© 2001 Nature Publishing Group http://structbio.nature.com

Nonprolyl *cis* peptide bonds in unfolded proteins cause complex folding kinetics

Günter Pappenberger[1,4], Hüseyin Aygün[2], Joachim W. Engels[2], Ulf Reimer[2,3], Gunter Fischer[3] and Thomas Kiefhaber[1]

Folding of tendamistat, an inhibitor of α-amylase, is a fast two-state process accompanied by two minor slow reactions, which were assigned to prolyl isomerization. In a proline-free variant, 5% of the molecules still fold slowly with a rate constant of 2.5 s^{-1}. This reaction is caused by a slow equilibrium between two populations of unfolded molecules. The time constant for this equilibration process, its sensitivity to LiCl and its temperature dependence identify it as a *cis-trans* isomerization of nonprolyl peptide bonds. Although nonprolyl peptide bonds have the *cis* conformation populating only ~0.15% in unfolded proteins, their large number generates a significant fraction of slow-folding molecules. This emphasizes that heterogeneous populations in an unfolded protein can induce complex folding kinetics on various time scales.

A major focus of protein folding studies has been the characterization of the rate limiting events for formation of the native structure. Several proteins were shown to fold rapidly according to a two-state mechanism without transiently populated intermediates[1,2], whereas folding of other proteins involves partially folded intermediates with various degrees of structure formation[3]. The elucidation of kinetic folding mechanisms and of the role of intermediates is essential to understand the energy landscape of protein folding. Analysis of the folding mechanism for almost all proteins is, however, complicated by heterogeneities in the unfolded chain caused by slow equilibration reactions. This gives rise to slow reactions coupled to the actual folding process. Most commonly, this heterogeneity involves the *cis-trans* isomerization of Xaa–Pro peptide bonds[4,5], but other reactions, such as heme religation, have been described[6,7]. To correctly assess complex folding kinetics, it is especially important to distinguish between the contributions of partially folded intermediates from those of heterogeneities in the unfolded state.

Here we analyze the slow folding reaction of the α-amylase inhibitor, tendamistat, a small all-β-protein of 74 amino acids[8]. It contains two disulfide bridges, which are kept intact during folding/unfolding, and three prolyl residues, which are in the *trans* conformation in the native state. For the majority of the tendamistat molecules, folding is a rapid two-state process[9]. However, a fraction of the molecules folds in two slower reactions. The slowest reaction, with a tryptophan fluorescence amplitude of 10% of the total fluorescence change between unfolded and native protein and a rate constant (k) of 0.08 s^{-1}, could be attributed to prolyl isomerization because of its slow equilibration in the unfolded state and its catalysis by human cyclophilin 18 (ref. 9). The intermediate reaction, with an amplitude of 5% and a rate constant of ~3 s^{-1}, was initially assigned to the formation of an intermediate with a non-native Xaa–Pro peptide bond because it could not be catalyzed. We could now assign this reaction to the *cis*→*trans* isomerization of nonproline peptide bonds, which statistically occur in the *cis* conformation in the unfolded polypeptide chain. This provides a general source of heterogeneity in unfolded proteins and can explain previously unaccounted minor folding reactions in many proteins.

Proline-free tendamistat

To simplify the mechanism of tendamistat folding, we replaced all prolyl residues with alanines (tendamistat P7A/P9A/P50A). The unfolding transition of tendamistat, induced by GdmCl, pH 7.0, at 25 °C (Fig.1*a*), shows a shift in the transition midpoint from 6.4 M GdmCl in the wild type protein to 5.0 M in the proline-free variant. This shift corresponds to a decrease in protein stability (ΔG°) of 11.6 kJ mol^{-1}. Despite the replacement of all prolyl residues, refolding of the proline-free variant is still complex (Fig. 2). The majority of the fluorescence change (95.5 ± 0.3%) occurs in a fast reaction with a rate constant of 10.1 ± 0.1 s^{-1}. The remaining 4.5% change occurs on a slower time scale with a rate constant of 2.3 s^{-1}; This reaction closely resembles the intermediate kinetic phase in the wild type protein. The slowest refolding reaction of wild type tendamistat, with a rate constant of 0.085 s^{-1}, is eliminated in the proline-free protein, corresponding to the lack of proline isomerization[9].

The two rates of the kinetic phases of the proline-free variant were further characterized by measuring their GdmCl-dependence. The GdmCl-dependence of the major kinetic phase is V-shaped (Fig. 1*b*), which is commonly observed for proteins that fold according to the two-state model[1,10]. The rate constant for the minor refolding reaction is independent of the GdmCl concentration (Fig. 1*b*). The rates of the two reactions are similar at ~1.5 M GdmCl. However, above 1.5 M GdmCl, the amplitude of the GdmCl-independent reaction rapidly approaches zero (Fig. 1*c*). The strong GdmCl-dependence of the major kinetic phase identifies it as a folding reaction associated with structure formation. The insensitivity to GdmCl of the minor folding reaction indicates that protein solvent interactions change little in this process.

[1]Biozentrum der Universität Basel, Abteilung Biophysikalische Chemie, Klingelbergstrasse 70, CH-4056 Basel, Switzerland. [2]Institut für Organische Chemie der Universität Frankfurt, Marie Curie Strasse 11, D-60439 Frankfurt/Main, Germany. [3]Max Planck Research Unit "Enzymology of Protein Folding", Weinbergweg 22, D-06120 Halle (Saale), Germany. [4]Present address: The Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London, SW3 6JB, UK.

Correspondence should be addressed to T.K. *email: t.kiefhaber@unibas.ch*

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

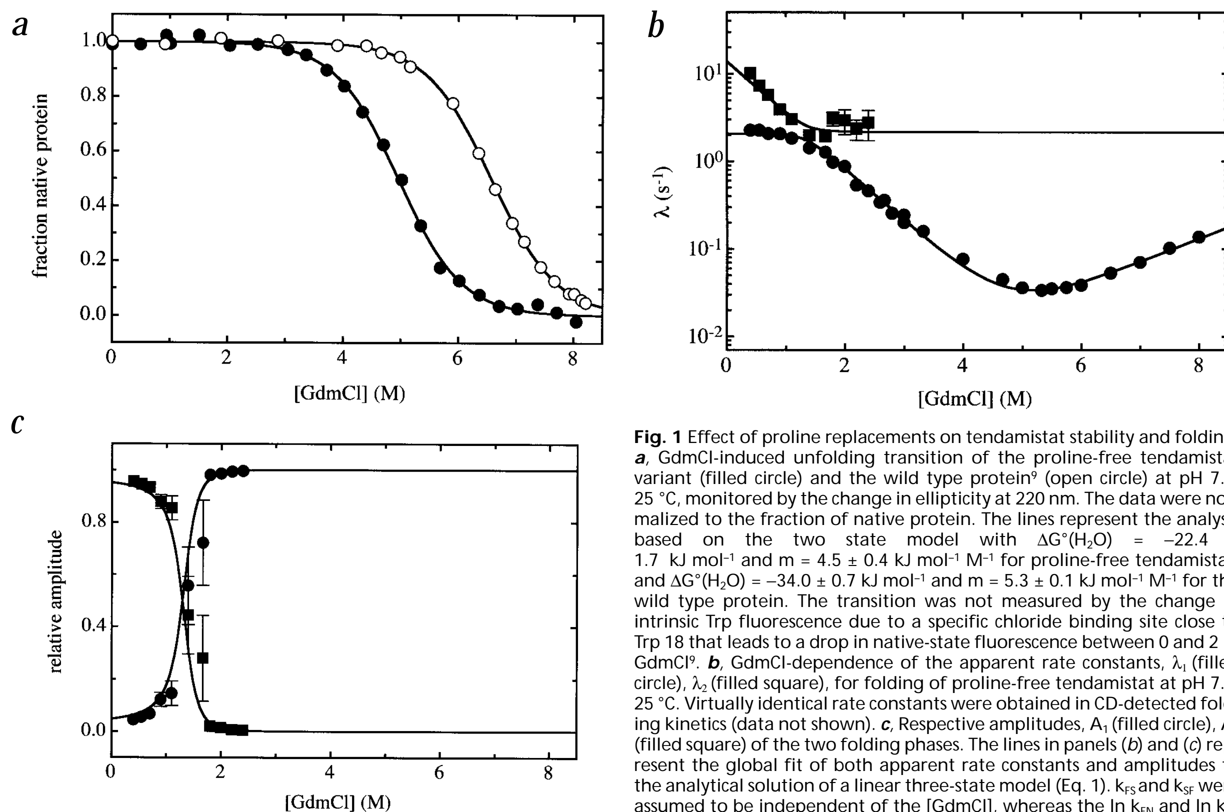
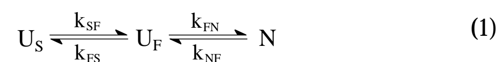


Fig. 1 Effect of proline replacements on tendamistat stability and folding. ***a***, GdmCl-induced unfolding transition of the proline-free tendamistat variant (filled circle) and the wild type protein[9] (open circle) at pH 7.0, 25 °C, monitored by the change in ellipticity at 220 nm. The data were normalized to the fraction of native protein. The lines represent the analysis based on the two state model with $\Delta G°(H_2O) = -22.4 \pm 1.7$ kJ mol^{-1} and m = 4.5 ± 0.4 kJ mol^{-1} M^{-1} for proline-free tendamistat, and $\Delta G°(H_2O) = -34.0 \pm 0.7$ kJ mol^{-1} and m = 5.3 ± 0.1 kJ mol^{-1} M^{-1} for the wild type protein. The transition was not measured by the change in intrinsic Trp fluorescence due to a specific chloride binding site close to Trp 18 that leads to a drop in native-state fluorescence between 0 and 2 M GdmCl[9]. ***b***, GdmCl-dependence of the apparent rate constants, λ_1 (filled circle), λ_2 (filled square), for folding of proline-free tendamistat at pH 7.0, 25 °C. Virtually identical rate constants were obtained in CD-detected folding kinetics (data not shown). ***c***, Respective amplitudes, A_1 (filled circle), A_2 (filled square) of the two folding phases. The lines in panels (*b*) and (*c*) represent the global fit of both apparent rate constants and amplitudes to the analytical solution of a linear three-state model (Eq. 1). k_{FS} and k_{SF} were assumed to be independent of the [GdmCl], whereas the ln k_{FN} and ln k_{NF} were assumed to have linear dependence with the [GdmCl] was assumed: $\ln k = \ln k(H_2O) + m / RT \times [GdmCl]$. The kinetic parameters obtained from the fits are $k_{FN}(H_2O) = 13.4$ s^{-1}, $k_{NF}(H_2O) = 9.9 \times 10^{-4}$ s^{-1}, m_{FN} = 3.42 kJ mol^{-1} M^{-1}, m_{NF} = −1.51 kJ mol^{-1} M^{-1}, k_{FS} = 0.075 s^{-1}, k_{SF} = 2.1 s^{-1}.

Heterogeneity in the unfolded state

To test whether the complex refolding kinetics in proline-free tendamistat are caused by a folding intermediate or by different populations of unfolded molecules, we monitored the formation of native tendamistat in interrupted refolding experiments[11,12] and compared the resulting kinetics to the direct fluorescence detected folding reaction. In interrupted refolding experiments the protein is allowed to refold for a certain time ('age time') and then is transferred to unfolding conditions to monitor the resulting kinetics. Because the native state has a characteristic stability and barrier for unfolding, the measured rate constant is characteristic of the unfolding reaction and is different from those of any partially folded intermediates. Interrupted refolding experiments measure the increase in amplitude of the unfolding reaction of the native protein as a function of the age time, which gives the time course of formation native protein. Thus, interrupted refolding experiments are able to detect whether a folding reaction produces native protein. The formation of native tendamistat proceeds in two kinetic phases (Fig. 3) with the same rate constants and amplitudes as those monitored directly by fluorescence. Thus, both kinetic phases directly produce native tendamistat, and there is no evidence for the transient population of partially folded states.

To detect putative slow equilibration processes in unfolded tendamistat, double-jump experiments[4] were performed. In these experiments the native protein is unfolded for a certain time ('age time') in a first mixing step. In a second mixing step the protein is allowed to refold, and the refolding kinetics are monitored. When a slow equilibration process occurs in the unfolded polypeptide chain, the amplitude of any refolding reaction is limited by this equilibration reaction and will develop slowly after unfolding. Any direct folding reaction will appear as soon as the protein is unfolded — that is, with the same rate constant as unfolding takes place. The appearance of the amplitudes of the two folding reactions was monitored as a function of the unfolding duration (Fig. 4). The fast-refolding species forms with virtually the same rate constant as measured for the unfolding reaction. This reaction loses ~3% of its amplitude at longer unfolding times (rate constant = 1.6 ± 0.4 s^{-1}). The slow-refolding molecules form significantly more slowly than the unfolding reaction takes place (Fig. 4). A single exponential fit gives a rate constant of 2.0 ± 0.2 s^{-1} and an amplitude of 3.9 ± 0.1% for the appearance of the slow-folding reaction, which correlates with the decay of the fast-folding molecules at longer times. These results show that the complex kinetics of refolding is caused by a slow equilibration process in the unfolded state between a fast folding population (U_F) and a slow-folding population (U_S):



$$U_S \underset{k_{FS}}{\overset{k_{SF}}{\rightleftharpoons}} U_F \underset{k_{NF}}{\overset{k_{FN}}{\rightleftharpoons}} N \quad (1)$$

The identical amplitudes for refolding measured by fluorescence and interrupted refolding experiments, combined with the lack of a burst phase reaction (Fig. 2), indicate that neither U_S nor U_F form partially folded states. Therefore, Eq. 1 could be used to describe the folding kinetics of the proline-free variant under all experimental conditions. This is confirmed by fitting the GdmCl-dependence of both apparent rate constants and their amplitudes (Fig. 1*b*,*c*) with the equilibrium unfolding tran-

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

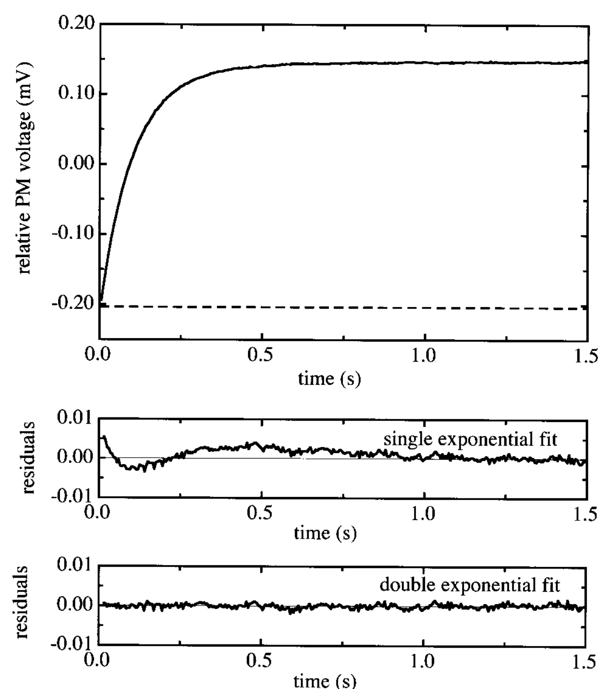


Fig. 2 Refolding kinetics of the proline-free tendamistat variant. Folding was monitored by the change in tryptophan fluorescence Refolding conditions were 0.40 M GdmCl, pH 7.0, 25 °C. The dashed line indicates the fluorescence of the unfolded protein extrapolated to refolding conditions from the GdmCl-dependence of the fluorescence signal at high denaturant concentrations. At time point zero the kinetic trace agrees well with the signal of the unfolded protein, which shows that no faster reactions occur in the dead time of mixing. The residuals for the single and double exponential fits show the existence of two kinetic phases with rate constants and relative amplitudes of: λ_1 = 2.26 ± 0.11 s^{-1}, A_1 = 4.5 ± 0.3%; λ_2 = 10.1 ± 0.03 s^{-1}; and A_2 = 95.5 ± 0.2%. The errors are from the uncertainty of fitting.

sition (Fig. 1*a*) to the analytical solutions of Eq. 1 (refs 13,14). This allowed a quantitative description of the experimental data and gave values of k_{SF} = 2.1 s^{-1} and k_{FS} = 0.075 s^{-1}. The fit also reproduces the decrease in the amplitude of the faster reaction (A2) and the concomitant increase in the amplitude of the slower reaction (A1) above 1.5 GdmCl when folding becomes slower than the equilibration process (Fig. 1*c*).

Eq. 1 also allows the global fit of the appearance of the slow- and fast-folding phases in the double jump experiments (Fig. 4) and gives values of k_{NF} = 27.1 ± 0.1 s^{-1}, k_{FS} = 0.093 ± 0.005 s^{-1} and k_{SF} = 2.6 ± 0.2 s^{-1}. These values are nearly identical to the respective values from fitting the GdmCl-dependence of folding (Fig. 1).

Activation parameters for the slow-folding reaction

Our results show that two slowly interconverting species of unfolded molecules account for the complex folding of proline-free tendamistat. In addition to isomerization of peptide bonds preceding proline residues[4,15] and ligand exchange of the heme moiety[6,7], threading of loops formed by disulfide bridges that overlap in sequence was suggested to cause slow reactions in protein folding[16,17]. None of these reactions can account for the heterogeneity found in the unfolded state of proline-free tendamistat, because it contains neither prolines nor a ligand. Proline-free tendamistat has two disulfide bridges between Cys 11 and Cys 27 and between Cys 45 and Cys 73, which are kept intact during unfolding/folding. Since they form sequential nonoverlapping loops, the disulfide bridges cannot give rise to loop threading reactions. This is confirmed by the folding kinetics of different single disulfide tendamistat variants, where entanglements of disulfide loops can be excluded. All single disulfide variants show the same GdmCl-independent folding reaction as the wild type protein with a relaxation time of ~2 s^{-1} (A. Bachmann & T.K., unpublished results). Disulfide isomerization reactions between a *syn* and an *anti* stereoisomeric form can also be excluded, since they occur on the 1–10 ms time scale in native proteins[18] and are significantly faster in unstructured polypeptides[19].

Cis-trans equilibria of nonprolyl peptide bonds could be another source of heterogeneities in unfolded proteins[4]. Nonprolyl peptide bonds have a *cis* content of ~0.5% in dipeptides and of ~0.15% in longer oligopeptides, which serve as a reference for unfolded proteins[20]. The activation parameters and the thermodynamics of the *cis-trans* equilibrium at nonprolyl peptide bonds have been investigated in NMR studies on model peptides[20] and on a slow *trans*→*cis* isomerization reaction in RNase T1 P39A (ref. 21). This protein contains a Tyr 38–Ala 39 *cis* peptide bond in the native state, which produces an extremely slow refolding reaction. Both systems show nearly the same rate constant for the *cis*→*trans* isomerization as the slow-folding reaction of proline-free tendamistat (Table 1).

To test if the activation parameters of the slow equilibration process in unfolded proline-free tendamistat are compatible with nonprolyl *cis*→*trans* isomerization, we measured the temperature dependence of folding. The two rate constants of folding (λ_1 and λ_2) and the equilibrium constant between U_S and U_F ($K = [U_F] / [U_S]$) were determined between 10 and 40 °C at pH 5.5 and 0.4 M GdmCl (Fig. 5). The temperature-dependence of λ_2 has a pronounced curvature commonly observed for protein folding reactions[22], confirming that this process is associated with a significant change in heat capacity ($\Delta C_p^{\ddagger}$ = –2.2 ±

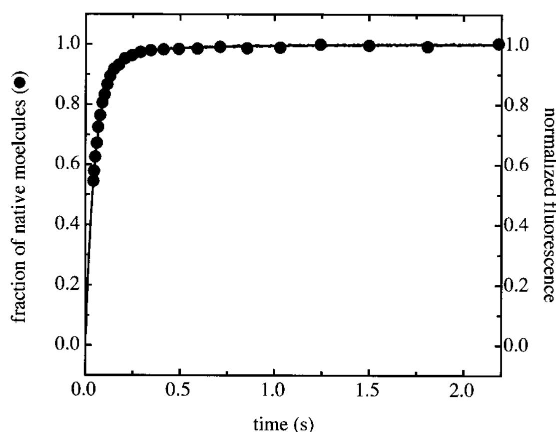


Fig. 3 Time course of formation of the native state of proline-free tendamistat. The folding reaction was monitored by the change in tryptophan fluorescence (solid line) and by the appearance of native molecules in interrupted refolding experiments (filled circle) under the same experimental conditions as described in the text. To obtain a better separation of the two reactions, refolding was carried out in 1 M GdmCl, 0.5 M Na_2SO_4, pH 5.5, 25 °C, which accelerates the faster reaction but does not affect the slower reaction. The apparent rate constants and relative amplitudes are: λ_1 = 2.63 ± 0.09 s^{-1}, A_1 = 4.75 ± 0.11% ; λ_2 = 19.7 ± 0.04 s^{-1}, A_2 = 95.3 ± 0.1% for the fluorescence experiments and λ_1 = 2.4 ± 0.5 s^{-1}, A_1 = 5.4 ± 1.0%; λ_2 = 20.9 ± 0.7 s^{-1}, A_2 = 94.6 ± 2.0% for the interrupted refolding experiments.

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

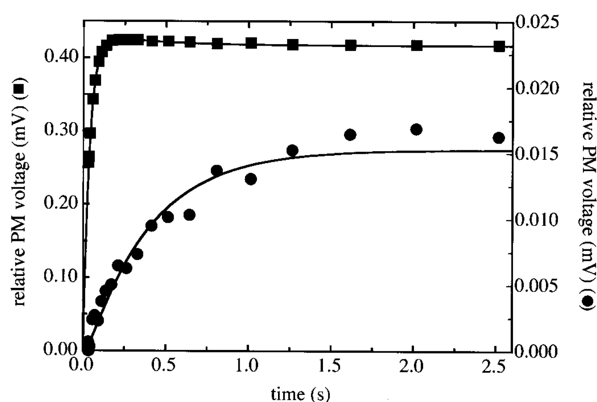


Fig. 4 Double jump experiments to monitor the slow equilibration process in unfolded proline-free tendamistat. Formation of the fast-refolding (filled square) and of the slow-refolding molecules (filled circle) are shown. The data show that the two refolding reactions are formed at largely different rates. The faster reaction forms with a time constant of 27.5 ± 1 s^{-1}, which is virtually identical to the rate constant for unfolding under these condition. The slower reactions is formed with a rate constant of 2 ± 0.2 s^{-1}. The amplitudes of the two refolding reactions are plotted *versus* the unfolding time. ('age time'). The lines correspond to a global fit of the data to the analytical solution of Eq 1. The values for the microscopic rate constants are: k_{NF} = 27.1 ± 0.1 s^{-1}, k_{FS} = 0.093 ± 0.005–s^{-1}, k_{SF} = 2.55 ± 0.15 s^{-1}. k_{FN} was set to 0, as unfolding under the given conditions is virtually irreversible.

0.1 kJ mol^{-1} K^{-1}). The temperature-dependence of λ_1, which corresponds to k_{SF} under these conditions, shows Arrhenius behavior as indicated by the linear dependence of ln λ on 1 / T. The fit gives (Fig 5*a*) an activation energy (E_A) of 51 ± 2 kJ mol^{-1} and a pre-exponential factor (A) of $10^{9.3 \pm 0.3}$ s^{-1}. These values are comparable to the respective values in the model peptides (E_A = 65 kJ mol^{-1}, A = $10^{11.7}$ s^{-1})[20], but they are significantly smaller than the activation energy of 107 kJ mol^{-1} found in unfolded RNase T_1 P39A[21] (Table 1). The lack of a curvature in the Arrhenius plot indicates a negligible $\Delta C_p^{\ddagger}$ (see Table 1).

The fraction of slow-folding molecules significantly increases with temperature; the reaction enthalpy (ΔH°) is –29 ± 1 kJ mol^{-1} (Fig. 5*b*). The resulting van't Hoff plot is linear, indicating similar C_p-values for U_F and U_S. These results qualitatively agree with the data from the model peptides (ΔH = –16 kJ mol^{-1}; $\Delta C_p \approx 0$). The absence of an appreciable ΔC_p between U_S and U_F further agrees with energy calculations that demonstrate identical solvation properties of *cis* and *trans* peptide bonds[23]. The significant temperature dependence of the equilibrium constant is in contrast to prolyl isomerization equilibria, which are approximately temperature-independent[24,25].

The properties of the slow-folding reaction of proline-free tendamistat agree remarkably well with the data on *cis-trans* isomerization of nonprolyl peptide bonds, making this the most likely source for the complex folding kinetics. The peptide studies have shown that increasing chain length leads (i) to an increase in the rate constant for the *cis-trans* reaction, (ii) to a decrease in activation energy and (iii) to an decrease in reaction enthalpy[20]. The respective values obtained for the slow equilibration reaction in proline-free tendamistat are in agreement with this trend (Table 1).

Test for peptide bond isomerization

Lithium chloride in anhydrous trifluoroethanol (TFE) was shown to increase the *cis* content in Xaa–Pro peptide bonds[26]. We used an Ala–Ala–Tyr model peptide to test if this effect is also observed in nonprolyl peptide bonds. 1D NMR spectra show that the Ala–Tyr bond in the tripeptide has 0.19% *cis* content in aqueous solutions. In 0.6 M LiCl /TFE the fraction of the *cis* isomer increases significantly to 0.55% (Fig. 6) in a Li^+-specific manner. When proline-free tendamistat is denatured in anhydrous TFE containing increasing concentrations of LiCl, an approximately linear increase of the fraction of slowly refolding molecules is observed (Fig 7). It rises by a factor of 2–3 between 0 M and 0.6 M LiCl, which corresponds to the three-fold increase in *cis* content of the Ala–Tyr bond in the model peptide under comparable conditions (Fig. 6). The apparent folding rate constants after jumping out of the LiCl/TFE mixture are not affected by the presence of LiCl (data not shown). These results

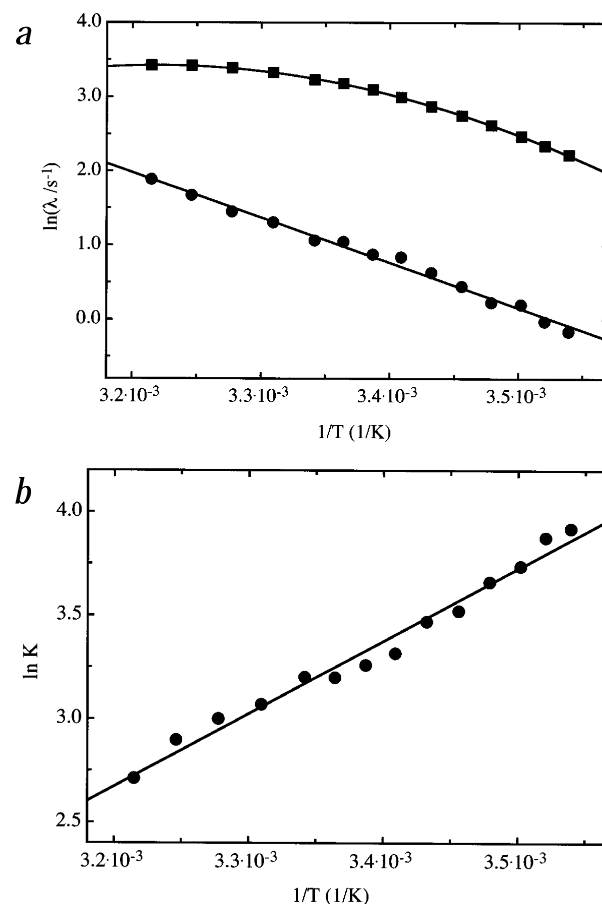


Fig. 5 Temperature dependence of the two folding phases. ***a***, Temperature-dependence of the apparent rate constants for refolding of the proline-free variant of tendamistat. Refolding conditions were 0.40 M GdmCl, pH5.5. The activation energy (E_A) and the pre-exponential factor A for the slow folding reaction were obtained using Eq. 3 (solid line). The data were additionally fit to Eq. 4 to obtain $\Delta H^{\circ\ddagger}$, $\Delta S^{\circ\ddagger}$ and $\Delta C_p^{\ddagger}$ (Table 1). The pre-exponential factor (k_0) of the Eyring equation (k_BT / h) was used for analysis of the peptide-bond isomerization reaction since the rate-limiting step is associated with the breakage of the partial double bond in the amide linkage. For the analysis of the folding reaction a temperature-independent pre-exponential factor of 10^8 s^{-1} (ref. 41) was used. The slow refolding reaction (filled circle) shows a temperature-independent activation energy of 51 ± 2 kJ mol^{-1}. The activation parameters for the fast folding reaction (filled squarse) are given in Table 1. ***b***, Temperature-dependence of the equilibrium between slow and fast folding species. The reaction enthalpy ΔH° for the equilibrium in the unfolded state was determined with the van't Hoff equation (Eq. 5). The equilibrium constant (K = $[U_F]$ / $[U_S]$; filled circles) was obtained from the fluorescence amplitudes of fast and slow refolding reactions. The fit gave a value of ΔH° = -29±1 kJ mol^{-1} (solid line), independent of the temperature ($\Delta C_p \approx 0$; see table 1).

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

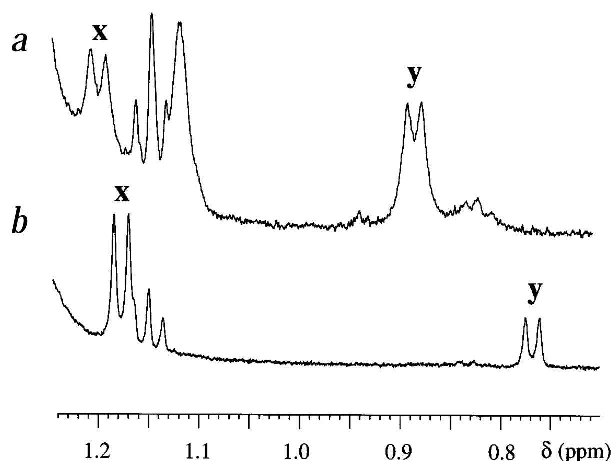


Fig. 6 Effect of LiCl on the *cis-trans* equilibrium of a nonprolyl peptide bond. ^{1}H NMR spectra of Ala-Ala-Tyr in a 0.6 M LiCl solution at 21.1 ± 0.1 °C in ***a***, TFE-d_3 and in ***b***, 10:1 (v/v) H_2O/D_2O. The resonances marked (x) are the ^{13}C satellites from the methyl group of Ala 2 of the *trans* isomer. These satellites have a signal intensity of 1.1% of the actual *trans* signal. The duplet resonating at higher field (y) is the actual methyl signal of the *cis* isomer[20]. At 37 °C, an exchange peak between the resonances from the *trans* and *cis* isomer becomes visible.

strongly support the assignment of the slow-refolding reaction to nonprolyl peptide bond isomerization.

The role of nonprolyl peptide bond isomerization

All experimental results on folding of proline-free tendamistat argue for nonprolyl *cis*→*trans* isomerization as the rate-limiting step in refolding of ~5% of the unfolded molecules. Commonly, the fraction of *cis* isomer of nonprolyl peptide bonds is considered to be negligible in accordance with literature values of ~0.15% for the fraction of *cis* isomer per peptide bond[20]. The high fraction of *cis* isomers in unfolded tendamistat can be explained if (i) a particular nonprolyl peptide bond exhibits a strongly increased propensity for the *cis* conformation in the unfolded state, or (ii) a *cis* peptide bond at any number of positions in the polypeptide chain prevents folding.

The fraction of *cis* isomer at prolyl peptide bonds was found to depend strongly on the amino acid residue preceding the proline residue; the highest fraction of *cis* conformation follows aromatic residues[27]. A fraction of ~0.15% *cis* isomer was found for the Ala–Tyr and Tyr–Ala peptide bond (Table 1), and similar values were recently obtained for a Gly–Gly bond (C. Schiene-Fischer & G.F., pers. comm.). This argues that amino acid sequence has a negligible contribution to the *cis*:*trans* ratio of nonprolyl peptide bonds; therefore, it is unlikely that a single peptide bond would populate 5% of the *cis* isomer in the unfolded state.

Tendamistat has 73 peptide bonds; thus, a fraction of 0.15% *cis* isomer per peptide bond would lead to ~10% of unfolded polypeptide chains with a *cis* peptide bond. The majority of these molecules will have a single *cis* isomer. The probability of finding two or more *cis* peptide bonds in the same unfolded chain is ~0.5%. The observed 5% slow refolding amplitude thus implies either that the average *cis* content of all peptide bonds in tendamistat is ~0.07% or that only about half of the peptide bonds prevent folding when they are in the *cis* conformation. At the other positions, a *cis* isomer is tolerated and is not rate-limiting for folding. A similar dichotomy in 'essential' and 'not essential' sites was postulated for proline residues on the basis of energy calculations[28].

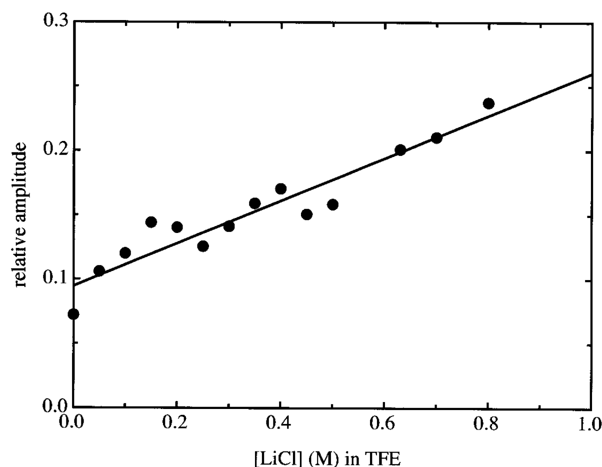
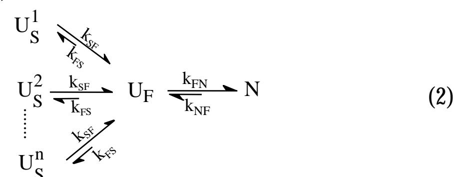


Fig. 7 LiCl dependence of tendamistat folding. The relative amplitude of the slow refolding reaction in the proline-free variant of tendamistat is shown as a function of the LiCl concentration. The protein was denatured in TFE containing the indicated LiCl concentrations. A slight increase in the amplitude of λ_1 for refolding starting from TFE unfolded protein compared to refolding starting from GdmCl-unfolded protein is observed. This is probably due to a TFE effect on the *cis-trans* equilibrium.

As discussed above, it is likely that a number of peptide bonds contribute to the observed slow-folding reaction in proline-free tendamistat. In this case, the $U_S \rightleftharpoons U_F$ reaction is complex and consists of many parallel reactions starting from a number of different U_S states and ending in a single folding-competent U_F state. The interconversion between the different U_S states requires two slow isomerization reactions and can thus be neglected (Eq. 2).



(2)

The *cis-trans* reaction is practically irreversible during folding, since it is followed by the fast folding reaction from U_F to N. Thus, there is no kinetic coupling between the parallel isomerization reactions, each of which has an individual apparent rate constant with an amplitude of ~0.15%. Since the *cis* content seems to be almost independent of the amino acids involved, we assume that the isomerization rate is also approximately sequence-independent. Thus, the kinetics of the individual isomerization reactions can not be resolved. The experiments, therefore, monitor a single slow reaction with an amplitude corresponding to the sum of the amplitudes of all parallel isomerization reactions.

Implications for protein folding

Ever since prolyl isomerization was identified as a slow step in protein folding, nonprolyl isomerization has been considered as a possible cause for heterogeneity in the unfolded state[4], especially when complex kinetics were observed in proline-free proteins[29–32]. However, there was no direct evidence for *cis-trans*

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

Table 1 Comparison of thermodynamic and kinetic parameters for nonprolyl peptide bond *cis-trans* isomerization[1]

	This work	RNase T_1 P39A[3]	Peptide AAYAA[4]	
Relevant bond	n.d.[2]	Tyr–Ala	Ala–Tyr	Tyr–Ala
Percent *cis* in U[5]	5.0 ± 0.3	0.17	0.14	0.11
$k_{cis \to trans}$ (s^{-1})	2.5 ± 0.2	1.4	2.4	1.8
$E_{A(cis \to trans)}$[6] (kJ mol^{-1})	51 ± 2	107 ± 5	65	66
$A_{cis \to trans}$[6] (s^{-1})	$10^{9.3 \pm 0.3}$	n.d.	$10^{11.7}$	$10^{11.8}$
$\Delta H^{\circ\ddagger}_{cis \to trans}$[7] (kJ mol^{-1})	47 ± 2	n.d.	n.d.	n.d.
$\Delta S^{\circ\ddagger}_{cis \to trans}$[7] (kJ mol^{-1} K^{-1})	$-7.5 \pm 0.5 \times 10^{-2}$	n.d.	n.d.	n.d.
$\Delta C^{\ddagger}_{P\,(cis \to trans)}$[7] (kJ mol^{-1} K^{-1})	−0.6 ± 0.3	n.d.	n.d.	n.d.
$\Delta H^{\circ}_{cis \rightleftharpoons trans}$[8] (kJ mol^{-1})	−29 ± 1	n.d.	−16	−16

[1]Determined at 25 °C.
[2]Not determined.
[3]Ref. 21. Parameter values were obtained for the isomerization in the unfolded protein.
[4]Ref. 20.
[5]The number of peptide bonds contributing to the observed fraction of molecules with a *cis* peptide bond is not known (see text).
[6]Calculated from Fig. 6*a*
[7]From a fit of the data in Fig. 6*a* to eq. 4.
[8]Calculated from Fig. 6*b*

isomerization of nonprolyl peptide bonds as a rate-limiting step in protein folding. Nonprolyl *cis* peptide bonds were found in the native structure of several proteins[33]. Bonds of this type can give rise to very slow refolding reactions[21]. We show that the reverse *cis→trans* reaction limits the folding of a fraction of unfolded molecules in proteins when all peptide bonds are *trans* in the native state. This heterogeneity of unfolded polypeptides will be of general relevance for protein folding because it applies to a majority of peptide bonds.

Slow folding reactions limited by nonprolyl peptide bond isomerization should be observable in a protein with a fast conformational transition. Because fast folding proteins are commonly small, the expected amplitude for the nonprolyl isomerization is ~5–10%. Additional refolding phases with rate constants of 1–10 s^{-1} and amplitudes of 5–30% are frequently observed for fast-folding proteins. They are commonly explained by prolyl isomerization accelerated in a partially folded intermediate. Whereas this possibility cannot be ruled out (except in the case of proline-free proteins), our results demonstrate that nonprolyl isomerization can give rise to such reactions. Addition of human cyclophilin did not increase the rate of the slow-folding reaction in proline-free tendamistat (data not shown), indicating that nonprolyl peptide bonds are not subject to PPIase catalysis.

Theoretical considerations have suggested that the conformational changes upon folding of small proteins can deviate from single exponential kinetics under certain circumstances[34–36]. If the barriers between local minima on the energy landscape become significantly larger than 3 k_BT, multiple kinetic phases are predicted and kinetic partitioning between productive and non-productive pathways has been proposed. Experiments are currently underway to test those theoretical predictions. Complex nonexponential kinetics have recently been reported for protein folding on the submillisecond time scale[37]. It is therefore important to distinguish kinetics due to folding through multiple intermediates from those due to heterogeneities in the unfolded ensemble. We demonstrate that such heterogeneities can occur on faster time scales than commonly conceived for proline isomerization and how to test for nonprolyl *cis-trans* isomerization as a rate-limiting step in protein folding.

Methods

Protein construction, expression and purification. Tendamistat P7A/P9A/P50A was constructed, expressed in *Streptomyces lividans* and purified as described[38]. Purity was checked by mass spectroscopy and exceeded 99%. GdmCl (AA grade) was from Nigu Chemie, and 2,2,2-trifluoroethanol (99.5+%) and anhydrous LiCl (99.9+%) from Aldrich. Sodium acetate and sodium cacodylate were from Fluka and sodium sulfate from Merck.

Equilibrium unfolding transition. The GdmCl-induced equilibrium unfolding transition of proline-free tendamistat at pH 7.0, 25 °C, was monitored by the change in ellipticity at 220 nm in an Aviv 62A DS spectropolarimeter. The data were fit following the procedure of Santoro and Bolen[39].

Folding kinetics. All single and sequential mixing stopped-flow experiments were performed on an Applied Photophysics SX.18MV instrument at 25 °C. The average of 10–20 individual kinetic traces was analyzed in all kinetic measurements. In all experiments, folding or unfolding was monitored by the change in fluorescence >320 nm after excitation at 276 nm. Single mixing refolding and unfolding experiments were carried out at pH 7.0, 25 °C. Interrupted refolding experiments (Fig. 3) were carried out in 1 M GdmCl, 0.5 M Na_2SO_4, pH 5.5, 25 °C. Lowering the pH to 5.5 (ref. 40) and adding SO_4^{2-} increases the stability of tendamistat compared to standard refolding conditions, which accelerates the faster reaction but does not affect the slower reaction. For the interrupted refolding experiments, completely unfolded tendamistat (in 6.0 M GdmCl, pH2.0) was refolded for the indicated time (t) before an unfolding step in 7.0 M GdmCl, 80 mM $NaSO_4$, pH 3.2 was applied. Unfolding was monitored by the change in tryptophan fluorescence. The amount of native protein present at the time t is proportional to the amplitude of its characteristic unfolding reaction. Unfolding can be described by a single exponential function with a rate constant of 0.58 s^{-1}, independent of the refolding time, t. The unfolding amplitude is normalized to an final value of 1. Lowering the pH to 5.5 (ref. 40) and adding SO_4^{2-} increased the stability of tendamistat compared to standard refolding conditions, which accelerated the fast-folding phase but did not affect the slow-folding phase. For the double jump experiments (Fig. 4), native protein was unfolded for the indicated time in 5.0 M GdmCl, 20% TFE, pH 1.8, before refolding in 0.8 M GdmCl, 3.3% TFE, pH 5.5 at 25 °C was monitored by the change in fluorescence above 320 nm. The addition of TFE and GdmCl speeds up the unfolding reaction to $\lambda = 23\ s^{-1}$ and leads to completely unfolded protein with identical spectroscopic properties as the GdmCl-unfolded state. Refolding can be described by a double exponential function with rate constants of 14 s^{-1} and 3 s^{-1}, independent of the age time >300 ms. For shorter age times, the slow refolding reaction had very small amplitudes and was difficult to fit. Thus, all kinetic traces were fit globally to the sum of two exponential functions, with the global rate constants k_1 and k_2. The dependence of the resulting amplitudes, A1 and A2, on unfolding time are plotted in Fig. 4. Refolding in the presence of TFE (Fig. 7) was performed by an 1+25 mixing step into final concentrations of 0.23 M GdmCl, 3.8% TFE and residual LiCl concentrations between 0 and 30 mM at pH 7.0.

Temperature dependence of the folding reactions. Refolding conditions were 0.40 M GdmCl, pH 5.5. The activation energy E_A and the pre-exponential factor A for the slow-folding reaction were derived from the temperature dependence of the rate constant according to the Arrhenius equation

$$k = A \times \exp(-E_A / RT) \tag{3}$$

The data were additionally fit to a general rate equation

$$k = k_0 \times \exp(-\Delta G^{\circ\ddagger} / RT) = k_0 \times \exp(-1 / RT\ (\Delta H^{\circ\ddagger}(T_0) - T \times \Delta S^{\circ\ddagger}(T_0) + \Delta C_p^{\ddagger} \times (T - T_0 - T \times \ln(T / T_0))) \tag{4}$$

© 2001 Nature Publishing Group http://structbio.nature.com

© 2001 Nature Publishing Group http://structbio.nature.com

The reaction enthalpy ΔH° for the equilibrium in the unfolded state was determined with the van't Hoff equation

$$d\ln K / d(1/T) = -\Delta H^\circ / R \quad (5)$$

NMR studies on the *cis* content in model peptides. NMR measurements were performed on a Bruker ARX-500 NMR spectrometer. For the determination of the *cis* content, the intensities of resonances corresponding to the *cis* and *trans* isomer were integrated using the FELIX NMR software[20]. Resonances were assigned as previously reported for the same peptide (ref. 20). The peptides (Fig. 6) were allowed to equilibrate for 1 h in a 0.6 M LiCl solution in TFE-d_3 or in H_2O/D_2O 10:1 (v/v). The assignment of the resonance corresponding to the methyl group of Ala 2 in the *cis* isomer was confirmed by a 2D ^{1}H exchange NMR experiment (EXSY).

Received 28 November, 2000; accepted 9 February, 2001.

1. Jackson, S.E. & Fersht, A.R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428–10435 (1991).
2. Jackson, S.E. How do small single-domain proteins fold? *Folding Des.* **3**, R81–R91 (1998).
3. Baldwin, R.L. & Rose, G.D. Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem. Sci.* **24**, 77–83 (1999).
4. 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).
5. Schmid, F.X. & Baldwin, R.L. Acid catalysis of the formation of the slow-folding species of RNase A: evidence that the reaction is proline isomerization. *Proc. Natl. Acad. Sci. USA* **75**, 4764–4768 (1978).
6. Colon, W., Wakem, L.P., Sherman, F. & Roder, H. Identification of the predominant non-native histidine ligand in unfolded cytochrome *c*. *Biochemistry* **36**, 12535–12541 (1997).
7. Yeh, S.-R., Takahashi, S., Fan, B. & Rousseau, D.L. Ligand exchange in unfolded cytochrome c. *Nature Struct. Biol.* **4**, 51–56 (1998).
8. Pflugrath, J., Wiegand, I., Huber, R. & Vértesy, L. Crystal structure determination, refinement and the molecular model of the α-amylase inhibitor Hoe-467A. *J. Mol. Biol.* **189**, 383–386 (1986).
9. Schönbrunner, N., Koller, K.-P. & Kiefhaber, T. Folding of the disulfide-bonded β-sheet protein tendamistat: Rapid two-state folding without hydrophobic collapse. *J. Mol. Biol.* **268**, 526–538 (1997).
10. Tanford, C. Protein denaturation. Part B. The transition from native to denatured state. *Adv. Prot. Chem.* **23**, 218–282 (1968).
11. Schmid, F.X. Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry* **22**, 4690–4696 (1983).
12. Kiefhaber, T. In *Methods in molecular biology*, vol. 40 (ed. Shirley, B.A.) 313–341 (Humana Press, Totowa, NJ; 1995).
13. Szabo, Z.G. In *Comprehensive chemical kinetics*, vol. 2 (eds Bamford, C.H. & Tipper, C.F.H.) 1–80 (Elsevier Publishing Company, Amsterdam; 1969).
14. Kiefhaber, T., Quaas, R., Hahn, U. & Schmid, F.X. Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry* **29**, 3053–3061 (1990).
15. Balbach, J. & Schmid, F.X. In *Protein folding: frontiers in molecular biology* (ed. Pain, R.) 212–249 (Oxford University Press, Oxford; 2000).
16. Nall, B.T., Garel, J.-R. & Baldwin, R.L. Test of the extended two-state model for the kinetic intermediates observed in the folding transition of ribonuclease A. *J. Mol. Biol.* **118**, 317–330 (1978).
17. De Young, L.R. *et al.* RhNGF slow unfolding is not due to proline isomerization: Possibility of a cystine knot loop-threading mechanism. *Protein Sci.* **5**, 1554–1566 (1996).
18. Otting, G., Liepinsh, E. & Wüthrich, K. Disulfide bond isomerization in BPTI and BPTI(G36S): An NMR study of correlated mobility in proteins. *Biochemistry* **32**, 3571–3582 (1993).
19. Fraser, R.R., Boussard, G., Saunder, J.K., Lambert, J.B. & Mixan, C.E. Barriers to rotation about the sulfur–sulfur bond in acyclic disulfides. *J. Am. Chem. Soc.* **92**, 3822–3823 (1971).
20. Scherer, G., Kramer, M.L., Schutkowski, M., Reimer, U. & Fischer, G. Barriers to rotation of secondary amide peptide bonds. *J. Am. Chem. Soc.* **120**, 5568–5574 (1998).
21. Odefey, C., Mayr, L. & Schmid, F.X. Nonprolyl *cis/trans* peptide bond isomerization as a rate-determinig step in protein unfolding and refolding. *J. Mol. Biol.* **245**, 69–78 (1995).
22. Pohl, F.M. Temperature-dependence of the kinetics of folding of chymotrypsinogen A. *FEBS Lett.* **65**, 293–296 (1976).
23. Radzicka, A., Pedersen, L. & Wolfenden, R. Influence of solvent water on protein folding: free energies of solvation of *cis* and *trans* peptides are nearly identical. *Biochemistry* **27**, 4538–4541 (1988).
24. Schmid, F.X. Proline isomerization in unfolded ribonuclease A. The equilibrium between fast-folding and slow-folding species is independent of temperature. *Eur. J. Biochem.* **128**, 77–80 (1982).
25. Fischer, G., Heins, J. & Barth, A. The conformation around the peptide bond between P1 and P2 positions is important for catalytic activity of some proline-specific proteases. *Biophys. Biochim. Acta* **742**, 452–462 (1983).
26. Kofron, J.L., Kuzmic, P., Kishore, V., Colon-Bonilla, E. & Rich, D.H. Determination of kinetic constants for peptidyl prolyl *cis–trans* isomerases by an improved spectrophotometric assay. *Biochemistry* **30**, 6127–6134 (1991).
27. Reimer, U. *et al.* Side-chain effects on peptidyl-prolyl *cis/trans* isomerization. *J. Mol. Biol.* **279**, 449–460 (1998).
28. Levitt, M. Effect of proline residues on protein folding. *J. Mol. Biol.* **145**, 251–263 (1981).
29. Brandts, J.F., Brennan, M. & Lin, L.-N. Unfolding and refolding occur much faster for a proline-free protein than for most proline-containing proteins. *Proc. Natl. Acad. Sci. USA* **74**, 4178–4181 (1977).
30. Herning, T., Yutani, K., Taniyama, Y. & Kikuchi, M. Effect of proline mutations on the unfolding and refolding of human lysozyme: the slow refolding kinetic phase does not result from proline *cis–trans* isomerization. *Biochemistry* **30**, 9882–9891 (1991).
31. Walkenhorst, W.F., Green, S. & Roder, H. Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease. *Biochemistry* **36**, 5795–5805 (1997).
32. Maki, K., Ikura, T., Hayano, T., Takahashi, N. & Kuwajima, K. Effects of proline mutations on the folding of staphylococcal nuclease. *Biochemistry* **38**, 2213–2223 (1999).
33. Weiss, M.S., Jabs, A. & Hilgenfeld, R. Peptide bonds revisited. *Nature Struct. Biol.* **5**, 676 (1998).
34. Thirumalai, D. In *Statistical mechanics, protein structure, and protein substrate interactions* (ed. Doniach, S.) 115–134 (Plenum Press, New York, NY; 1994).
35. Chan, H.S. & Dill, K.A. Transition states and folding dynamics of proteins and heteropolymers. *J. Chem. Phys.* **100**, 9238–9257 (1994).
36. Bryngelson, J.D., Onuchic, J.N., Socci, N.D. & Wolynes, P.G. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Prot. Struct. Funct. Genet.* **21**, 167–195 (1995).
37. Sabelko, J., Ervin, J. & Gruebele, M. Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci. USA* **96**, 6031–6036 (1999).
38. Haas-Lauterbach, S. *et al.* High yield fermentation and purification of tendamistat disulfide analogues secreted by *Streptomyces lividans*. *Appl. Microbiol. Biotech.* **38**, 719–727 (1993).
39. Santoro, M.M. & Bolen, D.W. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068 (1988).
40. Renner, M., Hinz, H.-J., Scharf, M. & Engels, J.W. Thermodynamics of unfolding of the α-amylase inhibitor tendamistat: correlations between accessible surface area and heat capacity. *J. Mol. Biol.* **223**, 769–779 (1992).
41 Bieri, O. *et al.* The speed limit for protein folding measured by triplet–triplet energy transfer. *Proc. Natl. Acad. Sci. USA* **96**, 9597–9601 (1999).

© 2001 Nature Publishing Group http://structbio.nature.com