# Multimedia in Biochemistry and Molecular Biology Education

# An Introductory Classroom Exercise on Protein Molecular Model Visualization and Detailed Analysis of Protein–Ligand Binding

**Poeylaut-Palena Andrés A. María de los Ángeles Laborde***

*Instituto de Química Rosario, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario—CONICET, Argentina*

## Abstract

A learning module for molecular level analysis of protein structure and ligand/drug interaction through the visualization of X-ray diffraction is presented. Using DeepView as molecular model visualization software, students learn about the general concepts of protein structure. This Biochemistry classroom exercise is designed to be carried out by following the detailed instructions that make software handling straightforward. Students learn about protein structure and gain insight into the molecular level of the interaction of two active compounds with their receptor. © 2013 by The International Union of Biochemistry and Molecular Biology, 41(2):118–124, 2013.

***Keywords:*** *computational chemistry; biochemistry; model; estradiol; protein structure; raloxifene*

## Introduction

The development of new equipment and the increasing number of research groups that are resolving protein structures have considerable increase in the available data collection. There are several well-known techniques to determine the structures of the macromolecules such as X-ray diffraction, nuclear magnetic resonance, or three-dimensional electronic microscopy. Analysis of these structures has become an important instrument to study the biological processes. Students need to gain capability to understand these data and learn how to manipulate them by using available software. In most cases, X-ray diffraction protein structures are available at public access databases [1,2]. This information can be efficiently used in exercises of Medicinal Chemistry undergraduate courses as well as Biochemistry and Molecular Biology educational activities. Therein, it is possible to find critical proteins for human health-forming complexes with substrates or inhibitors [3]. These molecular models can be analyzed by using a regular personal computer.

Therefore, we have developed a practical approach to teach students how to analyze protein structures using readily available software. During the teaching process, they will gain knowledge about the strengths and weaknesses of these methods. Students will observe how estradiol (**1**) (Fig. 1) binds to its α-receptor [3,4]. They will also learn many concepts related to protein structure such as primary, secondary, and tertiary structures, domains, hydrophobic cores, hydrogen bonds as key players, and many other issues. Finally, they will analyze how raloxifene (**2**) binds to the same estrogen receptor α (ER-α) and the factors involved in its biological activity as selective antagonist.

## Materials

To carry out this study, an ordinary personal computer is needed for every one or two students. The protein structure viewer software recommended is DeepView [5]. Software minimal requirements to run DeepView on PCs are Pentium or 486DX and Windows 95. As an example, this study was conducted on an AMD Sempron™//798 Hz//512 MB RAM, Windows XP Home Edition SP2®. The program's installation process is very simple. DeepView is available as an extractable file which generates a folder at any preferred location. Once the folder is created, the program does not need further installation and it could be run directly from the extracted folder.

Introductory information, discussion text, and software instructions are included in a Study Guide which is presented

**Address for correspondence to: Instituto de Química Rosario, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario—CONICET, Suipacha 531, S2002LRK Rosario, Argentina; E-mail: laborde@iquir-conicet.gov.ar.*

*Received 11 October 2012; accepted 30 December 2012*

*Additional Supporting Information may be found in the online version of this article.*

*DOI: 10.1002/bmb.20685*

*Published online 12 March 2013 in Wiley Online Library (wileyonlinelibrary.com)*

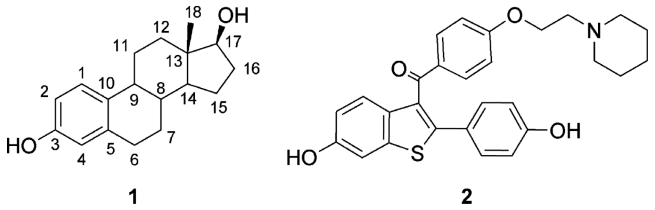


FIG 1 *Estrogen hormone estradiol (**1**) and a selective antagonist of its action, raloxifene (**2**).*

as Supporting Information material to this article. Crystallographic structures were obtained from the RCSB Protein Data Bank (PDB IDs 1ERE and 1ERR) [6]. Crystallographic water molecules and extra polypeptide chains were removed to leave just one protein structure. Copies of these modified files are included in the Supporting Information material.

## Experimental Methodology

### Organization of the Information

Our main efforts were focused on the Study Guide design. All software-related instructions were removed from the general introduction. Thus, the Study Guide is composed of two main sections: on the one hand, a discussion about molecular modeling and the analysis of protein structures and, on the other hand, a detailed list of actions to follow on DeepView (Supporting Information material). Commands to be executed were included in the general introduction section in a way that students can read the text and execute a command on the computer every time they find a call in the text. Each numbered command is fully explained and accompanied with the "expected effect" (Fig. 2 and Supporting Information material). The list of commands is extremely detailed, which allows most of the students to perform the exercise with almost no help. The "expected effect" for each action allows students to corroborate that each command was properly executed. Students should be asked to read the Study Guide and then proceed with the task, avoiding to go through the list of commands, ignoring the Study Guide discussion.

Based on this information, only concepts related to protein structural details are discussed within the general introduction, which is so simple that students can read it as an introduction prior to the class. This organization worked very well and became the essence of the Study Guide, reducing the complexity associated with program handling and focusing the student attention to protein structure which is the main subject of the exercise. There are stop points strategically placed (*ca.* every 10 commands) which bring homogeneity on the speed of the work granting that students complete the exercise stepwise. In our experience, without stop points it is very difficult for the teacher to keep the class under tight control. Different students could be at very distant parts of the exercise at the same time and it is hard to assess if they are properly doing and understanding their job. After all the students reach a stop point (Supporting Information material), a group discussion and a theoretical evaluation of the results are carried out; this helps to have a better control of the learning process.

### First Step

A general introduction to the DeepView [5,7] software is performed while few actions are executed. The software is fast and it has several independent windows which float around the screen that could be placed wherever the user chooses.

Immediately after opening the first structure, students can easily appreciate the great complexity and dimension of the molecular model. This constitutes an "information shock" and it is impossible to appreciate details in the middle of an atom jungle: sieving information is needed. When asked for some points to remark from this raw view, some students realize about the absence of hydrogen atoms in the structure. This event helps the instructor to open the first discussion: X-ray diffraction is unable to accurately solve the position of hydrogen atoms in the protein crystal structure. Coordinates are solved based on the X-ray diffraction upon the electronic density of each atom. As hydrogen atoms are poor in electronic density, it is hard to visualize them in the X-ray diffraction-derived structure [8,9]. Another concept to reinforce is that they are working just with the hormone-binding domain of the ER-α. The whole protein comprises two extra domains. Protein structure elucidation complexity is directly proportional to the size of the structure. Thus, for many proteins, the only way to obtain a good structure is to synthesize only the domain of

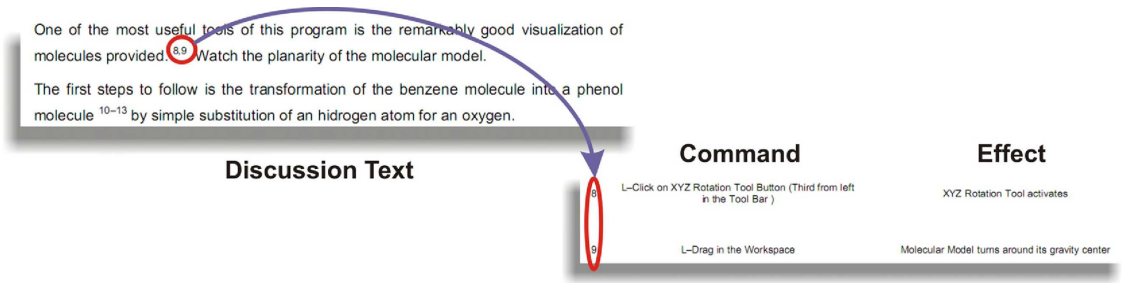


FIG 2 *Organization of information, command calls are placed as superscripts in the discussion text. Each call has an associated command placed in a list attached to the bottom of the text. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]*

interest. Often, a structural domain can maintain its structure and properties whether it is part of the whole protein or not. Of course, this strategy must be taken with care, as conclusions could be affected by this fact.

At the moment of performing visualizations, the DeepView core is the Control Panel (Fig. 3*a*). The polypeptide primary structure is fully detailed in this module and many visualization conditions can be varied for each amino acid residue with just one click. It is possible to leave just one amino acid in the workspace carrying out few actions, besides, coloring options are available. In this way, students begin to filter information in the screen. It was relieving for them to discover that structure visualization is under their command. The importance of the control panel will be revealed as the exercise progresses.

## Protein Structure Visualization: General Issues

Using the control panel, it is possible to replace for a ribbon which follows the path of each amino acid αC position all the protein amino acids (Fig. 3*b*). The ribbon is one of the most elegant and informative ways of protein visualization and it can be set with a couple of clicks. This is an easy way to observe questions related to secondary and tertiary structure imposing simplicity on the model filtering out useless information, leaving the workspace clearer (Students get surprised and begin to recognize structures, "there is an α-helix!.").

More information can be obtained by using colors. The ER-α hormone-binding domain structure has 12 α-helices and 1 β-sheet. Most students are able to recognize the presence of α-helices but not everyone can see the β-sheet at first glance. Therefore, coloring the ribbon in function of the secondary structure highlights the presence of a double-stranded β-sheet. Students also visualize secondary structures at molecular level. They leave just an α-helix on the workspace and then analyze the residues that take part in the structure. The polypeptide adopts a coiled conformation where the amino acid side chains spread out almost radially from the helix. Every backbone N—H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues before (Fig. 4*a*) [4,9]. As previously mentioned, hydrogen atoms are not visible in structures analyzed through X-ray diffraction, and hence the only moiety visible of an hydrogen bond is the heteroatoms involved (*i.e.* the O-atom from the carbonyl and the N-atom from the polypeptide chain). This section is of particular interest for students as they are able to view what had been explained to them in theory.

The same observation can be performed with the β-sheet. In this case, the secondary structure and the hydrogen bonds are too complex to be observed at first sight and hence applying the ribbon as a guide for a short while can simplify the observation. β-Sheet hydrogen bonds take place between two residues from different regions of the polypeptide, which is known as an intercatenary interaction (Fig. 4*b*) [4,9].

Additional data are hidden among the complex structure but can be revealed by looking at the control panel. For instance, the protein sequence begins at Ser305 which is the first amino acid of the crystalized domain but holds the numbering of the whole protein [3]. It is also possible to observe that there are some amino acids missing in the internal part of the sequence. These sequence gaps are owing to the high disorder of these amino acid residues on the X-ray diffraction. Atoms which do not uphold a periodic distance between every crystallographic cell are not detected in the diffraction pattern. Therefore, their absence in the crystallographic structure does not mean that they are missing in the real structure. This is another concept that can be hardly understood by students without any previous knowledge on protein structure.

The ER-α hormone-binding domain possesses two gaps which are visible in the Control Panel but would become even more evident in the workspace when coloring the extremes of each gap. Another way to point out the presence of gaps is to count how many polypeptide ends are present. However, not easy, the best way to search for polypeptide ends is to examine

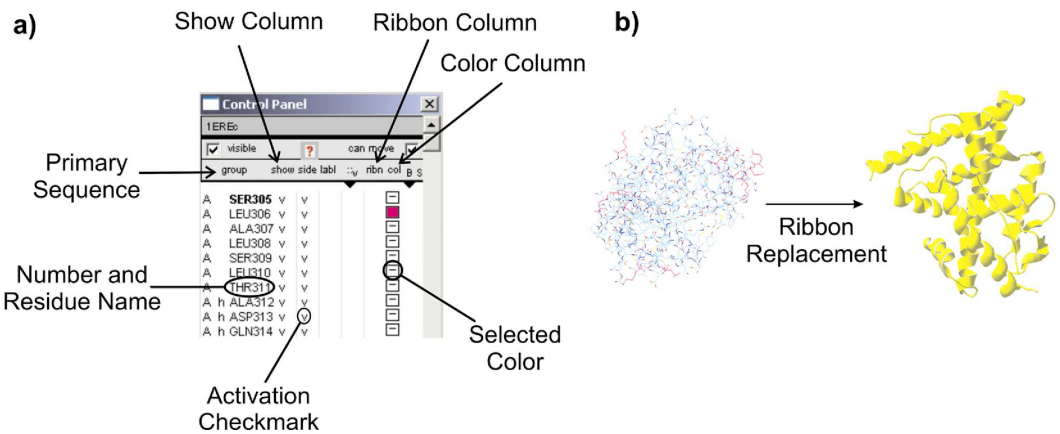


**FIG 3**

*(a) DeepView control panel. The amino acid sequence is shown and many visualization options are available in each column. A checkmark in the proper column assigns a visualization condition to the polypeptide chain. (b) A ribbon can be superimposed to the polypeptide chain. This makes clear the presence of secondary structures. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]*

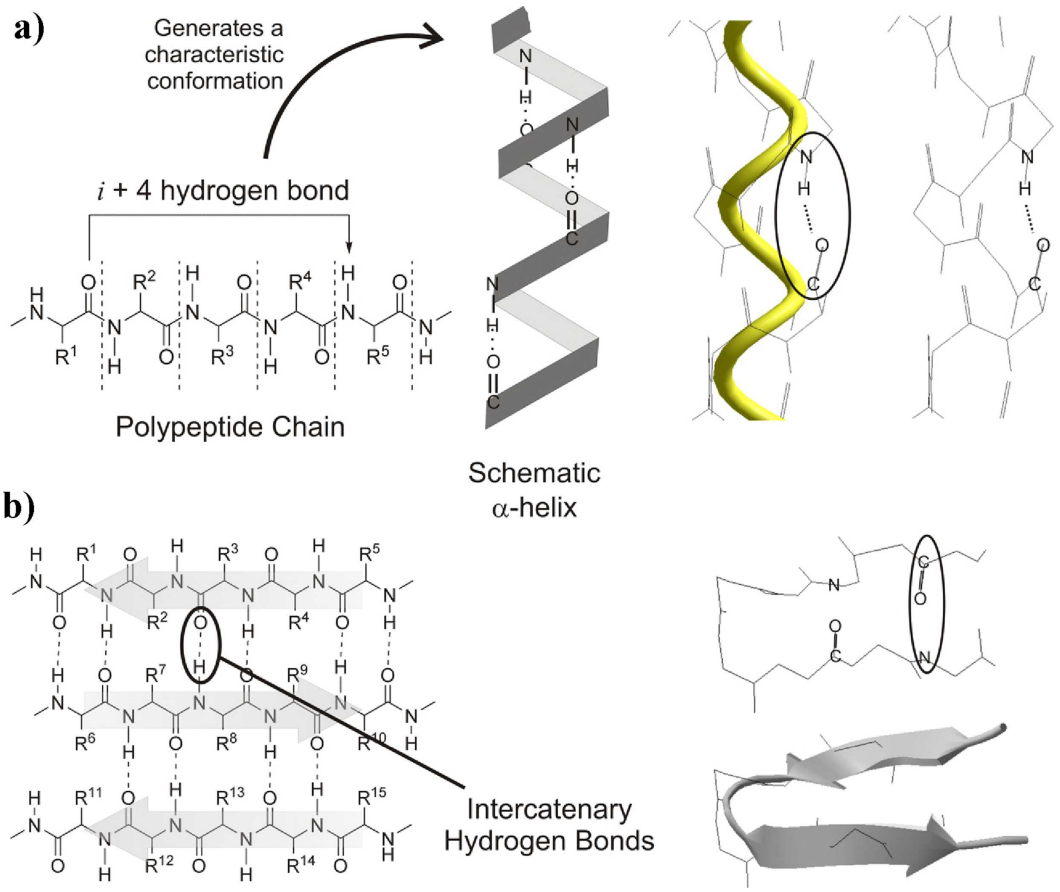


**FIG 4** *(a) Characteristics of α-helix structure. Intracatenary hydrogen bonds generate a coiled conformation. (b) Characteristics of β-sheet structure. Intercatenary hydrogen bonds maintain a sheet conformation. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]*

thoroughly the structure when it is represented as a ribbon. A regular polypeptide has two ends and there are two additional ends per gap (*i.e.* six ends in our case). After these observation techniques are used, students get the concept clearly.

## Analysis of the ER-α Hormone-binding Domain and Estradiol Interaction

Once the estradiol molecule is visible in the workspace, it is clear that the binding site is within the protein core, occluded in a hydrophobic pocket with just a couple of hydrophilic contact points. Coloring the structure by accessibility to the water is an effective way in which students assess the hydrophobicity of the binding pocket. Atoms colored in blue have minimal contact with water. Students are able to observe that estradiol is surrounded by amino acid side chains that are colored in blue which avoid contact with the water. Estradiol molecule can be seen as a hydrophobic plane with two hydroxyl groups which are of critical importance in its binding pattern. These hydroxyl groups are responsible for the hydrophilic contact points through hydrogen bonding to polar amino acid residues in the protein core.

The binding mode can be readily analyzed in detail. With just one click, students are able to leave the estradiol molecule alone in the workspace. Then, they can easily add groups that are close (<6 Å) to the estradiol hydroxyl groups to the view, find out which residues are taking part in a hydrogen bonding with estradiol, and measure the relative strength between them. Only the heteroatoms involved in the hydrogen bonds are visible. Consequently, strength has to be estimated as a function of the distance between two heteroatoms, being 2.50–2.80 Å a standard value [10]. At the moment of measuring the interatomic distances, students should be able to recognize the amino acid side chains properly, the use of the available amino acid labels helps with this task. Students notice that an OH in the position 3 of estradiol (Fig. 1) is donating a hydrogen bond to Glu353 (2.34 Å) and is accepting another from Arg394 (3.15 Å) (Fig. 5). It is important for the students to identify the atoms involved in the interaction correctly. This is a hard task for some of them. For instance, Arg394 possesses three nitrogen atoms in its side chain and only one is involved in the hydrogen bridge.

On the other side of estradiol, the hydroxyl group in position β17 (Fig. 1) is donating a hydrogen bond to a nitrogen

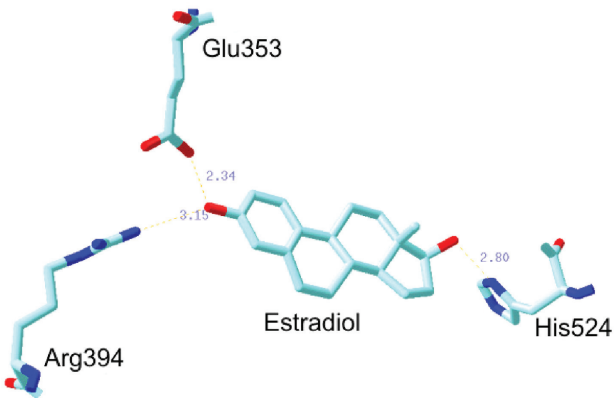


**FIG 5** *Hydrophilic interaction of estradiol with the ER-α hormone-binding domain.*

atom of the His524 imidazolic ring (2.80 Å). Of the three bond distances, 3-OH–Glu353 is the shortest one, which means that it is the strongest hydrophilic interaction in the hormone-binding process. Here, another discussion is carried out with the students: "Why is estradiol–Glu353 the strongest interaction?" A debate about the nature of the hydrogen bond is carried out and many of the students reach a correct explanation. In general terms, the interaction occurs between an electronically deficient hydrogen atom and an electron-rich heteroatom [10]. In this case, Glu353 possesses higher electron density than the other residues, in fact it is an anion, and is thus able to form the strongest hydrogen bond. This is a common situation with charged residues in a hydrophobic core: they tend to form strong interactions with available polar groups in the region [9].

## ER-α Hormone-Binding Domain in Complex With Raloxifene, Structural Comparison With ER-α–Estradiol Complex

Once general protein structure-related issues and estradiol binding have been carefully analyzed, the exercise continues with the study of raloxifene action. A quick examination of the ER-α–raloxifene complex reveals a couple of differences and students are invited to find out what they are. A major change is visible in the α-helix H12 position. At first sight, students observe it as a "flying α-helix"; this is owing to the presence of a new gap sequence close to H12 which leaves the α-helix disconnected in the solved structure. No further information could be obtained just by looking at the ribbon of ER-α–raloxifene complex. It is of greater use to superimpose structures of estradiol and raloxifene complexes. This can be easily achieved by the use of "Magic Fit" option. Students love this process!

Structures perfectly match each other almost in its whole sequence. Nevertheless, superimposition of structures brings to light some structural divergences between protein complexes. Again, the differences in the H12 region are clearly visible. Other changes include the appearance and disappear-

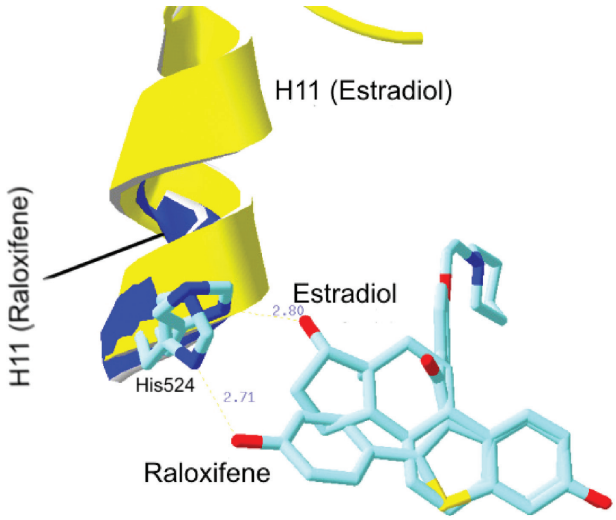


**FIG 6** *Hydroxyl displacement in the ER-α–raloxifene complex. His524 maintains the hydrogen bond and is displaced from the original position causing a disruption of H11 α-helix.*

ance of gaps and the dissolution of the β-sheet. Ribbon acts again as an efficient way of filtering information, providing a simplified view.

Working with the superimposition of structures, students leave just estradiol and raloxifene on the workspace and they will find that both molecules are at the same place in space, meaning that the binding site is the same. Estradiol and raloxifene share certain similarities: hydrophobic planes with two hydroxyl groups at each side although the distance between hydroxyl groups is not exactly the same and raloxifene has a side branch. These simultaneous similitudes/differences are responsible for the biological effect of raloxifene.

One hydroxyl of raloxifene binds through hydrogen bridges exactly to the same site to which estradiol does (Glu353 and Arg394). But the remaining hydroxyl of raloxifene is displaced from the original position of its hydrogen bond acceptor (His524). This variation leads to a deformation of the ER-α structure. His524 adopts a new conformation which causes a distortion in the structure of helix H11 (Fig. 6). Disruption of the H11 structure directly affects the position of H12 which is the neighboring secondary structure. Students can observe how raloxifene exerts its biological function at molecular level. Raloxifene binding changes the position of His524 imidazolic ring which follows the movement of the hydroxyl to maintain the hydrogen bridge.

Another important factor which contributes to the displacement of H12 is the raloxifene side branch. The side chain (not present in estradiol) occupies the same place originally occupied by H12 α-helix in the estradiol complex (Fig. 7). This is another important factor in the displacement of H12 from its original position. This may be the most exciting part of the exercise for students owing to the analysis of the action of a therapeutic agent at molecular level.

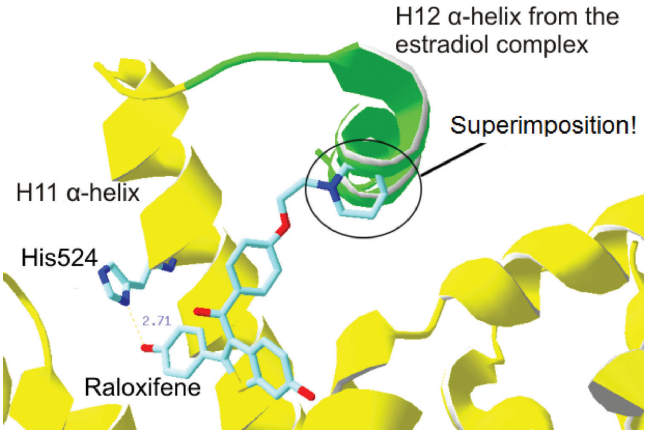


**FIG 7** *Superimposition of structures: raloxifene and estradiol complex ribbon. Raloxifene side-chain superimposition with H12 estradiol complex ribbon and location of His524 in H11 are highlighted.*

### Correlation Between the Visualizations and the Biological Activity

At this point, students are engrossed looking at a couple of atoms in the screen. They can see how atoms are displaced some Ångström away but they miss how this changes the activity of the same protein; at this point, a debate is carried out. Most of the students clearly assimilate the concept. Raloxifene binding modifies the topography of the ER-α hormone-binding domain in the region of H12 while maintaining the rest of the structure. These receptors exert their biological functions by association with other proteins. This happens owing to the interaction between the protein surfaces. Hence, if topography is altered, the biological function is altered as well.

Raloxifene acts as an agonist in some tissues, whereas as an antagonist in others. This situation is hard to explain, even researchers in the area have made some speculations. Students are encouraged to propose different hypotheses. Some of them explained it as follows: "ER-α–raloxifene maintains a great resemblance with the estradiol complex; therefore, it could be allowed to interact with the proper transcription factors and hence act as an agonist in some tissues. On the other hand, the perturbation in H12 region could be seen as the factor for antagonism preventing the correct interaction with other proteins." Regardless of the fact whether they are right or wrong, students put forward their own hypotheses. Debate could be lengthened by mentioning that estradiol is also capable of binding to at least another receptor (ER-β). What is clear is that the observations they have performed at the H12 region are a good molecular basis for the antagonic interaction of raloxifene with estradiol [3].

### Metrics of Success

This exercise was undertaken by 200 undergraduate students of Pharmacy in two semesters. Less than 4% of them failed in the multiple choice exam and no one failed when retested. Exams covered three main issues: (i) knowledge acquired by reading the study guide prior to the classroom exercise, (ii) the learning about general topics on protein structure during the exercise, and (iii) the learning about the interaction estradiol/raloxifene with their receptor and the biological consequences of it. Representative questions of each issue are as follows:

| | |
|---|---|
| Study Guide Knowledge | 1. From which database were obtained the structures used during this classroom exercise?<br>2. Which technique was employed to determine the structure we have used? Which other relevant techniques could be used to determine macromolecule structures? |
| General Protein Structure | 3. Describe and draw the hydrogen bonding pattern within an α-helix as you have observed during this classroom exercise.<br>4. Which is/are the origin/s of gaps formation in the structure determined by X-ray crystallography? |
| Receptor Interaction | 5. Which are the main hydrophilic contact points of estradiol inside the hydrophobic pocket? Given the hydrogen bond distances, which amino acid side chain is bound stronger to estradiol?<br>6. In which part of the structure occurs the maximum change upon binding of raloxifene? How this could account for its biological action? |

An anonymous survey was performed, revealing that about 95% of the students said the exercise was either interesting or very interesting (Fig. 8). Moreover, 81% of the students said that the classroom exercise was a not difficult at all. Most of 16% who said that the exercise was intellectually complex suggested that the experience had been interesting, which is also a sign of success.

## Conclusions

This article presents a learning module in molecular level analysis of protein structure and ligand/drug action that can be used in all areas related to Biochemistry. It is worth mention that there is no need for extraordinary computers or a squad of greatly qualified professors to carry out this classroom exercise. Students were delighted to be able to observe the action of an active principle at such depth. Through this activity, they realize how computational chemistry can be helpful for chemists and biochemists in the development of new active principles.

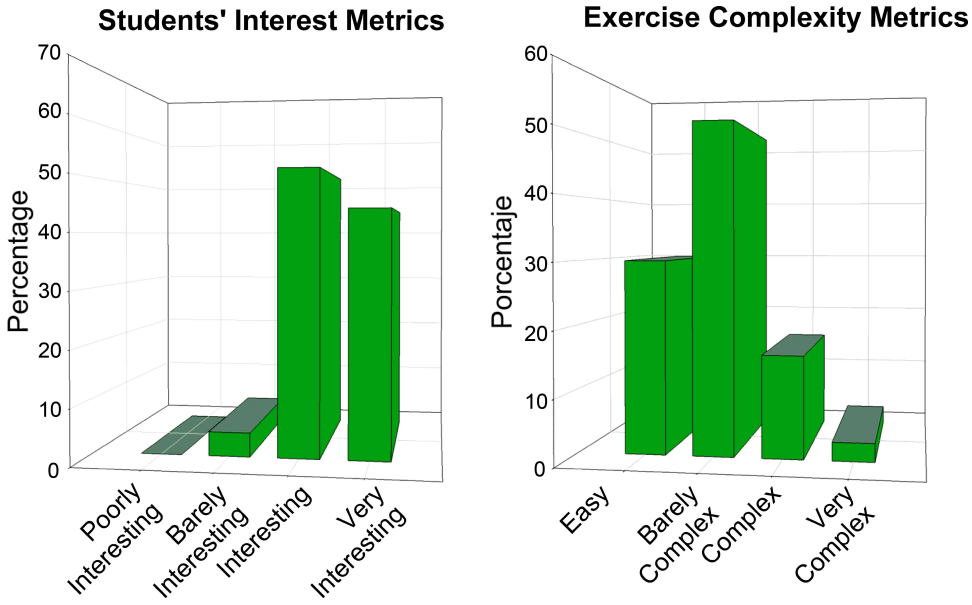


**FIG 8** *Results of the anonymous survey asking for interestingness and complexity of the exercise. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]*

## Acknowledgements

This study was supported by the Universidad Nacional de Rosario. Authors acknowledge M. Culasso and M. Robson for kindly reviewing the manuscript in search of English language corrections.

### References

[1] Ship, N. J. and Zamble, D. B. (2005) Analyzing the 3D structure of human carbonic anhydrase II and its mutants using deep view and the Protein Data Bank.J. Chem. Ed. 82, 1805.
[2] World Wide Protein Data Bank:www.wwPDB.org (revised 2012). PDB groups most of the relevant 3D structures.
[3] Brzozowski, A. M., Pike, A. C. W., Dauter, Z., Hubbard, R. E., Bonn, T., Engstrom, O., Ohman, L., Greene, G. L., Gustafsson, J.-A., and Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. Nature 389, 753–758.
[4] Stryer, L. (1999) Biochemistry, 4th ed., W.H. Freeman and Company, New York, pp. 18–44.
[5] DeepView:www.expasy.org/spdbv/ (revised 2012). DeepView is a program of free distribution which can be downloaded from the internet free of cost.
[6] RCSB Protein Data Bank:http://www.rcsb.org (revised 2012). RCSB Protein Data Bank is a free access database of 3D structures of biomacromolecules.
[7] Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis 18, 2714–2723.
[8] McPherson, A. (2009)Introduction to Macromolecular Crystallography, 2nd ed., Wiley-Blackwell, Hoboken, New Jersey.
[9] Fersht, A. (1999) Structure and Mechanism in Protein Science, W. H. Freeman and Company, New York.
[10] Pihko, P. M. (2009) Hydrogen Bonding in Organic Synthesis, 1st ed., Wiley-VCH Verlag GmbH & Co, Mörlenbach.