Protein Folding and Misfolding

Unlike collagen, many biologically important proteins are not extended but globular.

This is usually important for their function, enzyme activity etc.

How does the protein get from its undefined shape following biosynthesis, to a unique folded state?

This is a big question, and one which is not very well understood.

If there is a single minimum but billions of intermediate states which could be accessed, how does the protein rapidly find the minimum?

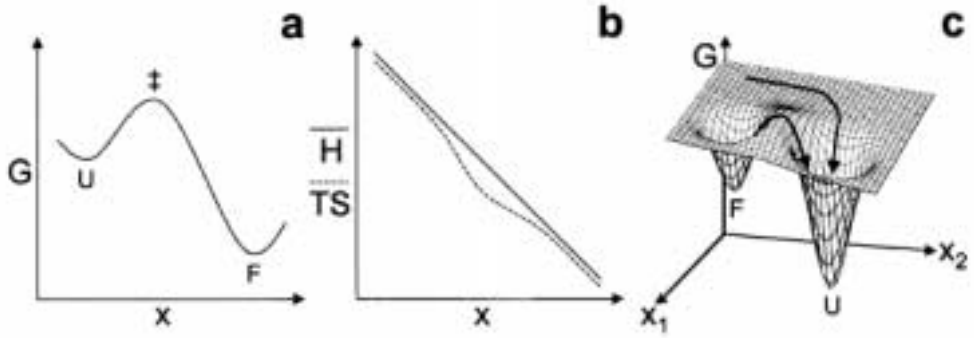


Fig. 2. (a) An example of a free energy diagram often used to describe protein folding and unfolding. The free energy, G, is plotted as a function of a generalized folding coordinate, x. The symbols U, ‡, and F represent the unfolded ensemble, the transition state ensemble, and the folded ensemble, respectively. (b) An example of how an imperfect compensation of entropy (TS) and enthalpy (H) leads to a (in this case, entropic) barrier on the folding coordinate. (c) A two-dimensional free energy diagram allowing for multiple unfolding pathways. x_1 and x_2 represent generalized unfolding reaction coordinates.

Current thinking favours funnels in a rough energy landscape with funnels.

There may be several folding pathways, but it is clear that many states will not be accessed at all.

Trying to understand what the key intermediate states is is one of the main current challenges.

Point mutations are used to see which regions of the chain are important for determining the correct folding states.

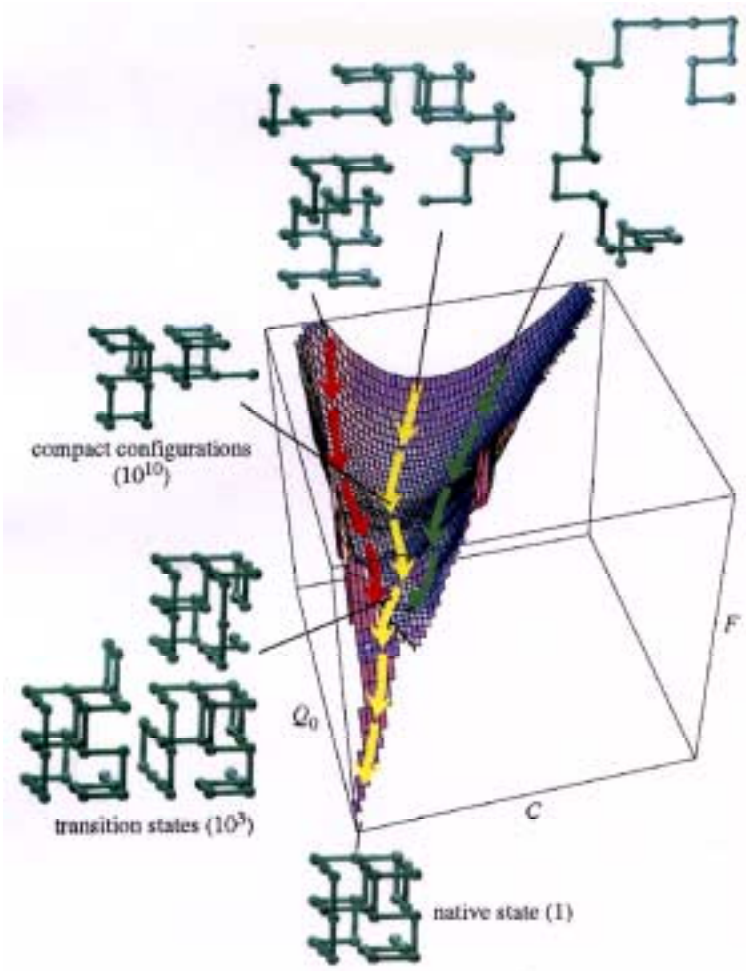


Figure 1. Free-energy ⟨F⟩ surface of a 27-mer model protein as a function of the number of native contacts (Q_0) and the total number of (native and non-native) contacts (C) obtained by sampling the accessible configuration space using Monte Carlo simulations. The yellow trajectory shows the average path traced by structures in 1000 independent trials that each began in a different random conformation. The other two trajectories (green and red) show a range of two standard deviations around the average and are thus expected to include *ca.* 95% of the trajectories. The structures illustrate the various stages of the reaction. From one of the 10^{16} possible random starting conformations, a folding chain collapses rapidly to a disordered globule. It then makes a slow, non-directed search among the 10^{10} semi-compact conformations for one of the approximately 10^3 transition states that lead rapidly to the unique native state. (From Dinner *et al.* 2000.)

Real protein folding may have more than one minimum, and different folding paths to the unique minimum are possible.

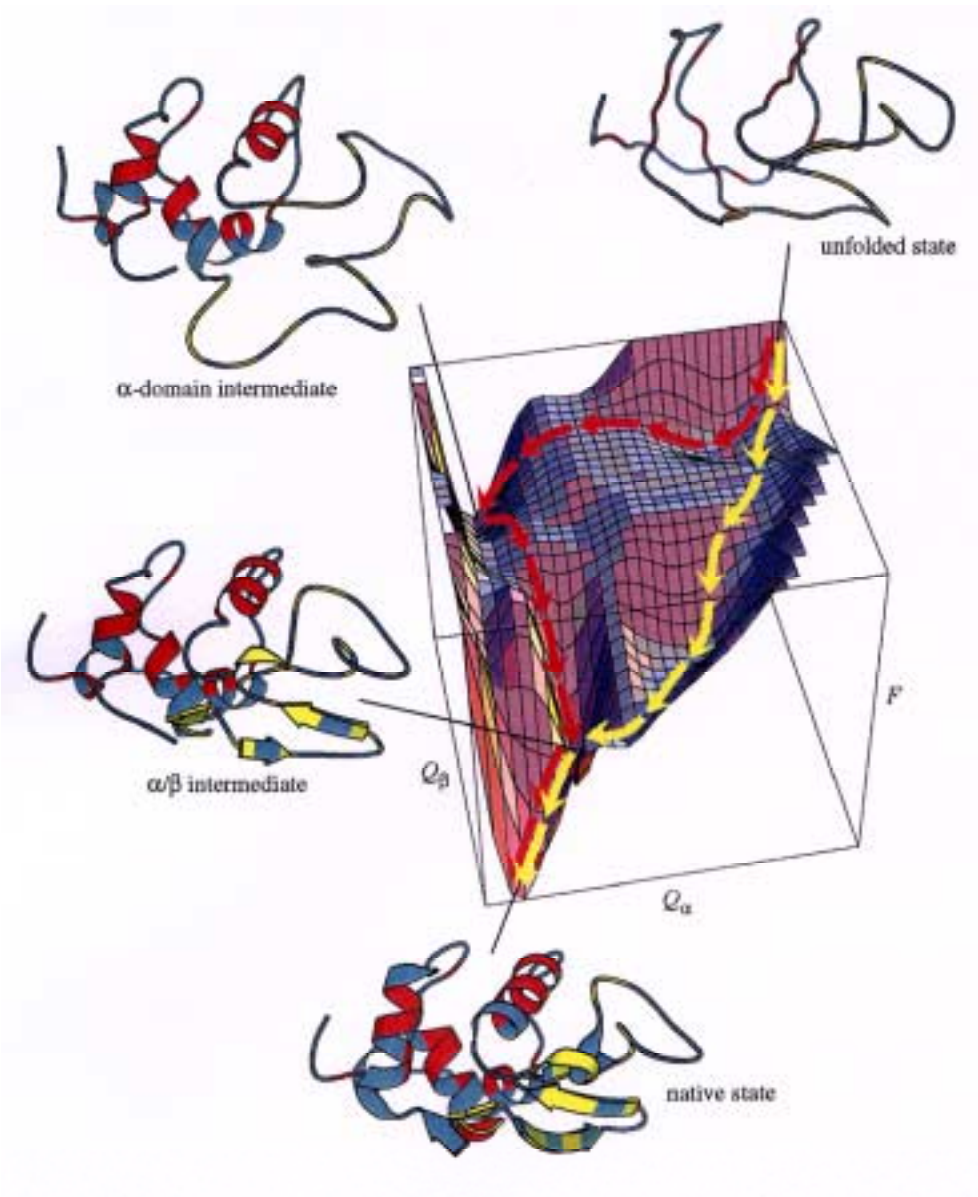


Figure 2. Schematic free-energy (F) surface representing features of the folding of hen lysozyme (a protein of 129 residues whose structure consists of two domains denoted α and β). Q_α and Q_β are the numbers of native contacts in the α and β domains. The yellow trajectory represents a 'fast track' in which the α and β domains form concurrently and populate the intermediate (labelled α/β) only transiently. The red trajectory represents a 'slow track' in which the chain becomes trapped in a long-lived intermediate with persistent structure in only the α domain; further folding from this intermediate involves either a transition over a higher barrier, or partial unfolding to enable the remainder of the folding process to occur along the fast track. Residues whose amide hydrogens are protected from solvent exchange in the native structure (as assessed by NMR) are coloured red (α domain) or yellow (β domain); all others are blue. In each case, regions indicated to be native-like by monitoring the development of hydrogen exchange protection during kinetic refolding experiments are drawn as ribbon representations of the native secondary structure elements (α-helices and β-sheets). (From Dinner *et al.* 2000.)

Misfolded Proteins

Nature has designed proteins to be pretty robust and they fold correctly most of the time.

But when things go wrong it can be catastropic – leading to diseases such as BSE/vCJD and Alzheimers.

These are known as amyloid diseases, as amyloid fibrils of misfolded protein are deposited in the body.

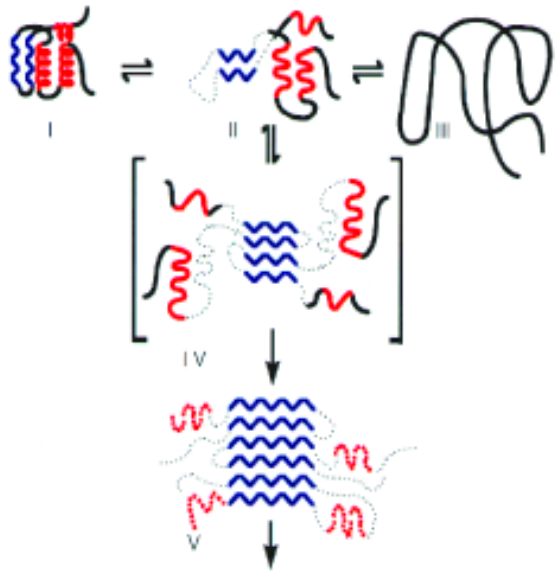
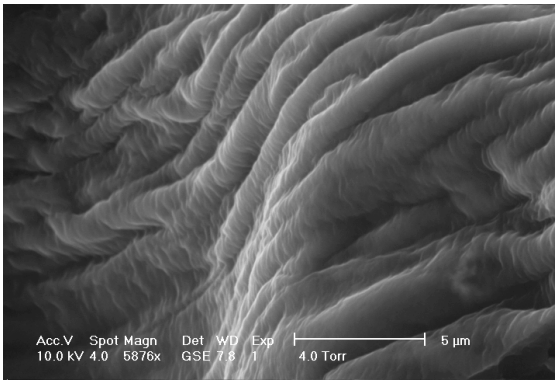


Figure 5. Proposed mechanism for lysozyme amyloid fibril formation. Blue, β-sheet structure: red, helical structure; dotted lines, undefined structure. A partially folded form of the protein (ii) self-associates through the β domain (iv) to initiate fibril formation. This intermediate provides the template for further deposition of protein and for the development of the stable, mainly β-sheet, core structure of the fibril (v). The undefined regions in (v) represent the possibility that not all of the polypeptide sequence is involved in the cross-β structure. The nature of this residual structure in (v) is not known, and the figure is not intended to represent any defined secondary structural type. (From Booth *et al.* 1997.)

Transmission electron micrograph of a dilute suspension of the misfolded proteins shows the individual fibrils quite clearly.



Undiluted have quite different structures. In the SEM show aggregates of plates.

In the polarising microscope show spherulites.