromes *c*, which was our original motivation to undertake the work on *Pseudomonas* cyt c_{551} [53], seems to have provided new insights. In particular, kinetic analysis of the refolding of cyt c_{551} at the isoelectric point (pH 4.7), as well as above and below, allowed us to identify an effect of the net charge of the protein on the overall properties of the collapsed and the transition state t_1, an original observation that was not accessible to those studying horse cyt *c*.

Acknowledgments

This work was partially supported by the Ministero Istruzione Università Ricerca of Italy (Programma di Ricerca Scientifica di Interesse Nazionale 2001, 'Structural dynamics of metalloproteins') and the Consiglio Nazionale delle Ricerche (Progetto strategico 'Folding and misfolding of monomeric proteins').

References

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

[2] S.E. Radford, Protein folding: progress made and promises ahead, Trends Biochem. Sci. 25 (2000) 611–618.
[3] H. Roder, W. Colon, Kinetic role of early intermediates in protein folding, Curr. Opin. Struct. Biol. 7 (1997) 15–28.
[4] A.R. Fersht, A. Matouschek, L. Serrano, The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, J. Mol. Biol. 224 (1992) 771–782.
[5] S.E. Jackson, How do small single-domain proteins fold?, Fold. Des. 3 (1998) R81–91.
[6] G.A. Elöve, A.F. Chaffotte, H. Roder, M.E. Goldberg, Early steps in cytochrome *c* folding probed by time-resolved circular dichroism and fluorescence spectroscopy, Biochemistry 31 (1992) 6876–6883.
[7] G. Pappenberger, H. Aygun, J.W. Engels, U. Reimer, G. Fischer, T. Kiefhaber, Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics, Nat. Struct. Biol. 8 (2001) 452–458.
[8] P.G. Wolynes, J.N. Onuchic, D. Thirumalai, Navigating the folding routes, Science 267 (1995) 1619–1620.
[9] W.A. Eaton, V. Muñoz, S.J. Hagen, et al., Fast kinetics and mechanisms in protein folding, Ann. Rev. Biophys. Biomol. Struct. 29 (2000) 327–359.
[10] A.P. Capaldi, M.C. Shastry, C. Kleanthous, H. Roder, S.E. Radford, Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate, Nat. Struct Biol. 8 (2001) 68–72.
[11] K. Kuwata, R. Shastry, H. Cheng, et al., Structural and kinetic characterization of early folding events in beta-lactoglobulin, Nat. Struct. Biol. 8 (2001) 151–155.
[12] M. Oliveberg, Y.J. Tan, M. Silow, A.R. Fersht, The changing nature of the protein folding transition state: implications for the shape of the free-energy profile for folding, J. Mol. Biol. 277 (1998) 933–943.
[13] C. Shastry, H. Roder, Evidence for barrier-limited protein folding kinetics on the microsecond time scale, Nat. Struct. Biol. 5 (1998) 385–392.
[14] J. Rumbley, L. Hoang, L. Mayne, S.W. Englander, An amino acid code for protein folding, Proc. Natl. Acad. Sci. USA 98 (2001) 105–112.
[15] T. Kiefhaber, F.X. Schmid, Kinetic coupling between protein folding and prolyl isomerization. II. Folding of ribonuclease A and ribonuclease T1, J. Mol. Biol. 224 (1992) 231–240.
[16] S.-R. Yeh, S. Takahashi, B. Fan, D.L. Rousseau, Ligand exchange during cytochrome *c* folding, Nat. Struct. Biol. 4 (1997) 51–56.
[17] T.R. Sosnick, L. Mayne, R. Hiller, S.W. Englander, The barriers in protein folding, Nat. Struct. Biol. 1 (1994) 149–156.
[18] W. Colon, L.P. Wakem, F. Sherman, H. Roder, Identification of the predominant non-native histidine ligand in unfolded cytochrome *c*, Biochemistry 36 (1997) 12535–12541.
[19] W. Colon, G.A. Elove, L.P. Wakem, F. Sherman, H. Roder, Side chain packing of the N- and C-terminal helices plays a critical role in the kinetics of cytochrome *c* folding, Biochemistry 35 (1996) 5538–5549.
[20] T.R. Sosnick, L. Mayne, S.W. Englander, Molecular collapse: the rate-limiting step in two-state cytochrome *c* folding, Proteins 24 (1996) 413–426.
[21] T.R. Sosnick, M.D. Shtilerman, L. Mayne, S.W. Englander, Ultrafast signals in protein folding and the polypeptide contracted state, Proc. Natl Acad. Sci. USA 94 (1997) 8545–8550.
[22] C.K. Chan, Y. Hu, S. Takahashi, D.L. Rousseau, W.A. Eaton, J. Hofrichter, Submillisecond protein folding kinetics studied by ultrarapid mixing, Proc. Natl. Acad. Sci. USA 94 (1997) 1779–1784.
[23] O.D. Monera, C.M. Kay, R.S. Hodges, Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions, Protein Sci. 3 (1994) 1984–1991.
[24] R. Gupta, S. Yadav, F. Ahmad, Protein stability: urea-induced versus guanidine-induced unfolding of metmyoglobin, Biochemistry 35 (1996) 11925–11930.
[25] R.A. Staniforth, M.G. Bigotti, F. Cutruzzolà, C. Travaglini-Allocatelli, M. Brunori, Unfolding of apomyoglobin from *Aplysia limacina*: the effect of salt and pH on the cooperativity of folding, J. Mol. Biol. 275 (1998) 133–148.
[26] M.M. Santoro, D.W. Bolen, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, Biochemistry 27 (1988) 8063–8068.
[27] C.N. Pace, Determination and analysis of urea and guanidine hydrochloride denaturation curves, Methods Enzymol. 131 (1986) 266–280.
[28] S. Gianni, C. Travaglini-Allocatelli, F. Cutruzzolà, M.G. Bigotti, M. Brunori, Snapshots of protein folding. A study on the multiple transition state pathway of cytochrome c_{551} from *Pseudomonas aeruginosa*, J. Mol. Biol. 309 (2001) 1177–1187.
[29] M. Silow, M. Oliveberg, Transient aggregates in protein folding are easily mistaken for folding intermediates, Proc. Natl. Acad. Sci. USA 94 (1997) 6084–6086.
[30] M.J. Parker, J. Spencer, A.R. Clarke, An integrated kinetic analysis of intermediates and transition states in protein folding reactions, J. Mol. Biol. 253 (1995) 771–786.
[31] A. Matouschek, A.R. Fersht, Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding, Proc. Natl. Acad. Sci. USA 90 (1993) 7814–7818.
[32] C. Travaglini-Allocatelli, F. Cutruzzolà, M.G. Bigotti, R.A. Staniforth, M. Brunori, Folding mechanism of *Pseudomonas aeruginosa* cytochrome c_{551}: role of electrostatic interactions on the hydrophobic collapse and transition state properties, J. Mol. Biol. 289 (1999) 1459–1467.

[33] S. Gianni, M. Brunori, C. Travaglini-Allocatelli, Refolding kinetics of cytochrome c_{551} reveals a mechanistic difference between urea and guanidine, Protein Sci. 10 (2001) 1685–1688.

[34] C.M. Johnson, A.R. Fersht, Protein stability as a function of denaturant concentration: the thermal stability of barnase in the presence of urea, Biochemistry 34 (1995) 6795–6804.

[35] O.B. Ptitsyn, Protein folding and protein evolution: common folding nucleus in different subfamilies of *c*-type cytochromes?, J. Mol. Biol. 278 (1998) 655–666.

[36] D.J. Detlefsen, V. Thanabal, V.L. Pecoraro, G. Wagner, Solution structure of Fe(II) cytochrome c_{551} from *Pseudomonas aeruginosa* as determined by two-dimensional ^{1}H NMR, Biochemistry 30 (1991) 9040–9046.

[37] F. Cutruzzolà, I. Ciabatti, G. Rolli, et al., Expression and characterization of *Pseudomonas aeruginosa* cytochrome *c*-551 and two site-directed mutants: role of tryptophan 56 in the modulation of redox properties, Biochem. J. 322 (1997) 35–42.

[38] J.M. Sauder, N.E. MacKenzie, H. Roder, Kinetic mechanism of folding and unfolding of *Rhodobacter capsulatus* cytochrome c_2, Biochemistry 35 (1996) 16852–16862.

[39] G.R. Moore, G.W. Pettigrew, Cytochromes *c*: Evolutionary, Structural and Physicochemical Aspects, Springer-Verlag, Berlin, 1990.

[40] C.M. Jones, E.R. Henry, Y. Hu, et al., Fast events in protein folding initiated by nanosecond laser photolysis, Proc. Natl. Acad. Sci. USA 90 (1993) 11860–11864.

[41] S. Takahashi, S.R. Yeh, T.K. Das, C.K. Chan, D.S. Gottfried, D.L. Rousseau, Folding of cytochrome *c* initiated by submillisecond mixing, Nat. Struct. Biol. 4 (1997) 44–50.

[42] A. Arcovito, S. Gianni, M. Brunori, C. Travaglini-Allocatelli, A. Bellelli, Fast coordination changes in cytochrome *c* do not necessarily imply folding, J. Biol. Chem. 276 (2001) 41073–41078.

[43] D.K. White, J.B. Cannon, T.G. Traylor, A kinetic model for R- and T-state hemoglobin. Flash photolysis of heme-imidazole-carbon monoxide mixtures, J. Am. Chem. Soc. 101 (1979) 2443–2454.

[44] E. Chen, M.J. Wood, A.L. Fink, D.S. Kliger, Time-resolved circular dichroism studies of protein folding intermediates of cytochrome *c*, Biochemistry 37 (1998) 5589–5598.

[45] M.C. Shastry, S.D. Luck, H. Roder, A continuous-flow capillary mixing method to monitor reactions on the microsecond time scale, Biophys. J. 74 (1998) 2714–2721.

[46] T. Kiefhaber, Kinetic traps in lysozyme folding, Proc. Natl. Acad. Sci. USA 92 (1995) 9029–9033.

[47] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, Nat. Struct. Biol. 4 (1997) 10–19.

[48] K. Gunasekaran, S.J. Eyles, A.T. Hagler, L.M. Gierasch, Keeping it in the family: folding studies of related proteins, Curr. Opin. Struct. Biol. 11 (2001) 83–93.

[49] J. Guidry, P. Wittung-Stafshede, Cytochrome *c*(553), a small heme protein that lacks misligation in its unfolded state, folds with rapid two-state kinetics, J. Mol. Biol. 301 (2000) 769–773.

[50] N. Ferguson, A.P. Capaldi, R. James, C. Kleanthous, S.E. Radford, Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9, J. Mol. Biol. 286 (1999) 1597–1608.

[51] R. Guerois, L. Serrano, The SH3-fold family: experimental evidence and prediction of variations in the folding pathways, J. Mol. Biol. 304 (2000) 967–982.

[52] C. Nishimura, S. Prytulla, H.J. Dyson, P.E. Wright, Conservation of folding pathways in evolutionarily distant globin sequences, Nat. Struct. Biol. 7 (2000) 679–686.

[53] M.G. Bigotti, C. Travaglini-Allocatelli, R.A. Staniforth, M. Arese, F. Cutruzzolà, M. Brunori, Equilibrium unfolding of a small bacterial cytochrome, cytochrome c_{551} from *Pseudomonas aeruginosa*, FEBS Lett. 425 (1998) 385–390.