TIPS, TRICKS & TECHNIQUES

Exploring Protein Structure with RasMol, a Free Online Tool

RECOMMENDED FOR *AP Biology*

Amber D. Fraley, Katherine E. Odegaard, Victor H. Vilchiz, ChulHee Kang, Cathy Lee

Abstract

Understanding the structural and functional relationships of proteins is an important concept of biology. Therefore, we had Gordon State College students in the biochemistry class of spring 2019 apply and evaluate a hands-on bioinformatics activity, using the free bioinformatics program, RasMol. This module was designed to help students understand the relationships of a protein to a specific disease, and in this paper, we have chosen to discuss apolipoprotein E (APOE) relationships to Alzheimer's disease since some of the students had an interest in neurobiochemistry.

During the learning module, students were asked to identify structural differences among the APOE isoforms. Overall, 20 out of 27 students (74.1%) evaluated the instructional value of the RasMol program positively after being surveyed. Student feedback also suggested they learned more about the molecular structure of APOE using RasMol. Therefore, utilization of free computer bioinformatics activities encourages students to apply logical, real-world solutions to biological problems.

Key Words: *bioinformatics; RasMol; apolipoprotein E; APOE; Alzheimer's disease; survey; introductory biochemistry.*

○ Introduction

The term *bioinformatics* can be formally defined as "the use of computers in collecting, storing, accessing, and analyzing biological data such as molecular sequences and structures" (Pratt & Cornely, 2015). Although bioinformatics is a field that has been around for many years, there is room for exploring the use of bioinformatics programs as a teaching tool in undergraduate biochemistry classrooms. Hands-on bioinformatics program activities can aid in teaching challenging biochemistry concepts. One of the many free bioinformatics programs available to the public is RasMol. In this teaching exercise, students applied RasMol to apolipoprotein E (APOE) isoforms to further their knowledge on the structural and functional relationships of APOE in relation to Alzheimer's disease (AD). APOE is a common protein discussed in biochemistry classrooms and may be linked to the brain response in the neurodegeneration of mammals (Poirier, 2000).

Using free bioinformatics modules in an introductory biochemistry class can help students better grasp the relationship between structure and function of critically important proteins.

It has been shown that different innate APOE isoforms can influence the time of onset and clearance of amyloid deposits, which are the neurological tangles that cause plaques in the brain and are associated with AD (Giau et al., 2015). Recent information in the field of AD research indicates that tau oligomers, or tau proteins, interact with APOE and may also play a key role in AD (Mroczko et al., 2019). APOE comes in three isoforms labeled APOE2, APOE3, and APOE4 (Zhao et al., 2018). In humans, these APOE isoforms have different interactions with the tau protein (Zhao et al., 2018). Further observations have been made that the tau protein can bind to APOE3 but does not bind to APOE4, and it is still unclear if it forms a complex with APOE2 (Zhao et al., 2018). In addition to tau protein interactions, evidence has shown that individuals with APOE4 have an increased risk of developing AD, while those with APOE2 have a decreased risk of developing AD (Leduc et al., 2011).

Therefore, students in biochemistry classes can benefit from viewing the different APOE isoform structures and understanding their possible role in the pathophysiology of AD through this module, which may aid them in a future health sciences career (Dodgen et al., 2017). In addition to biochemistry classes, biology classes typically incorporate structural molecular biology sections and emphasize the development of visualizing structures by starting with Watson and Crick's model of DNA (Davenport et al., 2017). Utilizing the wide variety of free bioinformatics tools available online encourages undergraduate students to apply logical, real-world solutions to biological problems that may help them prepare for their desired career field. In the context of this exercise, students can become familiar with the structure and function of APOE through bioinformatics programs (Luscombe et al., 2001). In a simple and fun way, this project was designed to support

The American Biology Teacher, Vol. 84, No. 6, pp. 369–373, ISSN 0002-7685, electronic ISSN 1938-4211. DOI: https://doi.org/10.1525/abt.2022.84.6.369.

Table 1. All websites and software used for the exercises.

Website/ Software	Web address
RasMol	http://www.rasmol.org/
National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov

the idea that using free bioinformatics modules in an introductory biochemistry class can help students better grasp the relationship between structure and function of critically important proteins, such as the APOE protein associated with AD.

Goals

The bioinformatics modules and corresponding exercises were created with three purposes:

- to have students with minimal bioinformatics experience successfully complete an exercise in RasMol
- for students to successfully apply RasMol to APOE isoforms
- to have students feel that they have enhanced their comprehension of how the APOE isoforms' structural and functional differences relate to AD after using RasMol

○ Materials & Methods

The RasMol exercise and survey were given to students to complete on their own, outside of class. Allowing students to conduct the exercise on their own would help to support the idea that bioinformatics exercises, with the right instructions, can be completed by anyone and therefore used in any introductory biochemistry class. In addition to discussions in class, the students were also briefed with information on APOE and the potential connection between the APOE isoforms and AD at the beginning of the exercise packet.

○ Part I: Visualizing the APOE Isoforms Using RasMol

The first step for students was to download the RasMol program to their computer. After downloading RasMol, the students were asked to go to the NCBI website to retrieve the Protein Data Bank identifier (PDB ID) for each of the APOE isoforms, one at a time, starting with the APOE2 isoform. The instructions below apply to the most current version upon publication of this article.

RasMol Instructions

1. Download the latest version of RasMol (http://www.rasmol.org). After the download is complete, RasMol may appear as "**RasWin**" on the computer's home screen. If RasMol does not open automatically, click on "**RasWin**."
2. Go to the National Library of Medicine (https://www.ncbi.nlm.nih.gov).
3. Type in the PDB ID "**1NFO**" into the search bar. Click on the name of the isoform. For "**1NFO**" the name should read "**Apolipoprotein E2 (Apoe2, D154a Mutation)**."
4. On the right side of the screen, there will be a blue box that says "**Download**" (Figure 1). Select "**PDB Format**" and click the "**Download**" box.
5. With RasMol already open, the downloaded file may automatically open with RasMol, and a spiral structure will be shown with a black background. If this occurs skip to step 6. If the structure does not automatically sync with RasMol, open and save the downloaded PDB file as "**1NFO**" and continue on to step 5.
6. Open RasMol and select "**File**" and then "**open**." Find and open "**1NFO**." Two windows will open, one with the structure display and one with the command line.
7. On the structure display window, under the "**colors**" tab select "**group**."
8. Change the display to "**Wireframe**" so that there is a thin wire shown as the structure. To do this, click "**Display**" and select the first option, "**Wireframe**." These wires signify bonds between atoms, and the different ends of the wires represent individual atoms.

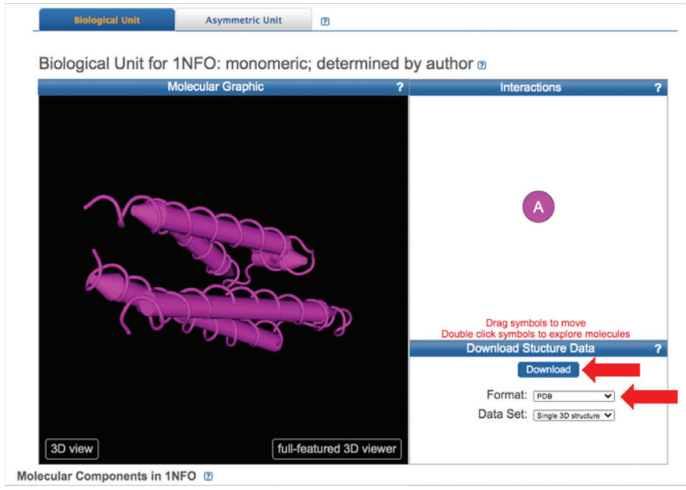

Figure 1. An image of what appears for the APOE2 PDB ID on the NCBI website.

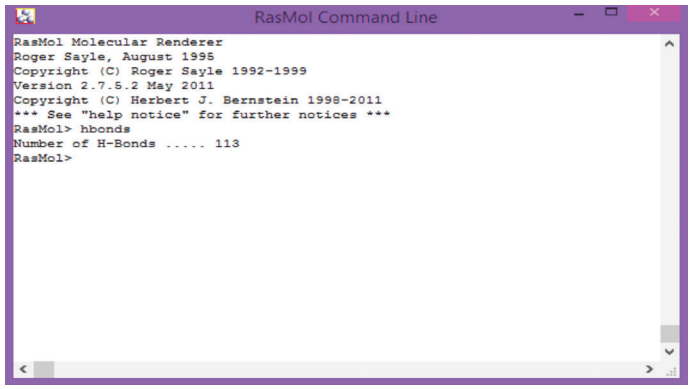

Figure 2. This is the RasMol command line and what happens when the command "**hbonds**" is typed into the command line.

9. Now, in the command line window, add hydrogen bonds by typing the command "**hbonds**" (Figure 2). A small, dotted line representing hydrogen bonds will appear, but it will be difficult to see. Note that the command line tells you how many hydrogen bonds are present.
10. To make the hydrogen bonds easier to see, type the command "**hbonds 200**" and they should be.

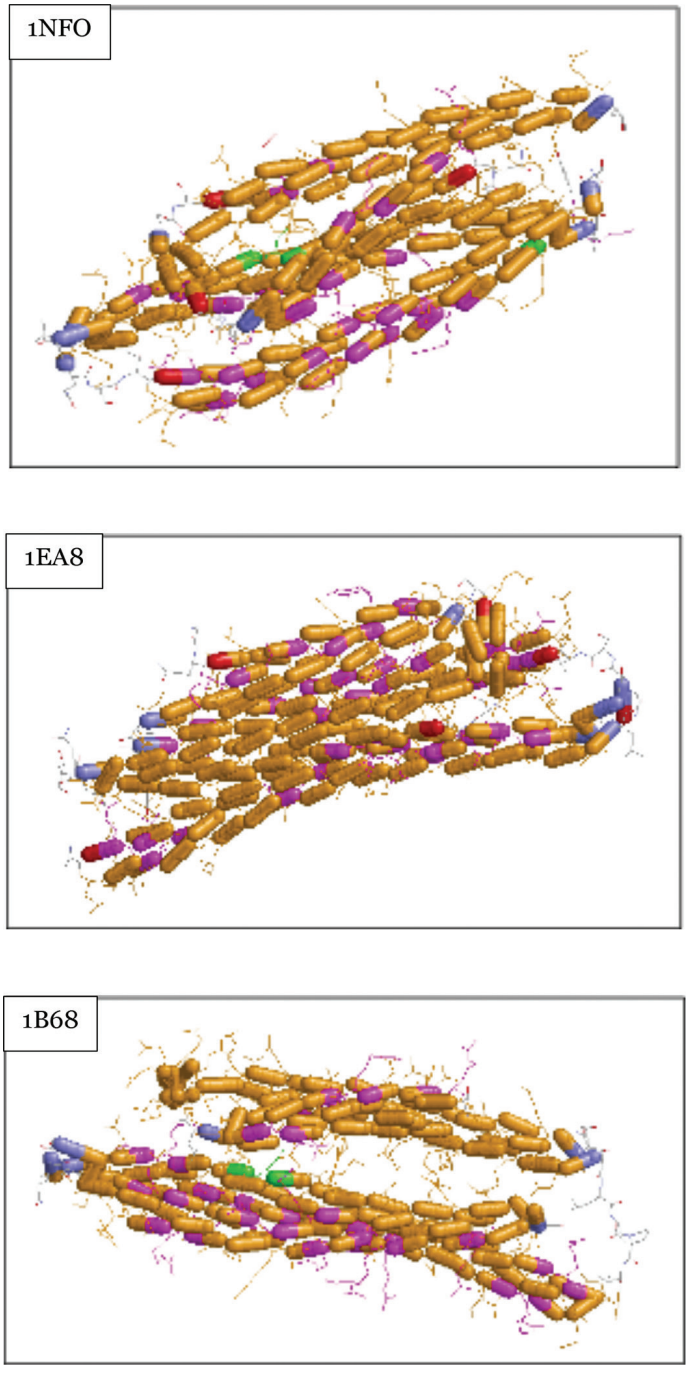

Figure 3. Completion of the RasMol module should yield the three APOE isoforms labeled with their appropriate PBD ID. In addition to the amino acids given designated colors (arginine and cysteine) and the helix residues, through the activity, RasMol by default assigns the color blue to the N terminus and the color red to the C terminus of proteins.

11. To highlight structural differences, we will highlight the helix structures. In the command line, type "**select helix**" and note how many atoms making up the helix are selected. Type "**color orange**" to color the helixes orange.
12. Highlight the number of arginine and cysteine residues found in the isoform. In the command line, type "**select arg**" and note that the command line tells you how many arginine residues are present. Type "**color magenta**" to color the arginine residues magenta.
13. Repeat step 11 by typing in the command line "**select cys**" and note that the command line tells you how many cysteine residues are present. Type "**color green**" to color the cysteine residues green.
14. Now, color the background white for clarity. In the command line, type "**background white**" to color the background white.
15. Return back to the structure display window. Then click on the windows icon and find "**Snip & Sketch**" by scrolling down to all applications that begin with the letter *S*. Click "**Snip & Sketch**" to open the program.
16. Put your structure display window side by side with the Snip & Sketch window and click "**New**" in the top left corner of Snip & Sketch. You will now be able to take a screenshot of the image of the protein structure by clicking and dragging a square around the structure. After surrounding the structure, release the mouse. Your screenshot will then be displayed on Snip & Sketch.
17. On Snip & Sketch, select the square floppy-disk icon in the top right corner to save your final image as "**1NFO LASTNAME**".
18. To ensure the correct steps were followed, repeat steps 14–16, but now saving an image of the command line window. Save this file as "**LASTNAME COMMANDLINE**." Upon the instructor's request, the images may be turned in digitally or by printing the images.
19. After saving both the image of the structure display window and the command line window, hit the "**X**" in the top right corner of each window to close the program.
20. Repeat steps 2–18 using PDB IDs "**1EA8**" for APOE4 until you have all three isoforms as shown in Figure 3.

○ Part II: Post RasMol Exercise Activity

Before completing the postactivity survey for the RasMol exercise, students were instructed to fill in Table 2 with information from the RasMol command dialog box and asked the following questions to help demonstrate if the students were benefiting from the RasMol exercise.

1. What was/were the most noticeable difference(s) among the three isoforms of apolipoprotein E?
2. Scientists have shown that one key difference between APOE3 and APOE4 is the substitution of a cysteine in place of an arginine in one position (Frieden & Garai, 2012). Do the data in Table 2 seem to be in agreement with that finding?
3. How could individuals benefit by identifying potential genetic factors for AD?

Table 2.

Apolipoprotein E Isoform	PDB ID	# Atoms in Helix	# Arginine	# Cysteine
APOE2	1NFO			
APOE3	1EA8			
APOE4	1B68			

○ Results

The students performed well on the postexercise activity by being able to fill in Table 2 using RasMol and to interpret the data by answering the questions with detailed responses. All the student's activity grades resulted in either an A or B. The grades were constituted by correctness and completeness but did not contribute to their class averages due to the activity being purely explorational. The post RasMol survey had questions that gauged understanding of APOE prior to the module, asked students about their bioinformatics experience, and asked the students if they enjoyed RasMol and would be willing to participate in more bioinformatics exercises. Some survey questions utilized the Likert scale ranging from 1 to 5, with 1 being "strongly disagree" and 5 being "strongly agree." Likert scale questions were then calculated into a mean score. The survey also included yes or no questions.

First, when students were asked if they understood the role of APOE in AD before encountering RasMol, only 2 out of 20 students were able to definitively answer "yes." Additionally, when students were asked if they had used RasMol before, 18 out of 20 students answered "no." Next, students were asked if they thought that RasMol was helpful in explaining the role of the APOE isoforms with their corresponding structures, to which students had an average response of 4.3 out of 5.0. Lastly, students were asked if they enjoyed using bioinformatics and if they would be willing to participate in more of these bioinformatics exercises in future biology courses, to which students had a mean of 4.4 out of 5.0 yes answers.

○ Final Thoughts

In conclusion, the results of this teaching module indicated that the goals of this project were met. According to the student surveys, students with minimal bioinformatics experience enjoyed successfully applying RasMol to APOE isoforms and felt that their comprehension of APOE isoform structural and functional differences, in relation to AD, increased after using RasMol. To ensure that the module was completely an immersive learning experience, the module was incentivized by awarding students with extra credit for a successful completion. Although RasMol was used to view APOE, RasMol can be used to show any protein structure. Therefore, the next step would be to expose students to even more proteins using RasMol as their guide. Although this project was tested in a small course, any teacher should be able to incorporate this activity into their lesson plans. In addition to helping students understand the importance of structural and functional relationships of proteins, this module can open students' eyes to a powerful tool in the bioinformatics field and increase their confidence in using this free online tool.

Table 3. RasMol module survey questions and student responses.

Statement	Response
Did you understand the role of APOE in AD before using this program?	18 answered no 2 answered yes
Have you ever used this program (RasMol) before?	18 answered no 2 did not answer
Using a scale of 1 to 5, were the instructions easy to follow? (1 = not easy, 5 = very easy)	Mean score: 4.6 (20 students total)
Using a scale of 1 to 5, how much better is your understanding of analyzing protein structure after using this program? (1 = no better, 5 = much better)	Mean score: 3.9 (20 students total)
Using a scale of 1 to 5, do you feel that this program did a good job explaining the role of APOE and its protein structure? (1 = not good, 5 = very good)	Mean score: 4.3 (20 students total)
Using a scale of 1 to 5, would you be interested in doing more of these exercises in future biology courses? (1 = not interested, 5 = very interested)	Mean score: 4.4 (20 students total)

○ Acknowledgments

We thank Jonathan Hughes of Gordon State College for helpful revision input and all the students enrolled in the biochemistry course of spring 2019 at Gordon State College. This work was presented at the 2019 Teaching Matters Conference at Gordon State College.

○ References

Davenport, J., Pique, M., Getzoff, E., Huntoon, J., Gardner, A. & Olson, A. (2017). A self-assisting protein folding model for teaching structural molecular biology. *Structure*, *25*(4), 671–78. https://doi.org/10.1016/j.str.2017.03.001.

Dodgen, C., Uwerosuo, U., Kang, C. & Lee, C. (2017). Exploring the use of free bioinformatics modules in an introductory biochemistry course. *Georgia Journal of Science*, *75*(2).

Frieden, C. & Garai, K. (2012). Structural differences between apoE3 and apoE4 may be useful in developing therapeutic agents for Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(23), 8913–18. https://doi.org/10.1073/pnas.1207022109.

Giau, V.V., Bagyinszky, E., An, S.S. & Kim, S.Y. (2015). Role of apolipoprotein E in neurodegenerative diseases. *Neuropsychiatric disease and treatment*, *11*, 1723–37. https://doi.org/10.2147/NDT.S84266.

Hagen, J.B. (2000). The origins of bioinformatics. *Nature Review Genetics*, *1*, 231–36. https://doi.org/10.1038/35042090.

Leduc, V., Domenger, D., De Beaumont, L., Lalonde, D., Bélanger-Jasmin, S. & Poirier, J. (2011). Function and comorbidities of apolipoprotein E in Alzheimer's disease. *International Journal of Alzheimer's Disease*, *2011*, 1–22. https://doi.org/10.4061/2011/974361.

Luscombe, N.M., Greenbaum, D. & Gerstein M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods of Information in Medicine*, *40*, 346–58. https://doi.org/10.1055/s-0038-1634431.

Martz, E. (1998). What can you learn about a protein/DNA molecule with RasMol?. https://www.umass.edu/microbio/rasmol/raswhat.htm.

Mroczko, B., Groblewska, M. & Litman-Zawadzka, A. (2019). The role of protein misfolding and tau oligomers (TauOs) in Alzheimer's disease (AD). *International Journal of Molecular Sciences*, *20*(19). https://doi.org/10.3390/ijms20194661.

Poirier, J. (2000). Apolipoprotein E and Alzheimer's disease a role in amyloid catabolism. *Annals of the New York Academy of Sciences*, *924*, 81–90. https://doi.org/10.1111/j.1749-6632.2000.tb05564.x.

Pratt, C.W. & Cornely, K. (2015). Essential biochemistry. 3rd ed. John Wiley & Sons.

Sayle R.A. & Milner-White, E.J. (1995). RASMOL: Biomolecular graphics for all. *Trends in Biochemical Sciences*, *20*(9), 374. https://doi.org/10.1016/s0968-0004(00)89080-5.

Zhao, N., Liu, C.C., Van Ingelgom, A.J., Linares, C., Kurti, A., et al. (2018). APOE ε2 is associated with increased tau pathology in primary tauopathy. *Nature Communications*, *9*(1), 4388. https://doi.org/10.1038/s41467-018-06783-0.

AMBER D. FRALEY (adf08123@uga.edu) is a former biochemistry student at Gordon State College, Barnesville, GA, pursing a doctorate in pharmacy at the University of Georgia College of Pharmacy, Athens, GA. KATHERINE E. ODEGAARD (katherine.odegaard@unmc.edu) is a former biochemistry student at Gordon State College, now a graduate from University of Nebraska Medical Center, Omaha, NE. VICTOR H. VILCHIZ (vvilchiz@gordonstate.edu) is the Gordon State College dean for the School of Nursing, Health and Natural Sciences. CHULHEE KANG (chkang@wsu.edu) is a professor of biochemistry and biophysics at Washington State University, Pullman, WA. CATHY LEE (clee@gordonstate.edu) is a biology professor at Gordon State College.