The nature of the initial step in the conformational folding of disulphide-intact ribonuclease A

Walid A. Houry, David M. Rothwarf and Harold A. Scheraga

Here we investigate conformational folding reaction of disulphide-intact ribonuclease A in the absence of the complicating effects due to non-native interactions (such as *cis/trans* proline isomerization) in the unfolded state. The conformational folding process is found to be intrinsically very fast occurring on the milliseconds time scale. The kinetic data indicate that the conformational folding of ribonuclease A proceeds through the formation of a hydrophobically collapsed intermediate with properties similar to those of equilibrium molten-globules. Furthermore, the data suggest that the rate-limiting transition states on the unfolding and refolding pathways are substantially different with the refolding transition state having non-native-like properties.

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301, USA

Correspondence should be addressed to H.A.S.

Protein folding proceeds from the unfolded state to the native state through a series of intermediates[1]. Those intermediates which form early on the refolding pathway provide insight into the nature of the interactions that initiate protein folding. Unfortunately, most of the folding processes studied are complicated by the presence of non-native peptide bond conformations (or non-native interactions) in the unfolded state[2]. These non-native conformations lead to the formation of kinetic barriers on refolding and, consequently, lead to the formation of intermediates that are not intrinsic to the folding process[3]. Hence, they cause the refolding kinetics to be slow and heterogeneous. In order to obtain a clear picture of the nature of the initial step in protein folding, the 'pure' folding process, which is not complicated by the presence of such non-native peptide bonds, has to be investigated. We designate this 'pure' folding reaction as the conformational protein folding reaction.

It is thought that the initiation of protein folding arises predominantly because of hydrophobic interactions[4–6]. Such interactions result in a hydrophobic collapse of the unfolded state leading to the formation of what is termed the molten globule[7–9]. The molten globule is generally defined as a partially folded compact state with a sizable hydrophobic core which has a high content of secondary structure but few tertiary contacts. It is characterized by a high degree of fluctuation compared to the native state. For many proteins, for example apomyoglobin[10], α-lactalbumin[11], β-lactamase[12] and carbonic anhydrase[13], the molten globule state has been observed as an equilibrium state. From studies of refolding kinetics, it has been suggested that the earliest detectable kinetic intermediates are identical to the equilibrium molten globule intermediates[14–16]. Nevertheless, there has been no direct experimental evidence from kinetic studies that clearly demonstrates that the initial step in protein folding is driven by hydrophobic interactions. In this paper, we provide such evidence from the conformational folding of disulphide-intact ribonuclease A (RNase A).

RNase A has two *cis* X–Pro peptide bonds in the native state[17]. In the unfolded state, these peptide bonds isomerize to adopt the *trans* as well as the *cis* conformation[18]. In a previous study[19], we have shown that these isomerizations result in the formation of at least four unfolded species. Only one of the unfolded species has the native *cis* X–Pro peptide bonds. We call this unfolded species U_{vf} - the very-fast folding species. The other unfolded species, which contain non-native X–Pro peptide bonds, are called fast and slow folding species. Unlike the fast and slow folding species, the refolding reaction of U_{vf} is not complicated by the presence of non-native X–Pro peptide bonds; hence, its refolding represents the conformational folding reaction of the disulphide-intact protein as defined above.

The possible presence of native-like structure in U_{vf} was investigated in our earlier study[19]. These studies indicated that U_{vf} is a completely unfolded species, and not partially unfolded, under the conditions and temperatures used. The same conditions are employed in the current investigation. Therefore, no significant native-like structure is present in U_{vf}.

For RNase A, there is no evidence that an equilibrium molten globule state can be populated by varying pH,

temperature, urea, or guanidine hydrochloride (GdnHCl) concentration. In this paper, we show that, when U_{vf} is refolded at low denaturant concentration, RNase A undergoes a hydrophobic collapse leading to the formation of an intermediate. The observed pK_a shifts, the extent of burial of the solvent-exposed suface area as a function of temperature, and kinetic and thermodynamic parameters are all consistent with hydrophobic collapse. Furthermore, the kinetic data indicate that the rate-limiting transition state on the refolding pathway is not identical to that on the unfolding pathway.

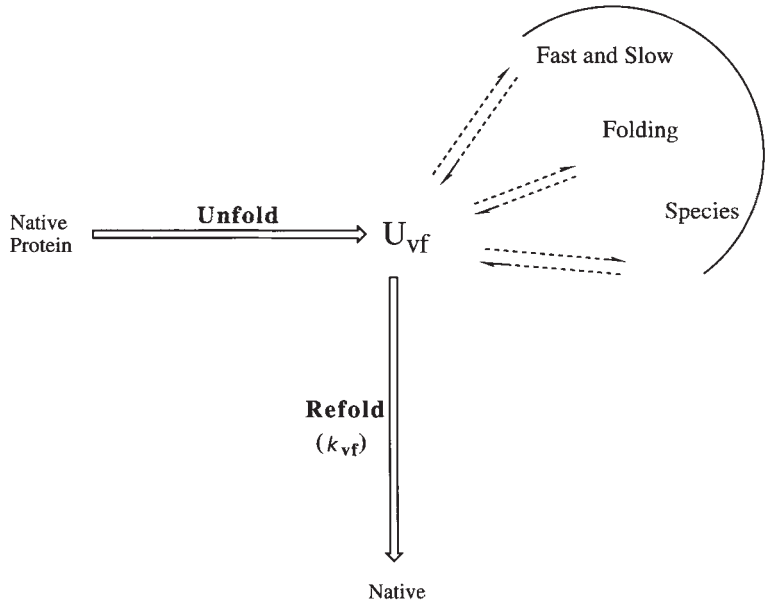


Fig. 1 A schematic representation of the double-jump experiments. In the first jump, the native protein is placed under unfolding conditions for a given time called the delay time. That delay time is long enough to allow for the formation of the very-fast folding species U_{vf}, but short enough not to allow for the formation of the fast and slow folding species. The fast and slow folding species are formed due to *cis/trans* proline isomerization reactions, and, hence, contain non-native X–Pro peptide bonds. It should be pointed out that each unfolded species folds to the native state along its own distinct pathway as discussed in detail by Houry *et al.*[19]. After the set delay time, the second jump is carried out. In this jump, the protein is placed under refolding conditions resulting in the formation of the native state from U_{vf}. The refolding conditions are varied, while the unfolding jump is always carried out under the same condition. Varying the conditions in the refolding jump has no effect on the formation of U_{vf} from the native protein in the unfolding jump. The rate constant for refolding from U_{vf} to the native protein (k_{vf}) is the parameter under investigation.

Conformational folding

The refolding reaction of U_{vf} was studied by using the same double-jump technique that we employed previously[19]. The folded protein was unfolded at high GdnHCl concentration and low pH for a delay time long enough to populate the U_{vf} species, but short enough not to form any of the fast or slow folding species (Fig. 1), so that, after this delay time, U_{vf} constituted > 99 % of the unfolded state. U_{vf} was then refolded under a variety of conditions. The GdnHCl- and pH-dependence of the rate constant for refolding of U_{vf} (k_{vf}) was studied at 5 °C and 15 °C. Seven sets of data were obtained (Fig. 2). The reaction under investigation is the refolding of the single species U_{vf}, which represents the conformationally unfolded state of RNase A containing the native X–Pro peptide bonds, to a well defined single species, namely the native state.

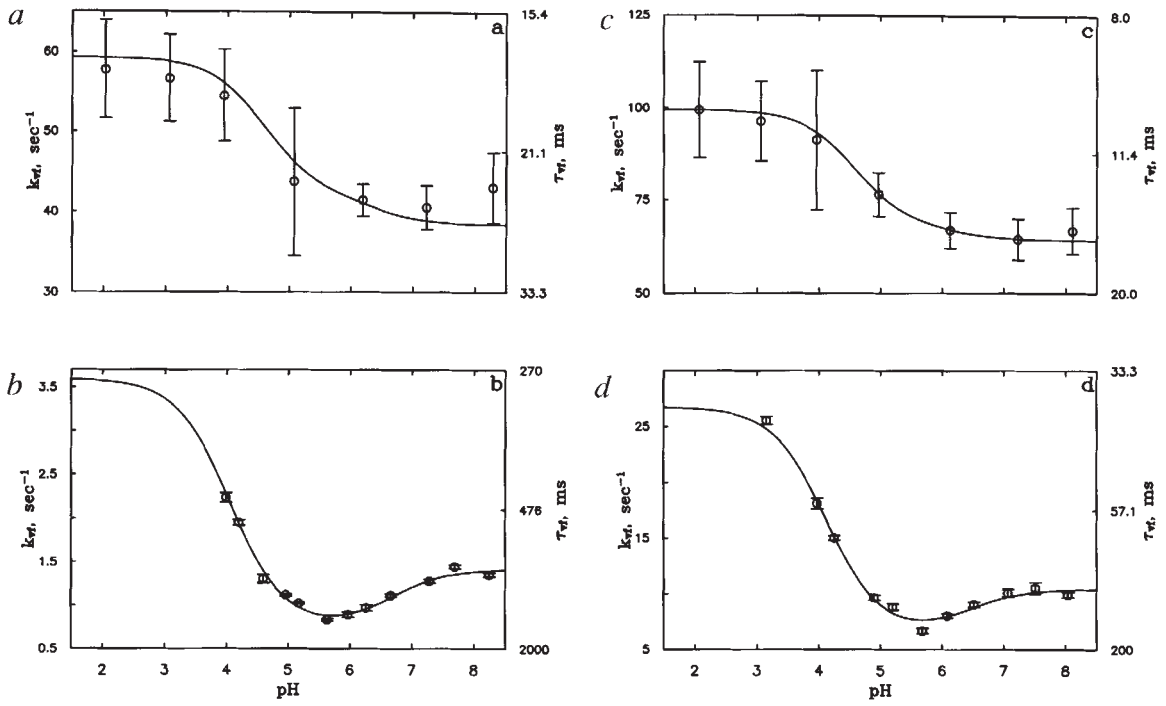
From the pH dependence data (Fig. 2*a–d*), the refolding rate constant is observed to decrease as the pH is raised from 2.0 to 6.0. This is an unexpected result, since, generally, the refolding rate is expected to be more rapid under conditions where the native protein is more stable. For RNase A, it is well known that the native state is more stable at neutral pH than at low pH (ref. 20). If the transition state along the refolding pathway of U_{vf} is native-like, then we would expect the rate constant to increase, rather than decrease, as the pH is raised. Futhermore, the variation of the rate constants with pH is much greater at the high GdnHCl concentration than at the low GdnHCl concentration (compare Fig. 2*a* with *b*, and Fig. 2*c* with *d*). This cannot be explained in terms of a simple two-state transition (see below). Hence, the decrease in the rate constant with increasing pH, and the effect of GdnHCl on the pH dependence of the rate constant, point to the presence of an intermediate with some unusual characteristics. Further support for the presence of the intermediate comes from the non-linearity of the GdnHCl-dependence data (Fig. 2*e–g*). Although the GdnHCl dependence data can be fitted by assuming a linear dependence of $\ln(k_{vf})$ on GdnHCl concentration, the curvature in the data is evident and reproducible under the three different conditions employed.

Based on the above arguments, the simplest kinetic model must take into account the presence of (at least) three species: the unfolded state U_{vf}, the intermediate I_Φ, and the native state N. We assume that, when the unfolded state U_{vf} is placed under folding conditions, it undergoes a rapid pre-equilibrium to form the intermediate. This intermediate then proceeds to the native state passing through the rate-limiting transition state (T.S.). Hence, the following sequential model is proposed:

$$U_{vf} \overset{K}{\rightleftharpoons} I_\Phi \overset{k}{\rightarrow} [T.S.] \rightarrow N \qquad (1)$$

The equilibrium constant is defined as $K = [I_\Phi]/[U_{vf}]$. The variation of the equilibrium constant with temperature, pH and GdnHCl concentration reflects the change in the relative stabilities of I_Φ and U_{vf} as a function of these three parameters. Similarly, the rate constant (k) reflects the behaviour of the T.S. relative to the intermediate.

Experimentally, the observed rate is the apparent rate of formation of N from I_Φ and U_{vf}. It can be ex-



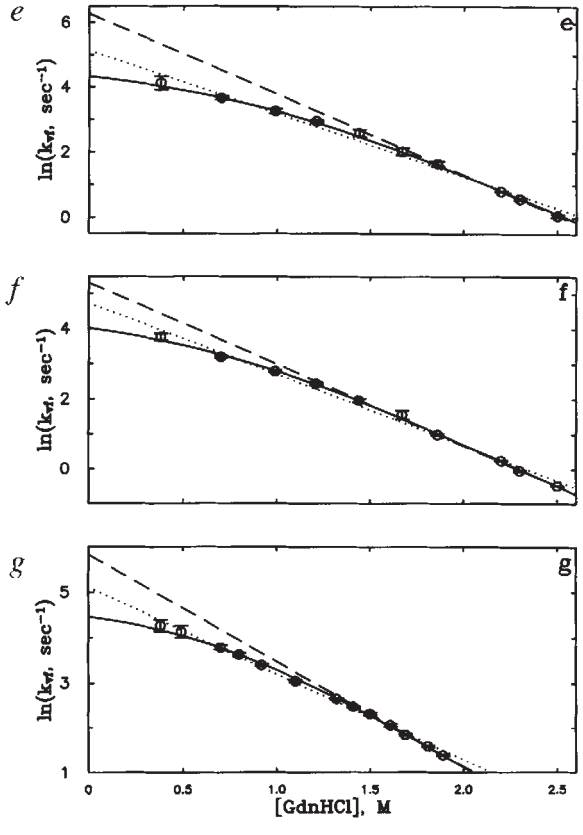


Fig. 2 *a–d*, give the pH dependence of k_{vf}; *e–g*, give the GdnHCl dependence of ln(k_{vf}) as obtained from the double-jump experiments. The error symbols give the experimental errors at the 95 % confidence limit. Each point is an average of 6–10 measurements. The conditions for the refolding step in each figure are as follows: *a*, 0.38 M GdnHCl, 5 °C; *b*, 2.20 M GdnHCl, 5 °C; *c*, 0.38 M GdnHCl, 15 °C; *d*, 1.50 M GdnHCl, 15 °C; *e*, pH 4.0, 5 °C; *f*, pH 7.2, 5 °C and *g*, pH 7.1, 15 °C. The solid lines are the theoretical fit to the experimental points based on the model of equation (1) and the results of Table 1. The dotted lines in *e–g* are the best linear fit to the experimental points with a correlation coefficient of better than 0.99. The dashed lines in these figures are obtained from the linear extrapolation of ln(k_{vf}) from the last four points at the high GdnHCl concentration to zero molar concentration.

pressed as follows:

$$\frac{d[N]}{dt} = k_{vf}([I_\Phi] + [U_{vf}]) \qquad (2)$$

where k_{vf} is the experimentally observed refolding rate constant. It can be shown that k_{vf} is related to k and K by the following equation:

$$k_{vf} = \frac{k\,K}{(1+K)} \qquad (3)$$

It should be pointed out that the model shown in equation (1) is kinetically indistinguishable from two other models: a parallel model, in which U_{vf} and I_Φ pre-equilibrate rapidly and then refold to the native state along separate pathways but pass through the same transition state, and a sink model whereby U_{vf} and I_Φ pre-equilibrate rapidly, but only U_{vf} can fold to the native state. The common feature of the three models is that I_Φ forms from U_{vf} in a rapid preequilibrium reaction, and that the rate-limiting transition state directly precedes the approach to the native state. The parameters obtained from fitting any one of these models can be used to fit the other two models. Therefore, in the interest of clarity, we shall use the model of equation (1) as our working example.

Table 1 Results of the fit to the kinetic model of equation (1)

	Temperature	
	5 °C	15 °C
pK_a (carboxylic acid side chain)		
Unfolded	4.01(0.07)	same
Intermediate	4.94(0.25)	same
Transition state	4.74(0.06)	same
pK_a (histidine)		
Unfolded	6.69(0.10)	6.46(0.12)
Intermediate	6.07(0.27)	5.91(0.28)
Transition state	6.40(0.11)	6.21(0.12)
'm' values (kcal mol^{-1} M^{-1})		
T. S. – U_{vf}	1.37(0.04)	1.42(0.04)
T. S. – I_Φ ($m^\ddagger$)	0.25(0.11)	0.08(0.12)
I_Φ – U_{vf} ($m_{pre\text{-}eq}$)	1.12(0.10)	1.34(0.11)
K^Θ	11.2(2.1)	11.6(1.9)
k^Θ (sec^{-1})	84(24)	126(30)

The kinetic data were fit to the model of equation (1) as described in the Methods. The numbers in parentheses give the standard deviations of the fitted parameters. Only a carboxylic acid residue and a histidine residue are considered to undergo pK_a shifts during the refolding of U_{vf}.

GdnHCl and pH dependence

In order to derive the GdnHCl- and pH-dependence of the rate constant for refolding of U_{vf} (k_{vf}), the GdnHCl- and pH-dependence of K and k in equation (3) must be determined. Several models have been used to describe the GdnHCl-dependence of the unfolding/refolding equilibrium process[21–23]. In the most commonly used model, the logarithm of the equilibrium constant, at a given temperature and pH, depends linearly on the GdnHCl concentration. The pH dependence of the equilibrium constant, at a given GdnHCl concentration and temperature, can be described by a simple binding model[24]. In this model, the value of the equilibrium constant changes because the apparent pK_a for a given charged group on the amino acid side chain changes as the protein goes from the unfolded state to the intermediate state. Hence, the GdnHCl- and pH-dependence of K at a given temperature can be expressed as follows:

$$K = K^\Theta e^{-m_{pre\text{-}eq}[\mathrm{GdnHCl}]/RT} \prod_j \frac{(10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a(aa_j, I_\Phi)})}{(10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a(aa_j, U_{vf})})} \quad (4)$$

where R is the gas constant and T is the temperature. K^Θ is the equilibrium constant at 0 M GdnHCl and infinite $[H^+]$. $m_{pre\text{-}eq}$ is a positive constant proportional to the surface area available for GdnHCl binding (the solvent exposed surface area) in U_{vf} that is buried in I_Φ. The product is evaluated over all the amino acid residues (aa_j) whose side chains have a pK_a in I_Φ different from that in U_{vf}.

In treating the rate constant k, we use transition state theory which assumes the presence of an equilibrium between the transition state and any stable intermediates preceding it[25]. This allows us to treat the rate constant in the same way as we treat the equilibrium constant. Therefore, we can write for the GdnHCl- and the pH-dependence of k at a given temperature:

$$k = k^\Theta e^{-m^\ddagger[\mathrm{GdnHCl}]/RT} \prod_j \frac{(10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a(aa_j, \mathrm{T.S.})})}{(10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a(aa_j, I_\Phi)})} \quad (5)$$

where k^Θ is the rate constant at 0 M GdnHCl and infinite $[H^+]$. m^* is a positive constant proportional to the solvent exposed surface area that becomes buried as the protein folds from the intermediate to the transition state.

In writing equations (4) and (5), we assumed that $m_{pre\text{-}eq}$ and m^* have no significant dependence on pH, and that the pK_as have no significant dependence on GdnHCl concentration. These assumptions are justified experimentally. Pace et al.[20] have shown that the m values of RNase A (between the unfolded and native state) at different pHs, ranging from pH 3.0 to 9.9, are the same within experimental error. Furthermore, Donovan *et al.*[26] have shown that the presence of GdnHCl lowers or raises the pK_a of model compounds by only 0.1–0.2 pK_a units which is within our experimental error.

At this point, we use the above equations to clarify why a simple two-state mechanism of the form:

$$U_{vf} \xrightarrow{r} [\mathrm{T.S.}] \rightarrow N \quad (6)$$

cannot be used to fit the experimental data. If the two state model is sufficient to explain the experimental data, then the GdnHCl and pH dependence of the experimentally observed rate constant (r) would be given simply by an equation similar to that of equation (5). Furthermore, to make the argument more general, it is sufficient to write *func*(GdnHCl) to describe the GdnHCl dependence of the rate:

$$r = r^\theta \, func\,(\mathrm{GdnHCl}) \prod_j \frac{(10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a(aa_j, \mathrm{T.S.})})}{(10^{-\mathrm{pH}} + 10^{-\mathrm{p}K_a(aa_j, U_{vf})})} \quad (7)$$

where r^θ is defined as the rate constant at 0 M GdnHCl and infinite $[H^+]$.

If the model of eq. (6) applies, it is clear from eq. (7) that, for a given temperature and at any given GdnHCl concentration, the ratio of the value of r at some pH to the value of r at another pH should be a constant independent of the GdnHCl concentration. Experimentally, we observe, for example, that at 15 °C and 0.38 M GdnHCl (Fig. 2*c*), k_{vf}(pH 3)/k_{vf}(pH 6) = (96.4 sec^{-1}/66.9 sec^{-1}) = 1.4, while at 15 °C and 1.5 M GdnHCl (Fig. 2*d*), k_{vf}(pH 3)/k_{vf}(pH 6) = (25.5 sec^{-1}/8.0 sec^{-1}) = 3.2. Hence, this ratio does depend on the GdnHCl concentration, and, consequently, the simple mechanism of eq. (6) cannot be used to fit the experimental data. The next simplest mechanism which can fit the data is that of equation (1).

An intermediate is needed to explain the pH dependence of the experimental data. Furthermore, by using a linear model to describe the dependence of ln(k) and ln(K) on [GdnHCl] (which is the simplest model available), the curvature in the GdnHCl dependence data of ln(k_{vf}) (Fig. 2*e, f, g*) arises as a normal consequence of the presence of the intermediate.

Hydrophobic collapse: evidence from kinetic analysis

Based on the above presentation, the GdnHCl- and pH-dependence of k_{vf} can be fitted using equations (3), (4)

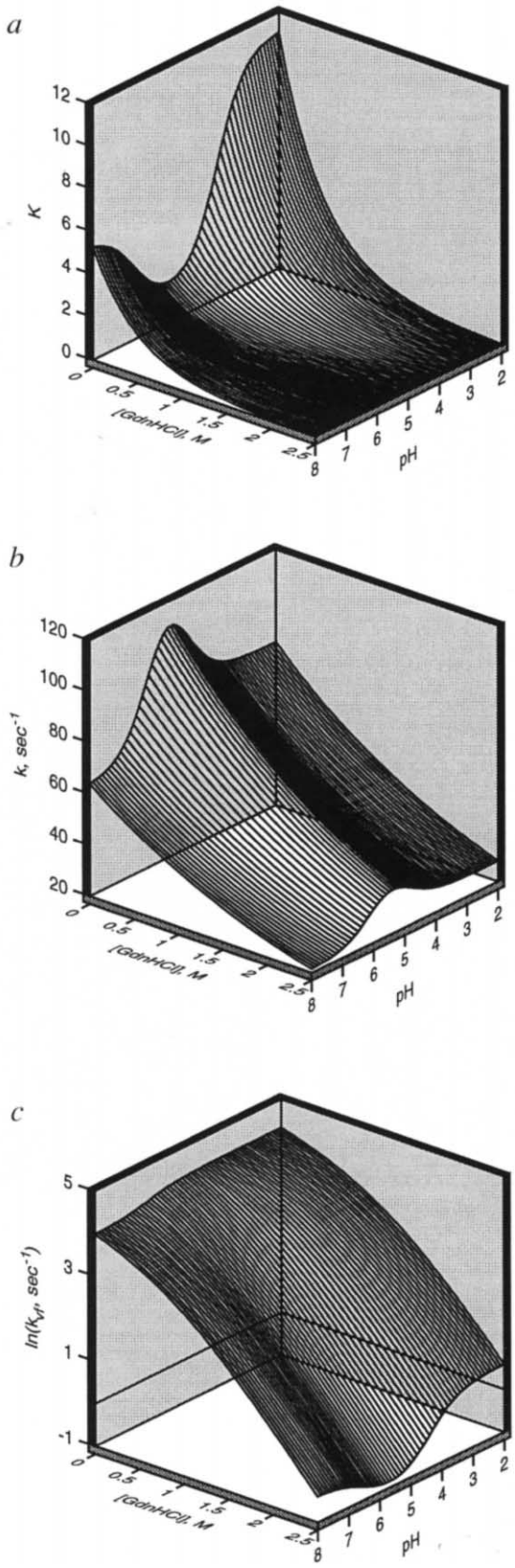


Fig. 3 The GdnHCl- and pH-dependence of *a*, *K*; *b*, *k*; and *c*, $\ln(k_{vf})$ at 5 °C. The curves are based on the model of equation (1) and the results given in Table 1. The curves at 15 °C (not shown) have similar shapes.

and (5). The pH-dependence data in Fig. 2*a–d* indicate that there are at least two residues which contribute to the pK_a shifts, one in the pH region of 4–5 which suggests that a carboxylic acid residue (Asp or Glu) is involved, and another in the pH region of 6–7 which suggests that a histidine residue is involved.

The results of the fit to the experimental data at 5 °C and 15 °C are given in Table 1, and the theoretical curves are shown as solid lines in Fig. 2. Good agreement is obtained between the experimental kinetic data and the theoretical curves. The GdnHCl- and pH-dependence of *K*, *k*, and $\ln(k_{vf})$ are given in Fig. 3.

As the protein proceeds from its unfolded state to the intermediate state, the pK_a of the side chain of the carboxylic acid residue (Asp or Glu) increases from 4.0 to 4.9, while that of the histidine residue decreases from 6.7 to 6.1 at 5 °C or from 6.5 to 5.9 at 15 °C. Both residues are expected to be solvent exposed (that is exposed to a polar environment) in the unfolded state. The pK_a shifts can then be explained, simply, if both residues become buried in a nonpolar environment in the intermediate state. In such an environment, a charged group prefers to remain in its neutral form. The pK_a values obtained for the unfolded state are in good agreement with those reported by Nozaki & Tanford[27] for GdnHCl-unfolded RNase A. The change in the pK_a of the histidine residue, from 6.7 to 6.5 in U_{vf} and from 6.1 to 5.9 in I_Φ, when the temperature is raised from 5 °C to 15 °C, is in the expected direction and has the expected magnitude[28]. Thus, the pK_as are consistent with the data reported in the literature.

The temperature dependence of the value of $m_{pre\text{-}eq}$ (between I_Φ and U_{vf}), shows that the solvent exposed surface area that becomes buried when the protein proceeds from the unfolded state to the intermediate increases with temperature (1.12 at 5 °C and 1.34 at 15 °C). This result suggests that the intermediate has more buried surface area at 15 °C than at 5 °C. This is consistent with hydrophobic interactions being the driving force for the formation of the intermediate[29,30].

It can be seen from Fig. 3*a* that the equilibrium constant is higher at low pH than at neutral pH. This indicates that, at low pH, the intermediate is more favoured than the unfolded state. This is a manifestation of the fact that the pK_a of the carboxylic acid residue is higher in the intermediate than in the unfolded state and, therefore, binds protons more strongly making it more stable than the unfolded state at low pH (eq. 4). This behaviour of the intermediate is a normal consequence of hydrophobic collapse, and hence is one of the defining characteristics of hydrophobically collapsed states.

Hydrophobic collapse: evidence from thermodynamic analysis

The thermodynamic parameters for the U_{vf} to I_Φ process can be obtained from the equilibrium constant and its temperature dependence. The GdnHCl- and pH-dependence of the apparent $\Delta G°$(10 °C), $\Delta H°$(10 °C), and $\Delta S°$(10 °C) are plotted in Fig. 4. Both the polar and nonpolar interactions contribute to these thermodynamic parameters.

As defined by Némethy & Scheraga[31], a hydrophobic interaction is considered to be formed if two or more

nonpolar side chains come in contact with each other, thereby decreasing their interaction with the surrounding water. Such a process is characterized by a positive enthalpy change and a positive entropy change as determined from theory[31,32] and experiment[32,33]. The term 'hydrophobic collapse' is defined as a process in which the protein collapses to form an intermediate that is stabilized predominantly by hydrophobic interactions. At 0 M GdnHCl, ΔG^{o}(10 °C) is negative, of the order of -1.4 to -0.4 kcal mol^{-1}, while ΔH^{o}(10 °C) and ΔS^{o}(10 °C), below about pH 6.5, are both positive (~1.0 kcal mol^{-1} and ~7 cal mol^{-1} K^{-1} respectively). The formation of the intermediate is, therefore, favoured at low GdnHCl concentrations; it is stabilized by entropic rather than enthalpic effects.

From equilibrium unfolding GdnHCl-dependence studies, Pace and coworkers[20] reported a value of m between the unfolded and the native state of 3.0 kcal mol^{-1} M^{-1}. Therefore, 37-45 % of the total area which is buried in the native state becomes buried in the intermediate; consequently, the conformational entropy of the U_{vf} to I_{Φ} process should be large and negative. The positive apparent entropy observed at 10 °C indicates that hydrophobic interactions are the dominant interactions in stabilizing the intermediate, and that extensive side-chain ordering has not yet occurred in I_{Φ}.

As the GdnHCl concentration increases, ΔG^{o}(10 °C) increases while both ΔH^{o}(10 °C) and ΔS^{o}(10 °C) decrease (Fig. 4). Such behaviour is expected for hydrophobic interactions since the presence of GdnHCl is known to weaken these interactions[34]. This provides further support for the dominance of hydrophobic interactions in stabilizing I_{Φ}. It is worth noting that the variation with GdnHCl concentration of the enthalpy and entropy changes for the U to N transition in RNase A and other proteins[35] is in the opposite direction of that observed for ΔH^{o}(10 °C) and ΔS^{o}(10 °C) for the U_{vf} to I_{Φ} transition.

The thermodynamic parameters obtained for the formation of I_{Φ} are very similar to those obtained for the formation of the equilibrium molten globule states observed in several proteins[36,37]. This suggests that I_{Φ} is a molten globule-like intermediate; however, further characterization of the secondary structure of this intermediate is required to confirm this suggestion.

The transition state

There has been much discussion about the nature of the transition state along the refolding pathway. It is usually argued that the transition state is native-like, and that the transition state is the same on unfolding and refolding[38–40]. Since the refolding kinetics are usually much more complicated than the unfolding kinetics, most of the experimental information available about the transition state comes from unfolding studies rather than refolding studies. Here, we are able to obtain information about the transition state directly from refolding studies.

The values of m* (0.25 and 0.08, Table 1), which represent the GdnHCl accessible surface area available in the intermediate that becomes buried in the transition state, are small positive numbers (8 % change at 5 °C

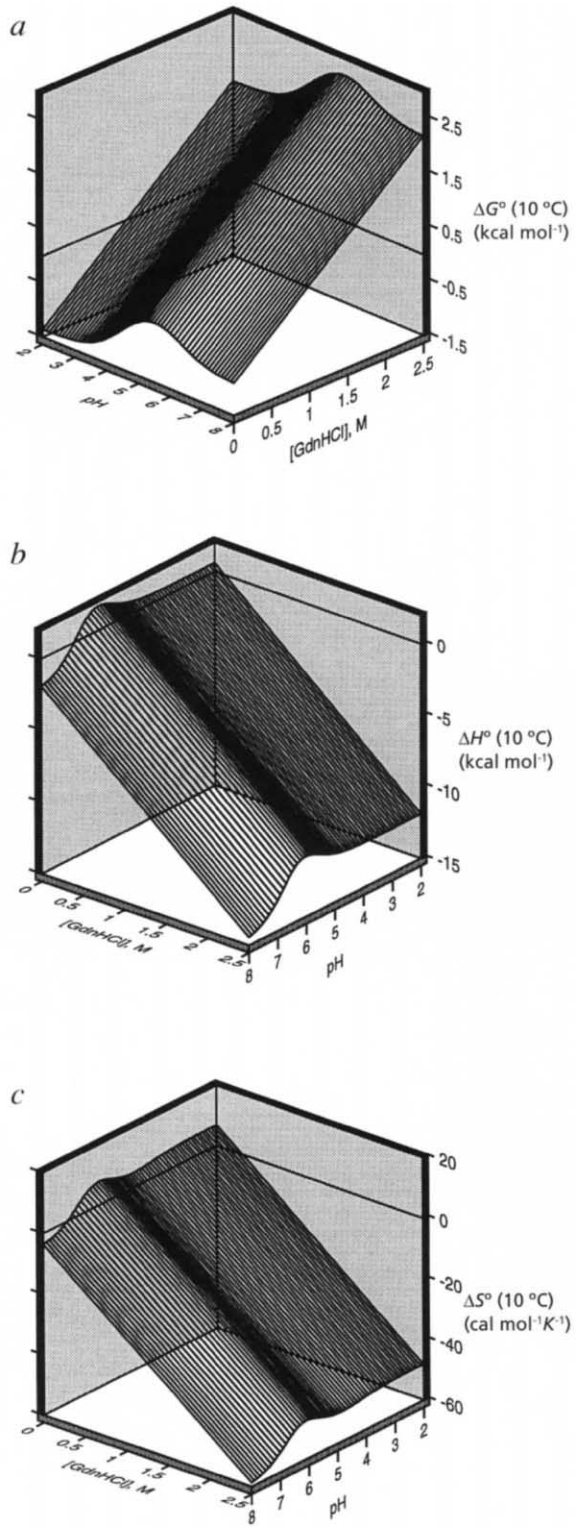


Fig. 4 The thermodynamic parameters for the pre-equilibrium reaction as a function of pH and GdnHCl concentration. *a*, The standard free energy change, *b*, enthalpy, and *c*, entropy, all at 10 °C.

and 3 % change at 15 °C when compared to the total - using the value of 3.0 kcal mol^{-1} M^{-1} of Pace *et al.*[20]). Furthermore, the difference in the pK_as between I_Φ and the T.S. are small (Table 1). This would indicate that there is no substantial difference in solvent exposure between I_Φ and the T.S. .

From the rate constants, we can obtain the activation free energy [$\Delta G^*(T)$], enthalpy [$\Delta H^*(T)$] and entropy [$\Delta S^*(T)$] as a function of pH and GdnHCl concentration (Fig. 5). The activation enthalpy, at 0 M GdnHCl and 10 °C, is positive and has a value of 5–7 kcal mol^{-1}. This is a small number if compared, for example, with the activation enthalpy of proline isomerization of about 20 kcal mol^{-1} (refs 2,19,41). In contrast to the behaviour of ΔH^o(10 °C) and ΔS^o(10 °C) (Fig. 4), both the activation enthalpy and entropy increase with GdnHCl concentration. This indicates that the transition state has more polar interactions (for example hydrogen-bonding or electrostatic) than I_Φ, and that the transition state is formed because of the appearance of ordered structure within the compact intermediate.

The solvent accessible surface area buried in the transition state is only about 47 % of the total area which becomes buried in the native state [(1.37 or 1.42)/3.0, Table 1]. On the other hand, the transition state on the unfolding pathway has about 90 % of its solvent accessible surface area buried at 22 °C and pH 5.8, and about 85 % of that area buried at 15 °C and pH 4.0 (from data of Lin *et al.*[42] and D.M.R. & H.A.S., unpublished results). This would argue against the hypothesis that the transition states on the refolding and unfolding pathways are the same. Furthermore, if the folding T.S. had 90 % of its solvent accessible surface area buried, then I_Φ would have ~85 % of that area buried. Since the difference in absorbance between the unfolded and the native state arises from tyrosine burial, we would expect the extinction coefficient of I_Φ to differ from that of the unfolded state, causing the total refolding absorbance amplitude to change as the population of I_Φ changes with pH and GdnHCl concentration. This is not observed experimentally which suggests that I_Φ and the T.S. do not have a large buried surface area.

The carboxylic acid pK_as reported for the native protein[43] are much lower than those observed for the transition state. This suggests that the tertiary interactions which cause the pK_as of the carboxylic acid residues to be low in the native state are not present in the T.S. This is another indication that the transition state on the refolding pathway might not be as native-like as previously thought.

The effect of the presence of the intermediate on the refolding rate

The intermediate I_Φ is highly populated at low pH and GdnHCl concentrations (> 90 % of the preequilibrium mixture, Fig. 3*a*). At high GdnHCl concentrations, I_Φ constitutes only about 5 % of the preequilibrium mixture. The hydrophobic forces dominate at low GdnHCl concentrations causing a collapse of the unfolded state to form the intermediate. But, when the GdnHCl concentration is high enough, such a collapse is prevented

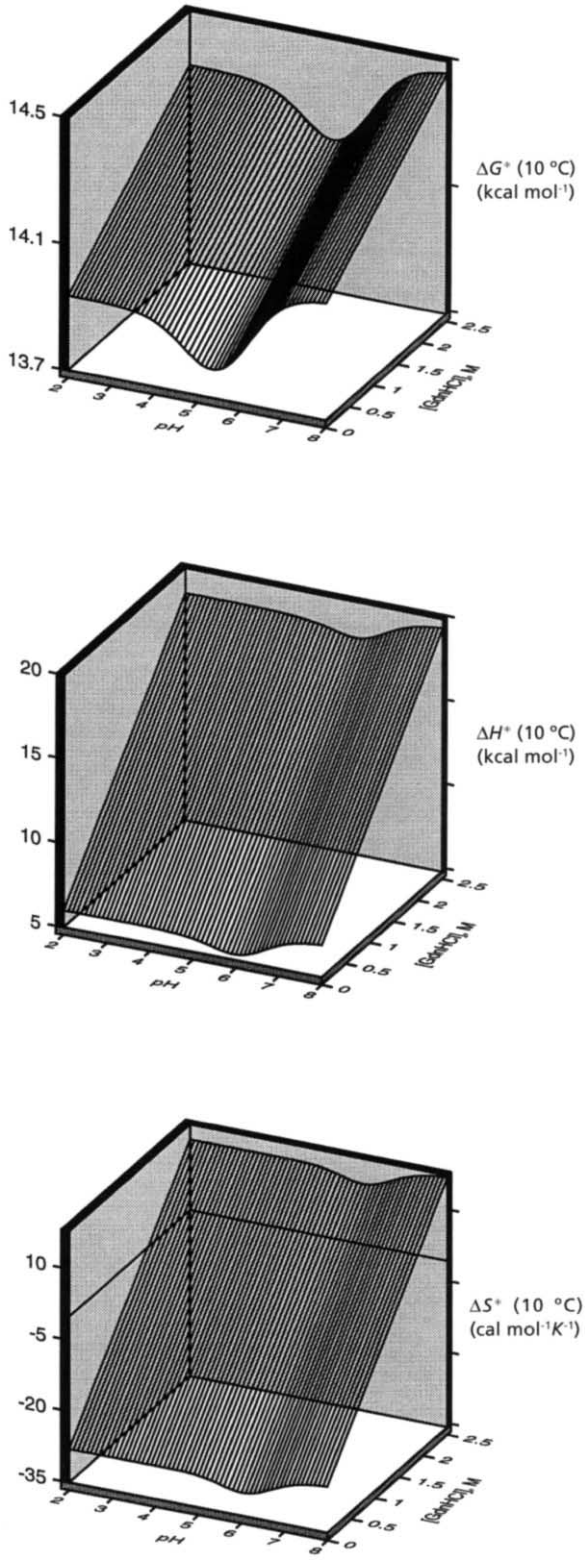


Fig. 5 The activation parameters for the formation of the transition state as a function of pH and GdnHCl concentration. *a*, The activation free energy, *b*, enthalpy, and *c*, entropy, all at 10 °C.

leading to the appearance of a two-state refolding process between U_{vf} and N.

The presence of a stable intermediate is expected to slow down the refolding rate. This is observed experimentally. If $\ln(k_{vf})$ is extrapolated linearly from high GdnHCl concentrations, where I_Φ is not appreciably populated, to 0 M [GdnHCl] (Fig. 2*e-g*), the extrapolated rate constant is larger than the experimentally measured rate constant. Therefore, the fastest mechanism by which a protein can fold is the simple two-state process. Nonetheless, whether the intermediate is highly populated or not, conformational folding of the protein, in the absence of denaturant, takes only a few milliseconds (2–20 ms; Fig. 2*e–g*) and not seconds or minutes. Hence, conformational folding is an intrinsically very-fast process.

We have been able to study the conformational folding reaction of RNase A in the absence of the complicating effect of *cis/trans* proline isomerization. The data indicate the presence of a hydrophobically collapsed intermediate which is highly populated when the protein is refolded at low GdnHCl concentrations. The evidence for a hydrophobic collapse comes from the observed pK_a shifts, the burial of solvent exposed surface area, and, most importantly, from the thermodynamic parameters. Furthermore, the experimental data suggest that the rate-limiting transition state on the refolding pathway results from the formation of ordered structure within the hydrophobically collapsed intermediate, and that the T. S. is not native-like and is not identical to the T.S. on the unfolding pathway.

Methods

Kinetic experiments. The instrument used and the experimental procedures employed have been described previously[19]. The native protein in 1.50 M GdnHCl and pH 5.9 was first unfolded at 4.20 M GdnHCl and pH 2.0 for a set delay time, typically 0.9 s at 5 °C and 0.4 s at 15 °C. After such delay time, > 99 % of the unfolded state of the protein consisted of the U_{vf} species. The unfolded protein (U_{vf}) was then refolded under a variety of conditions as shown in Fig. 2. The refolding step was monitored by absorbance at 287 nm. The refolding decay curves were fitted to a single exponential. The total absorbance refolding amplitude did not change with pH or GdnHCl concentration. Appropriate buffers were used for each pH range, and the temperature dependence of the buffers was taken into account[44,45]. The buffers used at or near pH 2, 3, 4, 5, 6, 7 and 8 are glycine, citric acid, formic acid, acetic acid, MES, BES or MOPS and EPPS, respectively. All pHs are given at the temperatures employed in the experiments. The error in the pH measurement is estimated to be ± 0.1 pH units. The protein is 100 % folded under all refolding conditions used. This was checked by obtaining thermal transition curves under the extreme refolding conditions (for example, at 2.5 M GdnHCl and pH 4.0 where the melting temperature is 24.9 °C). Hence, all the refolding rates are measured outside the unfolding transition zone and, therefore, the unfolding rate constant does not contribute to the observed reaction rate.

Fitting to the kinetic model of equation (1). The kinetic data (Fig. 2) were fitted to the model of equation (1) using the expressions of equations (3), (4) and (5) for the GdnHCl- and pH-dependence of k_{vf}. Only two residues were considered to undergo pK_a shifts during the refolding of U_{vf} : a carboxylic acid residue and a histidine residue, however it should be emphasized that we do not rule out the possibility that more than one carboxylic acid or histidine residue might be involved in the pK_a shifts. Nevertheless, it can easily be shown that, if two amino acid residues which undergo large shifts in pK_a can fit the pH-dependence data adequately, then the other ionizable groups in the protein must undergo only small changes in their pK_as. In addition, even if more than two ionizable groups are considered in fitting the pH-dependence data, the 'net' shift in the pK_a would still be in the direction expected for hydrophobic collapse. The possible presence, therefore, of these other ionizable groups in no way affects the conclusions of this paper. There is some indication, from the points at the high pH end (Fig. 2*b,d*), that a third residue might be involved. This third residue is either a Tyr or a Lys.

From experimental studies of the temperature dependence of the pK_as of Asp, Glu, and His, it has been observed that the pK_a's of Asp and Glu have a very weak temperature dependence, while that of His has a strong temperature dependence[46]. Hence, in considering the temperature dependence of k and K, the pK_as of the carboxylic acid residue were fixed with temperature, while those of the histidine residue were allowed to change with temperature. To avoid a multiple-minima problem, the assumption was made that the structure of the transition state has no significant temperature dependence between 5 °C and 15 °C. Hence, the sum $(m_{pre\text{-}eq} + m^*)/RT$ was assumed to be independent of temperature, while both $m_{pre\text{-}eq}$ and m^* were allowed to change with temperature.

The theoretical fit to the experimental kinetic data was obtained by using a simplex algorithm[47] to minimize $\chi^2 = [y(\text{experimental}) - y(\text{theoretical model})]^2 / [\sigma + 0.015\ k_{vf}]^2$, where σ represents the experimentally determined sample standard deviation, and the factor of $(0.015\ k_{vf})$ was added to the standard deviation to take into account the systematic errors. All the data were fitted simultaneously (a global fit). The reduced χ^2 of the fit was 0.98. The standard deviations of the fitted parameters were obtained by using the Monte Carlo procedure described previously[19]; 1,200 simulated data sets were used in the procedure.

Thermodynamic parameters. By using the equilibrium constants (K) obtained at 5 °C and 15 °C and by assuming that there is no temperature dependence for the heat capacity change, the enthalpy change at 10 °C was derived from the van't Hoff equation:

$$\frac{d\ln(K)}{d(1/T)} = -\frac{\Delta H^\circ(T)}{R}$$

while the entropy change at 10 °C was calculated from:

$$\frac{d\,[T\ln(K(T))]}{d[T]} = \frac{\Delta S^\circ(T)}{R}$$

The standard free energy change at 10 °C was then calculated using the relationship $\Delta G^\circ(T) = \Delta H^\circ(T) - T\,\Delta S^\circ(T)$. The activation free energy, enthalpy, and entropy at 10 °C were calculated using the above equations; however, $K^* = k\,\frac{h}{k_B T}$ (k_B is the Boltzmann constant and h is Planck's constant)[25] was used, instead of K, in these equations.

Received 22 November 1994; accepted 11 April 1995.

Acknowledgements
We thank J. -Y. Trosset and K. - H. Cho for their help in using the Cornell Supercomputing Facility. This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health. Support was also received from the National Foundation for Cancer Research.

1. Kim, P.S. & Baldwin, R.L. Intermediates in the folding reactions of small proteins.*A. Rev. Biochem.* **59**, 631–660 (1990).
2. Nall, B.T. Proline isomerization as a rate-limiting step. In *Mechanisms of protein folding* (ed. Pain, R. H.) 80–103 (Oxford University Press, New York; 1994).
3. Sosnick, T.R., Mayne, L., Hiller, R. & Englander, S.W. The barriers in protein folding. *Nature struct. Biol.* **1**, 149–156 (1994).
4. Matheson, R.R. Jr. & Scheraga, H.A. A method for predicting nucleation sites for protein folding based on hydrophobic contacts. *Macromolecules* **11**, 819–829 (1978).
5. Baldwin, R.L. How does protein folding get started? *Trends biochem. Sci.* **14**, 291–294 (1989).
6. Dill, K.A., Fiebig, K.M. & Chan, H.S. Cooperativity in protein-folding kinetics. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1942–1946 (1993).
7. Ptitsyn, O.B. Protein folding: hypotheses and experiments. *J. prot. Chem.* **6**, 273–293 (1987).
8. Kuwajima, K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins Struct. Funct. Genet.* **6**, 87–103 (1989).
9. Christensen, H. & Pain, R. H. The contribution of the molten globule model. In *Mechanisms of protein folding* (ed. Pain, R. H.) 55–79 (Oxford University Press, New York; 1994).
10. Barrick, D. & Baldwin, R.L. The molten globule intermediate of apomyoglobin and the process of protein folding. *Prot. Sci.* **2**, 869–876 (1993).
11. Yutani, K., Ogasahara, K. & Kuwajima, K. Absence of the thermal transition in apo-α-lactalbumin in the molten golbule state: a study by differential scanning microcalorimetry. *J. molec. Biol.* **228**, 347-350 (1992).
12. Goto, Y. & Fink, A.L. Conformational states of β-lactamase: molten-globule states at acidic and alkaline pH with high salt. *Biochemistry* **28**, 945–952 (1989).
13. Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E. & Razgulyaev, O.I. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **262**, 20–24 (1990).
14. Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. Comparison of the transient folding intermediates in lysozyme and α-lactalbumin. *Biochemistry* **24**, 874–881 (1985).
15. Garvey, E.P., Swank, J. & Matthews, C.R. A hydrophobic cluster forms early in the folding of dihydrofolate reductase. *Proteins Struct. Funct., Genet.* **6**, 259–266 (1989).
16. Jennings, P. A. & Wright, P. E. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* **262**, 892–896 (1993).
17. Wlodawer, A., Svensson, L.A., Sjölin, L. & Gilliland, G.L. Structure of phosphate-free ribonuclease A refined at 1.26 Å. *Biochemistry* **27**, 2705–2717 (1988).
18. Schmid, F. X. Fast-folding and slow-folding forms of unfolded proteins. *Meths Enzymol.* **131**, 70–82 (1986).
19. Houry, W.A., Rothwarf, D.M. & Scheraga, H.A. A very fast phase in the refolding of disulphide-intact ribonuclease A: implications for the refolding and unfolding pathways. *Biochemistry* **33**, 2516–2530 (1994).
20. Pace, C.N., Laurents, D.V. & Thomson, J.A. pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry* **29**, 2564–2572 (1990).
21. Pace, C.N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Meths Enzymol.* **131**, 266–280 (1986).
22. Schellman, J.A. The thermodynamic stability of proteins. *A. Rev. Biophys. biophys. Chem.* **16**, 115–137 (1987).
23. Alonso, D.O.V. & Dill, K.A. Solvent denaturation and stabilization of globular proteins. *Biochemistry* **30**, 5974–5985 (1991).
24. Tanford, C. Protein denaturation. Theoretical models for the mechanism of denaturation.*Adv. protein Chem.* **24**, 1–95 (1970).
25. Glasstone, S., Laidler, K.J. & Eyring, H. *The Theory of Rate Processes* 1–27 (McGraw-Hill Book Company, New York; 1941).
26. Donovan, J.W., Laskowski, Jr., M. & Scheraga, H.A. Carboxyl group interactions in lysozyme. *J. molec. Biol.* **1**, 293–296 (1959).
27. Nozaki, Y. & Tanford, C. Proteins as random coils. II. Hydrogen ion titration curve of ribonuclease in 6 M guanidine hydrochloride. *J. Am. chem. Soc.* **89**, 742–749 (1967).
28. Greenstein, J.P. Studies of the peptides of trivalent amino acids. *J. biol. Chem.* **101**, 603–621 (1933).
29. Kauzmann, W. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1–63 (1959).
30. Tanford, C. *The hydrophobic effect: Formation of micelles and biological membranes* 16–23 (John Wiley & Sons, New York; 1973).
31. Némethy G. & Scheraga, H. A. The structure of water and hydrophobic bonding in proteins. III. The thermodynamic properties of hydrophobic bonds in proteins. *J. phys. Chem.* **66**, 1773–1789 (1962).
32. Scheraga, H.A. Treatment of hydration in conformational energy calculations on polypeptides and proteins. In *Structure and reactivity in aqueous solution* (ed. Cramer, C. J. & Truhlar, D. G.) 360–370 (ACS Symposium Series No. 568, Washington, DC; 1994).
33. Privalov, P.L. & Gill, S.J. Stability of protein structure and hydrophobic interaction.*Adv. prot. Chem.* **39**, 191–234 (1988).
34. Creighton, T.E. Electrophoretic analysis of the unfolding of proteins by urea. *J. molec. Biol.* **129**, 235–264 (1979).
35. Makhatadze, G.I. & Privalov, P.L. Protein interactions with urea and guanidinium chloride. A calorimetric study. *J. molec. Biol.* **226**, 491–505 (1992).
36. Haynie, D.T. & Freire, E. Structural energetics of the molten globule state. *Proteins Struct. Funct. Genet.* **16**, 115–140 (1993).
37. Xie, D. & Freire, E. Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. *Proteins Struct. Funct. Genet.* **19**, 291–301 (1994).
38. Creighton, T.E. Toward a better understanding of protein folding pathways. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5082-5086 (1988).
39. Matouschek, A., Kellis Jr, J.T., Serrano, L. & Fersht, A.R. Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122–126 (1989).
40. Fersht, A.R., Kellis Jr., J.T., Matouschek, A.T.E.L. & Serrano, L. Folding pathway enigma. *Nature* **343**, 602 (1990).
41. Brandts, J.F., Halvorson, H.R. & Brennan, M. Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of proline residues. *Biochemistry* **14**, 4953–4963 (1975).
42. Lin, S.H., Konishi, Y., Nall, B.T. & Scheraga, H.A. Influence of an extrinsic cross-link on the folding pathway of ribonuclease A. Kinetics of folding-unfolding. *Biochemistry* **24**, 2680–2686 (1985).
43. Antosiewicz, J., McCammon, J.A. & Gilson, M.K. Prediction of pH-dependent properties of proteins. *J. molec. Biol.* **238**, 415–436 (1994).
44. Good, N.E., Winget, G.D., Winter, W., Connolly, T.N., Izawa, S. & Singh, R. M. M. Hydrogen ion buffers for biological research. *Biochemistry* **5**, 467–477 (1966).
45. Good, N.E. & Izawa, S. Hydrogen ion buffers. *Meths Enzymol.* **24**, 53–68 (1972).
46. Edsall, J.T. Dipolar ions and acid-base equilibria. In *Proteins, amino acids and peptides as ions and dipolar ions* (eds Cohn, E.J. & Edsall, J.T.) 75–115 (Reinhold Publishing Corporation, New York; 1943).
47. Caceci, M.S. & Caheris, W.P. Fitting curves to data. *BYTE* **9**, 340–362 (1984).