Local and Global Cooperativity in the Human α-Lactalbumin Molten Globule

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · MAY 2004

Impact Factor: 4.33 · DOI: 10.1016/j.jmb.2004.02.045 · Source: PubMed

5 AUTHORS, INCLUDING:

Christina Redfield
University of Oxford
129 PUBLICATIONS **4,688** CITATIONS

SEE PROFILE

doi:10.1016/j.jmb.2004.02.045 J. Mol. Biol. (2004) 338, 149–158

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Local and Global Cooperativity in the Human α-Lactalbumin Molten Globule

Cindy M. Quezada[1]†, Brenda A. Schulman[2]†, Joanna J. Froggatt[1] Christopher M. Dobson[3] and Christina Redfield[1]*

[1]*Chemistry Research Laboratory, Oxford Centre for Molecular Sciences, University of Oxford, Mansfield Road Oxford OX1 3TA, UK*

[2]*Department of Biology Howard Hughes Medical Institute and Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology 9 Cambridge Center Cambridge, MA 02142, USA*

[3]*Department of Chemistry University of Cambridge Lensfield Road, Cambridge CB2 1EW, UK*

**Corresponding author*

NMR spectroscopy has been used to follow the urea-induced unfolding of the low pH molten globule states of a single-disulfide variant of human α-lactalbumin ([28–111] α-LA) and of two mutants, each with a single proline substitution in a helix. [28–111] α-LA forms a molten globule very similar to that formed by the wild-type four-disulfide protein, and this variant has been used as a model for the α-lactalbumin (α-LA) molten globule in a number of studies. The urea-induced unfolding behavior of [28–111] α-LA is similar to that of the four-disulfide form of the protein, except that [28–111] α-LA is less stable and has greater cooperativity in the loss of different elements of structure. For one mutant, L11P, the helix containing the mutation is highly destabilized such that it is completely unfolded even in the absence of urea. By contrast, for the other mutant, Q117P, the helix containing the mutation retains its compact structure. Both mutations, however, show significant long-range destabilization of the overall fold showing that the molten globule state has a degree of global cooperativity. The results reveal that different permutations of three of the four major α-helices of the protein can form a stable, locally cooperative, compact structural core. Taken together, these findings demonstrate that the molten globule state of α-LA is an ensemble of conformations, with different subsets of structures linked by a range of long-range interactions.

Keywords: human α-lactalbumin; molten globule state; proline mutagenesis; protein folding; NMR

Introduction

Molten globules are compact, partially folded proteins that have a native-like complement of secondary structure but lack fixed tertiary packing interactions.[1–4] For a number of proteins experiments have indicated that there is a close similarity between molten globules formed at equilibrium under mildly denaturing conditions and species formed during the early stages of refolding.[5–10] α-Lactalbumin 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) (Figure 1).[11] At low pH this protein forms a molten globule, the A-state. A number of studies have shown that the overall architecture of the α-LA molten globule has extensive native-like character, particularly in the α-domain where native-like helices are arranged in a manner that bears a close

† C.M.Q. and B.A.S. contributed equally to this work.

Present addresses: C. M. Quezada, Laboratory of Structural Microbiology, The Rockefeller University, 1230 York Avenue, Box 52, New York, NY 10021, USA; B. A. Schulman, Department of Structural Biology, St. Jude Children's Research Hospital, 322 North Lauderdale, Memphis, TN 38105, USA.

Abbreviations used: α-LA, human α-lactalbumin; 4-SS α-LA, α-LA containing the four native disulfide bonds; all-Ala α-LA, α-LA containing no disulfide bonds; [28–111] α-LA, α-LA containing a single disulfide bond between Cys28 and Cys111; L11P, [28–111] α-LA with a proline replacing a leucine at position 11; Q117P, [28–111] α-LA with a proline replacing a glutamine at position 117; NMR, nuclear magnetic resonance; HSQC, heteronuclear single quantum correlation; CD, circular dichroism.

E-mail address of the corresponding author: christina.redfield@bioch.ox.ac.uk

0022-2836/$ - see front matter

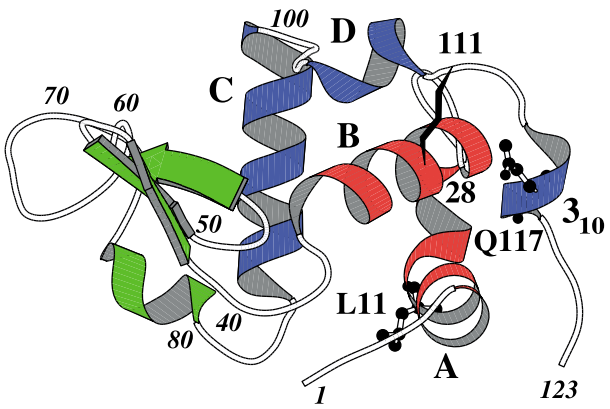

Figure 1. Schematic representation of the native structure of human α-lactalbumin. The 28–111 disulfide bond and the α-domain helices are indicated. L11 and Q117, the two residues mutated to proline, are shown in a ball-and-stick representation. In all-Ala α-LA residues 28 and 111 are replaced by alanine. The α-domain is shown in red (residues 1–39) and blue (residues 82–123) and the β-domain (residues 40–81) in green. The diagram was generated using MOLSCRIPT[48] and the X-ray coordinates for the native protein.[11]

relationship to that found in the native structure.[12–28] In particular, it is an important system through which to probe the nature of interactions that define the fold of a polypeptide main-chain prior to the close packing of the side-chains that results in the fully native structure.[4]

The NMR spectra of the molten globules formed by α-LA and a range of other proteins are characterized by broad resonances for most of the backbone amide groups.[12,19,29–32] In the case of α-LA this line broadening is interpreted as arising from constrained conformational fluctuations on the microsecond to millisecond time-scale.[12,19,33] We have found that the resonances broadened to such an extent that they are unobservable in $^{15}N-^{1}H$ 2D HSQC NMR spectra of human α-LA correspond to residues that are in compact folded regions of structure in the molten globule, whereas the sharp, well-resolved peaks observed in these spectra correspond to highly dynamic, largely unfolded parts of the protein.[19,24,26] The number of well-resolved resonances observed in the HSQC spectrum of the α-LA molten globule increases as the concentration of denaturant, such as urea or guanidine-HCl, is increased. The stepwise appearance of resonances is consistent with a non-cooperative denaturant-induced unfolding of the overall structure. Resonances appearing at the lowest concentrations of denaturant correspond to segments of the α-LA molten globule that are in less stable regions of the structure. The residues most resistant to denaturant are, in the native structure of α-LA, clustered together in the core of the protein, indicating that this is the most stable region of the molten globule.

Here we have examined the effects of proline mutations on the unfolding of individual residues in a variant of α-LA, termed [28–111] α-LA, which contains only the single disulfide bond between cysteine residues 28 and 111, with all other cysteine residues mutated to alanine (Figure 1). The features of the low pH α-LA molten globule are retained in [28–111] α-LA, even near neutral pH. Like the molten globule formed by wild-type α-LA containing four disulfide bonds, [28–111] α-LA is monomeric under a wide variety of conditions and is highly helical.[15] Other single-disulfide variants of α-LA have a significantly reduced helical content compared to [28–111] α-LA.[15] Moreover, the C28–C111 disulfide bond is the only disulfide bond with a propensity to form that is significantly higher than the statistical probability predicted by a random walk model or measured in a high concentration of denaturant.[15,17] Importantly, [28–111] α-LA is an excellent system for mutational studies because the correct disulfide bond is invariably formed under oxidizing conditions when only the two cysteine residues at positions 28 and 111 are present, even when folding is significantly disfavored.[18,21,22] By contrast, disrupting the structure in the presence of the four cysteine residues in the α-helical domain perturbs the ability to form native disulfide bonds.[18] Therefore, the use of the [28–111] α-LA species ensures that the correctly disulfide-bonded form of the protein is being studied. As well as the intrinsic interest in studying this species, a series of mutants of [28–111] α-LA has been described in which individual proline substitutions were introduced at positions within each of the five helices (four α and one 3_{10}) in the α-domain of the native structure.[18] For each helix, the effect on the measured circular dichroism (CD) signal of a single proline substitution at a number of sequence positions within the helix was similar, and the magnitude of the loss in helix content was approximately that expected if the proline substitution unfolded only the specific native-like helix containing the mutation.[18]

Here we describe the study of two of these mutant proteins using the $^{15}N-^{1}H$ HSQC method of probing the stability to urea-induced unfolding of individual regions of the molten globule structure. The two mutations, L11P in the A-helix and Q117P in the C-terminal 3_{10}-helix, were chosen for this study as they are in helices that are not linked by the C28–C111 disulfide bond, and also because the mutations cause similar losses of helix content relative to the wild-type [28–111] α-LA.[18] In addition, the mutations are in helices that have been found to differ significantly in their importance for forming native disulfide bonds. While mutations in the C-terminal 3_{10}-helix do not significantly affect the ability of a four-cysteine version of α-LA to form native disulfide bonds, mutations in the A-helix cause a dramatic reduction in the proportion of protein molecules containing native disulfide bonds.[18] Our results provide an explanation of these observations as the proline

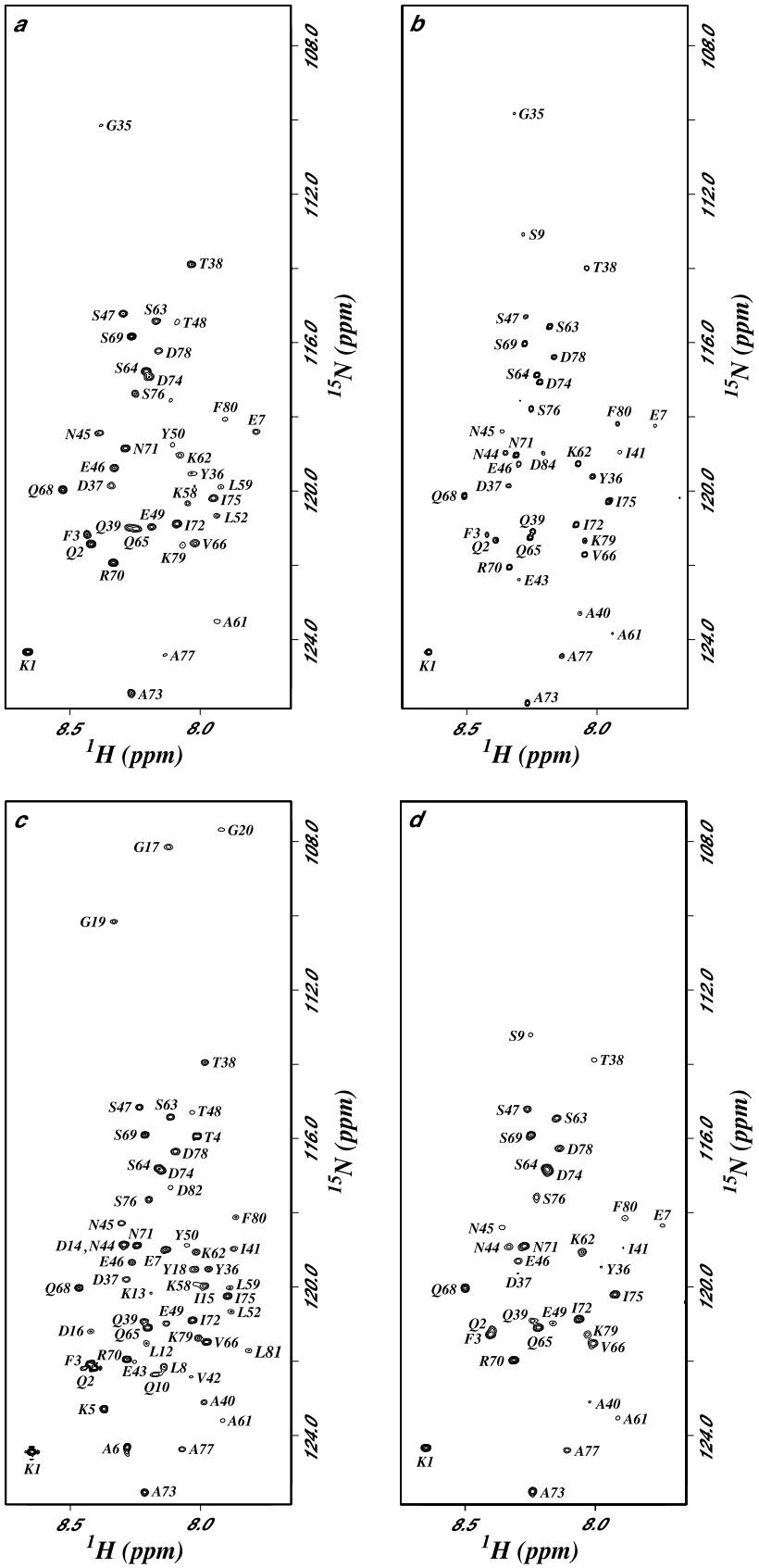

Figure 2. $^{15}N-^{1}H$ HSQC spectra of human α-lactalbumin variants at pH 2 and 20 °C in the absence of urea: a, [28–111] α-LA; b, all-Ala α-LA; c, L11P; d, Q117P. Peaks are labeled with their residue assignment. The spectra were contoured so that a cross-peak defined by a single contour has 20–25% of the intensity of a cross-peak corresponding to a fully unfolded residue.

mutations produce both long-range as well as short-range effects on the structure and stability of the α-LA molten globule structure.

Results

Effects on unfolding transitions for α-LA species with different numbers of disulfide bonds

The $^{15}N-^{1}H$ HSQC spectrum of [28–111] α-LA in the absence of urea at pH 2 and 20 °C (Figure 2a) contains cross-peaks from 37 of the 121 possible backbone amide groups. The spectrum of wild-type four-disulfide α-LA recorded under the same conditions contains only three detectable cross-peaks (Figure 1a of Schulman *et al.*[19]); these arise from the three N-terminal residues. The greater number of peaks in [28–111] α-LA indicates that the molten globule form of this protein contains a significantly greater number of residues that are unfolded than does that of the wild-type protein containing four-disulfide bonds. The HSQC spectrum of [28–111] α-LA is, however, very similar to that of all-Ala α-LA (Figure 2b), which differs from [28–111] α-LA only in the mutations of C28 and C111 to alanine that removes the one remaining native-like disulfide bond.[24]

The majority of resonances observed in [28–111] α-LA and all-Ala α-LA arise from residues that in the native structure are located in the β-domain; none of these resonances is observed in the spectrum of the four-disulfide protein. Thus, in [28–111] α-LA, as previously observed for all-Ala α-LA, the removal of the β-domain disulfide bond (C61–C77) and the disulfide bond linking the β-domain to the C-helix (C73–C91) leads to a loss of the stability of the collapsed structure in the β-domain of the protein.[24] There are, however, some differences between the residues giving rise to observable resonances in [28–111] α-LA and all-Ala α-LA. In the former, resonances from residues 9, 40, 41, 43, 44 and 84, which are observed in all-Ala α-LA, are not detected. In the latter species, resonances from residues 48–50, 52 and 59, which are all resolved in the spectrum of [28–111] α-LA, are not observed. Although the majority of the resonances arising from these two groups of residues are of very low intensity in the spectra, this suggests that there are subtle differences in the stability of the collapsed regions of the β-domain in the presence and absence of the C28–C111 disulfide bond located in the α-domain.

The unfolding of [28–111] α-LA was monitored as a function of urea concentration. The urea concentration at which a peak is first detected for each residue of [28–111] α-LA is shown in Figure 3a, along with similar data obtained previously for all-Ala α-LA[24] (Figure 3b) and wild-type α-LA (Figure 3e) at pH 2.[19] It can be seen from these plots that, although the extent of structure present in the molten globule states of these two proteins is similar, that of [28–111] α-LA is significantly more resistant to urea-induced unfolding than that of all-Ala α-LA; this result is consistent with CD measurements performed as a function of urea concentration.[24,34] In addition, the unfolding of all-Ala α-LA occurs over a smaller range of urea concentrations than that of [28–111] α-LA and appears to be somewhat more cooperative, a result consistent with conclusions by Luo & Baldwin[34] based on CD spectra. Figure 3 shows that for the α-domain residues of [28–111] α-LA unfolding occurs at urea concentrations ~3–5 M higher than required for all-Ala α-LA. By contrast, the effect of the presence of the C28–C111 disulfide bond on the stability of the β-domain is much smaller, although not negligible. The pattern of unfolding of [28–111] α-LA, however, shows much larger differences from that of the four-disulfide wild-type α-LA, particularly in the β-domain region of the sequence (residues 40–80). The latter observation can be attributed to the removal of the β-domain disulfide bond (C61–C77) and the disulfide bond linking the β-domain to the C-helix (C73–C91), as discussed for all-Ala α-LA.[24] In the α-domain of the protein there is a lower, albeit general, decrease in stability in [28–111] α-LA compared to wild-type α-LA, that can be associated primarily with the removal of the 6–120 disulfide bond. For [28–111] α-LA, unfolding of all regions of the protein, except for residues directly adjacent to C28 and C111, occurs at urea concentrations of 8–10 M while in four-disulfide α-LA the majority of residues in the native A, B, D and 3_{10} helices are not unfolded even in 10 M urea at 50 °C.[19]

The unfolding of [28–111] α-LA appears to be more cooperative than that of the four-disulfide protein, a result similar to that observed for all-Ala α-LA.[24] Thus, an increase in the number of disulfide bonds seems to be correlated with a decrease in the cooperativity of denaturant-induced unfolding, perhaps reflecting the need for a greater number of non-covalent interactions to stabilize a compact state in the absence of covalent cross-links.[24,34] Inspection of Figure 3 indicates that there is also a difference in the unfolding patterns of the N and C termini of the molten globule forms of wild-type four-disulfide α-LA and [28–111] α-LA. In the four-disulfide species, cross-peaks from residues 4 and 5 appear at a lower urea concentration than the peaks of other residues in the A-helix. By contrast, in [28–111] α-LA, resonances from residues 4 and 5 are observed at the same concentration of urea required to observe resonances from the remainder of the A-helix. Similar results are seen for residues 120–123 at the C terminus. It seems likely that in the absence of the C6–C120 disulfide bond, the A and 3_{10}-helices are more extended than in its presence. Indeed, at pH 2, [28–111] α-LA appears to have a higher helical content than the wild-type four-disulfide species as monitored by the ellipticity at 222 nm.[14,15] This

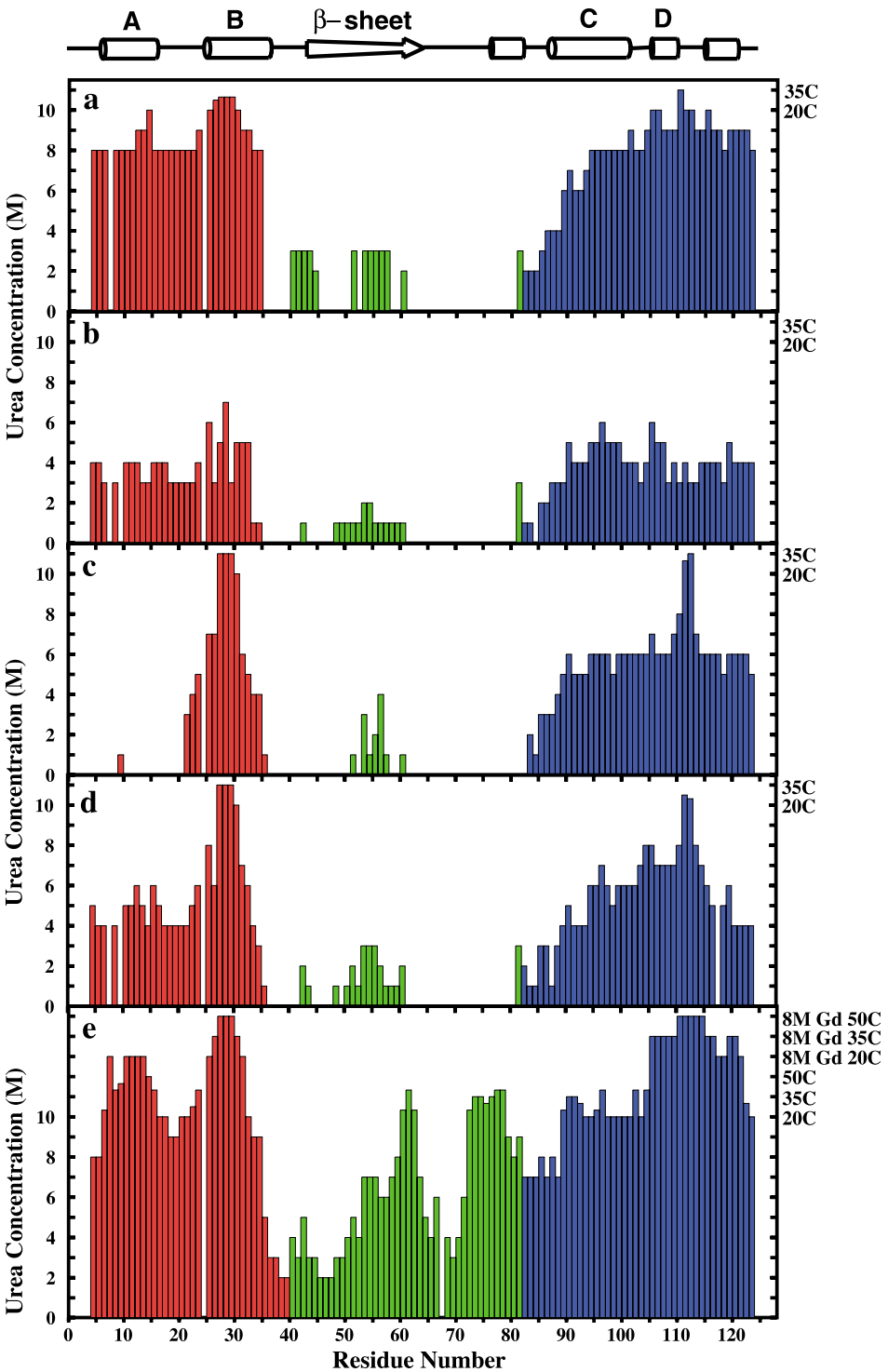

Figure 3. Unfolding behavior of a, [28–111] α-LA; b, all-Ala α-LA; c, L11P; d, Q117P; and e, 4-SS α-LA at pH 2 and 20 °C. The bars indicate the urea concentration at which an HSQC peak is first observed for each residue. The scale on the right indicates the higher temperatures required to unfold some residues in 10 M urea and the more strongly denaturing conditions of 8 M guanidine-HCl required to unfold some residues of the four-disulfide protein. The secondary structure found in native α-LA is summarized above. The α-domain is shown in red (residues 1–39) and blue (residues 82–123) and the β-domain (residues 40–81) in green.

may be accounted for by an increase in the lengths of the A-helix at the N terminus and the 3_{10}-helix at the C terminus. The C6–C120 disulfide bond has been shown to be strained in the structure of native α-LA.[15,17,35] A similar pattern of unfolding of the N and C termini is observed in all-Ala α-LA, which also lacks the C6–C120 disulfide bond.[24]

The effects of proline substitutions on the unfolding of [28–111] α-LA

The $^{15}N-^{1}H$ HSQC spectrum of the L11P mutant of [28–111] α-LA at pH 2 and 20 °C in the absence of urea is shown in Figure 2c. This spectrum contains substantially more cross-peaks than does that of [28–111] α-LA; 57 of the 120 possible backbone amide groups are seen for L11P in contrast to only 37 peaks for [28–111] α-LA. The additional resonances observed for L11P arise from residues that form the A-helix (residues 4–16) and AB loop (residues 17–20) in the native structure and from residues 40–44 and 57–59 of the β-domain. The latter are particularly interesting as they are sequentially and structurally distant from the mutation site (see Figure 1). The urea-induced unfolding transition of L11P also differs significantly from that of [28–111] α-LA; the pattern of unfolding of L11P is summarized in Figure 3c. Interestingly, although all of the resonances from the A-helix and AB loop are observed even in the absence of urea, the B, C, D and 3_{10}-helices remain in a compact structure, only unfolding under the strongly denaturing conditions of 6–7 M urea. These results indicate that the B, C, D and 3_{10}-helices can form a compact subdomain independently of the involvement of the A-helix, although this subdomain is less stable than the one that includes the A-helix. Resonances for all of the residues, except those directly adjacent to C28 and C111, in L11P are observed at a much lower urea concentration (7 M urea compared with 10 M urea for [28–111] α-LA). It is interesting to note, however, that in L11P residues in the B, C, D and 3_{10}-helices (except for the residues close to the C28–C111 disulfide bond) all unfold at very similar urea concentrations. Thus, the C-helix appears to be an essential element for the stability of a compact helical subdomain in a protein lacking a folded A-helix. By contrast, in the four-disulfide bonded protein the C-helix is found to unfold at a lower concentration of urea than the other α-domain helices (Figure 3e). The pattern of unfolding of the β-domain in L11P is similar to that seen for all-Ala α-LA and presumably reflects the lower overall stability of L11P compared to [28–111] α-LA.

The $^{15}N-^{1}H$ HSQC spectrum of the Q117P mutant of [28–111] α-LA at pH 2 and 20 °C in the absence of urea is shown in Figure 2d. This spectrum contains only 35 cross-peaks and is more similar to the spectra observed for [28–111] α-LA and all-Ala α-LA in the absence of urea than it is to the spectrum of L11P. In Q117P the residues in the vicinity of the proline mutation do not give rise to strong peaks in the HSQC spectrum as is the case in L11P. This finding indicates that although CD data show that the proline mutation reduces the total helical structure in the protein, it does not in fact lead to complete unfolding of the C-terminal region of the protein. Residues in this region of the sequence must still interact significantly with the remainder of the molten globule core, leading to a collapsed structure with broadened NMR resonances, despite the loss of secondary structure. The sequence of α-LA in the vicinity of Q117 contains several hydrophobic residues (L115, W118 and L119) and it is likely that their interactions with each other and with hydrophobic residues in the A and B-helices are the origin of the persistent but presumably otherwise relatively disordered structure.

The effects of the Q117P mutation on the urea-induced unfolding of [28–111] α-LA are illustrated in Figure 3d. The Q117P mutation has a general destabilizing effect on the protein as does the L11P mutation; as in L11P the B, C and D-helices in Q117P unfold at urea concentrations that are ~2 M lower than those needed to unfold these structures in [28–111] α-LA. As expected, the C-terminal region containing the mutation is destabilized; residues 115–123 unfold at urea concentrations ~3–5 M lower in Q117P than in [28–111] α-LA. The C-terminal region of Q117P also unfolds more readily than in L11P. In addition to these relatively local interactions, significant long-range destabilization can be observed in the other regions of the helical domain of Q117P. Most dramatically, the N-terminal region of Q117P, containing the A-helix and the AB loop, is considerably destabilized relative to [28–111] α-LA; like the C-terminal residues of Q117P the N-terminal residues unfold at urea concentrations 3–5 M lower than in [28–111] α-LA. Thus, resonances from both the N and C termini appear in the spectrum of Q117P at similar concentrations of urea indicating that these regions, that are adjacent to each other in the native fold, are both destabilized by the mutation in the C-terminal 3_{10}-helix.

Discussion

Two studies of hydrophobic to alanine mutations suggest the presence of a native-like hydrophobic core involving residues in the A, B, and C-terminal 3_{10}-helices in the α-LA molten globule.[21,22] Indeed, even though in the present study there are significant differences in the effects of the two proline mutations on the stability of specific regions of the structures, in both cases there is evidence for the disruption of native-like interactions. The A-helix has a number of long-range contacts to residues 117–120 and 123 of the C terminus and to residues 23, 26 and 27 of the B-helix in the native structure. As the introduction of a proline residue at position 11 leads to local unfolding of the A-helix, it must also lead to a loss of these long-range interactions. In the native structure, the C-terminal 3_{10}-helix has a number of interactions with the A-helix as described above and, in addition, Trp118 makes contacts with residues 28, 31 and 32 of the B-helix. That the unfolding of the A-helix in L11P does not lead to unfolding of the C-terminal 3_{10}-helix, although it does lead to a general destabilization

of the C-terminal regions of the protein, indicates that the contacts between the C-terminal 3_{10}-helix and the B-helix are sufficient to stabilize a folded conformation for the C-terminal region in the absence of interactions with the A-helix. Thus, L11P can form a stable molten globule core consisting of the B, C, D and C-terminal 3_{10}-helices in the absence of a folded A-helix.

Remarkably, and in contrast to the L11P mutation, the introduction of a proline residue at position 117 in the C-terminal 3_{10}-helix does not lead to complete unfolding of even this local region in the absence of urea. The hydrophobic residues within the C-terminal 3_{10}-helix appear to be able to maintain their stable interactions with the remainder of the α-domain even in the absence of local helical structure. However, the proline residue does lead to a degree of destabilization of the C-terminal region as shown by its unfolding at a lower concentration of urea. Importantly, however, this unfolding of the C-terminal region appears to be strongly coupled to the unfolding of the N-terminal region. In the absence of urea the A-helix and AB loop regions are folded and form an integral part of the molten globule structure. These residues are observed to unfold at similar concentrations of urea as observed for the C-terminal region. Thus, it appears that residues in the vicinity of the A-helix cannot maintain their folded conformation in the absence of a folded C terminus. This indicates that the contacts between residues of the C-terminal 3_{10}-helix and the A-helix are crucial for the stability of the A-helix in the molten globule. Thus, it appears from the analysis of the Q117P and L11P mutations that the residues at the C terminus of the [28–111] α-LA molten globule are more important for maintaining the molten globule conformation than are those at the N terminus.

The identification of the structured regions in L11P and Q117P reveals that the A and C-terminal 3_{10}-helices make different contributions to the native-like tertiary fold in the molten globule. It has been found that proline mutations in the A-helix, but not the C-terminal 3_{10}-helix, decrease the preferences for forming native disulfide bonds in the molten globule of a four-cysteine variant of α-LA that contains the α-domain cysteine residues (α-LA(α)).[18] Formation of such bonds depends both on factors favouring formation of the native-like topology, and on factors discouraging formation of other, non-native, topologies. Our residue-specific data help explain how both of these factors could contribute to the differences in disulfide bond preferences seen for the variants of α-LA. In 4-SS α-LA, [28–111] α-LA and all-Ala α-LA the structure in the vicinity of residues 6, 28, 111 and 120, corresponding to cysteine residues in the native protein, is highly resistant to unfolding by urea (Figure 3). This stability is reflected by the strong preference in reduced [28–111] α-LA, and (α-LA(α)) for formation of native disulfide bonds.[15,17] In Q117P, the N and C-terminal regions, containing residues 6 and 120, show a reduced stability relative to the rest of the structure (Figure 3d). However, in the absence of urea these regions are folded and form an integral part of the molten globule structure. This explains the observation that mutations in the C-terminal 3_{10}-helix have little overall effect on disulfide bonding preferences in α-LA(α).[18] On the other hand, in L11P the A-helix is completely unfolded even in the absence of urea. The loss of structure at the N terminus in L11P may decrease the steric hindrance of the chain around C6, possibly increasing its ability to form a non-native disulfide bond with its nearest neighbor, C28. This may explain the large decrease in the preference for forming native disulfide bonds in α-LA(α) that is associated with mutations in the A-helix.[18]

A number of studies have been carried out to assess the structure and folding of isolated peptide fragments from the human α-LA sequence. The isolated A, B and C-helices themselves show little helical propensity.[36–38] Peptides encompassing residues 101–111 and 101–120 of the α-LA sequence do show some degree of helicity in aqueous solution but the latter appears to be largely non-native in character.[39,40] A peptide consisting of residues 1–38 and 95–120 cross-linked by the C28–C111 disulfide bond, however, shows not only significant helical structure but also characteristics of a molten globule including the enhancement of ANS fluorescence.[41] Deletion of residues 95–100 has little effect on the structure in this peptide, indicating that the C-helix is not important for the stability of its molten globule structure. A significant loss of structure is by contrast observed as a result of deletion of either residues 101–110 or 112–120, indicating that the D-helix and C-terminal 3_{10}-helix regions are important.[41] The deletion of residues 1–19, which includes the A-helix, also leads to a significant decrease in structure.[41] The results of these studies have led to the identification of a 58-residue fragment, composed of residues 1–38 and 101–120, as the critical core required to form the α-LA molten globule.[41]

In our previous study of the four-disulfide protein we demonstrated that a highly stable core structure, composed of the A, B, D and C-terminal 3_{10}-helices but lacking the C-helix, was formed at high concentrations of urea. This stable core is similar to that identified in the peptide studies described above. Here, however, we have shown that disruption and unfolding of the A-helix by the introduction of a proline mutation at position 11 leads to a stable molten globule composed of the B, C, D and C-terminal 3_{10}-helices. Thus, it appears that the A-helix is not in itself essential for the formation of a molten globule in the full-length protein, as the presence of the C-helix can provide additional stabilization through long-range contacts that can compensate for the loss of interactions with the A-helix. These results suggest that the α-LA molten globule can exist as an

ensemble of conformations with different subsets of the various native-like helical regions. A recent fluorescence energy transfer study of cytochrome *c* also indicates the presence of an ensemble of conformations in the molten globule state of this protein.[42]

We find that the regions of structure that are most substantially destabilized by the proline mutations in α-LA are those that are in direct contact with the helix containing the mutation. In a comparative study of the urea-induced unfolding of the molten globules of human and bovine α-LA, differences in the pattern of unfolding were largely explained by differences in inter-helical contacts arising from amino acid substitutions.[26] Thus, native-like contacts between helices appear to be crucial in stabilizing the α-domain core of the α-LA molten globule. However, mutagenesis studies have shown that the removal of any individual side-chain from the hydrophobic core of a molten globule has a relatively minor effect on the overall stability of the molten globule and that these effects are small when compared to similar mutations in native structures.[21,22,43,44] Indeed it has been shown that simultaneous substitution of all the hydrophobic residues in the wild-type sequence of the α-domain of human α-LA with leucine residues does not decrease detectably the preference for forming native disulfide bonds, indicating that the mutated protein maintains a native-like fold.[20] Taken together, these results suggest that the native-like topology of the α-domain of α-lactalbumin in its molten globule state is stabilized by a relatively large number of rather non-specific hydrophobic interactions rather than a small number of highly specific native interactions. This conclusion is similar to that drawn from NMR studies of apomyoglobin where again the interactions of a subset of helices appear sufficient to generate the native-like fold within a core region of the structure.[5,6,43,45] As the formation of hydrophobic contacts between helices, particularly amphipathic ones, will serve in turn to stabilize the individual helices, this model for the molten globule state is consistent with the limited degree of both local and global cooperativity evident from the mutational studies described here.

Materials and Methods

Production of labeled protein

Human [28–111] α-LA differs from the wild-type protein in the replacement of the six cysteine residues (6, 61, 73, 77, 91 and 120) by alanine. L11P and Q117P differ from [28–111] α-LA in the replacement of a leucine by a proline at position 11 and of a glutamine by a proline at position 117, respectively. ^{15}N-labeled recombinant human [28–111] α-LA, L11P and Q117P were expressed from BL21(DE3)pLysS (Novagen) harboring the relevant plasmids under T7 control, grown in minimal M9 media containing 1 g/l [^{15}N]ammonium, and supplemented with 1.0 mg/l thiamine. Inclusion bodies containing the proteins were solubilized and purified as described for a related protein.[15,16]

Nuclear magnetic resonance spectroscopy

NMR samples contained 0.7–1.4 mM protein at pH 2 in 95% H_2O/5% ^2H_2O. A 500 MHz 2D gradient-enhanced ^{15}N–^{1}H HSQC spectrum[46] consisting of 128 complex t_1 increments of 1024 complex data points was collected for urea concentrations ranging from 0 M to 10 M at 20 °C and at temperatures ranging from 20 °C to 50 °C in 10 M urea. Sweep widths of 961.5 Hz and 5405.4 Hz were used in the ^{15}N (F_1) and ^{1}H (F_2) dimensions, respectively. Sixteen scans were collected for each t_1 increment. Resonances in the spectrum were assigned using standard procedures applied to ^{15}N-edited 3D gradient-enhanced NOESY-HSQC spectra[46,47] collected under various sample conditions (7 M, 8 M and 10 M urea at 20 °C and 10 M urea at 40 °C for [28–111] α-LA; 5 M and 7 M urea at 20 °C and 10 M urea at 35 °C for L11P; 4 M, 7 M and 10 M urea at 20 °C and 10 M urea at 35 °C for Q117P). Assignment is simplified because only intraresidue and sequential NOEs are observed and because the chemical shifts of side-chain resonances are close to random coil values. All NMR data were processed using Felix 2.3 (Accelrys, San Diego). The HSQC spectra shown in Figure 2 are contoured so that a cross-peak defined by a single contour has 20–25% of the intensity of a cross-peak corresponding to a fully unfolded residue. All the spectra were contoured in the same way to allow a direct comparison of the contour plots. The HSQC urea titration data are summarized in histogram format in Figure 3. The height of each bar in the histogram represents the denaturant concentration at which a peak corresponding to 20–25% of the intensity of a fully unfolded peak is first observed.

Acknowledgements

We thank P. S. Kim for his support and encouragement during this study. We also 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 and C.M.Q. by a Fulbright fellowship. 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.

References

1. Ptitsyn, O. B. (1995). Molten globule and protein folding. *Advan. Protein Chem.* **47**, 83–229.
2. Kuwajima, K. (1996). The molten globule state of α-lactalbumin. *FASEB J.* **10**, 102–109.
3. Dobson, C. M., Sali, A. & Karplus, M. (1998). Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.* **37**, 868–893.
4. Arai, M. & Kuwajima, K. (2000). Role of the molten globule state in protein folding. *Advan. Protein Chem.* **53**, 209–282.

5. Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990). Structural characterization of a partly folded apomyoglobin intermediate. *Science*, **249**, 1544–1548.
6. Jennings, P. A. & Wright, P. E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science*, **262**, 892–896.
7. Arai, M. & Kuwajima, K. (1996). Rapid formation of a molten globule intermediate in refolding of α-lactalbumin. *Fold. Des.* **1**, 275–287.
8. Forge, V., Wijesinha, R. T., Balbach, J., Brew, K., Robinson, C. V., Redfield, C. & Dobson, C. M. (1999). Rapid collapse and slow structural reorganisation during the refolding of bovine α-lactalbumin. *J. Mol. Biol.* **288**, 673–688.
9. Trouiller, A., Reinstadler, D., Dupont, Y., Naumann, D. & Forge, V. (2000). Transient non-native secondary structures during the refolding of α-lactalbumin detected by infrared spectroscopy. *Nature Struct. Biol.* **7**, 78–86.
10. Arai, M., Ito, K., Inobe, T., Nakao, M., Maki, K., Kamagata, K. *et al.* (2002). Fast compaction of α-lactalbumin during folding studied by stopped-flow X-ray scattering. *J. Mol. Biol.* **321**, 121–132.
11. Acharya, K. R., Ren, J. S., Stuart, D. I., Phillips, D. C. & Fenna, R. E. (1991). Crystal structure of human α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol.* **221**, 571–581.
12. Baum, J., Dobson, C. M., Evans, P. A. & Hanley, C. (1989). Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry*, **28**, 7–13.
13. Chyan, C. L., Wormald, C., Dobson, C. M., Evans, P. A. & Baum, J. (1993). Structure and stability of the molten globule state of guinea-pig α-lactalbumin: a hydrogen exchange study. *Biochemistry*, **32**, 5681–5691.
14. Peng, Z. & Kim, P. S. (1994). A protein dissection study of a molten globule. *Biochemistry*, **33**, 2136–2141.
15. Peng, Z., Wu, L. C. & Kim, P. S. (1995). Local structural preferences in the α-lactalbumin molten globule. *Biochemistry*, **34**, 3248–3252.
16. Schulman, B. A., Redfield, C., Peng, Z., Dobson, C. M. & Kim, P. S. (1995). Different subdomains are most protected from hydrogen exchange in the molten globule and native states of human α-lactalbumin. *J. Mol. Biol.* **253**, 651–657.
17. Wu, L. C., Peng, Z.-y. & Kim, P. S. (1995). Bipartite structure of the α-lactalbumin molten globule. *Nature Struct. Biol.* **2**, 281–286.
18. Schulman, B. A. & Kim, P. S. (1996). Proline scanning mutagenesis of a molten globule reveals non-cooperative formation of a protein's overall topology. *Nature Struct. Biol.* **3**, 682–687.
19. Schulman, B. A., Kim, P. S., Dobson, C. M. & Redfield, C. (1997). A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nature Struct. Biol.* **4**, 630–634.
20. Wu, L. C. & Kim, P. S. (1997). Hydrophobic sequence minimization of the α-lactalbumin molten globule. *Proc. Natl Acad. Sci. USA*, **94**, 14314–14319.
21. Wu, L. C. & Kim, P. S. (1998). A specific hydrophobic core in the α-lactalbumin molten globule. *J. Mol. Biol.* **280**, 175–182.
22. Song, J., Bai, P., Luo, L. & Peng, Z.-y. (1998). Contribution of individual residues to formation of the native-like tertiary topology in the α-lactalbumin molten globule. *J. Mol. Biol.* **280**, 167–174.
23. Polverino de Laureto, P., Scaramella, E., Frigo, M., Wondrich, F. G., De Filippis, V., Zambonin, M. & Fontana, A. (1999). Limited proteolysis of bovine α-lactalbumin: isolation and characterization of protein domains. *Protein Sci.* **8**, 2290–2303.
24. Redfield, C., Schulman, B. A., Milhollen, M. A., Kim, P. S. & Dobson, C. M. (1999). α-Lactalbumin forms a compact molten globule in the absence of disulfide bonds. *Nature Struct. Biol.* **6**, 948–952.
25. Bai, P., Song, J., Luo, L. & Peng, Z.-y. (2001). A model of dynamic side-chain–side-chain interactions in the α-lactalbumin molten globule. *Protein Sci.* **10**, 55–62.
26. Wijesinha-Bettoni, R., Dobson, C. M. & Redfield, C. (2001). Comparison of the denaturant-induced unfolding of the bovine and human α-lactalbumin molten globules. *J. Mol. Biol.* **312**, 261–273.
27. Last, A. M., Schulman, B. A., Robinson, C. V. & Redfield, C. (2001). Probing subtle differences in the hydrogen exchange behavior of variants of the human α-lactalbumin molten globule using mass spectrometry. *J. Mol. Biol.* **311**, 909–919.
28. Ramboarina, S. & Redfield, C. (2003). Structural characterisation of the human α-lactalbumin molten globule at high temperature. *J. Mol. Biol.* **330**, 1177–1188.
29. Jeng, M. F. & Englander, S. W. (1991). Stable submolecular folding units in a non-compact form of cytochrome *c*. *J. Mol. Biol.* **221**, 1045–1061.
30. Alexandrescu, A. T., Abeygunawardana, C. & Shortle, D. (1994). Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a heteronucluear NMR study. *Biochemistry*, **33**, 1063–1072.
31. Eliezar, D., Jennings, P. A., Dyson, H. J. & Wright, P. E. (1997). Populating the equilibrium molten globule state of apomyoglobin under suitable conditions for structural characterisation by NMR. *FEBS Letters*, **417**, 92–96.
32. Chamberlain, A. K. & Marqusee, S. (1998). Molten globule unfolding monitored by hydrogen exchange in urea. *Biochemistry*, **37**, 1736–1742.
33. Kim, S., Bracken, C. & Baum, J. (1999). Characterization of millisecond time-scale dynamics in the molten globule state of α-lactalbumin by NMR. *J. Mol. Biol.* **294**, 551–560.
34. Luo, Y. & Baldwin, R. L. (1999). The 28–111 disulfide bond constrains the α-lactalbumin molten globule and weakens its cooperativity of folding. *Proc. Natl Acad. Sci. USA*, **96**, 11283–11287.
35. Kuwajima, K., Ikeguchi, M., Sugawara, T., Hiraoka, Y. & Sugai, S. (1990). Kinetics of disulfide bond reduction in α-lactalbumin by dithiothreitol and molecular basis of superactivity of the Cys6-Cys120 disulfide bond. *Biochemistry*, **29**, 8240–8249.
36. Shimizu, A., Ikeguchi, M., Kobayashi, T. & Sugai, S. (1996). A synthetic peptide study on the molten globule of α-lactalbumin. *J. Biochem.* **119**, 947–952.
37. Kuhlman, B., Boice, J. A., Wu, W-J., Fairman, R. & Raleigh, D. P. (1997). Calcium binding peptides from α-lactalbumin: implications for protein folding and stability. *Biochemistry*, **36**, 4607–4615.
38. Demarest, S. J., Fairman, R. & Raleigh, D. P. (1998). Peptide models of local and long-range interactions in the molten globule state of human α-lactalbumin. *J. Mol. Biol.* **283**, 279–291.
39. Demarest, S. J. & Raleigh, D. P. (2000). Solution structure of a peptide model of a region important for the folding of α-lactalbumin provides evidence for the

formation of nonnative structure in the denatured state. *Proteins: Struct. Funct. Genet.* **38**, 189–196.
40. Demarest, S. J., Horng, J-C. & Raleigh, D. P. (2001). A protein dissection study demonstrates that two specific hydrophobic clusters play a key role in stabilizing the core structure of the molten globule state of human α-lactalbumin. *Proteins: Struct. Funct. Genet.* **42**, 237–242.
41. Demarest, S. J., Boice, J. A., Fairman, R. & Raleigh, D. P. (1999). Defining the core structure of the α-lactalbumin molten globule state. *J. Mol. Biol.* **294**, 213–221.
42. Lyubovitsky, J. G., Gray, H. B. & Winkler, J. R. (2002). Structural features of the cytochrome c molten globule revealed by fluorescence energy transfer kinetics. *J. Am. Chem. Soc.* **124**, 14840–14841.
43. Hughson, F. M., Barrick, D. & Baldwin, R. L. (1991). Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry*, **30**, 4113–4118.
44. Kay, M. S., Ramos, C. H. I. & Baldwin, R. L. (1999). Specificity of native-like interhelical hydrophobic contacts in the apomyoglobin intermediate. *Proc. Natl Acad. Sci. USA*, **96**, 2007–2012.
45. Eliezar, D., Yao, J., Dyson, H. J. & Wright, P. E. (1998). Structural and dynamical characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature Struct. Biol.* **5**, 148–155.
46. Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665.
47. Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T. & Gronenborn, A. M. (1990). Complete resonance assignment for the polypeptide backbone of interleukin 1(using three-dimensional heteronuclear NMR spectroscopy. *Biochemistry*, **29**, 3542–3556.
48. Kraulis, P. K. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

Edited by Dr. P. Wright

(Received 3 November 2003; received in revised form 10 February 2004; accepted 17 February 2004)