11

Protein folding *In vivo* and *In vitro*

Introduction

To address the mechanisms of protein folding it is necessary to understand thermodynamic and kinetic parameters governing the formation of the native state. Folding is a reversible reaction defined most simply as

$$\mathrm{U} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{N} \qquad (11.1)$$

where the equilibrium lies to the right ($K_{eq} > 1 = [N]/[U]$) for the native state (N) to be defined as more stable than the unfolded state (U). Quite clearly the respective rates of the forward (k_1) and reverse (k_{-1}) reactions are related to the value K_{eq} and to understand folding requires estimating the magnitude of these rate constants.

Although many proteins fold *in vitro* at comparable rates to those observed *in vivo* the environment and the presence of 'helper' proteins play critical roles in folding within the cell. The realization that some proteins are assisted to find the correct folded structure by molecular chaperones has expanded the area of protein folding.

Protein misfolding is no longer viewed as an experimental artefact. It results from gene mutation and underpins diseases such as Alzheimer's, collagen based diseases such as osteogenesis imperfecta, cystic fibrosis, familial amyloidotic polyneuropathy – a comparatively rare disease arising from mutations in the protein transthyretin – as well as more familiar diseases such as emphysema.

Studies of protein folding have revolutionized our ideas of disease transmission. Until recently it was thought that nucleic acid uniquely carried information 'directing' the transmission of a disease. This occurred either through a mutation in the host genome, integration of viral DNA into a genome or by infection directed by a pathogen's genome. Today there is compelling evidence that some diseases arise solely from a protein and a misfolded one at that!

Factors determining the protein fold

Globular proteins fold into conformations of ordered secondary and tertiary structure where hydrophobic side chains are buried on the inside of the protein and the polar/charged side chains are solvent accessible. Interactions governing the formation of secondary and tertiary structure involve the formation of hydrogen bonds, disulfide bridges, charge–pair interactions and non-polar or hydrophobic effects. The cumulative effect of these forces is that in *any* folded protein the magnitude of favourable interactions outweighs the sum of the unfavourable ones.

These interactions are disrupted by extremes of temperature, immersing proteins in acidic or alkaline solutions or adding solvents such as alcohol in a process known as denaturation that results in a loss

Proteins: Structure and Function by David Whitford

of activity and ordered structure. The disordered state is often described as a 'random coil' although this term should be used sparingly since unfolded states lack truly randomized structure. Denaturation results in a loss of compactness with the unfolded 'state' fluctuating between ensembles of iso-energetic, disordered and extended conformations.

In vitro studies of folding subject proteins to extreme conditions in the expectation that a progressive loss of native structure will be observed. Measuring parameters associated with the kinetics of folding or the stability of the native state allows the process of denaturation to be quantified. Loss of folded structure is usually measured by changes in absorbance or fluorescence although other techniques such as circular dichroism and NMR spectroscopy are increasingly being applied. In most cases folding is a cooperative process arising from simultaneous formation of multiple interactions within a polypeptide chain (Figure 11.1). Individually, each interaction is weak but their cooperative formation drives polypeptide chains towards folded states.

Denaturants like sodium dodecyl sulfate (SDS) are very effective at disrupting protein structure at low concentrations. This is useful for techniques such as SDS–polyacrylamide gel electrophoresis (PAGE) but less desirable in folding studies where a progressive and reversible effect is required. Denaturants such

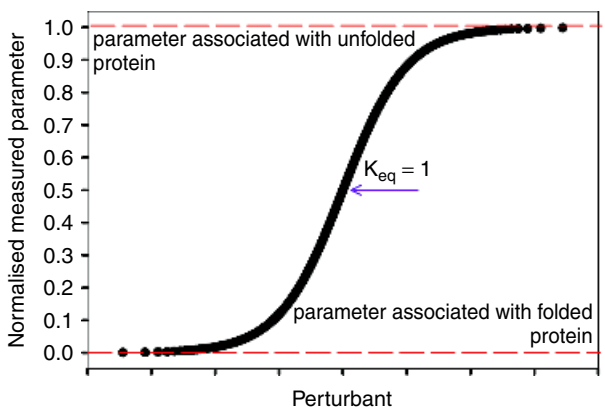

Figure 11.1 The denaturation of a protein showing a cooperative profile. The perturbant can be increasing concentrations of denaturant or increasing temperature

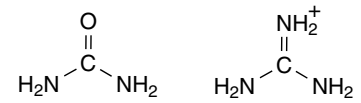

Figure 11.2 Urea is neutral whilst the charged guanidine is normally associated with chloride or thiocyanate anions

as the neutral diamide urea and guanidine (most commonly used as a chloride or thiocyanate salt) bind weakly to protein surfaces with typical affinities of $\sim$100 mM^{-1} and cause progressive loss of structure.

Urea and guanidine hydrochloride (Figure 11.2) are widely used as denaturants during studies of protein folding with their action arising from disruption of a large number of weak interactions. In view of their ability to unfold proteins there are more interactions with the unfolded state than with native forms. Both reagents are classified as chaotropes and in general effects are based on increases in the solubility of polar and non-polar regions of proteins. This effect can be measured by an increase in the partitioning of individual amino acids between water and denaturant. Studies with model compounds initiated by Charles Tanford in the 1960s showed that solubility of amino acids increased with elevated concentrations of urea and appeared to correlate with the size (accessible surface area) of a non-polar side chain. A linear correlation between accessible side chain surface area (Figure 11.3, shown in $Å^2$) and the free energy of transfer was apparent from model studies. However, the action of urea and guanidinium salts as denaturants is complex and preferential solubility of non-polar side chains does not account for all of the effects of these denaturants.

Denaturation caused by extremes of pH is better understood and involves protonation of side chains where these interactions underpin tertiary structure. Below pH 5.0 and above pH 10.0 many proteins denature as a result of a loss of stabilizing interactions. A destabilizing effect will include ionization of a buried side chain that is uncharged in the folded state. The presence of a charged group on the inside of a protein surrounded by hydrophobic residues is destabilizing and shifts equilibria towards the unfolded protein.

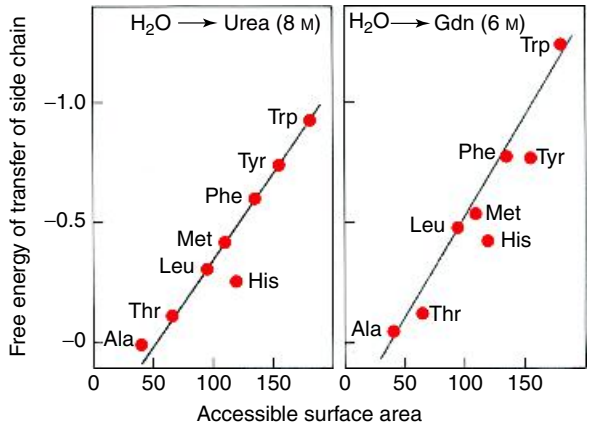

Figure 11.3 Denaturants such as urea and guanidinium chloride increase the solubility of amino acids measured by an increased transfer of amino acid from water to denaturant

Since Lys, Arg, Asp and Glu residues are frequently charged and lie on the surface of proteins the above effect is often exhibited by His and Tyr side chains as they are frequently accommodated within the interior.

Protein folding is a thermodynamically favourable process with a decrease in free energy from unfolded to folded states. This is shown by free energy profiles where the unfolded state (reactant) reaches the folded state (product) via an activated or transition state in a representation that is analogous to that seen for unimolecular reactions. The reaction is described as a two state system when only the folded and unfolded states are identified at any point during the reaction (Figure 11.4).

For a simple folding reaction (see Equation 11.1) the equilibrium constant is defined as the ratio of the concentration of products to reactants (unfolded). In a two state process and with a means of estimating the concentration of either products or reactant during the course of the reaction the free energy associated with unfolding is estimated from the equilibrium constant (K_{eq}) according to

$$\Delta G = RT \ln K_{eq} \tag{11.2}$$

where

$$K_{eq} = [F]/[U] \tag{11.3}$$

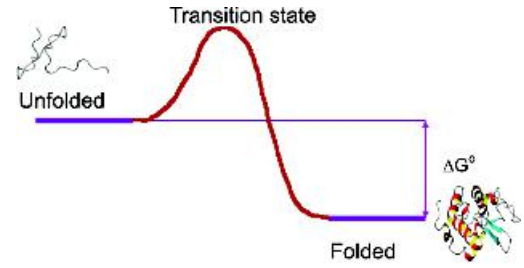

Figure 11.4 A simple reaction profile for two state reaction. The reaction can be described in terms entirely analogous to reaction kinetics where ΔG is the conformational stability of the folded protein ($G_U - G_F$). For folding to occur the change in free energy must be negative

At any point along a cooperative folding curve the concentration of [U] is determined and since the system is two state the fraction of folded protein (f_F) plus the fraction of unfolded protein (f_U) must be equal to unity ($f_F + f_U = 1$). The observed quantity (y_{obs}), usually absorbance or fluorescence, is equivalent to

$$y_{obs} = y_F.f_F + y_U.f_U \tag{11.4}$$

where y_U and y_F are the values characteristic of unfolded and folded states, respectively. Since $f_F = 1 - f_U$ this leads to the following equality

$$f_U = (y_F - y_{obs})/(y_F - y_U) \tag{11.5}$$

and a re-formulation of the equilibrium constant for protein folding as

$$K_{eq} = (y_F - y_{obs})/(y_{obs} - y_U) \tag{11.6}$$

$$\Delta G = -RT \ln (y_F - y_{obs})/(y_{obs} - y_U) \tag{11.7}$$

Transition midpoint temperatures, T_m, define thermal unfolding curves but are not useful estimates of conformational stability since they vary widely from one protein to another (Figure 11.5). However, the T_m remains useful for comparing closely related sequences such as homologous proteins or single site mutants. The unusually named 'cold shock' proteins from mesophilic *Bacillus subtilis* (*Bs*-CspB) and its thermophilic counterpart *B. caldolyticus* (*Bc*-Csp) contain 67 and 66 residues respectively with a high sequence homology

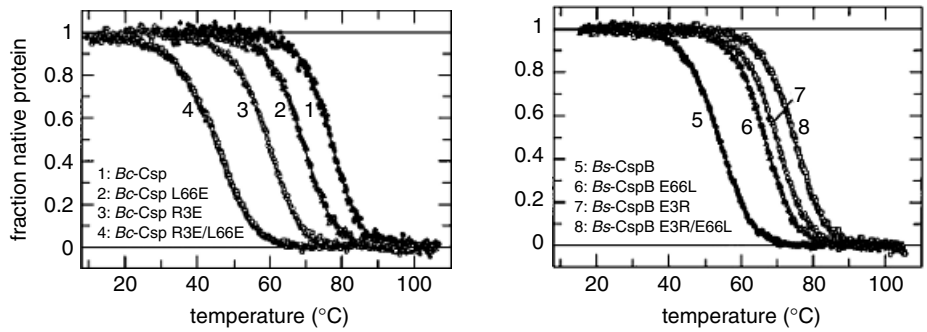

Figure 11.5 Thermal stability curves. Left: Unfolding transitions of wild type *Bc*-Csp and three destabilized variants. Right; Unfolding transitions of wild type *Bs*-CspB and three stabilized variants. The fractions of native protein obtained after a two-state analysis of the data are shown as a function of temperature with the continuous lines showing the results of the analysis (reproduced with permission from Perl, D. *et al. Nat. Struct. Biol.* 2000, **7**, 380–383. Macmillan)

seen by a difference in only 12 residues. The proteins have similar 3D structures and all of the different residues are located on the molecular surface. These residues contribute to a transition midpoint temperature of nearly 77 °C for *Bc*-Csp, approximately 23° higher than that observed in *Bs*-CspB. Using site directed mutagenesis the 12 residues were replaced in *Bc*-Csp by the corresponding residues found in *Bs*-CspB. Only two out of the 12 residues made significant contributions to thermal stability. These residues were Arg3 and Leu66, with each contributing to enhanced electrostatic (Arg) and hydrophobic (Leu) interactions. A recurring theme for elevated thermostability is increased numbers of charged residues and hydrophobic side chains whilst the number of polar uncharged residues usually decreases.

Analysis of thermal induced unfolding curves centres around substitution of the equation

$$\Delta G_{\mathrm{U}} = \Delta H_{\mathrm{U}} - T\Delta S_{\mathrm{U}} \tag{11.8}$$

Since $K_{\mathrm{eq}} = \mathrm{e}^{-\Delta G/RT}$ this leads to

$$K_{\mathrm{eq}} = \mathrm{e}^{-\Delta H_{\mathrm{U}}/RT + \Delta S_{\mathrm{U}}/R} \tag{11.9}$$

where ΔH and ΔS are the enthalpy and entropy associated with unfolding. The van't Hoff analysis of the variation in equilibrium constant with temperature

$$\ln K_{\mathrm{eq}} = -\Delta H/RT + \Delta S/R \tag{11.10}$$

is widely used in chemistry to estimate reaction enthalpy and entropy. Over a narrow temperature range a straight line defines the enthalpy and entropy in plots of $\ln K_{\mathrm{eq}}$ versus $1/T$. These plots in protein unfolding transitions are non-linear, indicating that ΔH is not temperature independent. In proteins ΔH and ΔS are temperature dependent with non-negligible variations over a typical range of 70–100 K.

The heat capacity is defined as the change in enthalpy with temperature

$$C_{\mathrm{p}} = \partial H/\partial T = T\partial S/\partial T \tag{11.11}$$

and during protein unfolding the product and reactants have different heat capacities leading to significant changes in enthalpy. Recasting Equation 11.11 leads to

$$\partial H/\partial T = C_{\mathrm{p}}(U) - C_{\mathrm{p}}(F) = \Delta C_{\mathrm{p}} \tag{11.12}$$

$$\Delta H(T_2) = \Delta H(T_1) + \Delta C_{\mathrm{p}}\ (T_2 - T_1) \tag{11.13}$$

where $C_{\mathrm{p}}(U)$ and $C_{\mathrm{p}}(F)$ are the heat capacities of the unfolded and native states respectively and ΔC_{p} is the change in heat capacity accompanying unfolding. From these equations calculating ΔG at any temperature T requires knowledge of ΔC_{p} and ΔH. The enthalpy is best estimated at a single temperature, and T_{m} provides the most convenient estimation point in the thermal unfolding curve since at this point

$$\Delta G(T_{\mathrm{m}}) = 0 = \Delta H_{\mathrm{m}} - T_{\mathrm{m}}\Delta S_{\mathrm{m}} \tag{11.14}$$

To calculate ΔG at any temperature T is given by

$$\Delta G(T) = \Delta H_m\ (1 - T/T_m) - \Delta C_p\ [(T_m - T) + T \ln\ (T/T_m)] \qquad (11.15)$$

where T_m is estimated from the unfolding curve; ΔH_m is estimated from the gradient of the straight line of $\ln\ K_{eq}$ versus $1/T$ around T_m – at this point $\Delta H_m = T_m\ \Delta S$, since $\Delta G = 0$, and this leaves ΔC_p as the only parameter remaining to be estimated.

Although generally true that proteins with a high T_m are more stable than those with lower T_m values this parameter is not a definitive measure of stability. Changes in protein stability reflected by small changes in T_m are approximated through the relationship

$$\Delta\Delta G_U \approx T_m\ \Delta S_m \approx \Delta H_m\ \Delta T_m/T_m \qquad (11.16)$$

If either the entropy or enthalpy associated with unfolding is unknown an estimate of $\Delta\Delta G_U$ may be obtained with reasonable precision from Equation 11.16. Intuitively it is clear that as the temperature increases a protein unfolds. The basic equations of state relating to protein folding reveal that as T increases the term $T\Delta S$ dominates to such an extent that it becomes greater than ΔH. At high temperatures the entropy of the unfolded state dominates and favours unfolding.

The heat capacity change is most accurately measured by differential scanning calorimetry (DSC). DSC enables direct measurement of T_m, ΔH and ΔC_p from increases in heat transfer that occur with unfolding as a function of temperature. The instrument is based around an adiabatic chamber where one cell contains the sample (protein plus solvent) at a concentration of $\sim$1 mg ml^{-1} whilst another cell acts as a reference and is normally filled with an identical volume of solvent (Figure 11.6). Both cells are heated with the temperature difference between the two cells constantly measured as a complex feedback loop increases or decreases the sample cell's power input via a heater in an effort to keep the temperature difference close to zero. Since the masses and volumes of the two cells are matched the power added or subtracted by the system is a direct measure of the difference between the heat capacity of sample and reference solutions – in other words the heat capacity of the protein. Despite first impressions it is surprisingly difficult to achieve perfect matching of sample and reference cells with the result that experimental scans are observed with a baseline offset. This 'constant' is usually subtracted from the data to give a corrected profile and accurate estimates of C_p (Figure 11.7).

Although ΔC_p can be estimated from a single thermogram it is generally obtained from a series of profiles obtained at different pH values. The T_m decreases as the pH is lowered from small enthalpy changes and a plot of ΔH_{cal} against T_m has a slope of ΔC_p. From Equation 11.11 integration of the heat

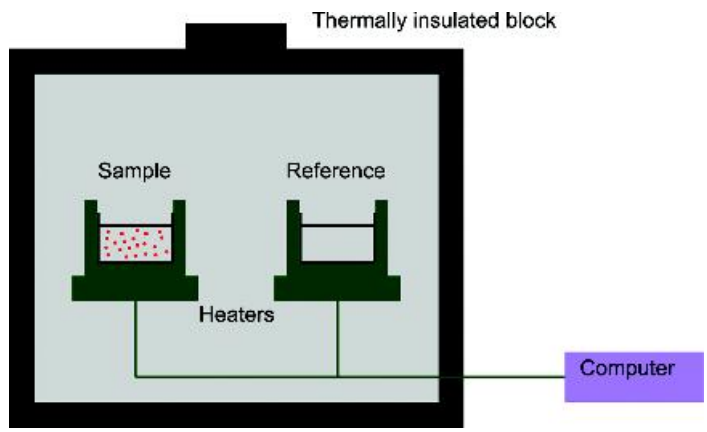

Figure 11.6 In a differential scanning calorimeter the sample cell contains protein plus solvent whilst the reference cell lacks protein. Volumes of $\sim$0.5 ml are used in each cell

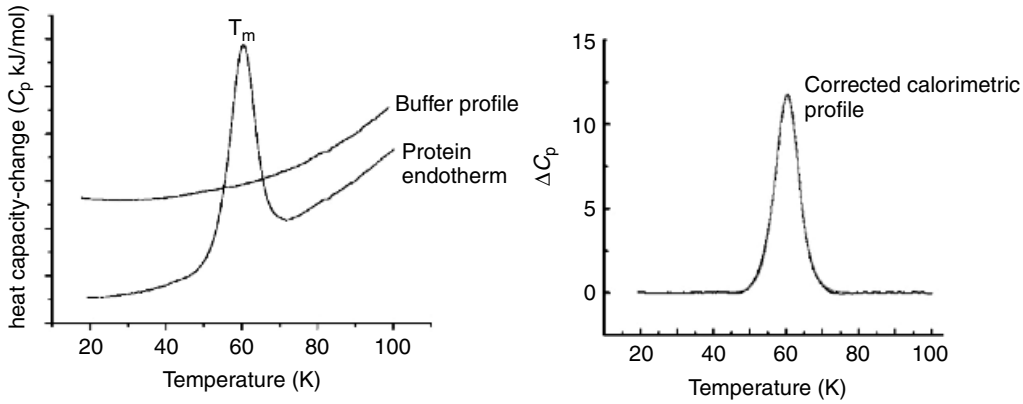

Figure 11.7 DSC profile for protein unfolding showing pre- and post-denaturational regions and baseline correction to achieve accurate estimation of $\Delta C_{p(U-F)}$

capacity curve yields the enthalpy (ΔH) and measures the total heat absorbed by the protein during the process of unfolding. Calorimetric studies of protein unfolding show that the heat capacities associated with unfolded states are larger than those of native states, although both are positive values.

In small soluble domains the value of ΔC_p ranges between 7000 and 10 000 J mol^{-1} K^{-1}. Systematic analysis of these values in many proteins suggests that a rule of thumb can be constructed for 'estimating' ΔC_p – it involves multiplying the number of residues in a protein by 50 J mol^{-1} K^{-1}. It is common to compare the enthalpy obtained from the van't Hoff plot (ΔH_{vh}) with that determined calorimetrically (ΔH_{cal}). A ratio ($\Delta H_{vh}/\Delta H_{cal}$) of 1.0 confirms that the process is two state whilst a value of less than 1 may be used as evidence of an intermediate and departure from a simple two-state analysis. A value of $\Delta H_{vh}/\Delta H_{cal} > 1$ often indicates polymerization at some point in the heating process.

Over a wide temperature range the variation in equilibrium constant is defined by a curve as opposed to a straight line. This curve has two points where the equilibrium constant is 1 and the concentrations of folded and unfolded forms are equal (Figure 11.8). The first point is the thermal denaturation temperature (T_m) whilst the second point reflects a low temperature denaturation point. The ΔG for unfolding is dominated by the entropy component at high and low temperatures

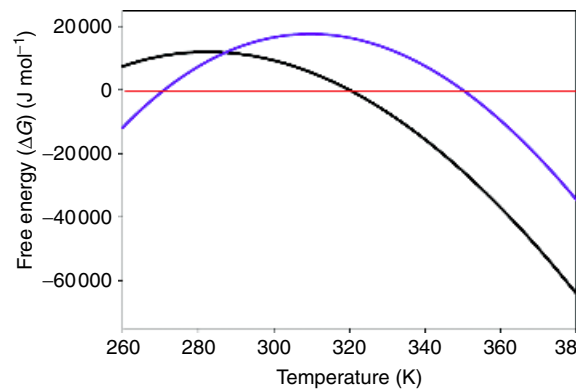

Figure 11.8 The temperature dependence of protein unfolding. The two curves show the variation in ΔG with temperature in two systems with different T_m, ΔH and ΔC_p

although the cold denaturation point lies below the freezing point of water and is not routinely accessible.

An alternative is to promote unfolding using denaturants such as urea or guanidine hydrochloride (GdnHCl) with the advantage that unfolding curves are generally simpler to interpret. In direct analogy to T_m a midpoint in the unfolding transition is given the symbol C_m and can be determined with reasonable accuracy (±0.05 M). For similar proteins such as those involving single site-specific mutations this approach offers a valid means of comparison but

different proteins exhibit a wide range of C_m and unfold under very different conditions. A logical step is to extrapolate estimations of protein stability to zero denaturant concentration. This allows conformational stability of proteins to be compared in the absence of denaturant ($\Delta G_{u(H_2O)}$).

In view of the correlation between denaturant concentration and accessible surface area changes in free energy are related to denaturation via the 'linear extrapolation method' (Figures 11.9–11.11). This model assumes without any strong theoretical basis that ΔG changes linearly with denaturant concentration, according to the equation

$$\Delta G_U = \Delta G_{U(H_2O)} - m_{eq}[D] \qquad (11.17)$$

where ΔG_U is the energy of unfolding in the presence of denaturant, D is the molar concentration of denaturant and m_{eq} is a constant reflecting the association between denaturant and protein. Values of m_{eq} (see Table 11.1) reflect the shape of denaturation curves, with high values leading to sharp transitions between folded and unfolded states. Values for m_{eq} are positive, with typical values ranging between 2000 and 15 000 J mol^{-1} M^{-1}.

A plot of ΔG versus denaturant concentration gives the conformational stability in the absence

Table 11.1 Experimentally determined values of m_{eq} and ΔC_p for selected proteins

Protein	m_{eq} (J mol^{-1} M^{-1})	ΔC_p (J mol^{-1} K^{-1})
Ubiquitin	7 300	5 690
Cytochrome b_5	8 950	6 000
Barnase	18 400	6 900
Thioredoxin	13 850	6 950
Lysozyme	14 480	6 610
Ribonuclease T1	10 700	5 300
Metmyoglobin	15 500	7 820
α-chymotrypsin	17 150	12 650
Phosphoglycerate kinase	40 580	31 380
Staphylococcal nuclease	28 580	9 700

All m_{eq} values were determined using GdnHCl as denaturant. Values of m calculated with urea are most frequently observed to be smaller than those obtained using GdnHCl.

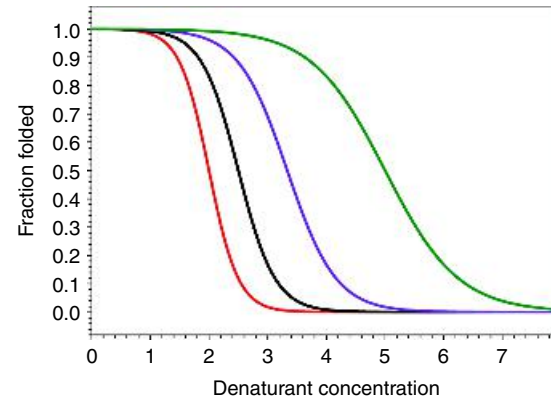

Figure 11.9 Four ideal denaturation curves representing a protein with a conformational stability $\Delta G_{(H_2O)} = 20$ kJ mol^{-1}. The curve reflects different values of m_{eq} ranging from 10 000 (red) 8000 (blue), 6000 (black) and 3000 (green) (J mol^{-1} M^{-1})

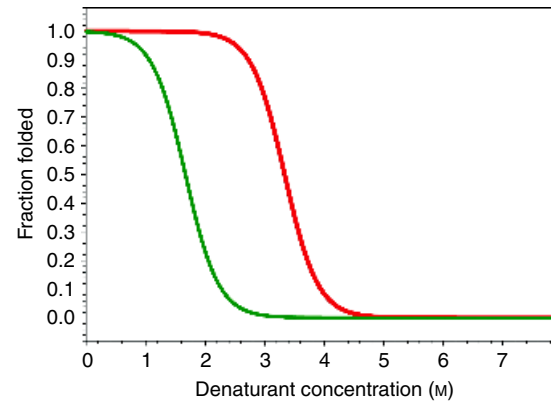

Figure 11.10 The denaturation of two proteins with different stabilities but comparable m_{eq} values. The first protein (red) has a stability of 30 kJ mol^{-1} whilst the second (green) has a stability of 12 kJ mol^{-1}. Denaturation leads to C_m values of ~1.8 M and 3.5 M. The two systems represent a relatively unstable protein and one of moderate stability

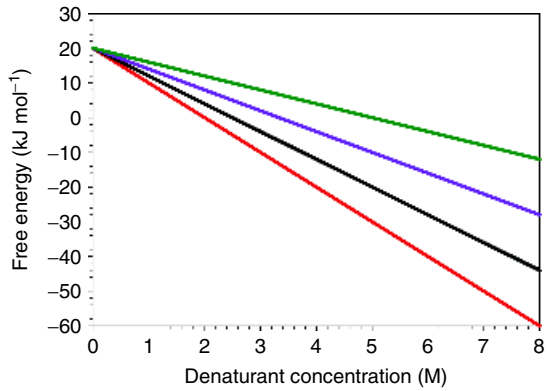

Figure 11.11 The linear extrapolation method applied to the data of Figure 11.9

of denaturant and experimental studies suggest that using either urea and guanidine hydrochloride and the same 'test' protein yields a common intercept at zero denaturant and a constant value of $\Delta G_{U(H_2O)}$ (Figure 11.11).

A conclusion from many denaturation studies is that overall stability of a protein (ΔG, $G_U - G_F$) is small with the folded state being marginally more stable than the unfolded state. This leads to values for protein conformational stability ranging between 10 and 75 kJ mol^{-1}, with most proteins exhibiting values towards the lower end of this range (Table 11.2).

Table 11.2 The conformational stability of proteins

Protein	Chain length	Conformationally stability (kJ mol^{-1})
λ repressor	80	12.5
Cytochrome c	104	74.0
Cold shock protein CspB	67	12.6
SH3 domain of α spectrin	62	12.1
CI2	64	29.3
U1A spliceosomal protein	102	38.9
Ubiquitin	76	15.9
CD2	98	33.5

Ribonuclease defines the 'outline' of protein folding in vitro

Ribonuclease has contributed much to our understanding of protein folding *in vitro* through the landmark studies of Christian Anfinsen who posed the question: 'what is the origin of the information necessary for folding?' Ribonuclease, with 124 amino acid residues and four disulfide bridges located between cysteines 26–84, 40–95, 58–110, and 65–72, catalyses the hydrolysis of RNA. Reduction of the disulfide bridges to thiols by mercaptoethanol in the presence of urea results in protein unfolding and a concomitant loss of activity.

Anfinsen noticed that when ribonuclease was oxidized (by standing in air) and the urea removed by dialysis that enzyme activity slowly recovered as a result of protein folding, the reformation of tertiary structure and most importantly the active site. Repeating these reactions in the presence of denaturant (oxidation of the thiols in 8 M urea) led to the regeneration of less than 1 percent of the total enzyme activity. Urea preventing correct disulfide pairings resulting in a 'scrambled' ribonuclease whilst in its absence correct disulfide bridge formation allowed the folded and thermodynamically most stable state to be reached. These classic studies showed that *all* of the information necessary for protein folding resides *within* the primary sequence. The intervening years have demonstrated the generality of Anfinsen's results and although a few caveats must be included the maxim 'sequence defines conformation' remains relevant today.

One of the first caveats in this picture of protein folding is the protease subtilisin E from *B. amyloliquefaciens*. The enzyme is synthesized as a preproenzyme of ~380 residues to permit secretion and to avoid premature proteolysis. Folding studies involving the removal of the pro sequence showed that its loss was coupled to a lack of subsequent protein refolding. In contrast unfolding the full length preproprotein resulted in a folded enzyme. The results were interpreted along the lines that the pro-sequence participated in folding reactions guiding subtilisin towards the native state. In its absence the native state could not be reached and emphasized that it is not necessarily the 'final' protein sequence that encodes folding information. The second caveat arose with the observation that 'helper' proteins assist folding *in vivo*

by preventing molecular aggregation. These helper proteins, molecular chaperones, derive their name from their ability to prevent unwanted interactions between newly synthesized chains and other proteins.

Factors governing protein stability

The native state is the most stable form of a protein but what factors contribute to this stability? From the previous section it might be thought that disulfide bridges or covalent bonds in general are major determinants of conformational stability. However, in the unfolded protein these bonds remain intact and do not contribute to conformational stability. To put this in a slightly different way the covalent bonds contribute equally to the stability of the folded and unfolded proteins and it is necessary when addressing this concept to focus on non-covalent interactions.

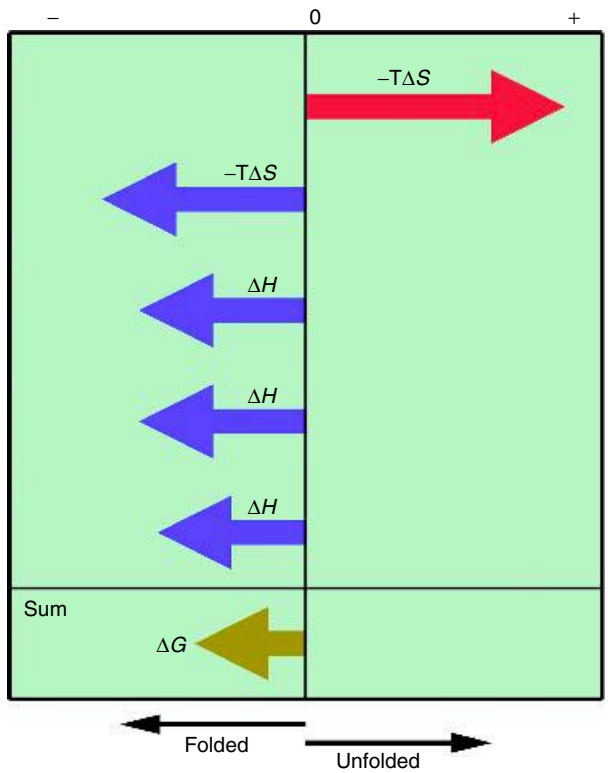

Figure 11.12 Contributions to the free energy of folding of soluble proteins

Non-covalent interactions include hydrogen bonds, hydrophobic forces and interactions between charged groups that cumulatively contribute to increased stability of the native state (Figure 11.12). The entropic contribution to folding is generally unfavourable since it involves transitions from large numbers of unstructured conformations to a single ordered structure. The decrease in entropy works against protein folding and this conformational entropy is largely responsible for favouring the unfolded state. To fold proteins must overcome conformational entropy from entropic and enthalpic contributions that derive from interactions between charged groups, hydrophobic effects, hydrogen bonding and van der Waals interactions.

A major contribution to thermodynamic stability arises from the hydrophobic interaction. Protein folding results in the burial of hydrophobic side chains away from water and their interaction with similar side chains on the inside of the molecule. In thermodynamic terms ordered water molecules specifically arranged in unfolded state are released from this state and are free to move with the result that the entropy increases and makes favourable contributions to the overall free energy change. Favourable enthalpic terms include the formation of stabilizing interactions in the native state via charge interactions and hydrogen bonding. The magnitude of ΔS and ΔH varies from one protein to another but the result appears to be largely compensatory with ΔG values confined to a narrow range generally between 15 and 40 kJ mol^{-1}.

Folding problem and Levinthal's paradox

The Ramachandran plot reveals that many dihedral angle combinations are compatible with stable secondary structure with the result that in an average protein of 100 residues the number of possible conformations is enormous. Some idea of the magnitude of this value is obtained by assuming that only three conformations exist for each residue in a protein of 100 residues and that each conformation can be sampled in 10^{-13} s or 0.1 ps. To sample all conformations would

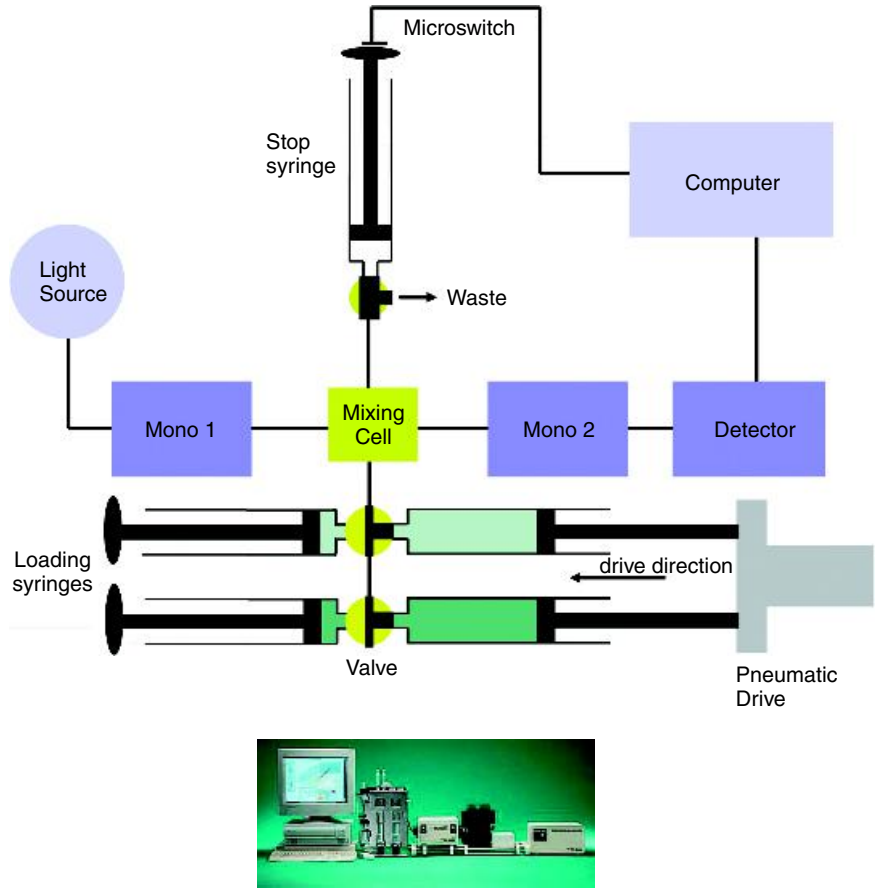

Figure 11.13 Schematic arrangement of a stopped flow system (lower figure reproduced with permission Applied Photophysics Limited)

take $3^{100} \times 10^{-13}$ s – an impossibly long period – that emphasizes that proteins do not sample *all* possible conformations.[1] This problem commonly known as Levinthal's paradox has focussed attention on the kinetics of protein folding.

Kinetics of protein folding

A consideration of Levinthal's paradox quickly eliminates random models of protein folding but a relevant question to ask is what is the time scale of protein folding? Experimentally, these studies involve placing unfolded protein into a buffer system that promotes folding and by mixing rapidly the reaction can be followed on time scales ranging from 1 ms to ~10 min. A common method is the stopped flow kinetic technique where reactants combine in a mixing cell in a process followed using changes in fluorescence or absorbance. In modern instrumentation it is possible to measure changes in CD signals with time but these experiments are technically difficult and less sensitive. In stopped flow experiments (Figure 11.13) pressure-driven syringes force reactants into a mixing cell where the reactants mix displacing resident solution within

[1]There are far more than three conformations per residue, some proteins contain more than 100 residues and 10^{-13} is an unrealistically short period in which to sample each conformation.

a dead time of ~1 ms. This initiates measurement of folding reactions and kinetic profiles usually follow single exponential decays that are fitted to obtain the observed rate constant (k_{obs}) and the maximum amplitude (A_i)

$$k_{obs} = A_i e^{-kt} \tag{11.15}$$

For a two-state unfolding system kinetic analysis is straightforward since the equilibrium constant K_{eq} is simply the ratio of forward (k_f) and reverse (k_u) rates defining the folding and unfolding reactions. Combining Equations 11.2 ($\Delta G_{U(H_2O)} = -RT \ln K_{eq}$) and 11.14 ($\Delta G_U = \Delta G_{U(H_2O)} - m_{eq}[D]$) with the relationship

$$K_{eq} = k_{fw}/k_{uw} \tag{11.16}$$

allows the dependence of the folding and unfolding rates on denaturant concentration to be expressed as

$$\ln k_u = \ln k_{uw} + m_u[D] \tag{11.17}$$

$$\ln k_f = \ln k_{fw} + m_f[D] \tag{11.18}$$

where D is the molar concentration of denaturant.[2] Plotting the logarithm of the rate of folding against denaturation concentration yields a linear relationship that when combined with the dependence of unfolding rates on denaturant concentration leads to a characteristic profile known as a chevron plot (Figure 11.14). C_m is defined by the point where $k_u = k_f$ and the unfolding and refolding rates in the absence of denaturant are extrapolated by extending each limb of the plot to zero denaturant. At any point the observed rate is calculated from

$$k_{obs} = k_u + k_f \tag{11.19}$$

$$k_{obs} = k_{uw} \exp(m_u[D]) + k_{fw} \exp(m_f[D]) \tag{11.20}$$

where

$$m_{eq} = (m_u - m_f)/RT \tag{11.21}$$

The parameters m_u and m_f are kinetic m values to distinguish from the equilibrium m value (m_{eq}) and are used to estimate the position occupied by the transition state along the reaction coordinate from U to F often

[2]The parameters $m_{ku} = RTm_u$ and $m_{kf} = RTm_f$ and k_{uw}/k_{fw} is the rate of unfolding/folding in the absence of denaturant.

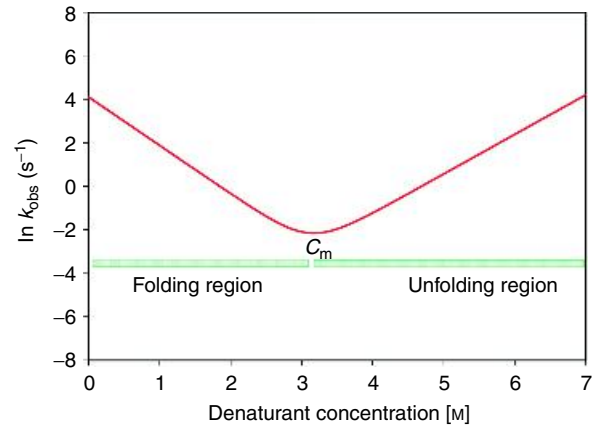

Figure 11.14 A chevron plot describes kinetics of protein folding for a simple soluble domain as a function of denaturant concentration

given the symbol α and equal to $m_f/(m_u + m_f)$. A value of 0.5 suggests a transition state midway between folded and unfolded states.

Barnase and CI2 are two proteins that have been actively studied by Alan Fersht to provide quantitative insight into the mechanism of folding. CI2 is a 64-residue inhibitor of chymotrypsin or subtilisin with a structure based around a single α helix from residues 12 to 24, and five β strands arranged in parallel and antiparallel sheets (Figure 11.15). The α helix lies between strands 2 and 3 and CI2 forms a single folding unit that exhibits two state kinetics and a $t_{1/2}$ of 13 ms at room temperature for the major phase of protein folding. Smaller contributions attributed to *cis–trans* peptidyl proline isomerization are observed but remain a minor element in the overall folding pathway that lacks intermediates in kinetic or equilibrium-based measurements. CI2 represents the simplest kinetic pathway for folding and one widely observed to occur in other small soluble domains.

Barnase is a small, monomeric, soluble, extracellular ribonuclease secreted by *B. amyloliquefaciens* containing 110 residues and lacking disulfide bonds. The protein exhibits reversible unfolding under a wide variety of denaturing conditions. The structure (Figure 11.16) shows a five-stranded β sheet with three helices (α_1–α_3) located between loops and outside of

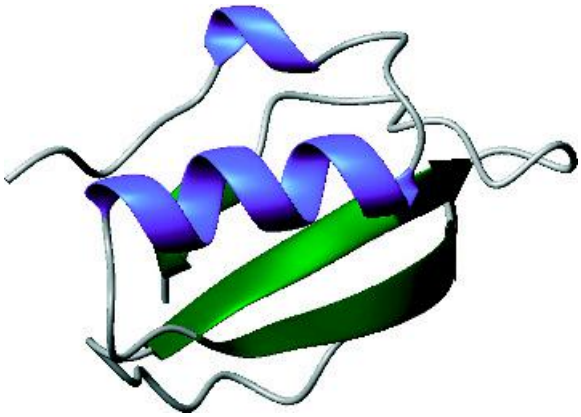

Figure 11.15 The structure of CI2 (PDB: 2CI2). Some strands are poorly defined

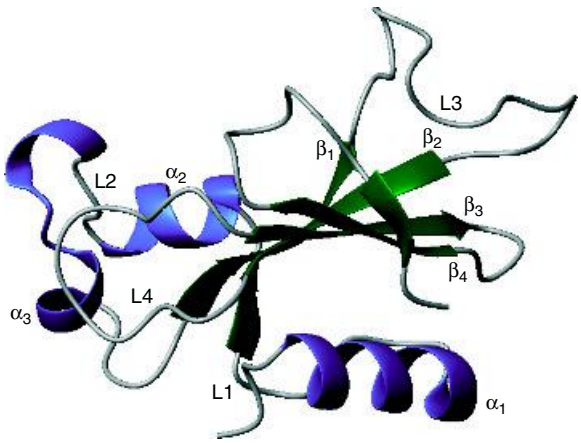

Figure 11.16 The structure of barnase showing distribution of secondary structural elements. L1–L4 are loop regions whilst β strands extend from residues 50–55, 70–76, 85–91, 94–99 and 106–108

the sheet region. The major helix (α_1) is located at the N-terminal and extends from residues 6–18 where its interaction with the β sheet forms the major hydrophobic core. A smaller core is formed from the interactions between loop2, α_2 and α_3 helices and the first β strand.

As a system for studying folding and stability barnase represents an excellent paradigm. With the α and β structures in different regions of the molecule, as opposed to a more mixed arrangement, it is possible to study either fragments of the protein or individual mutations confined to one secondary structural element. To understand protein folding, and in particular the pathway involved, it is necessary to determine the structure and energetics of the initial unfolded state, all intermediate states, the final folded states and importantly the transition states occurring along this pathway.

One approach that has shed considerable light on the mechanism of protein folding as well as the non-covalent interactions involved in these pathways introduced small mutational changes to barnase that removed specific interactions. By measuring changes in stability (i.e. ΔG) alongside kinetic measurements estimating the activation energy ($\Delta G^{\ddagger}$) and other properties associated with the transition state it is possible to begin to describe the energetic consequences of mutation.

The difference in conformational stability between the folded states of the wild type and mutant forms is difficult to measure directly but can be exploited via relationships that invoke Hess' law. The basis of the method is a mutational cycle (Figure 11.17) in which the introduction of specific mutations into barnase is used to estimate kinetic and thermodynamic parameters. The mutation acts as a sensitive reporter of local events during the pathway of protein folding. The mutational cycle measures the free energy of

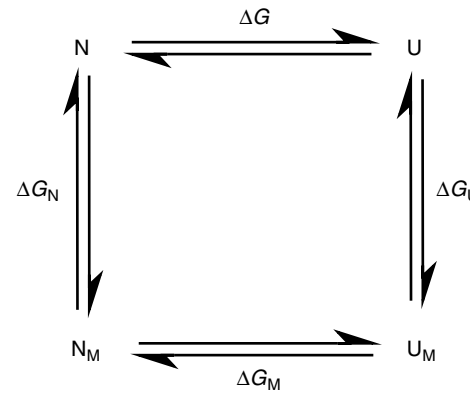

Figure 11.17 Application of Hess's law to the unfolding free energy of wild type and mutant proteins

unfolding for wild type (N) and mutant (M) proteins in the horizontal reactions whilst the vertical reactions represent experimentally indeterminable or 'virtual' free energy changes. They are the free energy changes that arise in the folded or unfolded protein as a result of mutation.

Hess's Law states that the overall enthalpy change for a chemical reaction is independent of the route by which the reaction takes place as long as the initial and final states are identical (Figure 11.17). It is simply a statement of the law of energy conservation. Using the mutational cycle approach it was shown that the energetic effects of mutation in barnase are evaluated from the relationship

$$\Delta G + \Delta G_{\mathrm{U}} = \Delta G_{\mathrm{N}} + \Delta G_{\mathrm{M}} \tag{11.22}$$

Re-arranging leads to

$$\Delta G - \Delta G_{\mathrm{M}} = \Delta G_{\mathrm{N}} - \Delta G_{\mathrm{U}} \tag{11.23}$$

where $\Delta\Delta G_{\mathrm{N-U}}$ is the change in free energy due to mutation (M) and the experimental quantity one would wish to access directly.

Mutations do not have a single effect on one 'folding' parameter but can alter solvent exposure, hydrophobic contacts, hydrogen bonding – the list is almost endless. As a result of the large number of parameters altered by a mutation it is very difficult to ascribe the changes to a single energetic parameter. One way to eliminate this problem is to borrow a trick from physical organic chemistry and use the Brønsted equation. The Brønsted equation relates changes in rate constant k and equilibrium constant K to a parameter β as a consequence of altering a non-reacting functional group

$$\log k = \text{constant} + \beta \log K \tag{11.24}$$

In physical organic chemistry this involves making series of derivatives whilst in the area of protein folding this involves the creation of mutants. In each case the effect on rate and equilibrium constants for folding are measured. For protein folding this equation was modified to

$$\Delta G^{\ddagger} = \text{constant} + \phi_{\mathrm{U}} \Delta G_{\mathrm{U}} \tag{11.25}$$

in view of the relationship of free energy terms $\Delta G^{\ddagger}$ and ΔG_{U} to log k and log K respectively whilst the parameter ϕ replaced β from the Brønsted relation. This scheme allows two measurable reactions to be compared since the constant can be cancelled by subtraction to give

$$\Delta G^{\ddagger}_{\mathrm{U}} - \Delta G^{\ddagger}_{\mathrm{N}} = \text{constant} + \phi_{\mathrm{U}}\ \Delta G_{\mathrm{U}} - \text{constant} + \phi_{\mathrm{U}}\ \Delta G_{\mathrm{N}} \tag{11.26}$$

$$\Delta\Delta G^{\ddagger} = \phi_{\mathrm{U}}\ \Delta\Delta G_{\mathrm{U}} \tag{11.27}$$

$$\phi_{\mathrm{U}} = \Delta\Delta G^{\ddagger}/\Delta\Delta G_{\mathrm{U}} \tag{11.28}$$

where $\Delta\Delta G_{\mathrm{U}}$ is the difference in conformational stability between folded and unfolded protein. A value of ϕ of 0 (measured from the unfolded state and in the direction of folding) implies that the structure at the site of mutation in the transition state is comparable to the structure in the unfolded state. Alternatively a ϕ value of 1 suggests that the structure in the vicinity of the mutation in the transition state is as folded as that found in the native structure. Fractional values of ϕ imply a mixture of states.

Kinetic analysis of the refolding of barnase was consistent with one major intermediate and a linear scheme $\mathrm{U} \rightleftharpoons \mathrm{I} \rightleftharpoons \mathrm{N}$, where the rate-limiting step was conversion of I to N. The $t_{1/2}$ for the observable step of refolding was ~30 ms, although the kinetics of the step preceding I occur within the dead time of stopped flow studies. Since ΔG_{U} and the forward and reverse rates were measured a partial energy diagram for barnase folding was constructed where kinetic and thermodynamic analysis exposed the properties of the transition state and major, late folding, intermediate in wild type and mutant forms of the protein (Figure 11.18).

Residues mutated at the C-terminal end of the α_1 helix show values of ϕ of ~1.0 for the transition and folded states. This implies that this region of the helix-ordered structure is formed rapidly. Interestingly, residues such as Thr6 at the beginning of the helix show intermediate values suggesting disorder is present before formation of native structure. In a similar fashion residues in loop 3 rapidly form structure comparable to the folded state whilst loops 1,2 and 4 remain disordered throughout the folding process until the later stages. Residues at the centre of the β sheet form rapidly but those located at the edges do not fold until later. More significantly, residues

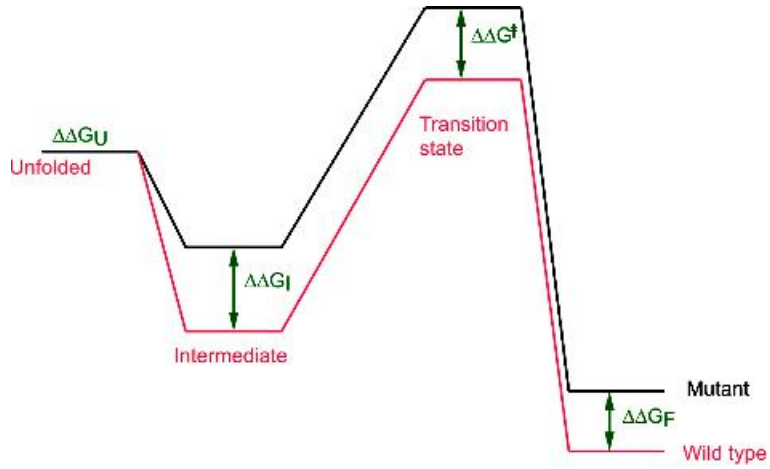

Figure 11.18 Partial free energy diagram describing the folding of wild type and mutant forms of barnase. Although not strictly necessary for simplicity the unfolded forms of the wild type and mutant barnase are shown as iso-energetic. The remaining energy terms describe the differences between the intermediate, transition state and folded forms of wild type and mutant barnase

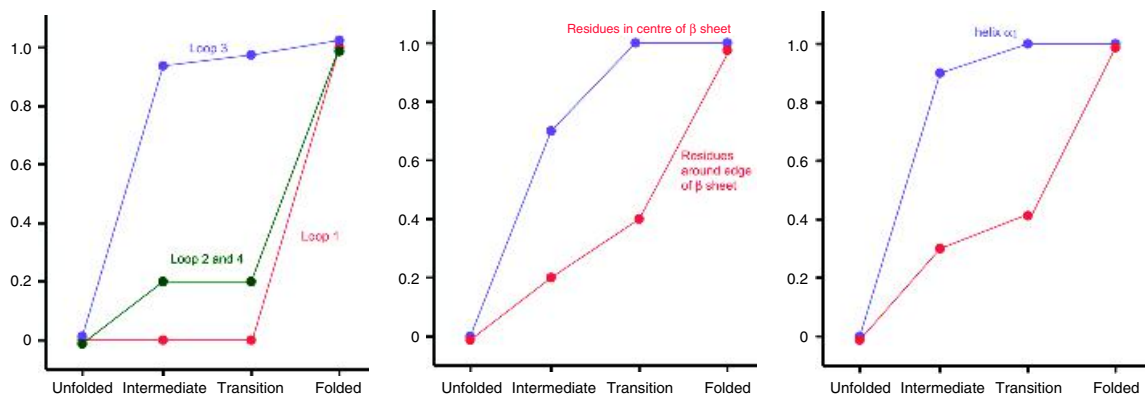

Figure 11.19 ϕ-values found for mutation of residues in the loop, β sheet and helix regions of barnase

located in the core region between the helix and sheet showed ϕ values dependent on location. Residues at the centre of the core exhibit ϕ values close to 1 in the intermediate and subsequent states whilst those located at the periphery of this core showed a gradation of values (Figure 11.19). In this fashion an extraordinary detailed picture of the pathway and energetics of protein folding and its sensitivity to mutation was established for barnase.

Models of protein folding

Exceptionally rates of protein folding are observed with time constants less than 5 ms but most domains with less than 100 residues fold over slower time scales ranging from 10 ms to 1 s. In most small domains folding is a cooperative process but for a smaller group of proteins folding is characterized by the transient population of detectable intermediates. These proteins

offer the possibility of characterizing partially folded states or sub-domains. Initial studies of protein folding focused extensively on helical domains or domains containing mixtures of helix and strand (α/β proteins). Gradually, it has been shown that small β sheet proteins also fold in a highly cooperative manner whilst larger β sheet domains especially those with complex topology frequently exhibit intermediates during folding. Thus, models of protein folding must account for processes spread over a wide range of time scales as well as the presence in some cases of intermediates or partially folded structures.

Helix formation

The formation of segments of helical structure occurs with time constants around 1 μs. The reactions proceed with a slow initial stage as the polypeptide chain struggles to form hydrogen bonds between donor and acceptor separated by three or four residues in a disordered chain. This is followed by faster rates that reflect progressively rapid extension of helices from an ordered structure sometimes called a nucleation site. Most experiments on helix formation have been derived using synthetic homopolypeptides and a major problem is that whilst considerable data has been accumulated for helical regions the formation of β structure relies on more than one element of secondary structure for stabilization and there is no satisfactory model system.

Although helices have regular repeating hydrogen bonds coupled with a uniformity of bond lengths and angles this periodicity masks their marginal stability. In an isolated state most helices will unfold and only synthetic polyalanine helices are reasonably described as stable. Initiation of helix formation is a slow and unfavourable process. This arises because five residues must be precisely positioned to define the first hydrogen bond between residues 1 and 5. Hydrogen bonds between residue i, i + 4 characterize regular α helices whilst the intervening three residues must be arranged to form a turn as well as arranged in a geometry favouring the formation of future hydrogen bonds. There is a large entropic penalty in forming the first turn whilst addition of a subsequent residue to the helix is far less severe in terms of energetic cost. At the same time dipoles associated with the peptide bond in a helix influence the energetics since for the first turn their alignment lies in the same direction (parallel) and is an unfavourable arrangement. With the completion of the first turn subsequent formation of peptide dipoles separated by 3 to 4 residues leads to a head to tail arrangement that may assist helix propagation.

The molten globule

During studies of protein folding it was recognized that intermediate states showing many of the properties of the native fold could be identified. This state became known as the molten globule and is characterized by a compact structure, containing most elements of secondary structure as helices, turns and strands, although many long range or tertiary contacts are lacking. The molten globule has a hydrophobic core consistent with burial of non-polar side chains away from the solvent.

Studies of α-lactalbumin using circular dichroism (CD) under denaturing low pH (<4) conditions revealed formation of a stable intermediate state that was distinguishable from the native *and* unfolded states by possessing far-UV CD spectra resembling the native fold but an unfolded-like near-UV CD spectrum. In other words it appeared to possess considerable secondary structure but no tertiary structure. This state was given the name molten globule and similar forms have been identified during the folding of other proteins. Proteins showing molten globule-like states include many familiar soluble proteins such as myoglobin, α-lactalbumin, β-lactoglobulin, BPTI, cytochrome c and azurin. These proteins differ considerably in their secondary structure content as well as tertiary folds. In contrast, some proteins fold without evidence of molten globule intermediates, and lysozyme is one interesting example, especially since lysozyme and α-lactalbumin are homologous proteins.

Although the molten globule state was identified by its stability and presence at equilibrium under partially denaturing conditions similar states have been detected in kinetic experiments. In particular stopped-flow CD studies appear to show kinetic intermediates populated during folding possessing high levels of secondary structure yet lacking complete side chain packing and extensive tertiary structure. These studies have led to the suggestion that molten-globule like intermediates form as part of the overall process of reaching the native state and have provided a picture in which the

formation of elements of secondary structure occurs prior to consolidation of the tertiary fold.

Hydrophobic interactions

Hydrophobic interactions are intrinsically weak forces between non-polar side chains that lead to dramatic ordering of water. In globular proteins these interactions lead to charged/polar side chains residing on the surface of a protein whilst non-polar side chains are buried on the inside of the molecule. The fluorescence probe anthraquinone naphthalene sulfonate (ANS) binds effectively to exposed hydrophobic surfaces and exhibits a characteristic λ_{max} that changes when the probe is placed in a polar environment. By measuring changes in fluorescence of ANS during protein folding in a rapid mixing experiment it is possible to monitor the formation of hydrophobic cores. From many studies of folding involving proteins with different topologies it has been recognized that different events can occur during the formation of the native state. These events include some or all of the following reactions:

1. Formation of all native contacts in a highly cooperative transition;
2. Folding via a molten globule intermediate;
3. Folding via kinetic intermediates resembling molten globules;
4. Condensation of hydrophobic cores prior to formation of secondary and tertiary structure.

A major problem for experimentalists and theoreticians has been to reconcile all of these different views within a single scheme of protein folding allowing formation of the folded state on time scales between 10 ms and 1 s.

A hierarchic model has been proposed where structure is determined by local interactions with conformational preferences among short sections of peptide chain guiding the polypeptide towards forming larger units of secondary structure. An obvious weakness with this model is that few short peptides in isolation form helices and none will form β strands. Although at first these elements of secondary structure are only metastable it is envisaged that interactions with neighbouring units reinforce stability. Early stages of protein folding are characterized by the interaction of secondary structure via a purely diffusional model in which elements collide to allowing favourable interactions and stable association. These interactions accumulate cooperatively to form larger collections of secondary structure that can acquire stabilizing tertiary structure and long-range order at longer time intervals. The detection of intermediates or transition states containing secondary but not tertiary structure is perhaps supportive of a hierarchic model whilst the fact that secondary structure can be predicated from sequence is also used as a supportive line of evidence.

Recently, considerable attention has focused on a concept called contact order. Although the length of a protein does not show a simple correlation with rates of protein folding the topology of a protein does seem to influence this process. To reflect the concept of topology the term contact order (CO) can be defined as the average sequence distance between all pairs of contacting residues normalized by the total sequence length of the protein and is described as

$$\mathrm{CO} = \frac{1}{LN}\sum^{N} \Delta S_{i,j} \tag{11.29}$$

where N is the total number of contacts, $\Delta S_{i,j}$ is the sequence separation in residues between contacting residues i and j, and L is the total number of residues within the protein. Providing the structure of a protein has been determined at a reasonable level of resolution the CO can be calculated. A correlation exists between CO and observed rates of protein folding. This leads, for example, to proteins that are extensively α helical having low CO and folding more quickly than proteins containing higher proportions of β strand. These proteins are characterized by higher CO. The observation that proteins with low CO tend to fold rapidly could suggest that local contacts are the most effective route towards rapidly reaching the native state.

Irrespective of the precise details during the early events of protein folding a common objective achieved by all schemes is to restrict the conformational space sampled by the protein and therefore to avoid the worst excesses of Levinthal's paradox. Restricting

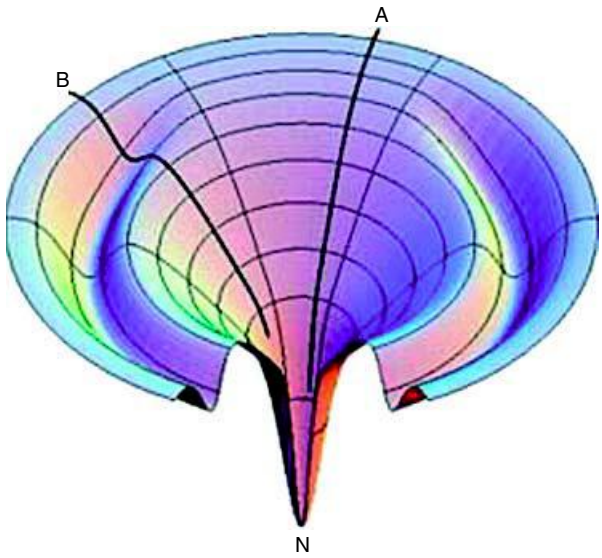

Figure 11.20 The route towards the folded conformation for protein A and B differ. Real energy landscapes are undoubtedly much more complex in profiles

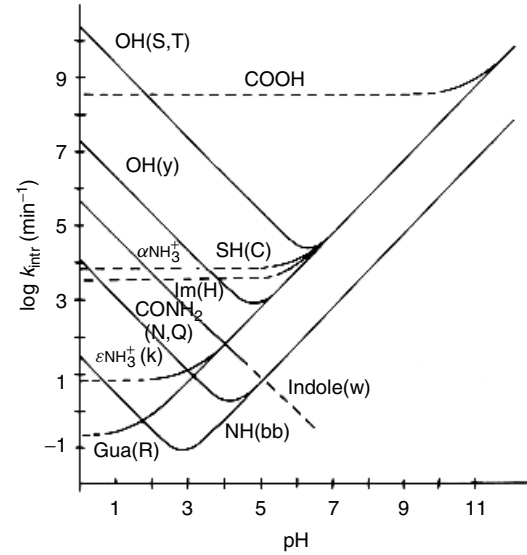

Figure 11.21 Rates of amide exchange for different functional groups in proteins. Backbone amide groups show minimum exchange at pH 3.5 with time constants of ~1 h whilst at pH 10 the rate is increased leading to exchange on the ms time scale

the number of conformations leads to remarkably rapid folding. A new picture of protein folding represented by contoured energy landscapes directs the unfolded protein towards the final (low energy) conformation by a 'funnel' (Figure 11.20). A three-dimensional plot suggests several routes may be used in reaching the native fold with local energy minima existing along pathways. The shape of the funnel directs all conformations towards the native state yet avoids the need to sample all of the possible conformations.

A new view of protein folding is one of energy landscapes, multiple pathways and multiple intermediates where folding leads towards the most stable state in a process influenced by kinetic events.

Amide exchange and measurement of protein folding

The measurement of amide exchange rates is a powerful way of assessing the protection (resistance to exchange) of these groups as a protein folds from denatured to native states. Protection occurs because the amide is involved a hydrogen bond or is buried within the hydrophobic interior of a protein. In NMR experiments based around the observation of a proton the substitution of H by ^{2}H (D) leads to a loss of a signal. Exchange rates are dominated by pH and proteins show profiles where the rate has a minimum around pH 3.5 but increases either side of this minimum due to acid and base catalysis (Figure 11.21). The intrinsic rate of exchange (k_{ex} or k_{intr}) reflects a summation of acid- and base-catalysed rates and differs according to the type of functional group.

$$k_{intr} = k_{OH}[OH^-] + k_H[H^+] + k_w \qquad (11.30)$$

The rates for each of the functional groups found in proteins have different pH optima and are strongly influenced by pH, local chemical environment, solvent, side chain identity, neighbouring residues and temperature. In proteins NH exchange can be extremely slow

and may take months or years to reach completion. A protection factor is described via a parameter, θ_p, where

$$\theta_p = k_{intrinsic}/k_{obs} \qquad (11.31)$$

and a slow rate of exchange is seen by values of θ_p of 10^6–10^7.

A theory of amide exchange in proteins developed by Kai Linderström-Lang involves open conformations where the protein is exposed to the solvent in equilibrium with a closed structure.

$$\text{closed} \underset{k_{clo}}{\overset{k_{op}}{\rightleftharpoons}} \text{open} \xrightarrow{k_{intrinsic}} \text{hydrogen exchange}$$

The closed state is often equated with folded molecules and the open state with denatured proteins. Protected amide hydrogens are 'closed' to exchange but become accessible to exchange through formation of the 'open' state at rates comparable to those observed for unstructured peptides. Two kinetic regimes are recognized for hydrogen exchange in the above scheme. The first regime is called the EX1 mechanism and occurs when $k_{intrinsic} \gg k_{clo}$. The observed rate is therefore approximated by the rate of protein 'opening' such that $k_{obs} = k_{op}$. In contrast the more common mechanism found in proteins occurs when $k_{intrinsic} \ll k_{clo}$ and the observed rate (k_{obs}) is therefore equivalent to the product $k_{op}k_{intrinsic}$. This is called the EX2 mechanism and occurs when a protein is relatively stable and exchange rates are comparatively slow.

In proteins the most important exchangeable proton is attached to the secondary amide of the polypeptide backbone. Amide exchange rates in proteins are readily measured in protein uniformly enriched in ^{15}N in a 'pulse hydrogen' exchange experiment. The principle of the experiment involves transferring a protein from water-supported buffers to those based on ^2H_2O (D_2O) and leads to a loss of cross-peak intensity in 2D NMR spectra as a result of proton/deuteron (H/D) exchange (Figure 11.22). By observing the rate of change in the intensity of cross peaks amide exchange rates are estimated.

Denatured protein in 6 M guanidinium HCl in solutions of ^2H_2O is allowed to fold in solutions containing ^2H_2O and leads to protein 'labelled' with deuterium. Placing the samples at high pH favours rapid exchange for accessible amides and is followed by 'quenching' and a lowering of the solution pH. Folding is allowed to continue to completion and the method samples amides that are rapidly 'removed' from exchange reactions by forming hydrogen bonds, or more probably from burial on the inside of the protein. After folding has been completed the pattern of NH and ND labels is analysed by 2D NMR spectroscopy (Figure 11.23). Increasing the refolding time (t_f) allows a greater number of exchangeable sites to be protected. Using this approach it was possible to sample the protection of over 50 different amide groups in the enzyme lysozyme. The results showed two groups of amides existed; approximately 50 percent were completely protected from exchange within about 200 ms whilst the remaining group remained accessible even after 1 s. Further analysis of rapidly protected amides revealed a location within a domain based around four α helices whilst the more accessible amides were contained in the β sheet rich domain.

This picture of amide protection provides a view of protein folding in which each domain of lysozyme acts as a separate folding unit (Figure 11.24). Studies with other proteins reinforced this view and domains are often defined as 'folding units'. It might be argued that the rates of protection simply reflect the units of secondary structure but working against this idea is the observation that 3_{10} helices found in each domain showed different rates of protection. The different technique used in the study of the kinetics of protein folding are listed in Table 11.3.

Kinetic barriers to refolding

Kinetic studies highlight the presence of transient intermediates during folding that represent partially folded forms along the pathway between denatured and folded states (Figure 11.25). In 'off' pathway intermediates it may be necessary to unfold before completing a folding pathway. When this type of reaction is performed *in vitro* the rate of reaction is slow. *In vivo* a different picture emerges with specific proteins catalysing the elimination of 'unwanted' folds.

Additional barriers to efficient protein folding exist. One major kinetic barrier to protein folding involves *cis*–*trans* isomerization of the amide bond preceding

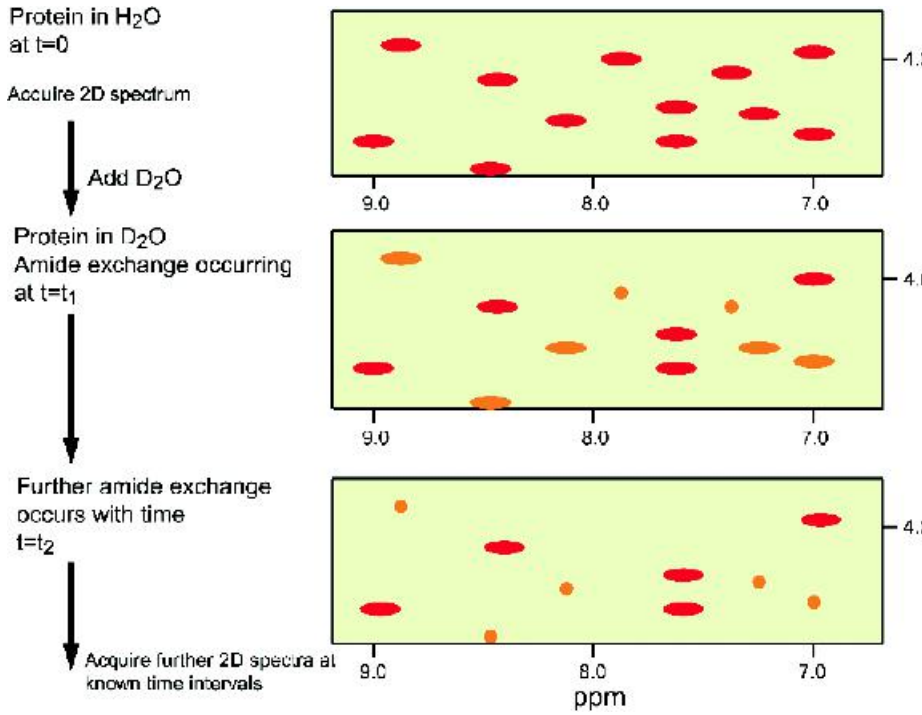

Figure 11.22 The principal of measuring H/D exchange by NMR spectroscopy. Initial cross-peaks are shown in red. Those exhibiting altered intensities are shown in orange and eventually exchange leads to a complete loss of intensity

Table 11.3 Techniques used to study kinetics of protein folding

Method	Structural information
Intrinsic fluorescence	Primarily environment of Trp and Tyr residues. Can also include environment of co-factors such as heme or flavin
Absorbance	Environment of aromatic groups or other conjugated systems
Near UV CD	Asymmetry of aromatic residues within tertiary structure
Far UV CD	Formation of secondary structure
H/D exchange NMR	Formation of persistent hydrogen bonds or amide groups protected from solvent
H/D exchange MS	Detection of folding intermediates and different populations
FTIR	Formation of hydrogen bonds
ANS fluorescence	Accessibility of hydrophobic residues

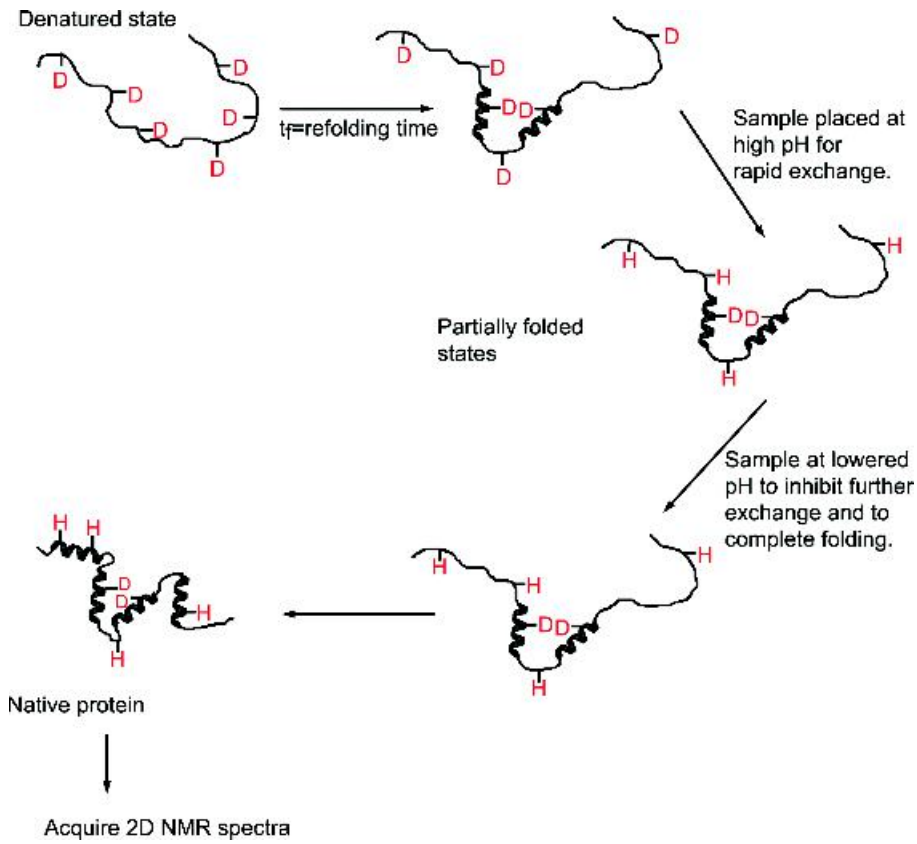

Figure 11.23 The principle of pulsed H/D exchange in a protein during folding

proline residues. Proline has unique properties among the 20 naturally occurring building blocks of proteins stemming from the cyclic side chain and the absence of an amide proton (NH). The peptide bond immediately preceding proline is unlike other peptide bonds in that the ratio of *cis–trans* isomers is ~4. Unfolding removes many of the constraints placed on peptide bonds and leads to randomization of prolyl peptide bonds so that in a protein containing five Pro residues it is expected that on average one prolyl peptide bond will exist in the less favourable *cis* conformation. The potential for proline *cis–trans* isomerization to inhibit the rate of protein folding arose from studies by Robert Baldwin showing that kinetically heterogeneous mixtures of molecules existed in refolding denatured ribonuclease. An explanation for this observation provided by John Brandts centred around the *cis–trans* isomeric state of prolyl peptide bonds. The refolding of a proportion of protein with prolyl peptide bonds in an unfavourable conformation is characterized by slow rates of formation of the native state superimposed on normal rates of folding. Kinetic heterogeneity is observed and arises because molecules with bonds arranged favourably fold rapidly whilst a fraction with unfavourable *cis* peptide bonds will fold more slowly. As a result protein refolding involving *cis–trans* isomerization is characterized by one or more slow refolding phases.

This view has been supported through the use of model peptides containing proline residues and the observation of slow refolding rates in proteins including insulin, tryptophan synthase, T4 lysozyme, staphylococcal nuclease, chymotrypsin inhibitor CI2,

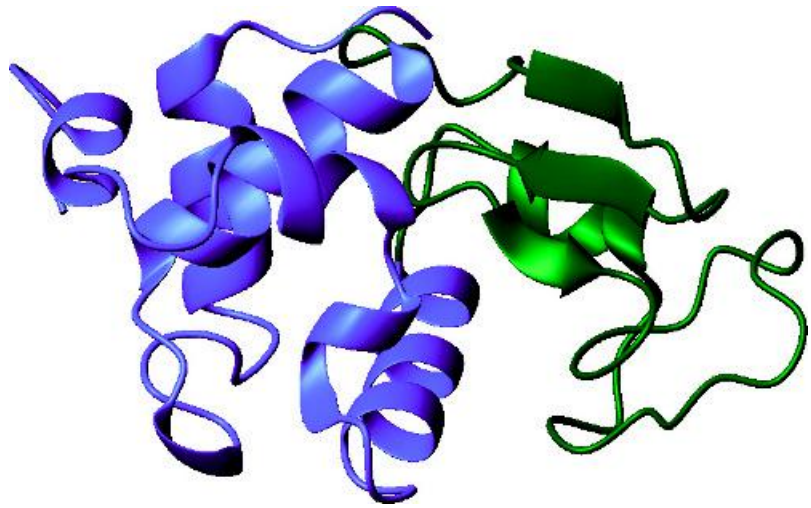

Figure 11.24 Structure of lysozyme showing the rapidly formed helix rich domain (blue) along with the slowly 'protected' sheet-rich region (green)

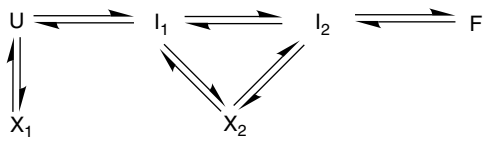

Figure 11.25 On and off pathway intermediates during protein folding. I_1 and I_2 represent on pathway intermediates such as the molten globule like state whilst X_1 and X_2 represent off pathway forms of the protein

barnase, yeast iso-1 and iso-2 cytochromes c, pepsinogen, thioredoxin as well as calcium binding proteins such as parvalbumin and calbindin. One direct observation is the refolding of three homologous parvalbumins (Ca binding proteins found in muscle). Two of the parvalbumins contain a single proline residue and exhibit refolding reactions with a slow component whilst this phase is absent in a third parvalbumin lacking proline.

Cis–trans proline isomerization is one of the major causes of kinetic heterogeneity in protein refolding and deviations away from a purely two-state system. The presence of such effects during refolding are normally confirmed by site directed mutagenesis, by kinetic analysis using 'double mixing' stopped flow techniques and from analysis of the temperature dependency associated with each folding reaction. Those associated with *cis–trans* isomerization of Xaa–Pro peptide bonds are characterized by activation energies ranging 80 to 100 kJ/mol.

In vivo protein folding

Proteins fold to form the native state *in vitro* but do similar pathways or mechanisms apply *in vivo*? Translation occurs at rates of ~20 residues per second, yet many small soluble proteins fold completely within 100 ms. The disparity between optimal translation rates and folding suggests that *in vivo* limitations exist in the latter process.

Peptidyl prolyl isomerases (PPI) catalyse one of two important but slow reactions associated with folding. The reaction is *cis–trans* isomerization of prolyl peptide bonds. The first enzyme was isolated from pig kidney in 1984 but during this period other studies looking at the binding of the immunosuppressant drug cyclosporin A (a vital immunosuppressant during transplant surgery preventing the rejection of

donor organs) discovered its intracellular target and named this protein cyclophilin. Subsequent protein sequencing showed that PPI isolated from pig kidney and cyclophilin were identical. The precise basis for cyclosporin A binding to PPIs remains to be established but it may be unrelated to the activity of the immunosuppressant in decreasing T-cell proliferation and activity. Peptidyl prolyl isomerase overcomes kinetic barriers presented by incorrectly oriented peptide bonds during protein folding.

Additional prolyl isomerases exist within cells and are grouped within one of at least three major protein families. Besides cyclophilins FK506-binding proteins and parvulins represent different families that are not homologous but share a common biochemical function. A plethora of PPIs exist in the human genome with at the last count 11 cyclophilins, 18 FK506 binding proteins and two parvulins encoded.

The second 'slow' reaction of protein folding is the formation of disulfide bonds. This reaction is catalysed by the enzyme protein disulfide isomerase (PDI). Since both reactions (*cis–trans* proline isomerization and disulfide bond formation) were shown to inhibit rapid protein folding *in vitro* the presence of enzymes within cells to specifically catalyse these 'slow' reactions is compelling evidence that similar mechanisms exist for folding *in vitro* and *in vivo*. Protein disulfide isomerase contains the conserved active site motif of Cys-Xaa-Xaa-Cys found in thioredoxin and shares a similar role to the Dsb family of proteins found in *E. coli*. Together with molecular chaperones these systems allow proteins to fold effectively in the cell.

Chaperones

The fundamental experiments of Anfinsen showed that denatured ribonuclease folds *in vitro* with the primary sequence directing folding to the native state. It was widely assumed that folding of all newly synthesized proteins *in vivo* would proceed similarly. Nascent polypeptide chains would not require additional assistance to fold efficiently. With hindsight this assumption seems foolish, and gradually observations established that specialized proteins within the cytosol assist formation of native states.

One of the first observations was that *E. coli* containing a defective operon called *groE* could not assemble wild type bacteriophage λ. This was despite the fact that all protein components of mature λ were encoded by the viral genome. A second observation was that in the mitochondrial matrix translocated polypeptide associated with a protein (Hsp60). The name reflected its mass and that production increased approximately two-fold after heat shock. The name is misleading because experiments showed that deletion of the Hsp60 gene in yeast was lethal and the protein was essential to the cell under all growth conditions. This is expected for a protein playing an important role in the folding of mitochondrial proteins.

The heat-shock or cell-stress response (i.e. changes in the expression of intracellular proteins) is a common event, with increased protein production an essential survival strategy that allows responses to diverse stimuli, including heat or cold, osmotic imbalance, toxins and heavy metals as well as pathophysiological signals. The proteins synthesized in response to such environmental stresses are collectively called heat-shock proteins (or HSPs), stress proteins or, more commonly today, molecular chaperones. The response of cells to stress is a primitive mechanism that appears to be evolutionarily ancient as well as essential to survival. The ancient nature is demonstrated by significant sequence homology in molecular chaperones from a wide range of organisms. Systematic work over the last 20 years has identified many proteins with chaperone-like activity. Chaperonins refer to a sub-group of these proteins that function as components of multimeric systems.

Many newly synthesized proteins reach their folded states *in vivo* spontaneously and without assistance, in processes analogous to the folding of ribonuclease *in vitro*. However, folding efficiency may be limited by side-reactions such as aggregation (Figure 11.26) promoted by transiently exposed hydrophobic surfaces. In *E. coli* aggregation of proteins is readily detected by SDS–PAGE after heat shock. Cells respond to heat shock and the production of significant amounts of unfolded protein by the synthesis of new systems designed to promote refolding. These systems are the molecular chaperones.

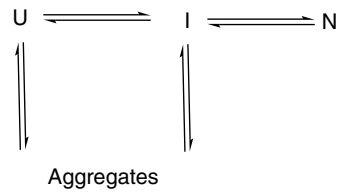

Figure 11.26 Possible routes of aggregation for unfolded and partially-folded protein. Chaperones limit the extent of aggregation reactions by binding to U and I forms

Several molecular chaperone systems have been characterized in *E. coli* including the GroES-GroEL, DnaK-DnaJ-GrpE and ClpB molecular chaperone systems. The acronym Hsp (heat-shock protein) is still used to describe these chaperones and the literature is full of terms such as Hsp60, Hsp70 and Hsp100 for families of homologous proteins. Today these terms are used interchangeably and the Hsp60 name could refer to GroEL – one component of the best-characterized chaperone.

Molecular chaperones are distributed in all cells ranging from archaebacteria through to complex eukaryotic cells. Although many chaperone systems have not been structurally characterized the GroEL-ES, the thermosome of *Thermoplasma acidophilum* and small HSPs from *Methanococcus jannaschii* represent exceptions.

The GroEL-ES system

The GroEL-ES chaperone system is essential for *E. coli* growth under all conditions with mutations in the *GroE* operon proving lethal. The GroEL-ES complex of *E. coli* has formed the basis of understanding chaperone function as a result of detailed structural studies using crystallography and electron microscopy. *In vivo* GroES is composed of seven identical subunits (M_r~10 000) whilst GroEL is composed of 14 larger subunits each with a mass of ~60 000. This leads to Gro-EL being called chaperonin60 (cpn60) and Gro-ES chaperonin10 (cpn10).

A model for the organization of GroEL obtained from negative staining EM studies showed a double ring cylindrical structure with a diameter of ~14 nm and height ~16 nm. The seven-fold symmetrical complexes were of comparable size to the ribosome with a central cavity of ~6 nm that bound polypeptide prior to ATP binding or GroES attachment. The first detailed structure for GroEL was determined by Paul Sigler and provided beautiful detail of the toroidal architecture indicated from electron microscopy. GroEL had a substantial central cavity in a cylindrical structure composed of two stacked heptameric rings arranged with seven-fold rotational symmetry. The rings are arranged back to back, contacting each other through a region known as the equatorial interface that is one of three distinct domains found in each GroEL subunit. The remaining domains are the apical and intermediate domains; these zones are simply designated as E, I and A domains (Figure 11.27).

GroEL contains a single polypeptide chain of 550 residues starting at the E domain extending to the intermediate region before the central part of the sequence makes up the A domain. The sequence then returns to form part of the I domain before terminating in the E domain. The E domain is a helical rich region that provides the ATP/ADP binding site whilst the apical domain shows greater mobility and has a lower percentage of regular secondary structure. Schematically the GroEL subunit is a tripartite structure (Figure 11.28).

The structure of GroES was determined independently of GroEL to confirm seven-fold symmetry arranged in a dome-shaped architecture. The dome contained seven subunits each of 110 residues and formed a core β barrel structure with two prominent loop structures (Figure 11.29).

One remarkable feature of the GroEL structure is a massive conformational change caused by binding co-chaperonin (GroES) and nucleotides. Revealed by cryo-EM studies of the ternary complex formed between GroEL/ADP/GroES image reconstructions showed a small elongation of the cylinder upon ATP binding but a very substantial elongation at one end of the GroEL complex upon GroES binding. This was accompanied by increases in diameter at the end cavities with the GroES ring observed as a

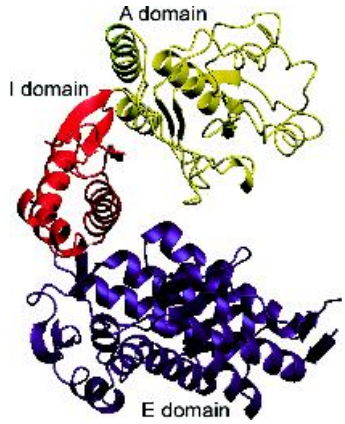

Figure 11.27 The monomeric structure of GroEL showing relative positions of E, I and A domains (PDB: 1GRL)

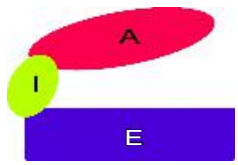

Figure 11.28 The tripartite structure of a GroEL subunit showing a schematic arrangement of apical (A), intermediate (I) and equatorial domains

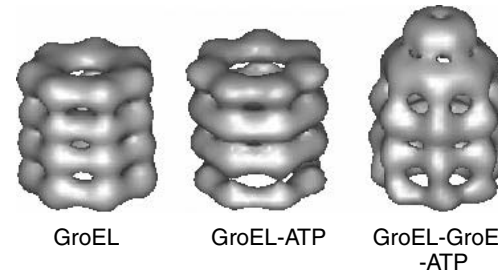

Figure 11.30 3D reconstructions of the surfaces of GroEL, GroEL-ATP and GroEL-GroES-ATP from cryo-EM. The GroES ring is seen as a disc above the GroEL and contributes to the bullet-like shape. (Reproduced with permission from Chen, S. *et al*. *Nature* 1994, **371**, 261–264. Macmillan)

flat, disc-like object above the neighbouring GroEL heptamer (Figure 11.30). In the presence of adenine nucleotides GroES binds to the stacked rings of GroEL forming a 'bullet-like' asymmetric structure with GroES attached to only one of the heptameric rings of GroEL (the *cis* ring).

The chaperonin-assisted catalysis of protein folding proceeded through cycles of ATP binding and hydrolysis. Nucleotide binding modulates the interaction between GroEL and GroES binding. ATP binding to GroEL occurs initially with low affinity ($K_d \sim 4$ mM)

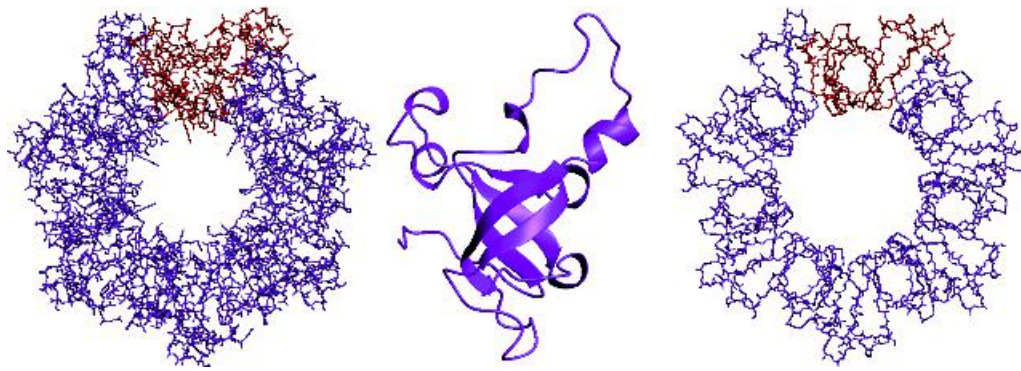

Figure 11.29 The structure of GroES within the heptameric ring seen in side and superior views. One subunit within the ring is shown in red. The arrangement of secondary structure showing the β barrel core is shown from the side in the centre for an individual subunit (PDB: 1G31)

to one subunit but as a result of positive cooperativity binding occurs with substantially higher affinity to the remaining subunits (K_d~10 μM). The cooperative nature of ATP binding establishes GroEL as an allosteric enzyme. In GroEL with a total of 14 subunits in two heptameric rings it is observed that tight nucleotide binding occurs only to one of the two available rings. ATP binding to the remaining ring is inhibited but promotes GroES association and a change in conformation for the *cis* ring subunits. The change in conformation is central to the switch in binding modes from an allosteric T state with low affinity for ATP and high affinity for non-native chains to an R state that shows the opposite trends.

The molecular basis for some of the conformational changes have been understood by comparing the structures of Gro-EL complexes in the presence of Gro-ES as well as ADP, ATP and AMP-PNP, using cryo-EM methods. Although the resolution of these studies was relatively low by crystallographic standards the large size of the complex meant that levels of resolution were sufficient to identify key catalytic features. Nucleotide binding in the equatorial domain pocket leads to substantial changes in the GroEL heptameric ring structures and movement in particular for the apical domain. The result of domain movement is the existence of several distinct structures.

Crystallographic analysis of the GroEL-GroES-$(ADP)_7$ complex showed a similar bullet shape and by comparing this structure with the un-liganded GroEL structure the significant conformational perturbations were shown to arise from 'en bloc' movements of I and A domains with respect to the E domain in the *cis* ring capped by GroES (Figure 11.31). The overall architecture of GroEL and the GroEL-GroES-$(ADP)_7$ complex emphasized the bullet shape (Figure 11.32) and allowed insight into the mechanism of binding partially folded proteins and preventing protein aggregation.

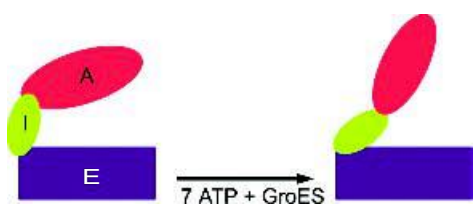

Figure 11.31 Domain movement in GroEL as a result of GroES and ATP binding

The reorganization of the *cis* ring of GroEL results from domain re-arrangement involving intermediate and apical domains with the I domain swinging downwards towards the E domain by approximately 25°. One effect of this movement is to close the occupied nucleotide (ATP) binding site located on the top inner surface of the equatorial domain. The A domain shows greater movements, swinging ~60° upwards relative to the equator, but also twisting about its long axis by ~90° to form new interfaces with neighbouring A domains. These movements lead to interactions between the A domain of GroEL and mobile loops on GroES. In comparison with the A domain the *cis* equatorial (E) domains do not show large conformational shifts, with smaller inward movement of the *cis* assembly by ~4°. Since these regions interact with the neighbouring heptameric ring there is a complementary outwards tilt in the E domains of the *trans* ring. These movements have important functional consequences in the overall cycle of catalysis. In contrast to the large conformational changes occurring in GroEL, the structure of the GroES ring within the complex is similar to that observed in the standalone structure. A minor exception to this rule would include mobile and disordered loops that become structured within the complex as a result of interaction with the A domain.

The chaperonin catalytic cycle

With the structure of the 'inactive' GroEL and 'active' GroEL-GroES-ATP complexes known it was possible to attempt to understand how this macromolecular complex assists in protein folding. At the simplest level the complex catalyses cyclic binding and release of target polypeptides. This process is divided into four distinct stages or phases (Figure 11.33). Phase I involves polypeptide binding; phase II the release of the polypeptide into the central channel and the initiation of folding; phase III involves hydrolysis of ATP, reorganization of the *cis* heptameric ring and the start of product (GroES, folded peptide and ADP) release; finally, phase IV involves ATP binding to GroEL subunits of the *trans* ring providing the trigger

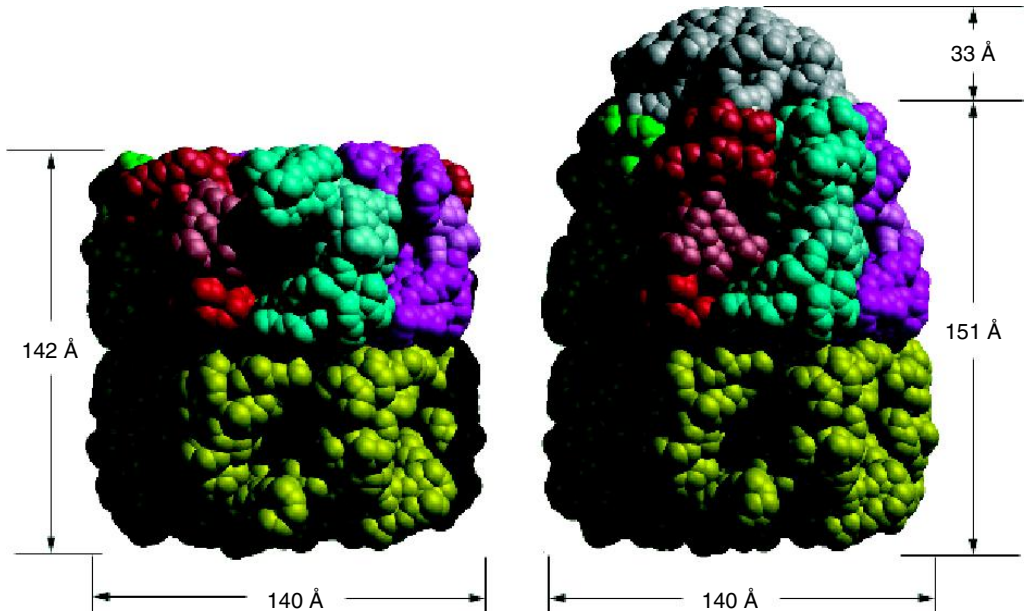

Figure 11.32 Structures of GroEL and GroEL-GroES-$(ADP)_7$ complex derived by X-ray crystallography. Different colours denote the individual subunits in the upper heptameric ring with the apical, intermediate and equatorial regions show in progressively darker hues. In the lower ring all subunits are shown in yellow. GroES is shown in grey (reproduced with permission Sigler, P.B. *et al. Annu. Rev. Biochem.* 1998, **67**, 581–608. Annual Reviews Inc)

to discharge GroES and entrapped folded polypeptide from the opposite side of the structure.

The cavity within GroEL provides a 'reaction vessel' in which the polypeptide chain can be 'incubated' until folding is complete. This vessel is called the 'Anfinsen cage' and estimations of internal volume place an upper limit of $\sim 85\,000$ Å^3 which is consistent with a spherical protein of 70 kDa.

In the first phase unfolded polypeptide binds to the apical domain largely through hydrophobic interactions and the direct involvement of at least nine residues (determined by site-directed mutagenesis) on each subunit. Eight of these residues have hydrophobic side chains and point inwards into the cavity creating a hydrophobic ring that binds target polypeptides. The central channel of GroEL functions as two separate cavities, one in each ring, as a result of a disordered 24 residue C-terminal region in each of the seven ring subunits that effectively 'blocks' communication between channels.

The second stage of the catalytic cycle involves nucleotide binding. From observations of GroEL-assisted refolding experiments it was noted that GroEL alone inhibits refolding whilst in the presence of K^+ ions, Mg-ATP and GroES efficient folding to the native state was seen. This led to a proposal that at least two distinct conformations of the complex existed; one that binds unfolded polypeptides tightly and ATP weakly and another in which the binding properties are reversed. Nucleotide-modulated conformational changes of GroEL are inherent to the protein folding cycle. ATP binding preceded GroES binding with the latter event occurring at near diffusion controlled rates (4×10^7 M^{-1} s^{-1}) in the presence of ATP, but approximately 100 times more slowly in the presence of ADP.

The third stage involves completion of polypeptide folding coupled with its release into the cavity of the *cis* ring of the GroEL-ES complex. Binding of unfolded polypeptide to the A domain has suggested

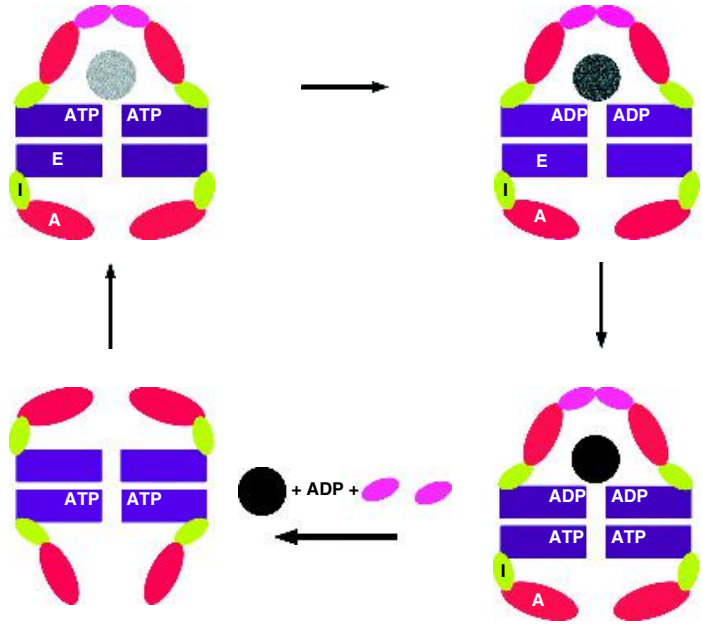

Figure 11.33 The chaperonin catalytic cycle (phases II–IV). ATP and GroES (purple subunits) bind to the GroEL ring causing change in conformation. ATP is trapped inside complex. In this state non-native polypeptide chains will no longer bind to apical domain but are released into central channel. ATP on *cis* ring is hydrolysed. Protein folds in central cavity. ATP binds to *trans* ring causing conformational change in opposite ring that releases GroES, ADP and folded peptide

the involvement of residues Leu234, Leu237, Val259, Leu263, and Val264. Mutation of any of these residues inhibits correct protein folding and in some instances proves lethal to *E. coli*. ATP and GroES binding cause 'en bloc' movements of the E, I and A domains and lead to competition for these sites within the complex and the rotation and displacement of the heptameric rings in opposite directions. As a result, unfolded chains are no longer bound and are released into the cavity of the *cis* ring where folding occurs. The non-polar side chains of residues responsible for binding the non-native polypeptide through hydrophobic interactions in the cavity of the un-liganded ring are now buried in the *cis* ring assembly. This leads to their replacement by polar residues on the walls of the cavity. This change is very significant because it now drives folding where the released polypeptide chain is able to re-initiate folding in an enlarged cavity whose lining is now hydrophilic.

The fourth phase represents the dissociation of the ligands (GroES and ADP) from the complex. The disassembly of the *cis* ring complex is triggered by hydrolysis of ATP to ADP with the latter forming weaker interactions at the nucleotide binding site as a result of the loss of a phosphate group. Weaker interactions breakdown the GroEL–GroES complex with the susceptibility to disassemble increased further by binding ATP to the opposite ring. This event also initiates the start of another cycle of binding of the non-native polypeptide chain and GroES association.

In summary, the GroEL-ES complex represents a remarkable molecular machine designed to ensure protein folding occurs efficiently and without the potential for complicating side reactions. As with many important protein complexes it appears that nature has seized the 'design' and used it in other chaperonin-based systems. Whilst GroEL and GroES are respectively the prototypical chaperonin and co-chaperonin

there exist in other cells chaperonins that form similar ring-shaped assemblies. The most convincing evidence derives from the structure of the corresponding complex from hyperthermophilic bacteria.

It is likely that only a fraction of all cytosolic proteins found in *E. coli*, or indeed in other cells, are processed via molecular chaperonins. Some proteins will be too large to be accommodated within the central cavity of chaperonins whilst other nascent polypeptides, particular small soluble domains, can fold 'safely' without aggregating and without the intervention of other proteins. Estimates of the number of proteins acting as substrates for GroEL vary but at least 300 translated polypeptides including essential enzymes of the cell have been identified as using this system. About one-third of these proteins unfold within the cell and repeatedly return to GroEL for conformational maintenance. It appears that GroEL substrates consist preferentially of two or more domains of the α/β class containing mixtures of helices and buried β sheets with extensive hydrophobic surfaces. Such proteins are expected to fold slowly and to be aggregation-prone whilst the hydrophobic binding regions of GroEL are well suited to interaction with non-native states of α/β proteins.

The thermosome of T. acidophilum

The characterization of the GroEL complex as a toroidal structure based on two heptameric rings capped by a dome like GroES ring defined a major group of chaperonins found widely throughout cells. Systems based on this design are called Type I chaperonins. Sequence homology studies show that eukaryotic cells have similar protein complexes. These complexes are located within the mitochondria and chloroplasts with Hsp60, and Cpn60 representing two prominent systems. This similarity is not surprising in view of the endosymbiont origin of these organelles from eubacteria.

In *T. acidophilum* comparable reactions are catalysed by thermosomes that show variations in organization and are example of Type II chaperonins. The structure of the thermosome from *T. acidophilum* was the first Type II chaperonin to be described in detail. It is a hexadecamer composed of two eight membered rings. Each ring contained an alternating series of two polypeptide chains denoted as $(\alpha\beta)_4(\alpha\beta)_4$. Group II chaperonins in archaea and in the eukaryotic cytosol have been shown to use the same mechanism as type I chaperonins, with the binding of substrate to a central cavity and ATP dependent substrate release. The thermosome shares many of the properties of the GroEL rings although the lid structure is derived from the secondary structure elements of the apical domains and does not involve GroES like subunits to cap a central cavity.

Strands S12 and S13 and most of the N-terminal part of helix H10 protrude toward the pseudo eight-fold molecular axis at the ends of the particle to block the entrance to the central cavity and form a lid domain. The β strand (S13) from each A domain forms a circularly closed β sheet of eight strands and leaves a central pore of diameter ~2.0 nm. The triangle of structural elements, strands S12 and S13 and helix H10, pack against corresponding elements in the neighbouring two subunits forming a hydrophobic core with nine non-polar side chains and a threonine residue from each subunit (Figure 11.34). The core is hydrophobic with a third of all lid segment residues buried completely and these interactions assist in maintaining the lid structure in a conformation comparable to that observed in the GroEL-ES complex.

Membrane protein folding

The preceding sections have dealt largely with the folding of soluble proteins *in vitro* and the assistance of macromolecular assemblies in governing folding *in vivo*. As with much of biochemistry the details of folding in soluble proteins has largely outpaced understanding of the same reactions in membrane bound systems. A key observation made in 1982 by H. Gobind Khorana was that bacteriorhodopsin could be denatured and refolded in a similar fashion to ribonuclease. This confirmed that membrane proteins possess all of the information necessary for folding and formation of the native state within their primary sequences. In the case of bacteriorhodopsin this involved the formation

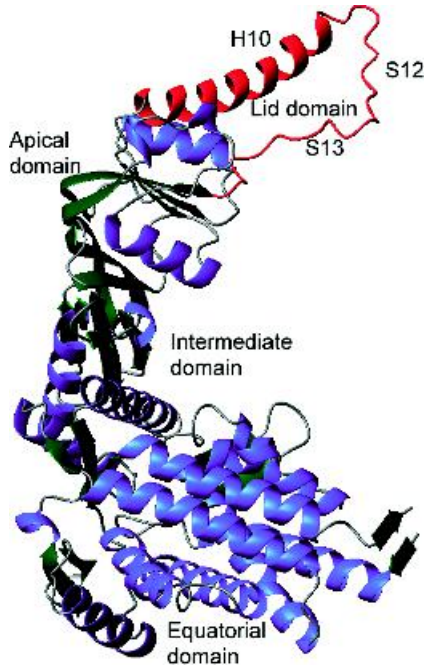

Figure 11.34 The structural arrangement of the thermosome E, I and A domains with the extended helix–strand–strand regions of H10, S12 and S13 shown in red (PDB: *1AG6*)

of seven transmembrane helices, but similar reactions have not been demonstrated for all membrane proteins. Most data currently available refers to the β-barrel rich domains of porins and the helix-rich domains of rhodopsin-like proteins.

Soluble proteins represent the minimum free energy conformation, as defined in Anfinsen's original experiments, and membrane proteins are no exception. However, the mechanism of reaching this folded state *in vitro*, and especially *in vivo*, is less obvious. In soluble proteins thermodynamic studies of the folding process in response to denaturing agents provided information on forces stabilizing tertiary structure. This approach is not successful for membrane proteins as they are very resistant to denaturation. Bacteriorhodopsin does not unfold completely in SDS, retaining helical structure, and denaturation arises primarily from subunit dissociation, loss of interactions between secondary structure elements and from unfolding of domains located either in the aqueous phase or at interfaces.

As a membrane protein bacteriorhodopsin is representative of proteins with seven transmembrane helices – a very large family – with the added advantage of a chromophore that is exquisitely sensitive to changes in physical properties. In the purple membrane bacteriorhodopsin is found as a two-dimensional crystalline lattice where the protein aggregates as collections of trimers. Two thermal transitions with enthalpies of denaturation (ΔH_d) of ~30 and ~400 kJ/mol have been linked to a dissociation of the lattice and dispersion into individual trimers at 80 °C followed by irreversible disruption of trimers via the loss of helix–helix interactions between monomers at 100 °C.

One framework for describing membrane protein folding is the two-stage model proposed by Donald Engelman that involves in the first step the independent formation of hydrophobic helices of ~25 residues upon insertion into the membrane. Stage II involves the ordering of structure where interhelix interactions drive proteins towards the native state. Insertion and ordering are believed to be separate and independent events. Supporting evidence for this model was obtained using fragments of the polypeptide chain of bacteriorhodopsin where five helices (A–E) assembled in the absence of the rest of the protein. This confirmed the ability of fragments to form stable secondary structure in membranes.

These studies emphasize the thermodynamics of protein folding, but an equally relevant aspect of the overall process is the kinetic parameters associated with formation of the native state. Folding kinetics of bacteriorhodopsin measured using CD and absorbance spectrophotometry indicate a rate limiting step is the formation of transmembrane helices with the formation of one or more helices allowing helix–helix packing and rapid completion of protein folding. Completion occurs upon retinal binding and a pathway defined by a series of intermediates with near native secondary structure but lacking the tertiary organization is observed (Figure 11.35).

Intermediate I_1 is formed within a few hundred ms whilst the transition from I_1 to I_2 is rate limiting and equated with the formation of transmembrane helices.

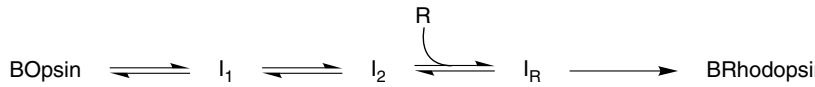

Figure 11.35 Sequential folding pathway for assembly of bacteriorhodopsin from co-factor (R) and apo-protein (bacterio-opsin)

It occurs slowly on a time scale of 10–100 s. These kinetics distinguish the process from those seen in soluble proteins that are frequently complete within 1 s. It seems unlikely that these reactions could occur so slowly *in vivo*, although data is lacking.

Fairly obviously, folding schemes derived from studies of helical membrane proteins may not be of great relevance to systems based around the β barrel, such as porins. Despite different structures experimental evidence suggests similarities extend to the processes of insertion and assembly. Porin OmpA is monomeric and this simplifies protein folding since there is no assembly into trimers. OmpA is synthesized with a signal sequence that is cleaved in the periplasm to leave a mature protein that inserts into membranes. *In vitro* OmpA can be completely refolded in artificial lipid bilayers and because of the moderate hydrophobicity of the individual strands of the β barrel these proteins are easily extracted from membranes in an unfolded state with urea or guanidinium chloride. In this state OmpA has been shown by CD spectroscopy to be unfolded at high denaturant concentrations. Rapid denaturant dilution in the presence of lipid leads to folded, membrane-inserted, conformations with several 'events' identified during a folding pathway distinguished by time constants of several minutes *in vitro*. The formation of individual β strands is unlikely and has not been described for any isolated sequence, suggesting that insertion of OmpA into the membrane may involve partial folding and the interaction of several strands. Despite a slow series of reactions the broad similarity to the folding of bacteriorhodopsin – a protein with very different topology – is reassuring and suggests that an outline of the basic pathway of membrane protein folding has been uncovered. Gradually, the folding of membrane proteins is becoming understood and it is clear that no new principles are involved in the formation of the native state (see Table 11.4).

Translocons and in vivo membrane protein folding

The details of membrane protein folding *in vivo* involve co-translational insertion of nascent proteins into the ER membrane at sites termed translocons. Translocons consist of membrane proteins that form a pore into which newly synthesized polypeptide chains enter as part of the normal cell trafficking pathways. Helical membrane proteins are targeted to the ER membrane in eukaryotic cells and to the plasma membrane in bacterial cells by a signal sequence composed of approximately 20–25 residues. Interaction between the signal sequence and the signal recognition particle (SRP) arrests translation and promotes binding to receptors with the direction of polypeptide chains into the translocon followed by integration into the membrane.

Uncovering the organizational structure of the translocon relied on yeast genetic studies of protein secretion (*SEC*) genes. Using electron microscopy it was possible to identify the cytological event perturbed by lesions and one set of experiments identified mutants (*SEC61*) defective in transport into the ER with unprocessed proteins accumulating in the cytosol. Subsequent cloning of the mammalian version of *SEC61* established a close homology with yeast and between prokaryote and eukaryote. The *SEC61* gene encoded the main channel-forming subunits within the translocon.

The translocon or Sec61 complex consists of three integral membrane proteins Sec61α, Sec61β and Sec61γ. Sec61α has many transmembrane helices whilst Sec61β and γ contain only single helical segments crossing the bilayer. In 1993 elegant reconstitution experiments by Tom Rapoport showed that it was possible to mimic protein translocation in artificial membranes containing the incorporated Sec61 complex although the catalytic efficiency of the reconstituted

Table 11.4 Some of the membrane proteins whose folding has been studied

Protein	Structure and function	Protein denaturation	Formation of folded protein
Bacteriorhodopsin	7 TM helices and retinal co-factor; light-driven protein pump	Apoprotein fully denatured in trifluoroacetic acid; apoprotein partly denatured in SDS, leaving ~55 % of native helix	Transfer from organic acid to SDS and then folded by mixing with lipid vesicles
LHC-II	3 TM α-helices and 1 short amphipathic helix at membrane surface. Photosynthetic light-harvesting protein	Apoprotein partly denatured in SDS, leaving about 30 % of native helix content	Reconstituted with thylakoid membrane extracts by freeze–thaw. Reconstituted in micelles containing pigments, lipid and detergent
E. coli DGK	Unknown structure	Apoprotein slightly denatured by SDS, leaving ~85 % of native helix content	Refolded in DM micelles
E. coli OmpA	Membrane domain of 170 amino acids in 8-stranded β barrel. Exact function unknown, but likely channel	Protein completely denatured in urea	Folding in lipid vesicles by mixing with urea-denatured state
E. coli OmpF	16-stranded β barrel. Trimer forms pore in outer membrane	Completely denatured in urea or GdnHCl	Poor folding in lipid vesicles on mixing the urea-denatured state. Folding increased if detergent used in mixed micelles

Adapted from Booth, P. *et al*. *Biochem. Soc. Trans.* 2001, **29**, 408–413. Abbreviations: TM, transmembrane; DM, dodecylmaltoside; GdnHCl, guanidine hydrochloride; LDS, lithium dodecylsulphate; OG, octylglucoside; PG, phosphatidyl-D,L-glycerol; DGDG, digalactosyl diacylglycerol; LHC-II, light harvesting complex of higher plants; DGK, diacylglycerol kinase

complex was low. With structural characterization of the SRP and its receptor greater emphasis has been placed on the organization of the translocon (Figure 11.36). Electrophysiological studies emphasize the presence of an ion channel in the translocon whilst studies with fluorescent labels attached to residues within a nascent polypeptide chain indicate changes in environment and conformation. Much work remains to

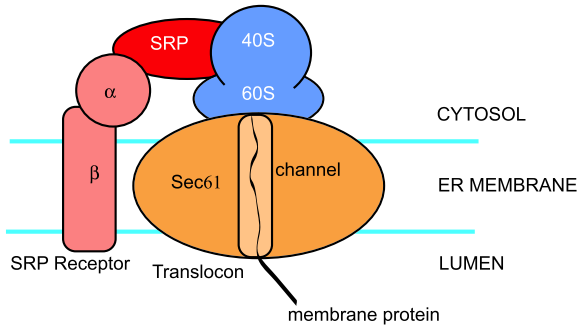

Figure 11.36 A model for the organization of the translocon. The signal recognition particle (SRP)–ribosome complex binds to receptor (α/β) in the ER membrane. The translocon consists of the Sec61 proteins (α, β, γ) and accessory proteins. Translocation requires GTP hydrolysis and the growing polypeptide chain is pushed across the membrane towards the lumen

be done in the characterization of the translocon and its role in membrane protein folding.

A key event in membrane protein folding is identification of a signal peptide by the SRP, but membrane proteins may contain many blocks of 20–25 non-polar residues each reminiscent of such a peptide. This could lead to unwanted transfer into the ER lumen and therefore requires specific recognition mechanisms. The translocon complex must also deal with integral proteins that differ considerably in topology. One answer to this problem lies in the presence of sequences within membrane proteins that control their assembly within the ribosome–translocon complex. These sequences are known as topogenic signals and include conventional N-terminal signal peptides, signal-anchor sequences, reverse signal-anchor sequences, and start-transfer/stop-transfer sequences.

Proteins with more than one transmembrane segment are possibly threaded into membranes through the effect of start- and stop-transfer sequences. In eukaryotes the main features of these transfer sequences are groupings of hydrophobic residues flanked by residues with positively charged side chains. The effect of these sequences is to promote and arrest the transfer of residues through the translocon (Figure 11.37). This process aids the formation of transmembrane helices but is also vital to the formation of important loop regions that frequently link such domains.

The mechanism by which transmembrane domains exit the translocon pore to reach the bilayer remains unclear. Evidence suggests that transmembrane helices fold within this pore and then exit to the bilayer either individually or in pairs. This scheme would lead to the final helix packing reactions occurring in the membrane bilayer after translocation.

Protein misfolding and the disease state

With a greater understanding of folding *in vivo* and *in vitro* has come the realization that diseases arise as a consequence of protein misfolding. Mutation within coding regions of genes can result in the insertion of a stop codon and the failure to synthesize full-length, folded, protein. Alternatively, mutations change the identity of one or more residues leading to a protein with altered folding properties. Ultimately, this research is traced back to studies of sickle cell anaemia identifying mutant haemoglobin as the basis for disease. Similar origins for disease (Table 11.5) occur in protein folding where mutations cause defective folding, aberrant assembly and incomplete processing.

Intracellular sorting and the emergence of defective protein trafficking

A key requirement for cellular function is that proteins are targeted to the correct compartment. In eukaryotic cells compartmentalization requires specialized sorting pathways and the ER and Golgi apparatus form part of the endocytic pathway where a key event is the budding (endocytosis) of membranes as part of the trafficking process. Defective protein folding often results in an inability to transfer mutant domains through the endocytic pathway.

One of the best examples described to date occurs in cystic fibrosis. Cystic fibrosis is very common in Caucasian populations with an incidence of approximately 1 in 20. In a homozygous state this results in a defective CFTR protein and a disease characterized by an inability to transport chloride across membranes effectively and problems achieving correct ion balance.

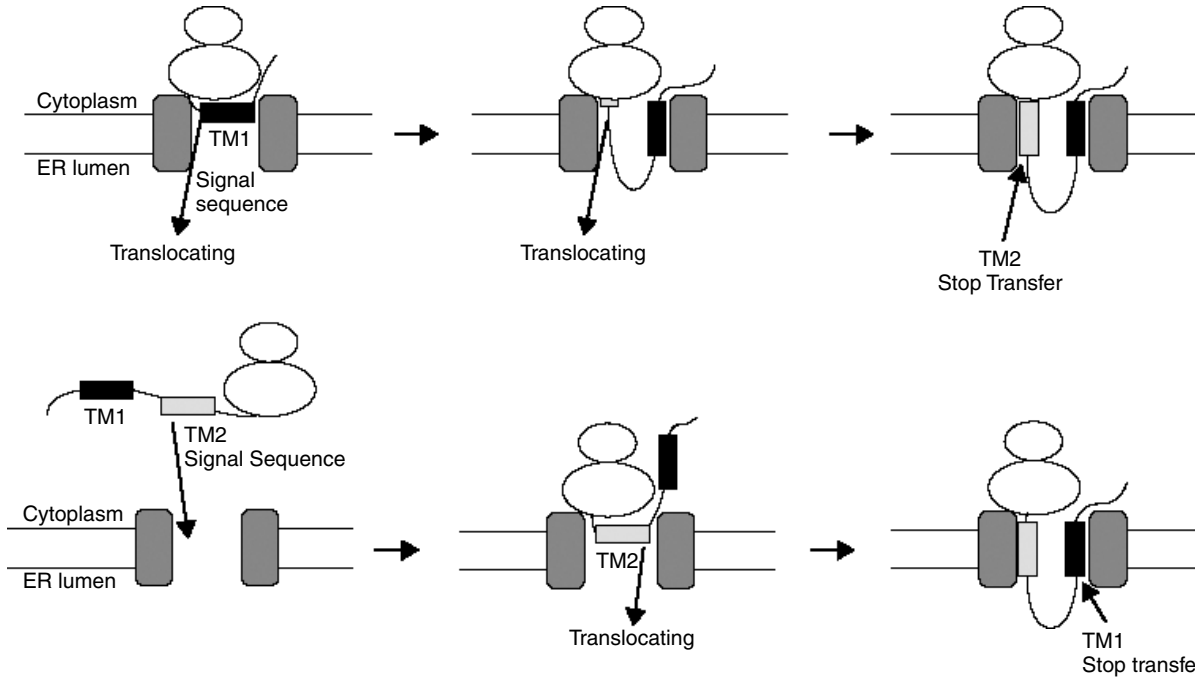

Figure 11.37 The function of topogenic sequences in membrane protein insertion. In the first mechanism TM1 functions as the signal sequence to initiate translocation of the carboxy-terminal region whilst TM2 functions as a stop transfer sequence halting translocation until the next start signal. Alternative mechanisms appear to exist such as that shown in the second pathway where TM2 initiates translocation of the N-terminal region and TM1 acts as the stop transfer signal (Reproduced with permission from Dalbey, R.E, Chen, M.Y., Jiang, F. & Samuelson, J.C. *Curr. Opin. Cell Biol.* 2000, **12**, 435–442. Elsevier)

Over 600 different mutations have been identified in the CFTR protein but the most common mutation involves deletion of one residue (Phe508). This does not lead to a truncated form of CFTR but instead results in a full-length protein missing a single residue. The mutant CFTR protein (ΔF508) is expressed and the translated product is easily detected. However, the protein is not processed further and is unable to complete post-translational glycosylation and does not transfer through the Golgi apparatus. Most of the protein becomes ubiquitinated and marked for destruction by the proteasome but a small amount (~1 percent) reaches the surface where it retains anion selectivity and conductance properties comparable to the wild type protein but has a decreased probability of ion channel opening. The result highlights how small 'defects' lead to major physiological impairment.

Protein aggregation and amyloidosis

Enormous interest has focussed on the appearance of large protein aggregates within tissues in diverse disease states. The appearance of protein deposits in close association with neurofibrillary networks in the cells of brain tissue has long been known to occur in Alzheimer's disease. Protein aggregation also underlies other diseases and may occur in tissues such as the liver and spleen, besides the brain. The protein aggregates are linked by formation of elongated fibrils (amyloid) and diseases showing this property are collectively grouped together by the term amyloidosis.[3]

[3]The term 'amyloid' was originally used to describe the resemblance of protein aggregates associated with disease states to the appearance of starch chains (amylose).

Table 11.5 Diseases arising from 'folding' defects

Disease	Protein or precursor involved in disease
Cystic fibrosis	CFTR
Primary systemic amyloidosis	Immunoglobulin light chain
Medullary carcinoma of the thyroid	Calcitonin
Osteogenesis imperfecta	Collagen (type I procollagen)
Lung, colon and other forms of cancer	p53 transcription factor
Maple syrup disease	α-ketoacid dehydrogenase complex
Amyotrophic lateral scoliosis	Superoxide dismutase
Creutzfeldt–Jakob disease, vCJD, scrapie, fatal familial insomnia	Prion protein
Alzheimer's disease	β-amyloid protein
Cataracts	Crystallins
Atrial amyloidosis	Atrial natriuretic factor (ANF)
Senile systemic amyloidosis, Familial amyloidosis	Transthyretin
Tay–Sachs disease	β-hexosaminidase
Hereditary emphysema	α-antitrypsin
Retinitis pigmentosa	Rhodopsin
Hereditary non-neuropathic systemic amyloidosis	Lysozyme
Transmissible spongiform encephalopathy (TSEs)	Prion protein
Huntington's chorea	Huntingtin
Familial hypercholesterolaemia	LDL receptor

Amyloidosis is apparent in many clinical conditions besides Alzheimer's and includes familiar diseases such as Parkinson's disease, type II diabetes together with less common conditions such as the spongiform encephalopathies. Some of these diseases are genetically determined (i.e. inherited) whilst others are acquired (sporadic) and others may be infectiously transmitted. In each amyloid disease a different protein or fragment aggregates, forming a fibril, and in systemic forms of amyloidosis aggregation leads to the deposition of kg quantities of protein. Ultimately, the progressive deposition of fibrils will cause death, especially when it occurs in a vital organ.

The association between formation of protein fibrils and irreversible protein aggregation was deduced from a number of studies but particularly those giving rise to a condition known as familial amyloidotic polyneuropathy (FAP). The disease was recognized accurately in the 1950s in a Portuguese family as an inherited condition. This disease centres around mutations in transthyretin, a protein involved in binding the hormone thyroxine.[4] Transthyretin is found in a binary complex with retinol binding protein (RBP) and acts as a soluble homotetrameric protein (Figures 11.38 and 11.39).

Each subunit contains 127 residues and a series of eight β strands that adopt a β barrel conformation. Seven of the strands (A–H) are seven to eight residues in length whilst the D strand is shorter and only three residues long (Figure 11.40). The eight strands form two sheets; the first is formed between strands *DAGH* and the second between *CBEF*. Unsurprisingly, the high content of β structure arranged in a barrel like topology contributes to significant stability.

Analysis of the molecular defects occurring in the transthyretin gene in individuals with FAP identified over 80 different mutations. The clinical symptoms of FAP usually begin from the third to fourth decades and are characterized by local neurologic impairment leading to wider autonomic dysfunction and death within 10 years. A common mutation noted in Portuguese, Japanese and Swedish kindreds is the transition Val30>Met. The identification and prevalence of this mutation led to an analysis of the structural and functional properties of Val30Met transthyretin especially since the mutant protein was known to be amyloidogenic.

Residue 30 lies on the inside of the monomer and the bulkier Met side chain displaces one set of four β strands from the remaining set of four. However, the mutant protein was folded with a conformation

[4]Transthyretin was formerly called prealbumin as a consequence of its migration pattern on gels just ahead of albumin.

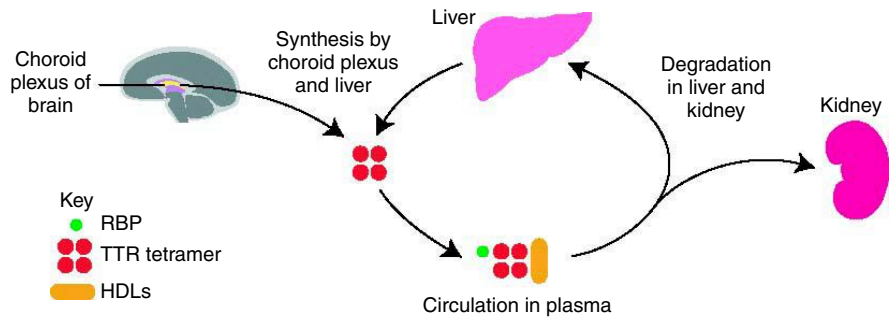

Figure 11.38 Synthesis, distribution and uptake of transthyretin. Transthyretin tetramer circulates in plasma bound to retinol-binding protein (RBP) providing a transport function for vitamin A and thyroxine with a small proportion binding high-density lipoproteins (reproduced with permission from Saraiva, M.J.M. *Expert Rev. Mol. Med.* Cambridge University Press, Cambridge, 2002)

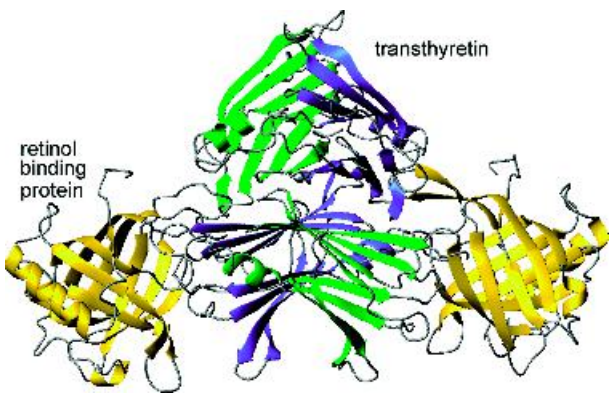

Figure 11.39 The complex of transthyretin and retinol binding protein. RBP is shown in yellow with the tetramer of transthyretin shown in blue and green. The larger RBP binds to a site formed by two of the subunits (PDB: 1RLB)

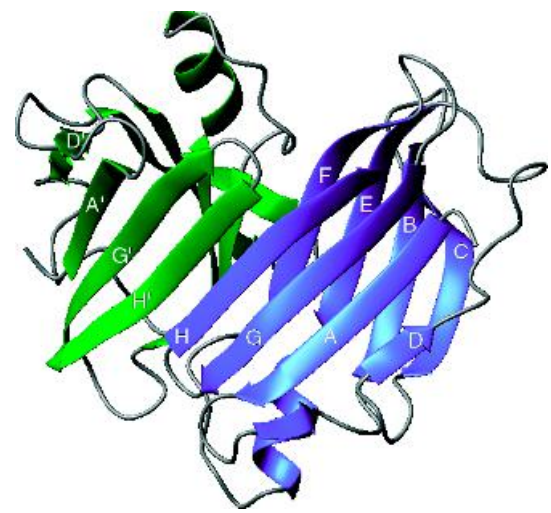

Figure 11.40 Transthyretin structure showing strands A–H and the sandwich formed between DAGH and CBEF in each monomer. Strands H and H′ interact at the interface (PDB: 1BMZ)

very similar to the wild type. This initially puzzling result was interpreted with the recognition that wild type protein aggregates to form fibrils under some conditions–an event associated with senile systemic amyloidosis, where fibril deposition in the heart leads to cardiomyopathy at the age of ~80. The results emphasized that amyloidosis is associated with *all* forms of the protein.

Mutations are viewed as exacerbating disease by promoting more active forms of amyloidosis. Structural studies of wild type protein at low pH (~4.5) show transthyretin dissociation to a monomeric but amyloidogenic intermediate of different tertiary structure involving rearrangement in the vicinity of the C strand–loop–D strand region. Dissociation of the tetramer is an unfavourable event under normal physiological conditions but these rates are

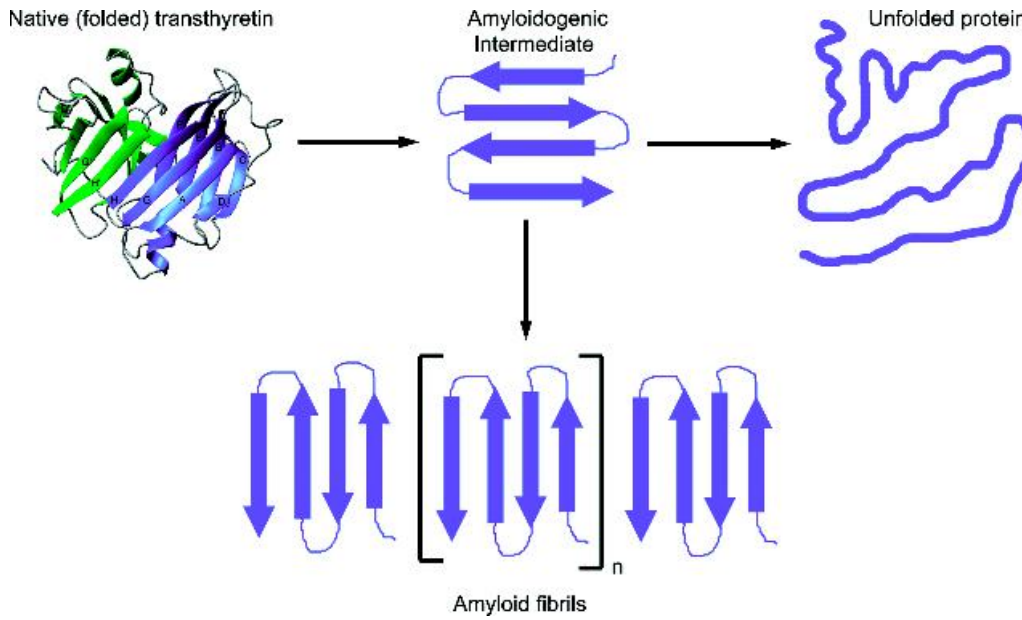

Figure 11.41 Scheme for initial formation of transthyretin-based fibrils

enhanced at low pH and for the Val30Met mutant a lower activation barrier increased the likelihood of amyloidogenic intermediates. Other mutants such as Leu55Pro confirmed this trend and led to the conclusion that mutants associated with FAP are less stable than wild type protein and form the amyloidogenic intermediate more quickly.

Individuals with the FAP mutations are kinetically and thermodynamically predisposed to amyloid fibril formation. Mutation increases the steady-state concentration of the amyloidogenic intermediate with the result that individuals with these defects are prone to amyloid formation and disease. The biophysical data mirrored the pathological outlook of the disease. Onset of the disease in the eighth, third and second decades correlated with the relative stabilities of wild-type, V30M and L55P forms of the transthyretin with the mutations shifting equilibria in favour of amyloidogenic intermediates and decreasing the age of appearance of FAP.

Fibril formation by the association of transthyretin results in comparable aggregates to those observed in other diseases and suggested a similar structure for all amyloid fibrils (Figure 11.41). Alzheimer's disease represents aggregates of a short fragment known as the A_β peptide. Since the A_β peptide and transthyretin do not share any sequence homology and are not related in any way the formation of similar fibril structures was remarkable and significant. The structure of amyloid fibrils examined using X-ray diffraction and cryo-EM methods reveal a β helix structure. In fibrils resulting from six precursor proteins present in different diseases a common structure existed and yet none of these proteins have sequence homology.[5] The morphology and properties of amyloid fibrils were based on a β scaffold with strands running transversely across a helix axis. In the direction of the helix axis the strands align to form β sheets. Four discrete chains are wound into the β helix (Figure 11.42).

[5]The systems studied were the A_β peptide; the λ immunoglobulin light chain (monoclonal protein systemic amyloidosis); lysozyme (hereditary lysozyme amyloidosis); calcitonin (medulary carcinoma of the thyroid); insulin (diabetes and insulin related amyloid disease); β_2 microglobulin (haemodialysis related amyloidosis), together with synthetic peptides of transthyretin (FAP) and the prion protein (transmissible spongiform encephalopathies).

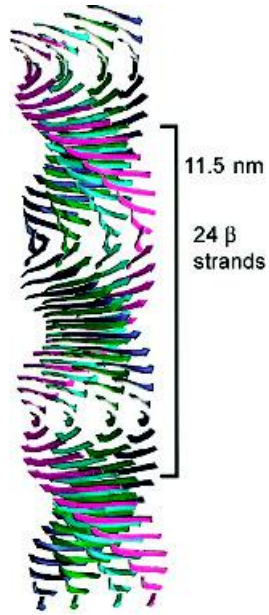

Figure 11.42 Model of a generic amyloid fibril structure. The structure of the amyloid fibrils is made up of four β sheets. Sheets run parallel to the axis of the protofilament with component β strands perpendicular. The twisting of β strands yields a series of parallel strands around a common helical axis that coincides with the axis of the protofilament and leads to a pitch of 11.55 nm containing 24 β strands (reproduced with permission from Sunde, M. *et al.*, *J. Mol. Biol.* 1997, **273**, 729–739. Academic Press)

These structures can be induced in protein not known to form amyloid fibrils in any documented disease phenotype. The SH3 domain (Figure 11.43) of the p85α subunit of phosphatidylinositol-3′-kinase was induced to form amyloid aggregates. The normal structure of the 84 residue domain is based around a β sandwich of five strands but under low pH conditions this domain unfolds in a reaction promoting fibril formation.

Mutating or 'mistreating' soluble proteins induces aggregation and the formation of fibrils. Fibrils derived from unfolded SH3 domains cannot form by simply linking together several units of the native fold because the diameter of the protein is too large to fit properly into the fibril structure. The SH3 domain must unfold to adopt a longer and thinner shape.

The assembly of β strands into a helix is a feature of amyloid fibrils and, although unusual, is known to occur in structural databases, with examples in folded proteins including alkaline protease, pectate lyases and the p22 tailspike protein. Collectively, these studies point to protein misfolding leading to aggregation and amyloid deposits.

As these studies progressed it became apparent that an obscure collection of neurodegenerative conditions resulted in a recognizable event, protein aggregation. These diseases were to become known as the prion-based diseases.

Prions and protein folding

Disease transmission requires the intervention of genetic material (DNA or RNA) in order to establish an infection. Most commonly, this event centres about bacterial or viral infections and the conversion of genetic information into protein with unpleasant consequences to the host. Gradually, convincing evidence has been acquired that a non-nucleic acid based agent is involved in transmission of a restricted number of diseases. These 'agents' were called 'proteinaceous infectious particles' or prions and were identified, after exhaustive analysis, as the active components in communicable and inherited forms of disease.

These ideas met with considerable scepticism and even hostility from some sections of the scientific and medical communities but there is good evidence that prions lie at the heart of neurodegenerative disorders associated with a pathological collection of diseases known as transmissible spongiform encephalopathies (TSEs). The common link between these diseases is development of gross morphological changes to tissue within the brain characterized by the formation of vacuoles and a sponge-like appearance (Figure 11.44). Disease progression is accompanied by the appearance of amyloid plaques within the tissue whilst in some states structural changes occur to cells leading to a loss of synaptic contacts. Three disease states showing these properties are scrapie, Creutzfeldt–Jakob disease (CJD) and Kuru.

Elucidation of the role of prions stems from attempts to understand the biochemical basis of these three obscure, and at first glance, unrelated diseases. Scrapie

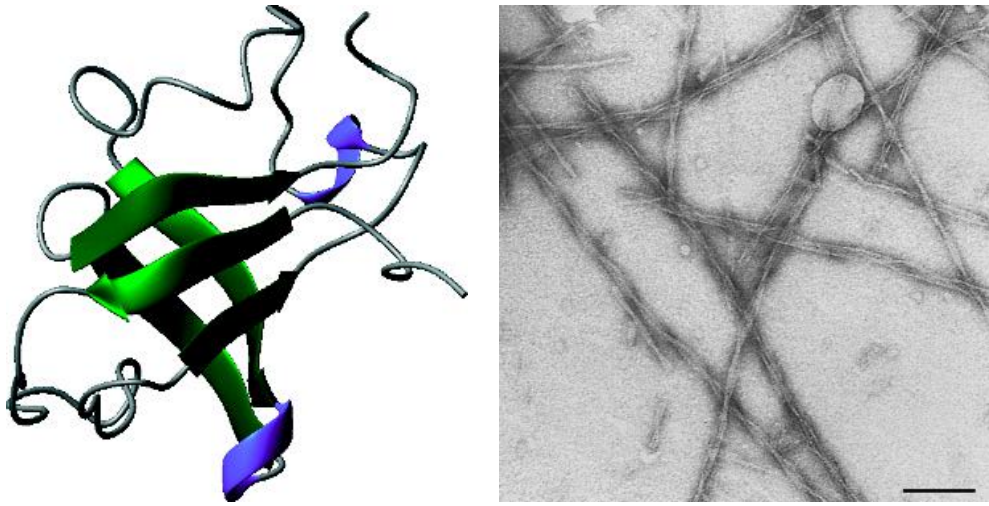

Figure 11.43 The normal fold of SH3 domains showing a sandwich based around five β strands (PDB:1PNJ) and fibrils induced to form via denaturation of SH3 domains at low pH. The solid scale bar is of length 100 nm (reproduced with permission from Guijarro, J.I. *et al*. *Proc. Natl Acad. Sci. USA* 1998, **95**, 4224–4228)

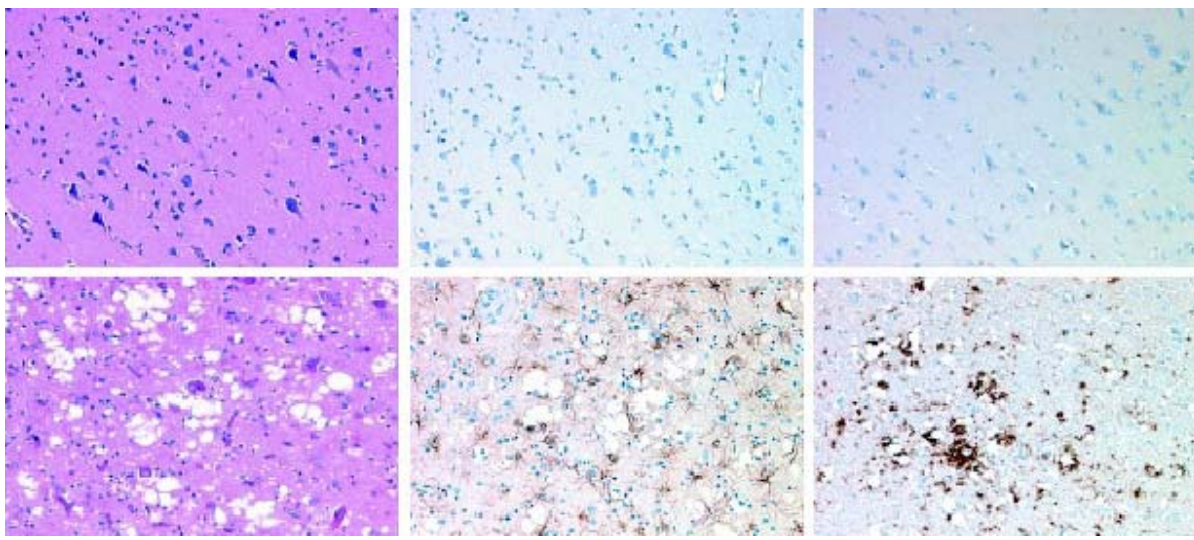

Figure 11.44 Neuropathological features associated with transmissible spongiform encephalopathies. Histological staining and immunohistochemical analysis of frontal cortex samples from a 'control' brain (top row) and of a patient suffering from CJD (lower row). Brain sections were stained with haematoxylin–eosin (left panel), with antibodies against glial fibrillary acidic protein (GFAP, middle) and with antibodies against the prion protein (PrP, right). Neuronal loss and vacuoles are visible in the H-E stain, proliferation of cross reactive astrocytes and prion protein deposits are detectable in the immunostains of the CJD brain samples (reproduced with permission from Aguzzi, A., Montrasio, F., & Kaeser, P.S. *Nature Rev* 2002, **2**, 118–126. Macmillan)

occurs in sheep and is defined by a progressive loss of motor coordination that leads to an inability to stand unsupported. Animals develop an intense itch that leads to wool scraping from skin and gives the disease its name. Scrapie has been recognized since the 17th century in the United Kingdom, with similar conditions noted more recently in animals such as mink, deer and elk.

CJD was first identified in the 1920s by H.G Creutzfeldt and then independently by A. Jakob. Patients exhibit dementia with poor motor coordination, poor perception and reasoning. The incidence of CJD is low, probably less than 1 case in two million, and typically occurs in individuals above 50 years of age. It is found throughout the world with little obvious association with economic, racial, or social patterns. Most frequently the disease arises sporadically, but in 10–15 percent of cases the disease is inherited in an autosomal dominant manner, and very rarely CJD is spread by iatrogenic transmission where contamination occurs most probably via donated tissue.

Kuru is an unusual disease confined to tribal people of the highlands of Papua, New Guinea where it is called the 'laughing death'. It was comprehensively described in 1957 as a fatal disease marked by ataxia and progressive dementia. A link was made between disease transmission and ritual cannibalism where tribes people honoured their dead by eating body parts including brain. A decline in cannibalism led to a decline in Kuru although the precise causative agent remained at that time unknown.

The three diseases are linked by the observation that transmission occurs when diseased brain extracts are injected into healthy animals. Of these diseases scrapie is experimentally accessible and initial studies attempted to identify the causative agent. Exhaustive analysis of scrapie infected brain tissue purified a single protein devoid of nuclei acid that was infectious when injected into hamster brains. It was called the prion protein (PrP) and the pioneering work in this area was attributable to the work of Stanley Prusiner and his extraordinary quest to define and characterize these agents.

Prion purification and the demonstration of infectivity raised a number of questions. How was PrP encoded? Was undetected DNA lurking in the 'pure' protein preparation? What is the origin of PrP? Major insight into the molecular events underpinning the generation of PrP arose with N terminal sequencing of the isolated protein. This sequence allowed the synthesis of degenerate oligonucleotide probes that hybridized to chromosomes found in a wide range of mammals including mice, hamsters and most importantly humans. The probes did not hybridize to the isolated PrP eliminating the possibility of trace amounts of DNA associated with this fraction. A gene encoding PrP was found on the short arm of chromosome 20 in humans where it coded for a glycoprotein of mass 33–35 kDa.

Most individuals never develop any form of the disease yet the prion gene is found in the human genome. Prusiner suggested one explanation. PrP was produced in two forms – a normal form and an abnormal one that generated disease. Immediate support for this view came with the demonstration that PrP found in infected brains was resistant to proteolysis whilst the normal form remained sensitive to proteases. The two forms are often called cellular PrP (normal) and scrapie PrP (infectious) – abbreviated as PrP^c and

Table 11.6 Structural differences between PrP^C and PrP^{Sc}

Property	PrP^c (cellular)	PrP^{Sc} (scrapie)
Polypeptide chain	1–231	1–231
Protease resistant	No	Stable core residues 90–231
Disulfide bridges	Yes (179–214)	Yes
Solubility	Soluble when expressed -GPI	Very insoluble except to all but the strongest 'solvents'
Aggregation state	Monomeric	Multimeric
Stability	25–40 kJ mol^{-1}	More stable than PrP^c
Structure	α helical	Increase in proportion of β strands

Adapted from Cohen, F. & Prusiner, S.B. *Annu. Rev. Biochem.* 1998, **67**, 793–819.

PrP^{Sc} (Table 11.6). The different proteolytic sensitivity of PrP^{c} and PrP^{Sc} suggested that each form of the protein possessed a different conformation.

Protein expression studies from the Syrian hamster, mice and humans showed monomeric species with little tendency to form aggregates whilst the full-length *infectious* conformer readily formed aggregates. PrP is composed of ~250 residues with a signal sequence of 22 residues, and close to the C-terminal residue is a GPI membrane anchor. Additional post-translational modifications include two N-linked glycosylation sites and a single disulfide bridge. In the absence of the GPI anchor the protein is soluble and does not partition in the membranes; this occurs when the protein is expressed in prokaryotic systems. As a result of its solubility it has been possible to subject the protein to detailed structural analysis.

The solution structure of a fragment containing residues 121–231 for the mouse PrP^{c} protein showed a folded domain with three α helices and two short β-strands (Figure 11.45). Structures of the analogous human protein (residues 23–231) and longer fragments from the mouse and Syrian hamster indicated that the prion protein consisted of a structured domain from residues 121–231 and a less ordered N-terminal domain. Very little secondary structure was detected in the N-domain. One reason for the absence of structure in the N-terminal region is seen in the primary sequence of the prion protein (Figure 11.46) where a series of five octapeptide repeats occur with the sequence PHGGGWGQ. The presence of this repeat is an unusual feature of the sequence.

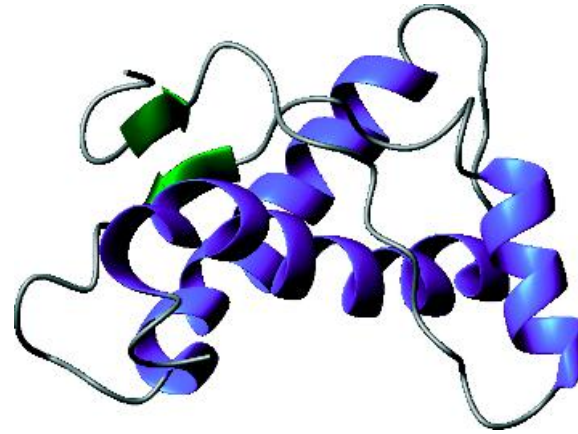

Figure 11.45 The structure of 121–231 fragment of the mouse PrPc determined using NMR spectroscopy (PDB: 1AG2)

Prions become infectious particles capable of transmitting disease (scrapie) by corrupting the conformation of other protein molecules and initiating a series of events that lead to the formation of protein fibrils or amyloid deposits within tissues.

Two additional TSEs, Gerstmann–Straussler–Scheinker (GSS) disease and fatal familial insomonia (FFI), also cause spongiform changes in the brain, neuronal loss, astrocytosis and amyloid plaque formation. GSS disease is an inherited neurodegenerative condition that offered the possibility of identifying mutations. Five different TSEs originate from different mutation sites within the prion gene (Table 11.7).

```
                10        20        30        40        50        60        70        80
PrPc : MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGGGWGQPHGGGWGQPHGGGWGQPHGGG : 80

                90       100       110       120       130       140       150       160
PrPc : WGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGAVVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQ : 160

               170       180       190       200       210       220       230       240
PrPc : VYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSSPPV : 240

               250
PrPc : ILLISFLIFLIVG : 253
```

Figure 11.46 Primary sequence of the prion protein (PrP^{c}). The N-terminal signal sequence is highlighted by the green block whilst the blue block represents the structured domain of 5 helices and three strands. Shown in yellow is the unstructured region and within this sequence are five octapeptide repeats highlighted by red text

Table 11.7 Mutations of the *PNRP* gene associated with inherited forms of transmissible spongiform encephalopathy

Mutation	Disease plus phenotype
Insertion of 24, 48, 96, 120, 144, 168, 192, or 216 base pairs between codons 51 and 91 of PNRP gene leads to repeat of octapeptide motif in PrP	CJD, GSS, or atypical dementias
Pro102Leu, Pro105Leu, Ala117Val, Gly131Val, His187Arg, Asp202Asn, Phe198Ser, Gln212Pro, Glu217Arg	GSS: classical ataxic form plus other phenotypes
Tyr145stop, Thr183Ala	Alzheimer-like dementia
Asp178Asn	CJD (where residue 129 is V on mutant allele)
Asn178Asn	FFI (where residue 129 is M on mutant allele)
Val180Ile, Val203Ile, Arg208His, Val210Ile, Glu211Gln, Glu200Lys, Met232Arg	CJD

FFI = fatal familial insomnia; CJD = Creutzfeldt–Jakob disease; GSS = Gerstmann–Straussler–Scheinker syndrome.

Most mutations (Figure 11.47) occur in the structured C-domain of PrP and GSS is caused primarily by the substitution of Pro102>Leu, whilst in CJD the prevalent mutation is at residue 178 and involves the replacement of Asp by Asn. A slight complication to this genetic pattern is the existence of polymorphisms within the *PNRP* gene where Met or Val is found at residue 129. By itself this change does not result in disease but when combined with mutation at residue 178 the polymorphism contributes to either FFI (Met) or CJD (Val).

The identification of all diseases as the pathological consequence of mutations within the *PNRP* gene has unified this area of study but still leaves the awkward question of how does the prion cause a change in conformation and spread the disease? A major problem in this area has been the unambiguous demonstration of conversion of PrP^{c} into PrP^{sc}.

Strong evidence that conversion of prion protein from normal to abnormal states was responsible for neurodegenerative disease was obtained using 'knockout' mice where the *PNRP* gene was deleted. With the normal *PNRP* gene present the injection of mice with PrP^{Sc} resulted in the transmission of the disease; mice showed familiar symptoms of ataxia and cerebellar lesions within approximately 150 days. When knockout mice were subjected to the same experiment they did not develop the disease. In other words the disease requires native PrP protein encoded by the relevant gene to propagate into the development of amyloid fibrils. Two hypotheses known as the 'refolding' and 'seeding' models attempt to explain prion propagation (Figure 11.48).

Summary

The folding of individual polypeptide chains from less structured states to highly organized topologies is vital to biological function. *All* of the information directing protein folding resides within the primary sequence.

Thermodynamically the folded state is the global energy minimum with the free energy decreasing in the transition from unfolded to native protein. Native proteins are marginally more stable than unfolded states with estimates of conformational stability ranging from 10–70 $kJ\,mol^{-1}$.

A number of quantifiable factors have been shown to influence protein stability in globular proteins. These factors include conformational entropy, enthalpic contributions from non-covalent interactions and the hydrophobic interaction. When the sum of all favourable interactions outweighs the sum of unfavourable interactions a protein is stable.

Protein denaturation is the loss of ordered structure and occurs in response to elevated temperature,

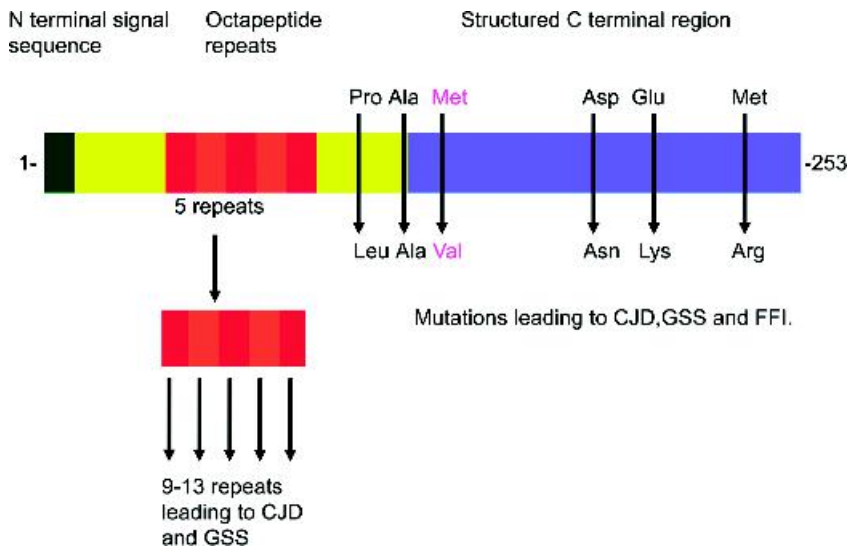

Figure 11.47 Schematic representing mutations and insertions in prion protein. Amplification of the number of octapeptide repeats leads to CJD or GSS

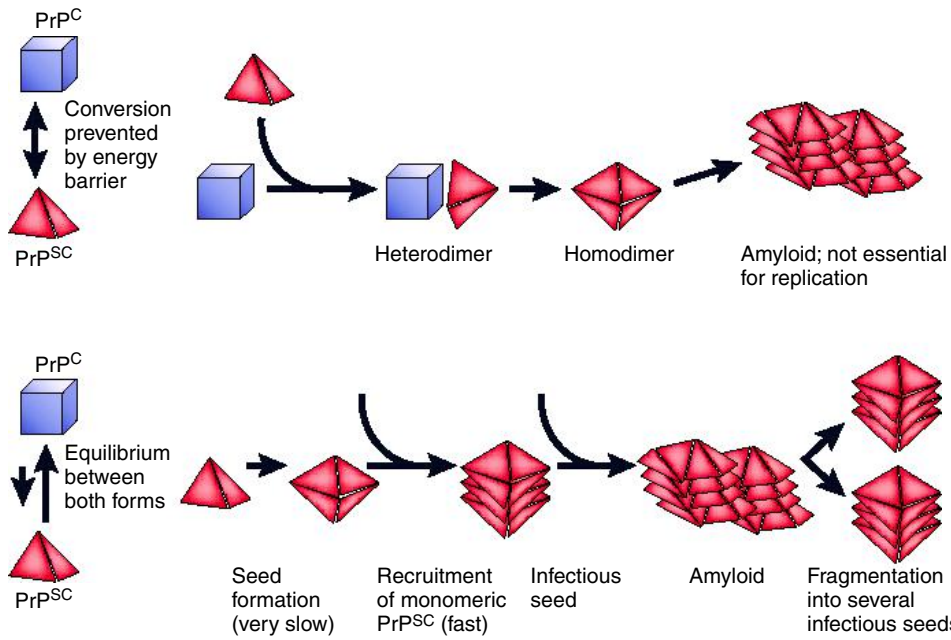

Figure 11.48 The 'protein-only' hypothesis and two popular models for the conformational conversion of PrP^{C} into PrP^{Sc} (reproduced with permission from Aguzzi, A., Montrasio, F., & Kaeser, P.S. *Nat. Rev.* 2002, **2**, 118–126. Macmillan)

extremes of pH or the addition of reagents such as urea or guanidine hydrochloride. Denaturation involves the disruption of interactions such as hydrophobic interactions, salt bridges and hydrogen bonds that normally stabilize tertiary structure.

Experimental measurements of protein stability include differential scanning calorimetry, absorbance and fluorescence optical methods, NMR spectroscopy and circular dichroism.

Kinetic methods allow protein folding to be followed as a function of time. Small soluble proteins (<100 residues) fold within 1 s and the observation of slower folding rates is normally associated with disulfide bond formation or cis-trans proline isomerization. To overcome slow *in vivo* folding specific enzymes catalyse reactions such as peptide prolyl isomerization and protein disulfide isomerization.

Critical events in the formation of the native fold are the acquisition of ordered stable secondary structure, the formation of hydrophobic cores, and the exclusion of water from the protein interior.

For proteins such as lysozyme, barnase and chymotrypsin inhibitor 2 the reaction pathway has been defined to identify properties of intermediates and transition states. In barnase it has proved possible to quantify the folding process in a pathway containing at least one intermediate between denatured and ordered states. A hydrophobic core and the formation of the β sheet are two early events in the folding pathway. In lysozyme each domain folds separately within the observation of multiple kinetic pathways for each folding unit.

Chaperones are multimeric systems found in all cells that prevent incorrect protein folding based around a toroidal structure of seven, eight or nine subunits. The toroidal structure is capped at either end and forms a cavity binding unfolded polypeptide via hydrophobic surfaces with ATP-driven conformational changes propelling the peptide towards the native state.

The kinetics of membrane protein folding *in vitro* are much slower than rates of folding for globular domains but the principles governing formation of the native state remain similar.

Incorrect folding leads to a loss of activity that is the basis of many disease states. Misfolding leads to incorrect protein trafficking within the cell and mutant proteins of *CFTR* are a good example.

Incorrect protein folding is also a key event in amyloidosis – the accumulation of long irreversibly aggregated protein within fibrils. Amyloidogenic diseases include many neurodegenerative disorders such as Alzheimer's, transmissible spongiform encephalopathies and hereditary forms of systemic amyloidosis. Fibrils from a range of amyloid proteins have a common structure based on collections of β strands that align into parallel sheets twisted into helical conformations known as β helices.

Neurodegenerative disorders such as CJD, BSE and scrapie are linked via the spongiform appearance of neuronal tissue and the accumulation of amyloid deposits. The disease arises from changes in the prion protein generating conformations with different secondary and tertiary structure. The abnormal form promotes amyloidosis by inducing other prion proteins to change conformation. These events occur spontaneously at very low frequency leading to sporadic occurrences of disease but are facilitated by mutations within the *PNRP* gene.

Problems

1. Draw the *cis* and *trans* conformations associated with a Xaa–Pro peptide bond.
2. The 3_{10} helix is named because there are three residues per turn and 10 atoms including hydrogen between the donor and acceptor atoms forming an intra-chain hydrogen bond. Describe the α helix using such notation. How does the helix compare with the 3_{10} helix? Repeat this analysis for the π helix.
3. Experimental studies of protein folding reveal the following trends of denaturation.
   Calculate the conformational stability and estimate any other relevant parameters.

Absorbance	Denaturant concentration (M)	Absorbance	Denaturant concentration (M)
0.985	0	0.375	1.8
0.985	0.1	0.325	1.9
0.985	0.2	0.285	2.0
0.980	0.3	0.220	2.1
0.975	0.4	0.175	2.2
0.970	0.5	0.140	2.4
0.960	0.6	0.120	2.5
0.945	0.8	0.115	2.6
0.890	1.0	0.110	2.7
0.760	1.2	0.105	2.8
0.625	1.4	0.105	2.9
0.510	1.6	0.100	3.0

4. Assume that in order to reach the native state a protein needs to sample only 10 conformations per residue with each conformation taking 0.1 ns. Estimate how long it might take for a 100 residue protein to fold. Now assume that the each block of 10 residues can fold independently of the remaining residues and by only sampling three conformations per residue. How long does folding take to occur? Comment on the two values?
5. Explain how disulfide bonds affect protein stability. What would you expect to be the effect on protein stability of introducing a disulfide bond using mutagenesis? What would be the effect of removing a disulfide bond on protein stability?
6. What factors contribute to the observation of heterogeneous folding kinetics of proteins? How would you attempt to evaluate the importance of these factors.
7. 'You cannot unscramble an egg'. Discuss this statement in the light of protein denaturation, refolding, modification and the known presence of chaperones.
8. Explain what might be the consequence of deleting the octapeptide repeats of PrP on prion infection in for example a host such as the mouse.