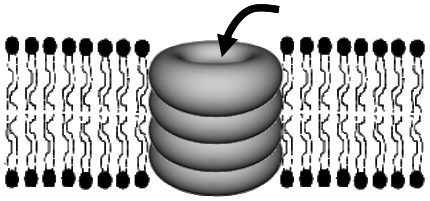
Molecular Biology Through Discovery

Problem Set 2: Protein + observations/assertions

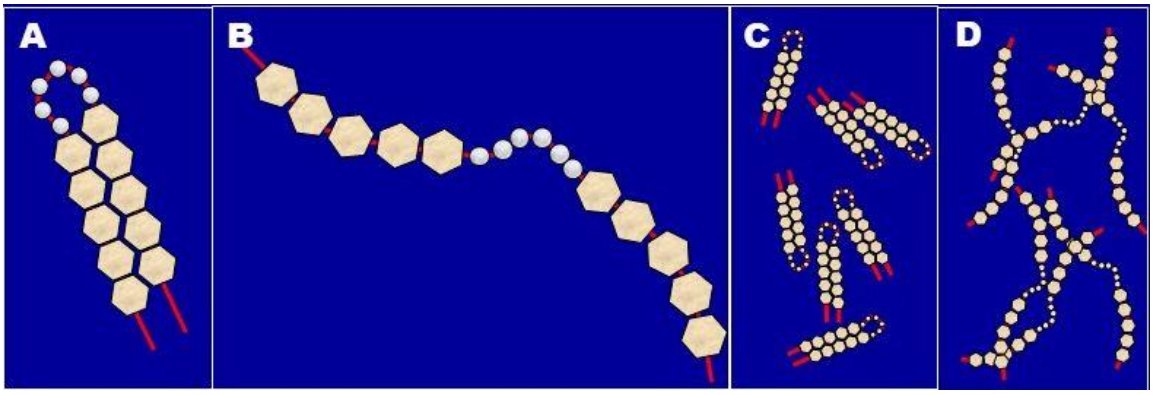
Submit reasoning and process, not merely answers!



2.1. Some antibiotics form rings that stack and create a pore through the membrane. Consider a cyclic polypeptide antibiotic in which each of four rings (shown as doughnuts) is composed of 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 a picture of the atoms of a ring, 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 – mostly ovalbumin -- 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)?

Consider a protein to be (grossly oversimplified) as a string of amino acids – some hydrophobic, some hydrophilic. A tiny protein might be visualized as in Panel A, below (ovalbumin actually has 391 amino acids, not just 16!). When the temperature goes up, it is unfolded, which you might visualize as Panel B. In egg white, the concentration of ovalbumin is very high (Panel C), and when it is heated up, all the ovalbumin is denatured (Panel D).



2.2a. Panel D is a snapshot in time, immediately after the egg suddenly reaches a critical temperature, and the state you see cannot remain for long. Armed with your experience with the Self Assembly tour, **draw how you think the ovalbumin chains will arrange themselves** once time resumes. Does it look like a precipitate?

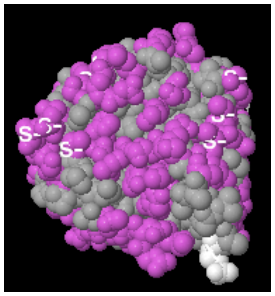
2.2b. Panel C is also not likely, but that's because it is too simple a model of ovalbumin. Let's consider a more realistic model.

- Go to FirstGlance, a program designed to help you visualize protein structures.
- Enter the Protein Data Base (PDB) identifier for ovalbumin – 1OVA – and click **Submit** (if the interface complains about "unusual components", click **Cancel**). You'll see a

* How big are nanometers? Try visiting Scale of the Universe and/or Proton to Protein.

rotating protein with four colored subunits. Actually ovalbumin is ordinarily monomeric (one subunit) but at the very high concentrations necessary to determine the protein's structure, the subunits aggregate.

- Click the **Spin** button to stop the protein from spinning.
- Click the **Views** button (upper left).
- Click the Isolate button (middle left), and click any of the four subunits (doesn't matter which). This gives you an image of just one subunit. You can see the alpha helices and other regions, but that's not why we're here.



- Click the Hydrophobic/Polar button (upper left). This shows you the hydrophobic amino acids (gray) and hydrophilic amino acids (pink). Grab the protein with your mouse and drag it around so you get a feel for what it looks like from all sides. End up so that the protein is facing you, so you're looking down its long axis (see panel at right).
- Click the **Slab** button (middle left). This view allows you to see a cross-sectional slice of the protein. Imagine that the protein projects outwards towards you, and the screen cuts it, allowing you to see what's inside. **What type of amino acid do you see in the interior of the protein?** Move the protein around to see other views.
- From what you've discover about the structure of ovalbumin, **in what way is Panel C (previous page) unrealistic?**

2.2c. Lactate dehydrogenase is an enzyme prominent in muscle that shunts glycolysis towards lactate when you are working so hard you have an oxygen deficit. Unlike ovalbumin, this protein really does exist as an aggregate at any natural concentration. 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. Based on what you've discovered, **create and describe a picture that illustrates why this should be**.

2.3. 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 described below), something like a geometric proof, appealing to observations – specific lines within the tables (axioms) -- and assertions you derive from them (theorems). A justification may consist of one observation or two prior assertions (assertions from Table 14 don't count towards this total). Make sure that your reasoning is transparent. For example:

	Assertion	Justification
A.	Thr-Pro*	Table 6, Line 8
B.	Thr-(Ala,Lys,Pro)¶	Table 9, Line 6
C.	Only one Pro	Table 14
D.	Thr-Pro-(Ala,Lys)	Assertions **A**+**B**+**C**†

*Meaning "The dipeptide N-Thr-Pro-C lies somewhere in the insulin polypeptide chain". The form N-Xxx-Yyy-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"

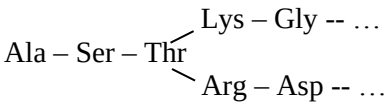
It would take a fair bit of time to reconstruct the structure of insulin by yourself. Big scientific problems, however, are generally solved not by individuals but by communities. Even though this question isn't too big of a scientific problem, we can still approach it in that spirit.

Accordingly, I've divided you into working groups, giving each person an area of expertise. Each person is associated with an amino acid. That person should consider all experiments from Sanger & Tuppy (1951) that bear on the specific amino acid, deducing all that's possible from the results, in the form shown above. When you are satisfied by a non-trivial deduction (the last line of your chain of reasoning), publish it, by posting it to the community bulletin board.

You'll be able to reach increasingly sophisticated deductions (perhaps even the entire sequence of the insulin B-chain!) as you combine your results with those of others in your group. You might also make some use of results from other groups, but beware! There might be all sorts of garbage posted to the community bulletin board – not intentionally, of course, but people do make mistakes. If you accept the results of others uncritically, you may accumulate mistakes and reach erroneous conclusions yourself.

How can you avoid contaminating your growing chain of deductions?

2.4. Suppose we lived in an alternate universe, and alternate Sanger and Tuppy used their methods to deduce the structure of a protein that was not a linear array of amino acid but rather had branch points:



Make up specific experimental results they might have obtained (i.e. make up results that might have appeared in one of their tables) that would have allowed them to detect this structure, consistent with branching and inconsistent with a linear structure.

2.5. Choose what you believe to be the two most contradictory observations in the research articles on which the four newspaper articles on Vitamin D are based. **If you believe the two observations to be truly irreconcilable, then describe the contradiction. If you believe that they can be reconciled, then describe how.** You can save yourself the task of reading all four articles by consulting with colleagues who have read articles different from the one you read. To that end, here's a list of who was assigned which newspaper article.