PROTEIN FOLDING – Dr. Alfonso De Simone

The majority of the protein molecules have been designed throughout evolution to spontaneously adopt a well-defined 3D structure after biological synthesis *via* the ribosome. The process of protein folding is mediated by a complex **energy landscape** that direct the unfolded state conformations into the encoded native structure.

Not all proteins are folded in the native state; some proteins are intrinsically disordered.
How to convert sequence into a structure? Native state is the most stable conformation in terms of thermodynamics.

The Protein Folding Problem

- What guides a linear polypeptide into a 3D structure?
- This code is more complex of the DNA code
- This knowledge is expected to largely impact our ability to understand and use biology

Some expected outcomes:

- Predict 3D structure from primary sequence
- Understand and combat misfolding related to human diseases
- Design proteins with novel functions

Anfinsen Experiment

Denaturation of ribunuclease A (4 disulfide bonds) with 8 M Urea and b-mercaptoethanol to totally unfold the protein in a random coil state having no activity.

- Remove b-mercaptoethanol only oxidation of the sulfhydryl group, then remove urea → scrambled protein, no activity
- Further addition of trace amounts of b-mercaptoethanol converts the scrambled form into native pattern.
- Conclusion: The native form of a protein corresponds to the thermodynamically most stable conformation.

Protein Stability

Structured proteins do fold into a native state characterised by a well-defined 3D structure because this state is **more stable** *(lower free energy)* than the unfolded state.
Protein stability normally refers to the **conformational** stability of the native state which is given by the difference in free energy between native and unfolded states DG_{NU}.

Chemical stability: what makes configuration of polypeptide chain stable, everything that makes bonds stable. Processes of chemical instability in proteins include:

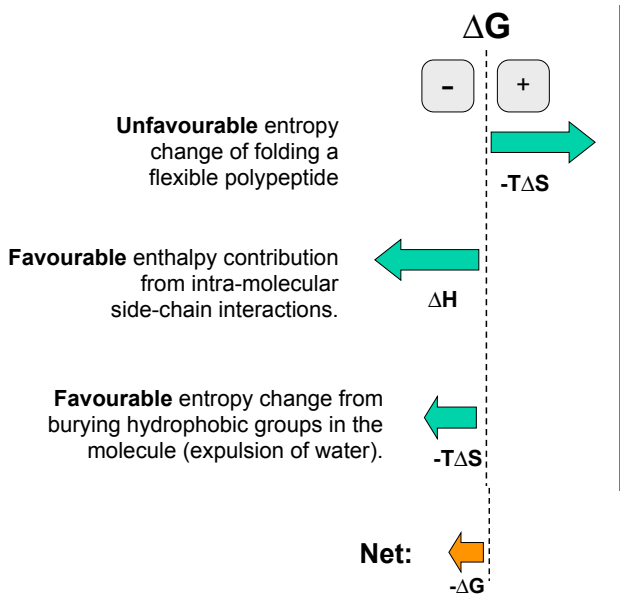
- deamination of asparagine and/or glutamine residues,
- hydrolysis of the peptide bond of Asp residues at low pH
- oxidation of Met at high temperature
- elimination of disulfide bonds
- disulfide interchange at neutral pH
- Other processes include thiol-catalyzed disulfide interchange and oxidation of cysteine residues

Conformational stability: Refers primarily to the ability of adopting a well-defined conformation rather than a random coil (unfolded) state. An important feature of a conformational state of a protein is the local structure adopted by backbone atoms, which can effectively been described in terms of phi and psi angles.

- Phi (Φ) and Psi (ψ) rotate, allowing the polypeptide to assume its various conformations
- some conformations are disallowed
- glycine has the greatest accessibility

Determinants of Protein Folding

There is an unfavourable entropy change of folding a flexible polypeptide. Favourable enthalpy contribution from intra-molecular side-chain interaction. Favourable entropy changes from burying hydrophobic groups in the molecule (expulsion of water). You want the protein not too rigid (delta g too strong); allows some flexibility. That is why the favourable change in free energy is small.
General rule is that 20% of AA identity there may be the same fold.



Parameters affecting protein folding (non covalent interactions):

interaction	nature	bond length	"bond" strength	example
ionic (salt bridge)	electrostatic	1.8-4.0 **Å** (3.0-10 **Å** for like charges)	1-6 kcal/mol	positive: K, R, H, N-terminus negative: D, E, C-terminus
hydrophobic	entropy	-	2-3	hydrophobic side chains (M,I,L,V,F,W,Y,A,C,P)
H-bond	H-bonding	2.6-3.5	2-10	H donor, O acceptor
van der Waals	attraction/ repulsion	2.8-4.0	<1	closely-spaced atoms; if too close, repulsion
aromatic-aromatic	p-p	4.5-7.0	1-2	F,W,Y (stacked)
aromatic-amino group	H-bonding	2.9-3.6	2.7-4.9	N-H donor to F,W,Y

Covalent interactions:

• disulfide bond formation is a covalent modification
• the oxidation process can either be intramolecular (within the same protein) or inter-molecular (within different proteins, *e.g.*, antibody light and heavy chains).
•The process is reversible.
•cellular enzymes (protein disulfide isomerases) assist many proteins in forming proper disulfide bond(s)

Compaction:
Proteins are compact

- Density comparable to organic crystals
- Helices and sheets esp. compact
- But packing also affects overall fold

Folding is largely directed by internal residues

- Modification of Lys residues in RNase A by poly-Ala did not interfere with folding
- Hydrophobic effect is important
- Surface mutations are accommodated without affecting fold (e.g. virus capsids)

Why is that ?

Hierarchy:
Protein structures (and folding) are hierarchical

- Large proteins have domains
- Domains have sub-domains
- Sub-domains fold independently
- Tertiary structure formed when these pack together

Adaptability
Protein structures are adaptable

- Packing of apolar side-chains in the core is unique
- But mutagenesis of T4 lysozyme reveals great adaptability
- Mutation can be accommodated with local shifts in packing without affecting fold or backbone

Is this always true?

Sequence versatility:
Conservation of sequence and folding

- 20% amino acid sequence identity between 2 proteins usually means they have the same overall fold
- 80% sequence difference can mean same fold
- However, mutation of ca. 50%of residues in GB1 completely changed the fold

Techniques for Measuring Protein Stability
Need to have a technique that gives you two different signal for folded and unfolded proteins.
Any methods that can distinguish between U and F

-*Absorbance* (e.g. Trp, Tyr, chromophoric probes)
-Fluorescence (Trp or Fluorophoric probes) difference in emission max & intensity
-CD
-NMR
-DSC (calorimetry)
-Urea gradient gels - difference in the migrating rates between F and U.
-Catalytic activity

Absorbance is not a very well sensitive method.
Circular Dichroism, absorbance level, alpha helices, beta sheets and random coils, look at the differences. Circularly polarised light is a chiral entity in a sense. Right-handed polarised light will interact differently with a protein than left-handed protein.

Protein Denaturation

- Loss of native structure integrity with accompanying loss of activity is called ***denaturation***
- Proteins can be denatured by
  - heat or cold; pH extremes; organic solvents
  - chaotropic agents: urea and guanidinium hydrochloride

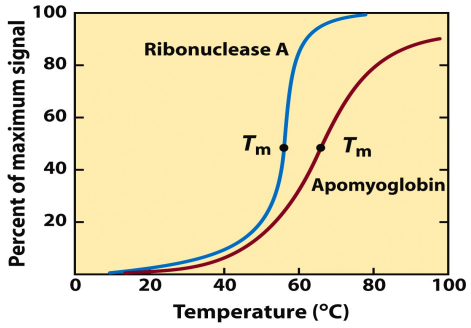


Figure 4-25a
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

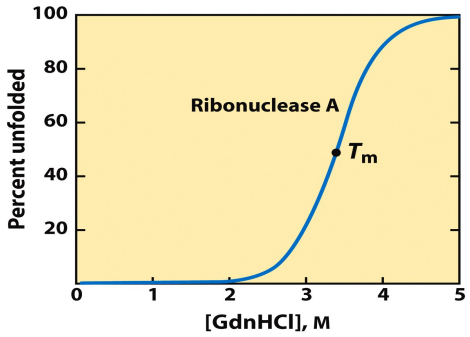


Figure 4-25b
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

Why denaturants such as urea and GdmCl cause proteins to denature may be considered empirically. Those denaturants solubilise all the constituent parts of a protein, from its polypeptide backbone to its hydrophobic side chains. To a first approximation, the free energy of transfer of the side chains and polypeptide backbone from water to solution of denaturant is linearly proportion to the concentration of denaturant. Because the denatured state is more exposed to solvent than the native state, the denatured state is preferentially stabilized by denaturant.

Most proteins denature when the pH is changed to very low or very high values. The primary reason for extremes of pH causing denaturation is that proteins usually have buried groups that have highly perturbed pKa's. A further reason for the change in stability is that at low pH, the protein becomes highly positively charged as the aspartates and glutamates become protonated; and at high pH, the protein becomes negatively charged as the lysine, tyrosine and eventually arginine residues become deprotonated. Electrostatic repulsion then causes destabilization.
Protein denaturation can be a simple two-state equilibrium of D→← N (cooperative all-or-none transition) or a multistate transition with the production of intermediates.

Circular Dichroism

- CD measures the molar absorption difference Δε of left- and right- circularly polarized light: $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$
- Chromophores in the chiral environment produce characteristic signals
- CD signals from peptide bonds depend on the chain conformation

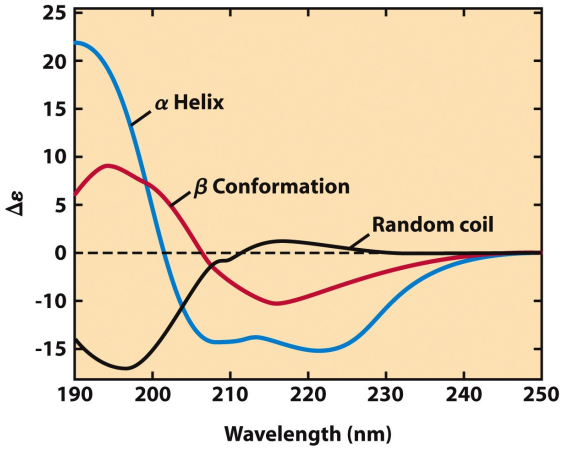
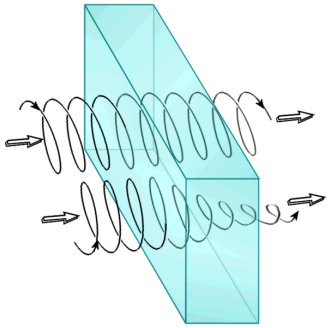


Figure 4-9
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

- CD arises from the differential absorption of left and right circularly polarized light by optically active media:
    1. Definition: $\Delta\varepsilon = \Delta\varepsilon_L - \Delta\varepsilon_R$ (where ε= molar absorbance)
    2. Also reported as ellipticity, Θ
- Chiral molecules (e.g. L-amino acids) are optically active (they have different refractive indices and different extinction coefficients for L and R circularly polarised light)
- Optical activity also refracted by environment (location in secondary structure elements, which are also chiral)



$$\varepsilon_L > \varepsilon_R$$

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R > 0$$

Pros:

- CD spectra recorded in the far-UV (190-240 nm)
- Characteristic spectra obtained for different secondary structures
- A particular 3D fold has a CD signatures (allow monitoring of folding/unfolding)

Cons:

- Lacks resolution, no insight at the amino acid level
- Often over-interpreted

Renaturation/Refolding is generally found to proceeds by a series of exponential phases. Many of these exponentials are a consequence of cis-trans isomerisation about peptidyl-prolyl bonds (Peptidylprolyl Isomerase)

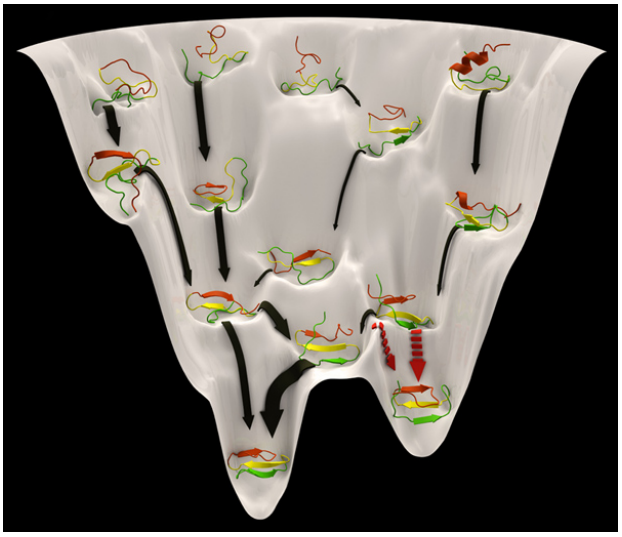
Protein Folding 2-3

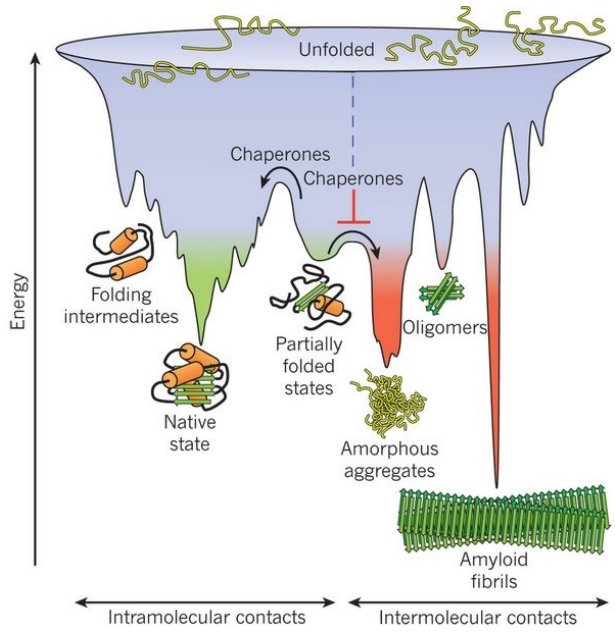
The Levinthal Paradox

- There are vastly too many different possible conformations for a protein to fold by a random search.
- Let's focus on just the backbone local structure in which we restrict to only 3 conformations per amino acid in the unfolded state, For a 100 a.a. protein there will be 3^{100} conformations!!
- If the chain samples 10^{12} conformations/sec, it will take 5 x 10^{35} sec (2 x 10^{28} year)
- Conclusion: Protein folding is not random, must have pathways.

What makes a protein fold so fast?

- Under physiological conditions proteins fold to the lowest-energy fold in the microsecond to second time scales. How can they find the right fold so fast?
- It is mathematically impossible for protein folding to occur by randomly trying every conformation until the lowest energy one is found (Levinthal's paradox).
- Search for the minimum is not random because the direction toward the native structure is funnelled.
- The energy landscape can have local minima in which partially folded proteins can be trapped.





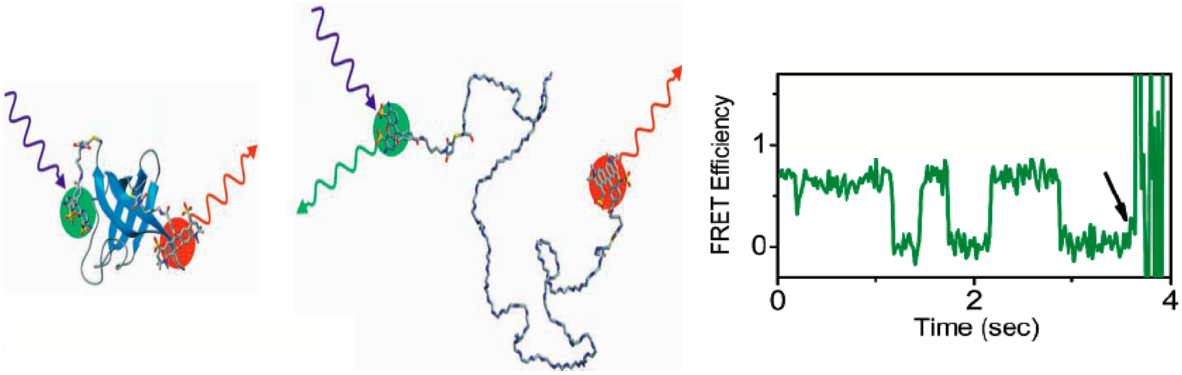
If a golfer was blindfolded, stood at the edge of a very large golf course, and was told to hit the golfball ina random direction, then the chances of sinking the ball in a single hole are infinitesimally small. This is analogous to a protein existing in a large number of non-native states of equal energy and having just one native state state of much lower energy. There is no driving force to push the protein in the direction of folding; this is the Levinthal paradox. Suppose, instead, that the golf course sloped down from all directions to the hole. Then gravity would funnel the bal to the hole, and the golfer would always score a hole-in-one. Theoreticians have likened the process of a protein overcoming the Levinthal paradox to a progression down a funnel.

At the top of the funnel, the protein exists in a large number of random states that have relatively high enthalpy and high entropy. There is a fight between the maximization of entropy keeping the protein as random as possible at the top of the funnel and the minimization of enthalpy trying to drag the protein down the funnel as it folds. Progress down the funnel is accompanied by an increase in native-like

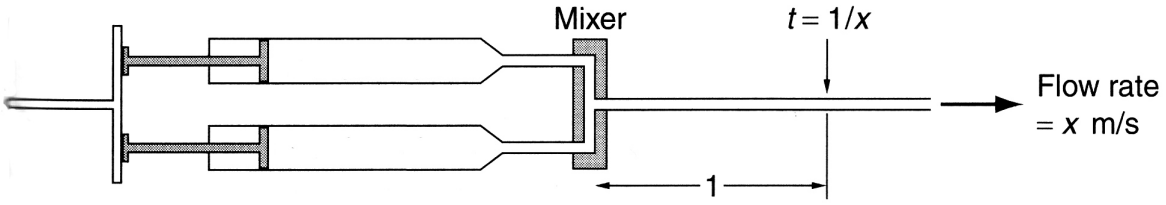
structure as folding proceeds. Proteins fold from the random state by collapsing and reconfiguring. The reconfiguring occurs by a Brownian type of motion between conformations that are geometrically similar and follows a general drift from higher energy to lower energy conformations. The funnel is thus a progressive collection of geometrically similar collapsed structures, one of which is more thermodynamically favourable than the rest.

2-state folding: Single Molecules

- The simplest "two-state model" for protein folding implies no stable intermediate species between the native and the unfolded state
- Two state folding can be studied in single-molecule experiments.
- In this FRET experiment, a green donor and a red acceptor dye are attached to the chain ends of a two-state protein. In the folded state, the transfer efficiency between the dyes is high because the chain ends are close. In the unfolded state, the separation between the dyes is larger, and the transfer efficiency is small.
- Folding and unfolding events then correspond to jumps in FRET efficiency



- Two-state folding is typically investigated in ensemble experiments. In this continuous flow rapid mixing experiment, the protein solution in the upper syringe contains a high concentration of denaturant, which stabilizes the unfolded state. In the mixer, the denaturant is diluted with water (from the lower syringe), and the proteins can start to fold. (Because of the continuous flow,) Time points of the folding process then correspond to distances from the mixer.
- A two-state folding process exhibits a characteristic single-exponential relaxation (into the new equilibrium after mixing). The relaxation rate corresponds to the folding rate.
- Rapid mixing to initiate folding:



  - Single-exponential relaxation for 2-state process:

Transition State Theory

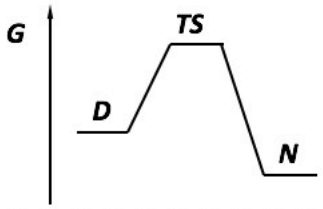
- Two-state folding is usually analyzed using the transition-state theory.
- The transition state TS is crucial for understanding the two-state folding kinetics.
- TS is extremely short-lived and cannot be observed directly.
- TS is located between D (denatured) and N (native).
- The folding rate is proportional to the exponential of the (negative) activation free energy, the free energy difference between the transition state TS and the denatured state D.

- The two-state folding kinetics is usually explored via mutational analysis.

- Experimental method to study of the structure of the transition state
- Using mutations as structural probes for TS
- Information at an aminoacidic resolution
- $\Phi = 1$, transition state has native like structure
- $\Phi = 0$, transition state has denatured like structure

- **Transition state** theory:

$$k \propto \exp(-\Delta G_{TS-D})$$



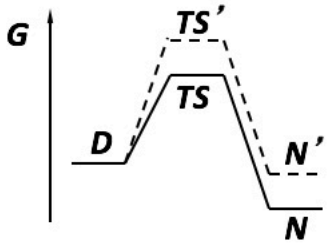
- **Mutations change** the folding **rate** k and **stability** ΔG_{N-D}

- Central quantities: **Φ-values**

$$\Phi \equiv \frac{\Delta\Delta G_{TS-D}}{\Delta\Delta G_{N-D}}$$

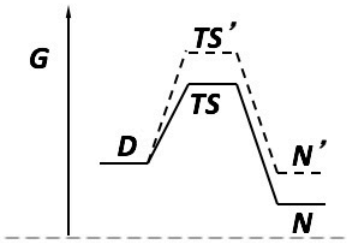
$\Delta\Delta G_{T-D} = \Delta G_{T-D}(mutant) - \Delta G_{T-D}(WT)$

$\Delta\Delta G_{N-D} = \Delta G_{N-D}(mutant) - \Delta G_{N-D}(WT)$

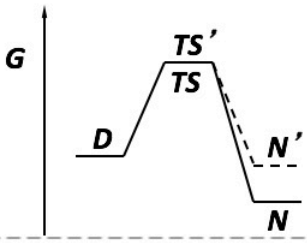


(The simplest model for a two-state process is transition-state theory.) Two-state folding is usually analyzed in transition-state theory. (For a two-state folding process,) The transition state is located between the denatured state D and the native (folded) state N of the protein. The folding rate is proportional to the exponential of the (negative) activation free energy, the free energy difference between the transition state T and the denatured state D.

- The transition state T is central for understanding the two-state folding kinetics. However, T is extremely short-lived and cannot be observed directly. Instead, the two-state folding kinetics is usually explored via mutational analysis.
- The effect of a mutation on the folding kinetics is typically quantified by a Phi-value. A Phi-value is defined as the mutation-induced change in activation free energy, divided by the change in equilibrium free energy.
- The central question: What do Phi-values tell us about the transition state?



Φ= 1: mutated residue is native-like structured in TS



Φ = 0: mutated residue is non-native-like and/or unstructured in TS

Use of Φ for structural refinement of TSE (transition state ensemble)

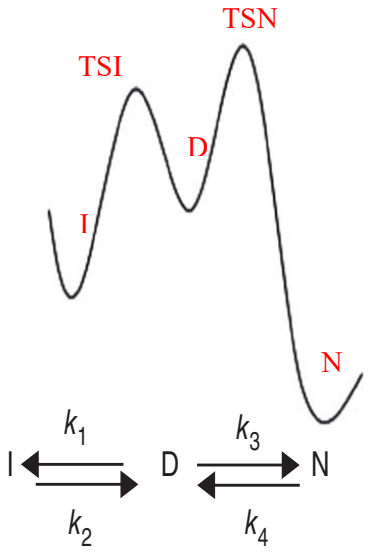
Structural characterization of a misfolded intermediate populated during the folding process of a PDZ domain, De Simone et al.

PDZ domains are small (90–100 amino acid) domains that adopt a six-stranded β-sandwich fold flanked by two α-helices, with the typical linear arrangement of the secondary structure elements being ββββαββαβ

D1pPDZ features two competing folding pathways

The PDZ domain of the D1 C-terminal processing protease (D1p) of the green alga *Scenedesmus obliquus*—involves a kinetic competition between folding to the native state and misfolding to a metastable off-pathway intermediate. J. Biol. Chem, 2008, 283:8954-60

To map the structural features of both the productive and nonproductive folding pathways of this small protein domain, we carried out a Φ-value analysis, whose results were then incorporated as structural restraints in molecular dynamics simulations of D1pPDZ.
The structural refinement showed that the misfolded intermediate is characterized by an alternative packing of the N-terminal b-hairpin onto an otherwise native-like scaffold.



TSI = Transition state D -> I
TSN= Transition state D -> N

Protein Folding in vivo

Many factors interplay with protein folding in vivo

- Crowding
- Protein aggregation
- Cellular components
- Compartments

Quality control mechanisms to assist protein folding *in vivo*

Proteins that (mostly) use ATP to disrupt misfolded proteins and set them on the folding path again
Variety of types:

- HSP70 (prok and euk.; e.g. DnaK in *E. coli*)
  - Reverses denaturation/aggregation
  - Works with HSP40 (DnaJ in *E. coli*)
- Trigger factor (ATP not needed), ribosome associated, acts ealy
- Chaperonins
  - Large, multi-subunit, cage-like
  - Type I: bacteria, mitochorndria, chloroplasts (e.g. GroEL/GroES)
  - Type II: archaea/eukaryotes

- HSP90 (euk)
  - Facilitates late-stage folding of signalling proteins
  - Unique regulatory role – induces active conformation

- Nucleoplasmins
  - Decameric acidic nuclear assemble nucleosomes

- Protein Disulfide Isomerase (PDI)

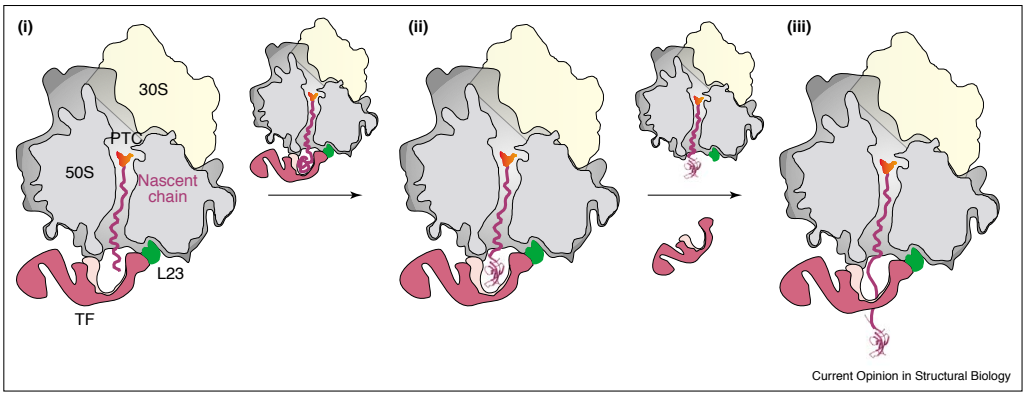
- Peptyl Prolyl Isomerase (PPI)

Trigger Factors

- Newly synthesized proteins leave the ribosome through a narrow tunnel in the large subunit (peptidyl transferase centre, **PTC**). During ongoing synthesis, nascent protein chains are particularly sensitive to aggregation and degradation because they emerge from the ribosome in an unfolded state.

- co-translational protein folding can occur in the cytosol of both prokaryotic and eukaryotic cells

- Trigger Factor is an ATP-independent chaperone and displays chaperone and peptidyl-prolyl-cis-trans-isomerase (PPIase) activities in vitro.



GroEL/GroES (Chaperonine Type I)

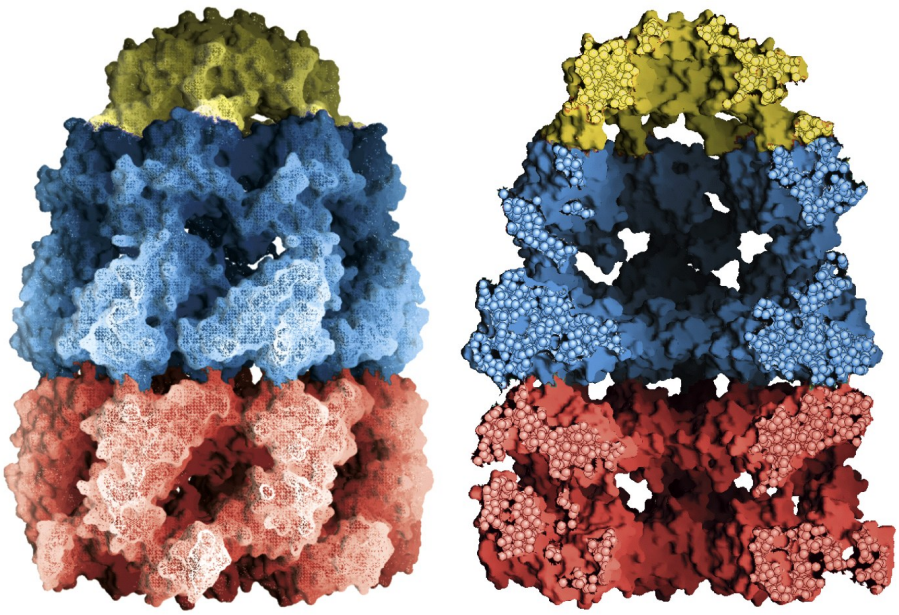
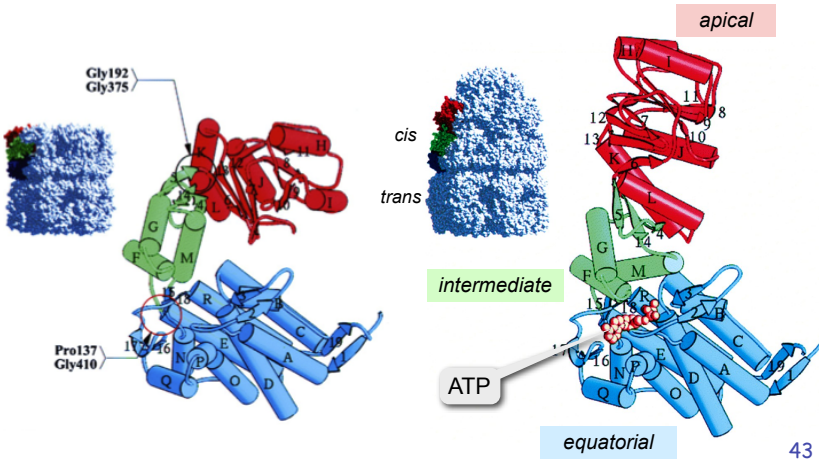


Figure 4-30b
Lehninger Principles of Biochemistry, Fifth Edition

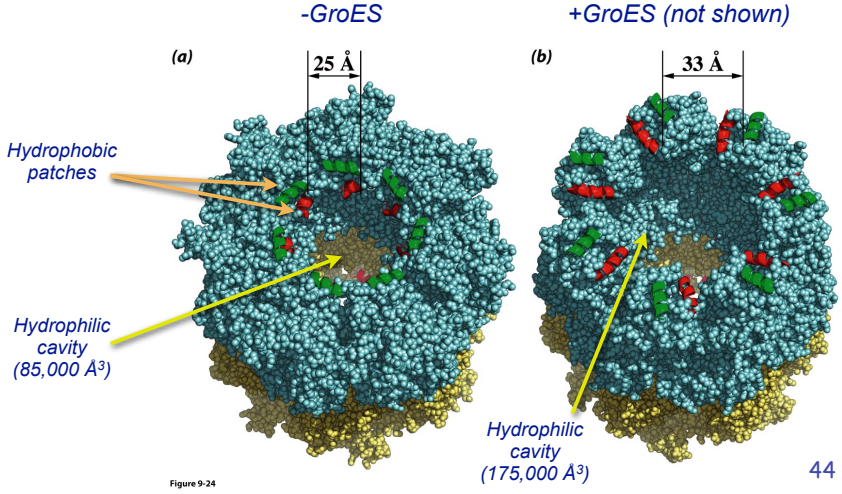
GroEL from *E. coli* is a typical member of the Hsp60 class of molecular chaperones, also known as chaperonins; GroEL is Cpn60. It works with a co-chaperone, GroES, that is a typical member of the Hsp10 or Cpn10 class of chaperonins. GroEL consists of 14 identical subunits, each of 57 kDa, that form a cylinder containing a central cavity. Each subunit consists of three domains. The equatorial domain has the binding site for ATP and forms most of the intersubunit contacts. The intermediate domain connects the equatorial domain to the apical domain, which binds polypeptides.

- GroEL: 3-domain protein

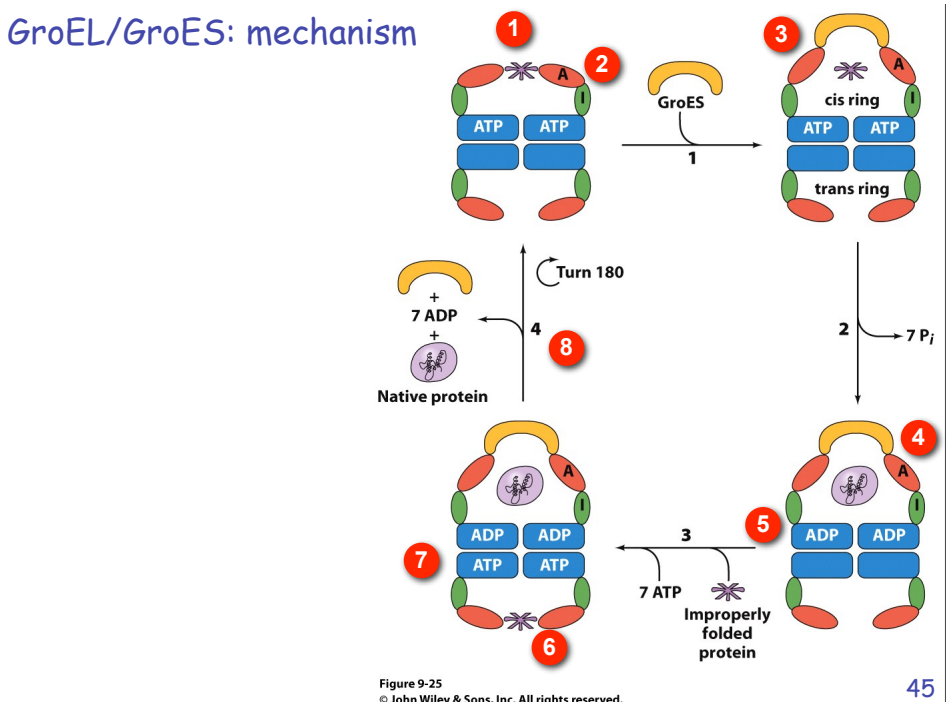


GroEL undergoes a series of conformational changes on the binding of its ligands, revealed at low resolution by EM and at high resolution by X-Ray crystallography. Free GroEL is a short symmetrical molecule; on the binding of ATP, it elongates. GroES binds as a heptamer to the apex of GroEL in the presence of ATP to give a bullet-shaped molecule. In doing so, it causes the cavity in the ring to which it is attached (the cis) to become larger. The other end of the cis ring is sealed off.
The hydrolysis of TP follows a classical sigmoid kinetics (allosteric proteins). GroEL undergoes a switch from a tense T state in the absence of ligands to the relaxed R state on the binding of ATP. The T state has a predominantly hydrophobic cavity. The conformational changes on the binding of ATP and GroES cause the hydrophobic surfaces to become buried and new hydrophilic ones to be exposed. Most important, the change in state is accompanied by a change in the affinity toward denatured proteins; GroEL binds them very tightly, GroEL-GroES much more weakly.

- Partially folded protein binds to at least 2 hydrophobic patches
- ATP-induced conformational change promotes unfolding
- GroES binds to same patches and displaces polypep. into cavity



Weaker binding of denatured states leads to faster folding
GroEL is targeted against denatured and partly denatured states of proteins in general, because eit binds hydrophobic side chains, which are the signatures of such states. Folding is slowed down because the bonds between the denatures state and GroEL are broken during folding. Thus the T stated of groEL is an enzyme that is complementary in structure to the denatured states of proteins. Its role is not to catalyse the steps of folding per se, but to act in the direction of unfolding to increase the yield of products.



Two.stroke mechanism: binding of STP and GroES to one ring (cis) prevents polypeptide binding to the other (trans). Trans ring can bind polypeptide only after ATP hydrolysis in cis ring. Polypeptide may be stretched by binding of ATP to the trans ring.
GroEL-GroES complex works on a subset of *E. coli* proteins , 82 cytosolic proteins are absolutely dependent on GroEL-GroES; many are α/β proteins (34% α/β barrels stabilised by many long-range interactions; why this selectivity ? How does a system distinguish between a misfolded protein and one that is partially folded on the right tracks?)

Sequence of loop on GroES that binds GroEL: GGIVLTGSA (chaperone substrates have similar motifs PXHHHXPXP)

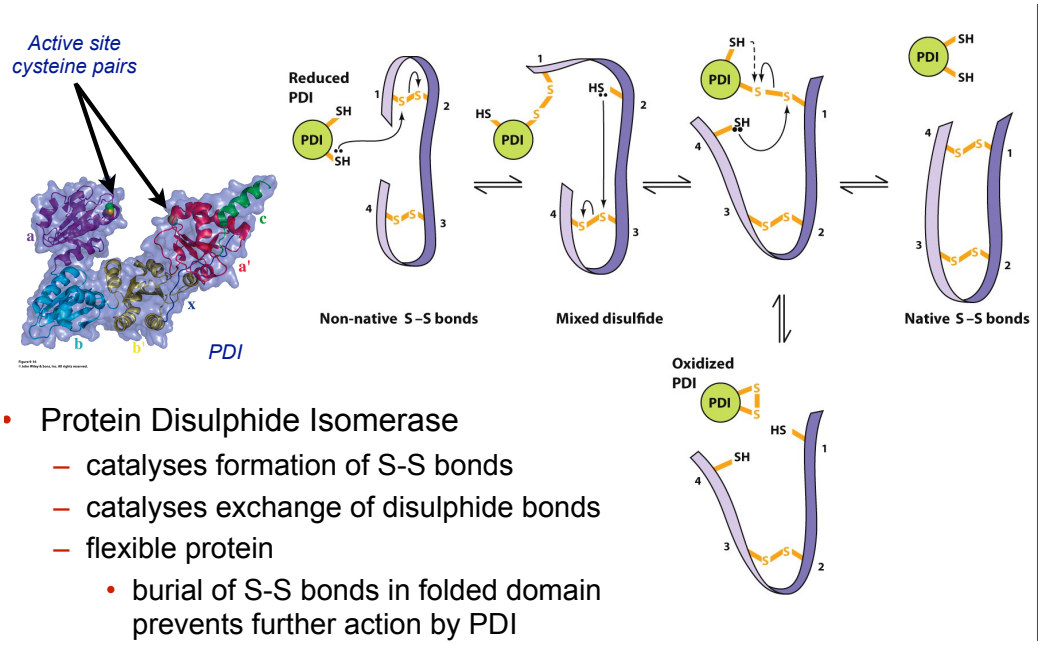
The essence of the mechanism of GroEL is first to prevent hydrophobic folding intermediates from sticking to one another and then, when necessary, to allow a misfolded state to unfold and have a fresh start for productive folding.
The initial stages of folding presented a puzzle, the Levinthal paradox, that was solved by the theoretician using a folding funnel. The final stages of folding present a problem for the experimentalists, especially the biotechnologists, because many desirable proteins that are produced by recombinant DNA technology do not fold successfully. Frequently, they precipitate as inclusion bodies. But the minichaperones provide an answer in some cases: attached to agarose, they provide a real folding funnel to which protein is added in denaturant at the top and pours out native at the bottom. The procedure works for proteins that have a low refolding yield of 5-10% or so in the absence of chaperones, but not, as yet, for the more recalcitrant ones.

Heat Shock Proteins

- Interplay when the cell has heat and stress damage to favour protein renaturation/degradation.
- Interplay with protein of proteins
- Prevent aggregation and misfolding
- Bind to nascent polypeptides to prevent premature folding
- Facilitate membrane translocation/import by preventing folding prior to membrane translocation
- Facilitate assembly/disassembly of multiprotein complexes (e.g. nucleosome)

Protein Disulfide Isomerase (PDI)

- The formation of correct disulfide bonds in nascent proteins is favoured in vivo by PDI.
- PDI has a broad substrate specificity for the folding of diverse disulfide-containing proteins
- By shuffling disulfide bonds, PDI enables proteins to find the thermodynamically most stable pairing those that are accessible.



Wrong disulfuide bridges are kinetics traps. It breaks just incorrect disulfide bonds. Correct disulfide will be shield from the PDI

Peptidyl Prolyl isomerase (PPI)

- Cis-trans Prolyl isomerization is the rate-limiting in the folding of many proteins in vitro.
- PPI accelerates cis-trans isomerization more than 300 fold by twisting the peptide bond so that the C,O, and N atoms are no longer planar.
- Prolines are the only residue that can be cis in the peptide bond. Isomerization is really slow usually.

Protein Turnover

Protein turnover is the balance between protein synthesis and degradation.

In the cell, proteins are degraded at different rates

- Ornithine decarboxylase has a half-life of 11 minutes.
- Hemoglobin lasts as long as a red blood cell.

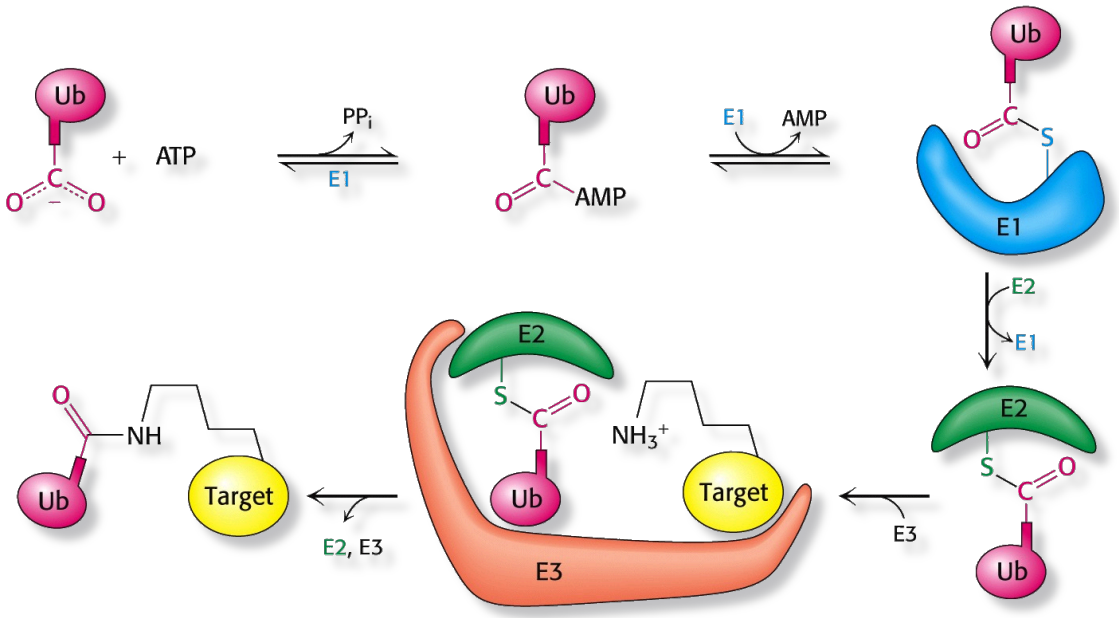
- Y-Crystallin (eye lens protein) lasts as long as the organism does.

Lysosomal degradation of proteins

- lysosomes are cellular vesicles containing proteolytic enzymes (e.g., papain-like cysteine protease, serine proteases, aspartic proteinases, etc., which are typically monomeric
- pH maintained at ~5.5 by proton-pumping ATPase
- account for 1-15% of cell volume (most abundant in liver and kidney)
- Hsc73 (constitutively-expressed Hsp70 chaperone) is involved in one pathway of lysosome-mediated degradation

Regulation of Protein Turnover

- *Ubiquitin* is a small protein of 76aa that is used to tag proteins for degradation in the proteasome pathway.
- In this pathway, proteins are chemically linked to ubiquitin via an isopeptide bond.
- - The ubiquitin is attached to the ε-amino groups of lysines by an isopeptide bond to the C-terminus of ubiquitin.
- - Ubiquitin is highly conserved throughout evolution: yeast and humans differ by only 3 out of 76 amino acids.
- - Reactions are similar to those that activate acyl groups in fatty acid degradation.
- - There are only one or a small number of distinct E1 proteins.
- - Eukaryotes have many distinct E2 and E3 proteins.
- - Only a single family of E2 proteins, but many families of E3 proteins.
- E3 provides most of the substrate specificity. Combination of E2 and E3 allow for fine tuning of specificity.
- -The human papilloma virus encodes for an E3 protein which targets the p53 tumor suppressor protein in its host. 90% of the cervical cancers are associated with this type of activity.



- *Ubiquitination can occur also via lysine residues (form UB), using the same mechanism E1, E2 and E3*

Protein Folding 3-4

Regulation of Protein Turnover
Ubiquitin: small protein of 76 aa that is used to tag proteins for degradation in the proteasome pathway; in this pathway, proteins are chemically linked to ubiquitin
Ubiquitin is activated and attached to proteins using a mechanism based in three enzymes; these enzymes are very variable but can be classified in three classes (E1, E2,E3); once a protein is tagged with a single ubiquitin molecule, this is signal to other ligases to attach additional ubiquitin molecules

Side Chains ubiquitination: ubiquitination can occur also via lysine residues
Ubiquitin-proteasome pathway: degradation pathway; misfolded proteins are bound by hsp chaperone (40, 70) and then these target them to E3 enzymes (E3 chip) → eventually will ubiquitinilate the protein with a series of 4 residues; tag then is responsible for the degradation by the proteasome; tag has affinity to bind region of the proteasome; it is an affinity tag

Proteasome: protein complexes inside all eukaryote and archaea, and in some bacteria. Located in the nucleus and the cytoplasm, the degradation process yields peptides of about seven to eight amino acid long.
It is a cylindrical complex containing a core of four stacked rings forming a central pore, each ring is composed of seven individual proteins. The core particle (destruction chamber contains active sites. The caps on the ends regulate entry into the destruction chamber). Leads can be variable, different proteins are degraded by specific proteasome (different leads).

Proteasome is a large collection of protease catalytic sites, the inner two rings are made of seven beta subunits hosting three to seven protease active sites. The target protein must enter the central pore before it is degraded. The outer two rings each contain seven alpha subunits that have a function of gate through which proteins enter the barrel. The alpha subunits are controlled by binding to cap structures or regulatory particles that recognise polyubiquitin tags attached to protein substrates and initiate the degradation process. It is a major way to recycle proteins.

The aggresome-autophagy pathway: has emerged as another crucial cellular defence system against toxic build-up of misfolded proteins. Basal autophagy mediates non-selective, bulk clrearance of misfolded proteins along with normal cellular proteins and organelles. It is a specialised type of induced autophagy that mediates selective clearance of misfolded and aggregated proteins under the conditions of proteotoxic stress. Heat shock induces a lot of aggregation within the cell (proteotoxic condition).
Misfolded protein can be brought back by some chaperones, however often they aggregate nd cause proteotoxicity→ induce aggresome pathway. Either a misfolded single proteins or a large aggregate proteins complexes form-> trigger this pathway to get rid of these very quickly. Series of polyubiquitin pput into different part of the aggregate; E2 and E3 system makes the ubiquitination, polyubiquitni has affinity to bind HDAC6 (dynein motor), walk into microtubule and form assembly near the nucleus, aggregated mass is formed as a black hole (aggresome). Aggresome then recruits membranes and form autophacig organelle that fuses with lysosome and is destroyed

Misfolding disease: associated with 20 aberrant diseases ranging from neurodegenerative disorders (Alzheimer's, Parkinson's …) to non-neuropatic diseases (Diabetes type II etc…). Despite starting by very different precursor proteins, amyloid fibrils share the same structural features.

Alzheimer's diseases: brain tissue contains amyloid plaques that form "neurofibrillary tangles" surrounded by dead neurons. Plaques formed by 40-42 aa amyloid-beta peptide (derived from a larger receptor protein (amyloid-beta precursor protein by proteolysis).
Are plaques the cause? injection of amyloid beta peptide into old-rhesus monkeys induces disease. Is this a infectious disease?? Cryo-EM managed to find the architecture of amyloid, these types of proteins are insoluble (no crystal for these). Only cryo-EM and solid state NMR (only recently).

Solid state NMR investigation, spectra of solid materials. Take fibrils from different patients with different phenotypes, grow some types of fibrils in the lab. Different types of phenotypes are associated with different types of structures.

Parkinson's Disease: neurodegenerative, alpha-synuclein in substantia nigra, neurons dye. Substantia nigra is associated with Lewy bodies and there are aggregates with alpha-synuclein. (don't know if aggregation alpha-synuclein induced Parkinson or something else induced alpha-synuclein aggregation). Neurotransmitter are in synaptic vesicles, fuse with plasma membrane and neurotransmitters are released. Synuclein seems to have a role in being a chaperon in SNARE machinery, diffuse and promotion of exocytosis of neurotransmitter. Dopaminergic neurons. Most of the functions of synuclein is involved in transient-membrane interaction.

Prion Diseases: Scrapie, Bovine Spongiform Encephalopacy, Creutfeld Jakob Disease
A new hypothesis for infection: an infectious protein, Stabley Prusiner, how do you prove it's just a protein? Some information can be transmitted to different cells, not via DNA or RNA but also with some proteins. In yeas info between mother and daughter cells can be transmitted via proteins.
Prion Protein (PrP): 280 aa; membrane anchored; widespread (mammals, drosophila); no known function, mouse knockout normal (until old Age) and resistant to infection-clue)
Prp exists in 2 conformational states
PrPC cellular protein ; normal
ProPSc – scrapie protein: chemical identical to PrPc, stable formational variant of Prpc, self-propagating

The amyloid pathway
Native state (folded in this case but could be disordered), activated state for the aggregation form dimers, oligomers of different size and evolve to protofibrils and assemble into larger fibrils. Not only accident of nature, also be used in functional purposes. Amyloids are rare states but can be used by some organisms for functional purposes. Some fungi use them to create some coats to isolate for attacks. Fibrils themselves are not toxic but the small oligomers are toxic, can be fragmented out of natural fibres, are able to diffuse and spread toxicity from one neurons from another, can disrupt cellular membrane and start the progression of the disease. They start o take holes in the membrane of neurons and slowly so kill them.
General features of amyloids are quite conserved. Fibre diffraction patter is not as regular as crystals. However, there is a regular distance between beta sheet (series of stacking of beta sheets in amyloids) –Z 4.7 Armstrong inter-stands packing. Inter-sheets distance: 10 A (series of beta sheets stacked together one in front of each other, the fibres move in the direction orthogonal to the single strand, different layers stacked at 10A between each other).
Recent amyloid structure, hierarchical assembly of large fibrils (Dobson). Treat amyloid fibres of chemical denaturant. Dry interfaces and in the space between protofibrils there is some void, possibly these types of structures are responsible for iron channel when fibrils are associated with the membrane.

Intrinsically Disordered Proteins (IDPs): 30% of Eukaryotic proteins are disordered
- Functions complementing the functional repertoire of ordered proteins
- Intermolecular interactions
- Remarkable binding promiscuity

Some proteins present a significant level of structural disorder. IDPs are often found to play the role of hubs in the protein-protein interactome, owing to a significant tendency to establish protein-proteins interactions. Some of the IDPs are extremely important for their connection to aberrant disorders such as Alzheimer's, Parkinson's and Cancer.
Intrinsic disorder can have multiple faces, can affect different levels of protein structural organization, and whole proteins, or various protein regions can be disordered to a different degree.
Because of their fundamental difference in structure and dynamics, folded and disordered proteins are associated to considerably different functions. The recognition that natural proteins can also be

disordered has opened to a large set of new disciplines in protein science. Disordered proteins are more flexible and adaptable shape. Structured protein has only 2-3 partners. P53 on the other side has hundreds of protein-protein interactions thanks to the disorder of the backbone, adapt the shape with different proteins. Involved in signalling as hub of pathways.

Energy Landscape of IDPs

- The folding energy landscape of a typical globular protein is funnelled
- The landscape of an IDP is composed of multiple minima, without any dominant global minimum (all the minima have favourable energy, so protein can jumb between all the different minima, thus structure)

General quartet Model: ordered – totally disordered protein .iN between there are intermediate states: Molten globule –pre-molten globule. Pre-molten maybe a single alpha helix in the middle of totally disordered structure. It is possible to physically separate the different types of structures. Size-exclusion chromatography is able to distinguish each of the four …

Sequence proteins of IDPs

Natively disordered proteins feature different distribution of amino acids …
If you do disordered proteins datasets minus ordered ones, negative values residues are more abundant globular proteins and positive values residues are more abundant in disordered. Bulky hydrophobic residues are more found in ordered protein (negative values) and in disordered proteins there are more charged residues, proline residues (does not like secondary structure).

Prerequisite for the absence of compact structure in a protein include:

- Low mean hydrophobicity (leading to low driving force for protein compaction)
- High net charge (leading to strong electrostatic repulsion)

Residual secondary structure in IDPs

It is not possible to crystallise disordered proteins, also not cryo-EM. NMR can provide some information. Can study the residual structure (helices, beta sheet etc…)Poly-proline structure is popular structure in disordered proteins: repetition of prolines.
Prion protein: divided in a structured C-term domain and a disordered N-term tail. The dual structural nature of this protein is accurately accounted by delta2D.
Another way to study structure is Circular Dichroism: different types of signatures (alpha helical, beta sheets and random coils).

Disordered proteins have huge tendency to form protein-protein interactions. Binding promiscuity is a distinctive feature of IDPs.
Fuzzy complexes: Static and dynamic disorder in protein-protein interactions are represented with arrow indicating increasing disorder. Folded protein partner, disorder proteins go in the top of the structure and form fuzzy complex. (not fit and lock mechanism where proteins form dimers).

Synuclein has been shown to be able to adapt to different types of membranes. It can adapt it shapes of membrane (curved but also more flat membranes). This is a result of the disorder of the proteins. In solution is a total random coil, when bound to membranes form amphipathic helices the bind different membranes.

Spider-silk and disorder-to-order transitions

Large number of different types of fibres, similar to amyloid. Different types of silk for different typed of net. The proteins inside spider are included in some ampullas and in that region they are disordered (form disordered micellar way NR3), proteins do not polymerise in this conditions; then it goes into a particular shear stress and at this point the amylogenic regions are released and form fibres. Play between order and disorder.