Exploring the folding energy landscape of immunity protein 7

Alice Bartlet, Stuart Knowling, Gareth Morgan, Clare Pashley, Sara Pugh, David Brockwell and Sheena Radford

Introduction

The fundamental problem of how an amino acid sequence determines protein structure is still not fully reconciled. Towards this aim, the folding of small, relatively simple proteins are being studied in atomistic detail to characterise all species encountered during the folding process. Over the last decade the folding mechanism of the colicin immunity protein, Im7, has been extensively interrogated by a range of biophysical approaches. Although we now have a detailed understanding of how this four-helical protein folds to its native stave, via a transiently populated intermediate ensemble, there are still many unexplored areas of this complicated mechanism. Current research has focussed applying new experimental approaches to glean new information about protein folding and exploring the relationship between in vivo and in vitro protein folding.

A three- helix intermediate is a necessary feature on the folding landscape

Although previous work demonstrated that Im7 folding proceeds via transient population of a three helix intermediate, it had not been clear whether this is a necessary feature on the folding landscape or simply a consequence of a sub-optimal amino acid sequence. Helix III is not fully formed, or docked, in the intermediate ensemble of Im7, and is only fully formed and correctly docked after the rate limiting transition state. However, helix III is also is also the shortest helix in Im7 and has a very low helical propensity. A polyalanine region was introduced into helix III, to extend and increase its helical propensity. Structural data indicates that the protein is still able to fold to the native state, and furthermore this alteration appears to have negligible effects on the kinetics of protein folding. These results show the three-helix intermediate is a necessity in the folding process, and even when the denatured state contains a preformed helical structure in this region, the protein folding mechanism is preserved.

Development of specific hydrophobic core packing during Im7 folding

Over-packing substitutions, in which side chain size is increased, were created to probe the specificity and malleability of core-packing during Im7 folding. In parallel, polar groups were introduced into the Im7 core to determine the solvation status of core residues at different stages of folding. Φ-value analysis demonstrated that the major changes in Im7 core solvation occur prior to the population of the folding intermediate, with key regions involved in docking of the short helix III remaining solvent exposed until after the rate-limiting transition state has been traversed. When these regions are over-packed Im7 fails to fold correctly and the intermediate species becomes highly populated at equilibrium indicating that Im7 does not achieve a specifically packed core until the very last step in the folding mechanism.

In vivo stability

Im7 has been used to develop a system that allows protein stability to be measured in vivo, allowing the directed evolution of protein stability independently of function. By fusing Im7 into the antibiotic resistance protein β-lactamase, we can infer the stability of the inserted protein by measuring the antibiotic resistance of cells expressing the fusion construct. The in vivo measurements of protein stability were compared with the in vitro measurements of the thermodynamic and kinetic stability of isolated Im7 variants. A striking correlation between in vivo and in vitro stability was observed. Evolved mutations that stabilized the protein were predominantly involved in the binding site of Im7 to its cognate nuclease, colicin E7. The results demonstrate a stark trade-off between stability and function in the evolution of Im7.

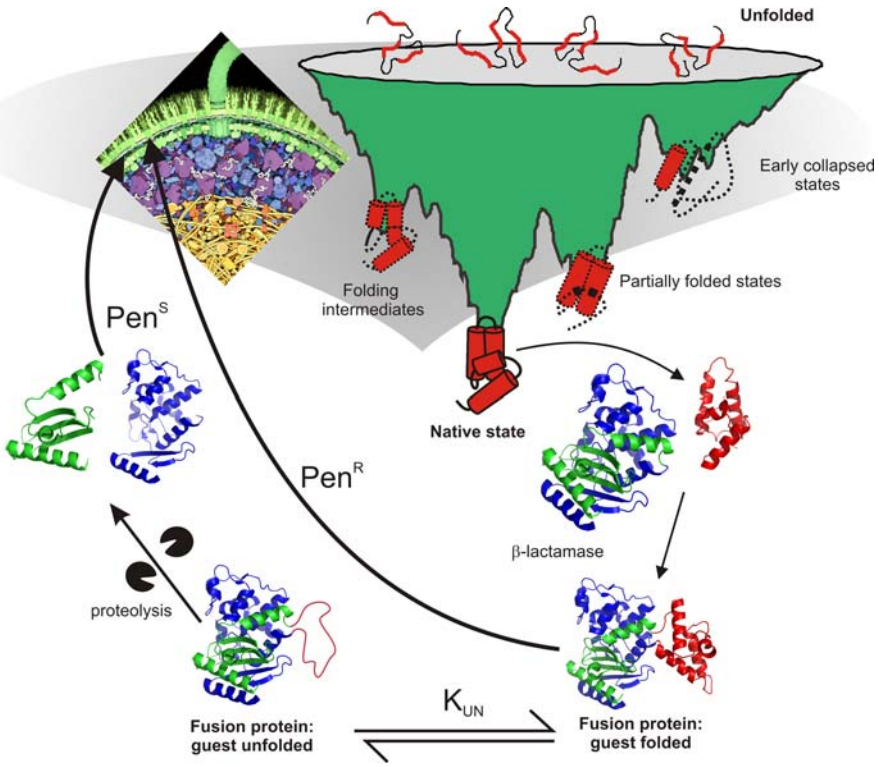

Figure 1: whilst much of our research is focused on dissecting protein folding events *in vitro*, we have recently developed a system to allow us to compare *in vivo* and *in vitro* measurements of protein stability. The β-lactamase enzyme tolerates insertions into a loop between its two domains, and hence a "guest" protein sequence can be inserted into the construct to create a fusion protein. The enzyme activity (against penicillin-type antibiotics) is used to report on the stability the guest protein, within *E. coli* cells. Inserts which do not fold correctly will be cleaved by the host's protein quality control machinery, severing the peptide chain between the two domains of the β-lactamase and inhibiting its activity. Hence, the resistance of the *E.coli* to penicillin is directly related to the stability of the guest protein within the cell, and in turn to its expression levels. This system offers us a convenient method to compare the stability of a protein *in vivo* with *in vitro* measurements of protein stability We can use this system to measure the stabilities of guest proteins, or to evolve proteins for increased stability *in vivo*.

Single molecule FRET studies of Im7 folding

Cysteine mutations were introduced into Im7 to allow fluorescent dye labelling of the protein for single molecule studies. The FRET efficiency of the dyes increases as the distance between the dyes decrease. The dyes are likely to, on average, have longer distances between them in unfolded species, compared with more compact states. Hence, this method allows the distributions of the unfolded, intermediate and native state to be assessed, and the protein folding of Im7 to be characterised at a single molecule level for the first time. A remarkable observation from these experiments was that the unfolded state of Im7 in the absence of denaturant was significantly more compact than the urea-denatured state, suggesting that significant structure could be present in the unfolded ensemble. How this structure influences folding and its precise nature are currently under investigation.

Current work

Although the correlation between *in vivo* and *in vitro* stability was striking, there were notable exceptions. In some cases a protein is more stabile *in vitro*, but apparently less well tolerated by the cell. How well a protein is tolerated by the cell is likely to have a fundamental affect on protein expression, and we are attempting to use the β-lactamase system to dissect the factors

involved in this process. We intend to further investigate the single molecule observations of the unfolded ensemble. The unfolded ensemble contains all conformations sampled at the beginning of the folding process, and is therefore crucial to our understanding of further folding events. However since the unfolded state is not significantly populated at equilibrium, it is difficult to characterise by biophysical methods. Mutations have been introduced into the Im7 sequence to destabilise the intermediate and native states, such that the unfolded state becomes the most stable species at equilibrium. Using this rationale, we have successfully created an unfolded variant of Im7, which populates an unfolded state under conditions favouring folding. Biophysical studies are on-going to characterise this ensemble in atomistic detail.

Publications

Foit L., Morgan, G., Kern, M., Steimer. L., von Hacht, A.-K. , Titchmarsh, T., Warriner, S., Radford, S. and Bardwell, J. (2009) Optimizing protein stability *in vivo*. *Mol Cell,* **36**: 861-871.

Knowling, S., Figueiredo, A., Whittaker, S., Moore, G. and Radford, S. (2009) Amino-acid insertion reveals a necessary three-helical intermediate in the folding pathway of the colicin E7 immunity protein Im7. *J Mol Biol*, **392**:1074-1086.

Bartlett, A. and Radford, S. (2010) Desolvation and development of specific hydrophobic core packing during Im7 folding. *J Mol Biol* **396**:1329-1345.

Pugh, S., Gell, C., Smith, D., Radford, S. and Brockwell, D. (2010) Single-molecule studies of the Im7 folding landscape. *J Mol Biol* **398**:132-145.

Funding

This work is funded by the BBSRC. We thank Keith Ainley for technical support.

Collaborators

Jim Bardwell (University of Michigan, USA), Christopher Gell (Max Planck Institute of Molecular Cell Biology & Genetics, Germany), Alastair Smith (Avacta Group Plc, York)