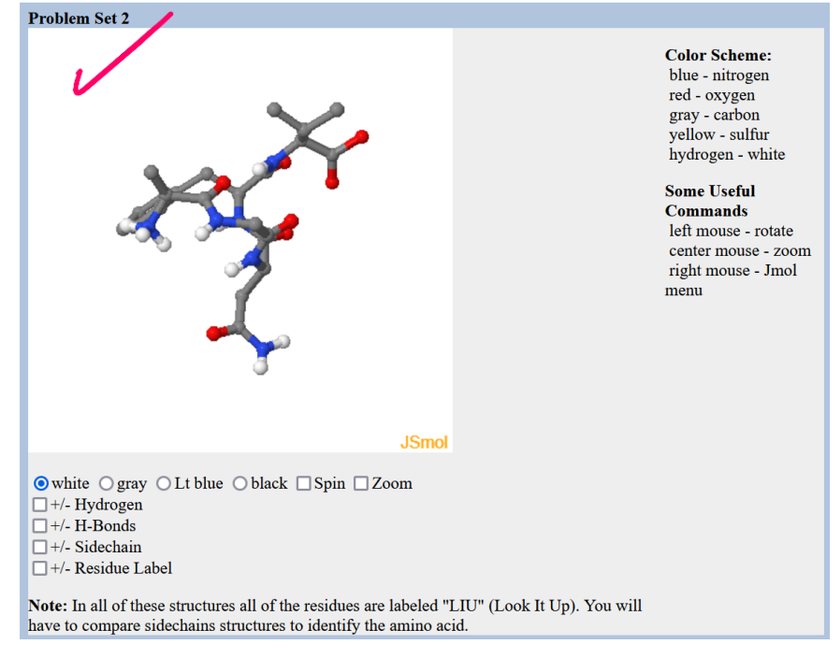
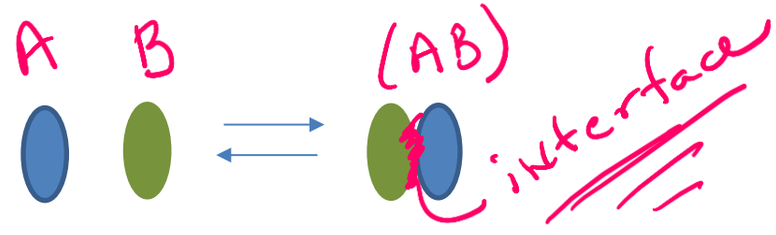
# Lecture 2
# Protein Structure and Function, Carbohydrates

- Protein Structure and Stability
- Ligand Binding
- Proteins as enzymes (PKU disease)
- Carbohydrates

Prior to lecture 3, please review the lecture material on introduction to nucleic acids (slides and video link posted on course web site)



# Molecular Interactions




$$\Delta E = E_{AB} - (E_A + E_B) \sim E_{A \cup B}$$

The energy change when two things come together can be approximated to be due to ***new*** inter-molecular interactions: $E_{A \cup B}$.

| Interaction | Interaction | Energy (kJ/mol) |
|---|---|---|
| Electrostatic interactions (in water) | Full charges | ~5 kJ/mol/single interaction |
| Van der Waals: Dipole-Dipole | Perm. partial charges | ~0.05 kJ/A$^2$ x 100 A$^2$ = 5 kJ/mol for 100 A$^2$ |
| Van der Waals: Induced-dipole | Induced partial charges | ~0.02 kJ/A$^2$ x 100 A$^2$ = **2 kJ/mol for 100 A$^2$** |
| H-Bonds | Electrostatic + e sharing | ~20 kJ/mol gross, **~5 kJ/mol net** |

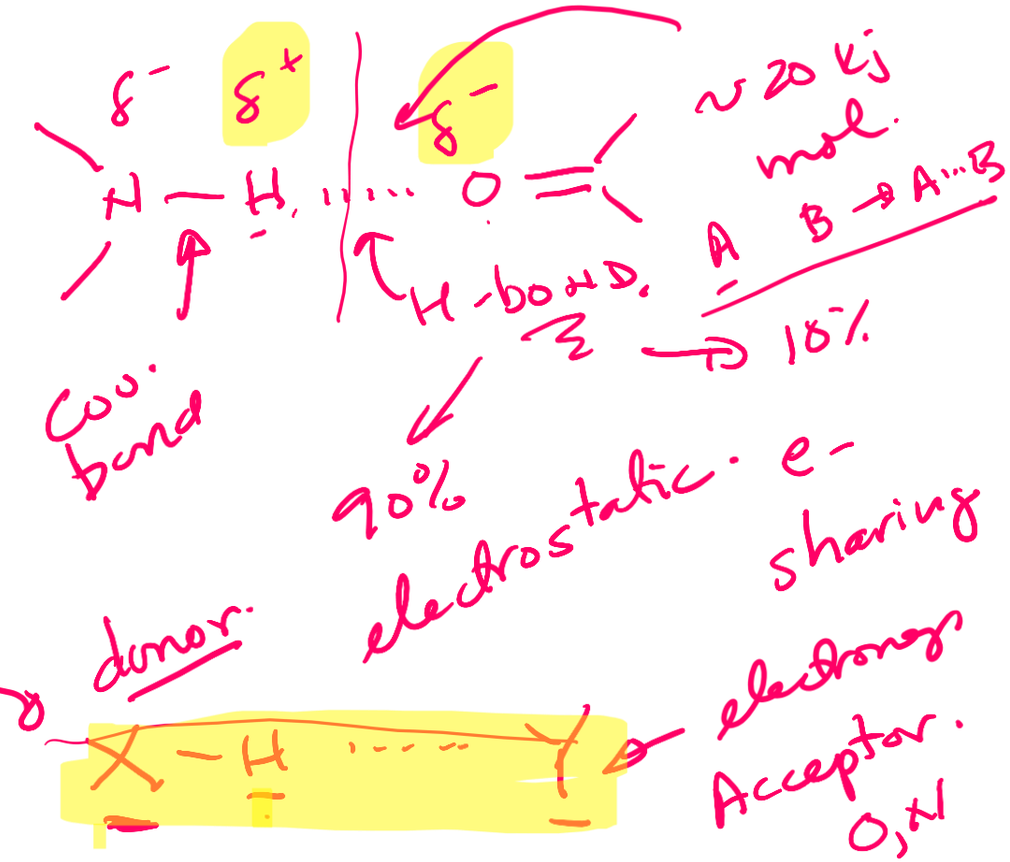
## iv) Hydrogen Bonds

- H-bonds are primarily (90%) an electrostatic attraction between:
  - Electropositive hydrogen, attached to an electronegative atom is the hydrogen bond donor (i.e. NH).
  - Electronegative hydrogen bond acceptor (e.g. the lone pairs of oxygen, or C=O group of an amide).

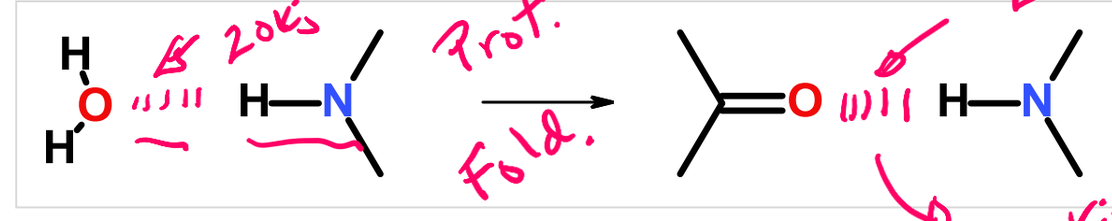
*A "bond" implies electron sharing – about 10% of the electron is shared from one molecule to the next in the case of H-bonds*

*Note that the proton is **NOT** transferred to the acceptor, it remains covalently bonded to the donor atom. The Hydrogen Bond is the **interaction** between the X-H donor and electronegative acceptor.*

| | $\lvert\Delta e\rvert$ |
|---|---|
| N(3.0)——H(2.1) | 0.9 |
| O(3.5)——H(2.1) | 1.4 |



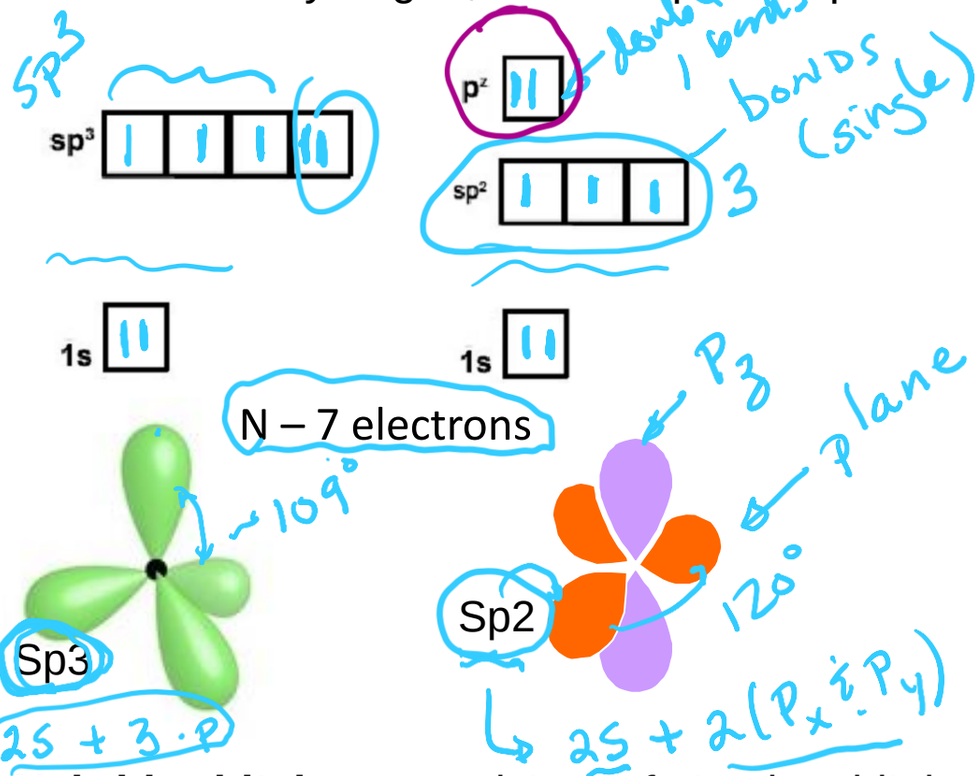

- The energy released when an H-bond forms depends on the distance and angle of the bond.
- Usually hydrogen bonds are exchanged, resulting in small ***net*** energy differences:



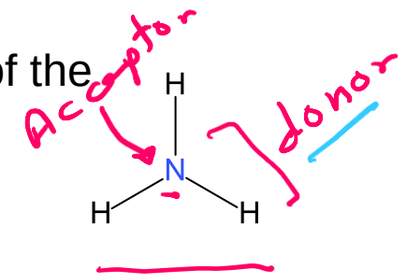
# How to Identify Hydrogen Bond Donor and Acceptors

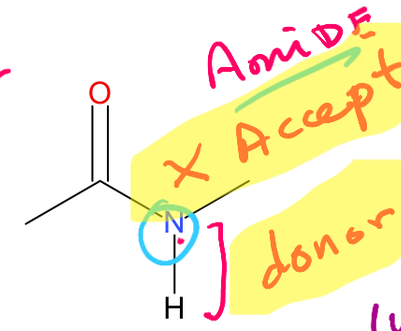
Exceptions, N in a delocalized system:

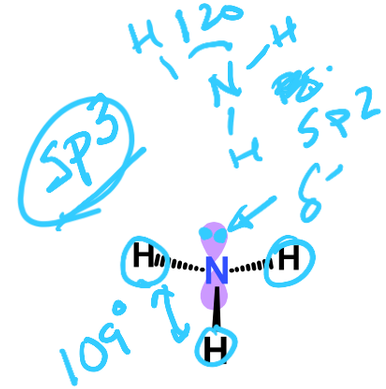
- Will not accept from above or below the plane of the system, because the lonepair is delocalized.
- Can accept in the plane of the ring if there is no attached hydrogen, via lone pair in sp2 orbital

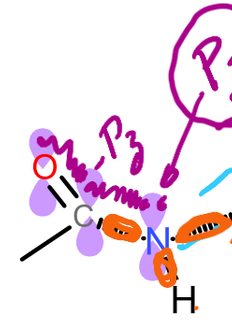



**Hybrid orbitals** are a mixture of atomic orbitals. Nitrogen can form two types of hybrid orbitals, sp3 (tetrahedral geometry) or sp2 (planer) + pz



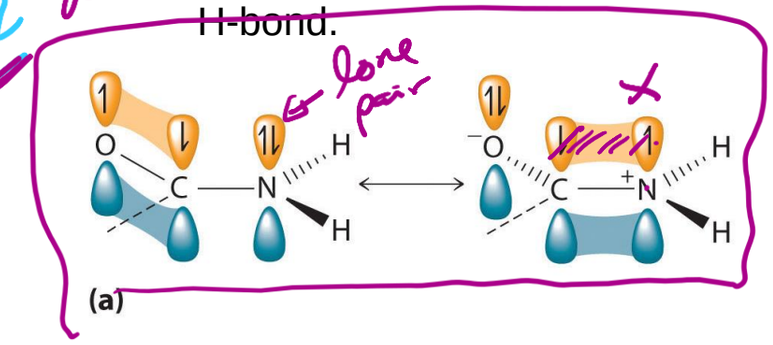



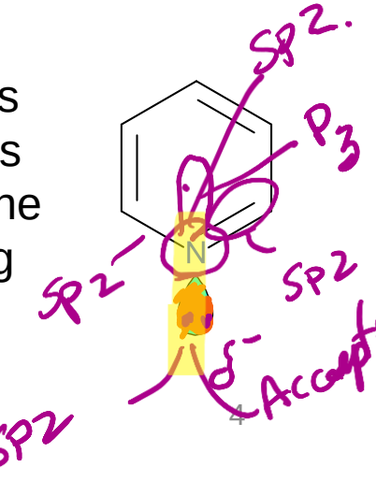




- Sp3 is used in ammonia, keeping the three hydrogen atoms as far from each other as possible. The fourth sp3 orbital is full with two electrons (lone pair).
- The lone pair is an excellent acceptor.

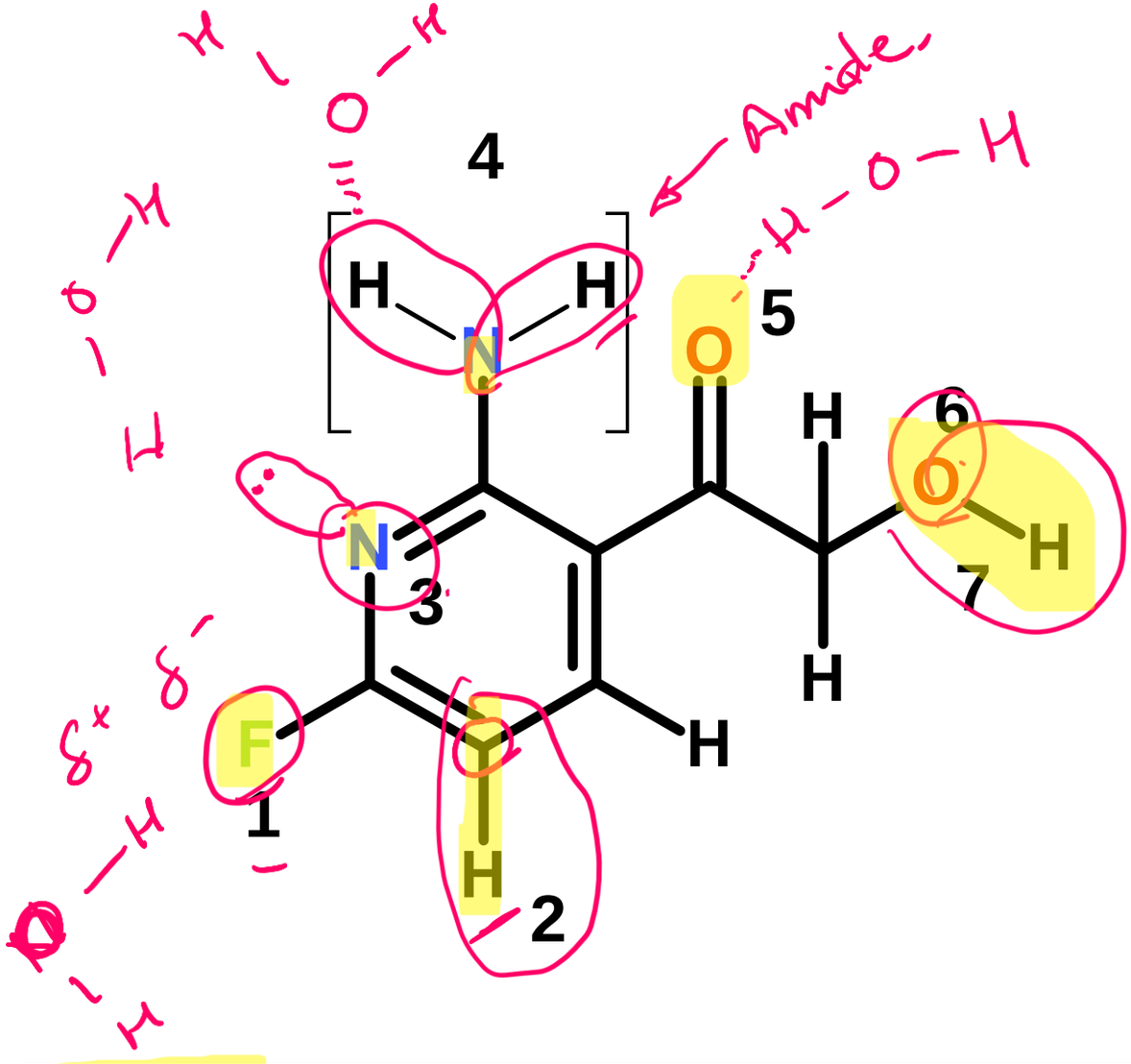
- Sp2 is used in amides, allowing favorable overlap of the full pz orbital with the pz on C and O
- The lone pair in the pz is shared with the pz electrons on carbon and oxygen.
- Due to electron sharing, there is only a slight neg. charge and the group does not accept an H-bond.




- Nitrogen in rings with three bonds can **accept** in the plane of the ring due to a filled lonepair orbital.



# 1. Indicate which atoms could donate an H-bond and which could accept and H-bond




| ATOM | Donor(D)? | Acceptor(A)? | Neither (N) |
|---|---|---|---|
| 1 (F) | | | |
| 2 ($C_{aro}$-H) | | | |
| 3 (N) | | | |
| 4 ($-NH_2$) | | | |
| 5 (C=O) | | | |
| 6 (O) | | | |
| 7 (O-H) | | | |

**Can you?**

- Identify groups that can donate or accept hydrogen bonds?

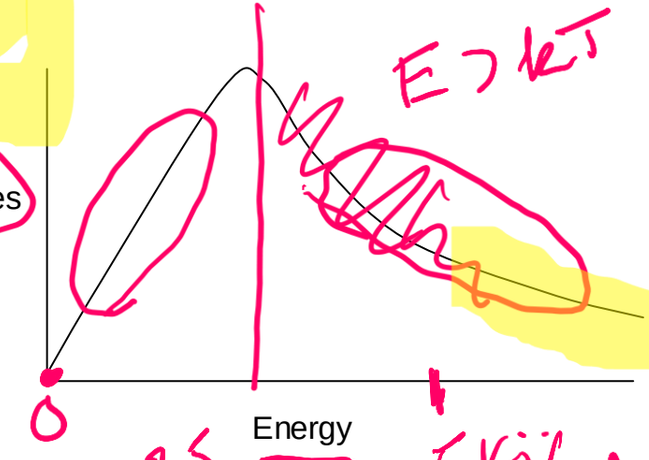
# Relative Energy of Interactions

| Interaction | Interaction | Energy (kJ/mol) |
|---|---|---|
| **Covalent Bond** | **Electron sharing** | **200-400 kJ/mol** |
| Electrostatic interactions (in water) | Full charges | ~5 kJ/mol/single interaction |
| VdW - Dipole-dipole (Keesom) | Perm. partial charges | ~0.05 kJ/A$^2$ x 100 A$^2$ = 5 kJ/mol for 100 A$^2$ |
| VdW – Induced dipole (London) | Induced partial charges | ~0.02 kJ/A$^2$ x 100 A$^2$ = **2 kJ/mol for 100 A$^2$** |
| H-Bonds | Electrostatic + e sharing | ~20 kJ/mol gross, **~5 kJ/mol net** |

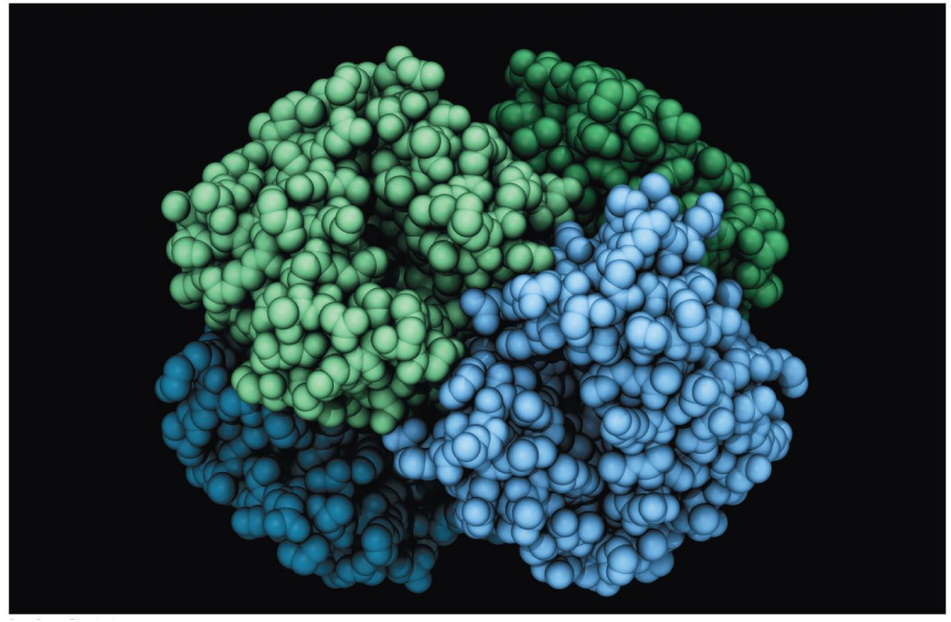
1. *How does the energy of the last four interactions compare to covalent bonds?*
   1. *Stronger* 2. *Weaker* 3. *The Same*

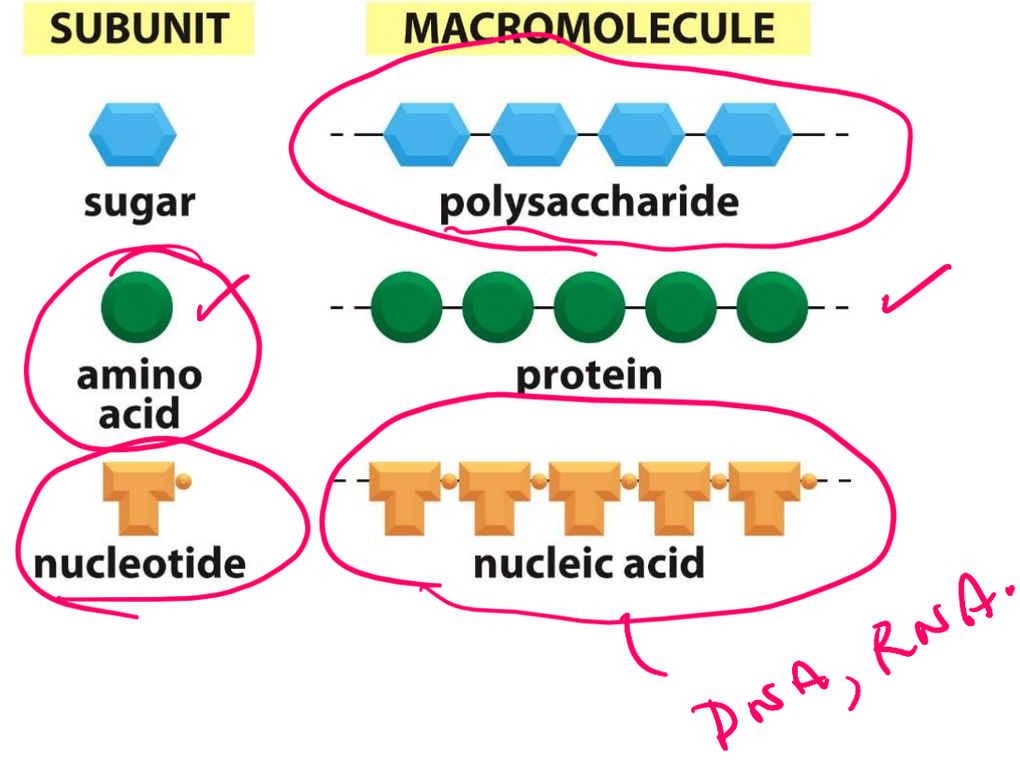
2. *Which of these are closer to thermal energy, kT = 2.5 kJ/mol @ room temp.*

3. *Are there significant molecules with enough energy at room temperature to break the interaction?*



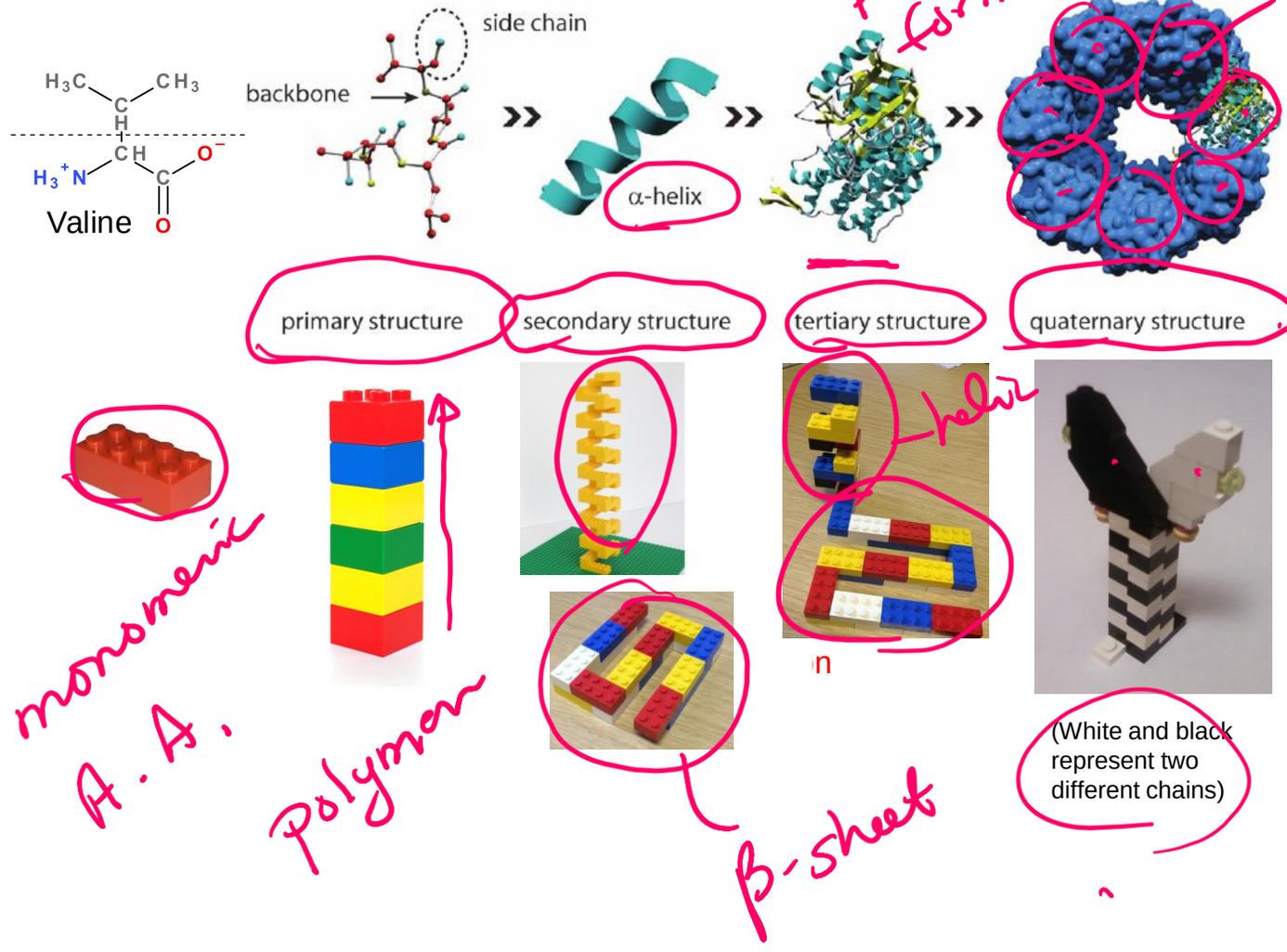
# Proteins and Amino Acids

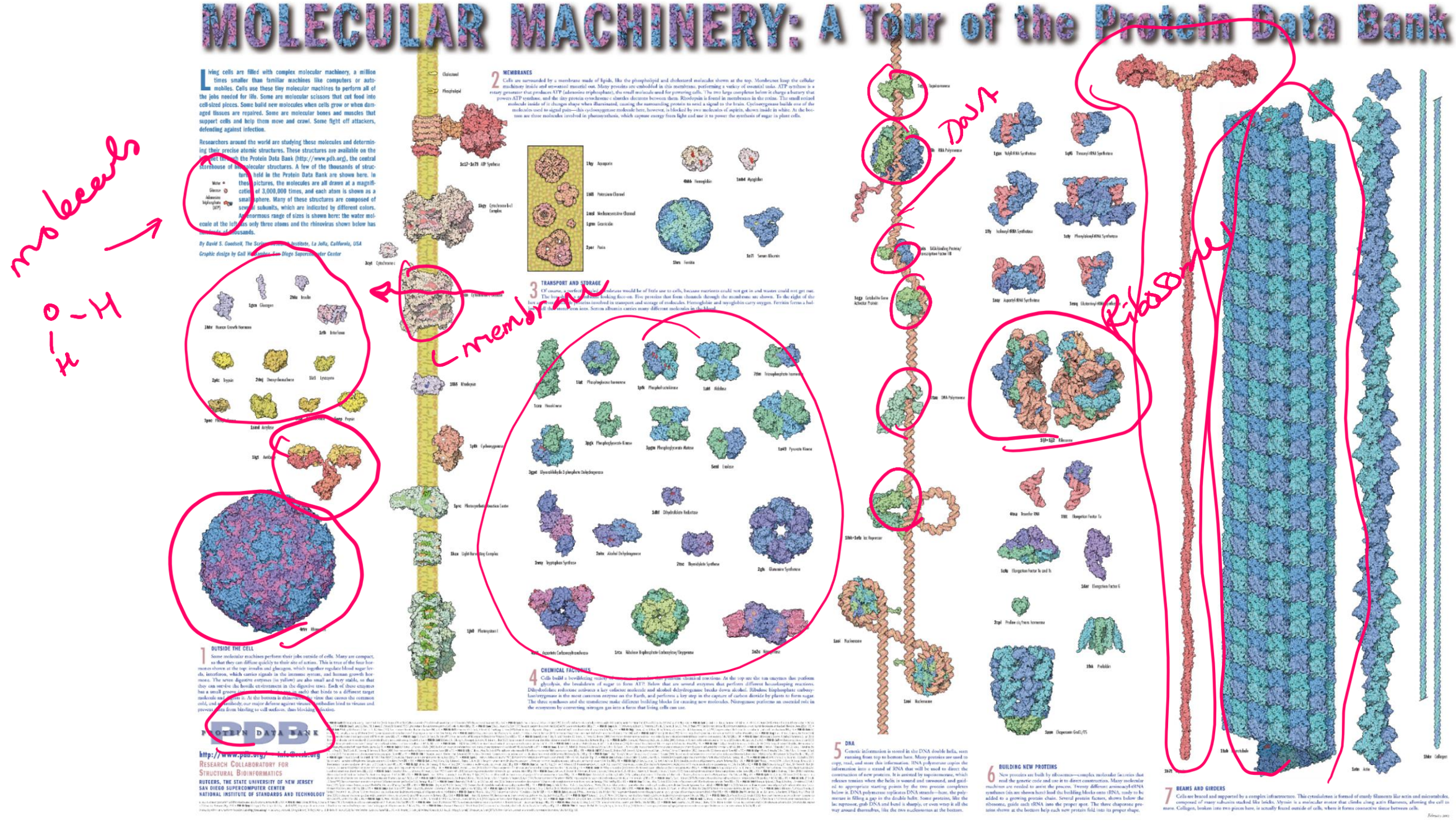




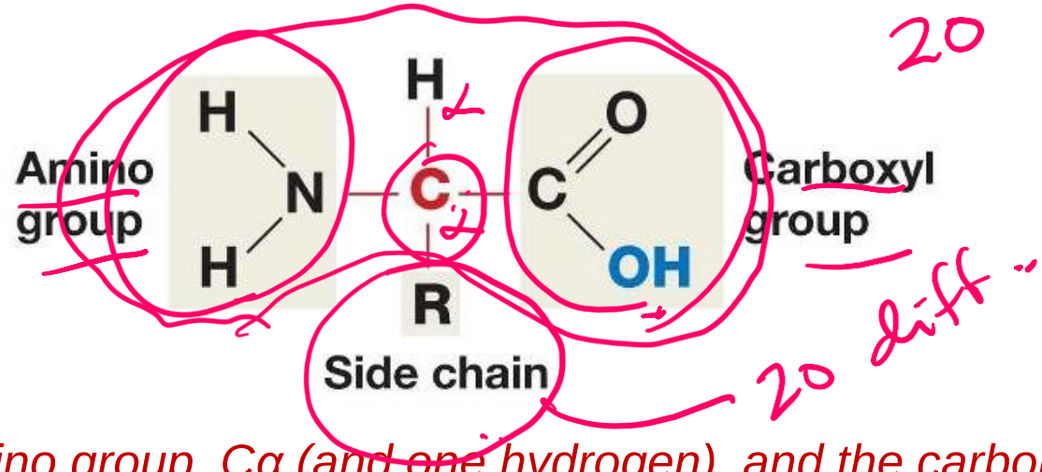
# Structural Hierarchy of Proteins

- Primary - sequence of amino acids, no 3D structural information
- Secondary - local structural elements, only mainchain atoms involved
- Tertiary - 3D position of ***all*** atoms, functional form of many proteins.
- Quaternary - multiple chains – multiple chains often required for function.

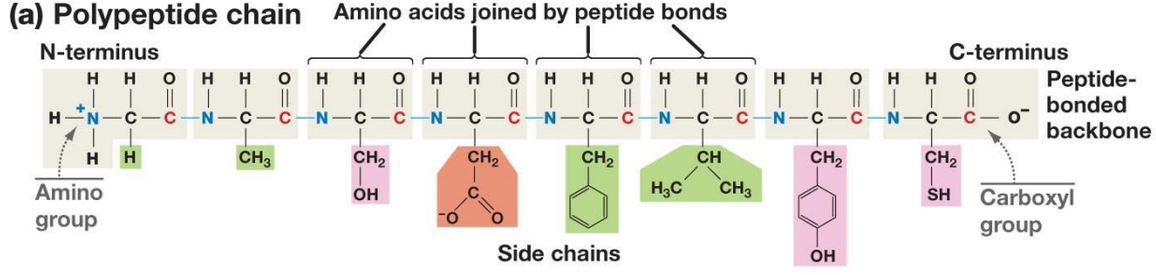



MOLECULAR MACHINERY: A Tour of the Protein Data Bank
Living cells are filled with complex molecular machinery, a million times smaller than familiar machines like computers or automobiles. Cells use these tiny molecular machines to perform all of the jobs needed for life. Some are molecular scissors that cut food into cell-sized pieces. Some build new molecules when cells grow or when damaged tissues are repaired. Some are molecular bones and muscles that support cells and help them move and crawl. Some fight off attackers, defending against infection.
Researchers around the world are studying those molecules and determining their precise atomic structures. These structures are available on the Internet through the Protein Data Bank (http://www.pdb.org), the central warehouse of biomolecular structures. A few of the thousands of structures held in the Protein Data Bank are shown here. In these pictures, the molecules are all drawn at a magnification of 3,000,000 times, and each atom is shown as a small sphere. Many of these structures are composed of several subunits, which are indicated by different colors. An enormous range of sizes is shown here: the water molecule at the left has only three atoms and the rhinovirus shown below has hundreds of thousands.
By David S. Goodsell, The Scripps Research Institute, La Jolla, California, USA
Graphic design by Gail Hodgkinson, San Diego Supercomputer Center
1 OUTSIDE THE CELL
2 MEMBRANES
3 TRANSPORT AND STORAGE
4 CHEMICAL FACTORIES
5 DNA
6 BUILDING NEW PROTEINS
7 BEAMS AND GIRDERS
PROTEIN DATA BANK
http://www.pdb.org/
RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS
RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
SAN DIEGO SUPERCOMPUTER CENTER
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

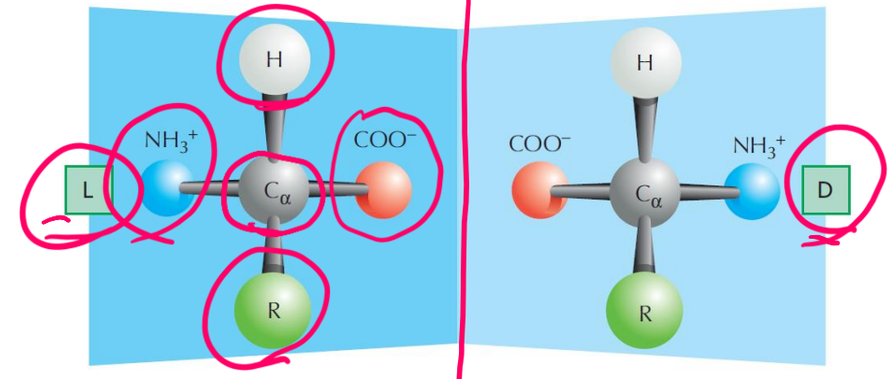
# The Structure of Amino Acids and Proteins




- *The amino group, Cα (and one hydrogen), and the carbonyl group are common to all amino acids*
- *The N-Cα-C=O are the mainchain of the protein polymer.*
- *The R groups are different –there are 20 common R groups they are the sidechain of the protein polymer – their* ***sequence*** *defines the properties of the protein.*




*Is there a chiral carbon on amino acids?*




Proteins consist exclusively of L-amino acids.

(as does the ribosome that make them)

# Primary Structure

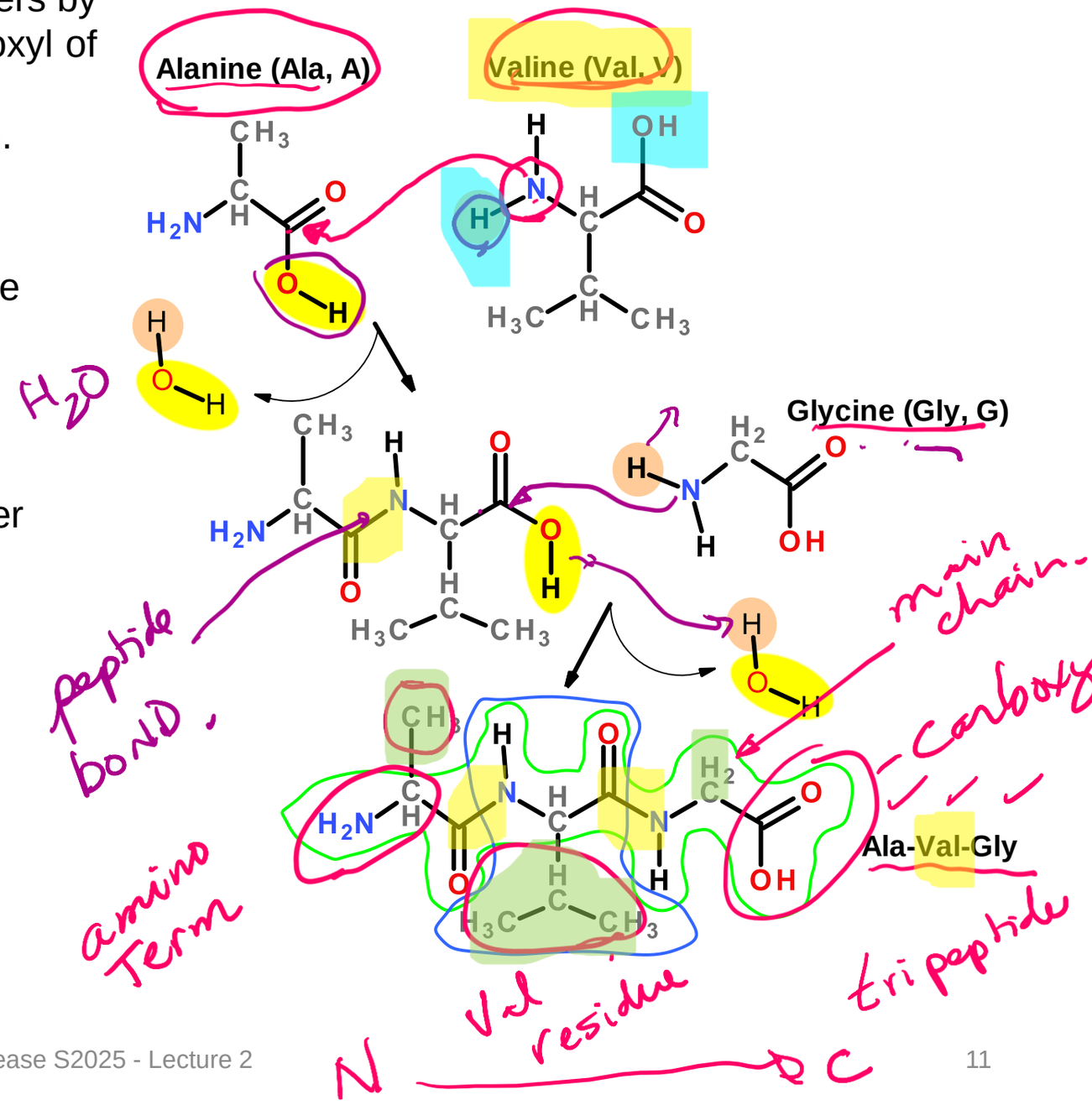
- Amino acids are joined together to form linear polymers by the formation of a **peptide bond** between the carboxyl of one amino acids and the amino group of the next.
- This reaction releases water: a **dehydration** reaction.
- The peptide bond can be broken (*lysis*) by the addition of water = **hydrolysis**.

Incorporated amino acid is called a *residue* (atoms are lost when the peptide bond is formed).

Polarity of chain direction – amino (N) terminus to carboxy(C) terminus = order of amino acids = *sequence* = *primary structure*

*Mainchain* (or backbone) – linear atoms of the polymer

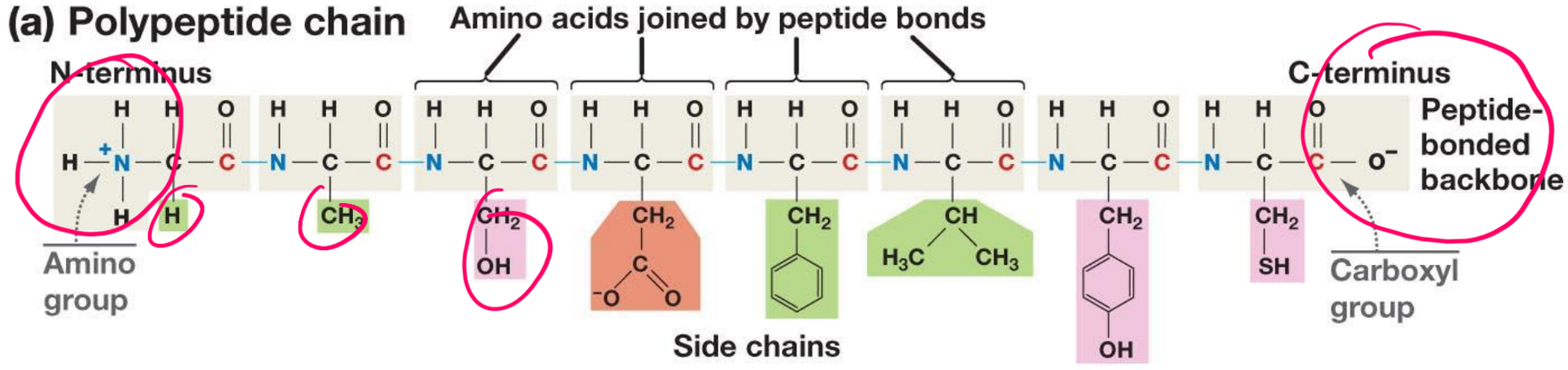
*Sidechain* – atoms off the Ca carbon




***Primary Structure – Expectations***

- Draw chemical structure given the sequence.
- Determine the seq. from chemical structure.
- Distinguish/identify:
  - **Mainchain** & **Sidechain** atoms,
  - **Residue** = aa in polymer,
  - N & C terminus,
  - Peptide bond(s).

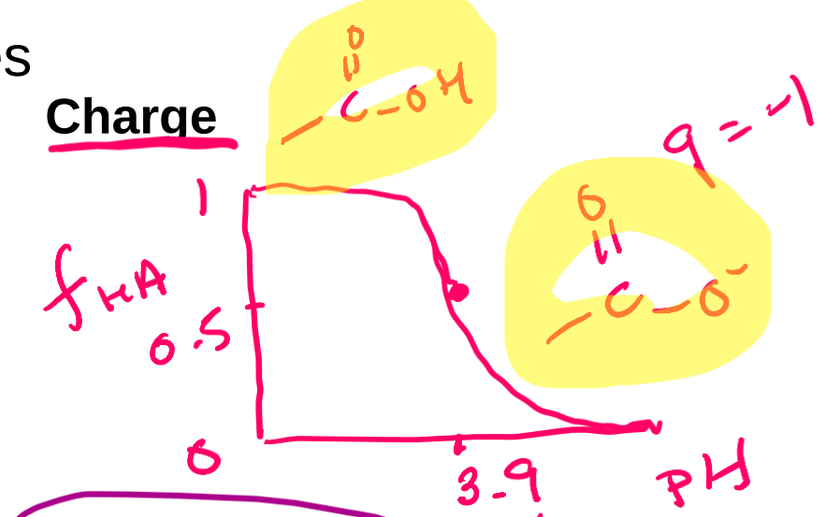
# Sidechain *Functional* Groups Affect Behavior (and the order is important)



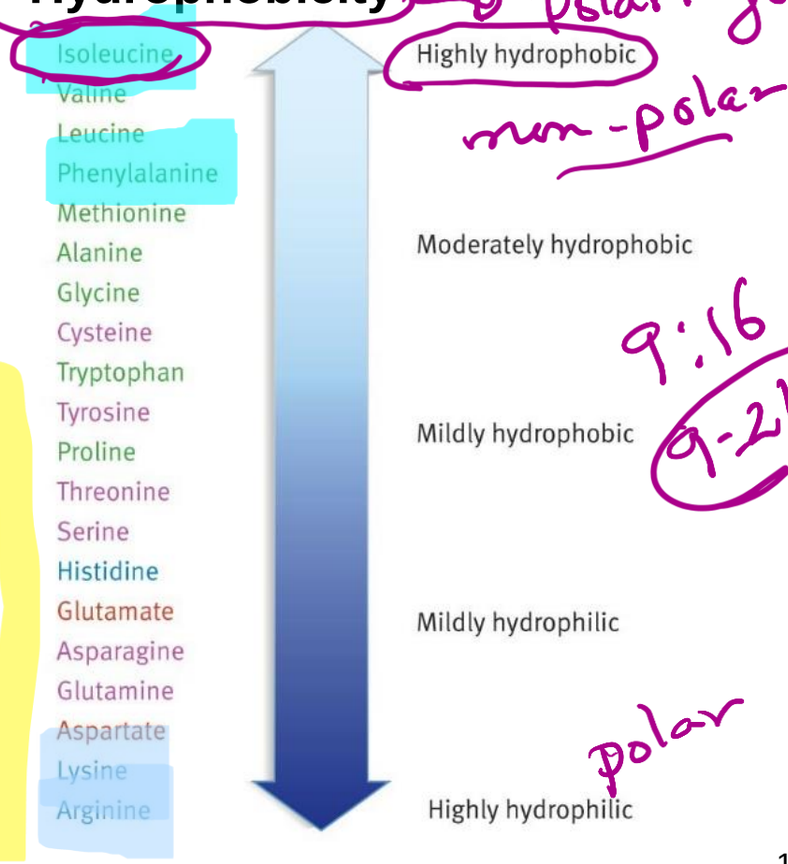

- Sidechains (R-groups) differ in their size, shape, reactivity, and interactions with water.
  1. Nonpolar Sidechains: **hydrophobic**; do not form hydrogen bonds; coalesce in water - typically form the core of folded proteins.
  2. Polar Sidechains: **hydrophilic**; form hydrogen bonds; readily dissolve in water
  3. **Ionizable** Sidechains: Can be charged at certain pH values. Interact strongly with water.

# Amino Acids – Structure and Properties

## Charge

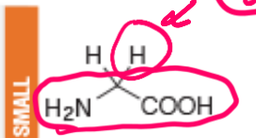



## Hydrophobicity

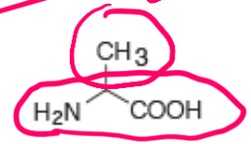



## Non-polar sidechains

SMALL

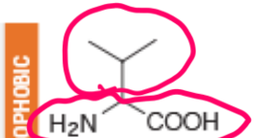


Glycine (Gly, G)
MW: 75.07

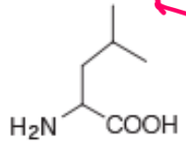


Alanine (Ala, A)
MW: 89.09

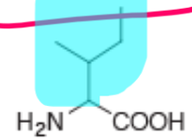
HYDROPHOBIC



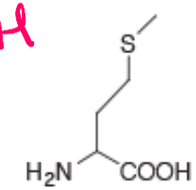
Valine (Val, V)
MW: 117.1



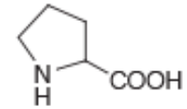
Leucine (Leu, L)
MW: 131.2



Isoleucine (Ile, I)
MW: 131.2



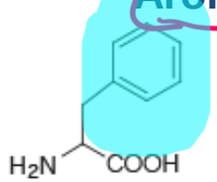
Methionine (Met, M)
MW: 149.2



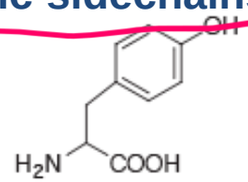
Proline (Pro, P)
MW: 115.1

## Aromatic sidechains

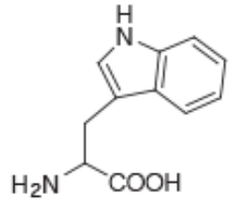
AROMATIC



Phenylalanine (Phe, F)
MW: 165.2



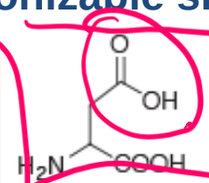
Tyrosine (Tyr, Y)
MW: 181.2, $pK_a = 10.46$



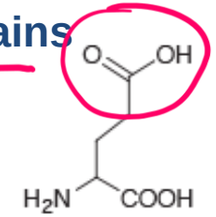
Tryptophan (Trp, W)
MW: 204.2

## Ionizable sidechains

ACIDIC



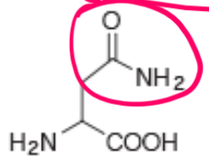
Aspartic Acid (Asp, D)
MW: 133.1, $pK_a = 3.9$



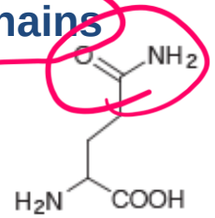
Glutamic Acid (Glu, E)
MW: 147.1, $pK_a = 4.07$

## Polar sidechains

AMIDE



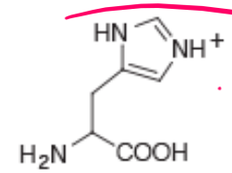
Asparagine (Asn, N)
MW: 132.1



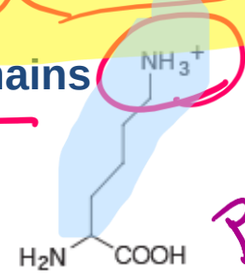
Glutamine (Gln, Q)
MW: 146.1

## Ionizable sidechains

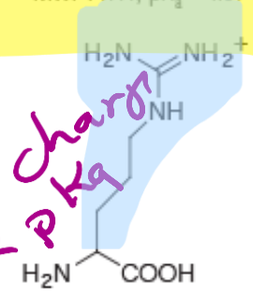
BASIC



Histidine (His, H)
MW: 155.2, $pK_a = 6.04$

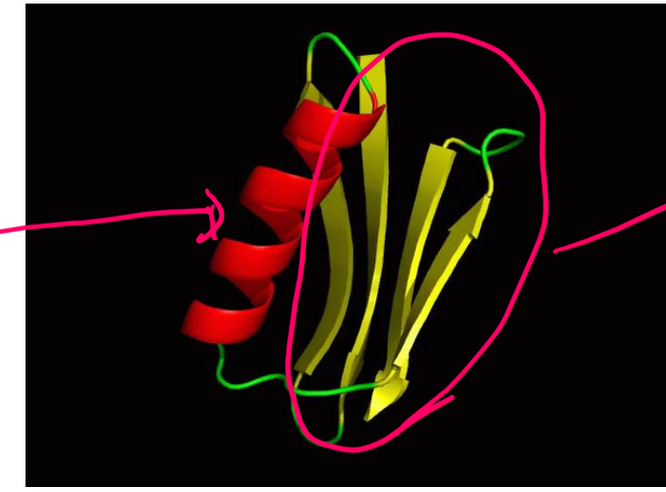


Lysine (Lys, K)
MW: 146.2, $pK_a = 10.79$



Arginine (Arg, R)
MW: 174.2, $pK_a = 12.48$

# Secondary Structure



"Building blocks of proteins"

- **Hydrogen bonds** between the *mainchain* carbonyl group of one amino acid and the *mainchain* amino group of another form a protein's **secondary structure**.
  - A polypeptide must bend to allow this hydrogen bonding, forming:
    - **α-helices**
    - **β-pleated sheets**

- The large number of hydrogen bonds in a protein's secondary structure increases its stability - each hydrogen bond that is formed releases some energy.
- All amino acids can be incorporated into either secondary structure

  (However, some are found more frequently in one structure)

**General Rule for Hydrogen Bonds:**

**X-H Y**

X & Y are electronegative (N and O usually)

**X-H = *Donor*** of the hydrogen bond

**Y = *Acceptor*** of the hydrogen bond

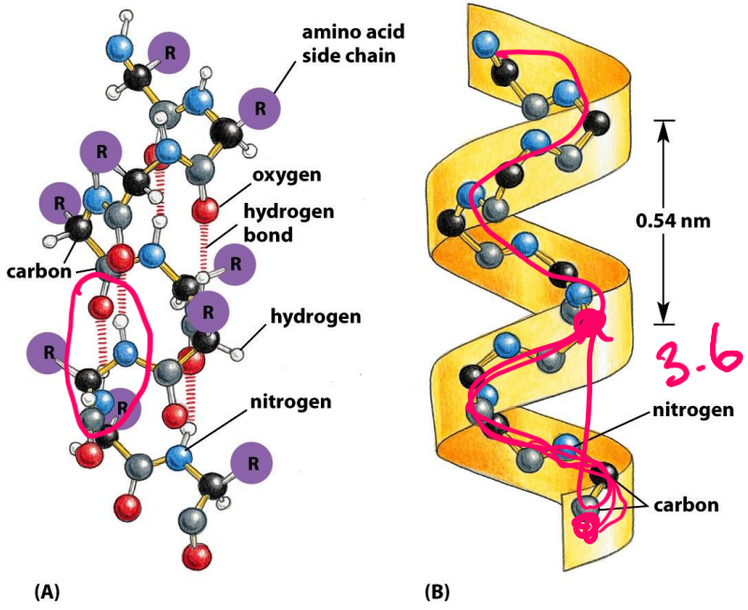
Mainchain hydrogen bonds

N-H ····O=C

The NH is the hydrogen bond_____.

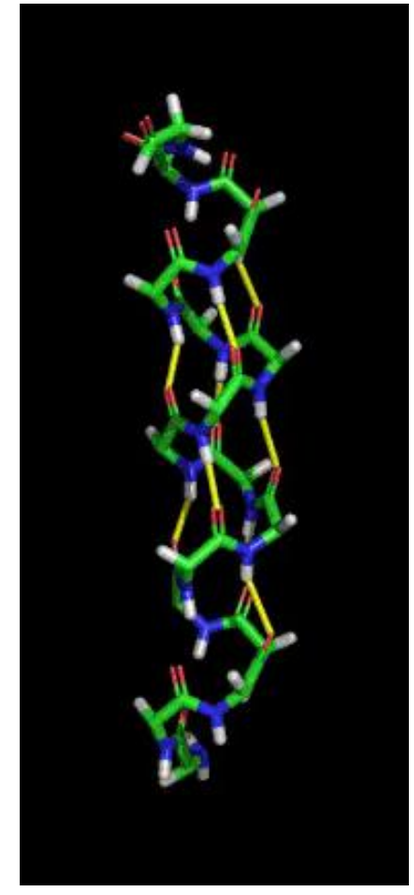
The C=O is the hydrogen bond _____.

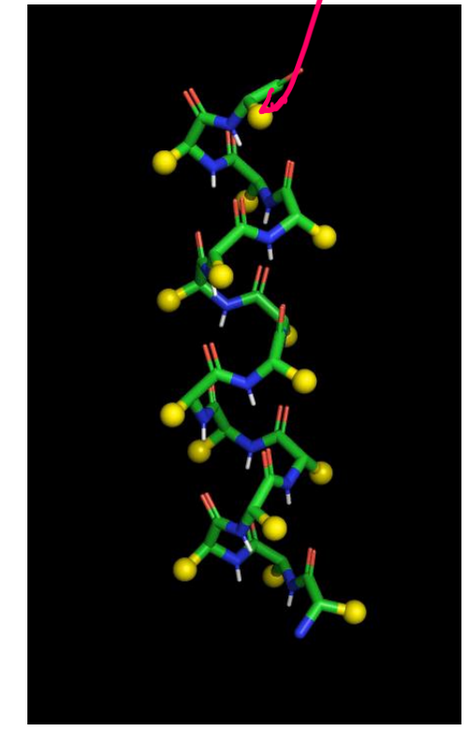
# Alpha Helix




H.bonds.

SIDE chains



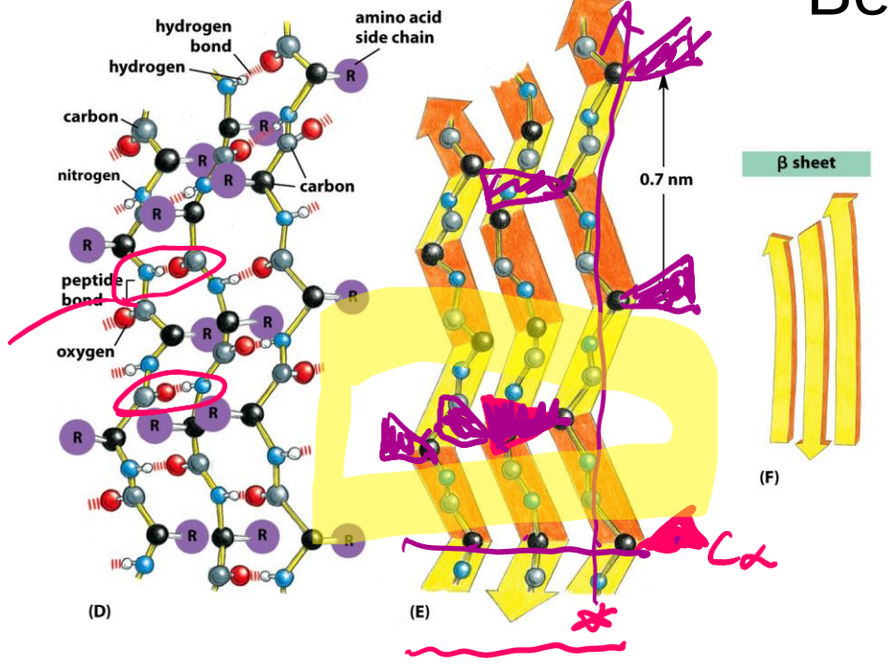


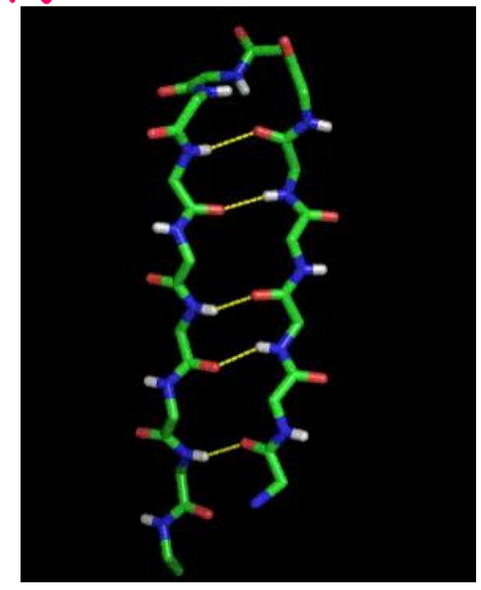
Spiral conformation (*helix*) in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier:

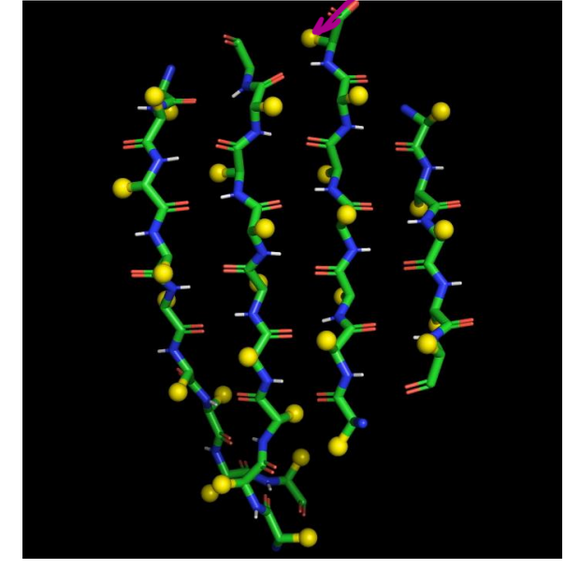
✓ *Intra-strand H-bonds, parallel to helix axis.*

✓ *Side-chains project outwards.*

# Beta Sheet

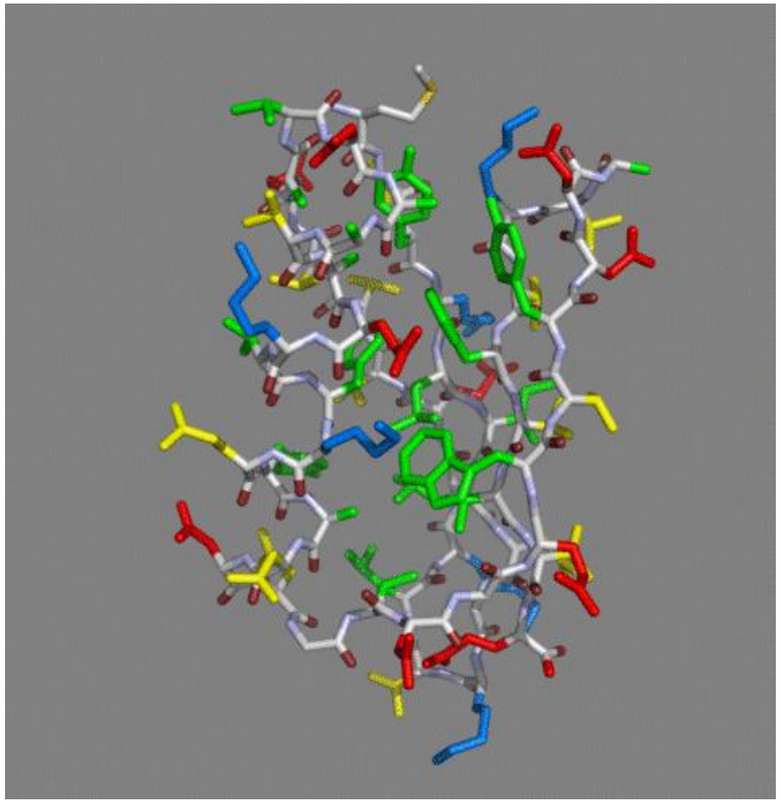




- Beta-Strands connected laterally by backbone hydrogen bonds that are perpendicular to the strand, forming a generally twisted, pleated sheet.
- Sheets can have two or more strands
- Side-chains:
  - project up and down along a strand.
  - project in the same direction going from strand to strand across the sheet.

# Tertiary Structure - Location of Residues in Globular Proteins

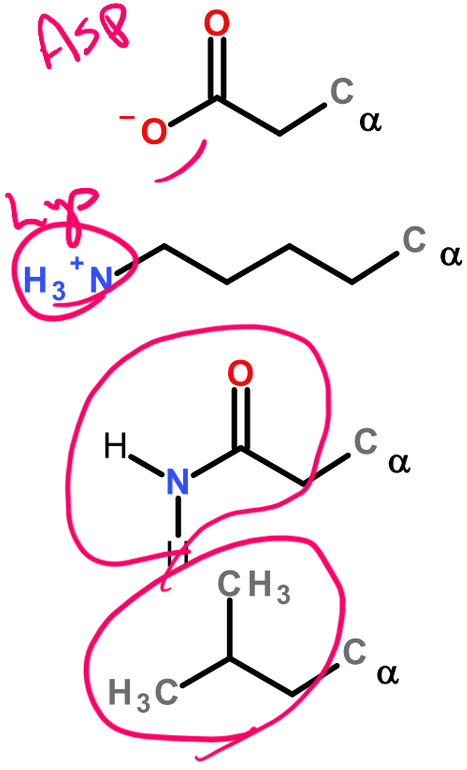


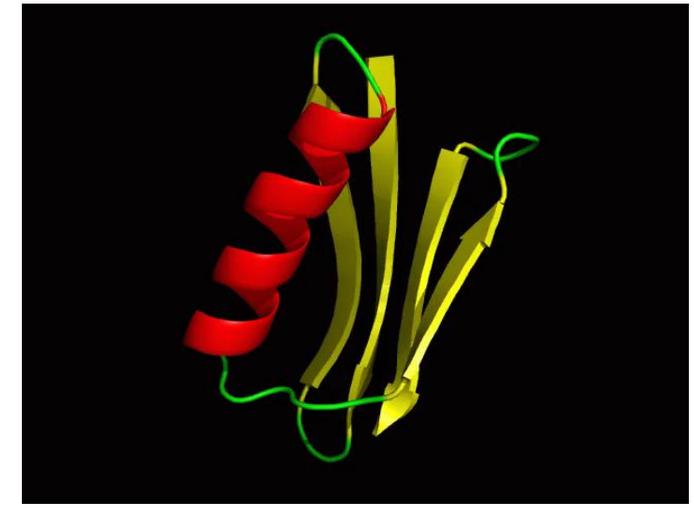
Red - amino acids with neg. sidechains (e.g. Asp)

Blue - amino acids with pos. sidechains (e.g. Lys)

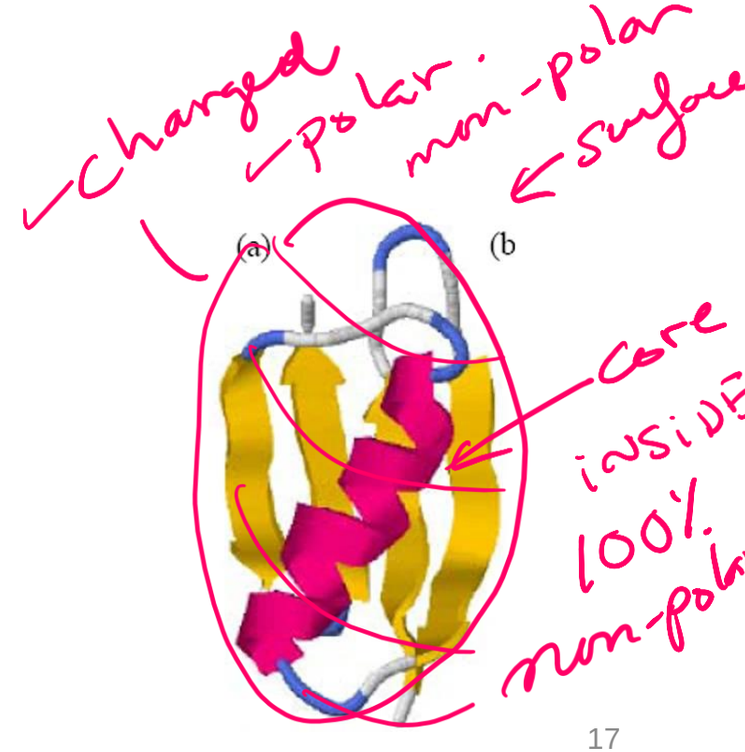
Yellow – amino acids with polar sidechain (e.g. Asn)

Green - amino acids with hydrophobic side chains (e.g. Leu)



| Amino Acid Type | | |
|---|---|---|
| | Inside (I) Core | Surface (S) |
| Charged | | |
| Polar | | |
| Non-polar | 100 % | also |



# Protein Stability:

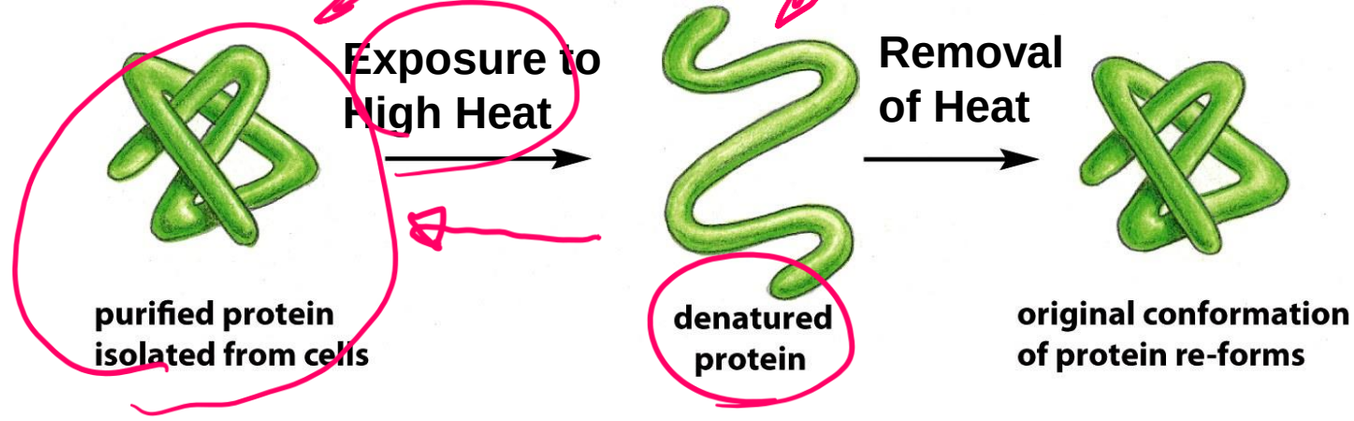
H-bonds
van der Waals
Hydrophobic effect

Native

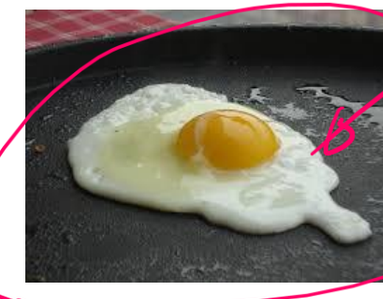
Unfolded

Chain disorder

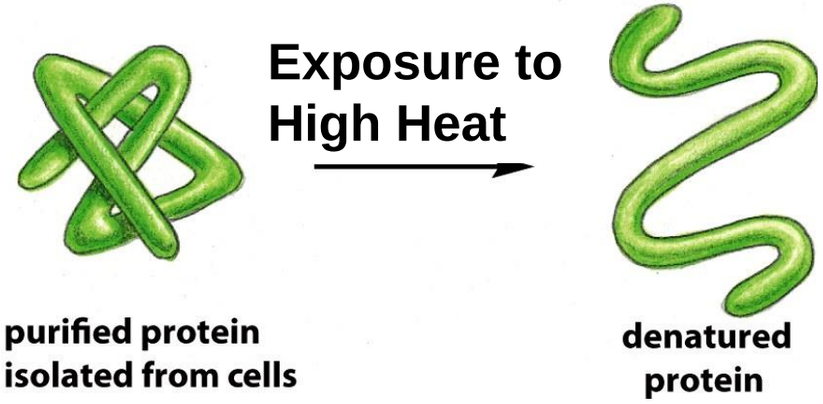
**Protein Denaturation**



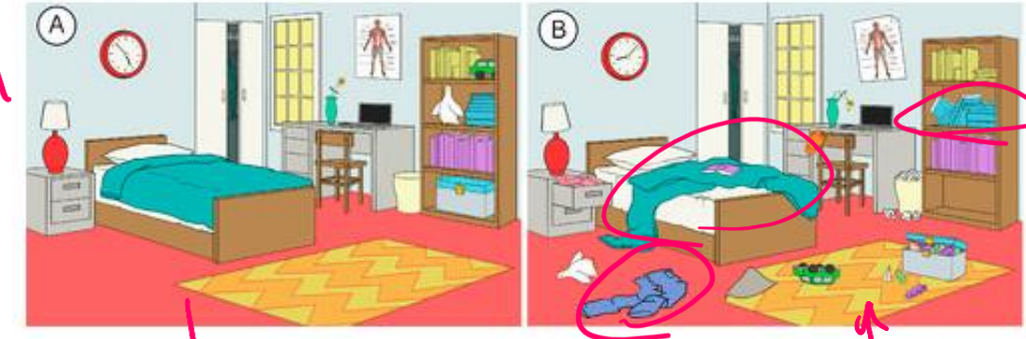

- Often, unfolded protein aggregate, which prevents refolding.



# Unfolded Polypeptides Are Flexible – High Entropy stabilizes the Unfolded state




Energy and Entropy = disorder!

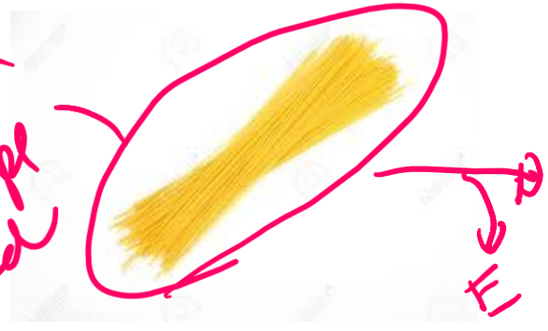


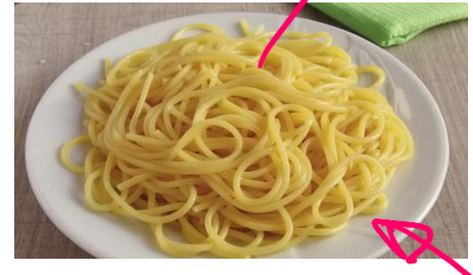

disorder

ordered

disordered

Same shape ordered



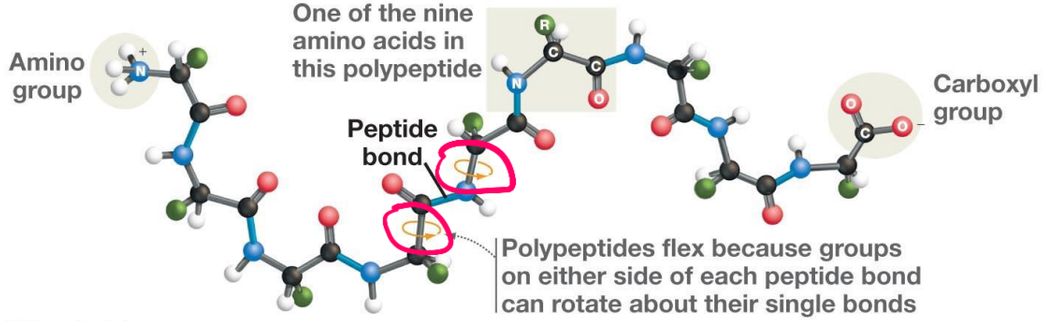


more stable

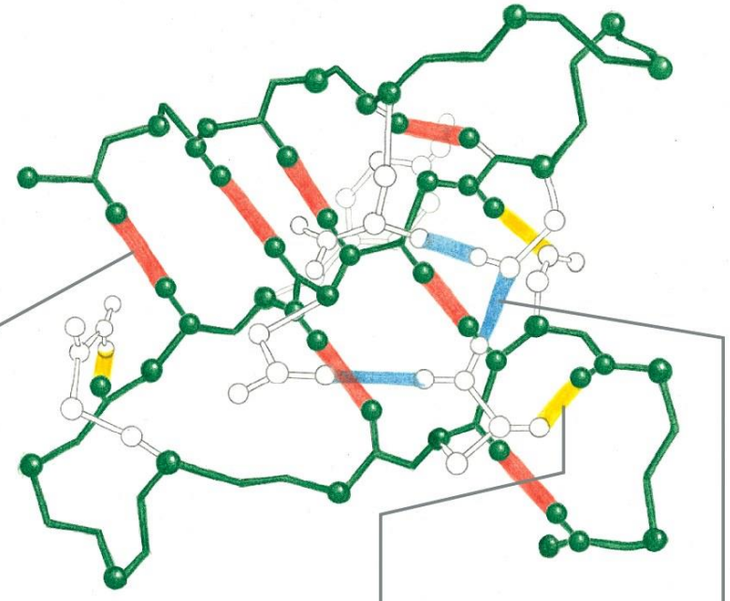
add energy

release energy

lower in energy more fav.



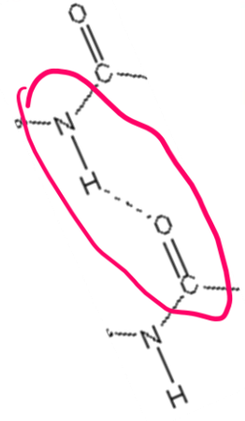
# Hydrogen Bonding Stabilizes the Tertiary Structure

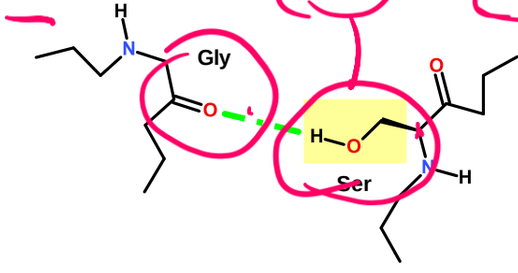


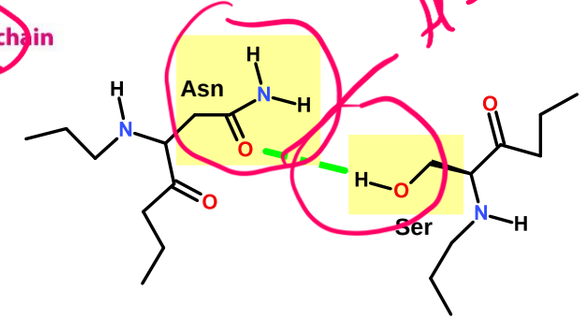
2ndy Structure

3ry Structure

| hydrogen bond between atoms of two peptide bonds | hydrogen bond between atoms of a peptide bond and an amino acid side chain | hydrogen bond between two amino acid side chains |
|---|---|---|
| backbone to backbone | backbone to side chain | side chain to side chain |



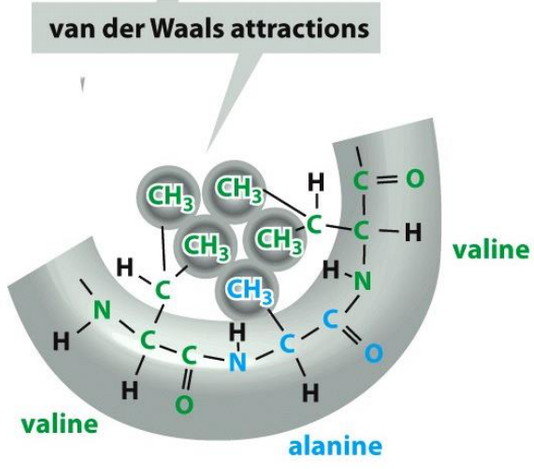


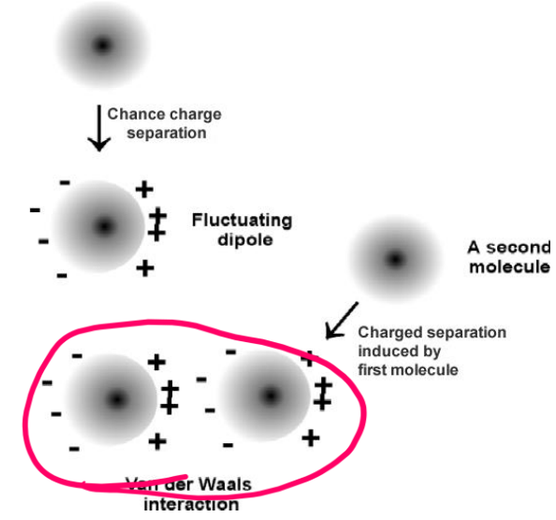



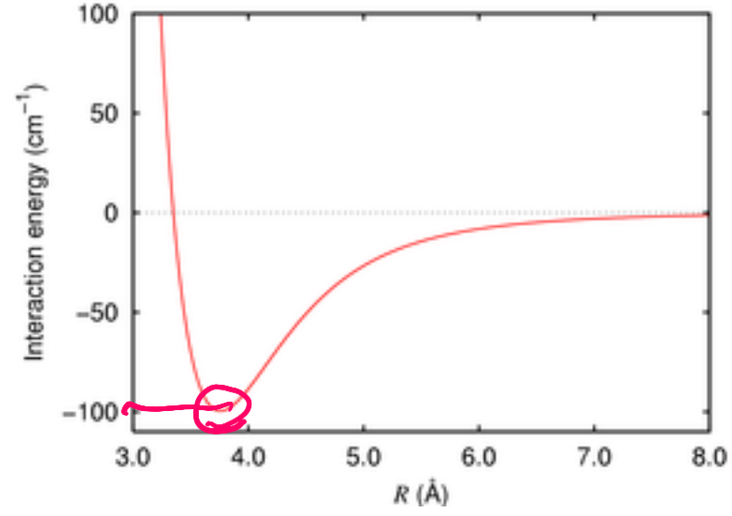

- **Hydrogen bonds** form between hydrogen atoms and the carbonyl group in the peptide-bonded backbone – secondary structure
- Hydrogen bonds are also found between hydrogen and electronegative atoms in side chains (sidechain-sidechain)
- Sidechains can form hydrogen bonds to the mainchain too.

# Van der Waals (VdW) interactions Stabilize the Folded State

- VdW are weak electrostatic interactions between side chains due to temporary (fluctuating) charges.
- Attractive from long distance
- Distance at lowest energy is at the van der Waals radii of the atoms.
- Optimized in the core of folded proteins by “knobs fitting into holes”
- Strength proportional to contact area.



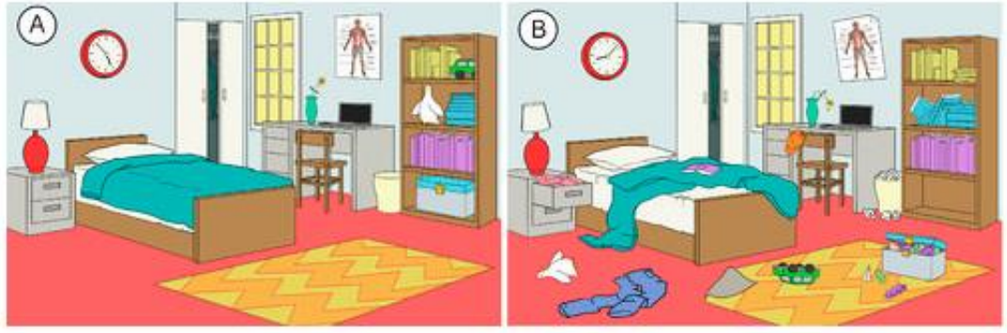



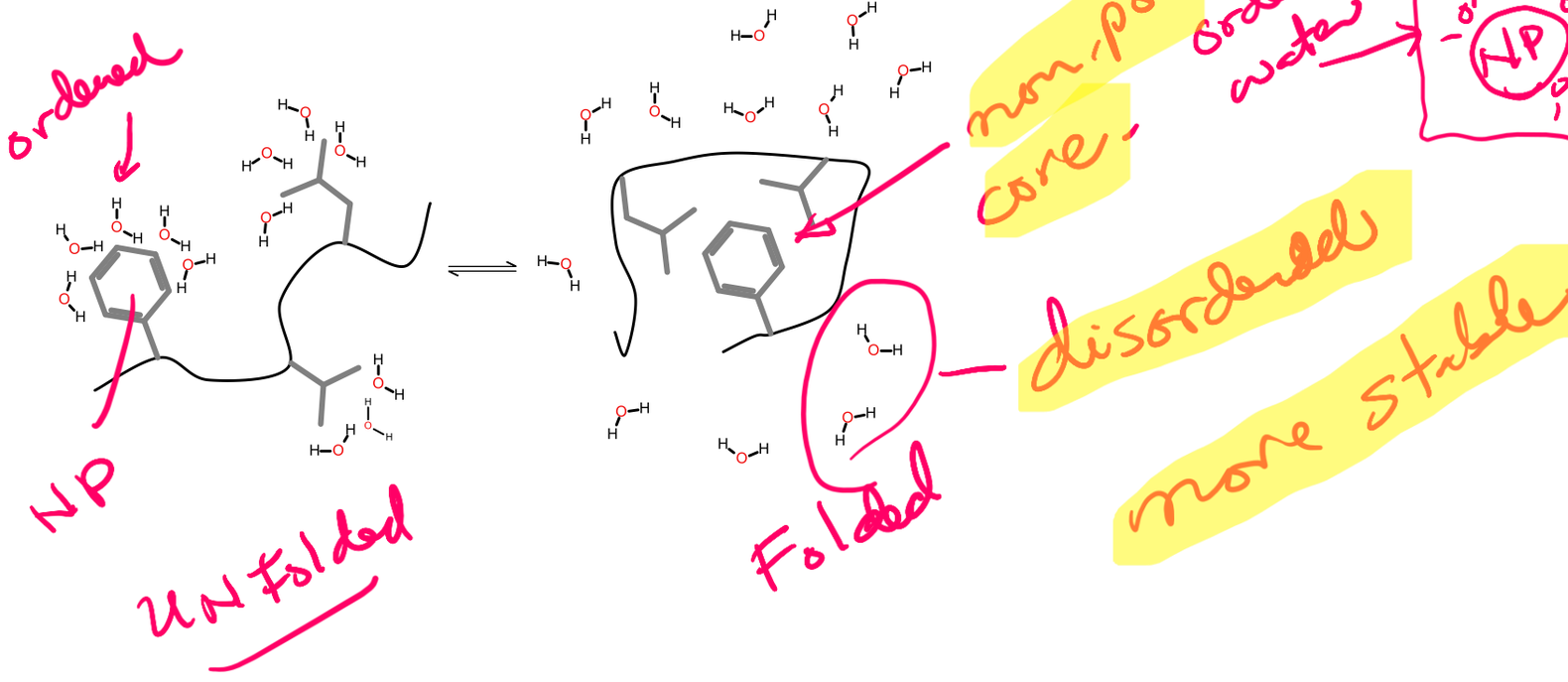




core => well packed => strong vdw

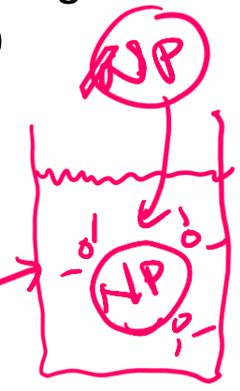
# Hydrophobic Interactions are Critical for Stabilizing the Folded Structure

Energy and Entropy



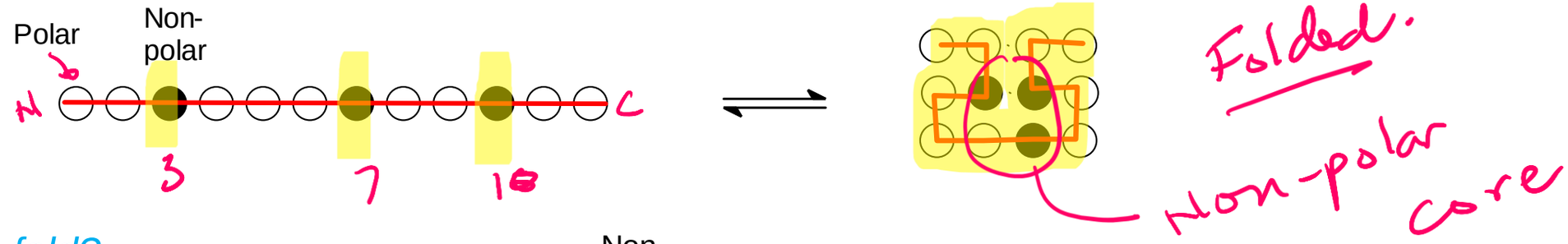

Ordered water hydrating a non-polar group




**Hydrophobic interactions** within a folded protein increase stability of the folded protein by releasing the ordered water that surrounded exposed non-polar groups in the unfolded protein. *Folding increases the entropy of the water – favorable.*

# Fold Depends on Amino Acid Sequence

The *position* of non-polar residues (filled circles) mostly affects the final fold:



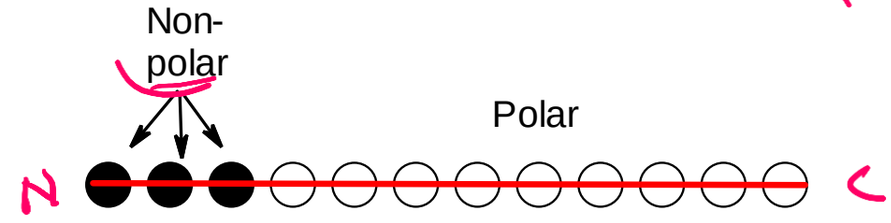

*Which is more stable fold?*

A

B

C

non-pola core



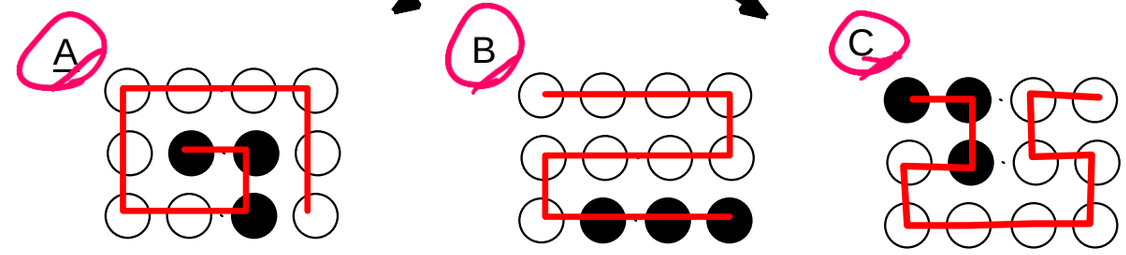

*Which is the least stable fold?*

A

B

C

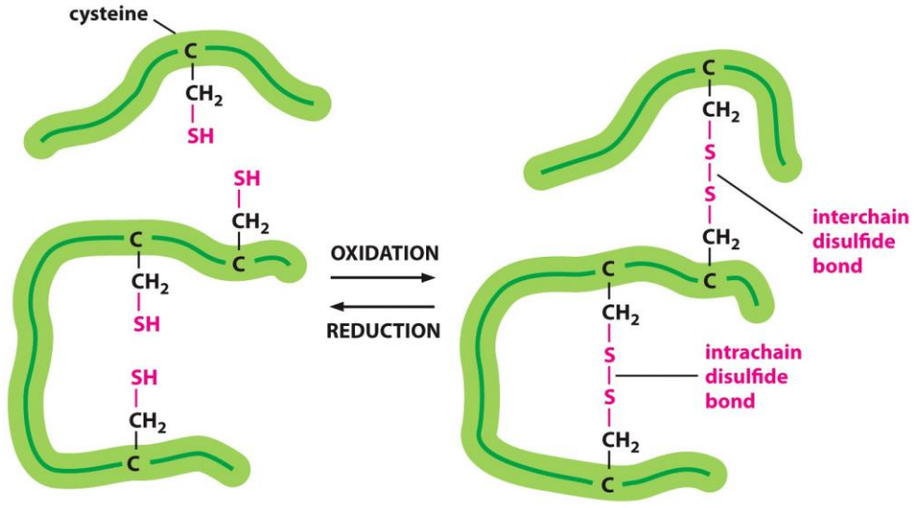
polar core
non-pola exposed
Still ordering water.

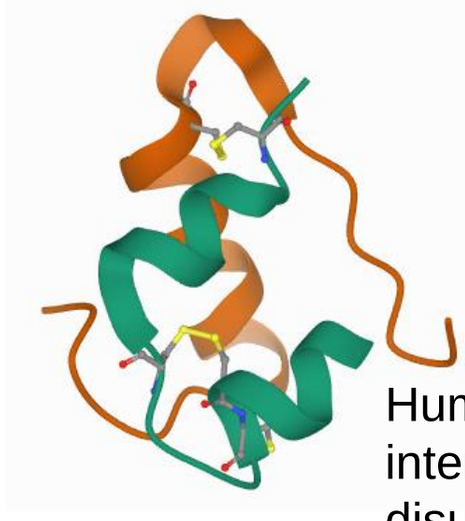



tertiary structure

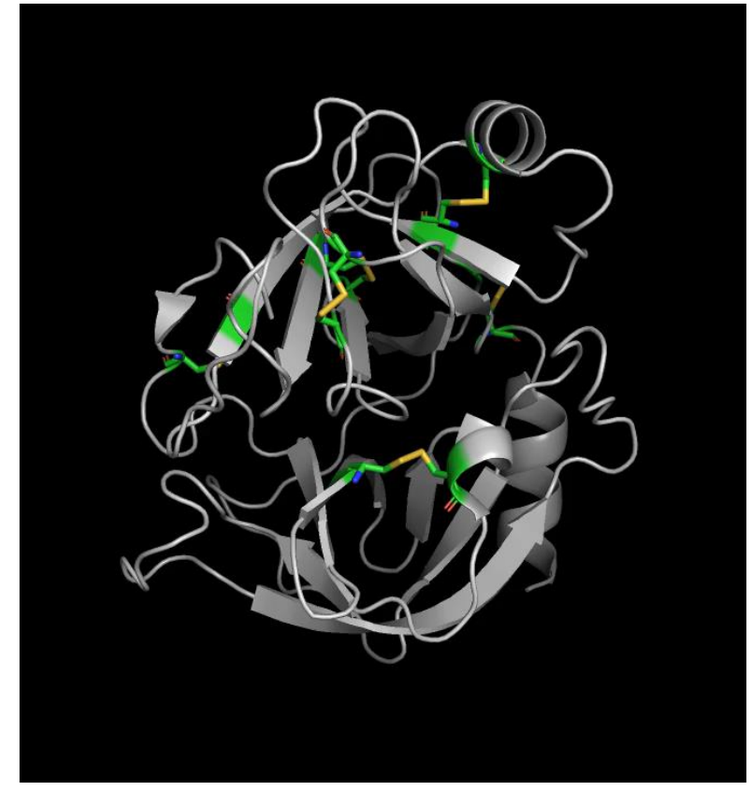
*Why?*

# Disulfide Bonds Stabilize Some Proteins Outside the Cell (and body)



Human Insulin – interchain disulfide bonds



Trypsin – a digestive enzyme produced in the pancreas, exported to the small intestine – disulfide bonds within a single chain.

# Summary - Interactions that Stabilize Folded Proteins.

- **Hydrogen bonds** form between hydrogen atoms (NH) and the carbonyl group in the peptide backbone (mainchain), and between and donors and acceptors on sidechains. *Mainchain-mainchain H-bonds are responsible for secondary structures.*
- **Hydrophobic interactions** within a protein increase stability of the folded state by *increasing entropy due to the release of water that was ordered by the exposed non-polar groups in the unfolded protein.*
- **van der Waals interactions** are *optimized in the well packed core of the protein*.

- **Covalent disulfide bonds** form between sulfur-containing cysteine residues stabilizing them (usually only exported, secreted proteins).

H-bonds
van der Waals
Hydrophobic effect

Chain disorder