IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Three-Dimensional Molecular Modeling of Amino Acid Sequences and Mutations to Enrich Biophysics Education

Alan Kim, Avi Lekkelapudi

Bellarmine College Preparatory, 960 West Hedding Street, San Jose, California 95126, USA; alanmhkim@gmail.com, avi.lekkelapudi@gmail.com
Mentor: Sarah Baysden

ABSTRACT: In traditional science high school curricula, biological and physical concepts are limited to abstract textbook descriptions and lectures, leaving students with superficial comprehension due to limited priority on experimentation and exploration. Although traditional teaching methods commonly isolate the concepts covered by physics, chemistry, and biology, the real world transcends those boundaries, namely in fields like biophysics, where these concepts converge. Hence, we developed a protein visualization app that transforms abstract ideas such as the protein structure, the central dogma of molecular biology, Deoxyribonucleic acid (DNA) mutations, and the sequence-to-structure relationship into interactive learning experiences using 3D molecular models. Our study implemented the app within the curriculum of an honors biology class, and 53 students rated their comprehension of three biophysics topics on a scale of 1-5 before and after using the app. The results highlight that the app enhanced students' comprehension of amino acid structures, DNA to protein relationships, and mutations, particularly for those with a moderate foundation of prior knowledge in biophysics. Overall feedback highlighted the app's potential as an educational tool for enriching learning in biophysics and related fields by fostering comprehension and creative problem-solving skills. By bridging the gap between abstract education and molecular-level understanding, our protein visualization app has the potential to elevate biophysics education.

KEYWORDS: Physics and Astronomy, Biological Physics, Online Learning, 3D Protein Structure Visualization, Amino Acid Mutation Modeling.

Introduction

In high school classrooms, the multifaceted and intricate nature of biophysical processes is often simplified into simple conceptual explanations due to a curriculum based on building a broad conceptual understanding of scientific principles, as well as time constraints. As a result, with little to no experimentation, students are often left with a surface-level understanding of key biophysical concepts.[1-3] For example, students abstractly learn that Deoxyribonucleic acid (DNA) codes for proteins, but lack the depth of visualizing dynamic processes by which linear sequences of nucleotides fold into 3D protein structures.[4] Additionally, students perceive the central dogma of molecular biology, a flow of genetic information from DNA to RNA to proteins, through a linear series of steps with limited exploration of interactions at an atomic level that catalyze those processes, vital to everyday life. When learning about mutations, discussions commonly emphasize changes within genetic sequences in a series of letters and numbers without the combination with alterations on 3D protein structure and function, which are necessary to understand the full-scale impact of seemingly minor genetic changes.[5]

Research highlights that traditional biology education tends to emphasize memorization of facts rather than helping students obtain a deeper understanding of underlying mechanisms, reinforcing an abstract and superficial comprehension of biological principles.[6,7] However, understanding abstract biophysical concepts requires more than theoretical knowledge; it necessitates visualization at a molecular level. Visual aids are especially crucial in biophysics, a field that explains biological functions through the physical properties of molecules. This gap in biophysics education stresses the need for new, comprehensive styles of teaching that embrace the integration of 3D visualizations and experimentation with biophysical processes. Visualization of complex processes helps students connect the dots between concepts in biology, chemistry, and physics by making abstract ideas more tangible.[8-10] For instance, understanding the biophysical mechanisms governing protein function requires a deep grasp of protein structure.[11] Implementation with Artificial Intelligence and Machine Learning visualization tools, including AlphaFold and ESMFold, enabled accurate predictions of millions of 3D protein structures, leading to a surge in available data, specifically over 74,000 models in the Protein Data Bank (PDB) archive.[12-14] Furthermore, tools like Jmol and PyMOL are implemented in academic settings to elevate the learning experience by deepening understanding, curiosity, and passion for science and technology.[4,15-17] While such tools are comprehensive and effective, their complexity often makes them less suitable for high school students, who need more user-friendly tools due to differing comprehension levels of core concepts. Consequently, this research article, where we developed a student-friendly protein visualization tool, aims to explore how advanced 3D visualization tailored for biophysics education helps high school students obtain valuable skills necessary to fluently navigate biophysics and related fields.

DOI: 10.36838/IJHSR84.15

Protein Visualization App Overview

We developed a protein visualization app and introduced it within a high school biology class curriculum, consisting of a total of 53 students across two classes. The app enabled students to learn and engage with biomolecular interactions through three key concepts. First, students visualized DNA sequences to protein structure relationships, learning how linear sequences of nucleotides translate into functional, 3D protein structures used throughout the body. Secondly, students visualized the central dogma of molecular biology, describing the flow of genetic information: DNA is transcribed into RNA, and RNA is translated into proteins. Third, students explored DNA mutations and their effect on protein structure and function after changing nucleotide sequences, resulting in diversity, altered functions, and disease. Afterward, students rated their understanding of these three concepts on a scale from 1 to 5 both before and after using the app, with 1 indicating minimal understanding and 5 indicating a thorough understanding.

The app was developed using advanced biophysical tools to create detailed and interactive protein visualizations. Specifically, the app was developed using protein sequences obtained from the PDB, an extensive repository of 3D structural data for biological molecules. The app utilized Py3Dmol and ESMFold libraries for visualization and Streamlit, a framework to host our interactive app on a local website.

The app features two modes. First, Sandbox mode was designed for exploration and experimentation. Specifically, students can select proteins, visualize them in 3D, customize 3D protein structure settings, and insert mutations. Second, Puzzle mode offers a more targeted and guided learning experience. Specifically, students are tasked with applying their knowledge through skill-based activities, such as transcribing DNA to RNA, translating RNA into proteins, and identifying the different types of mutations. This mode is designed to challenge students to apply their knowledge in problem-solving scenarios. Together, these modes provide an interactive, educational, and engaging experience. Table 1 below summarizes the key concepts, app features, and student tasks in more depth, along with the associated biophysical learning outcomes.

Table 1: Overview of Key Concepts, App Features, and Biophysical Learning Outcomes. A summary of the relationship between biophysical ideas and the interactive features of the visualization app employed by students.

Concept	App Features/Student Tasks	Biophysical Learning Outcomes
DNA Sequence to Protein Structure	**Protein Selection and Visualization**: Visualizing a PDB protein's 3D structure using ESMFold. **Interactive Exploration**: Manipulating the model to explore folding patterns and structural features. **Mutation Exploration**: Alter the protein sequence with mutations and visualize the structural impact. **Hydrophobicity Analysis**: Analyze how mutations affect hydrophobic regions	Understanding molecular forces that guide protein folding (e.g. hydrogen bonds, hydrophobic interactions). Analyzing the impact of mutations on protein stability and function.
Central Dogma	**Central Dogma Simulation:** Transcribe DNA into RNA and translate RNA into proteins using both wild-type and mutated sequences. **Mutation Identification:** Identify the type and location of mutations and understand their impact on protein structure.	Ribosome mechanics in protein synthesis. Comprehending protein structure formation and the dynamic process of translation.
DNA Mutations	**Mutation Exploration:** Explore specific mutations' effects on protein structure using 3D visualization. Compare mutated protein structures with the wild-type to see the impact of DNA sequence changes.	Analyzing how mutations affect proteins' sequences, stability, and function, showing how these changes may lead to disease or altered function.

Sandbox Mode:

In Sandbox mode, students are first given an instructions page which outlines various features, as shown in Figure 1A. This page informs students how to change the protein, highlight attributes, visualize amino acid mutations, and generate a van der Waals surface of the protein. The instruction page also outlines the limitations of certain features, such as the van der Waals surface not being compatible with a sphere style. On the page, there are also two tabs named "General Settings" and "Visualization Settings." To select a protein, students click on the "General Settings" tab and can either click on one of the preset options or input a PDB ID to select any protein in the RCSB protein database. The selected protein's basic structure and characteristics are then visualized by ESMFold, as shown in Figure 1B. Students can interact with the protein model by using the mouse to see it in 3D. To visualize the amino acid sequence as a protein structure, our code retrieves data from the Py3Dmol library to match the given protein structure. The protein is then predicted and visualized in an overlay through the ESMFold algorithm.

After selecting a protein, students may switch to the "Visualization Settings" tab. Under the model type, students may choose to visualize the protein through either the Ball-and-Stick model, the Ribbon Diagram, or an overlay of both. Additionally, by checking a box named "Generate Surface," students can visualize the van der Waals surface on top of the model with an option to change opacity (Figure 1C). For the ribbon diagram, students may also generate a surface (Figure 1D). The ribbon may also be highlighted with different amino acid hydrophobicity values. For the model type selected as "Both" (Figure 1E), students may generate the surface and highlight hydrophobicity while adjusting for both ribbon and surface opacity. A hydrophobicity key is also provided. Specifically, the regions with a larger positive value are more hydrophobic and are depicted with a red color, while the less positive values are more hydrophilic and depicted with a yellower or greener color. The code assigns the scale to each amino acid, which the function separates by color. Furthermore, another important feature is the ability to visualize subchains. Under the visualization settings, students may select a subchain to visualize by highlighting specific amino acids. The website then visualizes the subchain instead of the full protein chain (Figure 1F). Finally, under General Settings, students may also visualize ML-predicted structures through ESMFold by altering the amino acid sequence to create a "mutation." Specifically, students may alter the sequence by either adding, deleting, or inserting an amino acid. After visualizing the mutated amino acid, students have the option to overlay the mutated protein with the original (Figure 1G).

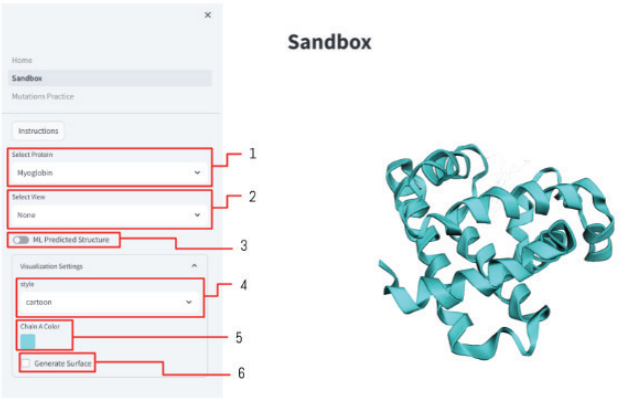


Figure 1A: An instruction guide for the available features in Sandbox mode. Users can select a protein (1), choose a structural view (2), toggle between ML-predicted or known structures (3), and customize visualization settings such as style (4), chain color (5), and surface rendering (6). On the right, the selected myoglobin protein is visualized in cartoon style with a cyan chain color. This exploration allowed students to interact with 3D protein models, reinforcing understanding and recognition of protein structure.

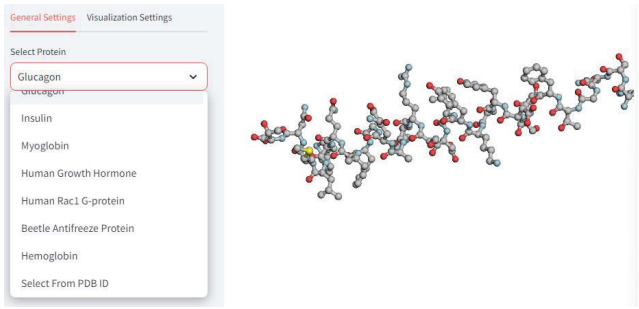


Figure 1B: In the Visualization Settings tab, students can select a protein to visualize from the dropdown menu. Students can choose among the present options or insert a protein via PDB ID. This freedom of selection enabled students to explore proteins of interest while enhancing exploration and engagement.

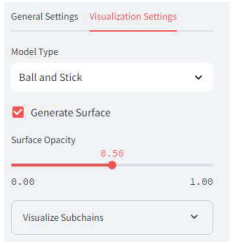


Figure 1C: Ball-and-Stick model of Glucagon protein with van der Waals surface generated. Surface Opacity is set to 0.50. This representation helped students comprehend important features of intermolecular forces and interactions.

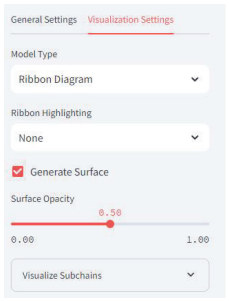


Figure 1D: Ribbon diagram of Glucagon protein with van der Waals surface and no hydrophobicity highlighting. Surface Opacity is set to 0.50. This visualization allowed students to understand secondary structure through a different model type.

Figure 1E: Both Ball-and-Stick and ribbon models overlaid with the hydrophobicity ribbon highlighting, and no van der Waals surface. Students utilized this feature to relate molecular structure to function by identifying hydrophobic regions.

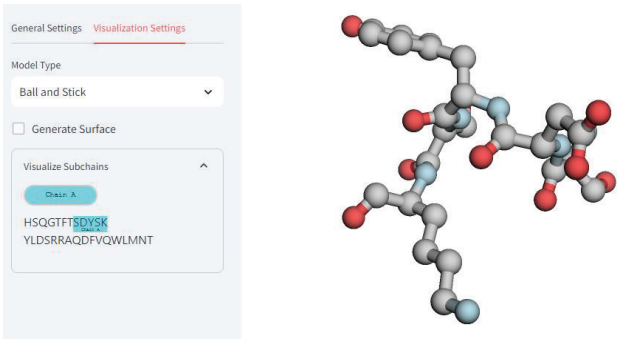


Figure 1F: A subchain of glucagon is visualized. No van der Waals surface is generated. This enabled students to zoom in on specific subchains within a protein for closer examination.

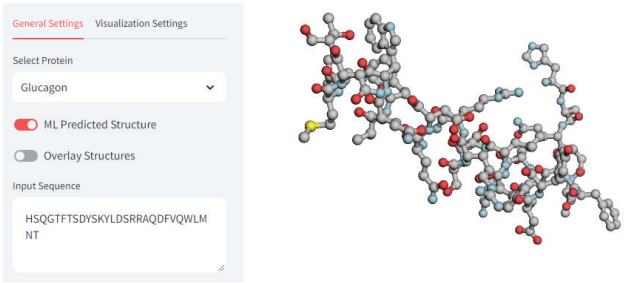


Figure 1G: Using ESMFold under the General Settings tab to predict amino acid changes to the protein. This feature allowed students to model mutations and observe their predicted effects on protein structure.

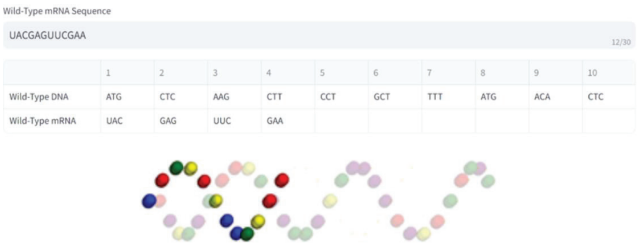
Puzzle Mode:

In Puzzle mode, students are first given a summary of the central dogma and mutations. Then, they are redirected to a page where they are tasked with transcribing a wild-type DNA sequence to RNA. They are also provided a visualization of the original DNA sequence and their input RNA sequence, shown in Figure 2, which is color-coded for different nucleotides to help them understand the concept of corresponding nucleotides from DNA to RNA. Specifically, blue represents uracil, red represents adenine, green represents cytosine, and yellow represents guanine. For reference, a codon chart is provided. As a result, students are able to "check" their input through the appropriate color coding, as seen by comparing Figure 2 and Figure 3. Next, students are given the same DNA sequence, but with mutations. First, in Figure 4, an insertion of the nucleotides "CTCGT" has occurred starting in codon 6. Students

DOI: 10.36838/IJHSR84.15

transcribe the DNA sequence to mRNA and are able to see the visualization of the helical mRNA structure with the appropriate colors for each nucleotide. Afterwards, students move on to the translation aspect of the central dogma, where they again translate both the wild-type and mutated mRNA sequence. While translating the mRNA sequence, students are given a visualization of the wild-type and mutated protein, as seen in Figure 5. Once both mRNA sequences are translated properly into amino acids, students are asked to identify the type of mutation (insertion, deletion, or substitution) and to highlight the location of the mutation (Figure 6). For substitution mutations, students are asked to first highlight the original DNA nucleotides that were mutated, then type the nucleotides that they were replaced with. In the example of Figure 6, students would type "C."

In Puzzle mode, the proteins whose nucleotide sequences were used for the puzzle mode were obtained from the RCSB protein database, based on a list of proteins with specific characteristics, such as an approximate sequence length. For the purposes of the puzzle mode code, a random protein is selected, and a "reading frame" of 10 codons, or 30 nucleotides, is extracted for the student to interact with. In this context, a reading frame refers to the way nucleotides in a DNA or RNA sequence are paired into sets of three, called codons, each of which codes for an amino acid. Afterwards, a random mutation is selected out of insertion, deletion, and missense substitution. The function for creating an insertion mutation is executed in a series of steps: a list of nucleotides is first created. Then, the indices of the beginning and end of the mutation are found; afterwards, the RNA string is altered to remove the nucleotides between the indices and replace them with the nucleotide string. Afterwards, the new RNA string is cut to keep the length as a multiple of 3 and returned. The deletion and missense substitution functions are similar, respectively deleting and replacing a section of the nucleotide string. In some cases of insertion or deletion where a frameshift mutation occurs, a nonsense mutation would usually occur in the protein, causing a premature stop codon to be read. However, the code does not include nonsense mutations, where a mutation would introduce an early stop codon, because the reading frame window of 30 nucleotides automatically selects a string of 30 nucleotides from the mutated RNA sequence, thus already artificially truncating the sequence and removing the need for a nonsense mutation function. This simplification is a limitation of the app's biological realism, but through this system, all puzzles are consistently set to the same sequence length, allowing students to focus on understanding how different types of mutations affect one amino acid sequence in different ways.



UACGAGUUCGAA

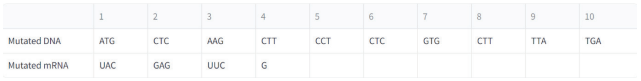
	1	2	3	4	5	6	7	8	9	10
Wild-Type DNA	ATG	CTC	AAG	CTT	CCT	GCT	TTT	ATG	ACA	CTC
Wild-Type mRNA	UAC	GAG	UUC	GAA						

Figure 2: A sample puzzle where the wild-type DNA sequence is given for students to transcribe to wild-type mRNA. Students are also able to visualize the helical DNA and mRNA structures as they input the mRNA sequence. In this example, the DNA has been successfully transcribed to mRNA, as seen through the correct corresponding colors.

CCCCCCCCCCCC

	1	2	3	4	5	6	7	8	9	10
Wild-Type DNA	ATG	CTC	AAG	CTT	CCT	GCT	TTT	ATG	ACA	CTC
Wild-Type mRNA	CCC	CCC	CCC	CCC						

Figure 3: A sample puzzle where the DNA has been incorrectly transcribed to mRNA, as seen through the incorrect colors in the mRNA sequence. Students were able to self-correct errors and improve the accuracy of transcription.



	1	2	3	4	5	6	7	8	9	10
Mutated DNA	ATG	CTC	AAG	CTT	CCT	CTC	GTG	CTT	TTA	TGA
Mutated mRNA	UAC	GAG	UUC	G						

Figure 4: After correctly transcribing the DNA sequence to mRNA, students are given the mutated version of the same DNA and are asked to transcribe it to mRNA. This exercise reinforced the concept of mutation and the central dogma principles.

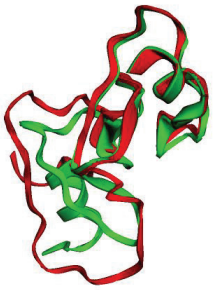


Figure 5: Both the wild-type and mutated protein structures are visualized while students translate both mRNA sequences to amino acid sequences. Students can zoom in and out and rotate the proteins. This gave students a 3D understanding of structural changes caused by mutations.

TGGGCGCGGCTTCTCCTCTCTTCCAGGCCC

Figure 6: Students are asked to identify and highlight all mutated nucleotides. This encouraged close examination of nucleotide-level differences and their relevance to protein structure.

Methods

The app was developed in Python 3.9 using Py3Dmol for 3D molecular visualization and Streamlit to host the app on a local website. Structure predictions for proteins were generated via ESMFold, and visual overlays of hydrophobicity utilized the Kyte-Doolittle scale alongside pLDDT confidence scoring, common frameworks for measuring protein features, including hydrophobicity. All source code was kept in a GitHub repository and a Google Colab Notebook, integrated automatically into Streamlit for testing. In addition, the app was divided into two distinct modules: (1) Sandbox mode, where students experiment freely with protein structures and their hydrophobicity ribbon structures selected from a preset list or by PDB ID; and (2) Puzzle mode, where students interact with transcription, translation, and mutation-identification exercises using color-coded DNA/RNA helices and protein visualizations.

53 high school AP or honors-level biology students from two classes (27 AP students, 26 Honors students) that underwent identical protocols participated in the study. Each student completed a baseline survey to self-assess understanding of three biophysical concepts on a 1-5 rating scale: DNA sequence to protein structure, the central dogma, and DNA mutations. 1 indicated minimal understanding, while 5 indicated thorough mastery. Following a 30-minute orientation to the app, students engaged in two 45-minute sandbox sessions exploring protein folding and mutation effects, then two 45-minute puzzle sessions focused on molecular biology tasks. A post-survey, identical to the baseline survey, captured quantitative changes or improvements through self-assessed understanding ratings. Students' qualitative feedback at the end of the study was also collected.

Students' quantitative feedback was also statistically analyzed to assess whether the observed improvements were significant or not. First, as students rated their understanding on a 1–5 scale, we used a nonparametric test, the Wilcoxon signed-rank test, a statistical hypothesis test that does not assume the data follows a specific probability distribution or normality, and is appropriate for comparing two related samples. In this case, the test measures the differences between each student's pre- and post-survey ratings, ranks the absolute differences, and determines whether there is a consistent increase or decrease in scores across participants. We also statistically evaluated the data using a paired t-test, a parametric test that assumes normally distributed differences and compares the mean pre- and post-scores. To verify this assumption, we conducted a Shapiro-Wilk test on the paired differences, which indicated non-normality ($p < 0.001$), justifying the use of the Wilcoxon test as the primary analysis. Then, the effect sizes were computed using Cohen's d and Hedges' g for the t-tests and $r = |Z|/\sqrt{N}$ for the Wilcoxon tests to assess the change of magnitude. Confidence intervals (CI) of 95% were calculated to estimate the range within which the true mean difference likely falls. Statistical analyses were performed in Python 3.9 using standard scientific libraries.

Results and Discussion

Protein Game Learning Outcome Results:

Student survey results show the potential for visual aids in complex topics to significantly improve students' learning, comprehension, and preparation for advanced fields like bioengineering and biophysics. According to survey bar graphs, the results support that the app was highly effective in increasing students' understanding to higher levels (4 and 5). For example, the number of students who rated their understanding of the DNA sequence to protein structure relationship as "4" or "5" increased substantially, from 16 students before using the app to 30 students afterward. Similarly, for the central dogma, while the number of students who rated their understanding a "4" stayed the same, the number of students who rated their understanding a "5" increased from 2 to 8. Finally, for DNA mutations, the shift was less dramatic for the "5" level ratings, yet the number of students who rated their understanding a "4" increased from 17 to 25. On the other hand, the number of students who rated their understanding at lower levels (1 to 3) decreased on all levels across the three different topics, indicating significant improvement in understanding after the study.

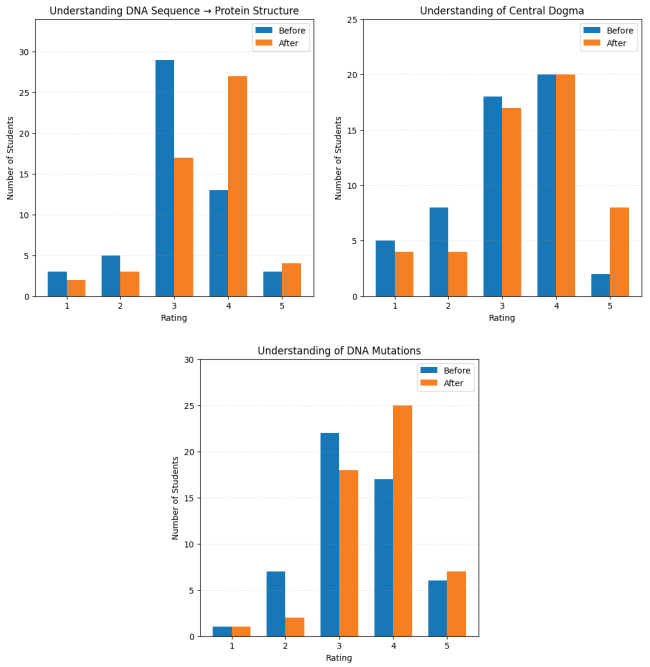


Figure 7: Impact of a Protein Visualization App on Students' Understanding of Key Biophysical Concepts. The bar graphs compare students' self-assessed levels of understanding before and after using the app, across three biophysical concepts. Students reported improvements in understanding protein structure, mutation impact, and sequence-structure relationships.

The results demonstrate varying levels of progress in conceptual understanding among students. The most significant finding was that the DNA sequence to protein structure and central dogma concepts showed improvement with 47% and 43% of students, respectively, demonstrating a one-point increase in their understanding. The app proved useful for creating modest 1-point improvements, but larger increases were less common. It is also important to note that students who began at Level 3 achieved the greatest improvement in their understanding. The DNA sequence to protein struc-

DOI: 10.36838/IJHSR84.15

ture relationship received substantial improvement from 32% of students who started at Level 3. The DNA Mutations understanding improved significantly for students who began at Levels 2 and 3, since 15% of students in each group showed progress. However, the app failed to produce meaningful improvements in DNA Mutations understanding for students who entered with minimal prior knowledge at Level 1. Specifically, the lowest level of understanding had little to no change for all three topics after the study. Students who began at Level 4 (15%) and Level 2 (13%) demonstrated similar improvements in their comprehension of the Central Dogma. In essence, while the app's potential was promising in elevating students' baseline understanding, it may not have benefited students equally due to different learning patterns and prior knowledge going into the study. Moreover, three students reported a decrease in understanding, primarily due to technical errors within the tool reported via qualitative feedback. Most of the student responses expressed positive and satisfactory feedback regarding the app, but a few students encountered bugs that prevented them from using all the features of the app. These technical issues suggest that the tool's overall effectiveness in enhancing understanding might have some challenges, and further refinement of the app can improve its reliability and effectiveness for all users.

Furthermore, spaghetti plot graphs of students showed improvement in scoring after the study, supporting that the app can improve overall student understanding across all three biophysical concepts (Figure 8). The "Understanding DNA Sequences" graph shows the largest average gain (Δ = 0.38) because students who began at level 3 achieved levels 4 or 5, aligning with the previously established relationship that the app proved most beneficial to students who already had previous knowledge of DNA sequences. Similarly, the "Understanding of Central Dogma" graph demonstrated a comparable upward slope (Δ = 0.34). The number of students who reached level 5 understanding increased from 2 to 8 students after using the app, which proves its ability to explain transcription-translation processes through interactive visualizations. The improvements in this section were more evenly spread across different starting levels, which indicates that the students successfully understood the Central Dogma process using the app. The "Understanding DNA Mutations" graph showed minimal improvement (Δ = 0.28), which mainly benefited students who started at levels 2 or 3. The spaghetti graphs also show that students who began at level 1 demonstrated minimal progress, which indicates the app may not work well for beginners or that this module requires additional development. Therefore, additional work, such as clearer instructions and simpler molecular visualization, is necessary to enhance support for struggling students and enhance understanding.

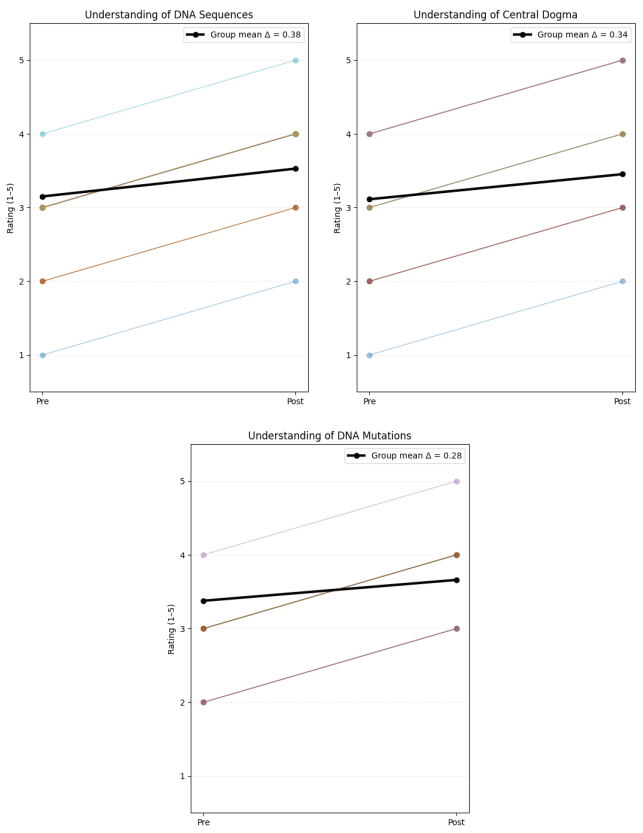


Figure 8: Spaghetti Plot of Student Gains in Understanding of Biophysical Concepts After Using Protein Visualization App. Colored lines represent an individual student (overlapping lines present), and bold lines represent mean improvement. The data shows improvements across all three concepts, highlighting the app's effectiveness in enhancing comprehension.

We statistically assessed three concepts: DNA sequence to protein structure, the central dogma, and DNA mutations. Students showed statistically significant improvements in self-assessed understanding after using the app. Specifically, for DNA sequence to protein structure, the Wilcoxon signed-rank test showed a significant improvement in ratings ($W = 56.0$, $p = 0.0002$, effect size $r \approx 0.63$). Mean understanding scores rose from 3.15 ± 0.88 to 3.53 ± 0.86 ($\Delta = +0.38$, 95% CI = [0.17, 0.58]). For the central dogma, Wilcoxon results showed significant improvement in ratings ($W = 51.0$, $p = 0.0006$, effect size $r \approx 0.62$). Mean understanding scores rose from 3.11 ± 1.02 to 3.45 ± 1.07 ($\Delta = +0.34$, 95% CI = [0.10, 0.58]). For DNA mutations, Wilcoxon results showed significant improvement in ratings ($W = 36.5$, $p = 0.007$, effect size $r \approx 0.57$). Mean understanding scores increased from 3.38 ± 0.92 to 3.66 ± 0.82 ($\Delta = +0.28$, 95% CI = [0.09, 0.48]).

Paired t-tests resulted in all three topics having consistent results (all $p < 0.01$). Effect sizes (Cohen's $d = 0.39–0.50$) indicate moderate improvements in students' self-assessed understanding. After Bonferroni correction for multiple comparisons ($\alpha = 0.017$), all results remained significant. These statistical results demonstrate that students' conceptual understanding significantly improved after learning biophysical concepts using the protein visualization app, confirming a strong positive relationship between the app's utilization for learning and students' comprehension of three complex biophysical concepts: DNA

sequence to protein structure, the central dogma, and DNA mutations.

Discussion:

Our study adds a critical layer to the field of biophysics education for younger students passionate about science and technology. Unlike complicated protein visualization tools like PyMOL, UCSF Chimera, and VMD, our app makes protein visualization simple, accessible, and flexible for students, creating an optimal environment for obtaining new knowledge through 3D interaction and visualization while improving comprehension through check-for-understanding exercises. Our study demonstrates quantitative and qualitative evidence that supports the integration of interactive 3D molecular visualization into high school biology education to boost understanding of core biophysical concepts. Utilizing the app, students explored DNA sequences, protein structures, molecular processes such as transcription and translation, and mutations. Across the Sandbox, Puzzle, and Contest instructional modes, the app provided real-time visual feedback, enabling students to manipulate biological models and engage with abstract content more concretely and intuitively.

To quantify, the pre- and post-intervention survey data from 53 AP and honors biology students demonstrated consistent learning progression. Students who initially demonstrated a moderate understanding level (self-rated level 3) achieved the most significant progress. The number of students who assessed their knowledge of DNA sequence-to-structure relationships at level 4 or 5 increased from 16 to 30. The number of students who assessed their understanding of the central dogma at level 5 increased from 2 to 8 following the intervention. The DNA mutations section showed limited improvement in student understanding, but students who started at levels 2 or 3 demonstrated progress to higher ratings, which indicates the app's effectiveness for students with some basic knowledge. That said, our data also reveals limitations. Students who started at the lowest level of understanding (level 1) experienced minimal progress, especially for mutations, and only a few students improved by more than one point. It is also important to note that 3 participants reported a decrease in the level of understanding because of technical issues highlighted in the qualitative feedback section of the survey. The findings helped us conclude that the app successfully enriches knowledge and comprehension, but its current design struggles to adapt to fit the needs of students with limited prior knowledge.

Future directions for the protein visualization app could focus on enhancing its user interface and experience by correcting technical issues and expanding customizability for the user. Integrating the app into educational curricula in schools can further enhance the learning experience for students and improve the generalizability of the app. Additional features could include a wider variety of biological processes and disease models, certain structures' evolutionary biology, settings that include language options, side-by-side comparison for viewing multiple proteins at once, and settings to directly modify 3D protein structures. Additionally, by adding collaborative and competitive learning features, such as group puzzles or leaderboard competitions, students can get an added incentive for learning while navigating a more interactive, adaptable learning environment. Incorporating automated assessment tools to analyze data would help track student progress over time and provide long-term feedback, making the app a more reflective educational resource. Finally, even integrating emerging technologies like virtual and augmented reality, as well as artificial intelligence virtual assistants, boosts the immersive and engaging feeling while learning, helping students obtain a passion for the complexities of modern scientific careers.

■ Conclusion

Traditional biology education teaches students about DNA structures, the central dogma of molecular biology, and mutations without sufficient integration of visual experimentation and interactive elements. The educational protein visualization app we created transforms abstract, linear biological processes into interactive visual experiences that help students understand biological functions down to a molecular level. Students can now use dynamic simulations to explore biological materials at the 3-dimensional molecular level through interactive protein model manipulation and challenging yet engaging puzzles to test their knowledge. Quantitative data from student surveys exemplify significant improvements in understanding across three biophysical concepts after using the protein visualization app, particularly for those with a moderate foundation of knowledge. More importantly, such simulations have a greater impact. Interactive visuals and learning experiences can inspire students, the next generation of bioengineers and biophysicists, to seek new questions, think outside the box, and bridge these scientific concepts to the real world through innovations and advancements. With more enhancements, the app can become a fundamental resource for teachers and students around the globe, passionate about biophysics, bioengineering, or related fields, making it an indispensable resource in scientific education.

■ Acknowledgments

We would like to formally acknowledge Ms. Sarah Baysden of Bellarmine College Preparatory for her dedicated mentorship and guidance throughout the entirety of this process. We would also like to thank Professor Prashanth Asuri of Santa Clara University for his assistance in addressing technical inquiries related to the website and processing.

■ References

1. Schönborn, K. J.; Anderson, T. R. The Importance of Visual Literacy in the Education of Biochemists. *Biochemistry and Molecular Biology Education* 2006, *34* (2), 94–102. https://doi.org/10.1002/bmb.2006.49403402094.
2. Dries, D. R.; Dean, D. M.; Listenberger, L. L.; Novak, W. R. P.; Franzen, M. A.; Craig, P. A. An Expanded Framework for Biomolecular Visualization in the Classroom: Learning Goals and Competencies. *Biochemistry and Molecular Biology Education* 2016, *45* (1), 69–75. https://doi.org/10.1002/bmb.20991.
3. Burgin, S. R.; Oramous, J.; Kaminski, M.; Stocker, L.; Moradi, M. High School Biology Students' Use of Visual Molecular Dynamics

DOI: 10.36838/IJHSR84.15

as an Authentic Tool for Learning about Modeling as a Professional Scientific Practice. *Biochemistry and Molecular Biology Education* 2018, *46* (3), 230–236. https://doi.org/10.1002/bmb.21113.
4. Craig, P. A.; Michel, L. V.; Bateman, R. C. A Survey of Educational Uses of Molecular Visualization Freeware. *Biochemistry and Molecular Biology Education* 2013, *41* (3), 193–205. https://doi.org/10.1002/bmb.20693.
5. Lundquist, K.; Herndon, C.; Harty, T. H.; Gumbart, J. C. Accelerating the Use of Molecular Modeling in the High School Classroom with VMD Lite. *Biochemistry and Molecular Biology Education* 2016, *44* (2), 124–129. https://doi.org/10.1002/bmb.20940.
6. Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Žídek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; Velankar, S.; Kleywegt, G. J.; Bateman, A.; Evans, R.; Pritzel, A.; Figurnov, M.; Ronneberger, O.; Bates, R.; Kohl, S. A. A.; Potapenko, A. Highly Accurate Protein Structure Prediction for the Human Proteome. *Nature* 2021, *596* (596), 1–9. https://doi.org/10.1038/s41586-021-03828-1.
7. Maria, M.; Guillaume, V.; Guionnière, S.; Dacquay, N.; Cyprien Plateau-Holleville; Larroque, V.; Lardé, J.; Naimi, Y.; Jean-Philip Piquemal; Guillaume Levieux; Lagarde, N.; Stéphane Mérillou; Montes, M. Interactive Visualization of Large Molecular Systems with VTX: Example with a Minimal Whole-Cell Model. *Frontiers in Bioinformatics* 2025, *5*. https://doi.org/10.3389/fbinf.2025.1588661.
8. Terrell, C. R.; Listenberger, L. L. Using Molecular Visualization to Explore Protein Structure and Function and Enhance Student Facility with Computational Tools. *Biochemistry and Molecular Biology Education* 2017, *45* (4), 318–328. https://doi.org/10.1002/bmb.21040.
9. Jenkinson, J. Molecular Biology Meets the Learning Sciences: Visualizations in Education and Outreach. *Journal of Molecular Biology* 2018, *430* (21), 4013–4027. https://doi.org/10.1016/j.jmb.2018.08.020.
10. Canning, D. R.; Cox, J. R. Teaching the Structural Nature of Biological Molecules: Molecular Visualization in the Classroom and in the Hands of Students. *Chemistry Education. Research and Practice* 2001, *2* (2), 109–122. https://doi.org/10.1039/b1rp90013g.
11. Pine, P.; Paina, L. I. Computational Methods in Chemistry and Biochemistry Education: Visualization of Proteins. *Computing in Science & Engineering* 2020, *22* (4), 45–49. https://doi.org/10.1109/mcse.2019.2962118.
12. 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. Highly Accurate Protein Structure Prediction with Alphafold. *Nature* 2021, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.
13. Mihály Váradi; Bertoni, D.; Magaña, P.; Urmila Paramval; Ivanna Pidruchna; Radhakrishnan, M.; Maxim Tsenkov; Nair, S.; Milot Mirdita; Yeo, J.; Oleg Kovalevskiy; Tunyasuvunakool, K.; Laydon, A.; Augustin Žídek; Tomlinson, H.; Hariharan, D.; Abrahamson, J.; Green, T.; Jumper, J.; Birney, E. AlphaFold Protein Structure Database in 2024: Providing Structure Coverage for over 214 Million Protein Sequences. *Nucleic Acids Research* 2023, *52* (D1). https://doi.org/10.1093/nar/gkad1011.
14. Rubach, P.; Sikora, M.; Aleksandra I Jarmolinska; Agata P Perlinska; Joanna I Sulkowska. AlphaKnot 2.0: A Web Server for the Visualization of Proteins' Knotting and a Database of Knotted AlphaFold-Predicted Models. *Nucleic Acids Research* 2024, *52* (W1), W187–W193. https://doi.org/10.1093/nar/gkae443.
15. Berry, C.; Baker, M. D. Inside Protein Structures: Teaching in Three Dimensions. *Biochemistry and Molecular Biology Education* 2010, *38* (6), 425–429. https://doi.org/10.1002/bmb.20434.
16. Hanson, R. M.; Prilusky, J.; Renjian, Z.; Nakane, T.; Sussman, J. L. JSmol and the Next-Generation Web-Based Representation of 3D Molecular Structure as Applied to Proteopedia. *Israel Journal of Chemistry* 2013, *53* (3-4), 207–216. https://doi.org/10.1002/ijch.201300024.
17. Rigsby, R. E.; Parker, A. B. Using the PyMOL Application to Reinforce Visual Understanding of Protein Structure. *Biochemistry and Molecular Biology Education* 2016, *44* (5), 433–437. https://doi.org/10.1002/bmb.20966.

Authors

Alan Kim is a senior at Bellarmine College Preparatory with a profound interest in biotechnology and neuroscience, particularly in the study of neurodegenerative diseases. He is committed to advancing his understanding through scientific research and aspires to pursue a career in neurosurgery, where he hopes to make meaningful contributions to the field.

Avi Lekkelapudi is a freshman at the University of Illinois Urbana-Champaign majoring in Chemical Engineering. He is passionate about exploring chemical processes and biotechnology to develop transformative innovations.