Molecular Biology Through Discovery
Problem Set 2: Protein Mutation*

2.1. 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:

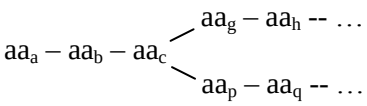
	Assertion	Justification
A.	Thr-Pro*	Table 6, Line 8
B.	Thr-(Ala,Lys,Pro)¶	Table 9, Line 6
C.	Only one Pro	Table 14
D.	Thr-Pro-(Ala,Lys)	**A+B+C**†

*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"

2.2. 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?

2.3. A child presents to you, her pediatrician, with all the classical symptoms of diabetes. Upon testing, you find that antibody against insulin detects only very low levels of insulin in her blood, but she responds normally to administered insulin. You are surprised to find, however, that the same antibody detects levels of insulin in the pancreas that are grossly higher than normal. What mutation might account for these findings?

2.4. An enzyme has a molecular weight of 60,000 daltons. When it is exposed to detergent, the protein breaks up to identical inactive components with molecular weights of 20,000 daltons. If the detergent is removed by dialysis, the 60,000-dalton protein reforms and regains enzymatic activity. You have isolated two mutant proteins. Mutant 1 shows no enzymatic activity and has a molecular weight of 20,000 daltons whether or not detergent is present. Mutant 2 has a molecular weight of 60,000 without detergent and 20,000 with detergent but shows no enzymatic activity in either case.

a. Suggest defects to explain the behavior of each of the mutant enzymes.

b. A person is heterozygous for Mutant 2 (i.e., has 50% Mutant 2 polypeptide and 50% normal polypeptide). How would you explain an observation that the person has 87.5% of the enzymatic activity of a normal person? How would you explain an observation of 12.5% activity?

c. Ascribe the terms "dominant" or "recessive" to the mutation leading to Mutant 2, according to the two situations presented in **b**.

---

* See the ***Structure and Function of Proteins*** topic page for links to possibly useful resources

2.5. Many proteins that form channels through membranes pass through the membrane multiple times. For example, rhodopsin, the light receptor protein in the rod cells of the retina, passes through the membrane seven times as alpha-helical chains. Below is a cartoon showing the side view of part of a hypothetical channel-forming protein -- call it rhodopsin. The circles are amino acid residues, the number of each corresponding to the amino acid's position in the chain. The roman numerals refer to membrane-spanning alpha-helical segments of the protein (only the first two are shown here). The top view shows how the seven α-helices participate in the formation of a pore through the membrane. The pore serves as the means by which protons can pass the membrane in response to light.

Congenital retinitis pigmentosa is a genetic disease leading to night-blindness. The disease exhibits a variety of symptoms of different severities, which, in many cases, have been linked to specific mutations in rhodopsin. For each given molecular outcome, choose one or more plausible amino acid mutations that could account for it. In each case, explain, briefly, why your choice(s) would lead to the outcome.

a. Rhodopsin found in cytoplasm, fails to insert in membrane.

b. Radical change in structure of rhodopsin. Channel doesn't form properly.

c. Overall structure of rhodopsin normal, but channel does not conduct protons.

d. Structure and function of rhodopsin normal.

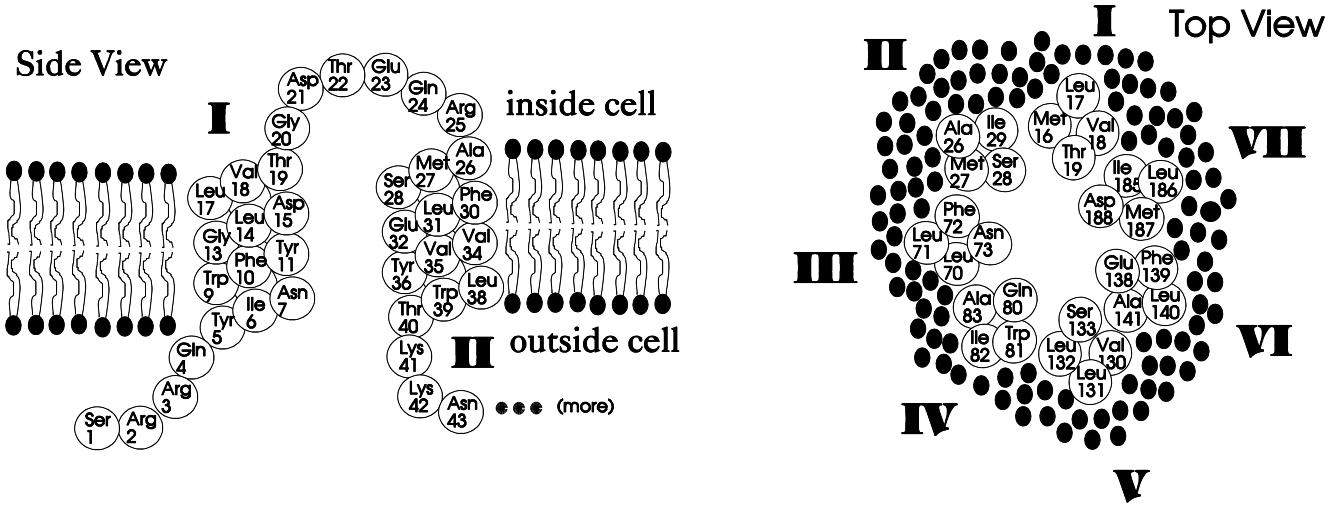
A. Insertion of three glutamates between Thr_{22} and Glu_{23}.

B. Insertion of three glutamates between Phe_{30} and Leu_{31}.

C. Glu_{138} mutated to arginine.

D. Asp_{188} mutated to leucine.

E. Mutation in amino acid not found in mature rhodopsin.



Abbreviations: **Ala**=alanine, **Arg**=arginine, **Asn**=asparagine, **Asp**=aspartic acid, **Cys**=cystine, **Gln**=glutamine, **Glu**=glutamic acid, **Gly**=glycine, **His**=histidine, **Ile**=isoleucine, **Leu**=leucine, **Lys**=lysine, **Met**=methionine, **Phe**=phenylalanine, **Pro**=proline, **Ser**=serine, **Thr**=threonine, **Trp**=tryptophan, **Tyr**=tyrosine, **Val**=valine

2.6. In Problem Set 1, problem 7, you considered a patch of hydrophobic amino acids that appears on the surface of a single polypeptide chain of the lactate dehydrogenase (LDH). You may have come up with a hypothesis as to why that patch is present. Let's see if you're right, by looking at the experimentally determined structure of the enzyme .

a. Find a Protein Data Base ID for the structure of human LDH.† You'll find several candidate structures. Avoid those that refer to domains or individual chains, because they will show only part of the protein.

b. Take the PDB ID you found in 2.7A and use it to display within Protein Explorer the structure of human LDH. Stop the spinning and hide water molecules. How many polypeptide chains does LDH have? (Here's a useful tool: Clicking any portion of the protein will identify in the message box (lower left corner) the atom you clicked, the amino acid containing that atom (and it's linear coordinate), and the polypeptide chain containing that amino acid). Which chains are most closely associated with each other?

c. Now to look for hydrophobic patches of a single polypeptide chain, let's say Chain A. Hide Chains B, C, and D, by clicking the Quick tab (to open up the QuickView selection interface) and doing the following:

   i. On the SELECT line, choose Chain A

   ii. On the SHOW AS line, choose Hide (several check boxes will appear)

   iii. Choose Hide **non-selected atoms and bonds** to hide everything ***except*** Chain A

d. Make a space-filling model of Chain A by selecting Chain A as before (actually, it's already selected), and SHOW AS Spacefill.

e. Color the amino acids by polarity by selecting hydrophobic (Hphobic) amino acids, COLOR BY gray. Then select Polar amino acids, COLOR BY yellow. Then select Acidic amino acids, COLOR BY red. Finally select Basic amino acids, COLOR BY green. Use the mouse to rotate the chain. Can you find one or more patches of gray, uninterrupted by hydrophilic amino acids?

f. What is the relationship between the hydrophobic patch(es) and the other chains? To answer this question, redisplay Chain B by selecting it and then SHOW AS Cartoon. Where does Chain B lie relative to the hydrophobic patch(es). Chains C and D?

g. With all this in mind, answer Problem Set 1, number 7 again: Why does a single polypeptide chain of LDH possess one or more hydrophobic patches on its surface? Of course you will make use of all the evidence you can, based on the work you've just accomplished.

Extra – in case you have the time and inclination

2.7. Plants and photosynthetic bacteria have been enjoyed enormous success in part because of their ability to harness sunlight to power the reduction of CO_2 to sugar. Organisms that can use sunlight as to drive the reduction of N_2 to biologically useful nitrogen compounds are far more rare. One reason for this is that the enzyme that catalyzes the nitrogen reduction is extremely sensitive to O_2, greatly limiting the environments in which nitrogen fixation can take place. Some have said that this is just the way it is – it is not possible for an enzyme to fix N_2 without also being killed by O_2. Why? Well if such an enzyme could exist, it would have arisen some time over the last 4 billion years of organismal evolution and organisms with this capability would have taken over the world!

Let's examine this reasoning. Have all possible proteins been tried sometime during the lifetime of the earth? If not, then what is a reasonable estimate for an upper limit on what fraction have arisen?

† The companion to Perutz et al (1965) tells you how to do this.