Protein Folding Problem

Structure of protein is determined by **only** the amino acid sequence! - Proved by Christian Anfinsen, 1957, using RNAse A - show characteristics of protein folding

- Has 4 S-S bonds (8 cysteines total - all cysteines are S-S bonded in RNAse A). The di-S-S pairs are 26-84, 40-95, 58-110, 65-72

- Cleave non-covalent bonds using chemical denaturants, such as SDS, urea (8 M) and guanidiuium chloride (6 M). They compete for the Hydrogen bonds and perturb the solvent.

- Cleave S-S bonds using reducing agents. The best known is β-mercaptoethanol, which has the chemical structure $HO\text{-}CH_2\text{-}CH_2\text{-}SH$

- First experiment is to treat with both urea and β-ME. After this treatment, there was a randomly coiled chain devoid of enzymatic activity (reduced & unfolded).

- Allow to reform (reoxidize) S-S bonds under two conditions. Since there were 8 cysteines, there are 7*5*3*1 = 105 possible combinations of S-S bonds. Only one of these arrangements can yield an active enzyme; the other 104 combinations give inactive enzymes. Thus, if the S-S bonds were to reform randomly, only about 1% of the original activity would be restored, while if all were to reform to the active conformation, 100% of the original activity would be restored.

- Condition 1: Urea was dialyzed out, 100% of the enzymatic activity was restored!
  * Sulfhydryl groups of the denatured enzyme became oxidized by air.
  * Enzyme spontaneously folded up into catalytically active form.
  * Thus, the information needed to specify the catalytically active structure of ribonuclease is contained in its amino acid sequence.

- Condition 2: If S-S bonds were allowed to reoxidize while still in the presence of 8 M urea, and then the urea was dialyzed out of the preparation, only 1% of the enzymatic was restored. However, if after this, a trace amount of β-ME was added to the preparation, 100% of the enzymatic activity was restored!
  * Wrong disulfides formed pairs in the presence of urea - called “scrambled” ribonuclease. Protein was unable to begin to form right structure.
  * β-mercaptoethanol catalyzed the rearrangement of disulfide pairings until the native (acitve) arrangement was regained in about 10 hours.
  * This process was driven by the decrease in free energy as the scrambled conformation were converted into the stable, native conformation of the enzyme.

Sequence Specifies Conformation

Protein Folding Problem

- Similar refolding experiments have been performed on other proteins. Often, they work the same way. However, sometimes they cannot reform - usually, in this case, when the protein was denatured, it became tangled with itself or other denatured protein chains such that it could not untangle itself.

Levinthal's Paradox:

- Say you have a protein with 100 aa; there are 3 conformations possible per aa. Thus, there are 3^{100} possibilities = $5x10^{47}$.
- Given that the atomic vibration is ~ 10^{-13} seconds (1 conformation) it would take $5x10^{34}$ seconds, or $1.6x10^{27}$ years, for the protein to sample all possibilities.
- Thus, the protein is obviously not sampling all possibilities - how is it folding?

First folds into a "molten globule" ("molten" - fluctuating; "globule" - compact); then folds into native tertiary structure.

Folding is cooperative/Proteins that assist in folding

- Protein Disulfide Isomerases (PDI):These enzymes promote the rapid reshuffling of disulfide (S—S) bonds in a protein. In order for a protein to reach its lowest energy folded state, the correct disulfide bonds must be in place. If incorrect disulfide bonds have formed, they must be broken to allow the correct bonds to form. This process would take an extremely long time to happen at random. The PDIs break and form disulfide bonds rapidly, reducing the amount of time it takes for the correct disulfide bonds to form and for the protein to properly fold (reach its native conformation).
- Peptidyl Proline Isomerases (PPI): These enzymes catalyze the interconversion of cis and trans proline.
- Chaperones: These proteins bind nascent polypeptide chains as they are synthesized from the ribosome. They protect these partially synthesized polypeptide chains from improper interactions with themselves or other proteins in the endoplasmic reticulum (ER). This protection prevents improper folding and the formation of aggregates. Once protein synthesis is complete, the chaperones release the chain and protein folding can occur.

Forces that drive folding

- Electrostatic bonds (salt bridges)
- Hydrogen (H-) bonds
- Van der Waals Interactions
- Hydrophobic Interactions

The Loss of Chain Entropy Opposes Folding

Before a protein is folded, it is simply a long chain of amino acids. This long chain can move at random in many different ways. Once a protein folds, it must remain fixed in a particular form. As a result, randomness is lost, meaning the chain has lost entropy.