α-Lactalbumin forms a compact molten globule in the absence of disulfide bonds

Christina Redfield[1], Brenda A. Schulman[2,3], Michael A. Milhollen[2], Peter S. Kim[2] and Christopher M. Dobson[1]

[1]Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, UK. [2]Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA. [3]Present address: Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, Box 576, New York, New York 10021, USA.

Human α-lactalbumin (α-LA) is a four disulfide-bonded protein that adopts partially structured conformations under a variety of mildly denaturing conditions. At low pH, the protein is denatured but compact, with a high degree of secondary structure and a native-like fold. This is commonly referred to as a molten globule. A variant of α-LA, in which all eight cysteines have been mutated to alanine (all-Ala α-LA), has been studied using NMR spectroscopy. At low pH all-Ala α-LA is nearly as compact as wild type α-LA. Urea-induced unfolding experiments reveal that the residues that remain compact in the absence of disulfide bonds are those that are most resistant to unfolding in the wild-type α-LA molten globule. This is particularly remarkable because this stable core is formed by segments of the polypeptide chain from both the N- and C-termini. These results show that the overall architecture of the protein fold of α-LA is determined by the polypeptide sequence itself, and not as the result of cross-linking by disulfide bonds, and provide insight into the way in which the sequence codes for the fold.

α-Lactalbumin (α-LA) is a 14 kDa protein whose native structure is divided into two domains, one is largely helical (the α-domain) and the other has a significant content of β-sheet (the β-domain)[1]. At low pH this protein forms a compact globular structure (the A-state) that lacks the specific side chain packing that is characteristic of native structures. This is commonly termed a molten globule[2-5]. A number of studies have shown that the residual structure in the α-LA molten globule has extensive native-like character. This is particularly evident in the α-domain of the protein where the native-like helices are all present and arranged in a manner that bears a close relationship to that found in the native structure[3,6–14].

α-LA contains four disulfide bonds, two in the α-domain, one in the β-domain and one cross-linking the two domains (Fig. 1*a*). In order to understand the extent to which topological constraints imposed by these disulfide bonds influence the compact, native-like structure found in the α-LA molten globule, we have studied a variant of α-LA in which all eight cysteines have been replaced with alanine (all-Ala α-LA)[9]. Our approach is based on NMR experiments used previously to characterize the wild type protein at low pH[12]. The principle of this approach is that the NMR resonances of residues located in compact regions of the low pH A-state are extremely broad due to conformational fluctuations on a millisecond time scale that result from local side chain disorder[6,12,14,15]. By exposing the partially folded protein to increasingly destabilizing conditions, the progressive unfolding of different regions of the structure can be followed by the appearance of sharp, well-resolved resonances in the NMR spectrum[12]. Comparison of the behavior of wild type and all-Ala α-LA will then reveal the effects of removal of the disulfide bonds on the stability of the various structural elements.

All-Ala α-LA is helical and compact

The circular dichroism (CD) spectrum of all-Ala α-LA, at pH 2 and 20 °C, is similar to that of wild-type α-LA at pH 2 (refs 3,9). Thus, like the wild-type protein at pH 2 all-Ala α-LA has a native-like helical content but lacks a close-packed environment for the aromatic side chains characteristic of a native protein. All-Ala α-LA, like other α-LA variants lacking the 61–77 and 73–91 disulfides, does not adopt a native structure at any pH value[13]. As the concentration of urea is increased, the ellipticity at 222 nm decreases markedly for all-Ala α-LA, until at 8 M urea it

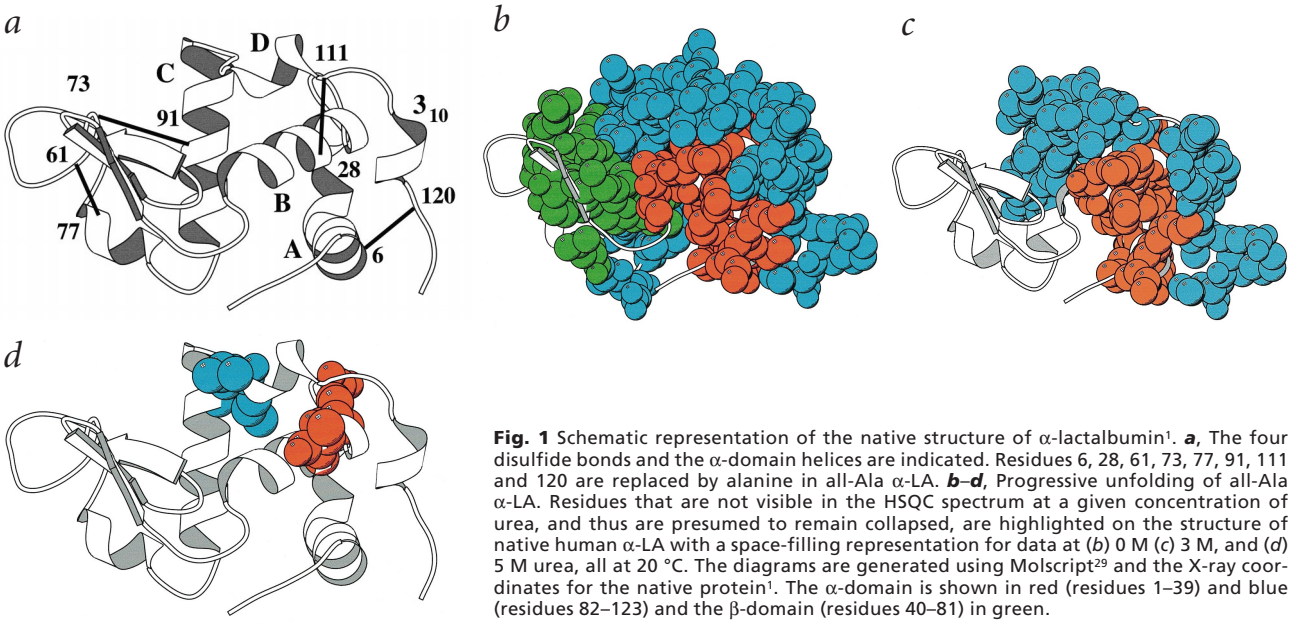

Fig. 1 Schematic representation of the native structure of α-lactalbumin[1]. ***a***, The four disulfide bonds and the α-domain helices are indicated. Residues 6, 28, 61, 73, 77, 91, 111 and 120 are replaced by alanine in all-Ala α-LA. ***b–d***, Progressive unfolding of all-Ala α-LA. Residues that are not visible in the HSQC spectrum at a given concentration of urea, and thus are presumed to remain collapsed, are highlighted on the structure of native human α-LA with a space-filling representation for data at (*b*) 0 M (*c*) 3 M, and (*d*) 5 M urea, all at 20 °C. The diagrams are generated using Molscript[29] and the X-ray coordinates for the native protein[1]. The α-domain is shown in red (residues 1–39) and blue (residues 82–123) and the β-domain (residues 40–81) in green.

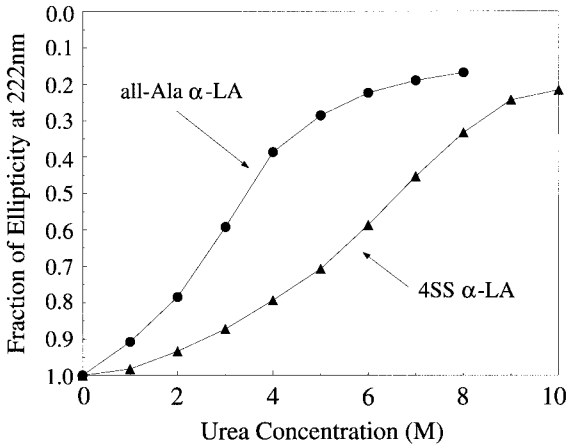

Fig. 2 Urea-induced unfolding of the helical secondary structure in all-Ala α-LA (circles) and wild type α-LA (triangles) at pH 2 and 20 °C, as monitored by CD at 222 nm.

is only ~15% of that observed in the absence of urea (Fig. 2). The apparent mid-point of unfolding occurs at ~3 M urea, showing a significant resistance to the loss of helical character. The helical structure is, however, disrupted substantially more readily than for wild type α-LA where the mid-point of unfolding is ~6 M urea[3,12].

Despite the lack of close packing of its side chains, the molten globule formed by wild type α-LA at pH 2 is remarkably compact; experiments suggest that its radius of gyration is less than 10% greater than that of the native state[3,16]. To examine the role of disulfide bonds on compactness we have measured the hydrodynamic radius of all-Ala and wild type α-LA using NMR diffusion methods[17] (Table 1). For native α-LA, at pH 6.2 and 20 °C, the experimental hydrodynamic radius is 19.7 Å, corresponding to a radius of gyration, R_g, of 15.2 Å, in good agreement with the value of 15.7 Å measured previously by X-ray scattering[18]. The hydrodynamic radius of the A-state of wild type α-LA, at pH 2 and 20 °C, is found to be 20.9 Å, a value only 6% larger than that of the native protein. At pH 2 all-Ala α-LA has a hydrodynamic radius of 21.7 Å, within 10% of that of the native protein. It is clear, therefore, that all-Ala α-LA is nearly as compact as the wild type α-LA molten globule. In 8 M urea all-Ala α-LA has a hydrodynamic radius of 33.3 Å, a value that is consistent with a prediction of 35.1 Å for a fully unfolded protein containing 124 amino acids (Wilkins, D.K., Grimshaw, S.B., Receveur, V., Smith, L.J. & C.M.D., unpublished data). Therefore, it is the intrinsic properties of the polypeptide sequence, and not the presence of disulfide bonds, that gives rise to a high degree of compactness of the α-LA molten globule.

All-Ala α-LA has a native-like core

The ^{15}N-^{1}H HSQC spectrum[19] of all-Ala α-LA at pH 2 and 20 °C (Fig. 3*a*) contains cross peaks from only 37 of the 121 possible backbone amides. The spectrum of wild type α-LA recorded under the same conditions contains only three cross peaks (Fig. 3*g*)[12]. The absence of cross peaks in the spectra of the A-state can be attributed to exchange effects resulting from the location of residues in compact regions of the structure in which fluctuations between different conformational states are slow[6,12,15]. Although the NMR data indicate that all-Ala α-LA has less of its chain in compact structure than the wild type protein, more than two-thirds of the polypeptide chain is involved in such structure. In order to identify the regions of the protein that are sufficiently unfolded in all-Ala α-LA to give rise to sharp resonances, and by inference the regions that remain collapsed, the peaks observed in the spectra of all-Ala α-LA were assigned using 3D NMR[12,19].

The most intense peaks observed in Fig. 3*a* arise from several regions of the protein sequence. Peaks arise from the three N-terminal residues, from residues that link the B-helix in the α-domain to the β-sheet region (36–39), and from residues that in the native structure are located in the β-domain (62–79). A group of lower intensity peaks is also seen in the HSQC spectrum; these arise mainly from residues of the β-domain. No peaks are observed from residues located in the native structure in the B, C, D or 3_{10}-helices or from residues in the C-terminal region of the protein although the latter are not located in well defined secondary structure in the native state. These C-terminal residues do, however, interact closely with residues from the A, B and D helices in the native state. Taken together, these results strongly suggest that a compact native-like core region (Fig. 1*b*), corresponding to a major section of the α-domain, persists in all-Ala α-LA. By contrast much of the β-domain is less structured although not completely disordered.

In order to understand how this compact structure is stabilized in all-Ala α-LA, the number of peaks corresponding to unfolded regions of the protein was monitored as a function of urea (Fig. 3). The urea concentration at which a peak is first observed for each residue of all-Ala α-LA is shown in Fig. 4, along with similar data for wild type α-LA at pH 2 (ref. 12). In accord with the CD results, all-Ala α-LA can be seen from the NMR data to be much less resistant to unfolding by urea at pH 2 than the wild type protein (Fig. 3*h*). The most notable difference in the unfolding behavior of all-Ala and wild type α-LA at pH 2 is seen in the vicinity of residues 61, 73 and 77. In wild type α-LA these are the β-domain cysteine residues involved in two disulfide bonds, 61–77 and 73–91; these three residues, and their neighbors, are the β-domain residues most resistant to unfolding in wild-type α-LA. In all-Ala α-LA these residues are completely unfolded in the absence of urea. Thus, the disulfide bonds in the β-domain do restrict the structure in this domain. It is interesting to note that residue 91 is much more resistant to unfolding than 61, 73 and 77, in all-Ala α-LA. Thus, although removal of the 73–91 disulfide bond significantly destabilizes compact structure in the vicinity of residue 73 in the β-domain there are other, more important, factors that maintain the stability of compact structure surrounding residue 91 in the α-domain.

Despite the difference in overall stability, there are marked similarities between the unfolding events in the molten globule states

Table 1 NMR diffusion data for wild type and all-Ala α-LA

	Conditions	$D_{diox}/D_{\alpha\text{-LA}}$[1]	R_g (Å)[2]	R_s (Å)[3]
wild type α-LA	pH 6.2	9.12 ± 0.03	15.2 ± 0.1	19.7 ± 0.1
wild type α-LA	pH 2.0	9.69 ± 0.10	16.2 ± 0.2	20.9 ± 0.2
all-Ala α-LA	pH 2.0	10.06 ± 0.06	16.8 ± 0.1	21.7 ± 0.1
all-Ala α-LA	pH 2.0, 8M urea	15.43 ± 0.11	(25.8 ± 0.2)[2]	33.3 ± 0.2

[1] $D_{diox}/D_{\alpha\text{-LA}}$ is the ratio of the measured diffusion coefficients of 1,4-dioxan and α-LA.
[2] R_g is the radius of gyration of the protein. This is calculated as the product of $D_{diox}/D_{\alpha\text{-LA}}$ and the R_g of dioxan (1.7 Å). R_g is not a meaningful parameter for an unfolded protein; a value is given in parenthesis for comparison.
[3] R_s is the effective hydrodynamic radius defined as $R_g = (3/5)^{1/2} R_s$.

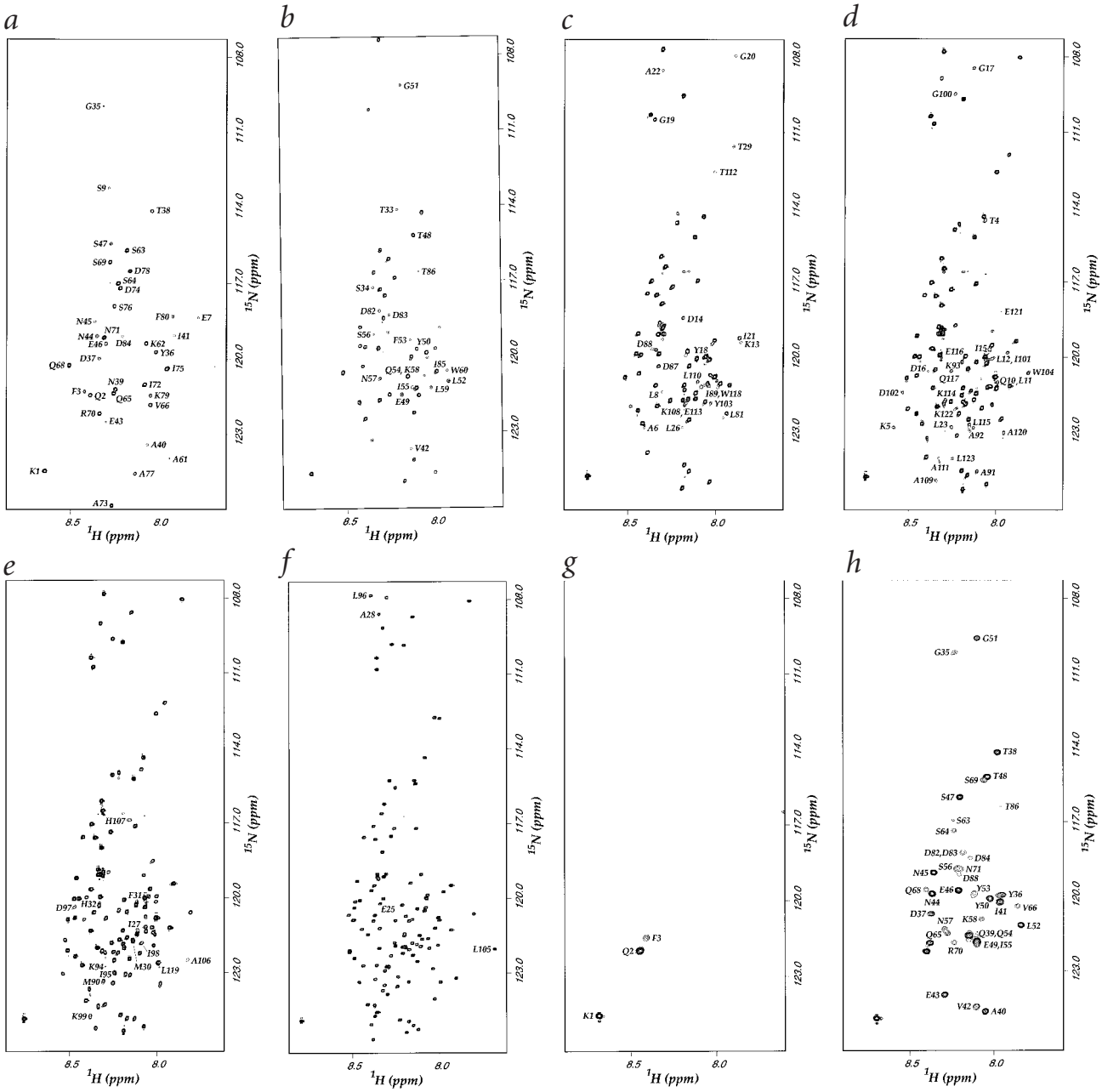

Fig. 3 Appearance of NMR resonances as all-Ala α-LA is unfolded. HSQC spectra are shown for all-Ala α-LA at pH 2 in ***a***, the absence of urea, ***b***, 2 M urea, ***c***, 3 M urea, ***d***, 4 M urea, ***e***, 5 M urea, and ***f***, 7 M urea, all at 20 °C. Peaks are labeled with their residue assignment as they become visible in the spectra. Peaks corresponding to Ala 22, Ala 28, Ala 73 (not in (*a*)) and Leu 96 are folded in the ^{15}N dimension. HSQC spectra of wild-type α-LA at pH 2 in ***g***, the absence of urea, and ***h***, 7 M urea are shown for comparison.

of wild type and all-Ala α-LA. In both proteins a urea concentration can be identified at which the complete central region of the polypeptide chain, corresponding to the β-domain, is unfolded but where both parts of the chain that form the α-domain in the native structure remain compact (Fig. 4). This phenomenon occurs at 2 M urea for all-Ala α-LA, and at 10 M urea and 40 °C for wild type α-LA; if residues surrounding Cys 61, Cys 73 and Cys 77 are ignored in wild type α-LA, this occurs at 9 M urea. It indicates that the helical core of the α-domain is retained in the absence of all four native disulfide bonds (Fig. 1*b*). This is particularly remarkable because this domain is formed by the N- and C- terminal regions of the polypeptide chain which in the native state are linked by two disulfide bonds, 6–120 and 28–111 which are, of course, absent in all-Ala α-LA.

Local cooperativity contributes to core stability

In all-Ala α-LA cross peaks from the four α-domain helices of the native structure become observable in the NMR spectra over a relatively narrow range of urea concentrations (Figs 1*c*,*d*, 4). This suggests that the stable core of the helical domain unfolds with a degree of at least local cooperativity. This contrasts with the results for wild type α-LA[12] where the majority of peaks from residues in the C-helix in the native structure become visible in the spectrum under substantially less destabilizing conditions

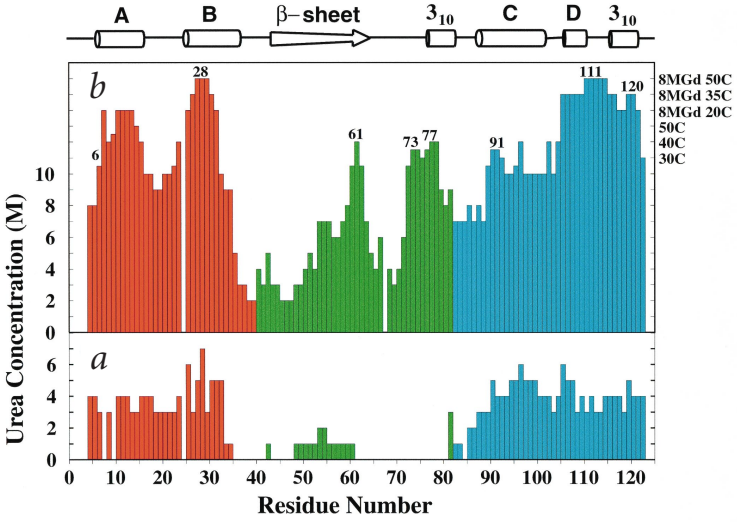

Fig. 4 Unfolding behavior of ***a***, all-Ala α-LA and ***b***, wild type α-LA at pH 2. The bars indicate the urea concentration at which an HSQC peak is first observed. In (*b*) the scale on the right indicates the more destabilising conditions required to unfold wild type α-LA, 10 M urea at 30, 40 and 50 °C and 8 M guanidine HCl at 20, 35 and 50 °C. The secondary structure found in native α-LA is summarized above.

than peaks from residues located in the other helices of the native structure. Even in the wild type protein, however, the final unfolding step, involving residues located in helices at both the N-and C-termini, appears to occur in a relatively concerted manner. This indicates that a significant number of residues must interact in order to stabilize a compact state. In all-Ala α-LA the data suggest that, in the absence of the stabilizing effect of the disulfide bridges, residues from all the major helices in the α-domain of the protein are required in order to stabilize a compact structure relative to the entropically favored fully unfolded state. It is well established that a major factor in the stabilizing effect of disulfide bridges in native proteins is that they reduce significantly the entropy of the unfolded state relative to the folded state[20]. The present data suggest that a similar effect stabilizes partially folded structures relative to fully unfolded ones.

The region of the molten globule that is collapsed in all-Ala α-LA in the absence of urea, corresponds very closely to the helical domain in the structure of native α-LA (Fig. 1*b*). In α-LA, the helical domain is not contiguous, with the central segment of the polypeptide chain forming the distinct β-domain. Because of the high entropic penalty for ordering the ends of a chain, it is remarkable that it is the helical domain that is folded in the α-LA molten globule. The finding that in all-Ala α-LA this structure is formed even in the absence of disulfide cross-links indicates that the amino-acid sequences in these regions of the structure have a high affinity for each other. Both the N- and C-terminal helices are highly amphipathic, and the burial of hydrophobic residues will induce helical structure in these amphipathic sequences. It is likely, therefore, that long-range interactions between the chain termini in α-LA are a key feature in the establishment of the globular core of the protein.

Native disulfide bond formation

In previous studies the propensity of native and non-native disulfide bonds to form during oxidative folding has been used to probe the topology of the α-LA molten globule[9,11]. Of the six possible disulfide bonds that can be formed between the four cysteines in the helical domain of α-LA, only the native 28–111 disulfide has a propensity to form that is significantly higher than the statistical probabilities predicted by a random walk model or measured in strong denaturant[11]. Even in the absence of any disulfide bonds, the region around residues 28 and 111 is extremely stable within the compact structure of all-Ala α-LA (Fig. 4). The four residues most resistant to unfolding are located in close proximity in the α-domain in the native fold (Fig. 1*d*). These observations are readily rationalized provided that the structure within the compact state is native-like. If this is the case, then the intrinsic interactions within the polypeptide chain bring these regions together, increasing the effective concentration of the cysteine residues and favoring formation of the disulfide linkage[21–23].

Recently, mutation to alanine of residues that form the hydrophobic core in native α-LA was shown to decrease the effective concentration of the 28–111 disulfide bond[24,25]. These residues, Leu 8, Leu 12, Ile 21, Leu 23, Leu 26, Ile 27, Met 30 and Trp 118, are all highly resistant to unfolding by urea in all-Ala α-LA. This supports the notion that the stability of the reduced form of the protein influences the effective concentration of the 28–111 disulfide bond. It is interesting in this regard that the effective concentration of 28–111 is markedly increased in a mutant of the α-domain in which every hydrophobic group has been replaced with leucine[26]. This strongly supports the idea that the structure of the compact region of the molten globule state does not involve specific side chain interactions, because the replacement of the native residues by leucines will increase their average hydrophobicity but is likely to reduce significantly the complementary packing interactions characteristic of the native sequence.

Thus, the 28–111 disulfide bond is formed in a region of the structure in which the intrinsic stability of long-range interactions is high. The 6–120 disulfide bond is also within the α-domain of the protein, and it is likely that its formation will be aided by the intrinsic stability of this region of the structure detected in the present study. In accord with this the effective concentration of the 6–120 disulfide bond, while ~10-fold lower than the enhancement for the 28–111 disulfide, is ~100-fold higher under native compared to denaturing conditions. It is interesting, however, that three of the other cysteine residues, 61, 73 and 77, are in a region of the polypeptide chain that is among the least structured in all-Ala α-LA (Fig. 4). In accord with this, the effective concentrations of the 61–77 and 73–91 disulfide bonds are not significantly different from those expected in a random chain[11]. It is clear that these disulfide bonds are essential to the stability of the native fold in the β-domain, as calcium-dependent tight side chain packing does not occur in variants of α-LA lacking these disulfide bonds[13,27]. There is relatively little entropic penalty in forming these disulfide links and it is tempting to speculate that their role is to stabilize the β-domain and its link to the α-domain in the absence of strong intrinsic interactions in this domain whose hydrophobic core is small relative to the α-domain. Thus, the data suggest that the character of the different disulfide bonds, and their role in stabilizing the fold of the protein may differ significantly.

How the sequence codes for the fold

The results described in this paper demonstrate that the topological constraints provided by disulfide bonds are not the source of

the native-like three-dimensional fold of α-LA. Rather, the data demonstrate that the latter arises because of the intrinsic stability of interactions between regions of the polypeptide chain that are distant from each other in the sequence. Indeed, the strongest interactions are between residues at the two ends of the polypeptide chain and the weakest involve those in the central region of the chain that are closest together in sequence. These strong interactions are likely to be associated with the complementary matching of hydrophobic and hydrophilic residues. Moreover, formation of such interactions can stabilize, in a synergistic manner, elements of secondary structure, for example in the case of amphipathic helices, and satisfy the electrostatic requirements of a polypeptide chain in a hydrophobic environment generated by the formation of a compact state[28]. The correct formation of other elements of structure, particularly β-strands where the correct alignments are essential in order to form native-like sheets, is likely to be aided substantially by the template formed by a helical core region such as that observed here. The role of disulfide bonds appears to be to stabilize the overall fold, at least in part through entropic factors, and to constrain regions of the chain where the intrinsic non-covalent interactions between residue side chains are less stable.

Methods

NMR spectroscopy. ^{15}N-labeled all-Ala α-LA was expressed and purified as described previously for related proteins[9,10]. NMR samples contained 0.4 mM protein at pH 2 in 95% H_2O/5% D_2O. 750 MHz 2-D gradient-enhanced ^{15}N-^{1}H HSQC spectra[17] consisting of 128 complex t_1 increments of 2K complex data points were collected for urea concentrations ranging from 0 M to 8 M. Sweep widths of 1,441 Hz and 8,130 Hz were used in the ^{15}N (F_1) and ^{1}H (F_2) dimensions, respectively. A total of 96 scans were collected for each t_1 increment. Resonances were assigned as described[12] using 3D NOESY-HSQC spectra[17] collected in 4 M, 6 M and 8 M urea at 20 °C. The HSQC spectra shown in Fig. 3 are contoured so that a cross peak defined by a single contour has ~25% of the intensity of a cross peak corresponding to a fully unfolded residue.

CD spectroscopy. Far-UV CD measurements for all-Ala α-LA at 222 nm were carried out at 20 °C using a Jasco J720 CD spectrophotometer . The protein concentration was ~20 μM and a cuvette of 0.1 cm path length was used. The reported ellipticities are the average of 300 measurements collected over a 5 min period. The measured ellipticities at each urea concentration were divided by the ellipticity in the absence of urea to obtain the fraction of ellipticity at 222 nm plotted in Fig. 2. The CD data for the wild-type protein are taken from a previous study[12].

NMR diffusion measurements. Diffusion measurements were carried out using the PG-SLED (pulse-gradient spin longitudinal echo diffusion) sequence. A small amount of 1,4-dioxan was added to each protein sample as the radius standard. A total of 20 spectra were collected with the gradient strength varying linearly between 5 and 100%. The diffusion coefficients for the protein and the dioxan were determined as described[15]. For all-Ala α-LA measurements were made with 0.4 mM samples in D_2O and 8 M urea, both at pH 2. For wild-type α-LA measurements were made at pH 6.2 on samples containing 1 mM, 1.5 mM and 3 mM protein, and at pH 2 on samples containing 0.32 mM, 0.65 mM and 1.3 mM protein. The reported values of $D_{diox}/D_{\alpha\text{-LA}}$ were obtained from an extrapolation to zero protein concentration.

Acknowledgments

We dedicate this paper to the late Oleg Ptitsyn. We thank Z.-y. Peng and L. Wu for helpful discussions. C.R. acknowledges the BBSRC for an Advanced Research Fellowship. B.A.S. was supported by an NIH training grant. This work was supported by the Howard Hughes Medical Institute and the Wellcome Trust. The Oxford Centre for Molecular Sciences is supported by the BBSRC, EPSRC and MRC.

Correspondence should be addressed to C.R. *email*: *redfield@bioch.ox.ac.uk*

Received 19 May, 1999; accepted 13 July, 1999.

1. Acharya, K.R., Ren, J.S., Stuart, D.I., Phillips, D.C. & Fenna, R.E. *J. Mol. Biol.* **221**, 571–581 (1991).
2. Ptitsyn, O.B. *Adv. Protein Chem.* **47**, 83–229 (1995).
3. Dolgikh, D.A. et al. O.B. *FEBS Lett.* **136**, 311–315 (1981).
4. Kuwajima, K. *Proteins* **6**, 87–103 (1989).
5. Kuwajima, K. *FASEB J.* **10**, 102-109 (1996).
6. Baum, J., Dobson, C.M., Evans, P.A. & Hanley, C. *Biochemistry* **28**, 7–13 (1989).
7. Chyan, C.L., Wormald, C., Dobson, C.M., Evans, P.A. & Baum, J. *Biochemistry* **32**, 5681–5691 (1993).
8. Peng, Z.-y. & Kim, P.S. *Biochemistry* **33**, 2136–2141 (1994).
9. Peng, Z.-y., Wu, L.C. & Kim, P.S. *Biochemistry* **34**, 3248–3252 (1995).
10. Schulman, B.A., Redfield, C., Peng, Z.-y., Dobson, C.M. & Kim, P.S. *J. Mol. Biol.* **253**, 651–657 (1995).
11. Wu, L.C., Peng, Z.-y. & Kim, P.S. *Nature Struct. Biol.* **2**, 281–286 (1995).
12. Schulman, B.A., Kim, P.S., Dobson, C.M. & Redfield, C. *Nature Struct. Biol.* **4**, 630–634 (1997).
13. Wu, L.C., Schulman, B.A., Peng, Z.-y. & Kim, P.S. *Biochemistry* **35**, 859–863 (1996).
14. Smith, L.J., Dobson, C.M. & van Gunsteren, W.F. *J. Mol. Biol.* **286**, 1567–1580 (1999).
15. Wang, Y. & Shortle, D. *Protein Sci.* **5**, 1898–1906 (1996).
16. Ptitsyn, O.B. In *Protein folding.* (ed. Creighton, T.E.) 243–300 (W.H. Freeman and Co., New York; 1992).
17. Jones, J.A., Wilkins, D.K., Smith, L.J. & Dobson, C.M. *J. Biomol. NMR* **10**, 199–203 (1997).
18. Kataoka, M., Kuwajima, K., Tokunaga, F. & Goto Y. *Protein Sci.* **6** 422–430 (1997).
19. Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. & Skelton, N.J. *Protein NMR spectroscopy*. (Academic Press, New York; 1996).
20. Betz, S.F. *Protein Sci.* **2**, 1551–1558 (1993).
21. Creighton, T.E. *Biopolymers* **22**, 49–58 (1983).
22. Goldenberg, D.P., Bekart, L.S., Laheru, D.A. & Zhou, J.D. *Biochemistry* **32**, 2835–2844 (1993).
23. Lin, T.-Y. & Kim, P.S. *Biochemistry* **28**, 5282–5287 (1989).
24. Song, J., Bai, P., Luo, L. & Peng, Z.-y. *J. Mol. Biol.* **280**, 167–174 (1998).
25. Wu, L.C., & Kim, P.S. *J. Mol. Biol.* **280**, 175–182 (1998).
26. Wu, L.C., & Kim, P.S. *Proc. Natl. Acad. Sci. USA* **94**, 14314–14319 (1997).
27. Ewbank, J.J. & Creighton, T.E. *Biochemistry* **32**, 3694–3707 (1993).
28. Dobson, C.M., Sali, A. & Karplus, M. *Angew. Chem. Int. Ed. Engl.* **37**, 868–893 (1998).
29. Kraulis, P. *J. Appl. Crystallogr.* **24**, 946–950 (1991).