FEATURE ARTICLE

Integrating AI in Research & Project-Based Learning in Biology Instruction

SARMISTHA GHOSH

Abstract

This article describes classroom-tested strategies for integrating artificial intelligence (AI) tools into high school biology through research and project-based learning. Focusing primarily on protein folding and structure-function relationships, it provides concrete lesson models using AlphaFold, UniProt, PDB, and visualization tools. It also discusses student challenges, required scaffolds, assessment methods, and ethical considerations. Broader applications of AI in CRISPR and biotechnology are briefly introduced to highlight future directions in science education.

Key Words: *artificial intelligence (AI); protein folding; AlphaFold; deep learning; CRISPR-Cas9; Guide RNA Design; gene expression regulation; structure-function relationship; project-based learning.*

AI isn't just an abstract technology, it's transforming real-world biology, from predicting folding of disease-related proteins to designing life-saving drugs.

○ Introduction

For decades, understanding protein structure from amino acid sequences (a key outcome of gene expression) was one of biology's grand challenges. Traditional methods such as X-ray crystallography and cryo-electron microscopy are time-consuming and expensive, and unimaginable in a public school science lab.

In 2021, DeepMind's AlphaFold (Figure 1), an AI model, revolutionized the field by accurately predicting protein 3D structures from sequences. This development leveraged deep learning, a type of artificial intelligence, to solve a central dogma problem: linking sequence (DNA → mRNA → amino acid chain) to function (3D protein shape). DeepMind's AlphaFold has demonstrated unprecedented accuracy in predicting protein three-dimensional structures from amino acid sequences, marking a major breakthrough in computational biology (Jumper et al., 2021; Varadi et al., 2022).

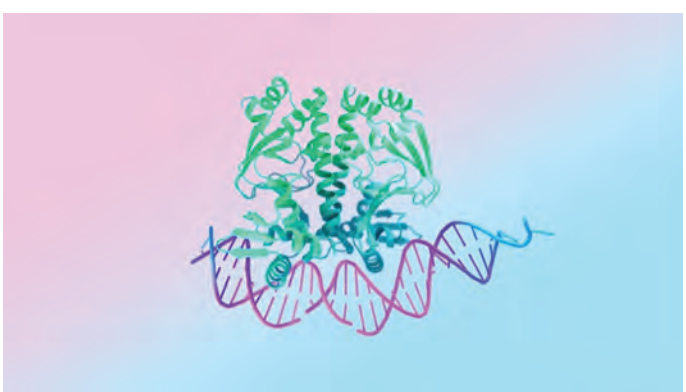

Figure 1. AlphaFold 3 by Google DeepMind.

○ Section I: Driving Questions

The unit was structured around the following driving questions, which guided student inquiry and discussion throughout the protein folding and AI activities:

- How does AI, such as AlphaFold, use data from thousands of known protein structures to predict new ones?
- Why is accurate protein folding essential for proper cellular function?
- What are the implications of AI-driven protein prediction on medicine, such as designing drugs or understanding genetic diseases?
- How might errors in protein folding (e.g., misfolding diseases such as Alzheimer's) be better understood with AI predictions?

Students responded to these questions during class discussions and presentations. For example, after modeling CFTR protein misfolding, they debated why accurate folding is crucial for channel function and how AI could guide drug design. The discussion was based on a foundational concept in biology:

The American Biology Teacher, Vol. 88, No. 3, pp. 159–165, ISSN 0002-7685, electronic ISSN 1938-4211. © 2026 by National Association of Biology Teachers. All rights reserved. Please direct all requests for permission to photocopy or reproduce article content through the University of California Press's Reprints and Permissions web page, https://online.ucpress.edu/journals/pages/reprintspermissions. DOI: https://doi.org/10.1525/abt.2026.88.3.159.

"The relationship between structure and function is fundamental in biology, especially when it comes to proteins, whether they are enzymes, hormones, regulatory proteins, inducers, or other types of proteins. The specific shape and arrangement of a protein determine how effectively it performs its role in the cell."

Protein Folding

In my biology class, students grasp the fundamentals of protein folding by exploring how disulfide bonds, hydrogen bonds, ionic charges, and polar versus nonpolar functional groups contribute to a protein's three-dimensional shape. To make this concept tangible, I use engaging simulations with beads, pipe cleaners, and other craft materials (Figure 2). Activity Example: Students worked in small groups using the Amino Acid Starter Kit (commercially available from 3D Molecular Designs) (Figure 3) and compared their physical models with AlphaFold predictions. They needed scaffolds for interpreting polar vs. nonpolar interactions, but most could successfully identify hydrogen bonding sites by the end of the lesson. This combination of hands-on and visual learning helps students connect abstract molecular concepts to real-world biological functions. Hands-on modeling approaches such as physical protein construction have been shown to significantly improve student understanding of structure–function relationships in molecular biology (Cooper & Oliver-Hoyo, 2017; White, 2012).

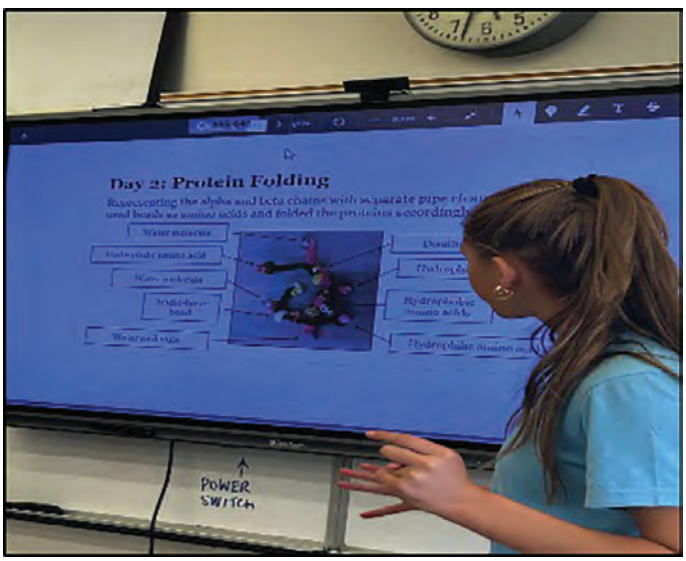

Figure 2. Student in my class presenting on the basics of protein folding with the help of a model she made with pipe cleaners and beads.

Figure 3. Amino Acid Starter Kit.

○ Section 2: Case Study—Cystic Fibrosis (CFTR)

We explored protein folding through a case study on cystic fibrosis. Students accessed the CFTR protein (UniProt accession P13569) and examined AlphaFold's predicted structure (PDB ID: 7TGZ) alongside bead and pipe-cleaner models they constructed. We discussed the ΔF508 deletion, how ER degradation occurs, and how AI predictions help visualize structural loss.

By examining how mutations in the CFTR gene lead to misfolding of the CFTR protein, and consequently its malfunction as a chloride channel, we connected the molecular details of protein structure to the physiological symptoms of the disease. Students used modeling activities and simulations to visualize how the ΔF508 mutation causes misfolding, degradation of the protein in the endoplasmic reticulum, and loss of function at the cell membrane. This case study allowed students to see first-hand how even a single amino acid deletion can disrupt protein folding, leading to devastating health consequences, and highlighted the critical relationship between structure and function in biology (reference: https://www.cell.com/trends/molecular-medicine/fulltext/S1471-4914(11)00191-2). To deepen students' understanding of protein synthesis, folding, and structure-function relationships, we incorporate a variety of interactive simulations and molecular visualization tools in our AP Biology class. We use the "Transcribe and Translate a Gene" simulation from Learn Genetics (https://learn.genetics.utah.edu/content/basics/transcribe/), which allows students to practice transcription and translation processes. For 3D molecular visualization, we utilize Jmol, an open-source Java viewer for chemical structures in three dimensions, and its HTML5 version, JSmol, which enables embedding interactive models directly into web pages (source: Jmol: http://www.jmol.org/). These tools empower students to manipulate and explore protein structures, reinforcing their grasp of how the unique arrangement of amino acids drives folding and determines biological function.

○ Section 3: Using Data Sources & Visualization Tools

To support analysis of protein structure and folding, students worked with several data sources and visualization tools, including the following:

- UniProt (protein sequences and AlphaFold predictions)
- PDB (experimentally determined structures with IDs)
- Jmol/JSmol (interactive visualization)
- Tinkercad (student-designed 3D representations)

Visual molecular representations are essential for helping students connect abstract molecular concepts to cellular processes across different biological scales (Goodsell, 2009).

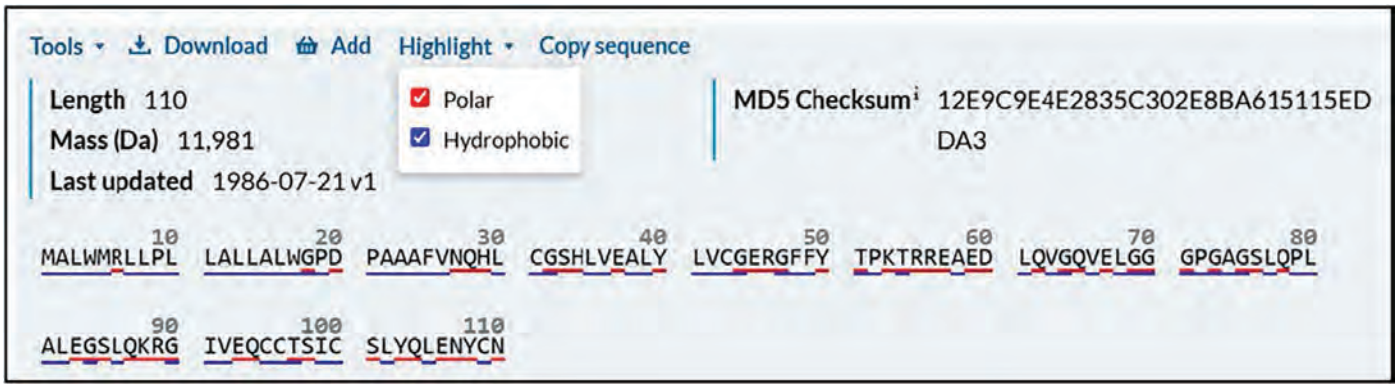

Figure 4. Amino acid sequence of human insulin displayed from the UniProt database, with polar and hydrophobic residues highlighted to illustrate how differences in amino acid properties contribute to protein folding and structure–function relationships.

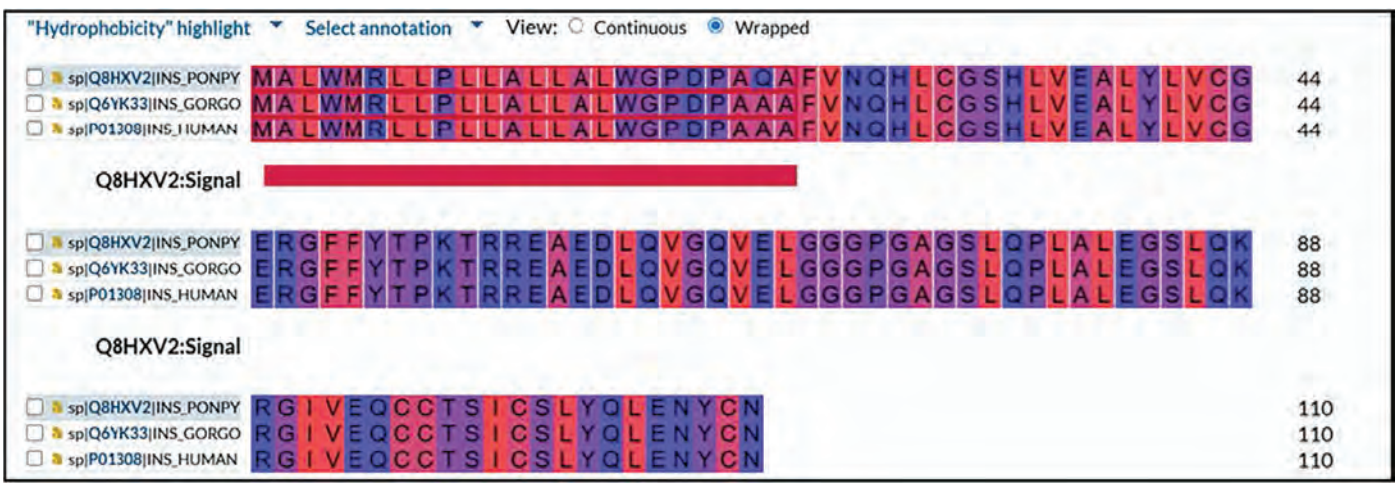

Figure 5. Insulin A and B Chains Comparison—Human, Western Lowland Gorilla, and Bornean Orangutan.

Lesson Flow

This sequence allows students to move from raw sequence data to AI prediction, comparison with experimental structures, and finally physical and virtual modeling. It highlights how AI complements, but does not replace, laboratory science. We find the amino acid sequences from UniProt and the NCBI Protein Database, for example, for human insulin (Figure 4).

A comparison of the insulin A and B chain sequences across humans, Western lowland gorillas, and Bornean orangutans highlights the strong conservation of amino acid composition and hydrophobic regions essential for insulin structure and function (Figure 5).

To illustrate the conservation of disulfide bond, forming residues critical for insulin folding, students examined aligned insulin sequences from multiple species with cysteine residues highlighted (Figure 6).

Students used the Protein Data Bank (PDB; https://www.rcsb.org/) to examine experimentally determined protein structures and then employed Tinkercad (https://www.tinkercad.com/) to create, present, and explain three-dimensional models, allowing them to visualize protein folding patterns and articulate structure–function relationships to their peers (Figure 7).

○ How to Incorporate AI into Your Protein Folding Unit

AI-Powered Structure Prediction (AlphaFold)

Introduce students to AlphaFold (https://alphafold.ebi.ac.uk/ and https://alphafoldserver.com/welcome) Google DeepMind's AI model that predicts protein 3D structures directly from amino acid sequences and it's free for non-commercial research.

Introduce students to AlphaFold (https://alphafold.ebi.ac.uk/ and https://alphafoldserver.com/welcome), Google DeepMind's AI model that predicts three-dimensional protein structures directly from amino acid sequences and is freely available for non-commercial research. Students can input a sequence (e.g., insulin or CFTR) and view the predicted structure and compare AlphaFold's predictions with experimentally determined structures in the PDB (Figure 8).

Discussion Prompt

How does AI prediction compare to lab-based results? What does this mean for future drug development or disease research?

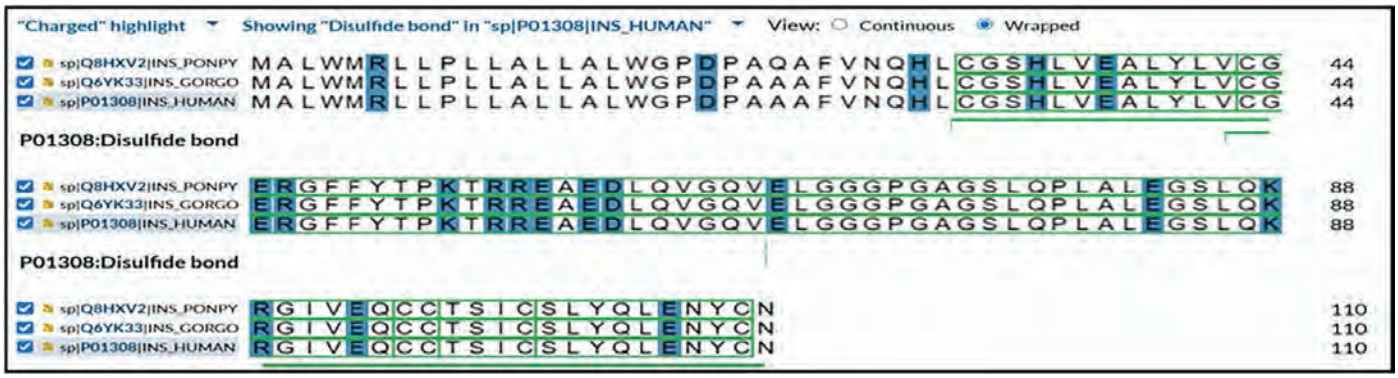

Figure 6. Multiple sequence alignment of insulin proteins from human (*INS_HUMAN*), gorilla (*INS_GORGO*), and pony (*INS_PONPY*) generated using UniProt. Conserved cysteine residues involved in intra- and interchain disulfide bond formation are highlighted, demonstrating their essential role in insulin structure and function.

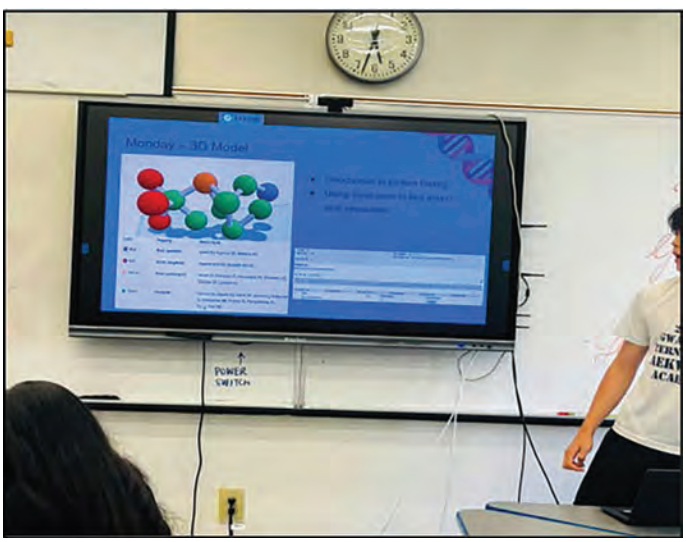

Figure 7. A student explaining protein structure and folding using Tinkercad.

Figure 8. AlphaFold-predicted three-dimensional structure of the cystic fibrosis transmembrane conductance regulator (CFTR) protein visualized using the Mol* 3D Viewer, with the corresponding amino acid sequence.

AI Image Analysis of Student Model

Use AI-powered image recognition tools such as Google Lens or Microsoft Azure Computer Vision API to analyze photos of student-built models. Students can label features (disulfide bridges, hydrophobic pockets) on their physical or Tinkercad models, upload photos, and let AI assist in checking whether the structures match expected folding patterns, enabling students to see how computational tools can analyze biological structures.

AI-Assisted Sequence Analysis

After retrieving sequences from UniProt or NCBI, students can use AI-powered sequence alignment tools such as ColabFold, which combines AlphaFold with Google Colab's GPU capabilities for fast, AI-assisted predictions. Students align human insulin with homologs in other primates (e.g., Figure 5 comparison) and generate AI-based structural models for each species. Students evaluate differences in folding between species and hypothesize about evolutionary or functional consequences.

○ Section 4: CRISPR & AI (Optional Extension)

This optional extension explores how artificial intelligence can enhance CRISPR-based gene editing by improving the design of guide RNAs (gRNAs) and predicting editing outcomes. The development of CRISPR-Cas9 genome editing technology has transformed biotechnology by enabling precise and efficient modification of DNA sequences in living organisms (Doudna & Charpentier, 2014; Hsu et al., 2014). Students investigate how AI tools support gene knockouts, activation, and silencing by increasing on-target efficiency and reducing off-target effects. This extension aligns with AP Biology Topic 3.5, which focuses on the regulation of gene expression and protein synthesis. These advances build upon foundational discoveries in CRISPR biology that established the system as a versatile tool for genome engineering and functional genomics research (Doudna & Charpentier, 2014; Hsu et al., 2014).

1. Designing Guide RNAs (gRNAs) with AI

CRISPR's power depends on the specificity of its guide RNA. Designing an effective gRNA that maximizes on-target activity while

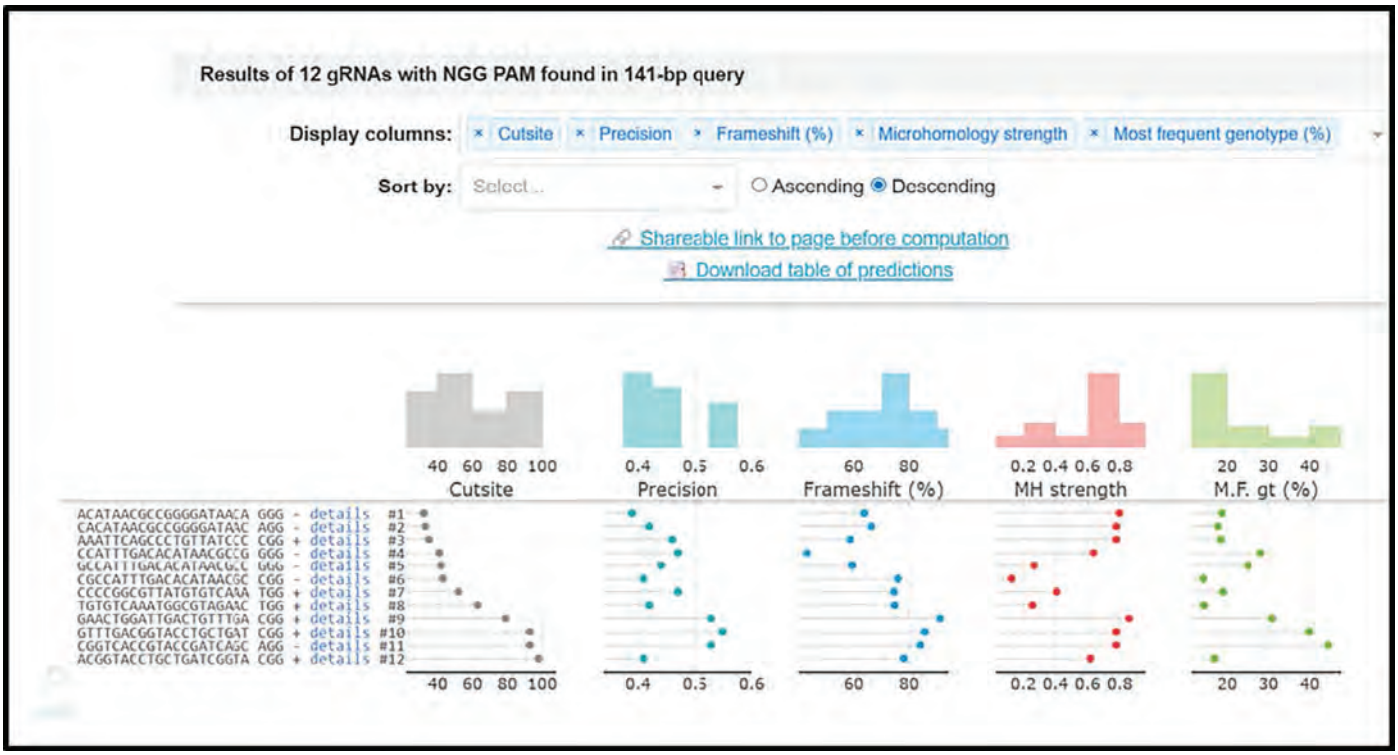

Figure 9. Tools such as inDelphi (https://indelphi.giffordlab.mit.edu/batch) and FORECasT use AI to predict DNA repair outcomes, which is crucial for reliably knocking out genes or avoiding unwanted mutations.

minimizing off-target effects is challenging. AI tools—especially deep learning models like DeepCRISPR, sgRNA Designer(Benching: https://www.benchling.com/ai), Azimuth, CRISPR-Net, and CRISPR-DO—can enable the following:

- Predict gRNA efficiency (how well it cuts the target)
- Evaluate off-target potential by scanning entire genomes
- Suggest gRNA sequences optimized for specific species or cell types
- Customize gRNAs for knockouts (frameshifts), activation (CRISPRa), or repression (CRISPRi)

The above dramatically reduces trial-and-error, saving time and resources.

2. Predicting CRISPR Outcomes

AI can forecast what happens after Cas9 cuts the DNA (Figure 9):

- Will it cause a frameshift mutation (knockout)?
- What kind of insertions or deletions (indels) might occur?

3. CRISPRa & CRISPRi Optimization

CRISPRa (activation) and CRISPRi (interference/silencing) don't cut DNA but modulate gene expression by recruiting activators or repressors. AI models can do the following:

- Recommend optimal target sites in promoter or enhancer regions
- Predict which gRNAs will most effectively boost or silence expression
- Adjust designs based on epigenetic data, since chromatin accessibility influences success

4. Prioritizing Target Genes

When we want to activate, silence, or knock out genes in a complex system (e.g., cancer cells), AI can do the following:

- Analyze multi-omics datasets (RNA-seq, proteomics, CRISPR screens)
- Identify genes critical for disease phenotypes using AI-based feature selection or network analysis (e.g., machine learning tools for synthetic lethality screens). This helps choose the best CRISPR targets for therapeutic or research purposes.

5. Off-Target Prediction & Minimization

AI models can scan genomes for sequences similar to the intended target to predict off-target sites more accurately than rule-based methods alone. This is vital for safety in therapeutic applications. Popular tools include the following:

- CRISPRoff (not just a tool but a method to silence genes epigenetically without cutting)
- DeepCRISPR (deep neural network for off-target risk assessment)

6. Integration with Lab Automation & Robotics

I clarified that AlphaFold provides predicted structures (with confidence scores), not experimental ones, and students compared these with PDB experimental structures where available.

In high-throughput CRISPR screens, AI-driven platforms can design experiments, analyze CRISPR screening data, and learn from results to iteratively improve gRNA libraries. Companies such as

Inscripta and Synthego use AI and robotics to automate CRISPR workflows, dramatically speeding up discovery.

The following are examples of AI-CRISPR integration in practice:

- Drug discovery: AI-guided CRISPR screens identify new drug targets.
- Functional genomics: AI analyzes CRISPR knockout/activation screen data to map gene functions.
- Precision medicine: Predict patient-specific responses to gene editing by combining AI with CRISPR.

AI supercharges CRISPR by doing the following:

- Making gRNA design smarter and more precise
- Predicting editing outcomes and avoiding off-target effects
- Guiding functional studies to knock out, activate, or silence genes reliably

This synergy accelerates research, therapeutic development, and synthetic biology, bringing us closer to truly personalized gene therapies.

○ Why This Matters

This integration shows students that AI isn't just an abstract technology, it's transforming real-world biology, from predicting folding of disease-related proteins to designing life-saving drugs. It also empowers students with future-ready skills.

The following considerations highlight the importance of setting thoughtful boundaries when integrating AI into biology education:

1. While AI can revolutionize protein modeling and disease research, it must not replace hands-on, conceptual learning essential for deep understanding.
2. The line should be drawn when AI tools discourage critical thinking or create overreliance on automated answers.
3. We must balance AI's power with ethical responsibility, ensuring technology serves humanity without enabling misuse.
4. Students should learn both the potential and the limitations of AI, recognizing when human oversight is crucial.
5. Ultimately, drawing the line means using AI as a tool to enhance, not overshadow, the curiosity, reasoning, and creativity central to science education.

○ AP Biology Connections

Topic 2.5 & 2.6: Structure and function of proteins, and enzymes
Topic 3.5: Regulation of gene expression and protein synthesis
Science Practices: Analyzing models, interpreting data, evaluating technological impacts on science

○ Suggested Classroom Activities/ Assignments

Interactive Exploration

Have students compare predicted AlphaFold models with experimentally determined structures (freely available at the AlphaFold Protein Structure Database).

Data Analysis

Assign students to analyze sequence-structure relationships by selecting a disease-associated protein and exploring how mutations could impact its predicted folding.

Discussion/Debate

Have students debate ethical considerations: Could AI-designed proteins pose biosecurity risks? What role should AI play in biotechnology regulation?

Model Evaluation

Students critique the strengths and limitations of using AI in biology, including whether the accuracy of AlphaFold predictions reduces the need for traditional laboratory experiments. This evaluation can be extended to CRISPR by examining how gene editing alters protein sequences and considering whether AI tools could predict whether such edits lead to protein misfolding. Students also explore AI's role in predicting viral protein structures, such as the SARS-CoV-2 spike protein, which has been critical for vaccine design.

Students can extend this analysis of protein sequence–structure relationships to CRISPR by exploring how gene editing alters protein sequences and considering whether AI tools could predict whether such edits lead to protein misfolding. This instructional analysis can be further broadened by examining AI's role in predicting viral protein structures, such as the SARS-CoV-2 spike protein, which has been critical for vaccine design.

○ Conclusion

Artificial intelligence has opened new avenues for making abstract molecular biology concepts accessible in high school classrooms. By integrating AI tools such as AlphaFold, UniProt, and PDB with hands-on modeling kits and interactive simulations, students can experience how data-driven methods illuminate the relationship between protein sequence, structure, and function. This approach not only strengthens conceptual understanding but also nurtures critical thinking, collaboration, and data literacy, skills essential for 21st-century science.

Importantly, AI should not replace traditional inquiry-based learning but rather complement it, providing students with opportunities to compare computational predictions with experimental evidence and classroom models. Structured case studies such as CFTR misfolding illustrate both the power and the limitations of AI, helping students recognize the importance of accuracy, ethical responsibility, and human oversight in science.

Extending this framework to CRISPR and other biotechnologies highlights the breadth of AI's applications and allows students to engage with current scientific debates about ethics, safety, and future directions. Ultimately, integrating AI into project-based learning prepares students not only to understand core AP Biology content but also to see themselves as contributors to an evolving scientific landscape shaped by emerging technologies.

References

Cooper, R. A., & Oliver-Hoyo, M. T. (2017). Creating three-dimensional protein models to promote student understanding of protein structure and function. *Journal of Chemical Education*, *94*(1), 72–78. https://doi.org/10.1021/acs.jchemed.6b00231

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096. https://doi.org/10.1126/science.1258096

Goodsell, D. S. (2009). Visual methods from molecules to cells. *The American Biology Teacher*, *71*(9), 549–556. https://doi.org/10.2307/20565382

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*(6), 1262–1278. https://doi.org/10.1016/j.cell.2014.05.010

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., … & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., … & Velankar, S. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, *50*(D1), D439–D444. https://doi.org/10.1093/nar/gkab1061

White, H. B. (2012). A structural bioinformatics approach to teaching the central dogma of molecular biology. *Biochemistry and Molecular Biology Education*, *40*(3), 149–158. https://doi.org/10.1002/bmb.20587

SARMISTHA GHOSH (sarmisthaghosh@yahoo.com) is an AP and Honors biology teacher at Hawthorn Academy of Health Sciences in Charlotte, North Carolina.

CALL FOR NOMINATIONS

We are looking for the next leaders of NABT.

Open positions in 2026 are listed below. Candidates for president-elect alternate between the pre-college, two-year college and four-year college/university communities.

Nominations from the pre-college level are sought for president in this election.

NABT is also looking for volunteers to serve on committees and in other capacities within the association.

Candidates for NABT Office should have: (1) evidence of active participation in the NABT such as previous service as an elected officer, committee chairperson or member, section of affiliate leader, etc. (2) at least five years of continuous membership in NABT; and (3) five years experience teaching biology, life science, or science education.

Positions Available

President-Elect (Pre-College)
Director-at-Large
Region IV Coordinator (IA, KS, MN, MO, NE, ND, SD)
Region V Coordinator (KY, NC, SC, TN, WV)
Region IX Coordinator (AK, CA, HO, OR, WA, Pacific Territories)

NOMINATE YOURSELF! Who else knows your interests and qualifications as well as you do?

Nomination deadline
March 15, 2026
Nominations accepted online at
www.nabt.org/About-Leadership-Opportunities