OLYMPIAD EDGE

Biology: Cell & Molecular Biology

From Foundations to Olympiad

Honors Bio ⟶ AP Bio ⟶ USABO Open ⟶ USABO Semifinal

8 Problems • 4 Tiers • Full Solutions • Original Problems

Joshua Gao

Olympiad Edge • Huntington, NY • Text: 631-833-0226

How to use this handout: Each problem is tagged with a difficulty tier. Work through each tier before moving to the next. The **Recognition Triggers** tell you what pattern to spot before you start solving. The **Flashy Tips** give you reusable techniques that transfer across problems. All problems are original, written in the style of the cited exams.

Difficulty Progression

Tier	Key Skill Tested	Prob.	Preparation Source
HONORS	Recall + basic application	1–2	Campbell Biology, Units I–III
AP BIO	Multi-step calculation + explanation	3–4	Campbell Ch. 23, AP CED Unit 7
OPEN	Quantitative reasoning + perturbation	5–6	Campbell + past USABO Opens
SEMI	Experimental design + synthesis	7–8	Alberts' MBoC + Semifinal archives

TIER 1 ★ Honors Biology — 2 Problems

Problem 1 | Organelle Function

Style: Honors Biology lab practical, typical of Regents-level or introductory biology courses.

A student observes a cell under a microscope and notices the following:

- The cell has a rigid outer boundary beyond the plasma membrane.
- The cell contains large, membrane-bound compartments filled with fluid.
- The cell has small green organelles clustered near the periphery.

(a) Identify whether this is a plant or animal cell. Justify your answer using all three observations.

(b) The student adds a concentrated salt solution to the slide. Predict what happens to the cell and explain why, using the term **osmosis**.

(c) Name the green organelles and state the overall equation for the process they carry out.

✓ Solution

(a) This is a **plant cell**. Three confirming features:

- **Cell wall** — the rigid boundary outside the plasma membrane (animal cells lack this).
- **Central vacuole** — the large fluid-filled compartment (animal cells have small or no vacuoles).
- **Chloroplasts** — the green organelles (animal cells do not photosynthesize).

(b) In a **hypertonic** salt solution, water moves out of the cell by **osmosis** (from high water potential inside to low water potential outside). The plasma membrane pulls away from the cell wall — this is called **plasmolysis**. The cell wall maintains the cell's shape, so the cell does not burst, but the cytoplasm shrinks inward. Importantly, the cell wall is freely permeable to solutes and water, so the space between the retracted protoplast and the cell wall fills with the external hypertonic solution (not with pure water).

(c) The green organelles are **chloroplasts**. They carry out **photosynthesis**:

$$6\,CO_2 + 6\,H_2O \xrightarrow{\text{light}} C_6H_{12}O_6 + 6\,O_2$$

▶ Recognition Trigger

Rigid boundary + green + large vacuole ⇒ Plant cell. This trio is the instant-ID pattern. If you see any of these three, lean plant. If you see all three, it's 100%.

Problem 2 | Mendelian Genetics

Style: Introductory Mendelian genetics, standard in Honors Biology curricula.

In pea plants, purple flowers (P) are dominant over white flowers (p). A student crosses two purple-flowered plants and observes the following offspring:

Phenotype	Count
Purple flowers	216
White flowers	72

(a) What are the genotypes of both parents? Justify your answer.

(b) What is the expected phenotypic ratio? Does the observed data fit the expected ratio?

(c) If one of the purple offspring is crossed with a white-flowered plant, what fraction of the F_2 offspring do you expect to be white? Show the Punnett square.

Solution

(a) Both parents are Pp (heterozygous). Reasoning: both parents show the dominant phenotype (purple), yet white offspring (pp) appear. For a pp offspring to exist, each parent must carry at least one recessive allele. Since both are purple, they must be Pp.
(b) Expected ratio from $Pp \times Pp$: **3 purple : 1 white**.
Total offspring $= 216 + 72 = 288$.

$$\text{Expected: } 288 \times \tfrac{3}{4} = 216 \text{ purple}, \quad 288 \times \tfrac{1}{4} = 72 \text{ white}.$$

Observed matches expected perfectly (a rare and clean result!).
(c) From a $Pp \times Pp$ cross, purple offspring are either PP $\left(\frac{1}{3}\right)$ or Pp $\left(\frac{2}{3}\right)$. Crossing a purple offspring with pp:
Case 1: If the purple offspring is PP (probability $\frac{1}{3}$): $PP \times pp \Rightarrow$ all Pp (all purple), so **0 white**.
Case 2: If the purple offspring is Pp (probability $\frac{2}{3}$): $Pp \times pp \Rightarrow \frac{1}{2}\,Pp$, $\frac{1}{2}\,pp$, so $\frac{1}{2}$ white.
Overall expected fraction white:

$$\frac{1}{3}(0) + \frac{2}{3}\left(\frac{1}{2}\right) = \boxed{\frac{1}{3}}$$

◇ Flashy Tip

The $\frac{2}{3}$ Trap: When a dominant-phenotype individual from a 3:1 cross is used in a test cross, remember that among the dominant phenotype, $\frac{2}{3}$ are heterozygous and $\frac{1}{3}$ are homozygous. This is a **conditional probability** — the ratio is 1:2, not 1:3, because you've already excluded the recessive phenotype. This shows up at every level from Honors Bio to USABO.

TIER 2 ★★ AP Biology — 2 Problems

Problem 3 | Hardy–Weinberg Equilibrium

Style: AP Biology Free Response, Unit 7 (Hardy–Weinberg). Cf. College Board 2023 AP Biology Exam guidelines.

Cystic fibrosis (CF) is an autosomal recessive condition. In a large population assumed to be in Hardy–Weinberg equilibrium, 1 in 2500 newborns is affected with CF.

(a) Calculate the frequency of the CF allele (q) and the normal allele (p).

(b) What percentage of the population are carriers (heterozygous)?

(c) A pharmaceutical company wants to screen the population. If they test 10,000 individuals, how many carriers do you expect to find?

(d) Name two conditions of Hardy–Weinberg equilibrium that are most likely violated for the CF allele in real human populations. Explain how each violation would affect allele frequency.

✓ Solution

(a) Affected individuals are homozygous recessive:

$$q^2 = \frac{1}{2500} = 0.0004, \qquad q = \sqrt{0.0004} = 0.02, \qquad p = 1 - q = 0.98$$

(b) Carrier frequency $= 2pq = 2(0.98)(0.02) = 0.0392$, i.e. approximately **3.92%**.

(c) Expected carriers in 10,000 people:

$$10{,}000 \times 0.0392 = \mathbf{392} \text{ carriers}$$

(d) Two violated conditions:

- **Natural selection is occurring:** CF homozygotes have reduced fitness (historically lethal before reproductive age), which should decrease q over time. However, there is evidence of **heterozygote advantage** — CF carriers may have increased resistance to certain infections (e.g., typhoid, cholera), which maintains the allele at a higher frequency than selection alone would predict (balancing selection).
- **Large population size is violated (genetic drift / founder effect):** In certain ethnic communities (e.g., Ashkenazi Jewish, Northern European), small ancestral populations experienced founder effects — a form of genetic drift — that elevated the CF allele frequency well above the global average. This violates the H–W requirement of an infinitely large population where stochastic fluctuations in allele frequency are negligible.

▶ Recognition Trigger

1 in N affected + autosomal recessive ⇒ Set $q^2 = 1/N$, take the square root for q. This is the single most common Hardy–Weinberg setup on AP Bio and USABO Open exams. The

carrier frequency $2pq$ is always the follow-up question.

Problem 4 | Cell Signaling & Feedback

Style: AP Biology FRQ, Unit 4 (Cell Communication) / Unit 8. Integrates endocrine physiology.

The hypothalamic-pituitary-thyroid (HPT) axis regulates metabolism through a negative feedback loop. Consider the following scenario:
A patient is found to have **elevated levels of TSH** and **low levels of T3/T4** (thyroid hormones).

(a) Is the primary defect most likely at the level of the hypothalamus, anterior pituitary, or thyroid gland? Explain your reasoning using the logic of negative feedback.

(b) If the patient were given exogenous T4 supplements, predict the effect on TSH levels and explain why.

(c) A second patient has **elevated TSH** and **elevated T3/T4**. Where is the most likely defect now? Explain how this differs from the first patient.

(d) Describe one molecular mechanism by which T3 could suppress TSH gene transcription in the anterior pituitary.

Solution

(a) The defect is at the **thyroid gland** (primary hypothyroidism).
Logic: In negative feedback, low T3/T4 should remove inhibition on the hypothalamus and pituitary, causing them to ramp up TRH and TSH production. The fact that TSH is elevated means the pituitary is responding correctly to low thyroid hormones. The thyroid itself must be failing to respond to the elevated TSH — it cannot produce adequate T3/T4.
(b) Exogenous T4 restores circulating thyroid hormone levels. T4 (and its active conversion product T3) exert **negative feedback** on the anterior pituitary and hypothalamus, suppressing TRH and TSH secretion. **TSH levels would decrease.**
(c) If both TSH and T3/T4 are elevated, the **negative feedback loop is broken at the pituitary**. Normally, high T3/T4 would suppress TSH. Since TSH remains high despite elevated thyroid hormones, the pituitary is **autonomous** — likely a **TSH-secreting pituitary adenoma** that does not respond to negative feedback.
(d) T3 enters pituitary thyrotroph cells and binds to **thyroid hormone receptors (TRs)**, which are nuclear receptors that function as ligand-activated transcription factors. The T3–TR complex binds to **negative thyroid response elements (nTREs)** in the promoter region of the TSH β-subunit gene, recruiting corepressor complexes (e.g., NCoR/SMRT) and **histone deacetylases (HDACs)**. This leads to chromatin condensation and transcriptional repression of the TSH gene.

◇ Flashy Tip

The Feedback Diagnostic Algorithm:

1. Check the hormone at the end of the axis (T3/T4, cortisol, etc.).
2. Check the tropic hormone (TSH, ACTH, etc.).
3. High tropic + low end-hormone ⇒ End-organ failure (**primary**).
4. Low tropic + low end-hormone ⇒ Pituitary/hypothalamic failure (**secondary/tertiary**).
5. High tropic + high end-hormone ⇒ Autonomous tropic secretion (adenoma) or receptor insensitivity.

Quick mnemonic: If tropic and end-hormone move in *opposite directions* (one up, one down), the defect is at the end-organ. If they move in the *same direction* (both up or both down), the defect is central or autonomous.

This algorithm works for the HPT, HPA, and HPG axes identically. Memorize the pattern, not the individual diseases.

TIER 3 ★★★ USABO Open Exam — 2 Problems

Problem 5 | Membrane Potential & Ion Channels

Style: USABO Open Exam (2016–2018 style), Cell Biology section. Cf. USABO 2016 Open Q1 (membrane potential) and 2018 Open Q7–Q10 (ion channels).

Consider the following information about a mammalian neuron at rest:

Ion	Intracellular [mM]	Extracellular [mM]
K^+	140	5
Na^+	15	145
Cl^-	10	110

(a) Using the Nernst equation $E_{\text{ion}} = \frac{61}{z}\log_{10}\frac{[\text{ion}]_{\text{out}}}{[\text{ion}]_{\text{in}}}$ at 37 °C, calculate the equilibrium potential for K^+ and Na^+.

(b) The resting membrane potential is approximately −70 mV. Explain why this value is closer to E_K than to E_{Na}.

(c) A researcher applies a drug that specifically blocks the Na^+/K^+-ATPase. Predict the short-term (seconds) and long-term (minutes to hours) effects on the membrane potential. Justify your answers.

(d) A separate experiment raises extracellular $[K^+]$ to 40 mM. Calculate the new E_K and predict the effect on neuronal excitability.

Solution

(a) For K^+ ($z = +1$):

$$E_K = \frac{61}{+1}\log_{10}\frac{5}{140} = 61 \times \log_{10}(0.0357) = 61 \times (-1.447) \approx \mathbf{-88}\text{ mV}$$

For Na^+ ($z = +1$):

$$E_{\text{Na}} = \frac{61}{+1}\log_{10}\frac{145}{15} = 61 \times \log_{10}(9.67) = 61 \times 0.985 \approx \mathbf{+60}\text{ mV}$$

(b) At rest, the membrane has much higher permeability to K^+ than to Na^+. The ratio of K^+ leak channels to Na^+ leak channels is roughly 100:1. Because the membrane potential is determined by the weighted average of equilibrium potentials (weighted by permeability), V_m sits much closer to E_K (−88 mV) than to E_{Na} (+60 mV). The **Goldman–Hodgkin–Katz equation** formalizes this: only ions with significant permeability pull the membrane potential toward their equilibrium value.

(c) **Short-term (seconds):** Minimal change. The Na^+/K^+-ATPase contributes only about −3 to −5 mV directly to the resting potential because it is electrogenic: pumping 3 Na^+ out for every 2 K^+ in creates a net outward positive current, making the interior slightly more

negative. Existing ion gradients are maintained by the enormous ion pools relative to the small fluxes per action potential.

Long-term (minutes–hours): The ion gradients dissipate. K^+ slowly leaks out, Na^+ slowly leaks in, and without the pump restoring them, $[K^+]_{in}$ decreases and $[Na^+]_{in}$ increases. The membrane depolarizes progressively, eventually reaching a state where the cell can no longer fire action potentials (**depolarization block**). Ultimately, intracellular and extracellular concentrations equilibrate and $V_m \to 0$ mV.

(d) New E_K with $[K^+]_{out} = 40$ mM:

$$E_K = 61 \times \log_{10}\frac{40}{140} = 61 \times \log_{10}(0.286) = 61 \times (-0.544) \approx \mathbf{-33}\text{ mV}$$

The resting membrane potential becomes much less negative (depolarized). Since the resting potential moves closer to the threshold for action potential firing (≈ -55 mV), the neuron becomes **hyperexcitable** — it is closer to threshold and requires a smaller stimulus to fire. This is clinically relevant in **hyperkalemia**, which can cause cardiac arrhythmias via the same mechanism.

▶ Recognition Trigger

Ion concentrations + membrane potential ⇒ Nernst equation. If they give you multiple ions and permeability info, think Goldman equation. On the USABO Open, the Nernst calculation itself is straightforward — the real test is your reasoning about what happens when you *perturb* the system (pump block, concentration change).

Problem 6 | Molecular Genetics & Gene Expression

Style: USABO Open (2017–2018 style), Genetics section. Requires Campbell Biology Ch. 17–18 depth.

A researcher isolates mRNA from human liver cells and human muscle cells. She uses reverse transcriptase to produce cDNA libraries from each tissue.

(a) Explain why the cDNA libraries from liver and muscle are different, even though both cell types contain the same genome.

(b) The researcher notices that a particular gene's cDNA from liver is 1.2 kb, while the corresponding genomic DNA is 8.5 kb. Account for the size difference.

(c) She discovers that the same gene produces a 1.2 kb cDNA in liver but a 0.9 kb cDNA in brain. The genomic sequence is identical. Propose a molecular mechanism that explains this observation.

(d) A mutation in the gene's **branch point sequence** (the conserved adenine in intron 3) is identified in a patient. Predict the effect on the mature mRNA and protein product. Be specific about the molecular mechanism affected.

Solution

(a) The cDNA libraries differ because they reflect the **transcriptome**, not the genome. Different cell types express different subsets of genes due to **differential gene expression** — mediated by tissue-specific transcription factors, epigenetic modifications (DNA methylation, histone modifications), and chromatin remodeling. Liver cells transcribe genes for albumin, cytochrome P450 enzymes, etc., while muscle cells transcribe genes for myosin, actin, and creatine kinase. The cDNA library captures only the mRNAs present in each tissue.
(b) The cDNA (1.2 kb) is synthesized from mature mRNA, which has been processed by **RNA splicing** to remove introns. The genomic DNA (8.5 kb) includes both exons and introns. The difference ($8.5 - 1.2 = 7.3$ kb) represents the total length of intronic sequence that was spliced out, plus portions of the transcript not captured in the cDNA.
(c) This is **alternative splicing**. The same pre-mRNA is processed differently in liver vs. brain: different exons are included or excluded depending on tissue-specific splicing factors (e.g., SR proteins and hnRNPs). In brain, one or more exons are skipped (exon skipping) or an alternative $5'/3'$ splice site is chosen, yielding a shorter mRNA. This produces **tissue-specific protein isoforms** from a single gene.
(d) The branch point adenine is essential for spliceosome-mediated splicing. During splicing, the $2'$-OH of the branch point adenine performs a **nucleophilic attack** on the $5'$ splice site, forming a **lariat intermediate**. If this adenine is mutated:

- The spliceosome (specifically U2 snRNP) cannot properly recognize or bind the branch point.
- **Intron 3 is retained** in the mature mRNA (intron retention).
- The retained intron likely introduces a **premature stop codon** (since intron sequences are not in the reading frame), producing a truncated, nonfunctional protein.
- Alternatively, the aberrant mRNA may be degraded by **nonsense-mediated mRNA decay (NMD)** before translation occurs.

◇ Flashy Tip

The Same Gene, Different Product Framework: When a problem gives you the same gene but different outcomes in different tissues or conditions, cycle through this checklist:

1. **Alternative splicing** (different exon combinations ⇒ different mRNA lengths)
2. **Alternative promoters** (different $5'$ exons / start sites)
3. **Alternative polyadenylation** (different $3'$ UTR lengths)
4. **Post-translational modification** (same protein, different processing)
5. **RNA editing** (enzymatic base changes in mRNA, e.g., APOB: C→U editing in intestine creates a premature stop codon, producing shorter ApoB-48 vs. full-length ApoB-100 in liver — a classic USABO topic)

On USABO, if the mRNA sizes are different but the gene is the same, **alternative splicing is almost always the answer**.

TIER 4 ★★★★ USABO Semifinal Exam — 2 Problems

Problem 7 | Experimental Design: Signal Transduction

Style: USABO Semifinal Part D (short answer/essay). Cf. USABO Semifinal format (2016–2018): interpretation of experimental data + experimental design. Requires Alberts' MBoC, Ch. 15.

A research team is studying a novel receptor tyrosine kinase (RTK) called **XKR1** in mammalian cells. When the ligand Factor-X binds to XKR1, cells proliferate. The team hypothesizes that XKR1 signals through the Ras–MAPK pathway.

They perform the following experiments:

Experiment 1: Cells are treated with Factor-X. Western blot shows phosphorylation of ERK1/2 (a downstream MAPK) within 5 minutes.

Experiment 2: Cells expressing a **dominant-negative Ras** (RasN17, which binds GDP but cannot exchange to GTP) are treated with Factor-X. ERK1/2 phosphorylation is abolished.

Experiment 3: Cells expressing a **constitutively active Ras** (RasV12, locked in GTP-bound state) show ERK1/2 phosphorylation even without Factor-X. However, proliferation is only 40% of the level seen with Factor-X treatment of wild-type cells.

(a) Interpret Experiments 1 and 2 together. What do they demonstrate about the pathway?

(b) Explain the molecular mechanism by which RasN17 acts as a dominant negative.

(c) Experiment 3 shows that constitutively active Ras alone is insufficient for full proliferation. Propose two distinct molecular explanations for the remaining 60% of proliferative signaling.

(d) Design a single experiment to test whether XKR1 also activates the PI3K/Akt pathway in parallel. Specify the treatment, control, readout, and expected result.

✓ Solution

(a) Experiment 1 establishes that Factor-X → XKR1 activation leads to ERK1/2 phosphorylation (the pathway is active). Experiment 2 shows this phosphorylation requires functional Ras, since dominant-negative Ras blocks it completely. Together, they demonstrate that **XKR1 signals to ERK1/2 through Ras** — Ras is a *necessary* intermediate in this pathway.

(b) RasN17 acts as a dominant negative through a **competitive sequestration mechanism**. Ras activation normally requires the GEF (guanine nucleotide exchange factor) Sos to catalyze GDP → GTP exchange. RasN17 has higher affinity for Sos than wild-type Ras but cannot complete the exchange to GTP. It therefore **titrates away the limited pool of Sos**, preventing endogenous wild-type Ras from being activated. Since the dominant negative is overexpressed, it outcompetes wild-type Ras for Sos binding, creating a dominant loss-of-function.

(c) Two explanations for the remaining 60% of signaling:

Explanation 1 — Parallel pathway activation: XKR1 likely activates additional signaling cascades beyond Ras–MAPK, such as the PI3K/Akt pathway or the PLCγ/PKC pathway. These pathways independently promote cell proliferation through distinct transcription factor activation (e.g., Akt phosphorylates and inactivates GSK-3β, stabilizing cyclin D1; PLCγ

generates IP_3 and DAG for Ca^{2+}/PKC signaling). Constitutively active Ras cannot activate these parallel arms because they branch at the level of the receptor, not downstream of Ras.
Explanation 2 — Signal dynamics matter: Constitutive Ras activation provides a tonic, sustained signal, whereas normal Factor-X stimulation produces **pulsatile, temporally regulated** ERK activation. Many proliferative genes require specific patterns of ERK activity (e.g., transient vs. sustained, oscillatory patterns) to properly activate transcription factors like Fos/Jun. Constant RasV12 signaling may trigger negative feedback loops (e.g., Sprouty, DUSP phosphatases) that attenuate the signal to suboptimal levels.

(d) Experimental design to test PI3K/Akt pathway:

Treatment:	Treat wild-type cells with Factor-X for 5, 15, 30 min.
Neg. control:	Untreated wild-type cells (no Factor-X).
Pos. control:	Cells treated with insulin (known PI3K/Akt activator).
Specificity ctrl:	Pre-treat with LY294002 (PI3K inhibitor) before Factor-X.
Readout:	Western blot for phospho-Akt (Ser473), a direct readout of PI3K pathway activation.
Secondary:	BrdU incorporation assay ± LY294002 to confirm PI3K/Akt contributes to the proliferation phenotype.
Expected:	Factor-X ↑ phospho-Akt (absent in untreated control, blocked by LY294002, present in insulin positive control). LY294002 should partially reduce Factor-X-induced proliferation.

▶ Recognition Trigger

Dominant negative or constitutively active mutant in a signaling pathway ⇒ You are being tested on **epistasis logic**. Dominant negatives block downstream signaling; constitutively active mutants bypass the need for upstream signal. The key question is always: *Is this component necessary? Is it sufficient?* Experiments 1+2 test necessity, Experiment 3 tests sufficiency.

Problem 8 | Molecular Evolution & Phylogenetics

Style: USABO Semifinal Part A/B, Genetics & Evolution section. Cf. USABO 2016 Semifinal Q34 (amino acid substitutions) and IBO theoretical exam phylogenetics problems.

You are given partial cytochrome *c* amino acid sequences from four species:

Species	Positions 39–48									
	39	40	41	42	43	44	45	46	47	48
Human	T	G	P	N	L	H	G	L	F	G
Chimpanzee	T	G	P	N	L	H	G	L	F	G
Dog	T	G	P	N	L	H	G	L	F	G
Tuna	T	G	P	N	L	H	G	I	F	G
Yeast	T	G	P	N	L	H	G	I	S	G

(a) Based on these data alone, which position(s) are most likely functionally constrained? Justify using the concept of purifying selection.

(b) At position 46, the human/chimp/dog lineage has Leu (L) while tuna has Ile (I) and yeast has Ile (I). Using the **principle of parsimony**, determine the most likely ancestral amino acid at this position and indicate on which branch the substitution occurred.

(c) Cytochrome c is one of the most conserved proteins across eukaryotes. Explain why the rate of amino acid substitution in cytochrome c is much lower than in fibrinopeptides (which evolve $\sim$100$\times$ faster).

(d) A student claims: "Since human and chimpanzee cytochrome c sequences are identical, humans and chimps must have diverged very recently." Critique this argument.

Solution

(a) Positions 39–44 (**TGPNLH**) and position 48 (**G**) are identical across all five species, spanning >1 billion years of divergence (from yeast to mammals). This extreme conservation indicates strong **purifying (negative) selection**: mutations at these positions are deleterious and are eliminated from the population. These residues likely participate directly in the protein's essential function — heme binding, interaction with cytochrome c oxidase/reductase, or maintaining the critical 3D fold.

Position 47 (**F** in all species except yeast) is also highly conserved and likely constrained.

(b) At position 46: Human/Chimp/Dog = L; Tuna = I; Yeast = I.

By parsimony, **Ile (I) is ancestral**. Three of the five species (including the most distant outgroup, yeast) retain Ile. The most parsimonious reconstruction requires only **one substitution event**: I $\rightarrow$ L on the branch leading to the mammalian clade (after the divergence of bony fishes and mammals, but before the divergence of dogs from the primate lineage). The alternative (L ancestral with two independent reversions to I in tuna and yeast) requires two events — less parsimonious.

(c) Cytochrome c functions as an electron carrier in the mitochondrial electron transport chain, physically interacting with both Complex III (cytochrome bc_1) and Complex IV (cytochrome c oxidase). It must maintain precise heme coordination (covalent attachment via CXXCH motif), correct surface charge distribution for electrostatic docking with its partner complexes, and a stable tertiary fold for proper redox potential. Nearly every residue is constrained by one of these requirements, so most mutations are deleterious.

Fibrinopeptides, by contrast, are cleaved off fibrinogen during blood clotting and discarded.

They have essentially no functional constraint after cleavage — nearly all amino acid substitutions are selectively neutral, so they accumulate mutations at close to the **maximum neutral rate**.

(d) The student's argument confuses protein sequence identity with divergence time. Several factors can produce identical protein sequences in species that diverged millions of years ago:

- **Strong purifying selection:** If the protein is highly constrained (as cytochrome c is), even millions of years of divergence may produce zero amino acid changes because all non-synonymous mutations are lethal or deleterious.
- **Synonymous (silent) substitutions:** The DNA sequences may differ at synonymous sites — particularly at third-codon (wobble) positions, where the degeneracy of the genetic code allows nucleotide changes without altering the amino acid. Comparing nucleotide sequences (especially K_s, the rate of synonymous substitution) would reveal the true divergence time.
- **Protein length matters:** A short segment (10 residues) provides very low resolution. The probability of observing zero substitutions by chance is higher for short sequences.

In fact, humans and chimps diverged ~6–7 million years ago, and their cytochrome c proteins are identical at all 104 positions, yet their DNA sequences differ at several synonymous sites.

◇ Flashy Tip

The Constraint = Conservation Principle: Across all of biology — from enzyme active sites to ribosomal RNA to HOX gene order — the pattern is universal: *the more functionally important a sequence is, the slower it evolves.*

On the USABO Semifinal, any question that gives you aligned sequences across species is testing whether you can: (1) identify conserved vs. variable positions, (2) infer function from conservation, and (3) reconstruct evolutionary history using parsimony. The tool is always the same: **compare, count changes, minimize substitution events.**

Summary of Difficulty Progression

Tier	Key Skill Tested	Preparation Source
HONORS	Recall + basic application	Campbell Biology, Units I–III
AP BIO	Multi-step calculation + explanation	Campbell Ch. 23, AP CED Unit 7
OPEN	Quantitative reasoning + perturbation	Campbell + past USABO Opens (2003–2018)
SEMI	Experimental design + synthesis	Alberts' MBoC + USABO Semifinal archives

Additional Resources

- **USABO Archive:** https://usabo-trc.org
- **IBO Paper Archive:** https://ibo-info.org/en/info/papers.html

Olympiad Edge
Joshua Gao • Huntington, NY
Text: 631-833-0226