The Denatured State under Native Conditions: A Non-native-like Collapsed State of N-PGK

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · APRIL 2006

Impact Factor: 4.33 · DOI: 10.1016/j.jmb.2005.12.080 · Source: PubMed

11 AUTHORS, INCLUDING:

Michelle A C Reed
University of Birmingham
12 PUBLICATIONS **211** CITATIONS
SEE PROFILE

Tooba Alizadeh
The University of Sheffield
9 PUBLICATIONS **159** CITATIONS
SEE PROFILE

Rosemary A Staniforth
The University of Sheffield
24 PUBLICATIONS **946** CITATIONS
SEE PROFILE

doi:10.1016/j.jmb.2005.12.080 J. Mol. Biol. (2006) 357, 365–372

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

COMMUNICATION

The Denatured State under Native Conditions: A Non-native-like Collapsed State of N-PGK

Michelle A. C. Reed[1], Clare Jelinska[1], Karl Syson[1], Matthew J. Cliff[1] Andrew Splevins[1], Tooba Alizadeh[1], Andrea M. Hounslow[1] Rosemary A. Staniforth[1], Anthony R. Clarke[2], C. Jeremy Craven[1] and Jonathan P. Waltho[1]*

[1]*Department of Molecular Biology and Biotechnology University of Sheffield, Sheffield S10 2TN, UK*

[2]*Department of Biochemistry University of Bristol, Bristol BS8 1TD, UK*

**Corresponding author*

The guanidinium-denatured state of the N-domain of phosphoglycerate kinase (PGK) has been characterized using solution NMR. Rather than behaving as a homogenous ensemble of random coils, chemical shift changes for the majority of backbone amide resonances indicate that the denatured ensemble undergoes two definable equilibrium transitions upon titration with guanidinium, in addition to the major refolding event. ^{13}C and ^{15}N chemical shift changes indicate that both intermediary states have distinct helical character. At denaturant concentrations immediately above the mid-point of unfolding, size-exclusion chromatography shows N-PGK to have a compact, denatured form, suggesting that it forms a helical molten globule. Within this globule, the helices extend into some regions that become beta strands in the native state. This predisposition of the denatured state to extensive non-native-like conformation, illustrates that, rather than directing folding, conformational pre-organization in the denatured state can compete with the normal folding direction. The corresponding reduction in control of the direction of folding as proteins become larger, could thus constitute a restriction on the size of protein domains.

Keywords: denatured state; nuclear magnetic resonance; chemical shift; molten globule; phosphoglycerate kinase

Unfolded forms of proteins play important roles in protein folding,[1] in protein misfolding events that have been associated with numerous diseases,[2] and in a wide variety of protein functions.[3] The traditional view of unfolded proteins as fully unfolded, randomly fluctuating polymers is being constantly re-evaluated,[4–6] with evidence of collapsed forms, commonly referred to as molten globules,[7,8] residual structure, particularly hydrophobic clustering in the vicinity of aromatic residues,[9–12] and native-like topology without compaction and at elevated levels of chemical denaturant.[13] The conformational ensembles that constitute unfolded states are commonly analysed in terms of features observed within the native protein,[14] including both secondary structure content and tertiary contacts, though some specific non-native contacts have been demonstrated to play an important role in the kinetics of folding.[12,15]

Here, we report the characterization of denatured species populated by the isolated N-domain of phosphoglycerate kinase (N-PGK), a 20 kDa protein fragment from *Bacillus stearothermophilus*[16,17] comprising an $(\alpha\beta)_6$ Rossmann fold.[18] In line with most other larger proteins,[19] N-PGK populates an intermediate state during refolding following dilution of

Present addresses: M. A. C. Reed, HWB-NMR, University of Birmingham, Vincent Drive, Edgbaston, Birmingham B15 2TT, UK; K. Syson, Department of Chemistry, University of Sheffield, Sheffield S3 7HF, UK.

Abbreviations used: GuHCl, guanidinium chloride; N-PGK, N-domain of phosphoglycerate kinase; HSQC, heteronuclear single quantum coherence; SEC, size-exclusion chromatography; RDC, residual dipolar coupling.

E-mail address of the corresponding author: j.waltho@shef.ac.uk

chemical denaturants. In N-PGK, this intermediate state attains a native-like topology on a millisecond or faster timescale and is highly stable, affording considerable protection from hydrogen exchange throughout the protein core.[20] While such kinetic intermediate states have been shown to share many features of equilibrium molten globule states,[21,22] a system permitting the direct characterization of the former as the most populated species at equilibrium has proved elusive.

NMR relaxation dispersion methods were used recently[23] to characterise ^{15}N chemical shift changes attributable to a low population intermediate present during the folding of an SH3 domain. These shifts were used to calculate a model of the intermediate state ensemble. The chemical shift changes for the N-PGK kinetic intermediate state should, in principle, be more readily detectable given its relatively high stability near the midpoint of denaturation using guanidinium chloride (GuHCl).[16,17] Here, the residual population of the kinetic intermediate state is predicted to be sufficiently high to perturb the chemical shifts of the denatured ensemble at moderate levels of chemical denaturant, and this perturbation will quickly diminish as the chemical denaturant concentration is raised. To investigate this and the nature of the highly stable folding intermediate of N-PGK, we measured the denaturant concentration dependence of ^{15}N and ^{13}C NMR parameters as reporters of the distribution of species within the denatured state ensemble.

Denaturant dependence of NMR chemical shifts

Under the conditions of the NMR experiments, the midpoint of the major unfolding transition for N-PGK occurs at 1.1 M GuHCl and is 99% complete by 1.4 M GuHCl (Figure 1), as monitored by intrinsic tyrosine fluorescence, amide circular dichroism (CD) and size-exclusion chromatography (SEC). At the midpoint of unfolding, the relaxation rate between the native state and the denatured ensemble is 0.4 s^{-1} at 25 °C,[17] and therefore the NMR signals of these forms are fully resolved and can be observed independently (i.e. the NMR signals are in a slow exchange regime). In contrast, within the species that make up the denatured ensemble, interconversion rates are much faster, leading to the averaging of NMR signals (i.e. a fast exchange regime). Complete backbone and CB NMR resonance assignment was determined for denatured N-PGK at pH 6.0 and 25 °C in 4.0 M GuHCl, on the basis of matching CA, CB and CO resonances, using the same strategy as that adopted for the native state of the intact 43 kDa protein.[20]

$^{1}H,^{15}N$-heteronuclear single quantum coherence (HSQC) spectra of the denatured ensemble were recorded over the denaturant concentration range of 1.2 M to 4.0 M GuHCl. If the denatured ensemble of N-PGK was composed of random coils, the observed NMR chemical shifts should be unaffected by the concentration of denaturant. However, the

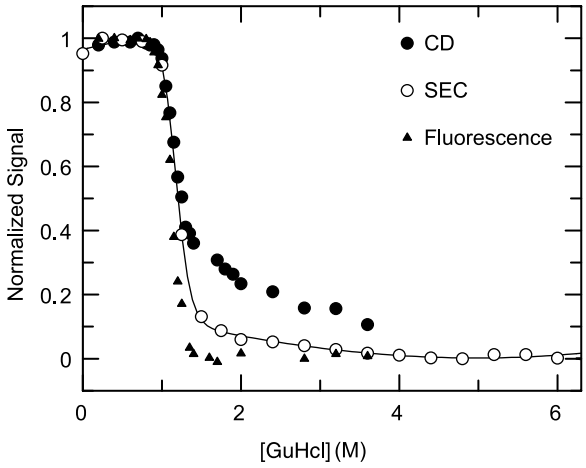

Figure 1. GuHCl denaturation of N-PGK monitored by intrinsic tyrosine fluorescence (filled triangles), far-UV CD at 222 nm (filled circles), and size-exclusion chromatography (open circles): N-PGK expressed from pET5alpha in BL21(DE3) cells was purified as published.[16] Fluorescence and circular dichroism measurements: samples containing 10 μM N-PGK under the standard conditions (20 mM Tris, 20 mM bis-Tris, 5 mM DTT, 1 mM NaN_3 (pH 6.0), 25 °C) were equilibrated for 30 min at the indicated concentrations of GuHCl. Fluorescence intensity measurements were made using a Varian Cary Eclipse fluorescence spectrometer using excitation and emission wavelengths of 272 nm and 305 nm, respectively. Molar ellipticity at 222 nm was measured on a Jasco J-810 spectropolarimeter. Size-exclusion chromatography: samples containing 50 μM N-PGK were pre-equilibrated for 1 h under the standard conditions plus an additional 100 mM NaCl, and the indicated GuHCl concentrations: 20 μl aliquots were injected onto a Phenomenex Biosep-SEC-S 3000 column, and the elution volume determined by monitoring absorbance at 280 nm. Elution volumes were corrected for the effects of GuHCl on the column matrix.[43]

NMR chemical shifts changed virtually throughout (Figure 2), indicating that the nature of the ensemble changes with denaturant concentration, and the species involved interconvert rapidly. An equivalent titration of the denatured ensemble with NaCl (in 1.4 M GuHCl) did not elicit such changes, indicating that it was the denaturing properties of the guanidinium ion that were the cause of the chemical shift perturbation. The $^{1}H,^{15}N$ resonance assignment for many residues was directly transferable throughout the range 4.0 M to 1.2 M GuHCl, and was confirmed following backbone and CB NMR resonance assignment for denatured N-PGK in 1.2 M GuHCl.

A simple, two-state transition within the denatured ensemble would produce a linear relationship between the ^{1}H and ^{15}N resonances, i.e. the $^{1}H,^{15}N$-HSQC spectra cross-peaks would move linearly with denaturant concentration until the transition ends. However, for numerous resonances the movement of ^{1}H and ^{15}N correlation peaks was non-linear as a function of denaturant

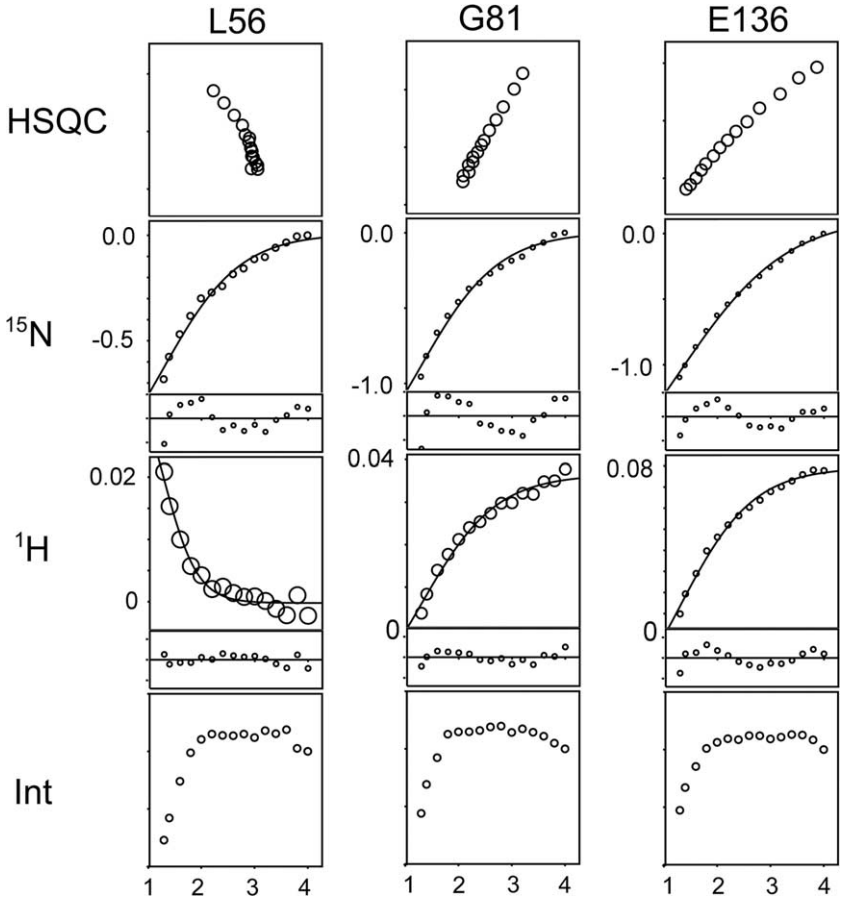

Figure 2. Behavior of representative $^{1}H,^{15}N$ correlation peaks in HSQC spectra of N-PGK as a function of GuHCl concentration. Columns 1–3 represent residues L56, G81 and E136, respectively. Row 1, change in chemical shift of both nuclei (vertical axis = ^{15}N, horizontal axis = ^{1}H shift), as observed in an overlay of a series of 2D-HSQC spectra. Rows 2 and 4, ppm change in ^{15}N and ^{1}H shift, respectively, as a function of GuHCl concentration, relative to the chemical shift at 4.0 M GuHCl. The size of the symbols used represents the differences in uncertainties of the measured chemical shifts. Rows 3 and 5, the residuals obtained following fitting of the data in rows 2 and 4 to a two-state transition. Row 6, normalized intensities of the $^{15}N,^{1}H$-HSQC correlation peaks as a function of GuHCl concentration. The tick marks on the *x*-axes correspond to 1–4 M GuHCl. NMR, a 1.2 mM sample of uniformly ^{15}N-labelled N-PGK was prepared as described,[17] equilibrated under standard conditions (20 mM Tris, 20 mM bis-Tris, 5 mM DTT, 1 mM NaN_3 (pH 6.0), 25 °C) plus 1.1 mM GuHCl, 10% $^{2}H_2O$, 0.4 mM 3-(trimethylsilyl)-propionic acid (TSP). Addition of aliquots of 6.0 M GuHCl in the standard buffer was followed by equilibration for 10 min before acquisition of 1D ^{1}H and 2D ^{1}H–^{15}N-HSQC NMR. The denaturant concentration at the end of the titration reached 4.0 M GuHCl. Water suppression was achieved using a pre-saturation pulse during the relaxation delay between scans. The 2D ^{1}H–^{15}N-HSQC spectra were recorded using sensitivity-enhanced pulsed-field gradient coherence selection and GARP ^{15}N decoupling during acquisition. Spectra were acquired on a Bruker DRX-600 spectrometer. Chemical shifts were referenced relative to the ^{1}H signal of TSP, using a $^{15}N/^{1}H$ ratio of 0.101329118. Resonance assignments were made at 1.2 M and 4.0 M GuHCl, using the standard triple resonance experiments[44] on a $^{1}H^{13}C^{15}N$-labelled sample.[17] Peak picking was performed semi-manually and the backbone assignment strategy used the simulated annealing program of the "asstools" suite of assignment programs, as described.[20]

concentration (Figure 2). Therefore, the chemical shift changes correspond to a minimum of three definable denatured species (termed here U, I_a and I_b) and two transitions. The non-linear movement of $^{1}H,^{15}N$-HSQC cross-peaks occurs as a result of the magnitude of the chemical shift changes being unequal between the two transitions,[24] including some cases where the ^{1}H chemical shift change alters its direction.

Correspondingly, the fitting of the chemical shift changes to a single transition within the denatured ensemble leads to a well-defined residual for many resonances (Figure 2). The data fit reliably to two transitions, one at high denaturant concentration (between states U and I_a), which is observed almost in its entirety between 4.0 M and 2.7 M GuHCl with a midpoint of ca 3.2 M, and one at low denaturant concentration (between states I_a and I_b), which is dominant below 2.0 M GuHCl. The transition between states I_a and I_b is also associated with significant intensity reductions, in contrast to the transition between states U and I_a (Figure 2). This is indicative of state I_b having regions that undergo conformational exchange on a slower timescale (in the intermediate exchange regime). Below 1.2 M GuHCl the intensities of resonances from the denatured ensemble start to diminish rapidly as the population of the native state becomes substantial. This transition to the native state prevents the observation of the complete I_a to I_b transition.

The denaturant concentration dependence of the free energies, or *m*-values, of the denatured state species involved can be correlated with the degree of desolvation of hydrophobic surface area in a way analogous to that used for kinetic folding intermediates and transition states.[25–27] An examination of the changes in NMR parameters on a per-residue basis provided no justification for involving more than two states in the analysis of each transition. Fitting of the behavior of individual resonances

during the U to I_a transition (to the relationship $K/(1+K)$, where $K=I_a/U$; $\ln K=\ln K_w+\Delta m_{U\text{-}Ia}\times[GuHCl]$)[26] indicates that the m-value of the I_a state differs only between 1 and 2 M^{-1} from that of the U state (i.e. $\Delta m_{U\text{-}Ia}<2\ M^{-1}$). The change in m-value for the I_a to I_b transition cannot be determined accurately, since the midpoint cannot be established unambiguously. However, an estimate of the maximum change in m-value can be made by assuming that the lowest denaturant concentration for which the chemical shift can be observed corresponds to the midpoint. This equates to a change in m-value for the I_a to I_b transition of $<5\ M^{-1}$ (i.e. $\Delta m_{U\text{-}Ib}<7\ M^{-1}$). Clearly, the conformational species dominating the I_a and the I_b ensembles have substantially decreased desolvation of hydrophobic side-chains, compared with the native state ($\Delta m_{U\text{-}N}=16.6\ M^{-1}$) and the previously characterized kinetic intermediate (I_k) state ($\Delta m_{U\text{-}Ik}=11.1\ M^{-1}$).[16]

Structural characterisation of the I_a and I_b states

Further evidence of the nature of the species involved in the transitions of the denatured ensemble comes from a closer examination of the other probes of protein denaturation (Figure 1). The intrinsic tyrosine fluorescence signal is not affected measurably at denaturant concentrations beyond the major native to denatured state transition, indicating that this probe is not sensitive to the transitions within the denatured ensemble. This is consistent with previous studies where no change in tyrosine fluorescence was observed during the refolding of N-PGK to its kinetic intermediate state.[16] In contrast, according to the amide CD response, residual secondary structure is apparent within the denatured ensemble beyond any contribution from the native state, as would be predicted from the behavior of the NMR signals. The U to I_a transition results in the formation of ca 20% of the native state CD response at 222 nm, while the I_a to I_b transition results in at least a further 40%. The SEC retention times indicate that the denatured state at high denaturant concentration is as expanded as other fully unfolded proteins,[28] but as the denaturant concentration is reduced, significant compaction is associated with the formation of the I_b state. An equivalent experiment performed on stefin A, a protein with a similar midpoint for the major folded to denatured state transition,[29] indicates that the change in retention time is not a result of the intrinsic behavior of the chromatography column. This experiment also eliminates a contribution from intermolecular association to the behavior of the denatured state ensemble, since the change in retention time for N-PGK as the denaturant concentration is reduced is in the opposite direction to that resulting from aggregation.

When viewed across the entire structure (Figure 3(a)), it is possible to determine the predominant changes in structure associated with the transitions. The chemical shift changes of CA, CB and CO resonances are reliably correlated with the backbone torsional angle distribution, and thus the secondary structure content of the protein.[30,31] For the majority of residues, the changes in

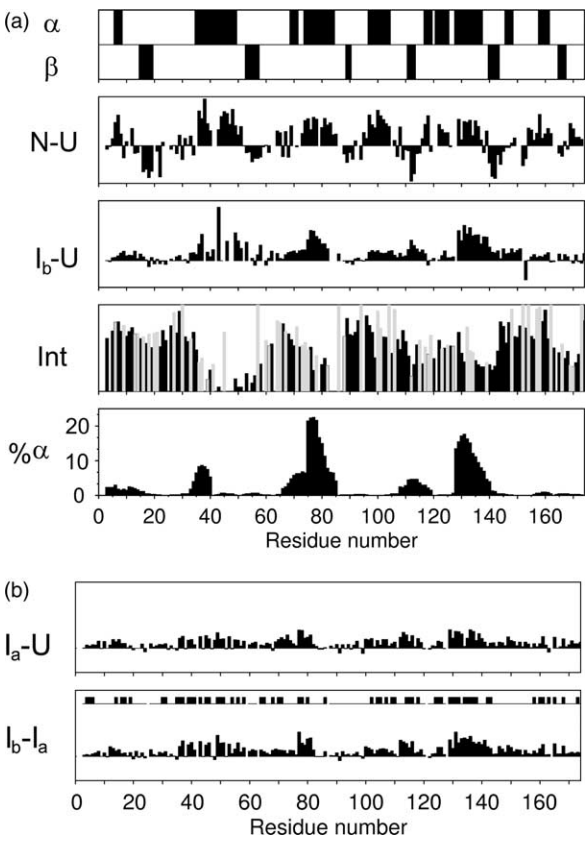

Figure 3. Structural properties as a function of the primary sequence. In (a), row 1 shows the secondary structure content of the native structure (pdb code 1PHP). Row 2 shows the change in chemical shift between the native and the 4.0 M GuHCl denatured ensemble of N-PGK. Each bar is based on a weighted combination ($\Delta(\delta CA)+\Delta(\delta CO)-0.5\Delta(\delta N)$) of changes in chemical shift from backbone ^{13}CA, ^{13}CO and ^{15}N resonances. Row 3 shows the equivalent data as row 2, but between 4.0 M and 1.2 M GuHCl, i.e. for the chemical shift changes between states U and I_b. For clarity, the vertical axis of row 3 has been magnified by a factor of 3 compared with row 2. Row 4 shows the relative intensity of $^{1}H,^{15}N$ correlation cross-peaks comparing data at 4.0 M and 1.2 M GuHCl (grey bars correspond to crosspeaks that suffer from partial overlap). Row 5 shows the per-residue helix propensity predicted using the program AGADIR. Chemical shifts were referenced relative to the ^{1}H signal of TSP, using $^{15}N/^{1}H$ and $^{13}C/^{1}H$ γ-ratios of 0.101329118 and 0.25144953, respectively. In (b), row 1 shows the changes in ^{15}N chemical shift between 4.0 M and 3.0 M GuHCl, where the transition between states U and I_a dominates the denatured ensemble. Row 2 shows the equivalent changes between 3.0 M and 1.2 M GuHCl, where the transition between states I_a and I_b dominates. For clarity, the vertical axis of row 1 has been magnified by a factor of 3 compared with row 2. The inset in row 2 demarks with blocks where non-linear movements of ^{1}H–^{15}N-HSQC cross-peaks are observed as the GuHCl concentration is changed.

chemical shift indicate an increase in helical conformation between the U-state and the I_b-state, which is particularly apparent in four main regions; namely, 33–53, 67–82, 97–117 and 129–151. Moreover, the regions with the greatest changes in helicity correspond with those with the greatest intensity reductions (Figure 3(a)).

The chemical shift changes of the ^{15}N resonances mirror closely the behavior of the diagnostic ^{13}C resonances (Figure 3(b)). Consequently, the individual transitions between states U and I_b can be examined separately. In the U to I_a transition a large proportion of the helical conformation is formed. In the I_a to I_b transition, the population of the same helical conformations is considerably increased, though the exact magnitude of the population at the end of the transition is masked by the denatured to native state transition. It is also apparent from the residues that show curvature in the ^{1}H,^{15}N-HSQC titration that the two states are distinguishable, and the widespread distribution of residues with such behavior (Figure 3(b)) is consistent with the differences between states I_a and I_b being global rather than local.

Extensive non-native character

What is most striking about the nature of the structure populated within the I_a and I_b state ensembles is that there is extensive non-native character (Figure 3). The helices that are formed are distinctly longer than those present in the native state, and encompass regions of the protein that become beta strands (notably strands 4 and 5) in the native structure. It is also striking that the distribution of helices within the denatured ensemble at lower denaturant concentration reflects quite closely that predicted for peptides on the basis of primary sequence,[32] implying that intrinsic properties of these sequences have a major impact on the conformational selection that is occurring as the denaturant concentration is reduced. The observed compaction at low denaturant concentration appears to reinforce this selection through tertiary contacts that inevitably must have a high degree of non-native character, and that most likely result in the formation of a helical bundle molten globule comprising four or five helices, rather than a direct precursor of the native alpha-beta topology.

The organization of the helical molten globule is not apparent from the changes in chemical shift. The lengths of the linkers connecting the helical regions are large enough to accommodate many relative orientations of the helices and the chemical shift changes in the linkers provide no further conclusive evidence. In principle, the relative orientations of the helices may be determinable by measuring residual dipolar couplings (RDCs) from within the denatured ensemble.[33] RDCs are a small fraction of the dipolar couplings observed for immobilized proteins and in the case of denatured states could arise from the rotational properties of the entire ensemble or a sub-population of that ensemble. For staphylococcal nuclease, the former interpretation was assumed and from the observed correlation between RDCs at low and high levels of urea, it was proposed that the highly expanded protein had a native-like topology.[13] Using similar methods,[33,34] we measured RDCs between amide ^{1}H and ^{15}N resonances for the denatured ensembles of N-PGK. Virtually all of the measured values for the resolvable resonances were negative (see Supplementary Data), which is a consistent feature of denatured ensembles in general,[35,36] rather than a reliable guide to the orientation of partially populated secondary structure elements. Hence, the exact relationship of the helices awaits confirmation, though it has proved very difficult with other molten globules to determine precise tertiary contacts.[37]

Implications for folding

The GuHCl-denatured state ensemble of N-PGK (Figure 4) shifts from a more fully unfolded, random coil-like U state ensemble at concentrations above 4 M to an I_a state ensemble that remains predominantly expanded but which contains the measurable population of at least four helices at a denaturant concentration of 2 M. The denatured ensemble shifts towards the more compact, I_b state, with considerable further population of the same helices by 1.2 M guanidinium chloride. The apparent paradox of the U to I_a transition, that short-range order is present and yet the ensemble otherwise has random coil behavior, has been

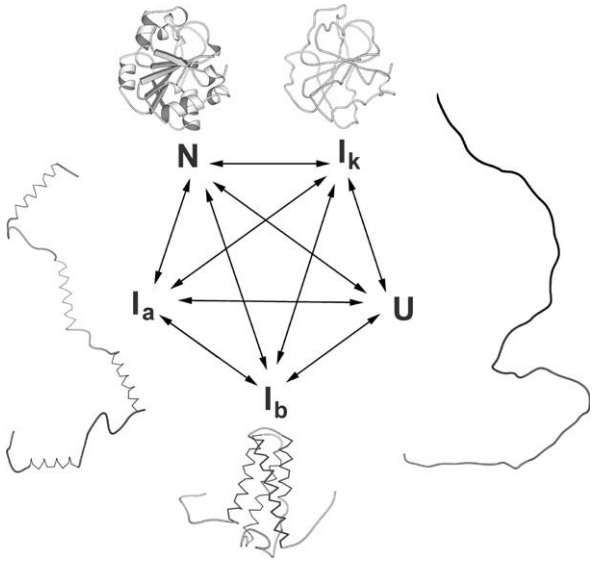

Figure 4. A schematic model of the potential interconversions of five definable states for N-PGK: unfolded (U), native (N), and kinetic intermediate (I_k) states previously defined, and the expanded (I_a) and collapsed (I_b) helical states. The relative orientation of helices depicted for state I_b is purely hypothetical. The measured parameters do not define an order in which the transitions occur during refolding and thus all possible transitions are denoted equally.

reconciled previously in terms of the insensitivity of hydrodynamic radii and radii of gyration to small populations of local order.[5,28] The I_b state ensemble, in contrast, displays the increased R_2 relaxation rates (and hence attenuated signal) widely observed for compact molten globule states,[38,39] which are proposed to result from the exchange of side-chain positions in the cores of the globules on a micro–millisecond timescale.[38] This provides further evidence that the I_b state is a relatively compact state, with extensive tertiary contacts. However, the I_b ensemble still has a substantially lower degree of desolvation (on the basis of *m*-value) than the previously identified, native-like, kinetic intermediate (I_k) state.[16]

Much of the behaviour of N-PGK in chemical denaturant is analogous to situations where molten globule states are the predominantly populated species in the absence of denaturants.[28,38,39] In these molten globules, the compact states are melted out gradually on addition of chemical denaturants, as observed here. The non-cooperative nature of denaturation observed in some molten globules[38,39] is less apparent for the I_b state of N-PGK, though the observation of this may be somewhat obscured by the lack of a full denaturation curve for the appropriate transition in the present case. The behavior of N-PGK, in other ways, is reminiscent of that of beta-lactoglobulin where, on dilution of denaturant, refolding leads to the transient formation of a kinetic intermediate with both a native-like core and some local non-native helicity prior to crossing the rate-limiting step to the native state.[40] In N-PGK, it appears that more extensive, non-native-like structure is already a substantial feature of the denatured state ensemble. There is some evidence that this is a more widespread phenomenon. In the denatured state of a much smaller protein, the SH3 domain of drkN, the measureable population of a short segment of turn conformations (i.e. the population of the alpha-region of phi-psi space) occurs in a region that becomes a beta-strand in the native state.[41,42] There is no evidence for long-range interactions in the denatured state of drkN SH3, and thus it would correspond most closely with the I_a state of N-PGK. The behaviors of these proteins illustrate that even where low-resolution probes of folding do not indicate drastic changes in conformation away from native-like or fully unfolded species, it may be inaccurate to assume that the result of the initial collapse in protein folding occurs with the acquisition of native-like topology.

The refolding of N-PGK from an expanded, chemically denatured state is thus not as straightforward as previously assumed. On refolding, the significant population of a non-native-like helical bundle species within the denatured ensemble will precede the formation of a native-like alpha-beta kinetic intermediate state. The conversion of non-native-like to native-like topology then occurs either directly *via* rearrangement of the helical bundle, or indirectly *via* a more fully unfolded form (Figure 4), all within a millisecond or faster timescale. Thus, N-PGK, a relatively large protein for which the denatured ensemble has been examined in detail, has an alternative folding route that is kinetically more preferable than the construction of a native-like topology. Proteins with higher molecular masses tend to form more stable intermediate states.[2,19] In circumstances where these intermediates have extensive non-native-like character, the denatured ensemble of a larger protein may have insufficient time to establish a native-like topology before progressing on an alternative trajectory. It may be this inability for larger protein units to dictate that folding proceeds in the chosen direction that is a major contributory factor in the preference to construct larger proteins from smaller domains, where the folding route can be controlled more tightly.

Acknowledgements

We thank M. J. Parker and A. Annila for many helpful discussions, M. Rossi for experimental assistance, and Accelrys for the provision of FELIX. This work was supported by the BBSRC (UK), the Wellcome Trust and the Royal Society (London).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.12.080

References

1. Baldwin, R. A. (2002). A new perspective on unfolded proteins. *Advan. Protein Chem.* **62**, 361–367.
2. Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, **426**, 884–890.
3. Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331.
4. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N. *et al.* (2004). Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl Acad. Sci. USA*, **101**, 12491–12496.
5. Fitzkee, N. C. & Rose, G. D. (2004). Reassessing random-coil statistics in unfolded proteins. *Proc. Natl Acad. Sci. USA*, **34**, 12497–12502.
6. Zagrovic, B., Snox, C. D., Khaliq, S., Shirts, M. R. & Pande, V. S. (2002). Native-like mean structure in the unfolded ensemble of small proteins. *J. Mol. Biol.* **323**, 153–164.
7. Kuwajima, K. (1989). The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct. Funct. Genet.* **6**, 87–103.
8. Dobson, C. M. (1994). Protein folding: solid evidence for molten globules. *Curr. Biol.* **4**, 636–640.

9. Neri, D., Billeter, M., Wider, G. & Wuthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science*, **257**, 1559–1563.
10. Saab-Rincon, G., Gualfetti, P. J. & Matthews, C. R. (1996). Mutagenic and thermodynamic analyses of residual structure in the alpha-subunit of tryptophan synthase. *Biochemistry*, **35**, 1988–1994.
11. Ropson, J. & Frieden, C. (1992). Dynamic NMR spectral analysis and protein folding: identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by ^{19}F NMR. *Proc. Natl Acad. Sci. USA*, **89**, 7222–7226.
12. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T. *et al*. (2002). Long-range interactions within a non-native protein. *Science*, **295**, 1719–1722.
13. Shortle, D. & Ackerman, M. S. (2001). Persistence of native-like topology in a denatured protein in 8 M urea. *Science*, **293**, 487–489.
14. Fersht, A. R. & Sato, S. (2004). Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl Acad. Sci. USA*, **101**, 7976–7981.
15. Di Nardo, A. A., Korzhnev, D. M., Stogios, P. J., Zarrine-Afsar, A., Kay, L. E. & Davidson, A. R. (2004). Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. *Proc. Natl Acad. Sci. USA*, **101**, 7954–7959.
16. Parker, M. J., Spencer, J., Jackson, G. S., Burston, S. G., Hosszu, L. L. P., Craven, C. J. *et al*. (1996). Domain behavior during the folding of a thermostable phosphoglycerate kinase. *Biochemistry*, **35**, 15740–15752.
17. Hosszu, L. L. P., Craven, C. J., Parker, M. J., Lorch, M., Spencer, J., Clarke, A. R. & Waltho, J. P. (1997). Structure of a kinetic protein folding intermediate by equilibrium amide exchange. *Nature Struct. Biol.* **4**, 801–804.
18. Hosszu, L. L. P., Craven, C. J., Spencer, J., Parker, M. J., Clarke, A. R., Kelly, M. & Waltho, J. P. (1997). Is the structure of the N-domain of phosphoglycerate kinase affected by isolation from the intact molecule? *Biochemistry*, **36**, 333–340.
19. Clarke, A. R. & Waltho, J. P. (1997). Protein folding pathways and intermediates. *Curr. Opin. Biotech.* **8**, 400–410.
20. Reed, M. A. C., Hounslow, A. M., Sze, K. H., Barsukov, I. G., Hosszu, L. L. P., Clarke, A. R. *et al*. (2003). Effects of domain dissection on the folding and stability of the 43 kDa protein PGK probed by NMR. *J. Mol. Biol.* **330**, 1189–1201.
21. 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.
22. Kay, M. S. & Baldwin, R. L. (1996). Packing interactions in the apomyglobin folding intermediate. *Nature Struct. Biol.* **3**, 439–445.
23. Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M. & Kay, L. E. (2004). Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature*, **430**, 586–590.
24. Baxter, N. J., Hosszu, L. L. P., Waltho, J. P. & Williamson, M. P. (1998). Characterisation of low free-energy excited states of folded proteins. *J. Mol. Biol.* **284**, 1625–1639.
25. Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M. & Fersht, A. R. (1990). Transient folding intermediates characterized by protein engineering. *Nature*, **346**, 440–445.
26. Parker, M. J., Spencer, J. & Clarke, A. R. (1995). An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J. Mol. Biol.* **253**, 771–786.
27. Staniforth, R. A., Dean, J. L. E., Zhong, Q., Zerovnik, E., Clarke, A. R. & Waltho, J. P. (2000). The major transition state in folding need not involve the immobilization of side chains. *Proc. Natl Acad. Sci. USA*, **97**, 5790–5795.
28. Millett, I. S., Doniach, S. & Plaxco, K. W. (2002). Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Advan. Protein Chem.* **62**, 241–262.
29. Zerovnik, E., Lenarcic, B., Jerala, R. & Turk, V. (1991). Folding studies of the cysteine proteinase inhibitor—human stefin A. *Biochim. Biophys. Acta*, **1078**, 313–320.
30. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1992). The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, **31**, 1647–1651.
31. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
32. Munoz, V. & Serrano, L. (1997). Development of the multiple sequence approximation within the AGADIR model of α-helix formation: comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers*, **41**, 495–509.
33. Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. & Bax, A. (1997). Use of dipolar ^{1}H–^{15}N and ^{1}H–^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nature Struct. Biol.* **4**, 732–738.
34. Tycko, R., Blanco, F. J. & Ishii, Y. (2000). Alignment of biopolymers in strained gels: a new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. *J. Am. Chem. Soc.* **9340**, 9341.
35. Louhivuori, M., Paakkonen, K., Fredriksson, K., Permi, P., Lounila, J. & Annila, A. (2003). On the origin of residual dipolar couplings from denatured proteins. *J. Am. Chem. Soc.* **125**, 15647–15650.
36. Mohana-Borges, R., Goto, N. K., Kroon, G. J., Dyson, H. J. & Wright, P. E. (2004). Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol.* **340**, 1131–1142.
37. Ptitsyn, O. B. (1995). Molten globule and protein folding. *Advan. Protein Chem.* **47**, 83–229.
38. 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.
39. McParland, V. J., Kalverda, A. P., Homans, S. W. & Radford, S. E. (2002). Structural properties of an amyloid precursor of β_2-microglobulin. *Nature Struct. Biol.* **9**, 326–331.
40. Kuwata, K., Shastry, M. C. R., Cheng, H., Hoshima, M., Batt, C. A., Goto, Y. & Roder, H. (2001). Structural and kinetic characterization of early folding events in beta-lactoglobulin. *Nature Struct. Biol.* **8**, 151–155.

41. Zhang, O. & Forman-Kay, J. D. (1995). Structural characterisation of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry*, **34**, 6784–6794.
42. Zhang, O. & Forman-Kay, J. D. (1997). NMR studies of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry*, **36**, 3959–3970.
43. Baskakov, I. V. & Bolen, D. W. (1998). Monitoring the sizes of denatured ensembles of staphylococcal nuclease proteins: implications regarding *m* values, intermediates, and thermodynamics. *Biochemistry*, **37**, 18010–18017.
44. Bax, A. (1994). Multidimensional nuclear magnetic resonance methods for protein studies. *Curr. Opin. Struct. Biol.* **4**, 738–744.

Edited by K. Kuwajima

(Received 28 September 2005; received in revised form 22 December 2005; accepted 26 December 2005)
Available online 11 January 2006