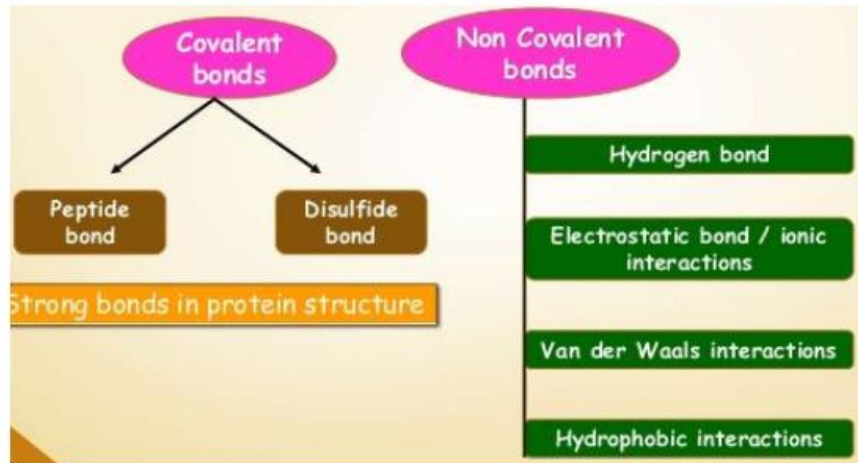


PROTEIN ENGINEERING

LECTURE 08: Forces that stabilize Protein Structure

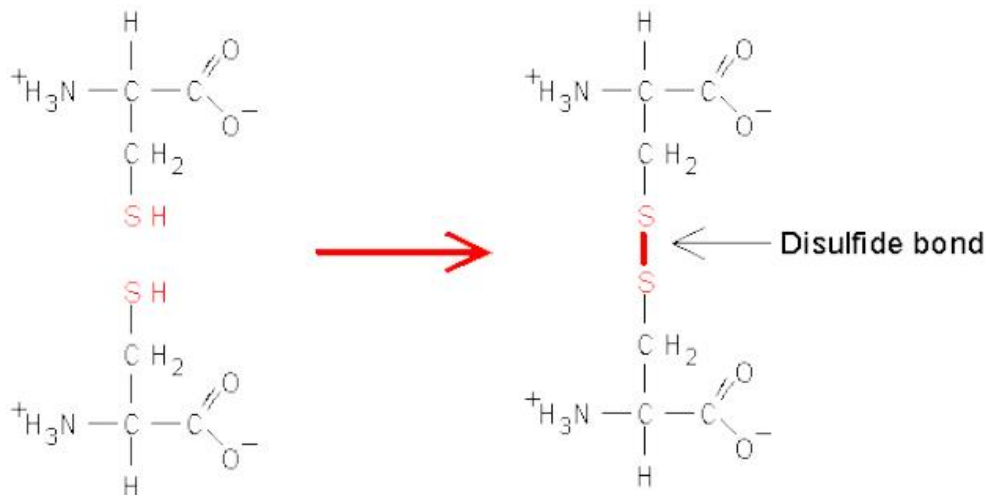
Proteins are formed of amino acids linked together by the following types of bonds



Covalent Bonds - Disulfide Bridges

Covalent bonds are the strongest chemical bonds contributing to protein structure. Covalent bonds arise when two atoms share electrons.

In addition to the covalent bonds that connect the atoms of a single amino acid and the covalent peptide bond that links amino acids in a protein chain, covalent bonds between cysteine side chains can be important determinants of protein structure. Cysteine is the sole amino acid whose side chain can form covalent bonds, yielding disulfide bridges with other cysteine side chains: $--CH_2-S-S-CH_2-$. A disulfide bridge is shown here:



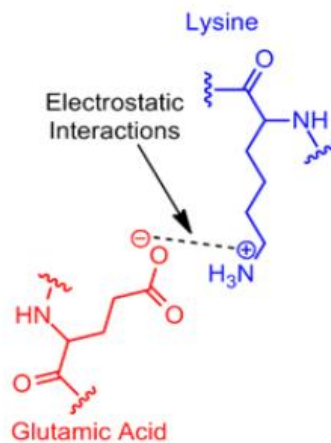
PROTEIN ENGINEERING

Non-covalent bonds

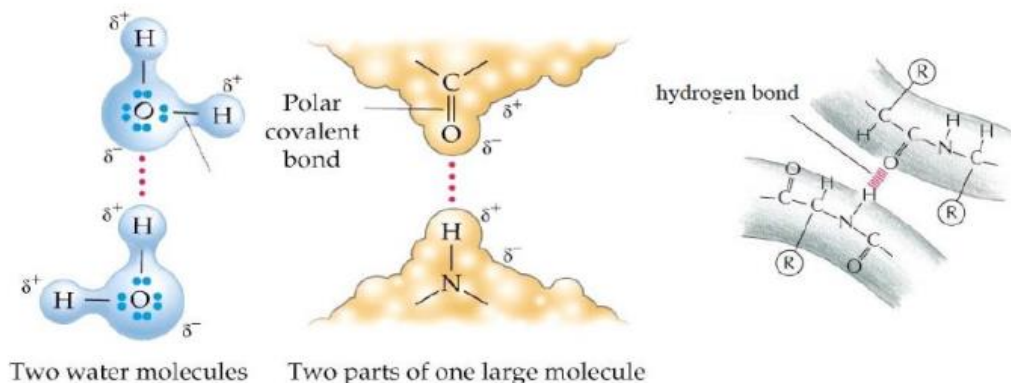
Electrostatic Interactions

A. Ionic Bonds - Salt Bridges

Ionic bonds are formed as amino acids bearing opposite electrical charges are juxtaposed in the hydrophobic core of proteins. Ionic bonding in the interior is rare because most charged amino acids lie on the protein surface. Although rare, ionic bonds can be important to protein structure because they are potent electrostatic attractions that can approach the strength of covalent bonds. A ionic bond-salt bridge between a negatively charged O on the sidechain of glutamic acid lies 2.8 Å from the positively charged N on the amino terminus (lysine) is shown here .



B. Hydrogen Bonds



PROTEIN ENGINEERING

Hydrogen bonds are a particularly strong form of dipole-dipole interaction. Because atoms of different elements differ in their tendencies to hold onto electrons -- that is, because they have different electronegativities -- all bonds between unlike atoms are polarized, with more electron density residing on the more electronegative atom of the bonded pair. Separation of partial charges creates a dipole, which you can think of as a mini-magnet with a positive and a negative end. In any system, dipoles will tend to align so that the positive end of one dipole and the negative end of another dipole are in close proximity. This alignment is favorable. Hydrogen bonds are dipole-dipole interactions that form between heteroatoms in which one heteroatom (e.g. nitrogen) contains a bond to hydrogen and the other (e.g. oxygen) contains an available lone pair of electrons. You can think of the hydrogen in a hydrogen bond as being shared between the two heteroatoms, which is highly favorable. Hydrogen bonds have an ideal X-H-X angle of 180° , and the shorter they are, the stronger they are. Hydrogen bonds play an important role in the formation of secondary structure. Alpha helices are hydrogen bonded internally along the backbone whereas beta strands are hydrogen bonded to other beta strands. Side chains can also participate in hydrogen bonding interactions. You should be able to list the side chains that can participate in hydrogen bonds now that you know the structures of the side chains. Because hydrogen bonds are directional, meaning the participating dipoles must be aligned properly for a hydrogen bond to form (another way of saying it is that the hydrogen bonding angle must be larger than about 135° , with an optimum of 180°), and because unfavorable alignment of participating dipoles is repulsive, hydrogen bonds between side chains play key roles in determining the unique structures that different proteins form.

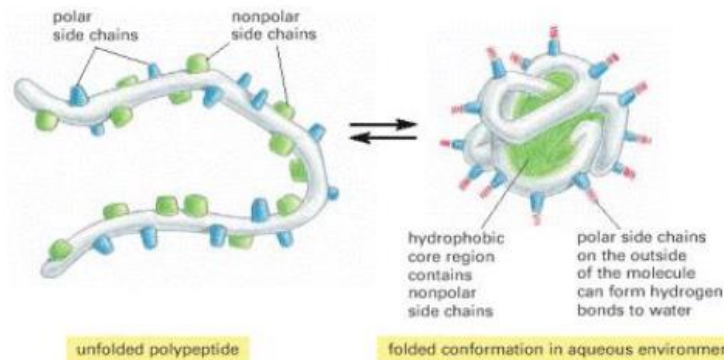
Hydrophobic Bonds

Hydrophobic bonds are a major force driving proper protein folding. Burying the nonpolar surfaces in the interior of a protein creates a situation where the water molecules can hydrogen bond with each other without becoming excessively ordered. Thus, the energy of the system goes down.

Therefore, an important factor governing the folding of any protein is the distribution of its polar and nonpolar amino acids. The nonpolar (hydrophobic) side chains in a protein such as those belonging to phenylalanine, leucine, isoleucine, valine, methionine and tryptophan tend to cluster in the interior of the molecule (just as hydrophobic oil droplets coalesce in water to

PROTEIN ENGINEERING

form one large droplet). In contrast, polar side chains such as those belonging to arginine, glutamine, glutamate, lysine, etc. tend to arrange themselves near the outside of the molecule, where they can form hydrogen bonds with water and with other polar molecules. There are some polar amino acids in protein interiors, however, and these are very important in defining the precise shape adopted by the protein because the pairing of opposite poles is even more significant than it is in water.



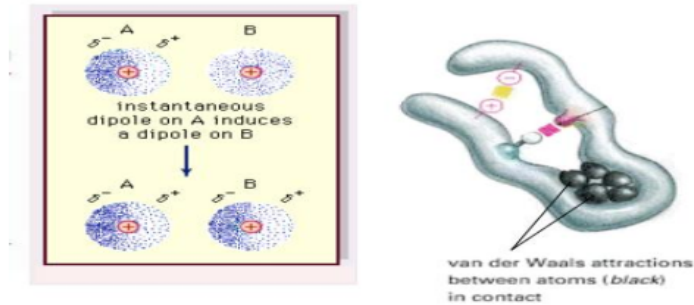
Van der Waals Forces

The Van der Waals force is a transient, weak electrical attraction of one atom for another. Van der Waals attractions exist because every atom has an electron cloud that can fluctuate, yielding a temporary electric dipole. The transient dipole in one atom can induce a complementary dipole in another atom, provided the two atoms are quite close. These short-lived, complementary dipoles provide a weak electrostatic attraction, the Van der Waals force. Of course, if the two electron clouds of adjacent atoms are too close, repulsive forces come into play because of the negatively-charged electrons. The appropriate distance required for Van der Waals attractions differs from atom to atom, based on the size of each electron cloud, and is referred to as the Van der Waals radius. The dots around atoms in this and other displays represent Van der Waals radii.

Van der Waals attractions, although transient and weak, can provide an important component of protein structure because of their sheer number. Most atoms of a protein are packed sufficiently close to others to be involved in transient Van der Waals attractions.

Van der Waals forces can play important roles in protein-protein recognition when complementary shapes are involved. This is the case in antibody-antigen recognition, where a "lock and key" fit of the two molecules yields extensive Van der Waals attractions.

PROTEIN ENGINEERING



Thermodynamics of protein folding

In contemplating protein folding, it is necessary to consider different types of amino acid side-chains separately. For each situation, the reaction involved will be assumed to be:



Note that this formalism means that a negative ΔG implies that the folding process is spontaneous.

First we will look at polar groups in an aqueous solvent. For polar groups, the ΔH_{chain} favors the unfolded structure because the backbone and polar groups interact form stronger interactions with water than with themselves. More hydrogen bonds and electrostatic interactions can be formed in unfolded state than in the folded state. This is true because many hydrogen bonding groups can form more than a single hydrogen bond. These groups form multiple hydrogen bonds if exposed to water, but frequently can form only single hydrogen bonds in the folded structure of a protein.

For similar reasons, the $\Delta H_{\text{solvent}}$ favors the folded protein because water interacts more strongly with itself than with the polar groups in the protein. More hydrogen bonds can form in the absence of an extended protein, and therefore the number of bonds in the solvent increases when the protein folds.

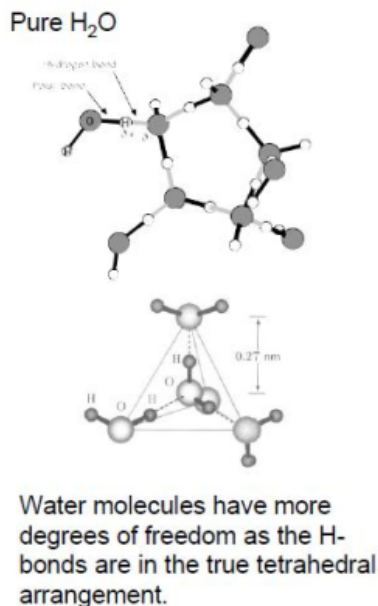
The sum of the ΔH_{polar} contributions is close to zero, but usually favors the folded structure for the protein slightly. The chain ΔH contributions are positive, while the solvent ΔH contributions are negative. The sum is slightly negative in most cases, and therefore slightly favors folding.

PROTEIN ENGINEERING

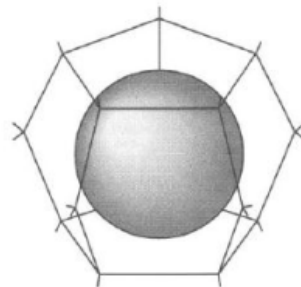
The ΔS_{chain} of the polar groups favors the unfolded state, because the chain is much more disordered in the unfolded state. In contrast, the $\Delta S_{\text{solvent}}$ favors the folded state, because the solvent is more disordered with the protein in the folded state. In most cases, the sum of the ΔS_{polar} favors the unfolded state slightly. In other words, the ordering of the chain during the folding process outweighs the other entropic factors.

The ΔG_{polar} that is obtained from the values of ΔH_{polar} and ΔS_{polar} for the polar groups varies somewhat, but usually tends to favor the unfolded protein. In other words, the folding of proteins comprised of polar residues is usually a nonspontaneous process.

Next, we will consider a chain constructed from non-polar groups in aqueous solvent. Once again, the ΔH_{chain} usually favors the unfolded state slightly. Once again, the reason is that the backbone can interact with water in the unfolded state. However, the effect is smaller for non-polar groups, due to the greater number of favorable van der Waals interactions in the folded state. This is a result of the fact that non-polar atoms form better van der Waals contacts with other non-polar groups than with water; in some cases, these effects mean that the ΔH_{chain} for nonpolar residues is slightly negative.



H₂O around a hydrophobic molecule



Water molecules have less degrees of freedom in the clathrate cage arrangements because some H-bonds cannot point inside toward the hydrophobic sphere

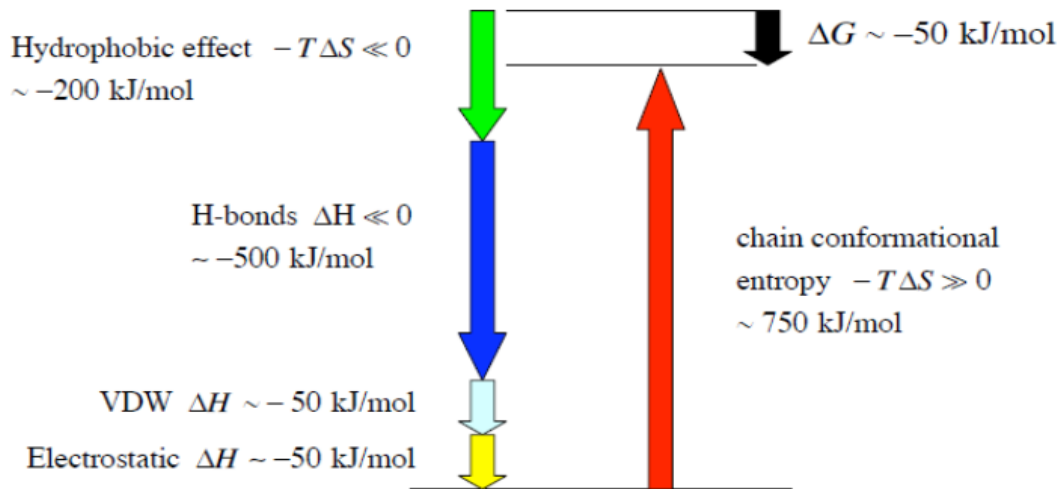
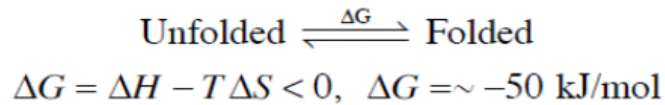
As with the polar groups, the $\Delta H_{\text{solvent}}$ for non-polar groups favors the folded state. In the case of non-polar residues, $\Delta H_{\text{solvent}}$ favors folding more than it does for polar groups, because water interacts much more strongly with itself than it does with non-polar groups.

PROTEIN ENGINEERING

The sum of the $\Delta H_{\text{non-polar}}$ favors folding somewhat. The magnitude of the $\Delta H_{\text{nonpolar}}$ is not very large, but is larger than the magnitude of the ΔH_{polar} , which also tends to slightly favor folding.

The ΔS_{chain} of the non-polar groups favors the less ordered unfolded state. However, the $\Delta S_{\text{solvent}}$ highly favors the folded state, due to the hydrophobic effect. During the burying of the non-polar side chains, the solvent becomes more disordered. The $\Delta S_{\text{solvent}}$ is a major driving force for protein folding which is called conformational entropy.

The $\Delta G_{\text{non-polar}}$ is therefore negative, due largely to the powerful contribution of the $\Delta S_{\text{solvent}}$. Adding together the terms for ΔG_{polar} and $\Delta G_{\text{non-polar}}$ gives a slightly negative overall ΔG for protein folding, and therefore, proteins generally fold spontaneously.



Raising the temperature, however, tends to greatly increase the magnitude of the $T\Delta S_{\text{chain}}$ term, and therefore to result in unfolding of the protein.

The folded state is the sum of many interactions. Some favor folding, and some favor the unfolded state. The qualitative discussion above did not include the magnitudes of the effects. For real proteins, the various ΔH and ΔS values are difficult to measure accurately. However,

PROTEIN ENGINEERING

for many proteins it is possible to estimate the overall ΔG of folding. Measurements of this value have shown that the overall ΔG for protein folding is very small: only about -10 to -50 kJoules/mol. This corresponds to a few salt bridges or hydrogen bonds.

Studies of protein folding have revealed one other important point: the hydrophobic effect is very important, but it is relatively non-specific. Any hydrophobic group will interact with essentially any other hydrophobic group. While the hydrophobic effect is a major driving force for protein folding, it is the constraints imposed by the more geometrically specific hydrogen bonding and electrostatic interactions in conjunction with the hydrophobic interactions that largely determine the overall folded structure of the protein.

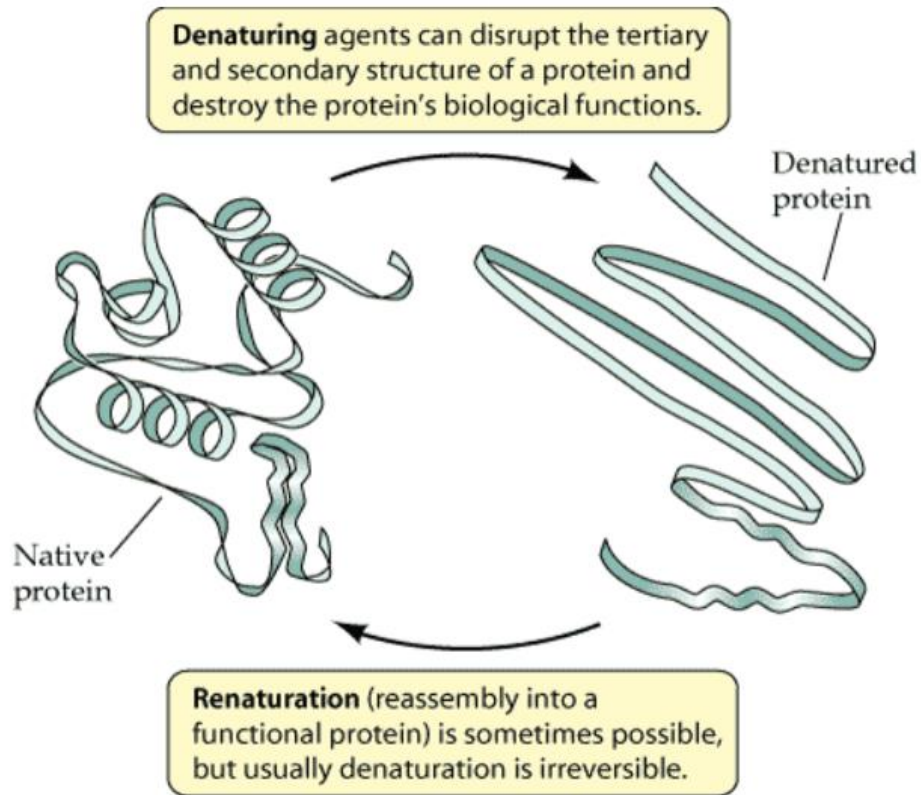
PROTEIN DENATURATION

Denaturation of proteins happens when there is a loss in the secondary or tertiary or quaternary structure of it. Rarely denaturation happens with loss of primary structure.

There are two types of protein denaturation:

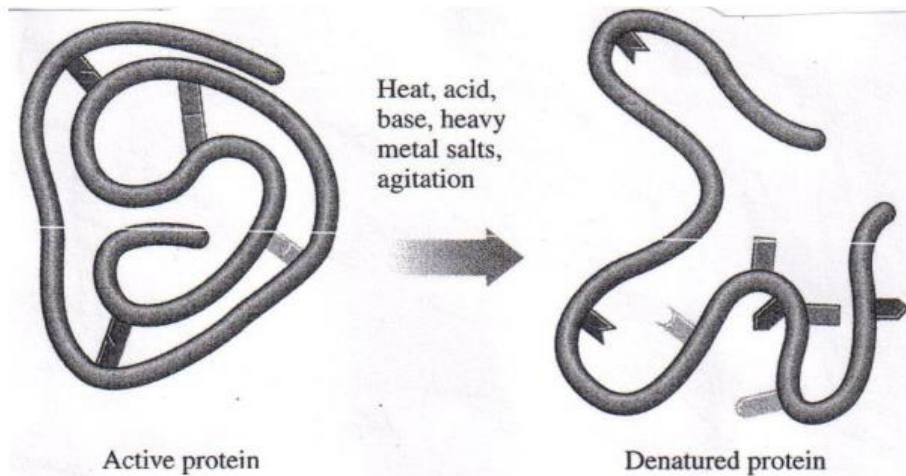
- 1- Reversible denaturation: when the effect of denaturation reagent is removed by dialysis, the enzyme will gain its activity again. If we store the protein below $0\text{ }^{\circ}\text{C}$, it will lose its activity, once the temperature is increased the protein is back active.
- 2- Irreversible denaturation: if the protein is at temperature over $60\text{ }^{\circ}\text{C}$ it will form an insoluble substance called coagulum as in heating the egg white. It cannot be returned to its original state.

PROTEIN ENGINEERING



Denaturation agents

Any mechanical or chemical agent that causes the denaturation of a protein is called a denaturing agent. Denaturation agents include heat, acids and bases, organic compounds, heavy metal ions, and mechanical agitation. Some denaturations are reversible, while others permanently damage the protein.



PROTEIN ENGINEERING

The most common factors that denatures proteins includes:

- **Heat:** This disrupts hydrogen bonds and non-polar hydrophobic interactions.
- **Alcohol:** This affects the hydrogen bonds that are formed between the amide groups of the secondary level and it also affects the hydrogen bonding between the side chains of the tertiary level.
- **Acids and Bases and Heavy Metals:** Strong acids and bases and heavy metals denatures the protein the same way by disrupting the salt bridge.
- **Urea:** Destabilizes internal bonds and unfolds the protein because it (urea) is a chaotropic.
- **UV:** The effects are similar to heat
- **Organic solvents:** They interrupt the intracovalent interactions of proteins. example Acetone.
- **Detergents:** Breaks up positive and negative interactions in the protein chains.

Thermal and cold denaturation

- Cold denaturation of proteins is a fascinating process, most ordered molecular systems are more stable at low temperatures where the thermal fluctuations are minimum. However, proteins are stable only over a limited range of temperature and pressure.
- A protein can be easily unfolded by heating and cooling, apart from different solvents and high/low pressures.
- There are multiple elements that contribute to protein stability, hydrogen bonds play a crucial role in driving secondary structures, the hydrophobic effect and other electrostatics contribute more towards stabilizing the tertiary structure.
- It is the disruption of this balance in the hydrophobic effect and the hydrogen bonding interactions that drive unfolding.

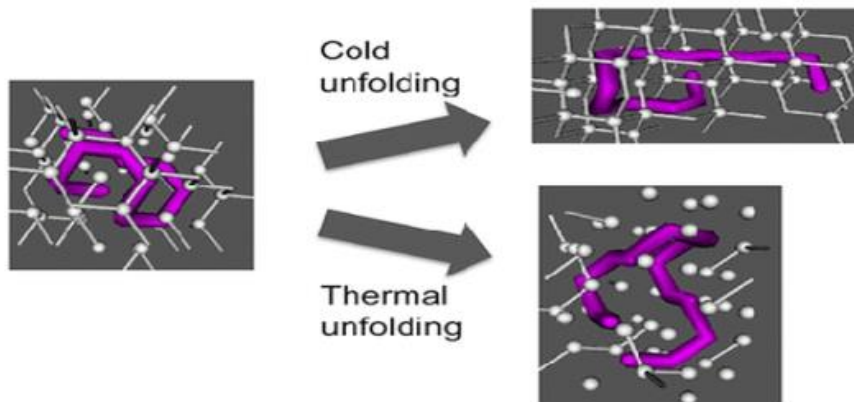
What is the main difference between thermal denaturation and cold denaturation?

Very simply, thermal denaturation is entropically driven, cold denaturation is enthalpically driven.

In elaborate:

PROTEIN ENGINEERING

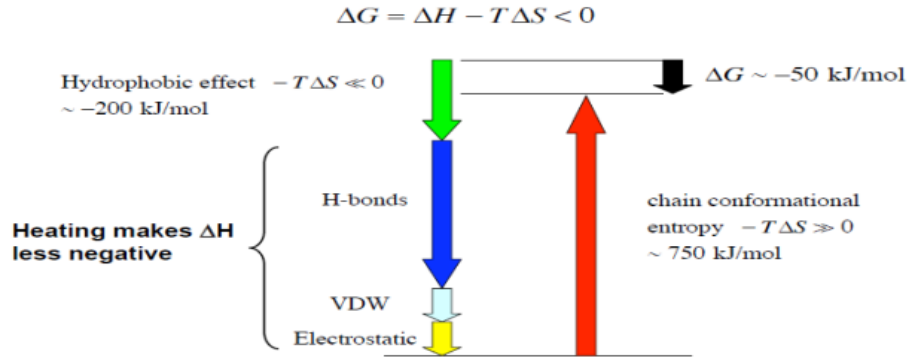
- Thermal and cold denaturation arises from the temperature dependence of a thermodynamic equilibrium. This is described by the Gibbs-Helmholtz equation, which theoretically predicts denaturation with heating and cooling.
- Thus changes in system enthalpy, entropy and heat capacity characterize unfolding. Based on the behavior of the Gibbs-Helmholtz equation, we know that heat and cold denaturation are driven by different mechanisms.
- When proteins are heated, they become more extended. The energy one supplies to the system in the form of heat, breaks enthalpically favorable interactions. As a result, **there is an increase in conformational entropy** and the polypeptide chain displays typical polymer behavior.
- Cold denaturation leads to partial unfolding of the polypeptide chain, and is a result of change in interaction between water and hydrophobic groups.
- With decrease in temperature, the **free energy cost for hydrophobic effect, or the unfavorable interaction of nonpolar residues with water decreases**, thus increasing their hydration.
- In most experiments, decrease in temperature leads to loss of tertiary structure, ergo, **weakening of the hydrophobic effect**.
- In cold denatured systems, **the hydrogen bonds involving interfacial water is more favorable than that in the bulk water**. And this precisely is what is hypothesized to drive unfolding.



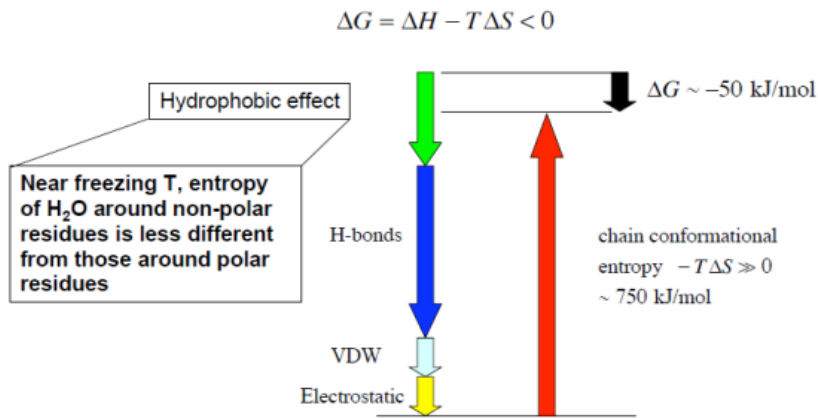
Ordered water molecules around unfolded state in the cold denatured system:

PROTEIN ENGINEERING

Denaturation by Heat -- break H-bonds and other enthalpically favorable interactions



Denaturation by Cold -- reduce the contribution from hydrophobic effect



pH

A protein consists of amino acids. Some of these amino acids are polar, having positively charged sides and negatively charged sides. A change in pH simply means a change in the amount of (H^+) atoms. As you can see these hydrogen atoms are positively charged, and attract the negative side of the polar amino acids. So a change in the pH changes the stability of a protein structure and can cause its denaturation

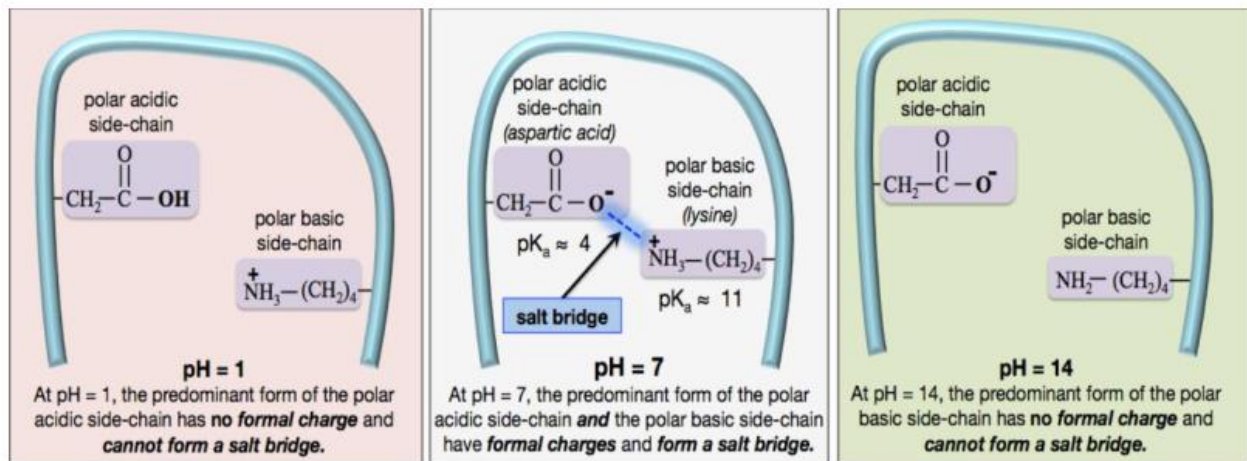
There are ionizable groups in the individual amino acids. The rate at which they ionize depends on the group and the pH. A high concentration of hydrogen ions (low pH) will result

PROTEIN ENGINEERING

in more groups being protonated. Carboxyl groups (aspartic acid, glutamic acid, the carboxy terminus) and phenolic groups are uncharged when protonated. The nitrogen groups (amines on lysine, guanidino of arginine, and imidazole in histidine, etc.) are charged when protonated.

Charged groups will tend to move towards the surface of the protein. Uncharged groups tend to move inwards. There may be a region that has an excess of like charges on adjacent chains which repel each other when there were mixed charges at physiological pH.

There is can be an increased ionic strength to the medium that impacts the electrostatic interactions of atoms, but the same can be achieved with neutral salts. This is the principle behind "salting in"/"salting out".

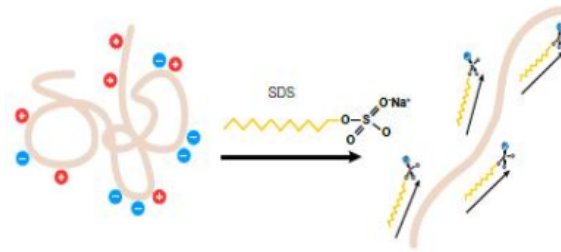


Detergent

Detergent has a hydrophobic side and a hydrophilic side. Proteins have hydrophobic and hydrophilic sides, the detergent is attracted to these and forces the protein apart. A protein's 3-D structure is partially created by hydrophobic and hydrophilic interactions to itself, the detergent substitutes this self bonding with detergent-amino acid bonding. Furthermore, detergent is a salt and breaks up positive and negative interactions of the 3-D shape as well.

Example: The detergent sodium dodecyl sulfate (SDS) binds noncovalently to proteins in a manner that imparts: An overall negative charge on the proteins. Since SDS is negatively charged, it masks the intrinsic charge of the protein it binds and leads to long, rod-like conformation on the proteins instead of a complex tertiary shape

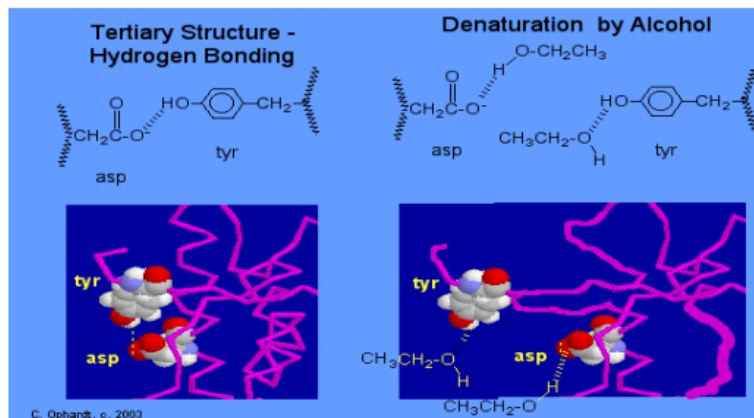
PROTEIN ENGINEERING



Alcohol denaturation

Hydrogen bonding occurs between amide groups in the secondary protein structure. Hydrogen bonding between "side chains" occurs in tertiary protein structure in a variety of amino acid combinations. All of these are disrupted by the addition of another alcohol.

A 70% alcohol solution is used as a disinfectant on the skin. This concentration of alcohol is able to penetrate the bacterial cell wall and denature the proteins and enzymes inside of the cell. A 95% alcohol solution merely coagulates the protein on the outside of the cell wall and prevents any alcohol from entering the cell. Alcohol denatures proteins by disrupting the side chain intramolecular hydrogen bonding. New hydrogen bonds are formed instead between the new alcohol molecule and the protein side chains.



Heavy Metal Salts:

Heavy metal salts act to denature proteins in much the same manner as acids and bases. **Heavy metal salts** usually contain Hg^{+2} , Pb^{+2} , Ag^{+1} , Tl^{+1} , Cd^{+2} and other metals with high atomic weights. Since salts are ionic they disrupt salt bridges in

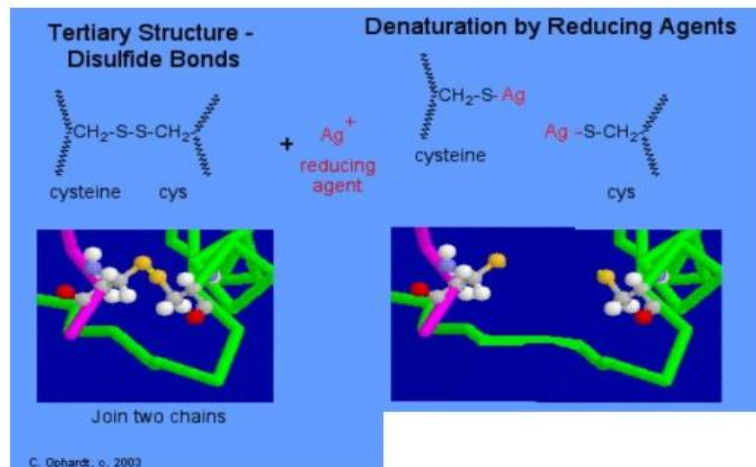
PROTEIN ENGINEERING

proteins. The reaction of a heavy metal salt with a protein usually leads to an insoluble metal protein salt.

This reaction is used for its disinfectant properties in external applications. For example AgNO_3 is used to prevent gonorrhea infections in the eyes of new born infants. Silver nitrate is also used in the treatment of nose and throat infections, as well as to cauterize wounds.

Mercury salts administered as Mercurochrome or Merthiolate have similar properties in preventing infections in wounds.

This same reaction is used in reverse in cases of acute heavy metal poisoning. In such a situation, a person may have swallowed a significant quantity of a heavy metal salt. As an antidote, a protein such as milk or egg whites may be administered to precipitate the poisonous salt. Then an emetic is given to induce vomiting so that the precipitated metal protein is discharged from the body.



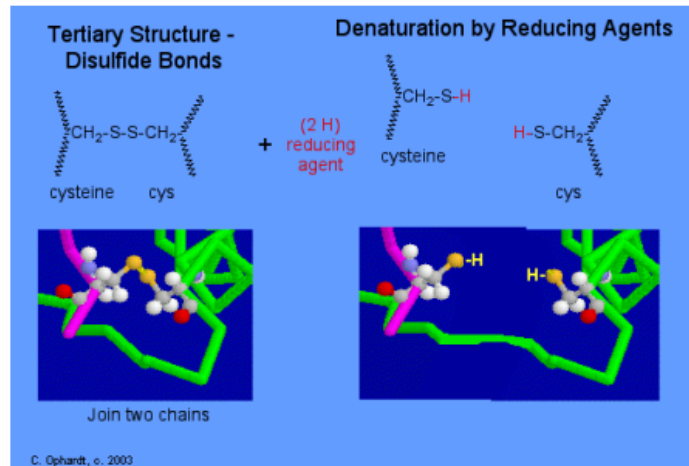
Reducing Agents Disrupt Disulfide Bonds:

Disulfide bonds are formed by oxidation of the sulfhydryl groups on cysteine. Different protein chains or loops within a single chain are held together

PROTEIN ENGINEERING

by the strong covalent disulfide bonds. Both of these examples are exhibited by the insulin.

If oxidizing agents cause the formation of a disulfide bond, then reducing agents, of course, act on any disulfide bonds to split it apart. Reducing agents add hydrogen atoms to make the thiol group, -SH. The reaction is:

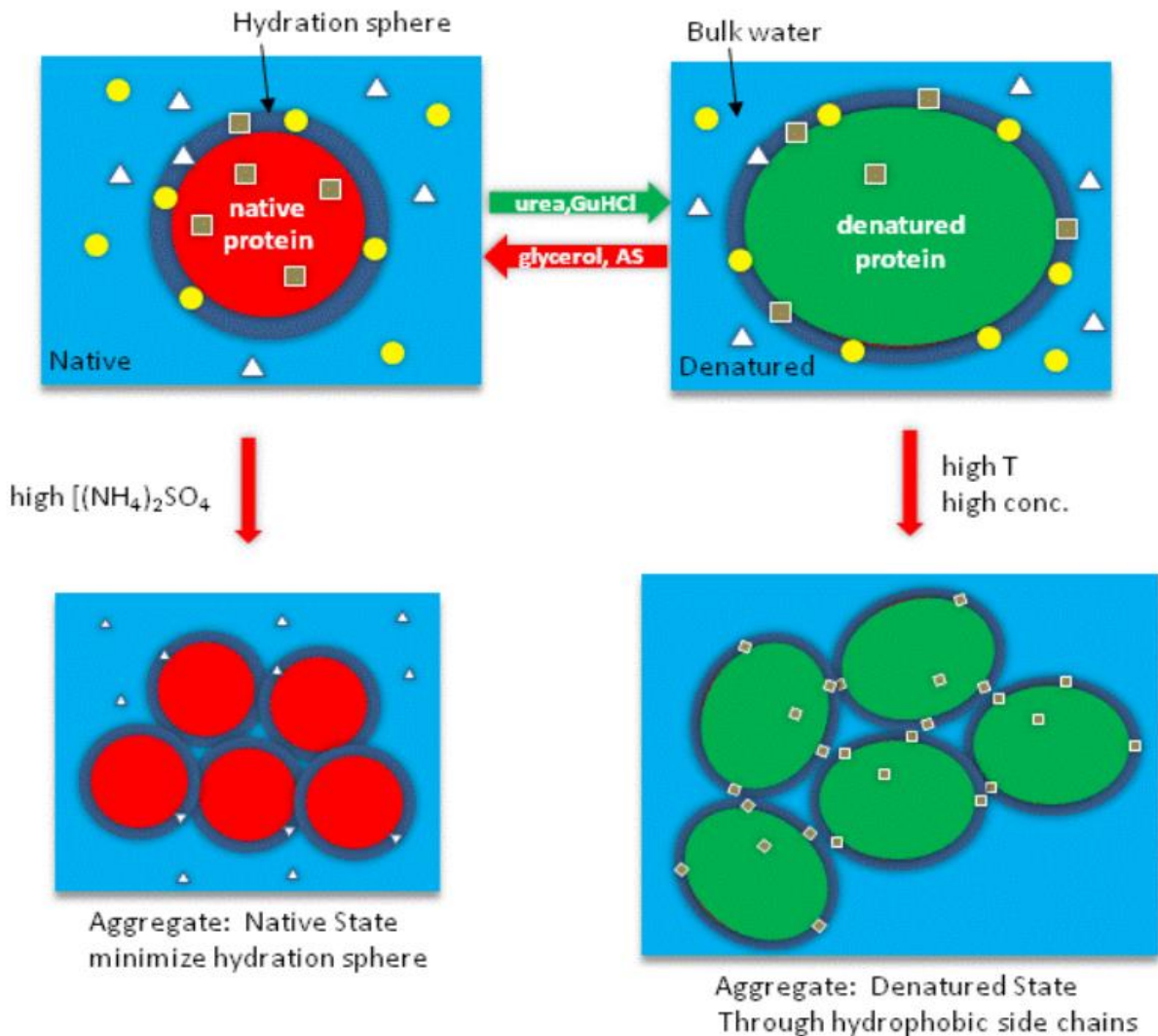


Urea/ guanidine HCl denatures proteins:

Apparently urea binds preferentially to the protein surface, and hence tends to increase the protein's surface area and hydrophobic exposure, and denature proteins. However, note in the figure below that glycerol, a bigger polar but uncharged molecule, stabilizes the native state. This pair of uncharged additives has correspondingly similar effects on protein stability as does the charged guanidine HCl/ammonium sulfate pair.

PROTEIN ENGINEERING

- Urea, GuHCl: selective partition into hydration sphere; shift eq. →
- △ Glycerol, $(\text{NH}_4)_2\text{SO}_4$ (AS): selective partition into bulk water; shift eq. ←
- Nonpolar side chains



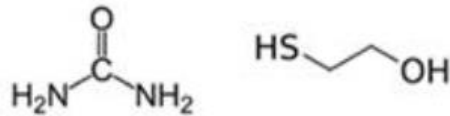
Reversibility and irreversibility

In very few cases, denaturation is reversible (the proteins can regain their native state when the denaturing influence is removed). This process can be called renaturation. This understanding has led to the notion that all the information needed for proteins to assume their native state was encoded in the primary structure of the protein, and hence in the DNA that codes for the protein, the so-called "Anfinsen's thermodynamic hypothesis".

PROTEIN ENGINEERING

Ribonuclease Refolding Experiment

- Ribonuclease is a small protein that contains 8 cysteins linked via four disulfide bonds
- Urea in the presence of 2-mercaptoethanol fully denatures ribonuclease



- When urea and 2-mercaptoethanol are removed, the protein spontaneously refolds (**self-assembly**), and the correct disulfide bonds are reformed
- The sequence alone determines the native conformation
- Quite "simple" experiment, but so important it earned Chris Anfinsen the 1972 Chemistry Nobel Prize

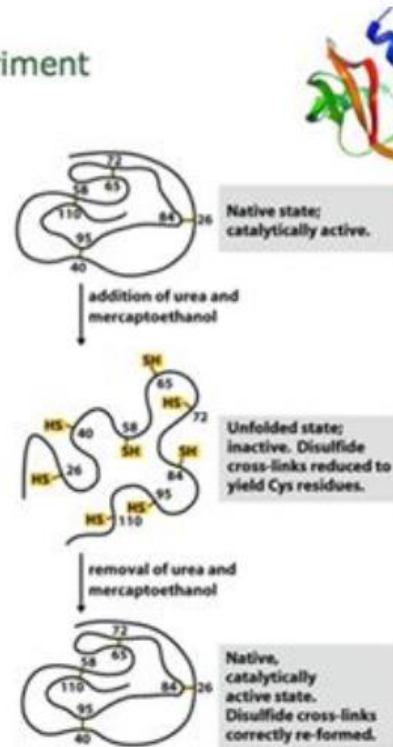


Figure 4-26
Lehninger Principles of Biochemistry, Fifth Edition

THE HYDROPHOBICITY

In a completely analogous fashion, a hydrophobic propensity or **hydropathy** can be calculated. In this system, empirical measures of the hydrophobic nature of the side chains are used to assign a number to a given amino acid. Many hydropathy scales are used. Several are based on the $\Delta\mu^\circ$ transfer of the side chains from water to a nonpolar solvent. Two commonly used scales are the **Kyte-Doolittle Hydropathy** and **Hopp-Woods scales** (used more like **hydrophilicity** index to predict surface or water accessible structures that might be recognized by the immune system)

PROTEIN ENGINEERING

Hydrophobicity Indices for Amino Acids

Amino Acid	Kyte-Doolittle	Hopp-Woods
Alanine	1.8	-0.5
Arginine	-4.5	3.0
Asparagine	-3.5	0.2
Aspartic acid	-3.5	3.0
Cysteine	2.5	-1.0
Glutamine	-3.5	0.2
Glutamic acid	-3.5	3.0
Glycine	-0.4	0.0
Histidine	-3.2	-0.5
Isoleucine	4.5	-1.8
Leucine	3.8	-1.8
Lysine	-3.9	3.0
Methionine	1.9	-1.3
Phenylalanine	2.8	-2.5
Proline	-1.6	0.0
Serine	-0.8	0.3
Threonine	-0.7	-0.4
Tryptophan	-0.9	-3.4
Tyrosine	-1.3	-2.3
Valine	4.2	-1.5

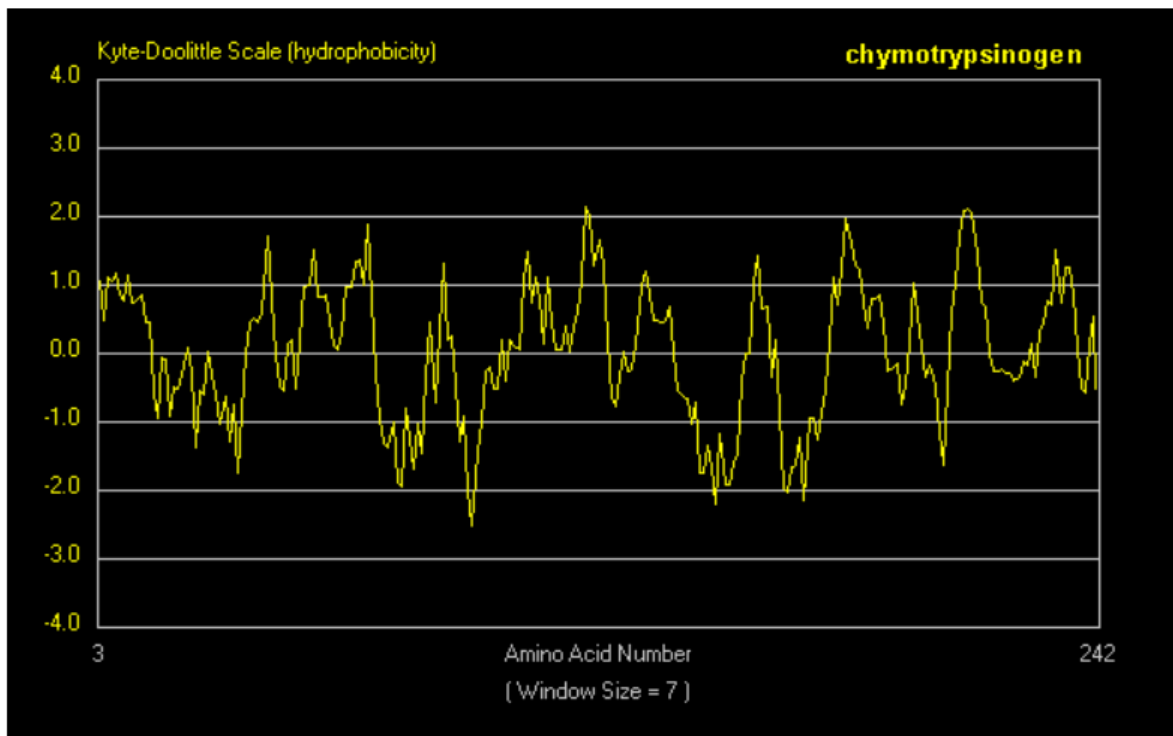
PROTEIN ENGINEERING

For a water-soluble protein, a continuous stretch of amino acids found to have a high average hydrophathy is probably buried in the interior of the protein. Consider the example of bovine α -chymotrypsinogen, a 245 amino acid protein, whose sequence is shown below in single letter code.

```
1 CGVPAIQPVLSGLSRIVNGEEAVPGSWPWQVSLQDKTGFHFCCGSLINENWVWVTAHCGV
61 TTSDVVVAGEFDQGSSEKIQLKIAKVKNSKYNLSLTINNDITLLKLSTAASFSQTVSA
121 VCLPSASDDFAAGTTCVTTGWGLTRYTNANTPDRLQQASLPLLSNTNCKKYWGTKIKDAM
181
ICAGASGVSSCMGDSGGPLVCKKNGAWTLVGIVSWGSSSTCSTSTPGVYARVTALVNWVQQ
241 TLAAN
```

A hydrophathy plot for chymotrypsinogen (sum of hydrophathies of seven consecutive residues) shows many stretches that are presumably buried in the interior of the protein.

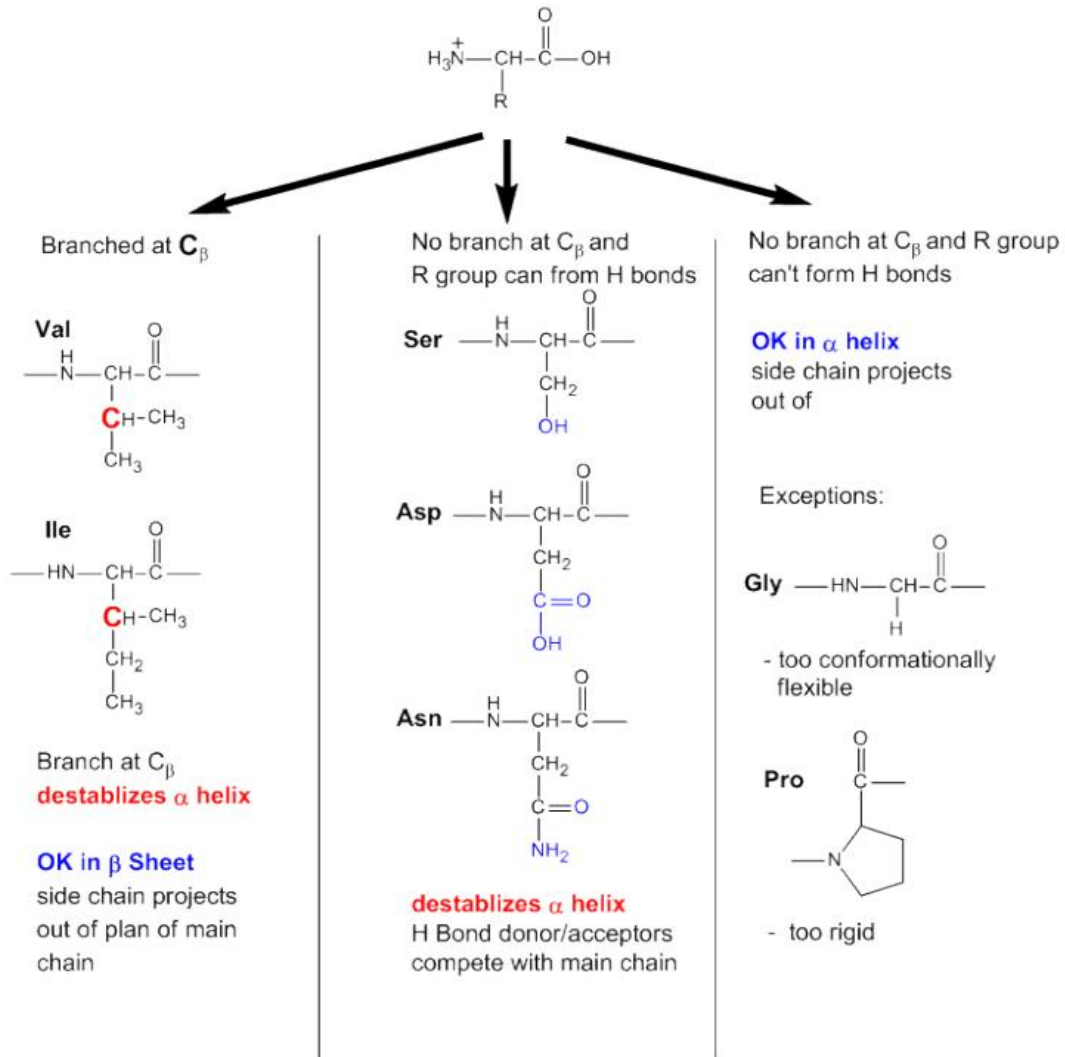
hydrophathy plot for chymotrypsinogen



PROTEIN ENGINEERING

AMINO-ACID PROPENSITIES

AMINO ACID PROPENSITIES FOR SECONDARY STRUCTURE



As we have seen previously, amino acids vary in their propensity to be found in alpha helices, beta strands, or reverse turns (beta bends, beta turns). These difference can be rationalized from the structure of each amino acid, as described before.

Table of standard amino acid alpha-helical propensities

Estimated differences in free energy, $\Delta(\Delta G)$, estimated in kcal/mol per residue in an alpha-helical configuration, relative to Alanine arbitrarily set as zero. Higher numbers (more positive free energies) are less favoured. Significant deviations from these average numbers are possible, depending on the identities of the neighbouring residues

PROTEIN ENGINEERING

Amino Acid	3-Letter	1-Letter	Helical Penalty (<u>Kcal/mol</u>)
<u>Alanine</u>	Ala	A	0
<u>Arginine</u>	Arg	R	0.21
<u>Asparagine</u>	Asn	N	0.65
<u>Aspartic acid</u>	Asp	D	0.69
<u>Cysteine</u>	Cys	C	0.68
<u>Glutamic acid</u>	Glu	E	0.40
<u>Glutamine</u>	Gln	Q	0.39
<u>Glycine</u>	Gly	G	1.00
<u>Histidine</u>	His	H	0.61
<u>Isoleucine</u>	Ile	I	0.41
<u>Leucine</u>	Leu	L	0.21
<u>Lysine</u>	Lys	K	0.26
<u>Methionine</u>	Met	M	0.24
<u>Phenylalanine</u>	Phe	F	0.54
<u>Proline</u>	Pro	P	3.16
<u>Serine</u>	Ser	S	0.50
<u>Threonine</u>	Thr	T	0.66
<u>Tryptophan</u>	Trp	W	0.49
<u>Tyrosine</u>	Tyr	Y	0.53
<u>Valine</u>	Val	V	0.61

PROTEIN ENGINEERING

Chou-Fasman is one of most commonly used algorithms

- Based on observation that certain amino acids tend to be enriched (or depleted) in different secondary structure classes
- The Chou, Fasman and co-workers calculated the propensity of each amino acid to adopt an alpha helix, beta-strand, or coil conformation (later turn as well)
- Propensity values (P_α or P_β) were calculated using a database of experimentally determined protein structures (first database had 15 proteins)
- Propensity for an amino acid of type i to form an α -helix (P_α^i) is calculated using the following formula:

$$P_\alpha^i = \frac{\text{fraction of residues of type } i \text{ in } \alpha \text{-helix}}{\text{fraction of all residues in } \alpha \text{-helix}}$$

Example calculation for α -helix propensity (P_α^i) for alanine (A):

- Data from protein structure database (29 proteins):
- Number of alanine residues in the database = 434
- Number of alanine that occur in alpha helix = 234
- Number of all residues in the database = 4741
- Number of all residues that occur in alpha helix = 1798

$$P_\alpha^A = \frac{234 / 434}{1798 / 4741} = \frac{0.539}{0.379} = 1.42$$

- $P_\alpha^i > 1$ = helix former; $P_\alpha^i \sim 1$ = indifferent; $P_\alpha^i < 1$ = helix breaker
- Beta strand propensities are calculated in a similar way:

$$P_\beta^i = \frac{\text{fraction of residues of type } i \text{ in beta strand}}{\text{fraction of all residues in beta strand}}$$

§ $P_\beta^i > 1$ = strand former; $P_\beta^i \sim 1$ = indifferent; $P_\beta^i < 1$ = strand breaker

PROTEIN ENGINEERING

The Chou-Fasman parameters for the 20 common amino acids.

P_α		P_β	
E	1.51	V	1.70
M	1.45	I	1.60
A	1.42	Y	1.47
L	1.21	F	1.38
K	1.16	W	1.37
F	1.13	L	1.30
Q	1.11	C	1.19
W	1.08	T	1.19
I	1.08	Q	1.10
V	1.06	M	1.05
D	1.01	R	0.93
H	1.00	N	0.89
R	0.98	H	0.87
T	0.83	A	0.83
S	0.77	S	0.75
C	0.70	G	0.75
Y	0.69	K	0.74
N	0.67	P	0.55
P	0.57	D	0.54
G	0.57	E	0.37

$\left. \begin{array}{l} E \\ M \\ A \\ L \end{array} \right\} H_\alpha$	$\left. \begin{array}{l} V \\ I \\ Y \end{array} \right\} H_\beta$
$\left. \begin{array}{l} K \\ F \\ Q \\ W \\ I \\ V \end{array} \right\} h_\alpha$	$\left. \begin{array}{l} F \\ W \\ L \\ C \\ T \\ Q \\ M \end{array} \right\} h_\beta$
$\left. \begin{array}{l} D \\ H \end{array} \right\} I_\alpha$	$\left. \begin{array}{l} R \\ N \\ H \end{array} \right\} i_\beta$
$\left. \begin{array}{l} R \\ T \\ S \end{array} \right\} i_\alpha$	$\left. \begin{array}{l} A \\ S \\ G \end{array} \right\} b_\beta$
$\left. \begin{array}{l} Y \\ N \end{array} \right\} b_\alpha$	$\left. \begin{array}{l} K \\ P \end{array} \right\} B_\beta$
$\left. \begin{array}{l} P \\ G \end{array} \right\} B_\alpha$	$\left. \begin{array}{l} D \\ E \end{array} \right\} B_\beta$