PROTEINS

Proteins: Functioning biological molecules that are polymers of amino acids.

Amino acid: the subunit or individual "link" in a protein "chain".

- There are twenty different amino acids from which all eukaryotic proteins are constructed (over 100,000 proteins ID'd in eukaryotes – amazing diversity from just 20 subunits!!).
- Nine amino acids are called **essential amino acids** because humans and most animals are unable to synthesize them, and is essential that we eat these in our diet.
- Every amino acid contains at least 1 N atom.

We will treat protein structure somewhat superficially in Bio101, because you will explore it in great detail in Bio103 (see box below). Please understand the following for Bio101

Protein structure

- Every protein has a very specific 3D shape, which is determined by the sequence of amino acids in a protein chain. Interactions among the aa's cause folding of the protein into its specific 3D shape.
- Every protein has one specific job (or function), that is determined by its 3D shape
- When a protein loses its shape, it loses its function and said to be **denatured**
- The sequence of amino acids in a protein chain is specific, critical, and coded for directly by DNA. DNA mutations can directly cause changes to protein structure. I hope we'll have time to discuss this at some point this semester, but I fear that we won't have time. You will cover it in Bio103.

Denatured: When a protein loses its shape, it loses its function. A protein in such condition is said to be **denatured**. Common causes are excessive heat, toxins, and cell chemistry imbalances (e.g., pH or ionic solutions).

Protein structure vocabulary not covered in Bio101 (covered in Bio103)

The following vocabulary and concepts for amino acid and protein structure are omitted from our conversation in Bio101. This brings me great pain (!), but you will cover this in Bio103 with Dr. Betancourt:

R group (aka: **amino acid side chain**): the atom or molecule bonded to the **alpha (*middle*) carbon** in an amino acid, and which determines the identity of an amino acid.

R groups are conceptually important because R groups in an amino acid chain (protein!) chemically attract or repel each other and cause the protein to fold into a specific 3-D shape. This shape then determines the job that the protein does. Form = function!

Polypeptide/polypeptide chain: Three or more amino acids bonded together in a chain (derived from many (poly) amino acids that are linked with peptide bonds). Often used as synonyms for protein, but this is not 100% accurate since some proteins are constructed of multiple polypeptide subunits (see quaternary structure, below).

Peptide bond: the bond that links amino acids together in a polypeptide / protein

Four levels of protein structure: primary, secondary, tertiary, quaternary

<u>Protein function</u>:
OK, so we now know a little bit about protein structure, but how do they function and what do they do in the tissues of organisms? <u>LOTS!!!</u> Remember – there are over 100,000 different proteins ID'd for eukaryotes!!

In understanding how proteins work, it is essential to remember the following two facts:
1) <u>A single protein has **one highly specific function** or job</u>.
2) <u>A protein's **3D shape** determines its **function**</u>**.**

<u>Proteins have two general functions:</u>
1) **Provide structure** in cells and tissues (mostly in animals – plant tissue is mostly cellulose and other polysacharrides and a molecule called lignin in wood…more later.)
2) **Do work** in cells (this work accomplished **by changing shape**). The shape change here is subtle, and so it is still accurate to say that a single protein has one specific 3D shape. The shape changes are akin to you opening and closing your fist to grasp an object. The anatomical shape of your hand does not change during the opening/closing.
3) **Storage proteins.** Proteins in seeds of plants that provide amino acids (and other benefits) to seedlings (young developing plants)

1) Structural proteins – Proteins that provide physical support for cells and organisms. (Prominent in animals, relatively sparse in plants).

As we discussed in lecture, structural proteins are a major component of animal tissues, but in plants structural proteins are rare but found in cell walls of certain cell types (if you want more detail, see this review: Keller, B. 1993. Structural cell wall proteins. Plant Physiology 101: 1197-1130). Having said this, recall that cellulose is a major component of plant cell walls. There are other important structural molecules, which we'll learn about after the first midterm.

Examples of important/common structural proteins in animals (although this is a plant biology course, you should be aware of know these examples):
EX: **keratin.** Two forms: **alpha helix** and **pleated sheets.** Common in hair, nails, skin of mammals, also claws, scales, and shells of reptiles; beaks and feathers of birds, etc…
EX: **collagen**. Helical polypeptides twisted into a triple helix (like a rope or cable). Gives strength to tendons and ligaments, also in bone, cartilage, skin; comprises 25-35% of all protein in human body.

2) Shape-changing proteins in plants
Some proteins do work by 'changing' between two (and only two!!) stable shapes. By changing back and forth between the two basic shapes, proteins can move materials and/or do work in cells.

2a) membrane transport proteins: transport proteins that are embedded in cell membranes and move materials into/out of cells through cell membranes. We'll revisit large complexes of membrane proteins when we discuss the light reactions of photosynthesis. There are many and diverse types of membrane proteins, but I defer coverage here as you will explore those in great detail in Bio103.

2b) Enzymes: Proteins that catalyze chemical reactions (the making/breaking of chem. bonds)
Note: Most enzymes have names that end with the suffix "–ase", while the root of the enzyme name hints at the identity of the molecule with which the enzyme reacts (EX: sucrase, cellulase, maltase, lactase, amylase, ATP synthase, etc…)

A single enzyme (like all proteins!) has **one highly specific function** or job. Organisms can do complex work (such as building large molecules) because they have "assembly lines" of many many specific enzymes – just like Henry Ford's innovation for building the first Model T. In organisms, these complex pathways are called **biochemical pathways**.

Biochemical pathway (metabolic pathway): a series of chemical reactions occurring within a cell, in which an initial molecule is modified by a sequence of chemical reactions. These reactions are catalyzed by enzymes, where the product of one enzyme acts as the substrate for the next.

You will not be made to memorize the reactants, products, and enzymes in biochemical pathways in this course (well, of course there will be a few exceptions to this rule!). You will no doubt explore and memorize a great number of biochemical pathways in Bio103 and other upper division biochemistry courses. Similarly, we will not cover the specifics of how enzymes function in Bio101, you will get to that in Bio103. (If you have been through Bio103 already, or if you want a preview, the enzyme-specific concepts covered in 103 but not this course are in the upper of the two boxes below.)

Specifics of enzyme kinetics that we will NOT cover in Bio101 (you'll study this in Bio103)
You do NOT need to know this material for Bio101

Enzymes: Proteins that catalyze chemical reactions (this definition you need to know, of course)
Activation energy (energy of activation): The energy required to initiate a chemical reaction.
Catalyst: substance that lowers the activation energy of a reaction by forming a temporary association of the reactant(s).
Substrate: The molecule acted upon by an enzyme.
Cofactor: Non-protein components required by some enzymes to function. Two types: 1) Ions and 2) coenzymes (non-protein organic molecules).
Active site: the region of an enzyme surface that binds to the substrate during the reaction catalyzed by an enzyme.
The "**lock and key**" concept for enzyme specificity.

A reiteration of enzyme basics you need to know for Bio101
(from this foundation, we will discuss important examples of enzymes in plant biology)

1) Enzymes are proteins that catalyze (accelerate or "make possible") chemical reactions.
2) Any one enzyme can catalyze one specific chemical reaction (many, many times!).
3) The names of enzymes end with the suffix "–ase".
4) The root word in the name of an enzyme often pertains to the molecule(s) upon which it acts to catalyze a chemical reaction (e.g., cellulase, lactase)

Applied concepts of proteins (mostly enzymes) in plant biology

Proteins such as enzymes are fundamentally important biomolecules, and they warrant our attention for that reason. We identified three specific and applied concepts in botany that have to do with proteins. These were:

1) Many plants are adapted to minimize heat stress in order to avoid protein denaturation
2) Glyphosate herbicides (inhibit construction of three amino acids – thus proteins can't form and plants die…).

   You should understand the following statement for exams: Glyphosates inhibit the enzyme EPSPS, which is one enzyme in a biochemical pathway that leads to synthesis of three specific amino acids (*no need to know which three, but they are tryptophan, tyrosine, and phenylalanine*). Because they block the synthesis of these amino acids, glyphosates such as *Roundup* therefore interfere with the construction of proteins in plants and cause their death. Understand why roundup can be applied as a postemergence herbicide on crops and not kill them.
3) Storage proteins (especially in legumes and grasses, the latter of which produce gluten)
4) Nitrogen fixation

Storage Proteins

Storage protein: Proteins in the seeds of plants, provide amino acids for developing seedlings (very young plants).

As seeds germinate (i.e., begin to grow) the developing embryo needs amino acids with which to construct proteins. Seeds contain storage proteins that provide these amino acids. The amount of protein varies taxonomically. Two important taxa are legumes and grasses, among the latter some species produce gluten.

Gluten: A general term for storage proteins in the seeds/fruits of many grasses (e.g., wheat, rye, oats, barley). Gluten causes coeliac disease, a serious autoimmune disorder that impacts the small intestine and causes discomfort, poor absorption of nutrients, diaorrhea, and stunted growth in serious cases.

Is gluten categorically "bad" for humans? It is unquestionably responsible for coeliac disease, but gluten has a bad reputation and is suggested by many to be something that many of us should avoid. Is this justified by scientific research? This would be a great review paper topic!

Plants and nitrogen fixation

Note that every third atom in a protein chain backbone is a nitrogen atom. Organisms must acquire the nitrogen needed to build amino acids, and by linking amino acids together, assemble proteins.

Q: Where do organisms acquire nitrogen to build amino acids?

A: Ultimately, the atmosphere!!…but… the way nitrogen gets form the atmosphere into organisms is an interesting story…

Composition of Earth's atmosphere:

~ 78% nitrogen gas (N_2, aka *molecular nitrogen*)

~ 21% oxygen gas (O_2, aka *molecular oxygen*)
~ 1% water vapor (H_20) at sea level
~ 0.03% carbon dioxide (CO_2)
(also~1% Ar, and other trace quantities of other gasses)

Based upon the composition of Earth's atmosphere, it would appear as though organisms can easily access nitrogen. This is not the case!! The two nitrogen atoms present in molecular nitrogen (N_2) are bonded together with a triple bond (you should be able to see why based upon what you know about the atomic structure and how covalent bonds are formed…what a GREAT exam topic!!).

This triple bond is extremely difficult to break, and therefore the N atoms in N_2 are inaccessible to nearly all biological organisms. Before N atoms can be accessed and used by most organisms, the N atoms must be liberated from the triple bond that holds them together. After this bond is broken, N atoms bond with hydrogen atoms to form ammonium (NH_4^+) , or with oxygen atoms to form nitrate NO_3^- that organisms (plants) can assimilate and/or subsequently dismantle to acquire N atoms.

As with carbon – ALL ANIMALS AQUIRE their N atoms form the plants they eat (or that some other heterotroph has eaten…).

This is a complicated process. For starters, know the following:

1) N_2 is triple bonded and only bacteria/archaea can break this bond and access N atoms in N_2
2) N_2 is converted to molecules that are usable to living organisms by a process called **nitrogen fixation** (more below, and in lecture).
3) All nitrogen in living organisms was once removed from the atmosphere through nitrogen fixation and subsequently absorbed into plant tissues before being eaten by heterotrophs or decomposed by decomposers (fungi, bacteria…)

Nitrogen fixation: Incorporation of ATM nitrogen (i.e., N_2) into N-containing molecules usable by plants (e.g., ammonium NH_4^+)

How does N fixation happen?

<u>Human mediated processes of nitrogen fixation</u>:
1) Haber-Bosch process (human performed, important source of N for "conventional" agriculture) complex chemistry…do not need to know!!

<u>Natural process of nitrogen fixation</u>
1) Lightning (natural, very slight contribution to global N cycle)
2) Bacterial fixation – most nitrogen fixation occurs due to special types of bacteria called *nitrogen fixing bacteria*!! **Nitrogen-fixing bacteria and archaea are the ONLY living organisms that can fix nitrogen**!! They are therefore critical to all life on Earth.

Nitrogen fixing bacteria contain the enzyme *nitrogenase*, which catalyzes the reaction of N_2 and H atoms to NH_3 (NH_3 subsequently converted to NH_4 and nitrate…)

nitrogenase: the enzyme in N-fixing bacteria & archaea that catalyzes the conversion of N_2 to NH_3

Many nitrogen fixing bacteria live symbiotically in the roots of plants called legumes (and other plants that are not legumes!). Examples of legumes are beans, peas, vetch, and clover.

Because legumes contain N-fixing bacteria in their roots, they are popular among farmers as a way to maintain high N levels in the soil. These plants are often planted by farmers - not to harvest them, but as a means to enhance soil quality – such plantings are known as **cover crops**.

Cover crops: crops planted primarily to manage soil fertility (N) and quality, water, etc…

<u>Humans and amino acids:</u>
Most (11) we can synthesize – (but to do this we need nitrogen in our diet – <u>remember that every third atom in a protein chain is nitrogen</u>!! - Itis also for this reason that plants must be fertilized with nitrogen-rich fertilizers, or “find” nitrogen in nature).

The acquisition of N atoms and amino acids is the major reason that humans (and other animals) must eat protein-rich foods. For vegetarians and vegans, that means substituting plant-based N sources for meat – legumes are an ideal part of the solution.

Essential amino acids: There are 9 amino acids that the human body cannot synthesize – we must eat these in our diet and they are therefore called the **essential amino acids**.

Animal-based foods usually contain all 9 essential amino acids. No single plant-based food*, including legumes, contains all of the 9 essential amino acids, and it has long been held true that a vegetarian/vegan diet must include sources that provide the “missing” amino acids (those not found in legumes). Historically this be accomplished by pairing legumes with corn (in the Americas) or rice (in Asia) – and thereby provide what is referred to as a “complete protein”.

Complete protein: Protein source that contains all 9 essential amino acids necessary for the dietary needs of humans or other animals.

Protein pairing: The practice of pairing plant foods to accomplish a complete protein.

As with many nutritional concepts, there is great debate and misunderstanding swirling about the concept of “complete protein”. The most sane and educated voices in the field seem to agree that vegetarian/vegan foods need not be perfectly paired at <u>every</u> sitting, but that a diverse diet that provides complete protein (and other nutritional needs!) should be a priority.

* One exception of this ‘rule’ appears to be quinoa, which is a non-cereal (i.e., non-grass family) grain native to South America where it has been cultivated for millennia. I find contradictory reports as to whether or not quinoa provides a complete protein, but the mounting consensus suggests that does indeed.