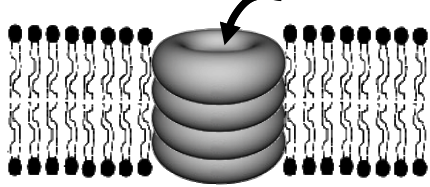


Molecular Biology Through Discovery

Problem Set 2: Protein

Submit reasoning and process, not merely answers!

- 2.1. Some antibiotics form rings that stack and create a pore through the membrane. Consider a cyclic polypeptide antibiotic in which each ring is composed one instance of each of the four amino acids: serine, glycine, threonine, and alanine. If the atoms of the backbone are approximated by touching spheres of about 0.2 nanometers in diameter,* estimate the circumference of the pore (presume it to be a circle) and the diameter of the largest molecule that could fit through it. Approximate the circumference (π -diameter) to be 3-diameter. (Show work: First read, then draw, then think, then draw some more, then calculate, then think some more)
- 
- 2.2. Before you cook an egg, the egg "white" is not at all white: it's clear. After you cook the egg, the "white" is white, because the large amount of globular protein has denatured (i.e., unfolded), and as a consequence, the protein has precipitated. Why should unfolding globular protein that are normally soluble in water cause them to stick to each other (which is what "precipitate" means)?
- 2.3. Lactate dehydrogenase (the last enzyme in human anaerobic glycolysis) is a soluble, multimeric protein. If you were to try to fold a single linear polypeptide chain of lactate dehydrogenase, you would find it impossible to do so without leaving a large number of hydrophobic amino acids exposed to water. Explain.
- 2.4. Use only the results of Sanger and Tuppy (1951) [Biochem J 49:463-481] to deduce as much of the structure of insulin you can. Do this ([as a group effort](#)) as if it were a geometric proof, appealing to lines within the tables (axioms) and truths you derive from them (theorems). For example:

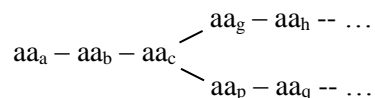
<u>Assertion</u>	<u>Justification</u>
A. Thr-Pro*	Table 6, Line 8
B. Thr-(Ala,Lys,Pro) [¶]	Table 9, Line 6
C. Only one Pro	Table 14
D. <u>Thr-Pro-(Ala,Lys)</u>	<u>A+B+C[†]</u>

*Meaning "The dipeptide N-Thr-Pro-C lies somewhere in the insulin polypeptide chain". The form N-XxxYyy-C means that the amino acids are read from amino end to carboxyl end.

[¶]Meaning "A tetrapeptide somewhere in insulin begins N-Thr and is immediately followed by Ala, Lys, and Pro in some unknown order"

[†]Meaning "The assertion on this line follows from the assertions on lines A, B, and C"

- 2.5. Suppose that Sanger and Tuppy tried used their methods to deduce the structure of a protein that was not a linear array of amino acid but rather had branch points:



What experimental results would they have obtained that would have allowed them to detect this structure?

* How big are nanometers? Try visiting [Scale of the Universe](#) and/or [Proton to Protein](#).

Extra – in case you have the time and inclination

- 2.6. Make a set of 20 different graphical symbols representing the 20 amino acids. You may use colors, shapes, fill style, etc, but no letters or numbers. The symbols should be organized so that they are *easy to write* and *easy to remember* and that if two amino acids share some important characteristics then their symbols also are similar in some respect (but not identical).