Parallel pathways in the folding of a short-term denatured scFv fragment of an antibody

Christian Freund[1,*], Peter Gehrig[1], Antonio Baici[2], Tad A Holak[3] and Andreas Plückthun[1]

Background: Antibodies are prototypes of multimeric proteins and consist of structurally similar domains. The two variable domains of an antibody (V_H and V_L) interact through a large hydrophobic interface and can be expressed as covalently linked single-chain Fv (scFv) fragments. The *in vitro* folding of scFv fragments after long-term denaturation in guanidinium chloride is known to be slow. In order to delineate the nature of the rate-limiting step, the folding of the scFv fragment of an antibody after short-term denaturation has been investigated.

Results: Secondary structure formation, measured by H/D-exchange protection, of a mutant scFv fragment of an antibody after short incubation in 6 M guanidinium chloride was shown to be multiphasic. NMR analysis shows that an intermediate with significant proton protection is observed within the dead time of the manual mixing experiments. Subsequently, the folding reaction proceeds via a biphasic reaction and mass spectrometry analyses of the exchange experiments confirm the existence of two parallel pathways. In the presence of cyclophilin, however, the faster of the two phases vanishes (when followed by intrinsic tryptophan fluorescence), while the slower phase is not significantly enhanced by equimolar cyclophilin.

Conclusions: The formation of an early intermediate, which shows amide-proton exchange protection, is independent of proline isomerization. Subsequently, a proline *cis–trans* isomerization reaction in the rapidly formed intermediate, producing 'non-native' isomers, competes with the fast formation of native species. Interface formation in a folding intermediate of the scFv fragment is proposed to prevent the back-isomerization of these prolines from being efficiently catalyzed by cyclophilin.

Addresses: [1]Department of Biochemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland. [2]Department of Rheumatology, University Hospital, CH-8091 Zürich, Switzerland. [3]Max-Planck Institute for Biochemistry, Am Klopferspitz 18a, D-82152 Martinsried, Germany.

*Present address: Department of Molecular Pharmacology and Biological Chemistry, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA.

Correspondence: Andreas Plückthun
E-mail: plueckthun@biocfebs.unizh.ch

Key words: single-chain Fv fragment, proline *cis–trans* isomerization, proton exchange, NMR, mass spectrometry

Received: **27 August 1997**
Revisions requested: **06 October 1997**
Revisions received: **14 November 1997**
Accepted: **14 November 1997**

Published: **17 December 1997**
http://biomednet.com/elecref/1359027800300039

Folding & Design 17 December 1997, **3**:39–49

Introduction

The *in vitro* folding of larger proteins consisting of more than one domain is often a slow process [1,2]. This is different from the behavior of several small proteins, some of which have been shown to fold fast and with two-state behavior [3–7]. Disulfide formation and proline isomerization are the best known examples of processes leading to slow folding [8,9], but additional factors related to the docking and correct annealing of individual domains have also been shown to be rate limiting for larger proteins [1,10,11]. It has been observed that refolding conditions can be crucial for the observation of kinetic intermediates [12] and that heterogeneity within the unfolded state can also contribute to the occurrence of intermediate states [13,14]. The question of whether certain intermediates actually represent kinetic traps of partially structured molecules [15], or even soluble aggregates of unfolded protein [16], has recently come into focus, and such steps might contribute to slow folding events, especially in larger proteins. In the case of proteins composed of distinct structural domains, local cooperativity is believed to exist prior to global cooperativity [17] and it has therefore been proposed that the formation of kinetic intermediates might simply be a consequence of the existence of more than one cooperative folding unit within the protein assembly [5].

Antibodies are prototypes of proteins, which consist of multiple domains, all belonging to the same structural family [18]. The individual domains can be covalently linked as in the heavy and light chains of natural antibodies, which show relatively few structural interactions between adjacent domains. In addition, the domains of different chains are usually non-covalently associated by the interaction of large, hydrophobic interfaces, as for example between the two variable domains V_H and V_L. The interaction of the domains during the refolding of an antibody light chain has been proposed to lead to a non-additive folding behavior [19]. For the Fab fragment of an antibody it has been shown that the association of the two chains takes place early on in the folding reaction and that proline isomerization is rate-limiting for the subsequent formation of the native structure [20]. Proline isomerization has also

been reported to be rate-limiting for the folding of isolated variable and constant domains [21,22], but no detailed information has been available for the formation of structure in the absence of additional slow steps such as disulfide bond formation or proline isomerization.

We have chosen the single-chain Fv (scFv) fragment (an antigen-binding fragment of the topology V_H–linker–V_L) of the phosphorylcholine-binding antibody McPC603 as a model system for our studies. The assignment of most of the backbone NH resonances of this two-domain β-sheet protein has been completed [23] and has allowed us to analyze the kinetic H/D exchange experiments by NMR methods. H/D exchange experiments and their analysis by NMR and mass spectrometry (MS) have become a very useful tool for the study of protein folding. MS provides information on the population of species with different amide-proton protection at certain time points [24], whereas NMR allows the kinetic analysis of individual amide-proton protection along the folding pathway [25,26]. Previously, we investigated the formation of H/D protection during the refolding of the equilibrium-denatured oxidized scFv fragment [27,28] and identified a rapidly formed intermediate, which shows some degree of amide-proton protection. The completion of folding from this intermediate was shown to be a slow, cooperative process involving both domains.

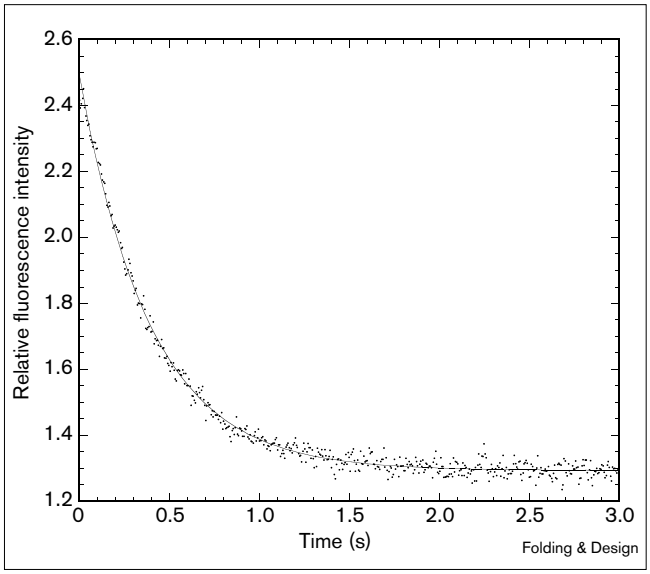
In the current study, we analyzed the refolding of a short-term denatured scFv fragment in order to delineate the role of proline *cis–trans* isomerization both for the formation of this intermediate and for the slow folding reaction. Because the V_L domain of the scFv fragment under investigation carries two *cis* prolines in the native structure [29], one of them being in CDR3 of the light chain where it makes several crucial contacts in the interface to the heavy chain, it seemed likely that proline isomerization might be involved in the rate-limiting steps of the folding reaction and native interface formation. Double-jump experiments, in which the protein is denatured only for a short time before refolding, in order to keep the prolines in a native-like conformation, were performed and analyzed by NMR, MS and fluorescence spectroscopy. Our results suggest that formation of an early intermediate is independent of proline conformations, but that folding from the intermediate with 'native-like' proline isomers competes with proline isomerization. As a consequence of the formation of non-native proline isomers, a folding intermediate with a non-native interface between V_H and V_L is formed. This trapped intermediate converts to the native state in an ~10-fold slower reaction than the formation of native species from the earlier intermediate with native-like proline isomers.

Results

Unfolding reaction

The unfolding reaction is biphasic at lower guanidinium chloride (GdnCl) concentrations, but becomes monophasic at 4 M GdnCl (Figure 1; the rate constant for unfolding $k = 2.54 \pm 0.02\ s^{-1}$ at 10°C). For the short-term denaturation (double-jump) experiments, the native scFv fragment was denatured for 20 s on ice in 6 M GdnCl so that the unfolding reaction and proton exchange of the unfolded protein [30,31] is completed, but proline isomerization should not have taken place to a significant degree [32–35].

Figure 1



A fluorescence trace for the unfolding of the wild-type (wt) scFv fragment in 4 M GdnCl at 10°C. The rate constant for unfolding $k = 2.54 \pm 0.02\ s^{-1}$ at 10°C.

Double-jump H/D experiments of the scFv fragment and their analysis by NMR

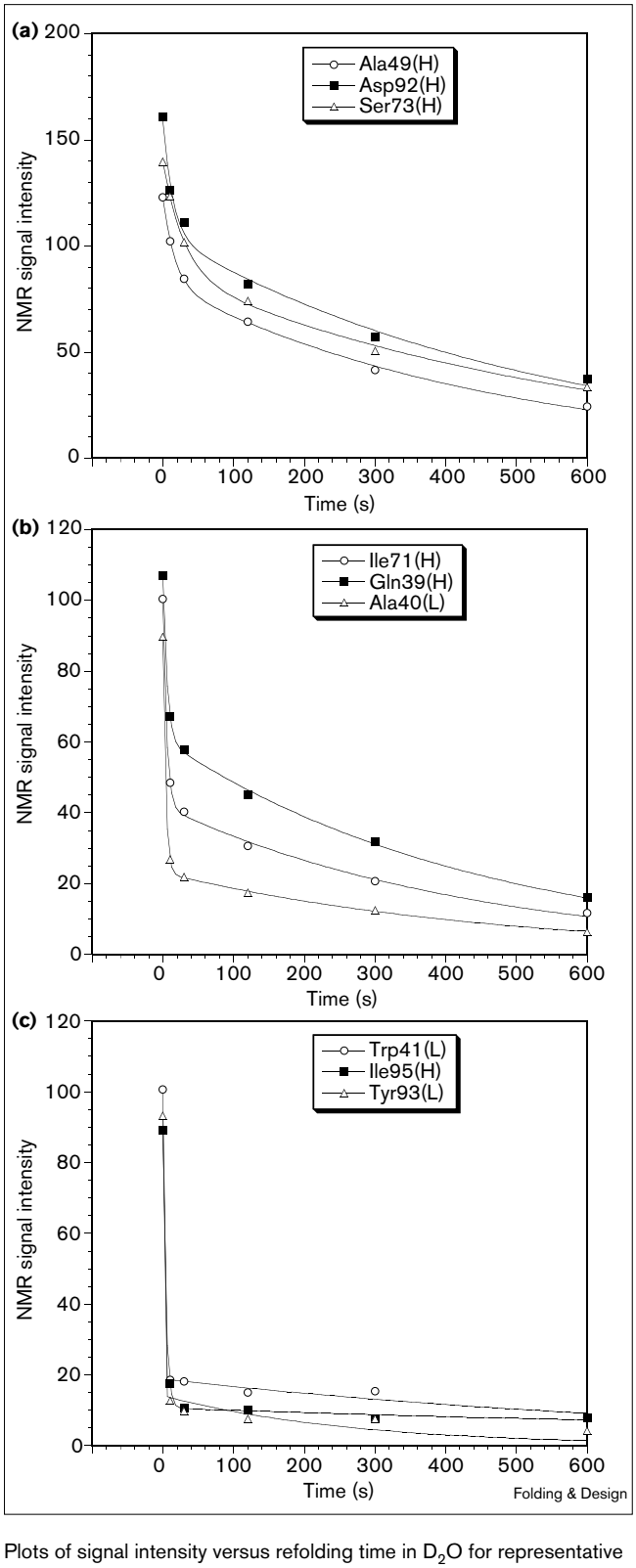
In order to resolve the folding pathway in terms of individual amide groups, we performed H/D exchange experiments and analyzed these experiments by NMR. A scFv fragment bearing the mutations Pro40→Ala, Ser63→Ala and Ala64→Asp was mostly used because this mutant was shown to be less prone to thermal aggregation [36]. NMR investigations of the native state of this mutant protein confirmed that the structure is changed only at the site of mutation and that the HSQC spectra of mutant and wild-type (wt) protein are otherwise superimposable, allowing the assignment to be transferred. It was also confirmed that the same folding intermediate with respect to amide-proton protection is populated for the mutant and the wt protein (see below). Differences in the rates of folding between mutant and wt protein are small but significant (see below).

For the kinetic H/D experiments, the native ^{15}N-labeled scFv fragment was diluted 10-fold into 6.6 M GdnCl/D_2O at pH 8.0 and incubated for 20 s on ice to denature

and deuterate the protein. Next, the protein was further diluted 100-fold into refolding buffer, containing 0.4 M arginine, 20 mM borate, 1 mM phosphorylcholine, pD_{read} = 8.0 (pD_{read}: the value directly read from the pH electrode), in D_2O. After various refolding times (10 s, 30 s, 60 s, 120 s, 300 s and 600 s), this solution was mixed with 6.25 volumes of 0.12 M potassium phosphate buffer, 1 mM phosphorylcholine, pH 4.0, to result in a final pH of 5.1. At this pH, exchange of amide protons with significant protection factors should be slow, while the exchange of unprotected amide groups should take place with a rate high enough, compared to the refolding reaction, to allow fast protonation of these residues. After completion of folding at this pH, samples were concentrated, dialyzed and lyophilized prior to dissolution in D_2O and acquisition of ^{15}N–^{1}H correlation spectra. In Figure 2, representative plots of the signal intensity of amide protons in the HSQC spectra versus time for refolding in D_2O (before dilution into H_2O) are shown. The curves were fitted by double-exponential functions, and a summary of the rates and amplitudes obtained is given in Table 1. For some residues (examples are given in Figure 2a) the six time points could be fitted well by a double-exponential function. For many residues, however, there is a large loss of signal intensity already within the dead time of the manual mixing experiment (examples are given in Figure 2b,c). The initial rapid decrease at some positions argues for the formation of an exchange-protected intermediate within the dead time of the experiment, as was observed for the equilibrium-denatured protein [27].

In Table 1, the proton occupancies after 10 s of refolding in D_2O are shown for the V_H and the V_L domains. The most prominent reduction of proton occupancies is observed for residues of the inner β sheet of the V_L domain (positions 40–43, 54 and 90–93). For the V_H domain, a few residues (positions 34, 82 and 95) also display largely reduced proton occupancies (Table 1). The observed pattern of proton occupancies is very similar to the pattern found within the intermediate of the refolding reaction of the equilibrium-denatured protein [27], suggesting that the same intermediate is formed in both experiments. In contrast to the refolding of the equilibrium-denatured protein, however, the reaction from the intermediate to the native state becomes biphasic for the short-term denatured protein. Due to the additional fast phase in the manual mixing experiments, the absolute values for the proton occupancies after 10 s of refolding are lower for the short-term denatured protein than for the equilibrium-denatured protein (see Table 1 and reference [27] for comparison). For the residues with no or small protection within the intermediate, this first rate constant could be estimated and an average rate constant of 0.05 s^{-1} was obtained. The second rate has an average value of 0.0014 s^{-1} and is similar to the rate of protection from exchange of the equilibrium-denatured protein (data not shown). This argues for the

Figure 2



Plots of signal intensity versus refolding time in D_2O for representative residues of the Pro40→Ala/Ser63→Ala/Ala64→Asp mutant scFv fragment. Double exponential fits are shown as lines. See text for more details. The residues are from the V_H (H) and V_L (L) domains.

Table 1

Rate constants and amplitudes for the decay of NMR signal intensity with time of refolding.

Residue	k_1 (s^{-1})	A_1 (signal intensity)	k_2 ($\times 10^{-3}$ s^{-1})	A_2 (signal intensity)	Proton occupancy after 10 s refolding
Leu4(H)	–	79.7 ± 3.0	1.4 ± 0.2	37.2 ± 2.1	0.41
Val5(H)	–	131.0 ± 6.3	1.4 ± 0.2	124.6 ± 4.5	0.59
Ser21(H)	0.05 ± 0.02	39.0 ± 9.0	1.3 ± 0.3	79.2 ± 8.0	0.86
Met34(H)	–	51.1 ± 3.1	–	22.9 ± 5.8	0.26
Gln39(H)	–	46.4 ± 2.0	2.2 ± 0.1	60.7 ± 1.4	0.63
Ile48(H)	–	29.3 ± 4.3	2.0 ± 0.5	30.0 ± 3.1	0.57
Ala49(H)	0.06 ± 0.01	40.0 ± 3.4	2.1 ± 0.2	82.6 ± 3.1	0.83
Ala50(H)	–	27.1 ± 6.2	2.2 ± 0.4	53.4 ± 4.5	0.69
Ser51(H)	0.03 ± 0.02	23.2 ± 10.6	2.4 ± 0.5	80.2 ± 10.4	0.98
Glu61(H)	0.05 ± 0.01	36.6 ± 4.2	0.7 ± 0.2	52.3 ± 3.5	0.79
Ala63(H)	0.05 ± 0.01	40.7 ± 5.1	1.8 ± 0.2	80.1 ± 4.7	0.85
Phe70(H)	0.09 ± 0.02	58.3 ± 4.9	1.6 ± 0.2	92.6 ± 4.0	0.75
Ile71(H)	–	58.6 ± 2.1	2.3 ± 0.2	42.0 ± 1.5	0.48
Ser73(H)	0.03 ± 0.006	52.0 ± 4.7	1.7 ± 0.2	87.7 ± 4.5	0.88
Asp75(H)	–	56.1 ± 8.9	1.8 ± 0.4	73.6 ± 6.6	0.65
Tyr82(H)	–	59.5 ± 3.4	–	15.6 ± 2.6	0.24
Gln84(H)	–	81.6 ± 4.7	2.1 ± 0.3	62.3 ± 3.4	0.5
Arg89(H)	–	46.7 ± 4.9	1.8 ± 0.3	64.0 ± 3.5	0.63
Asp92(H)	0.07 ± 0.03	53.6 ± 9.8	1.9 ± 0.3	105.7 ± 8.5	0.79
Ile95(H)	–	78.5 ± 1.3	–	10.7 ± 0.8	0.19
Val118(H)	–	30.3 ± 2.0	2.0 ± 0.2	39.7 ± 1.4	0.57
Thr119(H)	0.06 ± 0.02	53.0 ± 7.6	1.3 ± 0.3	73.1 ± 6.6	0.79
Val120(H)	0.07 ± 0.04	33.5 ± 7.9	1.1 ± 0.3	64.1 ± 6.6	0.78
Leu11(L)	0.05 ± 0.01	52.8 ± 4.2	0.6 ± 0.1	86.0 ± 3.5	0.84
Val13(L)	0.09 ± 0.03	57.4 ± 6.7	0.8 ± 0.1	135.8 ± 5.2	0.81
Glu17(L)	0.07 ± 0.02	74.6 ± 10.5	0.6 ± 0.2	130.0 ± 8.4	0.80
Thr20(L)	–	143.1 ± 10.6	1.5 ± 0.2	144.7 ±7.8	0.61
Cys23(L)	0.06 ± 0.02	35.9 ± 5.9	1.7 ± 0.3	58.0 ± 5.3	0.79
Ala40(L)	–	66.7 ± 0.6	2.1 ± 0.1	23.0 ± 0.4	0.30
Trp41(L)	–	81.8 ± 2.0	–	18.8 ± 4.2	0.18
Tyr42(L)	–	71.6 ± 2.8	–	11.6 ± 1.8	0.19
Gln43(L)	–	58.6 ± 1.0	1.9 ± 0.1	36.2 ± 0.7	0.43
Ile54(L)	–	63.7 ± 0.8	–	27.2 ± 0.5	0.52
Phe68(L)	0.08 ± 0.05	64.1 ± 16.3	1.2 ± 0.4	120.6 ± 13.5	0.78
Thr69(L)	0.04 ± 0.01	52.9 ± 5.8	1.5 ± 0.2	88.6 ± 5.3	0.89
Phe77(L)	–	26.0 ± 2.4	0.3 ± 0.1	63.7 ± 1.6	0.75
Thr78(L)	–	59.1 ± 3.2	1.0 ± 0.1	64.5 ± 2.4	0.66
Ile81(L)	–	92.0 ± 3.2	1.8 ± 0.3	43.0 ± 2.3	0.40
Val84(L)	0.03 ± 0.01	16.6 ± 1.5	1.4 ± 0.1	55.1 ± 1.4	0.93
Gln85(L)	–	63.6 ± 6.6	1.4 ± 0.2	86.4 ± 4.6	0.63
Ala86(L)	–	42.2 ± 8.0	1.3 ± 0.4	59.9 ± 5.9	0.69
Asp88(L)	–	110.4 ± 9.0	1.0 ± 0.2	147.9 ± 6.6	0.68
Ala90(L)	–	137.3 ± 2.2	0.9 ± 0.1	39.7 ± 1.4	0.25
Val91(L)	–	133.0 ± 1.8	–	18.9 ± 1.2	0.24
Tyr93(L)	–	76.5 ± 3.8	–	13.0 ± 7.7	0.14
Cys94(L)	–	53.1 ± 3.2	–	39.6 ± 2.0	0.49
Lys109(L)	0.04 ± 0.01	36.5 ± 5.4	0.6 ± 0.2	75.3 ± 4.6	0.89
Leu110(L)	0.09 ± 0.02	50.1 ± 5.2	1.4 ± 0.2	85.4 ± 4.2	0.76

Rates of the fast refolding phase (k_1) have been omitted if the loss of amplitude (A_1, measured as peak heights – signal intensity – of the respective NMR signal) due to the intermediate formation in the first time point is too large for an accurate determination. Rates of the second refolding phase (k_2) have been omitted if the corresponding amplitude (A_2) was very small compared to the overall amplitude. H, a residue in the V_H domain; and L, a residue in the V_L domain.

same folding process to take place for the single slow phase of folding for the equilibrium-denatured protein and the second phase of the double-jump experiments.

MS analysis of H/D experiments

In order to investigate the population of proton exchange-protected species along the folding pathway, we performed a similar series of H/D exchange experiments with the short-term denatured scFv fragment as described in the preceding section and analyzed the experiments by MS. From these experiments it should be distinguishable, whether the first phase of proton protection observed within the NMR experiments leads to a second, sequentially formed intermediate, or whether this phase corresponds to the fast formation of native or native-like protein for a subset of molecules. The results for the short-term denatured protein are shown in Figure 3. After 10 s of refolding a broad peak with an average mass of 27150 Da is observed, indicating that multiple species are populated. After 2 min, however, two major species remain, one 42 mass units higher than the protonated reference state and the second with the mass of the native species (27175 Da), which is 87 mass units higher than the protonated reference state. After 10 min, the species with the mass of the native protein is populated by ~85% of the molecules, leaving 15% of molecules with the mass of the intermediate.

These findings clearly favor the model of two parallel pathways. If a sequential mechanism were operative, a second highly populated intermediate should be observed after 1–2 min, according to the NMR experiments. The formation of a species with native-like protection should occur with a rate constant of 0.0014 s^{-1} (the rate constant of the second phase of the NMR experiments). This rate constant would predict that 50% of the molecules have native protection after 495 s. What we observe, however, is that half of the molecules are already fully protected after 120 s. This is exactly what one would expect, if two parallel pathways with rate constants of 0.06 s^{-1} (amplitude: 35%) and 0.0014 s^{-1} (amplitude: 65%) are assumed. Because in the NMR experiments similar rate constants and amplitudes were found from the average values for those residues that could be fitted well by a double exponential (see Table 1), and because all residues in the NMR experiments display a significant amplitude for the first phase, we propose parallel pathways of secondary structure stabilization, when folding is started from short-term denatured protein.

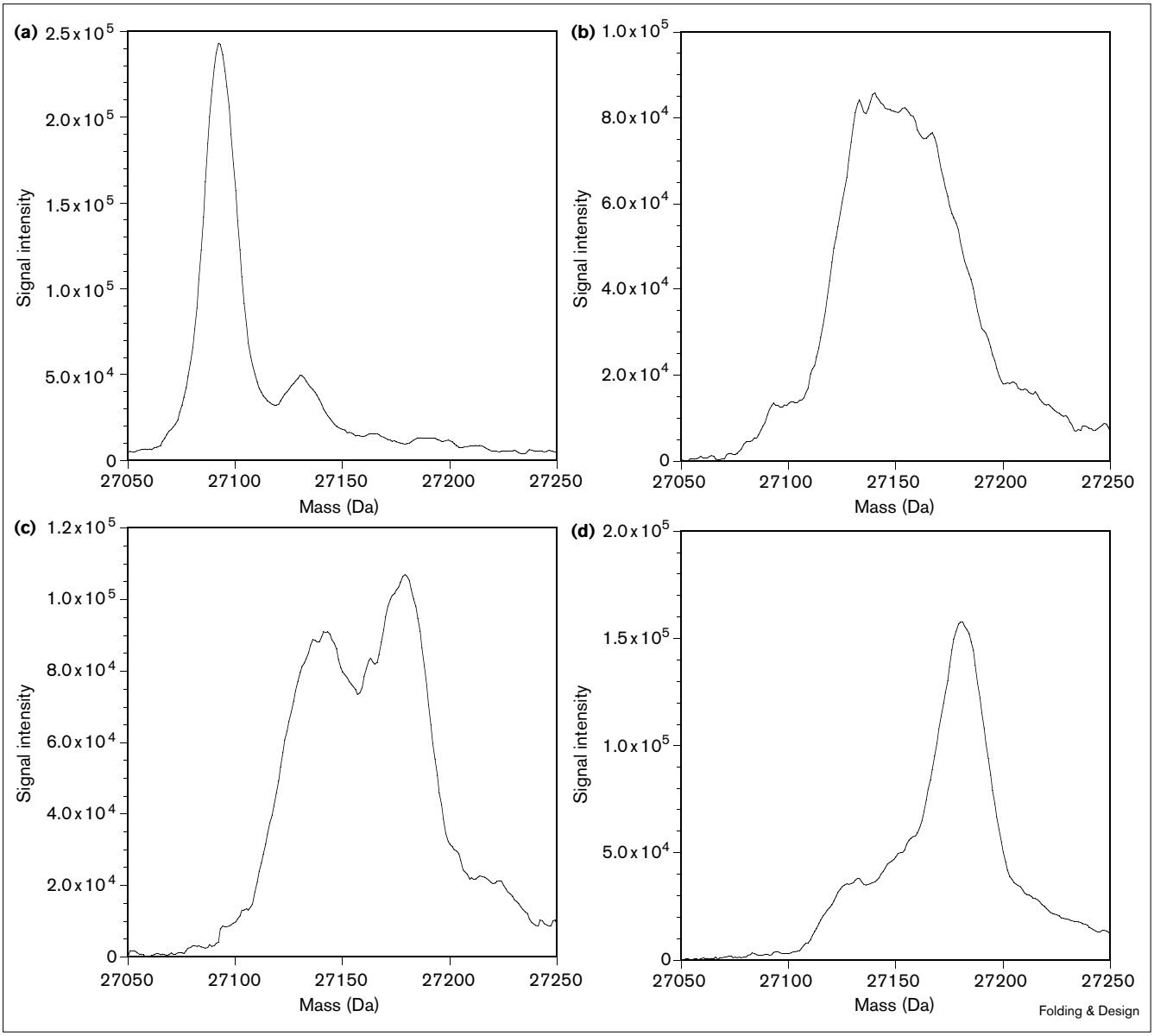
Mutant versus wild-type scFv fragment folding

The mutant used in this study, Pro40→Ala/Ser63→Ala/Ala64→Asp, was shown to result in higher expression yields in *Escherichia coli* than the wt scFv fragment [36]. In addition, the amount of scFv fragment obtained by preparative *in vitro* folding was also shown to be significantly enhanced for the mutant protein (C.F. and A.P., unpublished observations). Because the thermodynamic stability was shown to be the same within experimental error for wt and mutant protein [36], and because the thermal aggregation rate was higher for the wt protein, it was concluded that a difference in the rate of off-pathway reactions might be primarily responsible for different *in vivo* yields. We therefore wanted to compare the *in vitro* folding behavior of the two proteins in greater detail.

Because we wanted to be able to detect even small changes in the rate constants for the proton protection during the folding reaction, we measured both proteins under exactly the same conditions. This can be achieved by MS analysis of H/D exchange protection during folding if the mass difference of the two proteins is large enough to allow complete separation of the peaks but still allows the detection of the same number of charged species within a certain m/z range. Because the wt and mutant scFv fragment have almost identical molecular weight (1 Da difference), ^{15}N-labeled protein was used for the mutant in the experiments. A mass spectrum of a mixture of the equilibrium-denatured wt and mutant scFv fragment is shown in Figure 4a and shows that the peaks are well separated. The masses of the protonated reference samples correspond well with the theoretical masses. During the whole folding reaction, a total of 76 protons become protected. As can be seen from Figure 4, the protection proceeds with a higher rate for the mutant protein, for which the fully protected species is exclusively populated after 30 min, whereas there is still a significant amount of intermediate species for the wt protein. We therefore conclude that the three mutations in the V_H domain affect the net rate of secondary structure stabilization.

In addition, H/D exchange of the short-term denatured wt scFv fragment was monitored by NMR. The decay of NMR signal intensity of representative amide protons with the time of refolding is shown in Figure 5. As for the mutant protein (Figure 2), the protection of all residues shows a biphasic behavior. A close inspection of the degree of protection after 10 s of refolding reveals that the proton occupancies are very similar for the two proteins. What is different, however, is the second, slow phase of proton protection, which appears significantly slower for the wt protein. This was independently confirmed by MS analysis of the H/D exchange protection during the refolding of the short-term-denatured proteins (data not shown). The three mutations within the V_H domain might affect the slow folding rate either by stabilization of the domain itself, thus leading to a higher frequency of productive encounters with V_L to form the correct interface, or else indirectly by changing the nature or population of intermediates along the slow folding pathway. Another possibility is that the wild-type protein forms some reversible aggregates during these folding experiments, as has been demonstrated for a number of larger proteins with slow

Figure 3



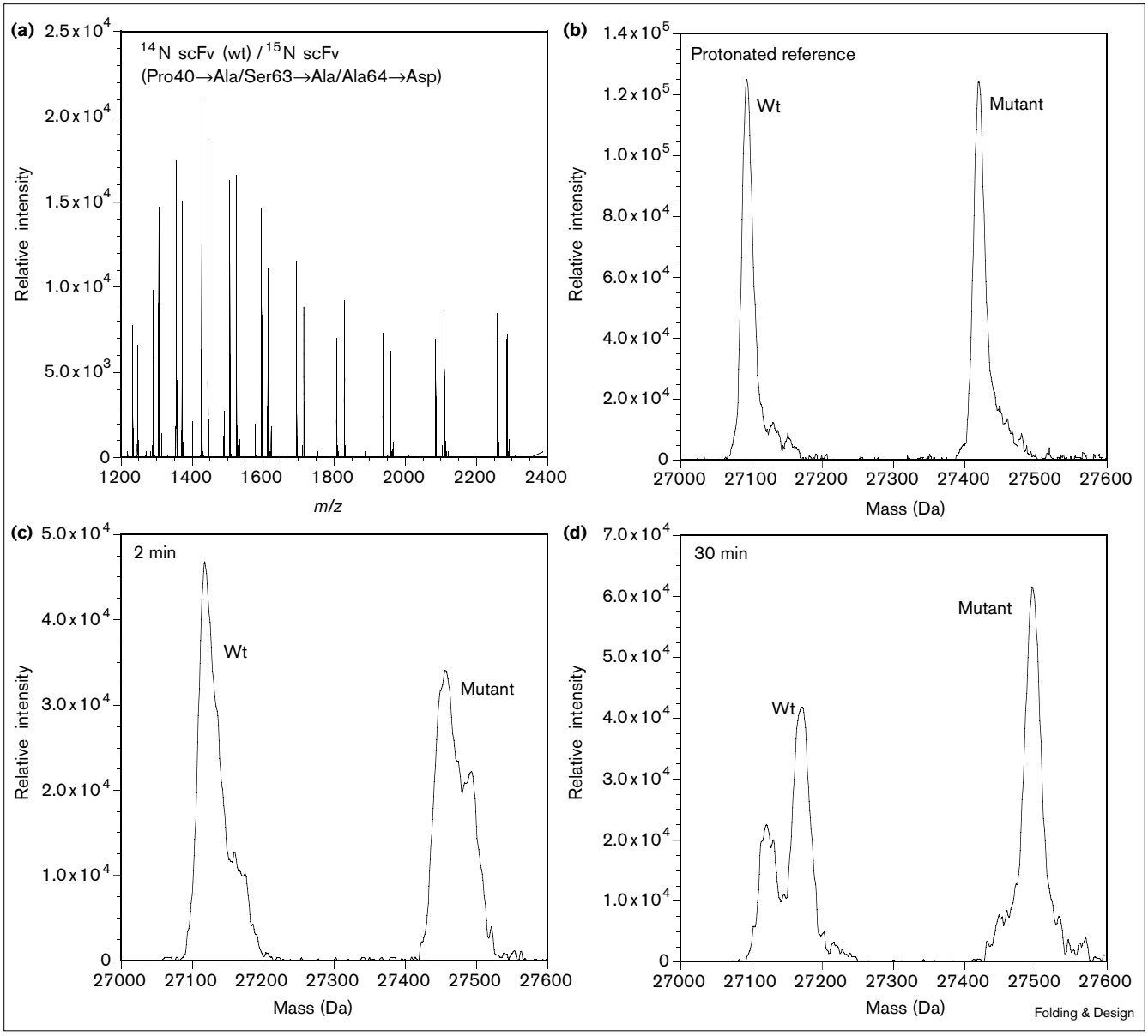
MS analysis of the short-term denatured Pro40→Ala/Ser63→Ala/Ala64→Asp mutant scFv fragment after various times of refolding in D_2O buffer. **(a)** Reconstructed hypermass of the protonated protein, and hypermass after **(b)** 10 s, **(c)** 2 min and **(d)** 10 min of refolding.

folding behavior [2,37]. Such aggregate formation would presumably slow down productive formation of the wt scFv fragment, whereas the mutant with the improved folding properties undergoes no such side reaction. Our results indicate, however, that these multimeric forms, if they exist, do not change the amide protection pattern of the initially formed folding intermediate.

Influence of cyclophilin on the rate of folding

Figure 6c shows the refolding reaction of the long-term denatured protein as measured by fluorescence at 331.4 nm, the wavelength maximum of the native scFv fragment in the presence of antigen. There is an initial loss of intensity, which cannot be resolved properly by manual mixing experiments. The subsequent gain of fluorescence can be fitted by a single exponential function ($k = 0.0008\ s^{-1}$), but the gain of fluorescence in the refolding of the short-term denatured protein becomes biphasic (Figure 6a) with rate constants of $0.0061\ s^{-1}$ and $0.00054\ s^{-1}$. Because the gain in fluorescence has been interpreted as formation of the native interface of the two variable domains [27,38], we suggest parallel pathways in the refolding of

Figure 4

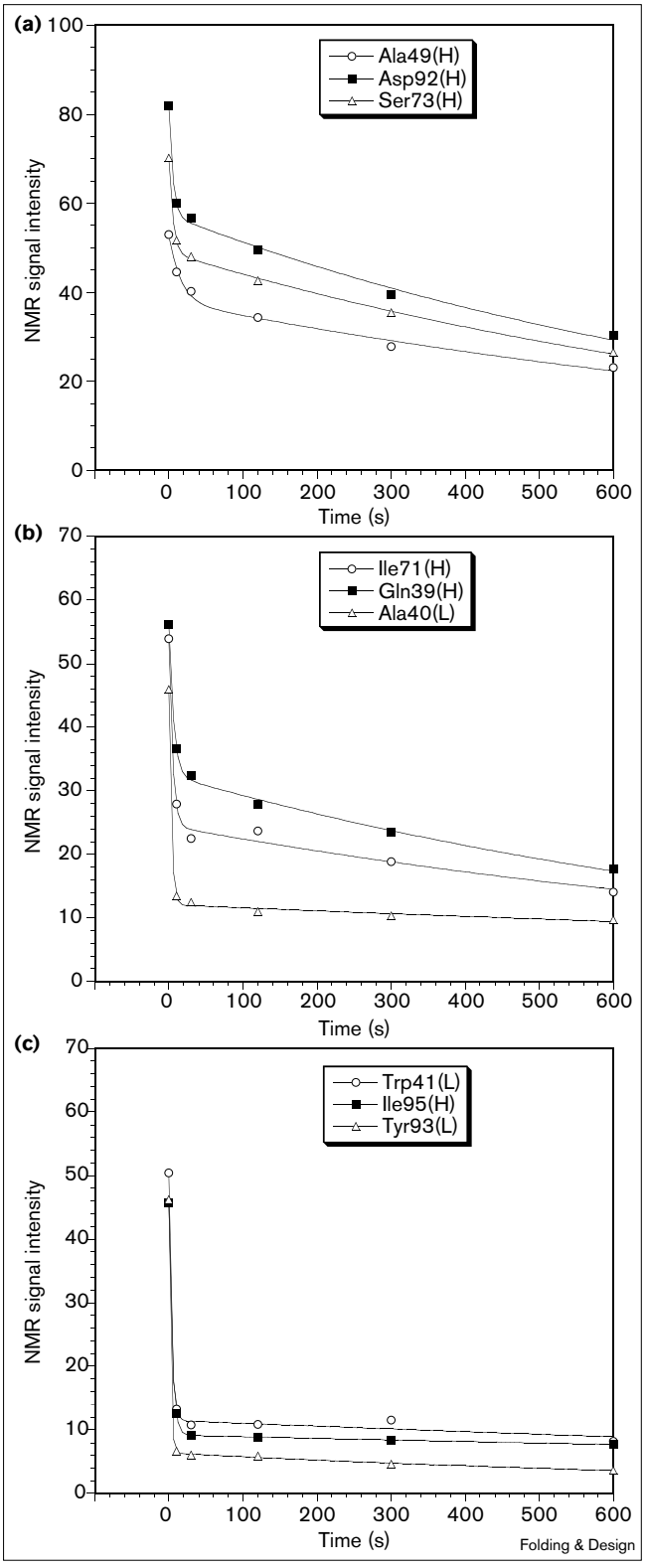


Analysis of H/D protection of the Pro40→Ala/Ser63→Ala/Ala64→Asp mutant in comparison to the wt scFv fragment by MS. **(a)** ^{15}N-labeled protein was used in the case of the mutant in order to achieve mass separation of both proteins; **(b)** the protonated reference samples; and **(c)** and **(d)** the hypermass of the species after 2 min and 30 min of refolding, respectively.

the short-term denatured scFv fragment to account for the observed biphasic behavior.

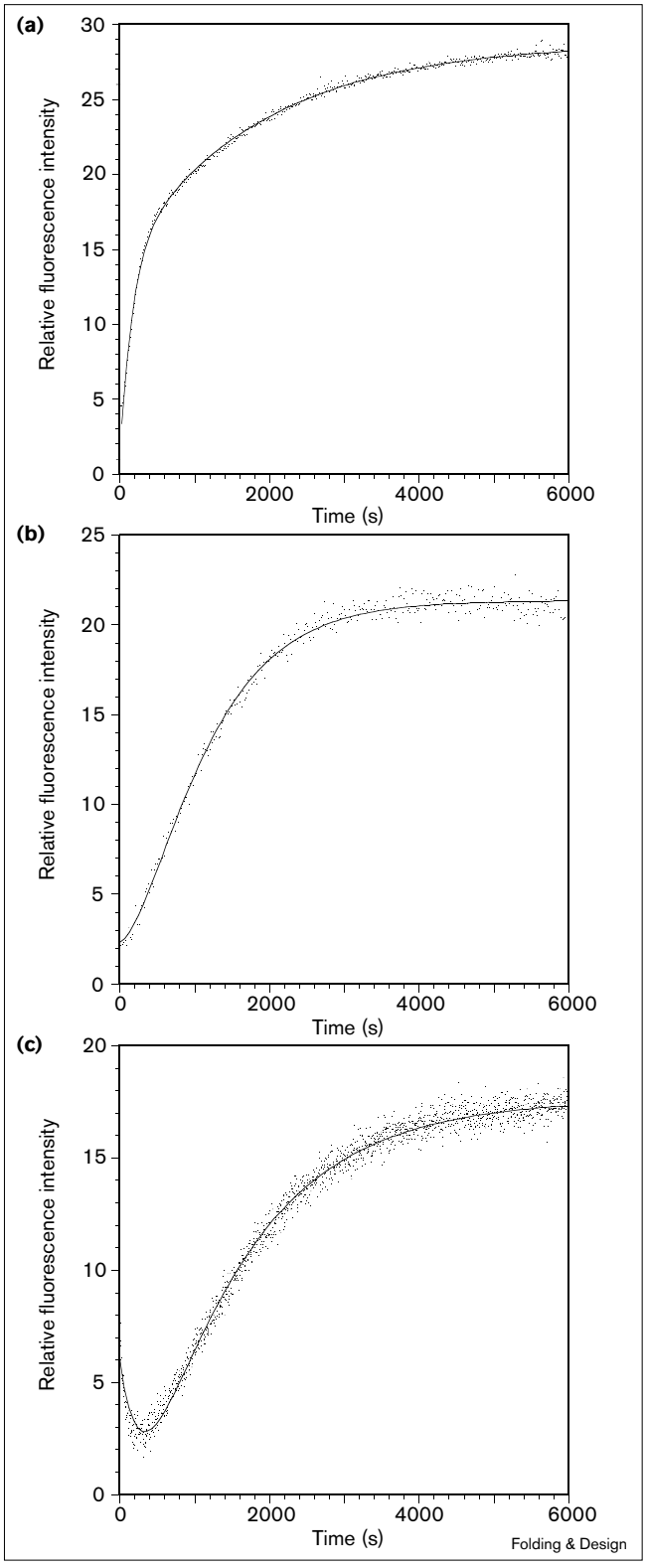
To test whether proline isomerization is involved in the folding of the short-term denatured protein, we added cyclophilin, a proline *cis–trans* isomerase, to the refolding buffer. In Figure 6b, the folding reaction of the short-term denatured scFv fragment in the presence of equimolar cyclophilin is shown. Interestingly, the gain of fluorescence becomes monophasic again ($k = 0.0013\ s^{-1}$), as it does for the long-term denatured protein (see Figure 6c). This suggests that the initially native-like prolines in the denatured scFv fragment isomerize to the 'wrong' conformer in competition with the refolding reaction and that this isomerization can be catalyzed efficiently by cyclophilin. A similar observation has been made for carbonic anhydrase [32] and a Fab fragment [39]. It has been proposed by Kern *et al.* [32] that the isomerization of prolines within an intermediate, rather than in the unfolded state, competes with the refolding reaction. We also examined these two

Figure 5



Plots of NMR signal intensity versus refolding time in D_2O for representative residues of the wt scFv fragment. Double exponential fits are shown as lines. See text for more details.

Figure 6



Fluorescence traces of the short-term denatured wt scFv fragment in **(a)** the absence and **(b)** the presence of a 2:1 excess of cyclophilin. **(c)** Refolding of the long-term denatured protein for comparison [25].

possibilities: on the one hand, the length of the denaturation time was altered, without affecting the amplitudes of the two phases or the rates significantly (data not shown); and on the other hand, when cyclophilin was added after the initiation of folding, the faster of the two phases could still be abolished to a large degree (data not shown). In addition, the effect of cyclophilin could be totally overcome by the addition of the inhibitor cyclosporin A to the refolding buffer, excluding non-specific interactions of cyclophilin with the scFv fragment to be the reason for our observations. We therefore conclude that proline isomerization takes place within a folding intermediate of the short-term denatured scFv fragment and competes with the refolding reaction. Because there are two *cis* prolines in the V_L domain of the native scFv fragment, it is most probable that one or both of these two prolines are involved in this isomerization reaction.

Discussion

The results of the current study show that the stabilization of secondary structure during the folding of the scFv fragment of an antibody after short-term denaturation can proceed along two parallel pathways. Measurements of the intrinsic tryptophan fluorescence of the short-term denatured protein during the folding reaction also reveals the existence of fast folding and slow folding species, although the rate constants are smaller than the rates obtained from the H/D experiments (0.06 s^{-1} versus 0.006 s^{-1} and 0.0013 versus 0.00054 s^{-1}). This difference at first appears to indicate that secondary structure formation precedes the formation of the native interface (which the fluorescence measurements are indicative of), but caution has to be taken to avoid overinterpretation of the rates of the H/D exchange experiments because they rely on a relatively small number of points. The H/D exchange experiments confirm the existence of an early folding intermediate, with an enhanced degree of amide-proton protection within certain regions of the β sheets [27,28], very similar to the results from the refolding of long-term denatured protein.

Upon addition of equimolar cyclophilin to the refolding buffer of the short-term denatured protein, the faster of the two phases of fluorescence gain vanishes completely. The enzyme obviously catalyzes a proline isomerization reaction that competes with the correct folding of the fast folding species. We suggest that the fast folding molecules of the uncatalyzed reaction refold from the early formed intermediate with all prolines in a native-like conformation. We further propose that the enzyme acts on one of the two native-like *cis* prolines (or both) in the early folding intermediate rather than the unfolded state, because the enzyme's action could still be observed if it was added after the start of the refolding reaction.

For the slowly folding molecules, no significant rate enhancement was observed in the presence of equimolar amounts of cyclophilin. In contrast, the slow folding of the isolated V_L domain, which carries two *cis* prolines (whereas V_H has none [29]), can be increased in the presence of equimolar amounts of PPI (peptidyl-prolyl *cis-trans* isomerase; [27]). Only at high concentrations of PPI is a rate acceleration seen in the refolding of both short-term denatured and long-term denatured scFv fragment [38]. We therefore propose that in the folding of the scFv fragment non-native domain association takes place after back-isomerization of native-like *cis* prolines. As a consequence, at least one proline becomes poorly accessible for the enzyme in this domain-associated intermediate. Because folding from this intermediate state with non-native proline isomers takes much longer than the folding from the earlier intermediate state with native-like prolines, we consider the former to be a kinetic trap along the folding pathway of the scFv fragment. At equimolar concentrations, cyclophilin efficiently catalyzes the formation of this trapped intermediate, but not its conversion to the native state. When equilibrium unfolded scFv is refolded, the future *cis* prolines of the V_L domain are largely in the non-native *trans* conformation and therefore escape the action of the enzyme due to non-native interface formation.

Early domain association of immunoglobulin domains has also been demonstrated for the Fab fragment of an antibody [39] and the rate-limiting step was shown to be proline isomerization within the association complex. Early formation of a non-native interface therefore seems to be a more general feature for the folding of larger fragments of immunoglobulins *in vitro*, especially if the effective concentration of domains is enhanced by covalent linkage.

Most strikingly, our results suggest that early domain association can hinder the fast folding of a two-domain protein. As a consequence of the formation of wrong proline isomers, an intermediate is formed that probably has an interface with (partly) non-native interactions and that prevents independent domain folding. Because similar rate constants for the slow phase of amide protection are observed for the residues of the two domains, folding from this kinetically trapped intermediate to the native state seems to be a committed step. In contrast to a truly cooperative folding behavior, however, in which the rates of folding for the scFv fragment should be enhanced as a result of favorable interactions of the individual domains, the folding process is in fact slowed down by the covalent linking of domains in the scFv fragment as a result of allowed unfavorable interactions early in the process. Similar observations of inhibitory domain interactions during folding have been made for phosphoglycerate kinase [40]. In contrast, single-chain Arc repressor folds faster than the non-covalent dimer by increasing the effective concentration of subunits [41].

The situation is different *in vivo* because the heavy and light chains are synthesized separately and interact with

chaperones like BiP and Grp94 [42]. It is therefore probable that the interaction of the two variable domains takes place from a native or native-like conformation of the individual domains. Our results suggest that for the fast formation of native antibodies it might be crucial that the interaction of V_L and V_H does not take place early during the folding of the individual domains. The interaction of the two native domains is in the micromolar range [43] and is made up primarily by hydrophobic interactions. The involvement of many hydrophobic sidechains may allow many energetically favorable, but non-native, interactions at the interface of the two domains and therefore slow down the folding rates. These non-native interactions probably compete with the correct folding of the single domains.

The formation of an exchange-protected intermediate is very fast [28] and, as our results suggest, is independent of the conformation of the prolines in the folding intermediate. The same protection pattern is observed in the dead time of manual mixing in this study as for folding studies of the long-term denatured protein [27,28]. Protection of amide protons of hydrophobic residues mainly located in the inner β sheets of the individual domains might therefore be an inherent feature of the single domains. On the other hand, formation of stabilized secondary structure within those β sheets that make up the future interface might facilitate the formation of stable, but kinetically trapped, intermediates, as observed for the scFv fragment investigated in this study.

Materials and methods

Protein preparation

The wt and mutant scFv fragments were obtained by refolding from inclusion bodies in *E. coli* as described before [44].

Fluorescence measurements

Fast unfolding kinetics were followed by stopped-flow fluorescence, using an SF61 spectrofluorimeter (HI-TECH Scientific, Salisbury, UK) equipped with a 0.25 ml and a 5.0 ml syringe, to produce a fixed dilution of 1:20. The excitation wavelength was set to 280 nm and the emission light passed through a 320 nm cut-off filter (WG-320, HIGH-TECH Scientific). The native protein in 50 mM borate, 20 mM NaCl, 1 mM phosphorylcholine, pH 8.0 was filled into the small volume syringe, and denaturant solutions containing 20 mM borate and different amounts of GdnCl were filled into the second syringe. The experiments were carried out at 10°C. The final protein concentration was 2 μM. The experimental curves were fitted by one or two exponential functions with the Kaleidagraph software on a Macintosh computer.

The fluorescence traces of the folding of short-term denatured scFv fragment were measured on a PTI machine with the excitation wavelength at 295 nm and the emission set to 331.4 nm, the wavelength maximum of the native protein. Native scFv fragment at a concentration of 100 μM was diluted 10-fold into 6.6 M GdnCl, stored for 20 s on ice and the mixture was further diluted 50-fold into the pre-cooled refolding buffer within the fluorescence cuvette containing 35 mM HEPES, 50 mM NaCl, pH 8.0. In a second experiment, an equimolar amount of human cyclophilin was added to the refolding buffer. Experiments were performed at 10°C.

H/D exchange experiments

For each time point of the H/D exchange kinetics purified scFv fragment (5–6 mg ^{15}N-labeled protein for the NMR analysis or a mixture of 0.5 mg of ^{14}N-labeled wt-scFv and 0.5 mg ^{15}N-labeled mutant scFv fragment for the MS analysis) in 50 mM borate, 20 mM NaCl, 1 mM phosphorylcholine, pH 8.0 was denatured by a 10-fold dilution into 6.6 M GdnCl, 20 mM borate, $pD_{read} = 8.0$, in D_2O for 20 s on ice. To initiate folding, the protein solution was then diluted 1:100 into 0.4 M arginine, 1 mM phosphorylcholine, $pD_{read} = 8.0$, in D_2O. After defined periods of time, a further 1:6.25 dilution into an H_2O buffer containing 0.12 M potassium phosphate, 1 mM phosphorylcholine, pH 4.0 was performed. The final pH was 5.1. For the protonated reference sample, the first dilution was done into H_2O buffer at pH 8.0 to allow all amide deuterons to exchange. The reaction was then allowed to proceed to completion (4 h at 10°C). After concentration of the solution using an Amicon A8200 cell, the protein was dialyzed against 5 mM KH_2PO_4 buffer for the NMR analysis and against 10 mM ammonium acetate buffer, pH 5.2, for the MS analysis. For the NMR analysis, the samples were then lyophilized and dissolved in D_2O 12 h prior to acquisition of the data in order to get rid of the fast exchanging amide protons. For the MS experiments, the dialyzed samples were further concentrated with a Centricon-10 concentrator and stored on ice before the measurement.

MS analysis

ESI-MS was used to determine the molecular mass of the scFv samples, which were injected into the ion source of a Sciex API III$^+$ instrument (Sciex, Ontario, Canada). The ion spray voltage was ~5000 V and the nebulizer gas pressure was 50 psi. The lyophilized samples were dissolved in ice-cold 10 mM ammonium acetate buffer, pH 5.0 and were mixed in a 1:1 ratio with an ice-cold mixture of 49.5% methanol, 49.5% H_2O, 1% formic acid to result in a final pH of ~3.0. 5 μl of the mixture were flow-injected into a cooled carrier solution, consisting of the abovementioned methanol/H_2O/formic acid mixture. The pH of 3.0 and the use of ice-cold solutions throughout the experiment was shown to minimize back-exchange of the protein samples upon denaturation within the mass spectrometer [28]. A *m/z* range of 1200–2400 was scanned with a step size of 0.15.

NMR spectra

All spectra were recorded on a Bruker DRX 600 spectrometer equipped with a Z-gradient unit. The temperature was 27°C and the samples were dissolved in D_2O. Relative protein concentrations were determined by comparing the methyl resonances at −1.0 ppm in 1D spectra recorded with 64 scans. ^{15}N–^{1}H correlation spectra were recorded using a gradient enhanced version of a HSQC experiment [45]. The carrier was positioned on the water resonance. Residual water suppression was achieved by the application of a WATERGATE 3–9–19 refocusing pulse [46] with a pulse interval of 200 ms in the final BACK INEPT step, allowing optimal inversion of the amide resonances. The gradient strengths were 25%, 10% and 80% of a maximum gradient power of 30 G/cm for the gradients G_1, G_2 and G_3, respectively. The delay for transferring proton magnetization to nitrogen in the INEPT step was set to 2.25 ms. Decoupling during aquisition was achieved using the GARP pulse sequence [47]. To obtain phase sensitive spectra, the TPPI method was used [48]. A data matrix of $4K \times 128$ points was acquired. Data were processed to a final size of $2K \times 256$ points using the in-house written software CC-NMR [49].

For each peak analyzed in the HSQC spectra, the corresponding peak in the reference spectrum was chosen to represent 100% proton occupancy. The intensities of the peaks were determined by measuring the peak heights because the line widths of individual peaks did not change in the spectra of different samples. Errors, which were estimated from the standard deviation of noise in 'empty' portions of the spectrum, were < 10%. Rate constants for the individual peaks were determined by applying double exponential fits to the plots of signal intensities (measured as peak heights) versus time of folding (until the dilution step into H_2O, see above), leaving out the 0 time point because the formation of an intermediate leads to a jump in signal reduction after 10 s for some protons, independent of the subsequent folding reactions.

Acknowledgements

We would like to thank Sabine Nieba-Axmann for providing human cyclophilin and Marcus Jäger for helpful discussions. This work was supported by the Swiss National Fund grant 3100-046624.96/1.

References

1. Jaenicke, R. (1987). Folding and association of proteins. *Prog. Biophys. Mol. Biol.* **49**, 117-237.
2. Jaenicke, R. (1995). Folding and association versus misfolding and aggregation of proteins. *Phil. Trans. R. Soc. Lond. Ser. B* **348**, 97-105.
3. Jackson, S.E. & Fersht, A.R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-10435.
4. Alexander, P., Orban, J. & Bryan, P. (1992). Kinetic analysis of folding and unfolding the 56 amino acid IgG-binding domain of streptococcal protein G. *Biochemistry* **31**, 7243-7248.
5. Viguera, A.R., Martinez, J.C., Filimonov, V.V., Mateo, P.L. & Serrano, L. (1994). Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* **33**, 2142-2150.
6. Huang, G.S. & Oas, T.G. (1995). Structure and stability of monomeric lambda repressor: NMR evidence for two-state folding. *Biochemistry* **34**, 3884-3892.
7. Schindler, T., Herrler, M., Marahiel, M.A. & Schmid, F.X. (1995). Extremely rapid protein folding in the absence of intermediates. *Nat. Struct. Biol.* **2**, 663-673.
8. Brandts, J.F., Halvorson, H.R. & Brennan, M. (1975). 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.
9. Creighton, T.E. (1984). Disulfide bond formation in proteins. *Methods Enzymol.* **107**, 305-329.
10. Seckler, R. & Jaenicke, R. (1992). Protein folding and protein refolding. *FASEB J.* **6**, 2545-2552.
11. Beasty, A.M., Hurle, M.R., Manz, J.T., Stackhouse, T., Onuffer, J.J. & Matthews, C.R. (1986). Effects of the phenylalanine-22→leucine, glutamic acid-49→methionine, glycine-234→aspartic acid, and glycine-234→lysine mutations on the folding and stability of the alpha subunit of tryptophan synthase from Escherichia coli. *Biochemistry* **25**, 2965-2974.
12. Kotik, M., Radford, S.E. & Dobson, C.M. (1995). Comparison of the refolding of hen lysozyme from dimethyl sulfoxide and guanidinium chloride. *Biochemistry* **34**, 1714-1724.
13. Sosnick, T.R., Mayne, L., Hiller, R. & Englander, S.W. (1994). The barriers in protein folding. *Nat. Struct. Biol.* **1**, 149-156.
14. Oliveberg, M. & Fersht, A.R. (1996). Thermodynamics of transient conformations in the folding pathway of barnase: reorganization of the folding intermediate at low pH. *Biochemistry* **35**, 2738-2749.
15. Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl Acad. Sci. USA* **92**, 9029-9033.
16. Silow, M. & Oliveberg, M. (1997). Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl Acad. Sci. USA* **94**, 6084-6086.
17. Miranker, A.D. & Dobson, C.M. (1996) Collapse and cooperativity in protein folding. *Curr. Opin. Struct. Biol.* **6**, 31-42.
18. Alzari, P.M., Lascombe, M.B. & Poljak, R.J. (1988). Three-dimensional structure of antibodies. *Annu. Rev. Immunol.* **6**, 555-580.
19. Goto, Y., Azuma, T. & Hamaguchi, K. (1979). Refolding of the immunoglobulin light chain. *J. Biochem. (Tokyo)* **85**, 1427-1438.
20. Lilie, H., Jaenicke, R. & Buchner, J. (1995). Characterization of a quaternary-structured folding intermediate of an antibody Fab-fragment. *Protein Sci.* **4**, 917-924.
21. Goto, Y. & Hamaguchi, K. (1982). Unfolding and refolding of the constant fragment of the immunoglobulin light chain. *J. Mol. Biol.* **156**, 891-910.
22. Tsunenaga, M., Goto, Y., Kawata, Y. & Hamaguchi, K. (1987). Unfolding and refolding of a type kappa immunoglobulin light chain and its variable and constant fragments. *Biochemistry* **26**, 6044-6051.
23. Freund, C., Ross, A., Plückthun, A. & Holak, T.A. (1994). Structural and dynamic properties of the Fv fragment and the single-chain Fv fragment of an antibody in solution investigated by heteronuclear three-dimensional NMR spectroscopy. *Biochemistry* **33**, 3296-3303.
24. Miranker, A., Robinson, C.V., Radford, S.E., Aplin, R.T. & Dobson, C.M. (1993). Detection of transient protein folding populations by mass spectrometry. *Science* **262**, 896-900.
25. Roder, H., Elove, G.A. & Englander, S.W. (1988). Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* **335**, 700-704.
26. Udgaonkar, J.B. & Baldwin, R.L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* **335**, 694-699.
27. Freund, C., Honegger, A., Hunziker, P., Holak, T.A. & Plückthun, A. (1996). Folding nuclei of the scFv fragment of an antibody. *Biochemistry* **35**, 8457-8464.
28. Freund, C., Gehrig, P., Holak, T.A. & Plückthun, A. (1997). Comparison of the H/D exchange behavior of the rapidly formed folding intermediate and the native state of the scFv fragment of an antibody. *FEBS Lett.* **407**, 42-46.
29. Satow, Y., Cohen, G.H., Padlan, E.A. & Davies, D.R. (1986). Phosphocholine binding immunoglobulin Fab McPC603. *J. Mol. Biol.* **190**, 593-604.
30. Bai, Y., Milne, J.S., Mayne, L. & Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75-86.
31. Connelly, G.P., Bai, Y., Jeng, M.F. & Englander, S.W. (1993). Isotope effects in peptide group hydrogen exchange. *Proteins* **17**, 87-92.
32. Kern, G., Kern, D., Schmid, F.X. & Fischer, G. (1995). A kinetic analysis of the folding of human carbonic anhydrase II and its catalysis by cyclophilin. *J. Biol. Chem.* **270**, 740-745.
33. Gratwohl, C. & Wüthrich, K. (1981). NMR studies of the rates of proline cis-trans isomerization in oligopeptides. *Biopolymers* **20**, 2623-2632.
34. Raleigh, D.P., Evans, P.A., Pitkeathly, M. & Dobson, C.M. (1992). A peptide model for proline isomerism in the unfolded state of staphylococcal nuclease. *J. Mol. Biol.* **228**, 338-342.
35. Thunecke, A., Kálmán, A., Kálmán, F., Ma, S., Rathore, A.S. & Horváth, C. (1996). Kinetic study on the cis-trans isomerization of peptidyl-proline peptides. *J. Chromatogr. A* **744**, 259-272.
36. Knappik, A. & Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its in vivo folding. *Protein Eng.* **8**, 81-89.
37. Pecorari, F., Minard, P., Desmadrils, M. & Yon, J.M. (1996). Occurrence of transient multimeric species during the refolding of a monomeric protein. *J. Biol. Chem.* **271**, 5270-5276.
38. Jäger, M. & Plückthun, A. (1997). The rate-limiting steps for the folding of an antibody scFv fragment. *FEBS Lett.*, in press.
39. Lilie, H., Rudolph, R. & Buchner, J. (1995). Association of antibody chains at different stages of folding: prolyl isomerization occurs after formation of quaternary structure. *J. Mol. Biol.* **248**, 190-201.
40. Parker, M.J., *et al.*, & Clarke, A.R. (1996). Domain behavior during the folding of a thermostable phosphoglycerate kinase. *Biochemistry* **35**, 15740-15752.
41. Robinson, C.R. & Sauer, R.T. (1996). Equilibrium stability and sub-millisecond refolding of a designed single-chain arc repressor. *Biochemistry* **35**, 13878-13884.
42. Melnick, J., Dul, J.L. & Argon, Y. (1994). Sequential interaction of the chaperones BiP and GRP94 with immunoglobulin chains in the endoplasmic reticulum. *Nature* **370**, 373-375.
43. Glockshuber, R., Malia, M., Pfitzinger, I. & Plückthun, A. (1990). A comparison of strategies to stabilize immunoglobulin Fv fragments. *Biochemistry* **29**, 1362-1367.
44. Freund, C., Ross, A., Guth, B., Plückthun, A. & Holak, T.A. (1993). Characterization of the linker peptide of the single-chain Fv fragment of an antibody by NMR spectroscopy. *FEBS Lett.* **320**, 97-100.
45. Mori, S., Abeygunawardana, C., Johnson, M.O. & van Zijl, P.C. (1995). Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn. Reson. B* **108**, 94-98.
46. Sklenar, V., Peterson, R.D., Rejante, M.R. & Feigon, J. (1993). Two- and three-dimensional HCN experiments for correlating base and sugar resonances in ^{15}N,^{13}C-labeled RNA oligonucleotides. *J. Biomol. NMR* **3**, 721-727.
47. Shaka, A.J., Barker, P.B. & Freeman, R. (1985). Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.* **64**, 547-552.
48. Marion, D. & Wüthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967-974.
49. Cieslar, C., Ross, A., Zink, T. & Holak, T.A. (1993). Efficiency in multi-dimensional NMR by optimized recording of time point-phase pairs in evolution periods and their selective linear transformation. *J. Magn. Reson. B* **101**, 97-101.