Alfred Nobel Symposium
Energy in Cosmos, Molecules and Life
June 18-22, 2005

Protein Folding

Arthur Horwich
Department of Genetics, Boyer Center, Yale University School of Medicine, HHMI, 295 Congress Avenue, New Haven, CT 06510, USA

In living systems, information transfer proceeds from DNA to RNA to protein. The final step of this process involves the folding of polypeptide chains, composed of characteristic sequences of amino acids, into unique three dimensional structures that mediate structural or catalytic functions. Studies of Christian Anfinsen and his coworkers in the 1950s recognized that the primary amino acid sequence of a polypeptide chain contains all of the information necessary for directing the polypeptide to the native state, which usually lies at a thermodynamic minimum. While many small proteins (150 or fewer amino acids) are able to fold spontaneously to their native state, a majority of proteins are of larger size and more complex domain structure; these proteins can encounter kinetic difficulties in crossing their folding landscape, particularly under cellular conditions of relatively high temperature and solute concentration, lodging in misfolded states that correspond to local energetic minima. Studies beginning in the 1980s have shown that cells have a means of "smoothing the energy landscape," specialized proteins known as molecular chaperones that use the energy of ATP to prevent or reverse misfolding.

The proteins that have come to be called chaperones were first observed in the 1970s as heat shock proteins, cellular components whose abundance was greatly increased during thermal treatment of cells. The nature of their action was initially not understood, but experiments of Hugh Pelham in the mid-1980s suggested that they were protecting other proteins from interacting improperly and aggregating, hence their name, chaperones, reflecting the role of the human counterpart. He proposed that hydrophobic surfaces were becoming exposed in the misfolding proteins, that such surfaces in different polypeptides interacted with each other to cause multimolecular aggregation, and that heat shock proteins were specifically binding to such surfaces and prying apart incipient aggregates through the consumption of ATP. Subsequent studies refined such understanding, indicating that molecular chaperones bind to monomeric polypeptide chains, indeed recognizing hydrophobic surfaces that are normally buried to the interior in the native state but that become exposed during normal folding or in a situation of misfolding. The energy of ATP binding is then employed to drive release of bound polypeptide, allowing further steps of folding or biogenesis to occur.

Our own studies have focused on a remarkable class of molecular chaperone, the double ring structures known as chaperonins. We originally uncovered the function of these components in supporting protein folding to the native state while studying the biogenesis of imported mitochondrial proteins, which occupy unfolded conformations during transit across the mitochondrial membranes. We identified a mutant of yeast in which proteins were imported into the mitochondria and processed to mature size but failed to properly fold. The affected gene

turned out to encode the subunit of a chaperonin in mitochondria called Hsp60. We have ever since been studying the mechanism of action of these machines, focusing on the bacterial chaperonin, GroEL, employing in vitro reconstitution (with Ulrich Hartl), biochemical, and structural studies (with Paul Sigler and Helen Saibil and more recently with Kurt Wüthrich).

We have found that the central cavity of a GroEL ring provides an environment that assists productive folding through actions associated with two major conformational states, and that ATP binding and hydrolysis drives the machine through these states. In a polypeptide binding-proficient state, the cavity of an open GroEL ring exposes hydrophobic surfaces that capture a non-native protein via its own exposed hydrophobic surface, involving multivalent interactions between the substrate protein and multiple surrounding subunits. When ATP cooperatively binds to the seven subunits in the same ring as polypeptide, followed by binding of a lid-like cochaperonin, GroES, the machine is switched to a folding-active state. ATP binding initiates and GroES binding completes a rigid body movement that causes the hydrophobic binding sites to be elevated and twisted away from facing the central cavity to form direct contact with the GroES lid. This movement is associated with release of polypeptide off the cavity walls into the GroES-encapsulated and now hydrophilic chamber where it commences to fold. Productive folding is supported both by solitary confinement of the polypeptide, preventing it from multimolecular aggregation, and by the now hydrophilic cavity walls, which favor burial of exposed hydrophobic surfaces and exposure of hydrophilic ones in the folding substrate protein, properties of the native state.

The lifetime of the folding-active state is approximately 10 seconds, governed by a "timer," ATP hydrolysis, which weakens the affinity for GroES and allows ATP binding to the subunits of the opposite GroEL ring, allosterically triggering ejection of GroES, the polypeptide substrate, and ADP. Polypeptide is ejected whether it has reached the native state or not. For many polypeptides, further rounds of binding and folding are required, with each round comprising an all-or-none process of reaching the native state, as reflected by observations that each time a non-native polypeptide is rebound in an open ring it comes to occupy a similar unfolded state.

We have recently focused on the role of ATP/GroES binding in triggering productive folding. Binding of ATP or ADP•metallofluoride complexes representing either the ground or transition states of ATP hydrolysis, along with GroES, is sufficient to trigger productive folding. Thermodynamic measurements of aluminum fluoride binding to an already-formed GroEL-GroES-ADP complex indicate that the _-phosphate of ATP provides an essential ~-45 kcal/mole (rings) "power stroke" that is needed to overcome a "load" imposed by bound polypeptide, enabling full elevation and twist of the peptide binding domains associated with release of substrate into the cavity. By contrast, in ADP, a complex is formed in which GroES associates with GroEL and encapsulates polypeptide, but the peptide binding domains fail to be elevated, and polypeptide fails to be ejected from the cavity walls. Yet such complexes are likely to lie on the productive pathway taken during the ATP-driven reaction because addition of beryllium fluoride or aluminum fluoride converts them to the folding-active state. The polypeptide-GroEL-GroES-ADP complexes may thus represent a GroES-GroEL collision state. Additional evidence for such a state comes from kinetic studies in ATP, where in the presence of polypeptide, the rate of GroES association is many-fold greater than the rate of apical domain movement. Further study of such states, for example resolving the site of initial GroES binding

on GroEL, will likely reveal why substrate polypeptides do not escape from the central cavity during activation of folding by ATP/GroES.