

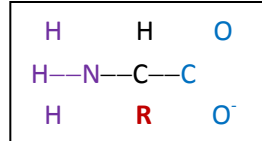
# PROTEINS – STRUCTURE AND FUNCTION (DR. TRAISH)

## Introduction to Proteins

- Proteins are abundant and functionally diverse molecules
- They participate in cell regulation at all levels
- They share a common structural feature: all are linear polymers of amino acids
- Examples of protein function: catalysis, transport and storage, coordinated motion, mechanical support (why we don't dissolve), organism host defenses, growth and differentiation, communication

## Structure and Function

### I. AMINO ACIDS



#### a. Chemical Structure

- i. Alpha carbon is bound to an amino group, a carboxyl group, and a side chain

#### b. Classification – based on characteristics of side chains

- i. Aliphatic, non-polar (alkyl) – hydrophobic, usually located on interior of protein

1. Gly – smallest amino acid; can be inside or outside
2. Ala
3. Val
4. Leu
5. Ile – side chain important for protein stability
6. Pro – shares many properties with aliphatic group

\* Alkyl chain is cyclized(imino acid)→unique properties

- ii. Aliphatic, polar – hydroxyl- or sulfur-containing side chains

1. Ser = Alanine + hydroxyl → polar, hydrophilic (can participate in H bonding); can be site of phosphorylation
2. Thr – has secondary alcohol group → hydrophilic, polar (H-bonding); can be found on the outside of a protein because hydrophilic
3. Cys – thiol (SH) group instead of hydroxyl = more acidic → more reactive
  - \* an enzyme with a lot of cysteine can form **disulfide bridges** → can exist at high temperatures
4. Met – hydrophobic → found on interior

- iii. Aromatic Amino Acids – bulky side chains; mostly nonpolar and hydrophobic

1. Phe – very hydrophobic → very important in folding (found on inside)
2. Tyr = Phe + hydroxyl group – can participate in H-bonding, but still hydrophobic (only ionizes at a high temperature)
3. Trp – has indole substituent
  - \* Deficiency of Trp → belagra (type of dermatitis)

- iv. Acidic side chains – have COOH group

1. Asp - α and β COOH groups → 2 negative charges → acid (polar, hydrophilic)
2. Asn – replaces COOH with NH<sub>2</sub> group – amide of Asp

## PROTEINS – STRUCTURE AND FUNCTION (DR. TRAISH)

3. Glu – like Asp but with an extra methyl (so gamma COOH)
4. Gln – like Glu but COOH replaced with NH<sub>2</sub> group (doesn't ionize but participates in H-bonding)
- v. Basic Amino Acids – accept proton; each plays a critical role in protein structure and function
  1. His – hydrophilic, basic, found on outside of protein
    - \* has imidazole (ring with 2 N) → can donate or receive H → functions in acid/base catalysis, found in enzymes
    - \* can bind to metal ions
  2. Lys – has NH<sub>3</sub><sup>+</sup> at epsilon C → highly protonated → doesn't lose H until very high pH; found in **histones** (bind DNA)
  3. Arginine – side chain has guanidino group
    - \* Arg + O<sub>2</sub> → nitric oxide, a vasodilator
- c. Ionic properties of Amino Acids
  - i. Amino acids are **ampholytes** – they have both acidic and basic weakly ionizable groups
  - ii. at a low pH, net charge is positive; at a high pH, net charge is negative
  - iii. but mostly they exist at a physiological pH:
    1. αCOOH ionizes and gives a H<sup>+</sup> and a conjugate base (COO<sup>-</sup>)
    2. αamino ionizes and retain a H<sup>+</sup>
    3. this is called a **zwitterion**
  - iv. an ampholyte's **isoelectric point** (point at which net charge is zero) = average of pKa<sub>1</sub> and pKa<sub>2</sub>

**\*\*The side chains of Amino Acids make them chemically diverse\*\***

## II. ORDERED STRUCTURE OF PROTEINS

- a. Overview
  - i. Primary Structure
    1. linear sequence of amino acids in a polypeptide chain
    2. Convention: amino terminus on the left (N-terminal), COOH terminus on the right (C-terminal)
  - ii. Secondary Structure
    1. Local folding of amino acids in primary sequence into a specific arrangement
    2. Different regions of the polypeptide can be different conformations – doesn't have to be all one or the other
    3. Carbonyl and amide groups of backbone are linked by hydrogen bonds, stabilizing the secondary conformation
  - iii. Tertiary Structure

# PROTEINS – STRUCTURE AND FUNCTION (DR. TRAISH)

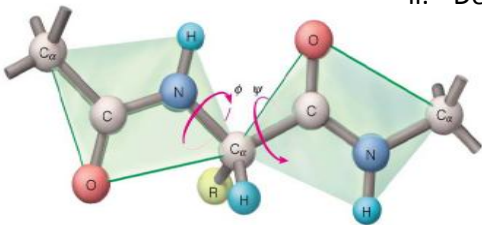
1. Spatial arrangement of various secondary structures (relationship of domains); the complete 3-D structure of polypeptide units
  2. hydrogen bonding, hydrophobic interactions, electrostatic interactions, van der Waals' forces all stabilize conformation
  3. For many proteins, tertiary is the highest level of structure
- iv. Quaternary Structure
1. When there is more than one subunit (even if they are all the same)
  2. Structure formed by monomer-monomer interaction in an oligomeric protein (can be hetero- or homo-oligomers)
  3. Forces: charge-charge, van der Waals,
  4. Example: hemoglobin: 2 $\alpha$  and 2 $\beta$  subunits

## b. Primary Structure – What is a Peptide Bond?

i. A condensation reaction between the  $\alpha$  amino group of one amino acid to the COOH group of another

ii. Delocalization of  $\pi$  electrons over O, C, N:

1. Partial double bond character of C–N and little twisting about C–N bond
2. C=O and C–N bonds are nearly parallel
3. O, C, N, and H are nearly coplanar – the *trans* form is favored
4. Peptide bond is **metastable** – proteins hydrolyze in aqueous solution when catalyst is present



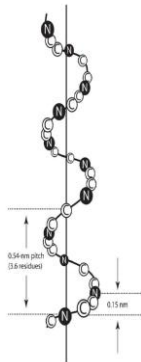
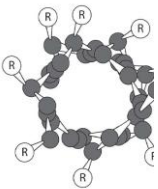
iii. Conformations of the polypeptide chain

1. Bond length and angles should be distorted as little as possible
2. no two atoms should be closer than their van der Waals radii
3. rotation only possible about bonds adjacent to carbons
4. hydrogen bond between amide H and carbonyl O necessary to stabilize regular folding

## c. Secondary Structure

i. Alpha helices

1. All are right-handed
2. Each amino acid contributes to rise:
  - \* **pitch (p)**– distance traveled to make 1 turn (.54nm)
  - \* **rise (r)** – distance between amino acid residues = .15nm/residue
  - \* **n** = residues per turn = 3.6
  - \* 3.6 residues x .15nm/residue = .54nm pitch
3. COOH of one residue H bonds to NH<sub>2</sub> of aa 4 residues ahead (i + 4)
4. Side chains project to outside
5. Amino acid residues that interrupt alpha helical structure:
  - \* Branched (ie Val, Ile, Phe) amino acids in succession (because of space)
  - \* bulky (Trp, Tyr) amino acids in succession
  - \* Charged aa in succession



## PROTEINS – STRUCTURE AND FUNCTION (DR. TRAISH)

- \* Prolyl residues (cyclic)
- 6. Ala, Cys, Met, Glu, Gln, His, Lys favor alpha helix
- ii. Beta sheets
  1. Polypeptide strands of at least 5-10 amino acids lay adjacent and interact to form sheet
  2. Can be parallel (N and N terminus together) or antiparallel (a bit more stable – H bond is at a 90° angle)
  3. Sheet is perpendicular to axis
  4. In alpha helix, peptides can still stretch; in beta sheet, they are already stretched (ex: silk)
  5. In alpha helix, hydrogen bonds within same polypeptide; in beta, bonding between different polypeptides
  6. Beta sheets can lie on top of each other – bound together by hydrophobic interactions
  7. Val, Ile, Phe, Tyr, Trp, and Thr favor beta sheet structure
- iii. Gly, Ser, Asp, Asn, Pro, and Arg favor turns (don't have to know about)
- d. Tertiary Structure
  - i. Thermodynamics of folding
    1. Spontaneous reaction:  $\Delta G$  must be negative
    2.  $\Delta G = H - T\Delta S$
    3. Stability depends on 3 factors:
      - \* Protein folding means negative  $\Delta S$  (unfavorable)
      - \* Lower  $\Delta H$  from side chain interactions (favorable)
      - \* Burying hydrophobic groups in protein means more room for solvent molecules to move around  $\rightarrow$  positive  $\Delta S$  (favorable)
    4.  $\Delta G \text{ total} = \Delta H \text{ chain} + \Delta H \text{ solvent} - T \Delta S \text{ chain} - T \Delta S \text{ solvent}$ 
      - \* Burying hydrophobic groups increases overall entropy  $\rightarrow \Delta G$  is negative  $\rightarrow$  proteins fold spontaneously

Information needed to determine the secondary and tertiary structure is all stored in the primary structure. This means the primary structure contains information pertinent to protein function.

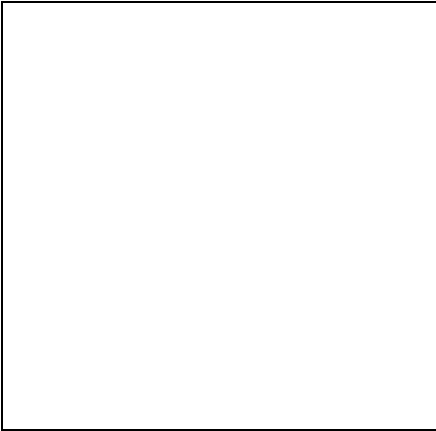
### III. DISEASE

About half of all human disease is attributed to protein misfolding.

- a. What happens when proteins misfold?
  - i. Misfolded intermediates and unfolded proteins can aggregate, forming **amyloid fibrils** – clog stuff up
  - ii. Misfolded proteins  $\rightarrow$  fewer normal proteins  $\rightarrow$  reduced biological activity  $\rightarrow$  disease
- b. Factors that predispose otherwise soluble proteins to misfold into  $\beta$ sheet-rich conformations:
  - i. Excessive production

## PROTEINS – STRUCTURE AND FUNCTION (DR. TRAISH)

- ii. Decreased clearance
  - iii. Mutations (missense, expansion) that result in misfolding
  - iv. Certain biochemical conditions (ie extreme pH)
  - v. Failure of chaperone systems
  - vi. Prolonged time (age)
- c. Diseases – Examples
- i. Alzheimer's Disease



1. "a progressive neurodegenerative disease characterized by cognitive and functional deterioration and death of neurons in areas of the brain important to memory and learning"
2. Diagnostic lesions of Alzheimer's
  - \* **Neurofibrillary tangles (NFT)** – composed of highly phosphorylated forms of tau (involved in signal transduction pathways, regulation of microtubule formation; are phosphorylated and dephosphorylated to accomplish)
  - \* **Amyloid plaques** – composed extracellular amyloid and peptide (A $\beta$ ) deposits =dense and insoluble (can be found in other subjects but more prevalent in AD patients)

- ii. Emphysema

1. "A chronic progressive lung disease characterized by abnormal permanent enlargement of airspaces as a result of destruction of alveolar walls"
2. Mostly from smoking, but 1-2% from genetic deficiency:
  - \*  **$\alpha$ 1-antitrypsin** (plasma proteinase inhibitor)
  - \* point mutations of  $\alpha$ 1-antitrypsin gene  $\rightarrow$  **S and Z deficiency variants**; Z results in more severe deficiency
    - i. polymerizes, aggregates in liver  $\rightarrow$  juvenile hepatitis, cirrhosis, hepatocellular carcinoma
    - ii. Mobile reactive loop of  $\alpha$ 1-antitrypsin is supposed to denature **elastase**; when Z variant forms elastase isn't inhibited, lungs lose elasticity...
3. \* Research with +6mer peptide – to prevent Z from polymerizing

- iii. Prion Diseases

1. Entirely different: not from mutations
2. Incl. Creutzfeldt-Jakob disease, Scrapie, Kuru, BSE? – prion proteins are affected
3. Alpha helices instead of beta sheets ("rogue conformer")  $\rightarrow$  protein becomes pathogenic no chemical difference, just conformation
4. Rogue protein also infects normal proteins – not really understood why, but normal protein conformation is affected (maybe thermodynamics?)