

Joanna F. Swain is in the Department of Biochemistry & Molecular Biology and Lila M. Gierasch is in the Department of Biochemistry & Molecular Biology and the Department of Chemistry at the University of Massachusetts, Amherst, Massachusetts 01003, USA. Correspondence should be addressed to J.F.S. email: feltham@nsm.umass.edu

- Brandts, J.F., Halvorson, H.R. & Brennan, M. *Biochemistry* **14**, 4953–4963 (1975).
- Schmid, F.X. & Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **75**, 4764–4768 (1978).
- Schiene, C. & Fischer, G. *Curr. Opin. Struct. Biol.* **10**, 40–45 (2000).
- Odefey, C., Mayr, L.M. & Schmid, F.X. *J. Mol. Biol.* **245**, 69–78 (1995).
- Pappenberger, G. *et al. Nature Struct. Biol.* **8**, 452–458 (2001).
- Schiene-Fischer, C., Habazettl, J., Schmid, F.X. & Fischer, G. *Nature Struct. Biol.* **9**, 419–424 (2002).
- Jabs, A., Weiss, M.S. & Hilgenfeld, R. *J. Mol. Biol.* **286**, 291–304 (1999).
- Stein, R.L. *Adv. Protein Chem.* **44**, 1–24 (1993).
- Scherer, G., Kramer, M.L., Schutkowski, M., Reimer, U. & Fischer, G. *J. Am. Chem. Soc.* **120**, 5568–5574 (1998).
- Lim, V.I. & Spirin, A.S. *J. Mol. Biol.* **188**, 565–574 (1986).
- Schiene-Fischer, C. & Fischer, G. *J. Am. Chem. Soc.* **123**, 6227–6231 (2001).
- Gragerov, A. & Gottesman, M.E. *J. Mol. Biol.* **241**, 133–135 (1994).
- Rüdiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau, B. *EMBO J.* **16**, 1501–1507 (1997).
- de Crouy-Chanel, A., Kohiyama, M. & Richarme, G. *J. Biol. Chem.* **271**, 15486–15490 (1996).
- Pierpaoli, E.V., Gisler, S.M. & Christen, P. *Biochemistry* **37**, 16741–16748 (1998).
- Zhu, X. *et al. Science* **272**, 1606–1614 (1996).
- Stoller, G. *et al. EMBO J.* **14**, 4939–4948 (1995).
- Teter, S.A. *et al. Cell* **97**, 755–765 (1999).
- Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A. & Bukau, B. *Nature* **400**, 693–696 (1999).
- Rüdiger, S., Buchberger, A. & Bukau, B. *Nature Struct. Biol.* **4**, 342–349 (1997).
- Kraulis, P. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Merritt, E.A. & Bacon, D.J. *Methods Enzymol.* **277**, 505–524 (1997).

Mini-proteins Trp the light fantastic

Samuel H. Gellman and Derek N. Woolfson

A new 20-residue peptide represents the smallest example to date of cooperatively folded tertiary structure. This achievement provides a new tool for elucidating protein conformational preferences. The mini-protein should serve as a fruitful platform for protein design.

On page 425 of this issue, Neidigh *et al.*¹ describe a remarkable peptide that adopts a small but well-defined globular shape. Formation of discrete tertiary structure is generally thought to be the exclusive province of much longer polypeptides. Although other short peptides with stable tertiary structure have been reported, the present case stands out in that only natural amino acid residues are employed and there is no crosslinking *via* disulfide formation, metal ion chelation or stabilization through oligomerization. The success of Neidigh *et al.*¹ appears to be primarily due to a structural motif the authors dub the ‘Trp cage’, in which the side chain of a Trp residue is penned in by several other residues, notably the side chains of prolines.

Development of the Trp cage began during examination of a natural 39-residue peptide from Gila monster saliva. The C-terminal portion of this peptide showed promise for folding, but structure was observed only in the presence of the fold-promoting cosolvent 2,2,2-trifluoroethanol (TFE). Recognizing the prospect for structural optimization, the authors pursued a series of incremental sequence modifications to enhance the stability of the folded conformation. These shrewd efforts culminated in a 20-residue Trp cage that appears fully structured at low temperature in water and displays several hallmarks of cooperatively folded proteins.

The Trp cage is cooperatively folded

The studies of protein structure and folding are mature fields of endeavor², and one is led to ask two fundamental questions regarding the system reported by Neidigh *et al.*¹: (i) Is the structure completely folded under accessible conditions? (ii) Is folding cooperative? This latter question is important because for most natural protein domains, all parts of the molecule undergo the folding transition in concert as conditions, such as temperature or denaturant concentration, are varied. This type of cooperativity is regarded as a hallmark of native protein structures. In such cases, only two ‘states’ of the protein — the folded and unfolded states — are present throughout the process. Both states are collections of conformations, also referred to as ensembles. The unfolded conformations are flexible and largely unrelated to one another, whereas the folded-state conformations are closely related. The work of Neidigh *et al.*¹ provides clear answers to these two questions.

Evidence for essentially complete folding into a specific tertiary structure near 0 °C comes from following independent global measures of folding — that is, characteristic CD signals and ¹H NMR chemical shifts — as a function of temperature. In both cases the measured parameters plateau at low temperature to values consistent with a highly folded state. In addition, amide protons of some buried residues are protected

from H/D exchange in the folded state, another classic sign of native-like folding. Further support for complete folding in water comes from the observation that there is very little change in the CD spectrum of the final Trp cage design when the sample contained 30% TFE. In contrast, intermediate designs showed substantial enhancements in folding upon addition of TFE. Although addition of TFE does not necessarily induce complete peptide folding³, it seems plausible that failure to see a TFE effect indicates that the extent of folding is already quite high.

What is the evidence for cooperative folding? Sigmoidal dependence of spectroscopic measurements with changes in temperature or denaturant concentration is a classic signature of cooperativity. The authors show such temperature dependencies for both CD and NMR signals. They also use a novel method⁴ to test for cooperativity: they compare the rates of change of characteristic ¹H NMR chemical shifts through the thermal unfolding transition ($\Delta\delta/\Delta T$) with the differences measured for these chemical shifts between the folded state and reference values for the unfolded state ($\delta_{\text{folded}} - \delta_{\text{unfolded}}$). A linear correlation between these two parameters, which is observed for the Trp cage, fits a cooperative unfolding transition between a single folded state and an unfolded ensemble. These comparisons support the hypothesis that the final Trp cage peptide undergoes a concerted unfolding transition.

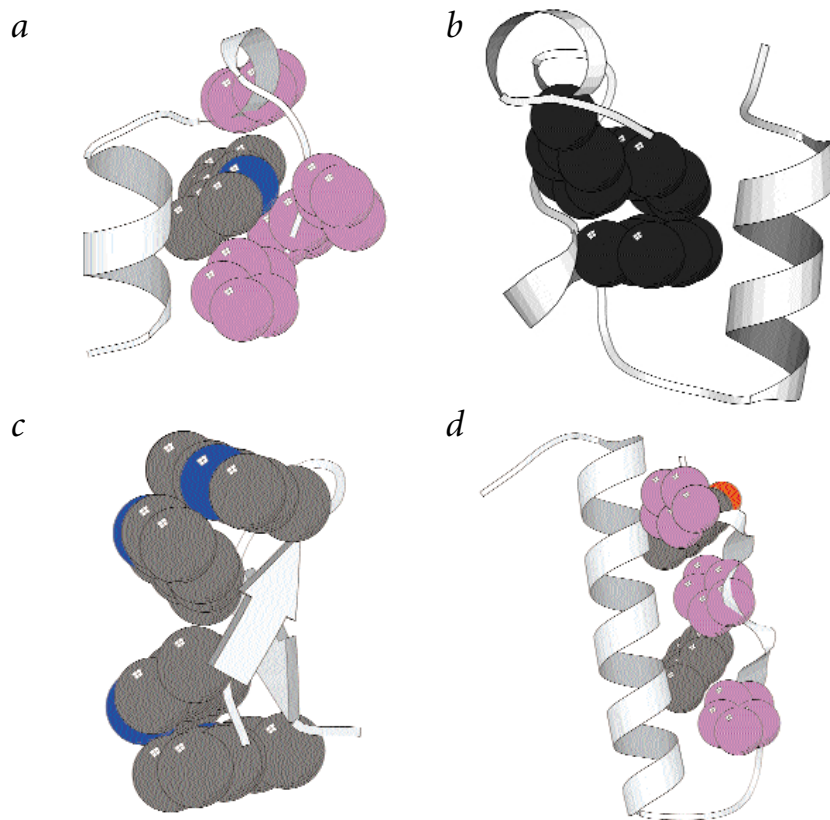


Fig. 1 Examples of mini-protein structures. **a**, The Trp Cage (PDB entry 1L2Y); **b**, the villin headpiece (PDB entry 1V7I); **c**, the Trp zipper (PDB entry 1HRW); and **d**, avian pancreatic polypeptide (PDB entry 1PPT). In each structure the backbone trace is shown as a ribbon. Aromatic and Pro side chains are space filled, with aromatic carbon and nitrogen atoms colored gray and blue, respectively, oxygen is colored red and the atoms of Pro are colored purple. These structures were generated using MolScript¹⁵.

What stabilizes the Trp cage?

Once we accept that the Trp cage is highly and cooperatively folded, at least at low temperature, we are eager to know what holds this little motif together. Neidigh *et al.*¹ oblige with a solution structure for the final Trp cage design (Fig. 1a). The critical tertiary contacts involve the Trp side chain, which is almost completely buried in the folded state. The atoms packed against the Trp side chain come largely from Pro residues in the mini-protein. The authors propose that the Trp–Pro contacts are hydrophobically favorable, and that Pro is an ideal partner for such contacts because the rigidity of this imino-acid residue limits the conformational entropy loss associated with folding.

This proposal merits a close examination. One noteworthy aspect of the truncated Trp cage sequence is its high Pro content (4 of the 20 residues are Pro), which is some four times greater than the average occurrence of the residue in natural proteins. In apparent contradiction to this, however, the sequence contains twice as much Gly (3 out of 20 residues) — the least rigid residue — as the average protein. This intrinsic flexibility is supposed to counter stable folding, but perhaps the Gly residues in the Trp cage are critical to allow contortion of the back-

bone at certain places in order to better define and, in effect, fine tune the tertiary fold. Indeed, measured across all protein structures, Gly residues adopt a much wider spectrum of backbone conformations than any other residue. No doubt related to this, Gly is also one of the most conserved residues at specific sites in protein structures. Consistent with these ideas, all three Gly residues in the Trp cage adopt conformations outside the α - and β -regions of the Ramachandran plot, which are the most populated areas for non-Gly residues.

Thus, an odd alliance of Trp, Pro and Gly in the Trp cage conspires to produce an organized fold cemented by tightly knit buried side chain interactions. Such arrangements are clearly important in specifying and stabilizing competently folded proteins. However, in systems as small as the Trp cage, the key is to make a limited number of such interactions count. The battle here is against the effects of conformational entropy, which is lost during folding because of restrictions placed on both the backbone and side chains. Clearly, the inclusion of Pro must help, but is it enough?

Another source of contributions to protein stability is from interactions in the unfolded state that may reduce the

entropic cost of folding. Neidigh *et al.*¹ present some evidence for this in both the early and final iterations of the Trp-cage design process. Again, Trp and Pro are involved; specifically, a cluster involving Trp 25–Pro 31 interaction may exist in some population of the unfolded ensemble. An embryonic and potentially native-like interaction like this should facilitate folding. Such clustering and possibilities have been observed and discussed for other systems⁵.

Neidigh *et al.*¹ note that Trp–Pro interactions play a key role in the formation of complexes between proline-rich ligands and WW domains. This protein–protein recognition theme is actually more general, involving other widespread recognition modules, such as SH3 domains, and other aromatic side chains (such as the highly conserved Tyr in WW domains⁶). So, are the favored side chain contacts between aromatic residues and/or Pro a general lesson in specifying and maintaining minimal tertiary folds? The answer is: possibly. For instance, a fragment of the most conserved residues at specific sites in protein structures — which at 36 residues is a close runner-up to the Trp cage for the smallest independently folded unit — has a cluster of three Phe residues in its core⁷ (Fig. 1b); many thermophilic proteins contain aromatic clusters, whereas their mesophilic counterparts do not⁸; and a recently designed β -hairpin based on two cross-strand Trp–Trp interactions (Fig. 1c) forms a highly twisted but very stable structure⁹. Turning to natural, intact structures and to aromatic–Pro interactions specifically, the 36-residue avian pancreatic polypeptide (aPP) and its related hormones represent the smallest tertiary folding units that are not stabilized through crosslinking¹⁰. The structure of aPP features a polyproline-II helix followed by an α -helix (Fig. 1d). Contact between these two helices occurs largely between the three Pro residues (positions 2, 5 and 8) and nonpolar residues on one face of the α -helix, including Leu 17, Phe 20, Leu 24, Tyr 27 and Val 30. The contacts between the aromatic and Pro

residues in this interface (Fig. 1d) provide food for thought on the role of such juxtapositions.

Prospects for mini-proteins

The development of the Trp cage is an important achievement for many reasons. We still have a great deal to learn about why specific amino acid sequences fold as they do, from both thermodynamic and kinetic perspectives. Small folding motifs offer excellent systems for detailed analysis because the number of factors that influence folded state stability is reduced, which benefits both experimental and computational studies of folding. In addition, a short sequence facilitates the use of chemical synthesis to generate analogs using not only the 20 standard amino acids, but also unnatural side chains and backbone elements. Such molecules would allow one to explore the origins of conformational stability in ways that are impossible with the proteinogenic amino acids¹¹. For example, replacing a methyl side chain (alanine) with ethyl can probe the contribution of incremental nonpolar

surface burial to conformational stability, and replacing a backbone amide linkage with an ester can probe the contribution of specific hydrogen bonds. In addition to their use for fundamental structural studies, small, robust tertiary motifs provide ideal scaffolds upon which interesting functions may be built. This utility has been recently demonstrated with aPP-derived peptides¹², and it is easy to imagine analogous studies with Trp cage peptides or other mini-proteins.

The Trp cage represents a new benchmark in protein design^{13,14}. It is hard to see how a well-defined tertiary fold could comprise fewer than 20 residues. The exciting challenge offered by the Trp cage is development of alternative folding patterns with comparable length and then endowing them with interesting functions.

Sam Gellman is in the Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706, USA. Dek Woolfson is in the Centre for Biomolecular Design & Drug Development, School of

Biological Sciences, University of Sussex, Falmer, BN1 9QG, UK. Correspondence should be addressed to S.G. email: gellman@chem.wisc.edu or D.W. email: dek@biols.susx.ac.uk

1. Neidigh, J.W., Fesinmeyer, R.M. & Andersen, N.H. *Nature Struct. Biol.* **9**, 425–430 (2002).
2. Branden, C. & Tooze, J. *Introduction to Protein Structure* 2nd ed. (Garland, New York; 1999).
3. Luo, P. & Baldwin, R. L. *Biochemistry* **36**, 8413–8421 (1997).
4. Andersen, N.H. *et al.* *J. Am. Chem. Soc.* **119**, 8547–8561 (1997).
5. Dyson, H.J. & Wright, P.E. *Methods Enzymol.* **339**, 258–270 (2001).
6. Zarrinpar, A. & Lim, W.A. *Nature Struct. Biol.* **7**, 611–613 (2000).
7. McKnight, C.J., Matsudaira, P.T. & Kim, P.S. *Nature Struct. Biol.* **4**, 180–184 (1997).
8. Kannan, N. & Vishveshwara S. *Protein Eng.* **13**, 753–761 (2000).
9. Cochran, A.G., Skelton, N.J. & Starovasnik, M.A. *Proc. Natl. Acad. Sci. USA* **98**, 5578–5583 (2001).
10. Blundell, T.L., Pitts, J.E., Tickle, I.J., Wood, S.P. & Wu, C.-W. *Proc. Natl. Acad. Sci. USA* **78**, 4175–4179 (1981).
11. Koh, J.T., Cornish, V.W. & Schultz, P.G. *Biochemistry* **36**, 11314–11322 (1997).
12. Chin, J.W. & Schepartz, A. *J. Am. Chem. Soc.* **123**, 2929–2930 (2001).
13. DeGrado, W.F., Summa, C.M., Pavone, V., Nastri, F. & Lombardi, A. *Annu. Rev. Biochem.* **68**, 779–819 (1999).
14. Imperiali, B. & Ottesen, J.J. *J. Pept. Res.* **54**, 177–184 (1999).
15. Kraulis, P.J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).

Love it or cleave it: Tough choices in protein quality control

Michael R. Maurizi

Housekeeping is not a humdrum process when it comes to protein quality control. Inside chaperone-protease machines, proteins are subject to dramatic life or death trials and must shape up rapidly to survive. Accumulation of aberrantly folded proteins is the basis of many diseases, so the outcome is vitally important to the cell.

Chaperone-protease machines have assumed a major position in the hierarchy of important biological molecules that are of interest to cell biologists, biochemists, and structural biologists. These complex proteins are associated with many fundamental processes required for the life of a cell. In addition, the molecules themselves have fascinating and intricate structural designs and mechanisms of action. Until recently the ATP-dependent molecular chaperone-protease machines, such as the proteasomes, the Clp proteases, Lon proteases and the membrane-associated FtsH/Yme1 proteases have received the most attention. Now, the crystal structures of two members of the DegP/HtrA protein

family — the mammalian HtrA2/Omi, reported by Li *et al.*¹ on page 436 of this issue of *Nature Structural Biology* and the *Escherichia coli* DegP/HtrA, reported by Krojer *et al.*² in a recent issue of *Nature* — have broadened the focus to include a novel class of ATP-independent chaperone-proteases.

DegP is an essential protein in the *E. coli* periplasm. It protects cells by degrading damaged or denatured proteins generated during heat or chemical stress³. Such damaged proteins would otherwise form aggregates that are highly toxic to cells. DegP homologs in a number of pathogenic bacteria have been shown to be virulence factors, helping bacteria survive within the host cell in

part by degrading periplasmic proteins damaged by the host's oxidative defense system⁴. Several years ago, Erhmann and colleagues⁵ reported an unusual property of *E. coli* DegP. At temperatures below 28 °C, DegP acts solely as a chaperone, protecting unfolded proteins from irreversible aggregation and increasing the yield of refolded protein. Above 28 °C, its protease activity increases dramatically and DegP efficiently degrades the unfolded proteins. Thus, DegP exhibits dual function and appears to combine chaperone activity with protease activity. While precedents exist for chaperone activity contributing to protein degradation, an important question is how this coupling is accomplished without the expenditure