Critical Reviews in Biochemistry and Molecular Biology, 40:181–189, 2005

ISSN: 1040-9238 print / 1549-7798 online
DOI: 10.1080/10409230591008143

Is There or Isn't There? The Case for (and Against) Residual Structure in Chemically Denatured Proteins

Evan R. McCarney
Department of Chemistry and Biochemistry, University of California, Santa Barbara Santa, Barbara, CA, USA

Jonathan E. Kohn
Interdepartmental Program in Biomolecular Science and Engineering, University of California, Santa Barbara, Santa Barbara, CA, USA

Kevin W. Plaxco
Department of Chemistry and Biochemistry and Interdepartmental Program in Biomolecular Science and Engineering, University of California, Santa Barbara, Santa Barbara, CA, USA

ABSTRACT First raised some 60 years ago, the question of whether chemically denatured proteins are fully unfolded has, in recent years, seen significantly renewed interest. This increased attention has been spurred, in large part, by new spectroscopic and computational approaches that suggest even the most highly denatured polypeptides contain significant residual structure. In contrast, the most recent scattering results uphold the long-standing view that chemically denatured proteins adopt random coil configurations. Here we review the evidence both for and against residual structure in chemically denatured proteins, and attempt to reconcile these seemingly contradictory observations.

KEYWORDS random-coil, protein folding, guanidine, urea

INTRODUCTION

Recent years have seen numerous and compelling reports of residual denatured state structure in even the most highly denatured proteins. For example, nuclear magnetic resonance (NMR) studies suggest that significant secondary structure and long-range hydrophobic clusters persist in unfolded proteins even at high concentrations of urea or guanidine hydrochloride (GuHCl) (Neri *et al.*, 1992; Kazmirski *et al.*, 2001; Hodsdon & Frieden, 2001; Klein-Seetharaman *et al.*, 2002). Similar recent reports have suggested that even the most highly denatured proteins exhibit residual long-range order similar to the native topology (Shortle & Ackerman, 2001; Ohnishi *et al.*, 2004). In contrast to these claims of significant residual denatured-state structure, however, intrinsic viscosity, hydrodynamic radii, and small-angle scattering experiments–some of which date back almost four decades–have been taken as evidence that the unfolded state is an effectively random coil ensemble (Tanford *et al.*, 1966; Wilkins *et al.*, 1999; Tcherkasskaya & Uversky, 2001; Kohn *et al.*, 2004). Given that residual denatured state structure would presumably play a significant role in both folding thermodynamics and kinetics (reviewed in Shortle, 1996), the way in which these seemingly mutually exclusive viewpoints might be reconciled remains a key question in protein biophysics. Here we review the evidence in favor and against the formation of significant structure in urea and GuHCl-unfolded polypeptides (which we will collectively

Editor: Michael M. Cox

Address correspondence to: Kevin W. Plaxco, University of California at Santa Barbara, Santa Barbara, CA 93106. E-mail: kwp@chem.ucsb.edu

refer to as the "chemically denatured state"), and attempt to reconcile this evidence into a coherent picture of the denatured state.

THE RANDOM COIL VIEW

As far back as 1936, Mirsky and Pauling noted that, based on the loss of activity, the solvent accessibility of functional groups and the large entropy associated with denaturation, "the denatured protein molecule we consider to be characterized by the absence of a uniquely defined structure" (Mirsky & Pauling, 1936). Some two decades later, Anfinsen and coworkers found that the oxidation of reduced ribonuclease in 8 M urea lead ultimately to the recovery of just ~1% of the protein's native activity, quite close to the amount expected were the protein's four disulfide bonds to reform randomly from a highly unstructured denatured state (reviewed in Anfinsen, 1973). Ten years later, Tanford used intrinsic viscosity measurements to put these speculations on even more quantitative grounds (as detailed below), concluding that, under highly denaturing conditions, proteins "are true random coils retaining no element of their original native conformation" (Tanford *et al.*, 1966). On the strength of these claims, the random-coil model became the standard reference state for interpretation of experimental data regarding unfolded proteins, and the starting point for most theoretical considerations of the folding process. Here we review the spectroscopic and scattering evidence supporting this view of the chemically denatured state.

Sequence-Local Structure?

Even in the most expanded, unfolded protein, sequence-local elements of the chain are held in proximity via the peptide bond. Taking this into account, Ramachandran (1968) observed that steric considerations alone restrict permissible ϕ/ψ angles and thus, at an arbitrarily local level, the conformation of all polypeptide chains is at least somewhat ordered. Questions remain, however, regarding the length scale over which the distribution of ϕ/ψ angles becomes uncorrelated, whether side chain-side chain interactions contribute to any deviations from a statistical population of backbone torsion angles, and whether heterogeneity in such sequence-local structure significantly affects the more global random-coil behavior of chemically denatured proteins.

Experimental probes of sequence-local structure in chemically denatured proteins are, unfortunately, relatively limited. For example, the traditional NMR signatures of sequence-local structure, such as medium-range nuclear Overhauser effects (NOEs), are rather rare in those proteins for which denatured state spectra have been assigned (Meekhof & Freund, 1999). The reasons for the paucity of medium-range NOEs remains unclear; it could arise due to excessive spectral averaging brought about by the highly dynamic nature of the unfolded state, or it could simply reflect a lack of significant structure even over the relatively short, ~5Å distance typically associated with NOEs (Meekhof & Freund, 1999). Consistent with the latter claim, Fiebig and colleagues have shown that the pattern of medium-range NOEs recorded for urea-denatured lysozyme corresponds reasonably well with the pattern expected for a random ensemble (Fiebig *et al.*, 1996). In the relative absence of these telltale medium-range NOEs, most authorities have relied on intra-residue signals, such as departures from random-coil chemical shifts (termed "secondary chemical shifts") and J coupling constants, in order to identify conformational biases that might produce significant deviations from random-coil behavior.

Intra-residue NMR parameters obtained from chemically denatured proteins are largely consistent with those expected for a random coil ensemble. For example, while the formation of well populated helices and sheets typically produce C_α secondary chemical shifts above 3 and below −1.5 ppm, respectively (Dyson & Wright, 2001), in chemically denatured proteins the vast majority of C_α secondary chemical shifts are less than one quarter of these values (Figure 1) (Frank *et al.*, 1995; Meekhof and Freund, 1999; Yi *et al.*, 2000; Bhavesh *et al.*, 2001; Kazmirski *et al.*, 2001; Garcia *et al.*, 2001; Tafer *et al.*, 2004; Bhavesh*et al.*, 2004). Consistent with the paucity of significantly perturbed backbone chemical shifts, the majority of the $^3J_{\mathrm{HN}\alpha}$ coupling constants for denatured proteins, which serve as a measure of the ϕ torsion angle, are indistinguishable from those expected for a random-coil ensemble, either as derived from a Boltzman distribution of energetically preferred conformations (Feibig *et al.*, 1996) or via observed ϕ/ψ preferences derived from native structures (Smith, 1996). Perhaps not surprisingly, side-chain χ_1 angles adopted by urea-unfolded lysozyme (derived from $^3J[\mathrm{N},\mathrm{C}^\gamma]$ coupling constants) also correspond closely to those expected for a random-coil ensemble of states (Henning *et al.*, 1999), suggesting minimal

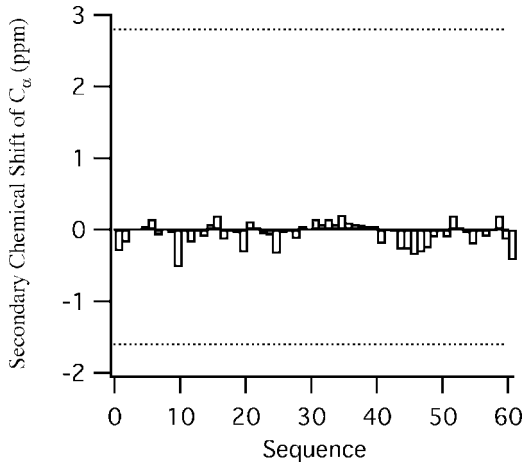

FIGURE 1 **The secondary chemical shifts of most chemically denatured proteins are consistent with those expected for a random-coil conformation. Shown here are the C_α chemical shifts of GuHCl-unfolded protein L. The observed secondary chemical shifts fall far below the values expected for well populated α-helices or β-sheets (upper and lower dotted lines respectively). Figure adapted with permission from Yi *et al.* (2000).**

residue-to-residue interactions are occurring in this denatured state. It thus appears that the large majority of residues in chemically denatured proteins adopt local conformations closely consistent with a statistical population of energetically preferred torsion angles.

Long-Range Structure?

Traditional spectroscopic methods provide relatively few clues to support interactions between sequence-distant amino acids (here termed "long-range structure") or other evidence for long-range order in chemically denatured proteins; the detection and assignment of long-range NOEs in the unfolded state has proven extremely difficult and no very compelling example of sequence-distant interactions has yet been reported for chemically denatured proteins (Dyson & Wright, 2004). Instead, studies of long-range structure in the denatured state have focused on more global, less site-specific measures.

Extensive small-angle X-ray and neutron scattering (SAXS and SANS, respectively) provide strong, qualitative support for the random-coil nature of chemically denatured proteins. For example, whereas the Kratky scattering profiles obtained from native and even molten-globule proteins exhibit a sharp peak indicative of long-range order (Kataoka & Goto, 1996; Kataoka *et al.*, 1997), the Kratkys of chemically denatured proteins almost invariably exhibit the monotonic rise expected for an expanded, random-flight ensemble of conformations (reviewed in Millet *et al.*, 2002). Consistent with this in finding, the scattering profiles of 4M GuHCl-denatured neocarzinostatin and phosphoglycerate kinase are well modeled as an excluded volume random coil (Russo *et al.*, 2000, 2001; Petrescu *et al.*, 1998).

More quantitative evidence in favor of the random coil model is provided by studies of the extrinsic properties and mean dimensions of chemically denatured proteins. A hallmark of ideal random-coil behavior is a simple power-law relationship between a polymer's length and both extrinsic properties, such as viscosity (η), and measures of overall dimensions, such as the radius of gyration (R_g):

$$\eta = (K/M_o)N^\nu \quad (1)$$

$$R_g = R_0 N^\nu \quad (2)$$

where N is the number of monomers in the polymer chain, M_o is the mean residue weight, K and R_0 are proportionality constants that are, among other things, functions of the monomer chemistry and persistence length of the polymer, and ν is an exponential scaling factor. For an excluded volume polymer (*i.e.*, a real polymer with non-zero thickness and non-trivial interactions between monomers) in a good solvent, such as a high denaturant, Flory estimated ν to be $\sim$3/5 (Flory, 1953), and more precise follow-up estimates stemming from renormalization group models indicate that $\nu = 0.588$ (LeGuillou & Zinn-Justin, 1977).

The formation of persistent denatured-state structure will lead to perturbations from the above described, idealized random-coil relationships. For example, the formation of hydrophobically stabilized clusters will cause a net compaction and a smaller R_0 (and thus smaller R_g) than that of a polymer lacking such interactions. Similarly, the presence of persistent local structure could increase (or decrease) a polymer's mean persistence length, leading to increases (or decreases) in both R_G and η. Differences in the magnitude of the residual denatured state structure, be it local (and thus affecting persistence length) or non-local, from protein to protein would thus be expected to produce significant scatter around any underlying power-law relationship.

Tellingly, the extrinsic properties and dimensions measured for nearly all urea- or GuHCl-denatured, crosslink-free proteins closely obey the theoretically expected random-coil scaling. Studies of this issue date from the late sixties, when Tanford and coworkers

(1966) employed viscosity measurements to determine that, for a set of 11 proteins unfolded in 6 M GuHCl, $\nu = 0.67 \pm 0.09$ (95% confidence interval).More recent, and somewhat more direct, studies have shown that the hydrodynamic radii of sets of 8 and 38 highly denatured, disulfide-free proteins fit power-law relationships with $\nu = 0.57 \pm 0.05$ and ~0.64 respectively (Wilkins *et al.*, 1999; Tcherkasskaya and Uversky, 2001). More recently still, we have critically re-evaluated the previous SAXS literature and determined the R_g of a number of previously uncharacterized, chemically denatured proteins and peptides in order to generate a dataset of 28 SAXS-derived R_g values (Kohn *et al.*, 2004). Notably, only 2 of these 28 chemically denatured, prosthetic group- and crosslink-free proteins (which span the range from 8 to 549 residues) represent experimentally significant outliers from the best-fit power law (Figure 2). Moreover, the experimentally determined exponent, $\nu = 0.598 \pm 0.028$, is well within error of that predicted by theory for a random coil polymer (Kohn *et al.*, 2004). Consistent with these results, simulations of random-flight polymers derived using only excluded volume constraints produce ensembles whose average dimensions closely match those of denatured CheY (Garcia *et al.*, 2001), phosophoglycerate kinase (Calmettes *et al.*, 1994), and several other proteins (Goldenberg *et al.*, 2003).

THE CASE FOR STRUCTURE

In contrast to the studies described above, a large body of spectroscopic and simulation-based work suggests that proteins retain sequence-local and, perhaps, long-range structure under even the most strongly denaturing conditions. Here we review the evidence in favor of such residual structure.

Local Structure

More than 30 years ago, Tiffany and Krimm (1973) suggested that the UV-CD of denatured proteins is best explained by the presence of a significant population of polyproline-II helix (P_{II}), even for sequences of low proline content. Recent years have seen significant additional evidence supporting this assertion, with NMR, Raman optical activity, vibrational CD and far-UV CD spectral features put forth in support of an important role played by this structure in the unfolded state (reviewed by Barron *et al.*, 2002; Keiderling and Xu, 2002; Shi *et al.*, 2002; Mohana-Borges *et al.*, 2004). Most of these studies, however, have been conducted using short, synthetic peptides unfolded in the absence of denaturant. In contrast, Kallenbach and coworkers recently employed far-UV CD to argue that the P_{II} helix population of the sequence $O_2A_7O_2$ (O = ornithine)

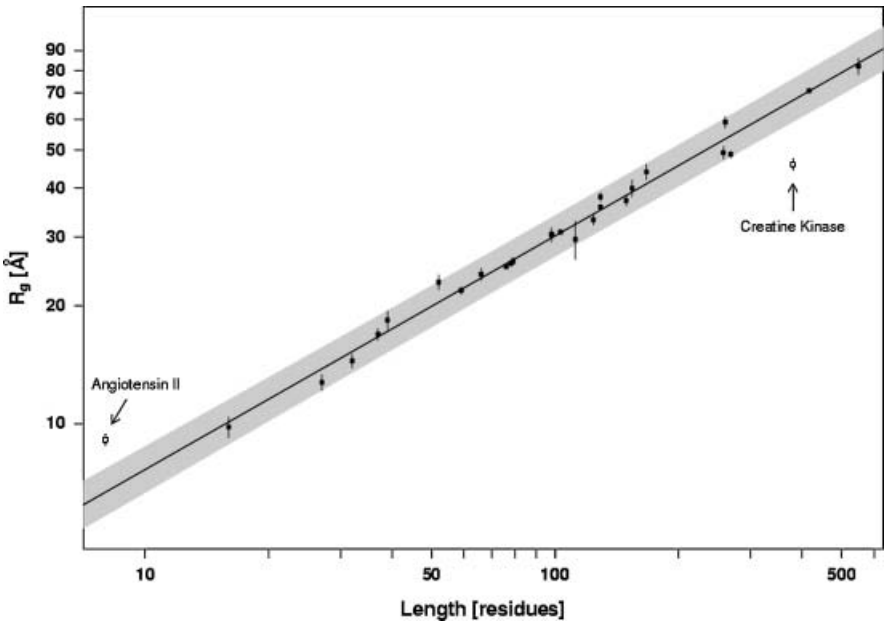

FIGURE 2 **The R_g of the large majority of chemically-denatured proteins scale with the power-law relationship expected for a random coil. The solid line, which is a least-squares fit ignoring the two noted outliers, produces an exponent $\nu = 0.598 \pm 0.028$, that is indistinguishable from the predicted 0.588 (Leguillou & Zinnjustin, 1977). The shaded region represents 95% confidence intervals around this relationship. Only 2 of 28 reported, non-crosslinked, chemically-denatured proteins represent statistically significant outliers from the observed scaling. Figure adapted with permission from Kohn *et al.* (2004).**

increases significantly with increasing GuHCl concentration (Liu *et al.*, 2004). The ensemble-averaged conformation of the peptide GGXGG, however, appears to be independent of GuHCl concentration when X is alanine, lysine, glutamate or valine (Plaxco *et al.*, 1997), and thus it is uncertain whether these observations are generalizable.

Tiffany and Krimm (1973) attributed the observed P_{II} helix to a Boltzman distribution of energetic states of dipeptide pairs along the polypeptide backbone. That is, in the absence of sequence-specific side chain interactions, the backbone of the polypeptide will preferentially populate ϕ/ψ angles in the lowest, broadest energy wells. These regions happen to be the adjacent energy wells in the upper left quadrant of the Ramachandran energy landscape, which typically correspond to the β-sheet and P_{II} helix. Experimental support for the enthalpically favorable nature of the P_{II} well includes the observation that, as urea denatured T7 RNA polymerase is heated, the P_{II} denatured structure will start to move toward a more randomized population of conformations, whereas upon cooling, P_{II} structure is stabilized (Griko *et al.*, 2001). Detailed analysis of this hypothesis via atomistic simulations parallels the experimental results. Pappu and Rose (2002) have tested the key characteristics that define an excluded volume random coil by minimizing chain packing with the only restrictions being excluded volume and chain connectivity. To achieve this, they employed a purely repulsive soft-core potential to an alanine dipeptide and a seven-residue alanine polypeptide. They observed that, under these conditions, there is a preference for ϕ/ψ values that correspond to the same global minimum characteristic of a P_{II} helix and its symmetry mate, a right-handed polyproline helix. A significant and still unanswered question, however, is whether the formation of P_{II} structure is sequence-dependent. If the formation of P_{II} is sequence independent, and all sequences populate P_{II} with equal frequency then we will simply observe a random coil comprised of P_{II} helices.

Rose and coworkers have also tested another characteristic of excluded volume random coils defined by Flory's isolated pair hypothesis (Pappu *et al.*, 2000). According to Flory, conformational restrictions are limited to nearest neighbor interactions (Flory, 1969). If this holds true, steric restrictions alone should be insufficient to stabilize local structure beyond that arising from nearest neighbor interactions. Modeling of the steric hindrance of bulky side-chains suggests that such interactions can propagate beyond nearest neighbor interactions and potentially stabilize regions of local structure (Pappu *et al.*, 2000). A notable result is that the only ϕ/ψ pairs that sufficiently satisfy the Flory isolated pair hypothesis are located in the P_{II} region, thus providing an additional mechanistic rational for the above described evidence in favor of such structure.

A potential bias towards P_{II} torsion angles is not the only sequence-local conformational bias reported for chemically denatured proteins. For example, Hosur and coworkers have used C_α secondary chemical shifts to identify regions in 8M urea-unfolded bar-star that may adopt α–helical or extended configurations (Bhavesh *et al.*, 2004). Perhaps surprisingly, the pattern of secondary chemical shifts observed appear to indicate that, while segments of the chain that are helical in the native protein are random-coil in the unfolded state, several contiguous residues that form a β-strand in the native state populate nonnative, α–helical ϕ/ψ angles in the urea unfolded state. A similar pattern of both native-like and non-native secondary chemical shifts is reported for the HIV protease in 6M GuHCl (Bhavesh *et al.*, 2001). Further evidence for helix formation in chemically-denatured proteins comes from Bruix and coworkers, who have reported detectable helical content in the urea-unfolded state of CheY (Garcia *et al.*, 2001), and Fersht and coworkers who report that the GuHCl-unfolded state of CI2 "includes a partially populated portion of the native helix" (albeit of just 3 residues) (Kazmirski *et al.*, 2001). Residual populations of β-turns have likewise been identified in the GuHCl-denatured states of protein L (Yi *et al.*, 2000) via chemical shift patterns, medium-range NOEs and spin labeling, and in urea-denatured Ig18′ via NOE and secondary chemical shift patterns (Fong *et al.*, 1998). Unfortunately, however, neither the equilibrium constant for the population of these torsion angles nor the extent to which the resides in question adopt regular torsion angles in concert (*i.e.*, the extent to which they represent a true element of secondary structure rather than a random mixture of rotamers, some of which are helical or turn-like at any given instant) have been established for any of the above denatured states. In contrast, Wuthrich and coworkers have used chemical shifts, NOE patterns and constrained simulations to identify a 10-residue sequence in the urea-unfolded state of OmpX that populates a helix an estimated 25% of the time (Tafer *et al.*, 2004). Despite these many examples, however, the generality of such structure has not been established,

and, indeed, the population of secondary structure appears to be undetectably low in many other chemically denatured proteins; for example, no significantly perturbed chemical shifts are observed in the urea-unfolded states of the helical protein apomyoglobin (Yao *et al.*, 2001), the sheet protein TnFNIII (Meekhof & Freund, 1999) or the mixed helix-sheet protein G (Frank *et al.*, 1995).

In addition to the evidence in favor of sequence-local secondary structure, limited evidence also support the formation of sequence-local structure arising due to hydrophobic interactions persisting even at high concentrations of denaturant. For example, Schwarzinger and colleagues (2002) have observed reduced chain dynamics (as monitored by T2 relaxation times) in urea-unfolded apomyoglobin that they attribute to the formation of persistent, sequence-local hydrophobic interactions, and Kazmirski and colleagues (2001) report perturbed chemical shifts and relaxation time constants consistent with the formation of "a small amount of hydrophobic clustering" across a stretch of 7 residues in GuHCl-denatured CI2. In contrast to these rather qualitative claims, Wuthrich and co-workers have observed similar sequence-local hydrophobic clustering in urea-denatured OmpX that they estimate is populated at the ~25% level (Tafer *et al.*, 2004). Here too, however, it remains unclear how general these effects are; no similar evidence of clustering is apparent in the urea-unfolded states of protein G (Frank *et al.*, 1995), apomyoglobin (Yao *et al.*, 2001) or TnFNIII (Meekhof & Freund, 1999).

Long-Range Structure

Spectroscopic evidence in favor of long-range structure persisting even at very high levels of chemical denaturant has been reported for a number of proteins. Often cited as the first example of such long-range structure in a denatured protein, Neri and colleagues reported in 1992 that long-range NOEs observed in the urea-unfolded state of 434-repressor are consistent with the formation of a well-populated, sequence-distant hydrophobic cluster spanning 6 residues across the protein's primary sequence. More recently, the occurrence of altered relaxation lifetimes in reduced hen egg-white lysozyme has been taken as evidence for well populated, long-range clusters of hydrophobic residues persisting in 8M urea (Schwalbe *et al.*, 1997; Klein-Seetharaman *et al.*, 2002). The clusters in question appear to be related to non-native interactions between the protein's six tryptophan residues and are abolished or reduced upon their mutation (Wirmer *et al.*, 2004). Similar hydrophobic clustering has been reported for urea-unfolded rat intestinal fatty acid binding protein (Hodsdon & Frieden, 2001) and, as indicated by slow exchange of histidine protons, in the 5M urea-unfolded, α-subunit of tryptophan synthase (Saab-Rincon *et al.*, 1996).

A second line of evidence that has been taken by some to indicate that long-range structure occurs even in highly denatured proteins has been the observation by NMR of residual dipolar couplings (RDC). RDCs can be observed when even a relatively small population of molecules adopts a net asymmetric orientation with respect to the externally applied magnetic field. In solution this occurs when non-spherical conformations experience an asymmetric environment, such as the pores of a mechanically stressed acrylamide gel or in the presence of liquid crystals. Under such circumstances, a protein structure that is asymmetric will tumble anisotropically, preventing the one-bond internuclear dipoles from achieving their rotationally averaged value of zero. This results in a net RDC signal that depends on the magnitude of the asymmetry and the orientation of the bond vector relative to the externally applied magnetic field (Bax, 2003). As such, RDCs are evidence for structural asymmetries that persist on the NMR timescale. To date RDCs have been observed for the chemically denatured states of a large number of proteins (Shortle & Ackerman, 2001; Louhivuori *et al.*, 2003, 2004; Mohana-Borges *et al.*, 2004; Fieber, *et al.*, 2004; Ding *et al.*, 2004; Ohnishi *et al.*, 2004), suggesting that even seemingly highly unfolded proteins may be hiding persistent long-range structure. The precise nature of this structure, however, remains controversial (Louhivuori *et al.*, 2003, 2004; Mohana-Borges *et al.*, 2004).

Shortle and coworkers (2004) have suggested that the observed denatured-state RDCs provide evidence that a protein's native topology persists even under highly denaturing conditions (Onishi *et al.*, 2004). They note that the RDCs of urea-denatured staphylococcal nuclease are correlated with those observed for an intrinsically-unfolded truncation mutation (Δ131Δ) that is known to be largely unfolded in the absence of denaturant and yet adopts a native-like topology (Shortle & Ackerman, 2001). Via this, slightly indirect line of reasoning Shortle and Ackerman (2001) argue that the urea-denatured

molecule also adopts a native-like topology, albeit they note that the RDCs of neither the truncation mutant nor the urea-denatured state correlate with those of native Snase. More recently, however, the RDCs from 8M urea-denatured Eglin C have been shown to correlate ($r = \sim 0.5$) with those of the native protein (Ohnishi *et al.*, 2004), providing more direct evidenced in favor of native-like long-range order in a chemically denatured protein.

RECONCILING THE RANDOM COIL WITH RESIDUAL STRUCTURE

How, then, should we view the chemically denatured state? We have argued that, at least with regard to chemically denatured proteins, Mirsky and Pauling's speculation of almost 70 years ago has withstood the test of time: to a very good first approximation, chemically denatured proteins "are characterized by the absence of a uniquely defined structure." That said, of course, we have also outlined the detailed arguments in favor of "significant" sequence-local and long-range residual structure. Here we conclude that, while this sequence-local structure is easily reconciled with the random coil model of the denatured state, the putative long-range structure is rather less so.

Sequence-local structure appears to be relatively rare in chemically denatured proteins and, when it does occur, is unlikely to affect significantly otherwise random-coil behavior. For example, while Liu and coworkers have argued that P_{II} helical ϕ/ψ preferences are generally adopted in denatured proteins (Liu *et al.*, 2004), spectroscopic studies indicate that sequence-specific local structure of the type that would lead to specific, protein-to-protein deviations from random-coil behavior is uncommon; as we have described above, extremely few of the C_α and $^3J_{HN\alpha}$ assignments reported for chemically-denatured proteins deviate significantly from random-coil values, and the few deviations that have been characterized quantitatively arise due to structure populated only at the 25% level (Tafer *et al.*, 2004). Moreover, even if sequence-local structure were well populated, Fitzkee and Rose (2004) have recently shown via simulation that it might have little effect on more global random coil behaviors; R_g, for example, appears to be surprisingly insensitive to the presence or absence of such structure.

It is perhaps more difficult to reconcile reports of long-range order with the random-coil scaling observed for the dimensions of the chemically denatured-state. This is because, in contrast to local structure, the formation of long-range structure would presumably affect R_g significantly. For example, the R_g of chemically denatured reduced Rnase A and lysozyme expand by 30% and 60%, respectively, upon reduction of their disulfide bonds (Sosnick & Trewhella, 1992; Hoshino *et al.*, 1997; Kohn *et al.*, 2004), and breaking a single (non-natural) crosslink increases the R_g of chemically denatured dihydrofolate reductase by 45% (Arai *et al.*, 2003). Few similarly large deviations from ideal random-coil scaling are observed in non-crosslinked, chemically denatured proteins, suggesting that well populated hydrophobic clusters, which would presumably pin together sequence-distant structural elements, are rare (Kohn *et al.*, 2004). More generally, Miller and Goebel (1968) have calculated that the formation of "irregular knots of compact structure" can be consistent with random-coil-like dimensional scaling but only if the fraction of residues that participate in the collapsed structure is the same in all proteins. Any sequence-specific variation in the size of these putative compact structures would produce deviations from random-coil scaling of approximately the same magnitude as the fraction of residues involved in the clusters. In this light, we have previously estimated that less than 3% of the residues in a typical chemically denatured protein are participating in compact, nonregular structures at any given instant (Kohn *et al.*, 2004). Similarly, we assume, the adoption of a native-like topology would produce significant scatter in the random-coil scaling of R_g. Given, however, that dipolar couplings can range up to many kHz and the observed RDCs are only a few Hz, it may be possible that the observed signals arise from rare deviations from random-coil distributions. This issue highlights a general means of reconciling spectroscopic evidence of long-range order with the ample evidence in favor of the random coil model. Namely, spectroscopic probes may be observing large signals arising from rare, highly ordered conformations instead of small signals arising from more representative structures.

It thus appears that, despite nearly 70 years of investigation, chemically denatured proteins remain very well approximated as random coils. And while the spectroscopic studies we have reviewed here may be revealing the first cracks in the model, we suspect the random-coil approximation will be with us for some time to come.

ACKNOWLEDGMENTS

The authors wish to thank David Baker for permission to adapt figure 1. Portions of work were supported by funds from the NIH (GM62868-01A2).

REFERENCES

Anfinsen, C.B. 1973. Principles that govern folding of protein chains. *Science* 181:223–230.

Arai, M., Kataoka, M., Kuwajima, K., Matthews, C.R., and Iwakura, M. 2003. Effects of the difference in the unfolded-state ensemble on the folding of *Escherichia coli* Dihydrofolate Reductase. *J Mol Biol* 329:779–791.

Barron, L.D., Blanch, E.W., and Hecht, L. 2002. Unfolded proteins studied by Raman optical activity. In: Rose, G.D. ed., Advances in Protein Chemistry, Vol. 62. pp. 51–90. Academic Press, San Diego.

Bhavesh, N.S., Panchal, S.C., Mittal, R., and Hosur, R.V. 2001. NMR identification of local structural preferences in HIV-1 protease tethered heterodimer in 6 M guanidine hydrochloride. *FEBS Lett* 509:218–224.

Bhavesh, N.S., Juneja, J., Udgaonkar J.B., and Hosur, R.V. 2004. Native and nonnative conformational preferences in the urea-unfolded state of barstar. *Prot Sci* 13:3085–3091.

Calmettes, P., Durand, D., Desmadril, M., Minard, P., Receveur, V., and Smith, J.C. 1994. How random is a highly denatured protein. *Biophys Chem* 53:105–113.

Ding, K., Louis, J.M., and Gronenborn, A.M. Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR. *J Mol Biol* 335:1299–1307.

Dyson, H.J. and Wright, P.E. 2001. Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Meth Enz* 339:258–270.

Dyson, H.J. and Wright, P.E. 2004. Unfolded proteins and protein folding studied by NMR. *Chem Rev* 104:3607–3622.

Fieber, W., Kristjansdottir, S., and Poulsen, F.M. 2004. Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings. *J Mol Biol* 339:1191–1199.

Fitzkee, N.C. and Rose, G.D. 2004. Reassessing random-coil statistics in unfolded proteins.*Proc Natl Acad Sci USA* 101:12497–12502.

Flory, P.J. 1969. Statistical Mechanics of chain molecules. Wiley, New York, N.Y.

Fong, S., Bycroft, M., Clarke, J., and Freund, S.M.V. 1998. Characterisation of urea-denatured states of an immunoglobulin superfamily domain by heteronuclear NMR. *J Mol Biol* 278:417–429.

Frank, M.K., Clore, G.M., and Gronenborn, A.M. 1995. Structural and dynamic characterization of the urea denatured state of the immunoglobulin binding domain of streptococcal protein-g by multidimensional heteronuclear nmr-spectroscopy. *Prot Sci* 4:2605–2615.

Garcia, P., Serrano, L., Durand, D., Rico, M., and Bruix, M. 2001. NMR and SAXS characterization of the denatured state of the chemotactic protein CheY: Implications for protein folding initiation. *Prot Sci* 10:1100–1112.

Goldenberg, D.P. 2003. Computational simulation of the statistical properties of unfolded proteins. *J Mol Biol* 326:1615–1633.

Griko, Y., Sreerama, N., Osumi-Davis, P., Woody, R.W., and Wood, A.Y.M. 2001. Thermal and urea-induced unfolding in T7 RNA polymerase: Calorimetry, circular dichroism and fluorescence study. *Prot Sci* 10:845–853.

Hennig, M., Bermel, W., Spencer, A., Dobson, C.M., Smith, L.J., and Schwalbe, H. 1999. Side-chain conformations in an unfolded protein: c_1 distributions in denatured hen lysozyme determined by heteronuclear ^{13}C, ^{15}N NMR spectroscopy. *J Mol Biol* 288:705–723.

Hodsdon, M.E. and Frieden, C. 2001. Intestinal fatty acid binding protein: The folding mechanism as determined by NMR studies. *Biochemistry* 40:732–742.

Hoshino, M., Hagihara, Y., Hamada, D., Kataoka, M., and Goto, Y. 1997. Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering. *FEBS Lett* 416:72–76.

Kataoka, M. and Goto, Y. 1996. X-ray scattering studies of protein folding. *Fold Des* 1:R107–R114.

Kataoka, M., Kuwajima, K., Tokunaga, F., and Goto, Y. 1997. Structural characterization of the molten globule of alpha-lactalbumin by solution X-ray scattering. *Prot Sci* 6:422–430.

Kazmirski, S.L., Wong, K.B., Freund, S.M.V., Tan, Y.J., Fersht, A.R., and Daggett, V. 2001. Protein folding from a highly disordered denatured state: The folding pathway of chymotrypsin inhibitor 2 at atomic resolution.*Proc Natl Acad Sci USA* 98:4349–4354.

Keiderling, T.A. and Xu, Q. 2002. Unfolded peptides and proteins studied with infrared absorption and vibrational circular dichroism spectra. In: Rose, G.D. ed., Advances in Protein Chemistry, Vol. 62. pp. 91–162. Academic Press, San Diego.

Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M., and Schwalbe, H. 2002. Long-range interactions within a nonnative protein. *Science* 295:1719–1722.

Kohn, J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., *et al.* 2004. Random-coil behavior and the dimensions of chemically unfolded proteins *Proc Natl Acad Sci USA* 101:12491–12496.

Leguillou, J.C. and Zinnjustin, J. 1977. Critical exponents for N-vector model in 3 dimensions from field-theory. *Phys Rev Lett* 39:95–98.

Liu, Z., Chen, K., Ng, A., Shi, Z., Woody, R.W., and Kallenbach, N.R. 2004. Solvent dependence of P_{II} conformation in model alanine peptides. *J Am Chem Soc* 126:15141–15150.

Louhivuori, M., Paakkonen, K., Fredriksson, K., Permi, P., Lounila, J., and Annila, A. 2003. On the origin of residual dipolar couplings from denatured proteins. *J Am Chem Soc* 125:15647–15650.

Louhivuori1, M., Fredriksson, K., Pääkkönen, K., Permi, P., and Annila, A. 2004. Alignment of chain-like molecules. *J Biomolec NMR* 29:517–524.

Meekhov, A.E. and Freund, S.M. 1999. Probing residual structure and backbone dynamics on the milli- to picosecond timescale in a urea-denatured fibronectin type II domain. *J Mol Biol* 286:579–592.

Millet, I., Doniach, S., and Plaxco, K.W. 2002. Toward a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins. In: Rose, G. D. Ed., Advances in Protein Chemistry, Vol. 62. pp. 241–262. Academic Press, San Diego.

Miller, W.G. and Goebel, C.V. 1968. Dimensions of protein random coils. *Biochemistry* 7:3925–3935.

Mirsky, A.E. and Pauling, L. 1936. On the structure of native, denatured, and coagulated proteins. *Proc Natl Acad Sci USA* 22:439–447.

Mohana-Borges, R., Goto, N.K., Kroon, G.J.A., Dyson, H.J., and Wright P.E. 2004 Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J Mol Biol* 340:1131–1142.

Neri, D. Billeter, M., Wider, G., and Wuthrich, K. 1992. NMR determination of residual structure in a urea-denatured protein, the 434-Repressor. *Science* 257:1559–1563.

Ohnishi, S., Lee, A.L., Edgell, M.H., and Shortle, D. 2004. Direct demonstration of structural similarity between native and denatured Eglin C. *Biochemistry* 43:4064–4070.

Pappu, R.V. and Rose, G.D. 2002. A simple model for polyproline II structure in unfolded states of alanine-based peptides. *Prot Sci* 11:2437–2455.

Pappu, R.V., Srinivasan, R., and Rose, G.D. 2000. The Flory isolated-pair hypothesis is not valid for polypeptide chains: Implications for protein folding. *Proc Natl Acad Sci USA* 97:12565–12570.

Petrescu, A.J., Receveur, V., Calmettes, P., Durand, D., and Smith, J.C. 1998. Excluded volume in the configurational distribution of a strongly-denatured protein. *Prot Sci* 7: 1396–1403.

Plaxco, K.W., Morton, C.J., Grimshaw, S., Jones, J.A., Pitkeathly, M., Campbell, I.D., and Dobson, C.M. 1997. The effects of guanidine hydrochloride on the "random coil" conformations and NMR chemical shifts of the peptide series GGXGG. *J Biomolec NMR* 10:221–230.

Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99.

Russo, D., Durand, D., Desmadril, M., and Calmettes, P. 2000. Study of thermally and chemically unfolded conformations of a small beta-protein by means of small-angle neutron scattering. *Physica B* 276:520–521.

Saab-Rincon, G., Gualfetti, P.J., and Matthews, C.R. 1996. Mutagenic and thermodynamic analyses of residual rtructure in the subunit of tryptophan synthase. *Biochemistry* 35:1988–1994.

Schwalbe, H., Fiebig, K.M., Buck, M., Jones, J.A., Grimshaw, S.B., Spencer, A., Glaser, S.J., Smith, L.J., and Dobson, C.M. 1997. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of Lysozyme in 8 M Urea. *Biochemistry* 36:8977–8991.

Schwarzinger, S., Wright, P.E., and Dyson, H.J. 2002. Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. *Biochemistry* 41:12681–12686.

Shi, Z., Woody, R.W., and Kallenbach, N.R. 2002. Is polyproline II a major backbone conformation in unfolded proteins? In: Rose, G.D. ed., Advances in Protein Chemistry, Vol. 62. pp. 163–240. Academic Press, San Diego.

Shi, Z., Olson, C.A., Rose, G.D., Baldwin, R.L., and Kallenbach N.R. 2002. Polyproline II structure in a sequence of seven alanine residues. *Proc Natl Acad Sci USA* 99:9190–9195.

Shortle, D. 1996. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J* 10:27–34.

Shortle, D. and Ackerman, M.S. 2001. Persistence of native-like topology in a denatured protein in 8 M Urea. *Science* 293:487–489.

Smith, L.J., Fiebib, K.M., Schwalbe, H., and Dobson, C.M. 1996. The concept of a random coil Residual structure in peptides and denatured proteins. *Fold Des* 1:R95–R106.

Sosnick, T.R. and Trewhella, J. 1992. Denatured states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry* 31:8329—8335

Tafer, H., Hiller, S., Hilty, C., Fernandez, C., and Wuthrich, K. 2004. Nonrandom structure in the urea-unfolded *Escherichia coli* outer membrane protein X (OmpX). *Biochemistry* 43:860–869.

Tanford, C., Kawahare, K., and Lapanje, S. 1966. Proteins in 6M guanidine hydrochloride demonstration of random coil behavior. *J Biol Chem* 241:1921–1923.

Tcherkasskaya, O. and Uversky, V.N. 2001. Denatured collapsed states in protein folding: Example of apomyoglobin. *Proteins Struct Funct Genet* 44:244–254.

Tiffany, M.L., and Krimm, S. 1973. Extended conformations of polypeptides and proteins in urea and guanidine hydrochloride. *Biopolymers* 12:575–587.

Wilkins, D.K., Grimshaw, S.B., Receveur, V., Dobson, C.M., Jones, J.A., and Smith, L.J. 1999. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38:16424–16431.

Wirmer, J., Schlorb, C., Klein-Seetharaman, J., Hirano, R., Ueda, T., Imoto, T., and Schwalbe, H. 2004. Modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations. *Angew Chem Int Ed* 43:5780–5785.

Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H.J. 2001. NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry* 40:3561–3571.

Yi, Q., Scalley-Kim, M.L., Alm, E.J., and Baker, D. 2000. NMR characterization of residual structure in the denatured state of protein L. *J Mol Biol* 299:1341–1351.